CONSERVATION GENOMICS IN SPECIES REINTRODUCTIONS: THE ASIATIC WILD ASS *EQUUS HEMIONUS* IN ISRAEL

Lilith Johanna Zecherle

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Doctor of Philosophy.

This research programme was carried out
in collaboration with Ben-Gurion University of the Negev, Israel.

February 2020

Table of Contents

Abstract

Conservation reintroductions are a frequently used management tool for the recovery of endangered species. However, many reintroductions fail to establish viable, self-sustaining populations. There are a multitude of factors that can impact the success of a reintroduction programme and population genetic aspects have been identified as an essential factor in the long-term persistence of reintroduced populations. However, due to a general lack of detailed long-term data sets, little is known about how different reintroduction strategies affect the genetic viability of a population and the long-term reintroduction success. In this thesis I apply high-resolution genomic tools to investigate the reintroduction of the Asiatic wild ass *Equus hemionus* in Israel. This case study provides a unique opportunity to investigate genetic impacts of conservation reintroductions, as it offers a long-term data set and a rare reintroduction protocol: founder individuals of the population were sourced from two different subspecies. I recovered a genome-wide set of genetic markers for the species using high-throughput sequencing techniques. Analyses based on this data set show that the populations display high levels of subspecies admixture and that population genetic parameters indicate a relatively high genetic variability compared with other reintroduced *E. hemionus* populations. These findings suggest that the highly controversial practice of subspecies admixture may be beneficial to reintroduction success in certain scenarios. Furthermore, I apply tools and methods from landscape ecology to uncover that habitat characteristics impact individual habitat selection but not genetic relatedness across the landscape. These findings suggest that current landscape configurations pose no barrier to gene flow in the reintroduced population. The presented results provide new insights on the population in Israel, relevant for its continued management. Furthermore, the outcome of this study has broader implications for conservation reintroductions in general.

Declaration

I declare that that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Lilith J. Zecherle

February 2020

Acknowledgements

I would like to thank my team of supervisors and advisors for their continued support and guidance, which was fundamental to the success of this project. **Serge Wich** was an important help, guiding me through the early stages of my PhD and helping me settle in at LJMU. My warmest thanks also to **Hazel Nichols** who has been a fantastic supervisor throughout my PhD. Hazel has also been a great mentor and helped me develop myself as a young scientist while providing me with the freedom to follow my own interests and ideas. I am also very grateful to my supervisor **Richard Brown** who has been extremely supportive and encouraging throughout my PhD. Particularly, his statistical and analytical support with the research have been critical to the success of my project. Furthermore, this project would not have been possible without the collaboration with **Shirli Bar-David**, my external advisor. Shirli has been a constant support throughout the entire project. I am grateful for her welcoming me to her lab and making me feel part of her research group. She has guided me in my research with many discussions, creative ideas and enthusiasm. I would also like to thank my external advisor **Alan Templeton** who has contributed to the success of my research from an early stage. I am very grateful for his advice and insights on population genetic theory.

This project was based on a collaboration with Ben-Gurion University of the Negev in Israel and I would like to thank my friends and lab members **Orly Brandeis**, **Yuval Cohen, Jonathan Tichon**, **Elyasaf Freiman**, **Omer Golan**, **Tal Halevy**, **Uri Roll**, **Gili Greenbaum** and **Ariel Altmann** for welcoming me to the department and making me feel at home. Specifically, I would like to thank **Naama Shahar** for her invaluable technical support during the lab work and sample shipments and for many enjoyable lunch breaks together. At the University of Sheffield, I would like to thank **Deborah Dawson**, for helping arrange everything for my visit there. I am also grateful to **Gavin Horsburgh**, for his guidance and training during the preparation of the sequencing libraries. The help from **Helen Hipperson** with the daunting task of the bioinformatic analysis has been invaluable and I am thankful for her long-term support. At LJMU, **George Zouganelis** and **Jerry Bird** have been a great support, during lab work, shipment and collection of samples. Furthermore, I would like to thank **Ineke Knot**, **Noémie Bonnin, Anne-Sophie Crunchant, Camille Giuliano** and the other members of room 306 for their company and hospitality during my time there. This research was based on GPS data and DNA samples, which have been kindly provided by the **Israeli Nature and Parks Authorities**, **Tiergarten Nürnberg**, **Zoo Rostock**, **Chester Zoo** and **Wildlands Adventure Zoo Emmen** and I am very grateful for their contribution and support. Specifically, I would like to thank **Gal Vine** for his time and advice on the ArcGIS data. **NERC** and **LJMU** have funded the bioinformatic sequencing and my PhD studentship for which I am most grateful.

I would like to thank my family and friends for their tremendous support throughout the past three and a half years. Special thanks go to my friend **Trevor Conroy**, for always providing me with a home away from home during my frequent visits. I would like to thank my parents, **Alfred** and **Meggie Zecherle**, for their support and

encouragement, for building me up during my lows and celebrating my successes with me. Finally, I would like to thank **Yoel Bitton** for being an incredibly supportive and understanding partner, for his countless pep talks and for always providing me with everything I needed.

1. General introduction and background

Conservation biology is a relatively young scientific field that draws on diverse techniques from multiple disciplines, such as behavioural ecology, population genetics and movement ecology, which are applied to the conservation of endangered species (Frankham, Briscoe, & Ballou, 2002). It is considered a crisis discipline, which means that due to the urgency of the matter, practical recommendations and decisions may be needed before a thorough empirical evidence base can be established (Soulé, 1985). This predicament is likely the underlying cause for many debates and controversies in the field of conservation biology. Since the fundamental purpose of conservation biology research is to find solutions for time-critical conservation threats to rare and endangered species, it is often not possible to test theories in controlled laboratory experiments, and much of the research relies on the opportunistic analysis of case studies.

The sub-discipline of conservation genetics, which applies traditional population genetics and evolutionary theory to the conservation of endangered species, has emerged recently. In this introductory chapter, I first present a review of the literature on genetic applications for conservation research and highlight areas of controversy in the field. Then, I draw specific attention to genetic applications in species reintroduction programmes and present the special study opportunity provided by the reintroduction of Asiatic wild ass in Israel. Finally, I conclude with an outline of the research aim and objectives of this thesis.

1.1 Genetics in the conservation of endangered species

Small populations, which are the focus of most conservation work, face an increased extinction risk due to environmental, demographic and genetic factors (Frankham, Briscoe, et al., 2002; Lande, 1988). There are two main genetic processes which affect the viability of small populations: i) inbreeding, the mating between relatives and ii) genetic drift, the random loss or fixation of alleles. When population size is reduced inbreeding becomes inevitable due to limited mating opportunities. High rates of inbreeding lead to an increase in homozygosity and an accumulation of deleterious recessive alleles. This can cause inbreeding depression, which results in a reduction in the reproductive rates and offspring survival, ultimately driving the population to extinction (Charlesworth & Charlesworth, 1987). Inbreeding accumulates faster in small populations with fewer mating opportunities and inbreeding depression can

occur in a population over a relatively short time frame (Frankham, Ballou, Briscoe, & McInnes, 2002). The extinction risk of small populations is further increased by stochastic genetic processes. Genetic drift, acts more strongly in small populations (Frankham, Briscoe, et al., 2002). Both inbreeding and genetic drift cause a reduction in neutral genetic diversity, thereby lowering the population's adaptive potential and increasing its risk of extinction (Reed & Frankham, 2003; Templeton, 2017).

The importance of considering genetic factors in species conservation, which had predominantly focussed on demographic and ecological processes, was first expressed by Otto Frankel and Michael Soulé (Frankel, 1974; Frankel & Soulé, 1981). However, the relationship between genetics and species extinctions has been challenged continuously since it was first presented (reviewed by DeSalle and Amato, 2004; DeSalle, 2005). The main criticism was that genetic processes occur over a very long timeframe, making them irrelevant to conservation (Lande, 1988). While the risk posed by inbreeding and genetic drift had been acknowledged for *ex situ* breeding stock, fast-declining wild populations were believed to become extinct from other pressures before any genetic erosion could occur (Caro & Laurenson, 1994; Lande, 1988). This criticism was largely based on a lack of empirical data from wild populations and resulting erroneous assumptions about the lethal effect of inbreeding and loss of genetic diversity (Frankham, 2005).

A major constraint to testing conservation genetic theory in wild populations was the difficulty of disentangling genetic factors from other potential causes for population extinction. A classic study by Saccheri et al. (1998) succeeded in highlighting the negative impact of inbreeding in a wild butterfly meta-population. The study compared 42 inbred and outbred groups of Glanville fritillary (*Meliaea cinxia*) which were released in a field setting. In the experiment all inbred populations become extinct and inbreeding accounted for 26% of variation in extinction rates after controlling for other stochastic and ecological factors. Similarly, a study on wild populations of white-footed mice (*Peromyscus leucopus noveboracensis*), reported a significantly reduced survival probability in inbred compared to outbred groups (Jiménez, Hughes, Alaks, Graham, & Lacy, 1994). More recently, a high impact meta-analysis of numerous empirical studies across taxa highlighted the importance of genetic factors to fight extinction risk: Spielman et al. (2004) showed a significant relationship between genetic diversity and

risk of extinction, with threatened taxa displaying on average a 35% lower level of heterozygosity than non-threatened taxa. These studies were central in demonstrating the link between genetic deterioration and extinction and helped establish genetics as a tool for species conservation.

Today there is ample evidence of the negative impacts of inbreeding depression in wild populations (reviewed by Crnokrak and Roff, 1999; Hedrick and Kalinowski, 2000; Keller and Waller, 2002). The importance of genetic considerations for conservation is now widely recognised and genetic analyses are increasingly applied to the study of endangered populations (Frankham, 2005; Frankham & Ralls, 1998). Furthermore, conservation genetics has developed into a diverse and continuously growing field. In the management of wild populations, where detailed demographic records are often not available, genetic analysis presents a powerful, comparatively easy method to gather diverse information, such as census and effective population size (Moore & Vigilant, 2014; Pelletier, Turgeon, Bourret, Garant, & St-Laurent, 2019; Solberg, Bellemain, Drageset, Taberlet, & Swenson, 2006), demographic history (Pilot et al., 2014; Stoffel et al., 2018) and population connectivity (Howell, Koen, Williams, Roloff, & Scribner, 2016; Latch & Rhodes, 2005; Riley et al., 2006). Other applications of genetics in conservation include the management of captive populations (Hammer, Schwammer, & Suchentrunk, 2008; Wisely, Mcdonald, & Buskirk, 2003; Witzenberger & Hochkirch, 2011), wildlife forensics (Alacs, Georges, Fitzsimmons, & Robertson, 2009; Mondol, Sridhar, Yadav, Gubbi, & Ramakrishnan, 2014; Ogden & Linacre, 2015), invasive species management (Berry et al., 2012; Rollins, Woolnough, Wilton, Sinclair, & Sherwin, 2009; Thresher et al., 2013) and reintroductions.

1.2 Conservation genetics in species reintroduction programmes

Genetic factors are a key element in species reintroductions. The successful establishment and long-term persistence of a population are strongly influenced by population genetic dynamics (Seddon & Armstrong, 2016). This is because by nature, most reintroduced populations are small and isolated during the early stages of establishment. Consequently, reintroduced populations have an increased risk of extinction due to inbreeding and genetic drift (Frankham, Briscoe, et al., 2002).

1.2.1 Selection of founders

There are two main applications for genetics in species reintroductions: i) the selection of suitable founder individuals prior to release and ii) post-release genetic monitoring of the population. Selecting founders is a critical step, which has long-term impacts on the outcome of a programme (Olsson, 2007; Schneider, 2011). Consequently, the International Union for the Conservation of Nature (IUCN) recommends that adequate source populations should be selected with respect to taxonomy and environmental adaptations and that individuals should be taken from populations that have evolved under environmental conditions similar to those at the release site, to reduce the negative impact of new climatic and environmental pressures (He, Johansson, & Heath, 2016; IUCN/SSC, 2013). An example of the damaging consequences of selecting poorly adapted founder stock has been described for white storks (*Ciconia ciconia*) (Olsson, 2007). Individuals from a North African source population were reintroduced in Sweden, where they mixed with local stock. North African pairs had significantly poorer breeding success than native breeding pairs. The author concludes that no sustainable population could have been established with North African stock alone.

The number of individuals released during a reintroduction programme deserves careful consideration. During the establishment phase, populations experience a founding event; a period of strong genetic drift during which genetic diversity is lost at an increased rate (Frankham, Briscoe, et al., 2002). By releasing large numbers of founders, the negative impact of the founding event can be minimised. Fischer and Lindemayer (2000) analysed published data from 116 reintroductions and found that definite successes were fewer (18%) than definite failures (30%), when less than 100 individuals were released. However, in reality, the number of individuals available for release is often limited by high costs, logistics and the potential impact of removing a large number of individuals on the source population (Seddon & Armstrong, 2016; Tracy, Wallis, Efford, & Jamieson, 2011).

While smaller numbers of founders may suffice to withstand demographic stochasticity and establish populations, they can result in long-term genetic deterioration. In New Zealand, bird reintroductions to offshore islands are commonly based on an average number of 30 founders (Taylor, Jamieson, & Armstrong, 2005). Most populations have established successfully and displayed post-release

demographic growth, however, pedigree analysis has revealed that in the long-term these small founder numbers retain insufficient genetic variation and continued management is required to counteract accumulative inbreeding (Jamieson, 2010; Taylor et al., 2005). The problem of small founder numbers is exacerbated by the fact that often not all individuals contribute genetically to the population (Biebach & Keller, 2012). Post-release mortality, uneven sex ratio and reproductive skew lead to fewer effective founders (Jamieson, 2010; Miller, Nelson, Smith, & Moore, 2009; Wilson, Nishi, Elkin, & Strobeck, 2005). For example, of 58 reintroduced North Island robins only 25 individuals (43%) contributed genetically to the population (Jamieson, 2010).

When founders are few, the chances of long-term persistence of the population can be enhanced by selecting individuals in a way that maximises genetic diversity in the founding stock (He et al., 2016). Sourcing individuals from different populations can increase the captured genetic diversity, resulting in greater genetic variability maintained in the established population. In fact, analysis of long-term genetic data of 40 reintroduced Alpine ibex (*Capra ibex ibex*) populations uncovered that expected heterozygosity was more strongly impacted by the level of admixture in the founders than by the number of released individuals (Biebach & Keller, 2012). Similarly, in reintroduced European beaver (*Castor fiber*), populations displayed higher genetic diversity and produced larger average litter sizes when the founders were obtained from different sources (Saveljev & Milishnikov, 2002). These results are extremely valuable for reintroduction management as they highlight how admixture holds enormous potential for species recovery. Nevertheless, if source populations are genetically too divergent, admixture can be harmful and the population may suffer outbreeding depression (Edmands, 2007; Frankham et al., 2011). The result is a loss of local adaptations, which could jeopardize successful population establishment (Templeton et al., 1986). A classic example is that of the ibex reintroduced to former Czechoslovakia (Greig, 1979). After extinction of the local population, Alpine ibex (*Capra ibex ibex*) were successfully reintroduced to the Tatra mountains. To supplement the population, additional Nubian ibex (*Capra ibex nubiana*) and Bezoar ibex (*Capra aegarus aegarus*) were released. This admixture between three different (sub)species rendered fertile hybrids, yet they displayed a fatal shift in their rutting

season. Kids were born too early, during winter months, with no chances of survival. This admixture resulted in the extinction of the entire population (Greig, 1979).

The above examples demonstrate that carefully evaluated admixture has the potential to improve reintroduction success, however, due to the associated risks it remains highly controversial (Allendorf, Leary, Spruell, & Wenburg, 2001; Shemesh, Shani, Carmel, Kent, & Sapir, 2018; Weeks et al., 2015). There are few well-documented case studies of reintroductions mixing different source populations (e.g. White *et al.*, 2018; Thavornkanlapachai *et al.*, 2019), hence practitioners are lacking data for informed decision making. Detailed evaluations of recovery programmes which sourced founders from different populations are important to create a scientific evidence base and guide future direction in conservation reintroductions.

1.2.2 Genetic monitoring

The second key application for genetics in reintroduction programmes is post-release monitoring. Genetic monitoring is an important tool to assess population viability and quantify the success of recovery programmes. Genetic methods have been applied successfully to detect inbreeding (Brekke, Bennett, Wang, Pettorelli, & Ewen, 2010), loss of genetic diversity (Bull, Heurich, Saveljev, Schmidt, & Förster, 2016) or slow demographic growth (De Barba et al., 2010) in reintroduced populations. Furthermore, by comparing alternative reintroduction strategies, protocols can be optimised (Schwartz, Luikart, & Waples, 2007; Tollington et al., 2013). For example, an extensive analysis of nine reintroduced ibex populations identified that genome-wide heterozygosity was reduced, and inbreeding levels elevated in populations that were sourced from previously reintroduced populations, compared to those sourced from autochthonous populations. Based on these results the authors recommended to avoid the former practice of stepwise reintroductions in the future (Grossen, Biebach, Angelone-Alasaad, Keller, & Croll, 2018).

In most recovery programmes the highest monitoring priority is given to demographic parameters such as population size and demographic growth rate. However, including population genetic measures can provide important insights which would otherwise be missed (Gitzen et al., 2016). For example, a genetic analysis of reintroduced Chinook salmon (*Oncorhynchus tshawytscha*) showed that observed demographic growth of the established population was largely dependent on immigration (Sard et al., 2016).

Genetic parentage assignment demonstrated that the rate of reproductive output of reintroduced individuals was well below population replacement. Hence, the applied reintroduction strategy failed to establish a viable population; critical information which regular census estimates failed to uncover.

In the early stages the goal of a reintroduction is the establishment of a self-sustaining population, however, long-term reintroduction success often depends on habitat connectivity and the integration of the reintroduced population into the wider meta-population (Armstrong & Seddon, 2008; Seddon & Armstrong, 2016). Genetic connectivity analysis is an important measure to detect future threats, even for seemingly successful reintroductions. For example, the return of wolves (*Canis lupus*) to Yellowstone National Park is widely regarded to be a conservation success story. A comprehensive population genetic analysis confirmed high genetic diversity and low inbreeding levels in the population, 10 years after establishment (vonHoldt et al., 2007). However, the authors also highlight that there was no indication of gene flow into the national park and estimated that population-wide inbreeding depression could arise in approximately 60 years. This study underscores the importance of long-term genetic monitoring of reintroduced populations and the need for continued management.

There is a clear discrepancy between the popularity of reintroductions as a conservation management tool and the high rate of programmes that fail. The application of reintroductions has consistently increased over the past decades (Fischer & Lindenmayer, 2000; Seddon, Armstrong, & Maloney, 2007) and this trend is expected to persist as global biodiversity continues to decline and an increasing number of species will require conservation interventions (IPBES, Diaz, Settele and Brondízio, 2019). However, there is no standardised measure of success and most programmes miss long-term monitoring data either due to financial constraints or a lack of awareness (Fischer & Lindenmayer, 2000). The IUCN included population genetic considerations in their species reintroduction guidelines only in the revised version from 2013 (IUCN/SSC, 2013) and consequently, very few case studies with long-term genetic records exist today (e.g. VonHoldt et al., 2007; Wisely et al., 2008; Biebach and Keller, 2012; White et al., 2018). Thus, rare long-term data sets are a particularly valuable source of information to understand how different management

protocols impact population development. There is an urgent need to analyse and learn from past species recovery programmes and identify methods to improve the success rates of reintroductions.

1.3 The reintroduction of the Asiatic wild ass in Israel

The reintroduction of Asiatic wild ass *Equus hemionus* in Israel provides an interesting case study for conservation. After the extinction of the local subspecies, the Syrian wild ass *E.h.hemippus*, individuals from two different subspecies, the Iranian onager *E.h.onager* and Turkmen kulan *E.h.kulan,* were reintroduced (Saltz & Rubenstein, 1995). Therefore, this species recovery presents a rare opportunity to investigate the impact of an unusual reintroduction protocol on the established population.

1.3.1 Biology of the Asiatic wild ass

The Asiatic wild ass *Equus hemionus* (Pallas 1775) is a large-bodied wild equid well adapted to semi-arid and arid environments (Schoenecker, King, Nordquist, Nandintsetseg, & Cao, 2016). Asiatic wild ass live in fission-fusion societies with highly variable group sizes and no lasting social bonds (Boyd, Scorolli, Nowzari, & Bouskila, 2016). Females form groups based on their reproductive status and resource requirements rather than genetic relatedness (Altman, 2016; Renan et al., 2018). Males either roam in non-stable bachelor herds or defend individual territories (Renan et al., 2018). Wild asses are highly mobile and travel an average cumulative daily distance of 21.8km. Mean 95% home range sizes reach 20-30km^2 (Giotto, Gerard, Ziv, Bouskila, & Bar-David, 2015; Moehlman, Shah, Masseti, & Feh, 2010).

Asiatic wild asses have a resource-defence polygyny mating system, with males establishing territories near permanent water sources. Breeding and foaling occurs during the spring and summer, with females first reproducing at 3 years of age and males at 5 years of age (Saltz & Rubenstein, 1995; Volf, 2010). Direct monitoring of the reintroduced population in Israel (see 1.3.3) revealed a reproductive rate of R_0=1.87, with females giving birth to one foal between every year and every two years (Saltz & Rubenstein, 1995). In the wild *E. hemionus* can reach a maximum lifespan of 16 years, with a generation time of 7.5 years (Ransom et al., 2016; Saltz & Rubenstein, 1995).

1.3.2 Taxonomy and conservation status of Asiatic wild asses

Asiatic wild ass formerly ranged across steppes and grasslands of Western and Central Asia. However, today the species' distribution is highly fragmented and populations persist only in isolated refugia (Kaczensky *et al.*, 2015, Fig. 1.1). Currently, the Asiatic wild asses are divided into 5 subspecies. One of these, the Syrian wild ass *E.h.hemippus*, was endemic to the Middle East including Israel and has become extinct, with the last wild specimen seen in 1927 (Groves & Mazák, 1967). The four extant subspecies consist of the Iranian onager *E.h.onager*, the Turkmen kulan *E.h.kulan*, the Indian khur *E.h.khur* and the Mongolian khulan *E.h.hemionus* (Kaczensky *et al.*, 2015, Fig. 1.1). The current taxonomic classification has been primarily based on differences in coat colour, skull morphology and geographic distribution (Groves & Mazák, 1967). However, a growing number of genetic analyses have investigated the phylogeny and evolutionary history of equids, triggering an on-going debate about the correct taxonomic classification of the Asiatic wild asses. In particular, the current distinction between the subspecies *E.h.onager* and *E.h.kulan* has been challenged (Bennett et al., 2017; Oakenfull, Lim, & Ryder, 2000).

Among the first to contest the current classification were Oakenfull et al. (2000) who used mitochondrial DNA sequences and maximum likelihood methods to investigate the phylogeny of extant equids. The authors report that unlike other equids the Asiatic wild ass did not cluster together by subspecies on the phylogenetic tree. Instead, kulan and onager individuals were found to cluster closely together and even shared a mitochondrial haplotype. These results suggest recent geneflow between the subspecies. The authors proposed an original divergence between kulans and onagers, followed by recent admixture before separating into the current populations. This possibility is further supported by polymorphism of the chromosome number in both subspecies. While the number of chromosomes differ between most *E.h.onager* (n2=56) and *E.h.kulan* (n2=54), the subspecies also share an intermediated karyotype (n2=55) (Ryder & Chemnick, 1990). Based on the genetic similarity between onagers and kulans, Oakenfull et al. (2002) concluded that there is not enough genetic distinctiveness to separate them into different subspecies. Vilstrup et al. (2013) provided further support for these findings. The authors re-sequenced whole mitogenomes of all extant equids and also reported a mixed *E.h.onager*/ *E.h.kulan*

clade. However, the analyses in both studies displayed low bootstrap support for the suggested grouping of Asiatic asses and other molecular studies contradict the reported findings. Investigations based on different nuclear DNA markers, including microsatellites (Krüger, Gaillard, Stranzinger, & Rieder, 2005), high-density SNP arrays (McCue et al., 2012) and short genomic sequences (Steiner, Mitelberg, Tursi, & Ryder, 2012), consistently report separate subspecies groups of *E.h.onager* and *E.h.kulan* and support the current taxonomic distinction. Nevertheless, these studies were based on small numbers of samples (N=2-18), from few captive populations, which might not accurately represent the genetic variability found in wild populations.

Recent investigations have aimed to increase sample sizes and number of source populations. Additionally, multiple studies have drawn on historical and ancient DNA samples from extinct lineages to resolve the phylogeny of modern-day equids (Bennett et al., 2017; Geigl & Grange, 2012; Vilstrup et al., 2013). The most comprehensive analysis to date used 253 ancient, historic and modern samples from wild populations (Bennett et al., 2017). The study combined phylogeographic and phylogenetic analyses based on mitochondrial DNA sequences and assigned *E.h.onager* and *E.h.kulan* samples to three different mitochondrial clades. There was a distinct onager clade to which most *E.h.onager* samples were assigned to (1). However, there was also a second clade shared between *E.h.onager* and *E.h.kulan* (2). A third clade (3) was identified containing mostly ancient Caucasian samples. The study also included samples from the reintroduced population in Israel (see 1.3.3), which were assigned to two different clades (1,2). The findings from Bennett et al. (2017) are consistent with genetic clustering analysis of extant populations (Kaczensky et al., 2018) and previous reports of a shared mitochondrial haplotype between onagers and kulans (Oakenfull et al., 2000; Vilstrup et al., 2013), which suggests recent gene flow between the subspecies.

To summarise, studies based on mitochondrial DNA markers consistently show shared mitochondrial haplotypes and close phylogenetic proximity between the subspecies (Oakenfull et al., 2000; Vilstrup et al., 2013). This is opposed by studies based on nuclear markers identifying distinct genetic clusters between onagers and kulans (Krüger et al., 2005; McCue et al., 2012; Steiner et al., 2012). Discrepancies between different analyses based on mitochondrial and nuclear genetic markers are commonly

reported in phylogenetic studies due to their different properties (Rubinoff & Holland, 2005). In the case of the Asiatic wild ass it is likely that secondary contact and incomplete mitochondrial lineage sorting resulted in the observed overlapping mitochondrial haplotypes (Rosenbom et al., 2015; Steiner et al., 2012). Conversely, nuclear DNA markers may represent recent divergence between the subspecies. However, the outcome of these nuclear DNA cluster analyses also depends on the sampling regime. Small sample sizes obtained from captive populations, which likely originated from a small number of founders, might not represent the overall genetic variability present within the subspecies (Garamszegi & Møller, 2010). Furthermore, observed differentiation between kulans and onagers might be exaggerated by recent genetic drift experienced by the captive populations rather than true divergence between the subspecies (Frankham, Ballou, et al., 2002; Weeks, Stoklosa, & Hoffmann, 2016).

Today, the intraspecific phylogeny of the wild ass has not yet been resolved, despite numerous genetic analyses. This has been attributed a highly complex evolutionary history including periods of gene flow, divergence and secondary contact (Jónsson et al., 2014; Rosenbom et al., 2015; Steiner et al., 2012). However, there is increasing evidence that the current taxonomic classification, which was defined in the absence of genetic data, is inadequate. Especially, the distinction between the Iranian and Turkmen subspecies has been criticised and some authors have advocated a revision of the current taxonomic classification (Oakenfull et al., 2000).

The disputed taxonomy of the species also impacts its conservation. The Asiatic wild ass is of considerable conservation interest and the species is currently classified as near threatened by the IUCN Red list of threatened species. Additionally, the 4 extant subspecies have been classified individually with Indian khur and Mongolian khulan listed as near-threatened and onagers and kulans both classified as endangered (Kaczensky et al., 2015). The population in Israel is listed as a hybrid population (*E.h.onager* x *E.h.kulan*) and classified as endangered based on a regional assessment (Moehlman et al., 2010). A major conservation concern for the endangered subspecies is the extreme fragmentation and small sizes of remaining wild populations (Bennett *et al.*, 2017; Kaczensky *et al.*, 2018, Fig. 1.1). To ensure effective conservation of the species some authors have recommended that the remaining populations of the

different subspecies should be managed as one metapopulation (Bennett et al., 2017). However, others have warned against losing genetic variability and unique local adaptations and recommend that onagers and kulans are managed separately (Kaczensky et al., 2018). Information from the potentially admixed population in Israel may provide important insights for future conservation management of the species.

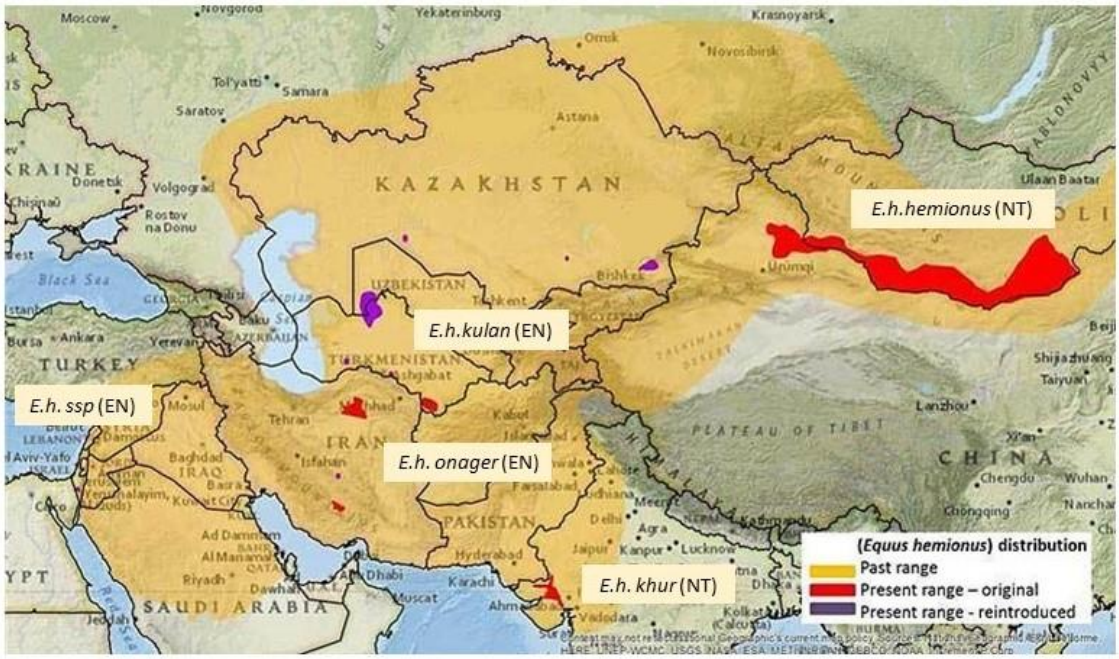


Fig. 1.1 Past and current range of original and reintroduced *Equus hemionus ssp.* populations and their IUCN Red List category. *E.h.khur* is native to India, *E.h.hemionus* is native to Mongolia and China. *E.h.onager* is native to Iran, *E.h.kulan* is native to Turkmenistan and has also been reintroduced in Uzbekistan and Kazakhstan. *E.h.hemippus* was endemic to the Middle East and has become extinct. *E.h.onager* and *E.h.kulan* have been reintroduced in Israel. The figure has been adapted from the Norwegian Institute for Nature Research (2019).

1.3.3 The population in Israel

In the 1960s, more than 30 years after the extinction of the local subspecies, the Israeli Nature and Parks Authorities (INPA) initiated a reintroduction programme and imported individuals from the most closely related subspecies to Israel. Six onagers (3 females, 3 males) from captive populations were brought to Israel in 1968. A year later an additional five kulans (3F, 2M) were imported from European zoos (Saltz & Rubenstein, 1995; Yoffe, 1980). With these eleven founders a captive breeding core was established in the Hai-Bar Yotvata Reserve, a 2km^2 fenced area in the South of Israel (Saltz & Rubenstein, 1995). The breeding core population was not managed, and

no known studbook or pedigree exist. The two-subspecies were allowed to interbreed, yet there are no records on whether interbreeding occurred (Yoffe, 1980).

Individuals from this breeding core were subsequently released into the wild. In 1982 five males were released in the Maktesh Ramon Nature Reserve in the Negev desert (Fig. 1.2; Saltz & & Rubenstein, 1995). However, most of these individuals dispersed quickly after release and their fate is unknown. Between 1983 and 1987 three additional releases followed, during which a total of nine males and 14 females were released to the nature reserve. Three of these individuals were found dead shortly after release, the cause of death is unknown. Between 1992 and 1993 an additional ten individuals (7F, 3M) were released in the Paran streambed during two events (Fig. 1.2). Three females were found dead shortly after release, the cause of death was not determined. In summary a total of 38 individuals were released of which at least six are known to have died shortly after release. In 1991, three generations after the establishment of the breeding core and prior to the release in the Paran streambed, blood samples were taken from 30 individuals of the captive population (Gueta, Templeton, & Bar-David, 2014; Sinai, 1994). These blood samples were preserved and have been used as representation of the captive breeding population in this thesis ("founder population").

Post-release population growth was initially slow, but it increased after two generations (Saltz & Rubenstein, 1995). Today the population is estimated at 300 individuals (Renan et al., 2018). The population has expanded its range across most of Southern Israel, with activity centres around three artificial water sources, which are the main permanent sources of water (Renan et al., 2018; Fig. 1.2).

The reintroduction of the Asiatic wild ass in Israel offers an excellent opportunity to study strategies for improving reintroduction success, because of the accidental mixture of individuals from different subspecies. Furthermore, this reintroduction provides a rare long-term genetic data set for a non-model species in a wild setting. To date, 18 reintroduction attempts of Asiatic wild asses have been recorded, yet only 28% (5) of these have succeeded in establishing large viable populations (>100 individuals and stable or increasing population trend, Kaczensky et al., 2016). While the Israeli programme is generally considered a success, many questions on past population genetic processes and their impact on future population viability remain.

This case-study is particularly interesting since the genetic consequences of the mixed founder stock have not been explored. Admixture may have improved genetic variability in the population or, alternatively, the introduction of different subspecies may have impaired interbreeding and caused a cryptic population differentiation.

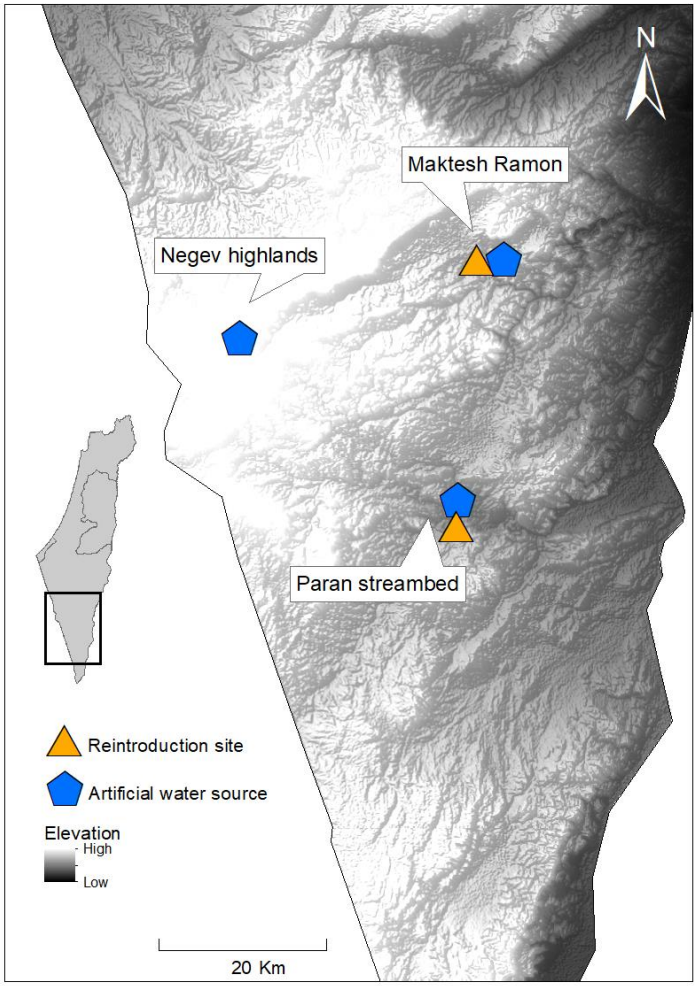


Fig. 1.2 Reintroduction sites and locations of artificial water sources in the Negev desert, Israel.

1.4 Aim and structure of the thesis

The broad aim of the thesis is to use modern high-throughput sequencing methods to investigate the conservation genomics of the reintroduction of the Asiatic wild ass in Israel. Specifically, I will address the following topics:

Chapter 2: SNP discovery in *Equus hemionus ssp.* via ddRAD sequencing. In this chapter I compare different parameter settings and alignment methods to optimise the bioinformatic processing and SNP calling.

Chapter 3: Genomic evaluation of the reintroduction in Israel. Here I, investigate subspecies admixture, measures of genetic diversity, inbreeding and effective population size pre- and post-release.

Chapter 4: Landscape genetics of the Israeli wild ass population. In this chapter I explore the genetic structure and landscape genetics of the reintroduced population.

Chapter 5: Final discussion and concluding remarks. A discussion of the study's findings in a broader context and suggested directions for future research.

2. SNP discovery in *Equus hemionus* via ddRAD sequencing

2.1 Abstract

Restriction site associated DNA sequencing (RADseq) protocols are widely used, fast and effective techniques to generate tens of thousands of SNPs. They have become especially appealing for conservation genetic projects due to their low cost and the fact that they do not require a reference genome. However, RADseq protocols are prone to errors accumulating during library preparation and sequencing and therefore raw sequences require careful bioinformatic processing, which should be optimised for the specific study organism and downstream analysis. For the Asiatic wild ass *Equus hemionus* I used double digest restriction-site associated DNA sequencing, followed by bioinformatic processing in the Stacks pipeline to generate a panel of SNPs suitable for investigating landscape and conservation genomics. I performed a parameter optimisation approach to identify optimal settings for *de novo* assembly of sequencing data. Then, I compared the output of *de novo* assembled loci with loci that were subsequently aligned to the reference genomes of either of two closely related species (domestic horse *E. caballus*, domestic donkey *E. asinus*). Finally, I performed a hierarchical filtering step to create a final data set of robust SNP markers. The optimal parameter settings produced 2,639 shared polymorphic loci while maintaining a low SNP error rate (1.08%). Alignment to either of the reference genomes resulted in a considerable reduction in the number of shared polymorphic loci (50% and 64% reductions in *E. asinus* and *E. caballus,* respectively) and thus, SNPs were called without prior alignment. Low success of reference alignment is likely due to phylogenetic distance between Asiatic wild ass and the other equids. However, high depth of coverage and low SNP error rates indicate robustness of the *de novo* alignment. I conclude with recommendations on how future projects could improve alignment and SNP recovery.

2.2 Introduction

Restriction site associated DNA sequencing (RADseq; Baird et al., 2008) and the different variations of the original protocol, i.e. double digest RADseq (ddRADseq) (Peterson, Weber, Kay, Fisher, & Hoekstra, 2012), ezRADseq (Toonen et al., 2013), and 2b-RAD (Wang, Meyer, Mckay, & Matz, 2012), have become widely used tools for genomic analyses in conservation research (e.g. Dierickx *et al.*, 2015; Martin *et al.*, 2016; Svengren *et al.*, 2017; Grossen *et al.*, 2018; Sovic *et al.*, 2019). Low costs, high

flexibility and the ability to discover tens of thousands of genome wide markers make this method an attractive alternative to whole genome sequencing (Andrews, Good, Miller, Luikart, & Hohenlohe, 2016). Furthermore, RADseq protocols do not require a reference genome, which makes them particularly suitable for studies on wild populations of non-model species, which are predominantly the focus of conservation research (Davey et al., 2011).

In ddRADseq, short DNA fragments that are flanked by two different restriction enzyme recognition sites are sequenced, theoretically resulting in homologous DNA sequences from across the genome in all samples of closely-related populations or species (Davey et al., 2011). The raw sequence reads are assembled into orthologous loci for each individual. Subsequently, these assembled loci are compared across individuals to call SNPs and individual genotypes (Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011). However, like other RADseq methods, ddRADseq is susceptible to errors being introduced during the different stages, from library preparation to sequencing and bioinformatic analysis (O'Leary, Puritz, Willis, Hollenbeck, & Portnoy, 2018). For example, polymorphisms in the restriction enzyme recognition sites can lead to allele drop out, the failure to sequence one of the alleles, resulting in the overestimation of homozygosity (Andrews et al., 2016). Furthermore, duplication errors occurring during PCR steps of the library preparation can lead to both false homozygote and false heterozygote loci. Additionally, during the bioinformatic processing step, errors in the assembly of sequence reads, such as over- or under-splitting of assembled loci (i.e., erroneous splitting of one locus into two or erroneous merging of two different loci into one) can lead to analytical artefacts (O'Leary et al., 2018). To minimise errors in the final data set, the bioinformatic processing of a ddRADseq project from raw sequence reads to a final SNP data set is a crucial and challenging step, which requires great care. Assembly of loci and SNP calling must be performed in a way that reduces potential errors introduced during library preparation and sequencing, while simultaneously minimising the risk of calling false SNPs. Finally, the resulting SNP data sets must undergo stringent filtering to guarantee a robust set of markers and avoid bias in downstream analysis.

Multiple bioinformatic pipelines are available to assist with the processing of raw RAD sequence reads (for example: RADtools, Baxter et al., 2011; PyRAD, Eaton, 2014;

AftrRAD, M. G. Sovic, Fries, & Gibbs, 2015). One frequently used pipeline is Stacks (Catchen et al., 2011; Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013; Hohenlohe, Amish, Catchen, Allendorf, & Luikart, 2011). The pipeline offers two different methods: reference alignment and *de novo* assembly. For species with an available reference genome, sequence reads which align to the same section in the reference genome are assembled into loci for each individual. Loci of all individuals are then compiled into stacks, from which SNPs are called. Alternatively, the *de novo* method can be used to assemble loci without previous alignment to a reference genome. This method consists of three steps performed by different programs within the Stacks pipeline: First, identical sequence reads are assembled into loci within each individual. Then, assembled loci from all individuals are compiled in a catalogue and finally individual loci are matched back against this catalogue. Only matching loci are retained and used for SNP calling.

Throughout the different steps of the *de novo* method, different parameters can be defined to optimise the process. This is a crucial step, as settings of these key parameters can directly impact downstream analyses, such as the results of population differentiation analysis (Díaz-Arce & Rodríguez-Ezpeleta, 2019). However, optimal parameter settings will vary based on genomic characteristics of the study species and downstream analyses, which may differ in sensitivity to missing data and number of samples (Andrews et al., 2016; Díaz-Arce & Rodríguez-Ezpeleta, 2019). Thus, while Stacks provide default parameters, there is no universal template for the analysis, rather the assembly process must be adjusted to fit the individual study system.

Whether to choose the *de novo* or reference alignment method depends predominantly on the availability of a suitable reference genome. The reference alignment approach is advantageous as it reduces the rate of false heterozygotes created by errors during library preparation (Rochette & Catchen, 2017). Furthermore, a comparative study reported lower genotype error rates in the reference alignment method, particularly when sequencing coverage was low (Fountain, Pauli, Reid, Palsbøll, & Peery, 2016). However, low quality or large gaps in the reference genome can impact alignment success and lead to loss of raw reads that cannot be aligned, resulting in markedly fewer assembled loci (Fountain et al. 2016). To exploit the benefits of both methods, Paris et al. (2017) have recommended a combined

approach. The authors suggest an *integrated* method for data processing in Stacks, which aligns *de novo* assembled loci to a reference genome and then reintegrates alignment positions into the catalogue.

After successful loci assembly and SNP calling, the final step in the processing of sequencing data is the application of different criteria to selectively remove SNPs from the data set, known as filtering. Rigorous filtering is important to improve the quality of the data set and remove artefactual SNPs from true variants (O'Leary et al., 2018). Furthermore, like assembly parameters, filtering thresholds can directly impact downstream analyses and thus should be selected carefully (Díaz-Arce & Rodríguez-Ezpeleta, 2019). Most studies include some generic filtering steps, however, if thresholds are not adjusted to suit the unique data set, this may result in the loss of true variants or the introduction of errors in the final data set.

There is currently no reference genome available for the Asiatic wild ass (*Equus hemionus*). However, the genomes of other members of the equid genus, the domestic donkey (*E. asinus*) and domestic horse (*E. caballus*), have been sequenced and are currently available at different stages. The aim of this analysis was to identify the optimal bioinformatic processing steps for the wild ass ddRADseq data. Specifically, the objectives were 1) to identify the optimal parameter settings for the *de novo* method 2) to compare outputs of the *de novo* and *integrated* method using both reference genomes and 3) to perform hierarchical SNP filtering on the output of the optimal method to generate a robust SNP panel suitable for downstream analysis.

2.3 Methods

2.3.1 Sample collection

Founder population

In 1991, prior to the second reintroduction event and three generations after establishment of the breeding core, whole blood samples were collected from 30 individuals in the captive breeding core population (Gueta et al., 2014). Samples were stored frozen (-80°C) in EDTA tubes (BD Vacutainer K2EDTA 18.0mg, Thermo Fisher Scientific, Waltham, MA, USA; Vacuette K3EDTA 3mg, Greiner Bio-One, Kremsmünster, Austria). At the time of my PhD study, 25 samples were still available for DNA extraction (Table 2.1).

Wild population

Blood and tissue samples of the reintroduced population ("wild population") were collected opportunistically during veterinary treatments, fitting of radio collars and from animals killed in traffic accidents. Between 2011-2017 a total of 33 samples were collected from across the population's range (Table 2.1). Whole blood samples were stored in EDTA tubes, tissue samples were either stored untreated in paper bags or in screw-cap tubes in 70% ethanol. All samples were stored frozen (-20°C or -80°C).

Zoo population

Samples of the onager *E.h.onager* and kulan *E.h.kulan* subspecies were obtained from captive individuals from European zoos (Table 2.1). The zoos provided tissue and whole blood samples collected opportunistically during veterinary treatments or from dead individuals. Blood samples were stored in either EDTA or Eppendorf tubes, tissue was stored in screw-cap tubes in 70% ethanol. All samples were stored frozen (-20°C).

2.3.2 DNA extraction

DNA was extracted from samples using commercial silica spin column-based extraction kits (DNeasy Blood and Tissue Kit, Qiagen, Hilden, Germany; GeneJET Genomic DNA Purification Kit, Thermo Fisher Scientific), following the manufacturers protocol. For DNA purification from blood samples 90µl (Qiagen) or 180µl (Thermo Fisher Scientific) of whole blood were used. For extractions from tissue samples, 20mg of tissue were used, which was cleared of hairs and cut finely using a surgical knife prior to lysis. DNA was eluted from each spin column in two consecutive elution steps using 100µl of the elution buffer (Buffer EB, 10mM Tris-Cl, pH 8.5, Qiagen). DNA concentration of eluates were measured using a spectrophotometer (NanoDrop 2000, Thermo Fisher Scientific). Where needed, several extractions were performed of the same sample to achieve a total yield of 1µg of DNA per sample. All eluates of the same sample were combined, and concentrations were measured using a fluorometer (FLUOstar OPTIMA, BMG Labtech, Ortenberg, Germany). Samples were adjusted to a DNA concentration of 25ng/µl through dilution with purified water or concentration using either centrifuge filters (Microcon DNA Fast Flow, Merck, Kenilworth, NJ, USA) or a vacuum concentrator at 60°C (Concentrator plus, Eppendorf, Hamburg, Germany).

Table 2.1 Information on DNA samples collected from the different study populations.

Population	Subspecies	N	Sex	Sample Type	Source
Founder	*E. hemionus ssp.*	25	13F, 12M	Blood	HaiBar Yotvata Reserve, Israel
Wild	*E. hemionus ssp.*	33	11F, 15M, 7 unknown	Blood (N=13) and Tissue (N=20)	Negev desert, Israel
Onager	*E.h.onager*	6	5F, 1M	Blood	Chester Zoo, UK (N=4) Wildlands Adventure Zoo, Emmen, Netherlands (N=2)
Kulan	*E.h.kulan*	15	5F, 6M, 4 unknown	Blood (N=11) and Tissue (N=4)	Nuremberg Zoo, Germany (N=11) Rostock Zoo, Germany (N=4)
Replicates	*E. hemionus ssp.*	7	3F, 4M	Blood (N=4) and Tissue (N=3)	HaiBar Yotvata Reserve (N=2) and Negev desert (N=5), Israel

2.3.3 ddRADseq library preparation and sequencing

The ddRAD sequencing libraries were prepared following the protocol by Peterson et al. (2012) with minor adjustments as detailed below.

Restriction Enzyme digestion

A total of 86 samples (including 7 replicates, Table 2.1) were digested with two restriction enzymes in a single reaction. High fidelity versions of *Eco*RI (R3101S, New England Biolabs, Ipswich, MA, USA thereafter NEB) and *Sbf*I (R3642L, NEB) were used for the digestion. The reaction had a volume of 35µl (0.5µl *Eco*RI; 1µl *Sbf*I; 2.5µl NEBuffer 4 B7004S, NEB; 0.525µg of DNA, purified water to adjust volume). Samples were digested in a thermocycler (DNA Engine Tetrad2, Bio-Rad, Hercules, CA, USA) for 30min at 37˚C, followed by 20min enzyme deactivation at 65˚C and a final cool-down to room temperature.

Adapter Ligation

Custom-made adapter sequences were ligated to each sample. Paired-end compatible P2 adapters (50nM), common to all samples, and P1 adapters (50nM) with a unique 8-base inline barcode for each sample, were ligated in a single reaction (for adapters synthesis see Peterson et al. 2012). Reactions had a volume of 35µl (25µl sample; 1µl NEBuffer 2, B7002S, NEB; 0.3µl 1000mM rATP, E601B, Promega, Madison, WI, USA; 2µl P1 adapter; 6µl P2 adapter; 0.2µl purified water; 0.5µl T4 ligase, M0202L, NEB). Using a thermocycler, adapters were ligated during a 30min reaction at 24˚C followed by 20min enzyme deactivation at 65˚C and cool-down to room temperature.

Size Selection

DNA fragments were manually selected by size using gel electrophoresis. A 2% low melt agarose (AG-LM2, Cambridge Reagents, Hessle, UK) gel was created using 1x lithium borate buffer (10mM, pH 8.5). 3µl of custom internal size standards (300bp and 450bp) and 2.5µl of Blue Loading Buffer (B7703S, NEB) were added to 50µl of sample. The gel was run for 150min at 110V. DNA fragments within the recommended size range for Illumina sequencing (300bp-450bp) were manually cut from the gel using an open UV-transilluminator (UVT-40M, Syngene, Bangalore, India; Illumina, 2020). Size-selected DNA fragments were extracted from gel pieces using a commercial extraction kit, following manufacturer's protocol (Min Elute Gel Extraction Kit, Qiagen).

PCR amplification and bead clean

Eluted DNA was amplified in a PCR reaction (Phusion High-Fidelity PCR Kit, NEB). The total reaction volume was 60µl (30µl Phusion Mix, 14µl purified water, 3µl forward primer, 3µl reverse primer, 10µl template DNA) and reactions were run on a thermocycler (30sec at 98˚C followed by 24 cycles of 10sec at 98˚C, 30sec at 60˚C, 40sec at 72˚C; after completion of the 24 cycles, reactions were held at 72˚C for 5min, followed by a slow cool-down to room temperature). Subsequently, samples were purified using magnetic bead cleaning following the manufacturer's protocol (AMPure XP beads, A63880-2, Beckman Coulter, Indianapolis, IN, USA).

Quantification and pooling of libraries

To accurately quantify DNA concentration of individual libraries, quantitative PCR reactions were performed. For each sample three independent 1,000-fold and 10,000-fold dilutions with sample dilution buffer (10mM Tris pH8.0, 0.05% Tween20) were produced. A no-template control was included and a library quantification kit with 6 DNA size standards was used (KAPA Library Quant Kit, 07960336001, Roche, Basel, Switzerland). Reactions had a volume of 10µl (8µl qPCR MasterMix with primers included, 2µl sample dilution) and were performed in a qPCR thermocycler (QuantStudio 12K Flex, Life Technologies, Carlsbad, CA, USA). The programme consisted of an initial denaturation step for 5min at 95˚C followed by 35 cycles of denaturation for 30sec at 95˚C and annealing/extension for 45sec at 60˚C. Finally, libraries were pooled in equimolar amounts (100nM).

Quality control and sequencing

Fragment size and DNA concentrations of the final library pool were assessed on a Tape Station (4200, Agilent Technologies, Santa Clara, CA, USA) and fluorometer (Qubit, 3.0, Life Technologies). The library pool was sequenced by paired-end sequencing on one lane of an Illumina HiSeq4000 flowcell (San Diego, CA, US). The quality of raw reads was assessed using the FastQC tool (Andrews, 2010). Reads were de-multiplexed and barcodes and Illumina adapters were trimmed using the *process_radtags* script in the Stacks pipeline (Catchen et al., 2013). Simultaneously, reads were subjected to initial quality filtering.

2.3.4 Bioinformatic processing

For this study, a two-step approach was adopted to optimise the bioinformatic processing of the sequencing data in the Stacks pipeline. First, optimal parameter settings for the *de novo* method were identified through parameter trials. Second, using optimal parameter settings, I compared the outputs for the *de novo* method and the *integrated* method for two different reference genomes. In addition, SNP error rates were estimated and compared between the different methods and parameter settings.

2.3.4.1 De novo parameter optimisation

The *ustacks* program within the *de novo* method assembles raw sequence reads into stacks within each individual. The minimum number of reads required to create a stack is defined by the -m parameter. For each individual these stacks are then assembled into putative loci, whereby the maximum nucleotide distance allowed between stacks is defined by the parameter -M (Table 2.2). Stacks with a nucleotide distance below the threshold are merged. Secondary reads (reads which were too few to be assembled into a stack) are then matched against compiled putative loci, allowing for a greater nucleotide distance, defined by the -N parameter. Here I set the parameters to -m3 and -N0. Setting the minimum number of reads to -m3 has been found to be optimal for a broad range of study systems (Paris, Stevens, & Catchen, 2017; Rochette & Catchen, 2017). Since the data in the present study had extremely high coverage, with a mean depth of coverage of 114x per individual (Table 2.3), I discarded all secondary reads by setting -N0. It has been recommended to set -m to a moderate level and exclude secondary reads when coverage is exceptionally high (>40x; Mastretta-Yanes et al. 2014).

The *cstacks* program creates a catalogue of putative loci assembled across individuals. Here the number of mismatches allowed between loci when building the catalogue is defined by the -n parameter (Table 2.2). To identify the optimal setting for the -M and -n parameter, the *de novo* method was run five times with altered parameter settings (M2n2, M3n3, M4n4, M5n5, M6n6). Paris et al. (2017) retrieved best outcomes when setting the -n and -M parameter to the same value or setting -n one iteration either side of -M (n=M, n=M+1, n=M-1). Hence, after the optimal run for n=M was identified, two additional runs were performed testing the variations (n=M+1, n=M-1).

Table 2.2 Parameter settings for the programs within the *de novo* method in Stack (Catchen et al., 2013).

Program	Parameters	Description	Value in parameter trials
ustacks Compiles identical raw sequence reads into stacks	-m	The minimum number of identical raw sequencing reads required for a new stack	3
	-N	The maximum nucleotide distance allowed between stacks and secondary reads	0
	-M	The maximum nucleotide distance allowed between stacks	2-6
cstacks Creates a catalogue of putative loci by compiling stacks identified within individuals. Stacks can be added to the catalogue as new putative loci or merged with existing ones	-n	allowed mismatches between Stacks during catalogue construction	2-6

To compare the performance of runs with different parameter settings I used the r80 method (Paris et al., 2017). This method involves filtering the output using the *populations* program within Stacks and retaining only loci shared among a minimum of 80% of samples. The number of assembled loci, polymorphic loci and SNPs shared among at least 80% of samples were compared between runs. In addition, I also calculated and compared SNP error rates between runs. SNP error rates were calculated using the seven replicate pairs of individuals included in the data set. The number of SNP mismatches within a replicate pair after r80 filtering was divided by the total number of SNPs present in the individual and the mean error rate for all seven replicate pairs was determined.

Due to restricted computing power all parameter trials were performed on the forward reads only. After optimal parameters were identified, the *de novo* method was run again with optimal settings and forward and paired-end reads were merged in the

gstacks program (Catchen et al., 2013). The final output was filtered using the r80 filter and the number of retained loci and SNP error rates was determined.

Table 2.3 Depth of coverage for individual samples as produced by the *ustacks* program within the *de novo* method in the Stacks. Depth of coverage is given for the identified optimal parameter settings (m3N0M4n4). Grey shaded samples were removed prior to filtering due to low coverage.

Individual ID	times coverage	Individual ID	times coverage	Individual ID	times coverage	Individual ID	times coverage	Individual ID	times coverage
4014	6.79	N10	87.49	4029	101.69	N6	130.21	N9	165.1
N4	9.4	F12	87.7	4035	103.41	4022	130.66	6005	167.62
4028	29.79	4037	87.99	N5	104.88	M11	132.64	4023	167.77
M10	33.97	R7	88.16	F6	105.43	6003	133.16	4036	169.55
RK3	34.2	4030	89.84	F1	106.29	6002	134.8	6004	170.57
F10	68.61	4034	90.39	N11	106.4	F7	135.75	N7	179.48
4012	69.36	M18	91.04	R5	106.91	R2	137.14	4011	179.56
N2	73.16	4031	91.54	M1	107.74	6001	139.24	F3	181.59
F11	76.84	R6	92.38	M6	108.77	4038	140.45	C2	228.27
4021	77.8	M5	95.85	4033	110.42	RK2	142.11	E2	248.04
4019	79.01	M15	96.27	M13	110.9	M16	143.16		
F8	81.25	4032	96.42	F4	111.92	N8	145.38		
4015	81.34	4024	96.56	4005	112.92	4017	145.46		
6000	82.93	4027	97.82	F0	114.99	M14	145.85		
F9	84.25	R4	97.96	C4	119.27	E1	147.1		
F2	84.26	C1	98.51	M2	119.67	4020	147.38		
M3	86.28	6006	98.96	RK4	119.79	N1	156.37		
R1	87.01	C3	99.5	4018	128.77	N3	156.79		
RK1	87.4	R3	100.76	4013	129.91	F5	161.84		

2.3.4.2 Comparison of de novo and integrated method

To compare performance between the *de novo* and the *integrated* method, *de novo* assembled loci produced with the optimal parameter conditions for the merged data set were aligned to a reference genome. Since there is no *E. hemionus* reference genome available, the domestic horse (*E. caballus,* NCBI accession GCF_002863925.1) and donkey (*E. asinus,* NCBI accession GCF_001305755.1) genomes were used for alignment in the *integrated* method.

Alignment was performed using the GSNAP program (version 2017-11-15; Wu and Nacu, 2010) with settings recommended by Paris et al. (2017): a maximum of five allowed nucleotide mismatches (-m 5), indel penalty (-i 2) and terminal alignments turned off (--terminal-threshold 10). Subsequently, aligned loci were re-integrated into the *de novo* catalogue of stacks using the *integrate_alignments.py* script distributed with the Stacks pipeline. The output of the *integrated* method was filtered using the r80 filter and number of loci and SNP error rates were determined in the same way as for the *de novo* method output.

2.3.5 SNP filtering

After identifying the optimal parameter settings and method, all SNPs called for this optimal approach were retained (no r80 filter) and subjected to a thorough filtering procedure. O'Leary et al. (2018) have suggested that datasets be divided based on biologically meaningful groups and filtering thresholds adjusted for each division. This approach seems appropriate for the present study since the data comprised distinct populations which are expected to be genetically differentiated. This in turn could lead to the loss of informative markers unique to one of these populations. Therefore, by filtering biologically meaningful groups separately, the number of retained SNPs can be maximised. Furthermore, physical linkage between markers can lead to bias in some downstream analyses such as genetic clustering analysis (Willis et al. 2017). Therefore, to optimise datasets for downstream analyses, three different working datasets were created prior to SNP filtering: A dataset retaining only the first SNP per locus for the complete set of samples (wds1) and for the wild population samples only (wds2) and a third data set containing >1 SNP per locus for the complete set of samples (wds3).

There is a trade-off in SNP filtering procedure between maximising the number of SNPs retained and maximising the number of individuals retained. If filters are set to keep

even rare markers, many individuals for which these markers have not been sequenced will be lost. On the other hand, if the number of individuals is to be maximised, only a small set of common SNPs will be retained. Consequently, less common yet defining SNPs, which may be crucial for downstream analysis, are likely to be lost (O'Leary et al., 2018; Puritz, Hollenbeck, & Gold, 2014). Therefore, to minimise the loss of markers and individuals, I performed multiple filtering steps, alternating between SNP-based and individuals-based filters while slowly increasing the filtering threshold.

First, five individuals with a mean coverage of <35x were removed, as recommended by Catchen et al. (2011; Table 2.3). The datasets were then filtered in the *vcftools* program (Danecek et al. 2011) in six consecutive steps: (1) the minimum minor allele count was set ≥ 3 (--mac 3) to retain only variants for which the minor allele was present in at least three heterozygotes or 1 heterozygote and 1 homozygote individual. (2) The maximum number of missing individuals for a variant was set to 50% (--max-missing 0.5). (3) The percentage of missing data per individual was assessed and individuals with >50% missing data were removed. (4) The minor allele frequency (--maf) threshold was set to 1% for the complete data sets (wds1, wds3). Due to the smaller size of the wild population dataset (wds2), a --maf filter of 1% would have had no effect and so it was increased to 2%, which resulted in a comparable absolute allele frequency filter. (5) SNPs that had been recorded in <80% of individuals were removed (--max-missing 0.80).

2.4 Results

2.4.1 Sequencing

Illumina sequencing produced a total of 803,092,446 raw sequence reads with a mean read length of 150bp. The FastQC tool identified a mean Phred+33 quality score >30 for all bases and hence reads were not trimmed. Quality filtering removed reads containing adapter sequence (1.88%), those with ambiguous barcodes (9.87%), low quality scores (0.1%) and ambiguous RADtags (1.49%). A total of 695,888,624 reads (86.65%) passed initial quality filtering.

2.4.2 Bioinformatic processing

2.4.2.1 De novo parameter optimisation

The number of assembled loci, polymorphic loci and SNPs shared among at least 80% of samples as well as SNP error rates were compared between the 6 runs of the *de novo* method (Table 2.4). Overall error rates were low (<2%) and variation between parameter trials was minimal (Fig. 2.1). Paris et al. (2017) recommended choosing the parameter settings which rendered the highest number of polymorphic loci shared among at least 80% of samples. Setting parameters to M4n4 retrieved the highest number of polymorphic loci (2,371) and was chosen as optimal (Table 2.4). The *de novo* program was then run for the merged forward and paired-end reads with the optimal parameter settings (m3N0M4n4), which produced a slight increase in polymorphic loci to 2,639 and a drop in the mean(±SD) SNP error rate to 1.08±0.31% (Table 2.4).

2.4.2.2 Comparison of de novo and integrated method

The *integrated* method resulted in markedly fewer assembled loci than the *de novo* method (Table 2.4). For the merged data set with optimal *de novo* parameter settings 2,639 polymorphic loci were shared among 80% of samples. This number dropped to 1,326 for loci aligned to the *E. asinus* reference genome, and 949 for loci aligned to the *E. caballus* genome.

2.4.3 SNP filtering

After filtering the complete dataset (wds3) retained 69 individuals and 5,981 SNPs, with an average of 2.43 SNP per locus. The reduced dataset retaining only the first SNP per locus (wds1) contained 2,203 SNPs and the reduced data set of the wild population (wds2) contained 30 individuals and 1,645 SNPs. The numbers of recovered SNPs are low when compared with other ddRADseq studies (Table 2.5).

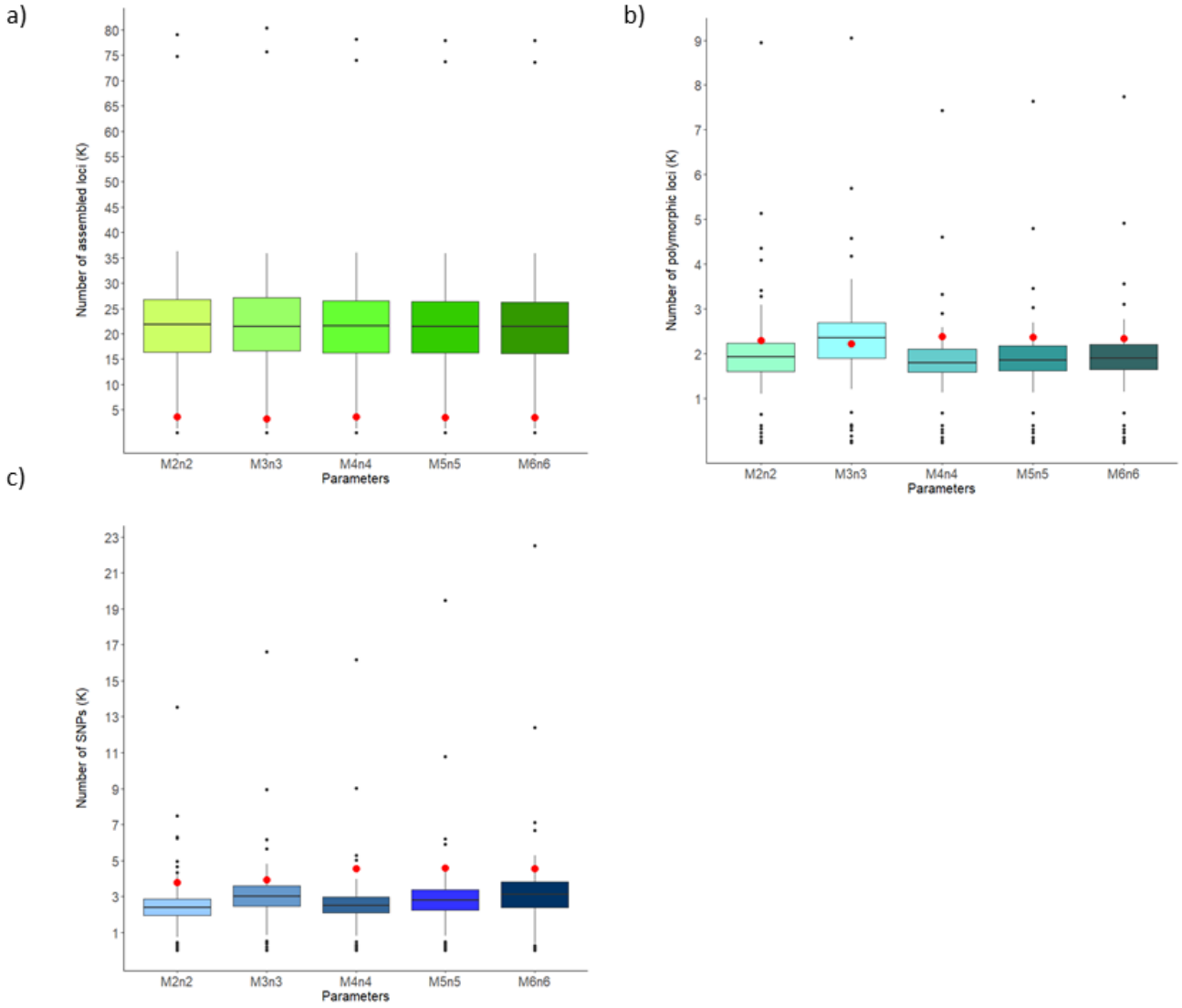


Fig. 2.1 Comparison of *de novo* outputs for different parameter settings. The median (Q1, Q3) number of assembled loci (a), polymorphic loci (b) and SNPs (c) are compared. Red points indicate the number shared among at least 80% of individuals.

Table 2.4 *De novo* outputs for different methods and parameter settings in Stacks. Number of assembled loci, polymorphic loci and SNPs shared among at least 80% of samples and SNP error rates for different parameter settings are presented. a) *de novo* parameter trials on forward reads identify M4n4 as optimal b) subsequent trials of variation of optimal parameter settings on merged paired-end reads identify M4n4 as optimal c) *de novo* assembled loci aligned to one of two reference genomes for optimal parameter settings. Grey shading indicates optimal parameter settings.

a)

***de novo* trial on forward reads**						
	Parameters	**M2n2**	**M3n3**	**M4n4**	**M5n5**	**M6n6**
r80						
	Assembled loci	3562	3182	3549	3529	3500
	Polymorphic loci	2288	2222	2371	2360	2337
	SNPs	3774	3926	4540	4597	4571
SNP error rate						
	Mean	0.0108	0.0146	0.0168	0.0164	0.0161
	SD	0.0038	0.0043	0.0089	0.0087	0.0087

b)

***de novo* trials on merged paired end reads**				
	Parameters	**M4n3**	**M4n4**	**M4n5**
r80				
	Assembled loci	3129	3135	3123
	Polymorphic loci	2628	2639	2625
	SNPs	5945	5981	6006
SNP error rate				
	Mean	0.0107	0.0108	0.0104
	SD	0.0034	0.0031	0.0031

c)

Reference aligned *de novo* reads for			
	Reference genome	***E. asinus***	***E. caballus***
	Parameters	**M4n4**	**M4n4**
r80			
	Assembled loci	1598	1182
	Polymorphic loci	1326	949
	SNPs	2867	1970
SNP error rate			
	Mean	0.0143	0.0144
	SD	0.0050	0.0052

Table 2.5 Comparison of the number of recovered SNPs with other ddRADseq studies. Shown are restriction enzymes used, the length of their recognition sites in parentheses, and the number of SNPs recovered after filtering (1 SNP per locus).

Organism	**SNPs recovered**	**Restriction Enzyme 1**	**Restriction Enzyme 2**	**Reference**
White-rumped shama *Copsychus malabaricus*	18,221	*Eco*RI (6-base)	*Msp*I (4-base)	(Ng et al., 2017)
Bighorn sheep *Ovis canadensis*	17,095	*Eco*RI (6-base)	*Mse*I (4-base)	(Jahner et al., 2019)
White-footed mouse *Peromyscus leucopus*	14,930	*Sph*I (6-base)	*Mlu*CI (4-base)	(Munshi-South, Zolnik, & Harris, 2016)
Caribou *Rangifer tarandus*	6,384	*Sbf*I (8-base)	*Msp*I (4-base)	(Gagnon, Yannic, Perrier, & Côté, 2019)
Pupfish *Cyprinodon diabolis*	4,679	*Sbf*I (8-base)	*Nla*III (4-base)	(Martin et al., 2016)
Asiatic wild ass *Equus hemionus ssp.*	2,203	*Sbf*I (8-base)	*Eco*RI (6-base)	Present study
Snow leopard *Panthera uncia*	511	*Sbf*I (8-base)	*Eco*RI (6-base)	(Janjua et al., 2019)

2.5 Discussion

Bioinformatic processing and SNP filtering are an essential part of reduced representation sequencing methods. To avoid introducing bias due to sequencing errors, raw sequence reads must be processed carefully. Here, I optimised the Stacks pipeline for the wild ass ddRADseq dataset through parameter trials. Overall, the analysis showed that parameter settings had only a small impact on retrieved number of polymorphic loci. Furthermore, SNP error rates were consistently low across different trial runs (<2%). Alignment of *de novo* loci to the horse and donkey reference genomes led to a steep drop in the number of recovered polymorphic loci (and SNPs), hence the *de novo* method was chosen as optimal. The applied stepwise hierarchical filtering approach of the called SNPs minimised the loss of individual samples.

2.5.1 Reference genome alignment

ddRADseq is commonly used for non-model species without an available reference genome. Genomes of closely related species can often serve as reference and many ddRADseq studies reported high success rates with this approach (for example: *Ovis canadensis*/ *Ovis aries*, Jahner *et al.*, 2019; *Cyprinodon diabolis*/ *Cyprinodon* variegatus, Martin *et al.*, 2016; *Salmo trutta/ Salmo salar* and *Aptenodytes patagonicus/ Aptenodytes forsteri,* Paris, Stevens and Catchen, 2017; Table 2.5). However, for the wild ass data set *de novo* assembled loci aligned poorly to either of the reference genomes. Only half of the assembled loci were retained after alignment to the donkey genome and even fewer after alignment to the horse genome. This is consistent with the phylogenetic relatedness of the species: The donkey is believed to have been domesticated from African wild ass *Equus asinus*, and is more closely related to Asiatic wild ass (Kimura et al., 2011). The lineages of African and Asiatic asses are estimated to have split ~1.7 million years ago, while the more distant horse and ass lineages split ~4.5 million years ago (Jónsson et al., 2014).

The results obtained in this study are supported by a large-scale RADseq study which investigated the performance of reference genome alignment for SNP discovery in different species of butterfly fish *Chaetodon sp.*. The authors report that for closely related species the use of the reference genome improved the number of recovered SNPs, however, in less closely related yet congeneric species the reference alignment resulted in a major reduction of recovered SNPs compared to *de novo* assembly (DiBattista et al., 2017). It is possible that the poor alignment success is due to a relatively large genomic distance between the Asiatic wild ass and the domesticated equids. The genus *Equus* has undergone an extremely rapid genome evolution, resulting in great differences between species in the number and structure of chromosomes (horses: 2n=64, donkeys: 2n=62, onagers: 2n=55-56, kulans: 2n=54-55; Eldridge and Blazak, 1976; Ryder and Chemnick, 1990; Trifonov *et al.*, 2008). These differences likely affected the successful alignment of assembled loci.

While the *integrated* method allows inclusion of positional information of the variants, it also acts as a very strong SNP filter (Paris et al., 2017; Rochette & Catchen, 2017). Using a reference genome of a more distantly related species might cause the loss of SNPs on novel or more divergent loci which failed to align. Yet, unmapped sequences

may contain important information of recent divergence and local adaptations (Gouin et al., 2015). Hence, when only the reference genome of a distantly related species is available, a *de novo* approach is more suitable.

In addition to general similarity due to phylogenetic history, the quality of the reference genome can also impact alignment success. The donkey reference genome available at the time of the analysis was only at the scaffold stage and included gaps. Scaffold size and the number and extent of gaps also directly impacts the number of sequences that can be successfully mapped to the genome (Catchen et al., 2013; DiBattista et al., 2017; Huang et al., 2015). A recently published new and improved donkey reference genome, would likely lead to improved alignment and increased number of sequences retained (Waller et al. 2018).

2.5.2 Number of retrieved SNPs

The number of SNPs retrieved in the present study was relatively low when compared with other ddRADseq studies, some of which succeeded recovering >10,000s SNPs (Table 2.5). This could be due to a species-specific low level of polymorphism (Paris et al., 2017). However, considering that 84% of the assembled loci were polymorphic (Table 2.4), the number of SNPs does not seem to be impacted by low rates of polymorphism, but rather overall low number of assembled loci. ddRADseq allows control over the level of genomic representation through selection of different combinations of restriction enzymes. Cutting frequency of restriction enzymes directly controls the number of loci in the library preparation (Davey et al., 2011; Peterson et al., 2012). The restriction enzymes used in this study were *Eco*RI and *Sbf*I, which have 6- and 8-base recognition sites, respectively (Table 2.5). However, other combinations of restriction enzymes, particularly with shorter recognition sites, would likely have rendered overall more cut sites, hence increasing genome representation and resulting in more assembled loci and SNPs.

Another factor impacting retrieved number of SNPs might be the condition of the DNA samples. Many of the samples analysed in this study have been stored for extended periods of time (>20 years) with occasional defrosting and refreezing (Bar-David pers. comm.). This likely caused degradation of the DNA and further reduced sequencing success and overall number of unique raw sequences (Graham et al., 2015).

2.5.3 Summary

The bioinformatic analysis showed low alignment success of Asiatic wild ass ddRADseq data to the horse and donkey reference genomes, likely due to phylogenetic distance between the species. The *de novo* approach produced markedly more assembled loci, however, the final number of SNPs retained was relatively low (2,203). The analysis could be improved by using higher quality reference genomes to increase alignment success. Additionally, in silico enzyme digestion trials could be used to identify optimal levels of genome representation and increase the number of discovered SNPs. Nonetheless, *de novo* parameter trials successfully identified the optimal parameter settings that maximised polymorphic loci while retaining low SNP error rates and producing high read coverage.

3. Genomic evaluation of the reintroduction in Israel

3.1 Abstract

Reintroductions are a powerful tool for the recovery of endangered species. However, their success is strongly influenced by the genetic diversity of the reintroduced population. Chances of long-term persistence can be improved by mixing individuals from different sources, thereby maximising the genetic diversity in the founders. However, a very diverse group of founders could also suffer from outbreeding depression or unsuccessful admixture due to behavioural or genetic barriers. For the reintroduction of Asiatic wild ass *Equus hemionus* ssp. in Israel, a breeding core was created from individuals of two different subspecies (*E.h.onager, E.h.kulan*). Today the population comprises approximately 300 individuals and displays no signs of outbreeding depression. The aim of this study was a population genomic evaluation of this reintroduction, including assessment of subspecies hybridisation and the potential effects on retained heterozygosity. I investigated subspecies admixture using maximum likelihood methods and analyses of genetic structure and spatial autocorrelation. Further I examined heterozygosity and effective population size in the breeding core and in the current wild population. Both populations displayed high levels of admixture, consistent with a significant heterozygote excess in the breeding core. Retained heterozygosity in the wild population was relatively high compared with similar reintroductions. However, the inbreeding and variance effective population size estimates were low. The study indicates no barriers to admixture between the subspecies. Further, results suggest that the reintroduction led to greater diversity from mixing individuals of different subspecies. Nonetheless, continued management will be necessary to increase effective population size and enhance chances of long-term success.

3.2 Introduction

Reintroductions are an important and powerful tool in the conservation and recovery of endangered species (Armstrong & Seddon, 2008). However, the goal of establishing a stable, self-sustaining population is strongly dependent on the genetic makeup of the population (Seddon & Armstrong, 2016). This is because of the small number of founder individuals, inherent to most reintroduced populations. Small populations face an increased risk of inbreeding and associated loss of genetic diversity (Frankham, Ballou, et al., 2002). Additionally, founder effects and genetic bottlenecks experienced

during the establishment phase further increase the loss of genetic diversity due to random genetic drift (Frankham, Ballou, et al., 2002). Consequently, reintroduced populations often display low levels of genetic diversity, which can result in reduced fitness and adaptive ability of the population (Seddon & Armstrong, 2016).

For the long-term persistence of a reintroduced population, genetic factors are critical. However, in many reintroductions the newly established population displays lower levels of heterozygosity than the source population and practitioners should aim to minimise this loss of genetic diversity (De Barba et al., 2010; Grossen et al., 2018; Mock, Latch, & Rhodes, 2004; Williams, Serfass, Cogan, & Rhodes, 2002). Particularly, extended periods of slow population growth during the establishment phase can result in strong genetic drift and the loss of genetic diversity. Therefore, it is important to increase effective population size quickly after release (Seddon & Armstrong, 2016).

Capturing sufficient genetic diversity in the founder individuals is paramount. For many endangered species, potential source populations are small and have a history of bottlenecks, hence the number of individuals available for translocation is often limited. Therefore, multiple authors have recommended that individuals should be selected to maximise genetic diversity. Specifically, by mixing individuals from different source populations genetic diversity can be increased even when the total number of individuals is small (Grossen et al., 2018; Jahner et al., 2019; Moodley & Harley, 2005; Weeks et al., 2015). However, other studies have highlighted the risks of outbreeding when introducing individuals from different source populations (Edmands, 2007; Huff, Miller, Chizinski, & Vondracek, 2011). If founders stem from different geographical or ecological regions, individuals may have developed local adaptations (Templeton et al., 1986). Population admixture is expected to break up these coadaptations, resulting in reduced fitness of the hybrid descendants, especially in later generations (Tallmon, Luikart, & Waples, 2004). An example of outbreeding depression has been described in a reintroduced population of Arabian oryx. A captive breeding programme was based on individuals from different source populations, the offspring of which have been reintroduced into a sanctuary. Despite population growth after release, genetic monitoring revealed a significant negative effect of outbreeding on juvenile survival (Marshall & Spalton, 2000).

Another potential risk when introducing individuals from different sources, is that of a barrier to admixture. Genetic incompatibility or behavioural differences may prevent successful interbreeding between individuals of the different sources (Gottsberger & Mayer, 2019; Rieseberg, Whitton, & Gardner, 1999). Complete or partial admixture barriers have been reported between different species, subspecies and populations of the same species (Gottsberger & Mayer, 2019; López-Rull, Lifshitz, Macías Garcia, Graves, & Torres, 2016; Soland-Reckeweg, Heckel, Neumann, Fluri, & Excoffier, 2009). For example, in brown boobies (*Sula leucogaster)* females actively selected against males of a different colour morph, thereby preventing hybridisation of different genetic clusters within the same species (López-Rull et al., 2016). Admixture barriers can also be influenced by environmental factors, as observed in two sympatric species of stickleback (*Gasterosteus sp.*). While the two species hybridised in the lab, differences in the selection of microhabitats by male hybrids prevented backcrossing with females of the parental species in the wild (Vamosi & Schluter, 1999). While a barrier to admixture prevents the emergence of a hybrid swarm of individuals with poor fitness, it could still impact the success of a reintroduction by effectively creating two cryptic populations of smaller size and resulting increased extinctions risk.

The reintroduction of the Asiatic wild ass in Israel is an interesting study system to investigate the impact of mixed source populations on a reintroduction. A captive breeding core was established from 11 individuals of two different subspecies, the Iranian onager *Equus hemionus onager* and the Turkmen kulan *E.h.kulan* (for details see chapter 1). Three wild ass generations later (7.5 years generation time, Ransom *et al.*, 2016), 38 descendants of this breeding core were released into the wild. The reintroduced population has rapidly increased in size and expanded its range across a large geographical area and habitat gradient in Southern Israel (Gueta et al., 2014). No decrease in reproductive output or individual fitness which could suggest outbreeding depression has been observed and the reintroduction is considered a success at the current stage (Gueta, Templeton and Bar-David, 2014; Bar-David, personal communication). However, it is not known whether the two subspecies did interbreed or whether genetic or behavioural barriers have led to distinct cryptic populations. A previous analysis on mitochondrial DNA suggested spatial differences in subspecies-specific haplotype distributions (Gueta et al., 2014). It is possible, that differences in

habitat selection between parental subspecies or between hybrids and parental subspecies could prevent complete admixture. Alternatively, if admixture between the subspecies occurred this may have increased genetic diversity and thereby benefited the established population.

The aim of this analysis was to test whether the two subspecies hybridised and whether the apparent reintroduction success occurred due to this hybridisation or despite it. First, I tested whether the two subspecies hybridised and to what extent. Next, I investigated population genomic parameters including effective population size estimates and heterozygosity levels in the breeding core and current wild population, to evaluate the genetic status of the populations before and after release. Finally, I attempted to compare what heterozygosity levels would be in the wild population under different reintroduction scenarios using coalescent simulations.

3.3 Methods

3.3.1 Admixture of the subspecies

To increase the power of the subspecies admixture analysis the complete data set with >1 SNP per locus was used (total N=5,581).

3.3.1.1 Estimation of individual hybrid indices and pairwise fixation indices

The outgroup samples (kulan N=9, onagers N=5) were used to identify diagnostic SNPs, which are fixed for opposite alleles in the two subspecies, to calculate hybrid indices for each individual and investigate levels of admixture in the Israeli population. Out of the total of 5,581 SNPs, 4 were ungenotyped in all onagers and were removed from the analysis. Of the remaining 5,577 SNPs 28 were found to be fixed for different bases in the two subspecies populations. Individual hybrid indices were calculated based on the proportion of alleles inherited from either subspecies. However, due to the extremely low percentage of diagnostic SNPs in the dataset (0.5%) individual hybrid indices were also calculated using all SNPs and maximum likelihood methods implemented in the R package *introgress* (R 3.5.3 R Development core team 2008; Gompert & Buerkle, 2010). This approach estimates parental allele frequencies and calculates individual hybrid indices using maximum likelihood methods and 95% confidence intervals. The program can account for uncertainty of the ancestry of

alleles when parental populations share alleles. The onager and kulan samples from the zoo populations were set as parental populations. Allele ancestry information was then used to calculate individual hybrid indices as the proportion kulan ancestry based on genotypes. Furthermore, pairwise fixation indices (*Fst*) were calculated for either parental population and the combined Israeli populations (founder & wild) using the *hierfstat* R package (Goudet, 2005). Pairwise *Fst* values were tested for significant deviation from zero using the function *boot.ppfst* and 10,000 bootstrap permutations.

3.3.1.2 Spatial autocorrelation based on hybrid indices

To investigate potential environmental barriers to admixture, the estimated hybrid indices of the wild population were tested for spatial autocorrelation in ArcGIS 10.0 (Environmental Systems Research Institute, 2011). A fixed distance threshold of 3,400m with row standardisation was used, which ensured that each individual had at least one neighbour.

3.3.1.3 Genetic structure

The admixture of the two subspecies was further investigated by analysing the genetic structure. The data set was initially explored using principal component analysis (PCA), which fits orthogonal principal components (PCs) that summarise overall variability between individuals. Subsequently, genetic structure was investigated in more detail using two different approaches: A discriminant analysis of principal components (DAPC) and a Bayesian clustering analysis implemented in the program STRUCTURE (Pritchard, Stephens, & Donnelly, 2000).

The DAPC is a multivariate approach which performs a PCA in a first step and then performs a discriminant function analysis (DFA) on a subset of the PCs. Unlike the PCA, the DFA fits orthogonal discriminant functions that maximise between group relative to within-group variation. Therefore, it is better suited to describing the relationship between different genetic groups (Jombart, Devillard, & Balloux, 2010). If the initial number of clusters is unknown, a K-means clustering approach is applied to test models for different numbers of clusters (K) in the data. The best supported model is identified using the Bayesian information criterion (BIC), where the lowest BIC is preferred.

The PCA and DAPC were performed in the *adegenet* R package (Jombart, 2008) with the first 15 PCs retained in both analyses accounting for 52% of the total variance. STRUCTURE was run with the admixture model and correlated allele frequencies. The correlated allele frequency model was chosen, despite potentially discrete populations of the subspecies, since it provides greater power in detecting closely related clusters while giving the same results in the absence of such a correlation (Porras-Hurtado et al., 2013). STRUCTURE was run for K=1-10, with 10 repetitions for each K. The runs were performed with $1x10^6$ iterations of the Markov Chain Monte Carlo (MCMC) chain preceded by $1x10^5$ burn-in iterations. STRUCTURE outputs were analysed for optimal value of K using the log likelihood method (Pritchard et al., 2000) and the Evanno method (Evanno, Regnaut, & Goudet, 2005) in the online version of STRUCTURE HARVESTER (Earl & vonHoldt, 2012). Pritchard et al. (2000) suggest that the value of K which maximizes the model log likelihood Ln P(D) is optimal. However, Ln P(D) often plateaus or continues to increase after reaching the optimal K-value and Evanno et al. (2005) proposed an improved method to estimate the optimal K based on the second order rate of change of the likelihood function.

3.3.2 Heterozygosity and inbreeding coefficient

To avoid any bias due to obvious linkage disequilibrium between SNPs, the reduced data set with only 1 SNP per locus (total N=2,203) was used for this analysis. Expected (*He*) and observed heterozygosity (*Ho*) across loci and the mean inbreeding coefficient across loci for each population were calculated in the *pegas* R package (Paradis et al. 2010). Individual heterozygosity was expressed as the proportion of heterozygote markers for each individual.

3.3.3 Effective population sizes

Effective population sizes were estimated for the wild population using the program NeEstimator V2 (Do et al., 2014). The variance effective population size (*Nev*) refers to the size of an ideal population which displays the same sampling variance in allele frequencies as the focal population and can be estimated using samples of the population from two different time points (Nei & Tajima, 1981).

Inbreeding effective population size (*Nef*) describes the size of an ideal population with the same probability of alleles being identical by decent as the population in question. Since deviations from the ideal population do not influence all evolutionary measures

in the same way, estimates of variance effective size and inbreeding effective size are expected to differ (Templeton, 2018). In fact, by comparing inbreeding and variance effective size of a population, a deeper understanding of the demographic processes can be gained.

To estimate *Nev* the temporal method (Waples, 1989) was applied, with the founder and the wild population used as two samples of the same population at generation 0 and generation 3. The method assumes genetic drift to be the sole cause for shifts in allele frequencies over time. Standardised variance in allele frequencies were computed using the method described by Pollak (1983). The generation time was determined based on the years of sample collection (founder: 1991; wild: 2011-2017) and a species generation time of 7.5 years (Ransom et al., 2016). Additionally, *Nef* was estimated for both the founder and the wild population, using the linkage disequilibrium method. This method is a single-sample estimator and is based on the observed linkage disequilibrium between unlinked loci (Waples & Do, 2008). Both methods were run with the lowest allele frequency set to 0.02, to avoid bias in estimates due to low frequency alleles. Jacknifed 95% confidence intervals (CIs) were calculated for both estimates of *Ne*, as recommend for larger numbers of markers (Do et al., 2014).

3.3.4 Heterozygosity simulations under alternative reintroduction scenarios

To investigate whether genetic diversity in the wild population was affected by the hybridisation of subspecies during establishment of the breeding core, genomic data were simulated for different reintroduction scenarios using the program fastsimcoal2 (Excoffier, Dupanloup, Huerta-Sánchez, Sousa, & Foll, 2013). This program implements a continuous-time coalescent simulator which simulates the site frequency spectrum (SFS) of a population for a given demographic scenario.

The SFS is influenced by past demographic processes, for example, populations which underwent a genetic bottleneck will display a deficiency in rare alleles (Gattepaille, Jakobsson, & Blum, 2013; Stoffel et al., 2018). Inferences about the population's demographic history can be made by comparison of the simulated and the observed SFS based on empirical data.

In the case of the reintroduction in Israel, detailed records of the demographic history are available, which can be used to create a reference scenario. Subsequently, changes can be made to this reference to simulate the SFS under alternative reintroduction scenarios. Finally, heterozygosity levels can be obtained from the simulated SFS and compared with empirical data. In this analysis the first step was to recreate the observed SFS through simulations, using a demographic model based on historical records. The demographic history of the Israeli population included an admixture event and two changes in populations size, during the creation of the breeding core and during the release into the wild (here referred to as bottleneck1 and bottleneck2, respectively) (Fig. 3.1, Appendix A, Table A4). This reference model was built in a stepwise approach, gradually increasing model complexity by adding historic events over five different models (Table 3.1).

For each model, specific parameters were simulated (Table 3.1). fastsimcoal2 first selects initial parameters at random from a user-defined search range. These search ranges were defined based on historic record and in the absence of detailed records a sensitivity analysis was performed (Appendix A, Table A4). The program then optimises parameters using the Brent algorithm which selects parameters that maximise the likelihood estimate for the observed SFS under the respective demographic model (Excoffier et al., 2013). Simulations were run with a mutation rate of 7.242e-9 per site per generation as described for the domestic horse (*Equus caballus;* Orlando et al., 2013) and a generation time of 7.5 years. Timing of demographic events were provided in model input and based on historic records (Appendix A, Table A4). Since model fit was not improved by including a population growth rate, and to simplify the model, the growth rate was set to 0 (Appendix A, Table A4).

For each model 10 repetitive runs of 40 ECM optimisation cycles, each with 100,000 simulations, were performed. Optimisation cycles use the ECM algorithm which in turn maximises one parameter of the model while keeping the other parameters at the value that was last estimated (Excoffier & Foll, 2011; Meng & Rubin, 1993). Runs with the highest maximised log likelihood were selected and model fit was assessed visiallu and by comparison of the maximised log likelihood estimates for the simulated and observed SFS. The analysis was based on the wild population dataset, with adjusted filtering steps (no -maf/-mac filters, see chapter 2) to avoid loss of information. All

other filtering steps were retained. The resulting data set contained 30 individuals and 2,058 SNPs. Since the ancestral allele state was not known the minor allele site frequency spectrum was generated using the software Arlequin ver. 3.5.2.2 (Excoffier & Lischer, 2010).

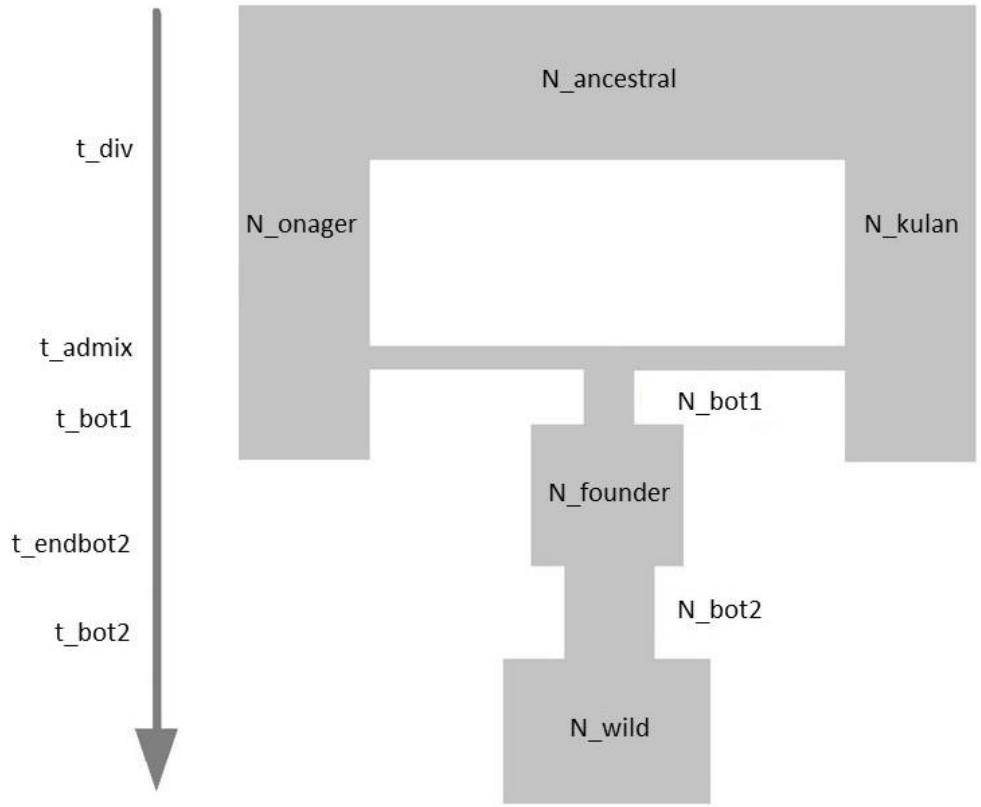


Fig. 3.1 Schematic representation of the demographic history of the Asiatic wild ass *E. hemionus ssp.* population in Israel. The captive breeding core was created from individuals of two subspecies, representing the first bottleneck and simultaneous admixture event. Two generations later, a subset of individuals was released resulting in the second bottleneck. Population sizes at different points in time, estimated using fastsimcoal2, are displayed. Timing of events was based on historical records (Appendix A, Table A4).

Table 3.1 Five different demographic models increasing in complexity and their simulated parameters. Model 5 resembles the demographic history of the Israeli Asiatic wild ass population, including 2 bottlenecks and an admixture event.

Model	Description	Parameters simulated
1	Single population of constant size (null)	N_wild
2	Single population, 1 bottleneck (release into the wild)	N_wild, N_bot2, N_founder
3	Single population, 2 bottlenecks (release into the wild and creation of the breeding core)	N_wild, N_bot2, N_founder, N_bot1
4	Hybrid population, created by admixture of onager and kulan population	N_wild, N_kulan, N_onager, N_ancestral
5	Hybrid population, created by admixture of onager and kulan population, 2 bottlenecks (release into the wild and creation of the breeding core)	N_wild, N_bot2, N_founder, N_bot1, N_kulan, N_onager, N_ancestral

3.4 Results

3.4.1 Admixture of the subspecies

3.4.1.1 Individual hybrid indices and pairwise fixation indices

Mean hybrid index, expressed as proportion of *E.h.kulan* ancestry, differed significantly between estimates based on diagnostic SNPs (Mean=0.504, SD=0.101) and those obtained by maximum likelihood using all SNPs (Mean=0.442, SD=0.055; paired samples Student's t-test: $t(54)=-3.520$, $p<0.001$, Fig. 3.2). Identified diagnostic SNPs were few and based on a small number of samples taken from zoo populations, which have likely experienced genetic drift since foundation and may not represent subspecies allele frequencies correctly. Furthermore, their locations in the genome are not known and the diagnostic SNPs could be clumped and poorly represent genome-wide admixture levels. Therefore, subsequent analyses used hybrid indices estimated by maximum likelihood methods.

The founder and the wild populations displayed high levels of individual admixture (Fig. 3.3). There was no significant difference in the mean hybrid indices, between the founder (Mean=0.454, SD=0.056) and the wild population (Mean=0.433, SD=0.054; independent samples Student's t-test: $t(50.66)=1.421$, $p=0.162$). These proportions of admixture were consistent with pairwise *Fst* values, which were all significant and identified higher level of differentiation between the Israeli population and the *E.h.kulan* population (Table 3.2).

The presence of Hardy-Weinberg equilibrium (HWE) outlier loci in the dataset, which potentially indicate sequencing error, could affected the results of this analysis. Repeating the analysis with prior removal of Hardy-Weinberg equilibrium (HWE) outlier loci, affected the results. Nevertheless, these differences made no major qualitative difference to the biological interpretation of the results (see Appendix A, Table A1, Table A2). Furthermore, in the case of an admixed populations with non-random mating HWE outlier loci can be a result of large allele frequency differences between the subspecies and therefore characteristic of the populations demographic history (Choudhry et al., 2006).

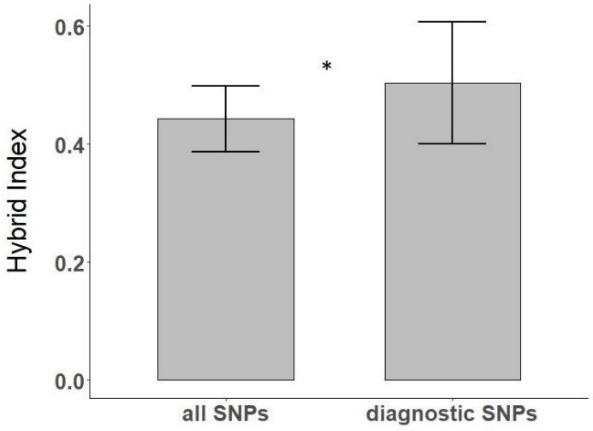


Fig. 3.2 Hybrid index as proportion *E.h.kulan* ancestry. The mean (±SD) hybrid index was significantly larger when based only on diagnostic SNPs (0.504 ±0.101) than when based on all SNPs and maximum likelihood methods (0.442 ±0.055; paired samples Student's t-test: $t(54)=-3.520$, $p<0.001$).

Table 3.2 Weir & Cockerham's pairwise *Fst* values between the parental populations and the combined Israeli population (founder and wild). All values differed significantly from 0 (10,000 bootstrap permutations, p<0.05).

	pairwise *Fst*
Kulan - Onager	0.278
Kulan - Israeli	0.207
Onager - Israeli	0.157

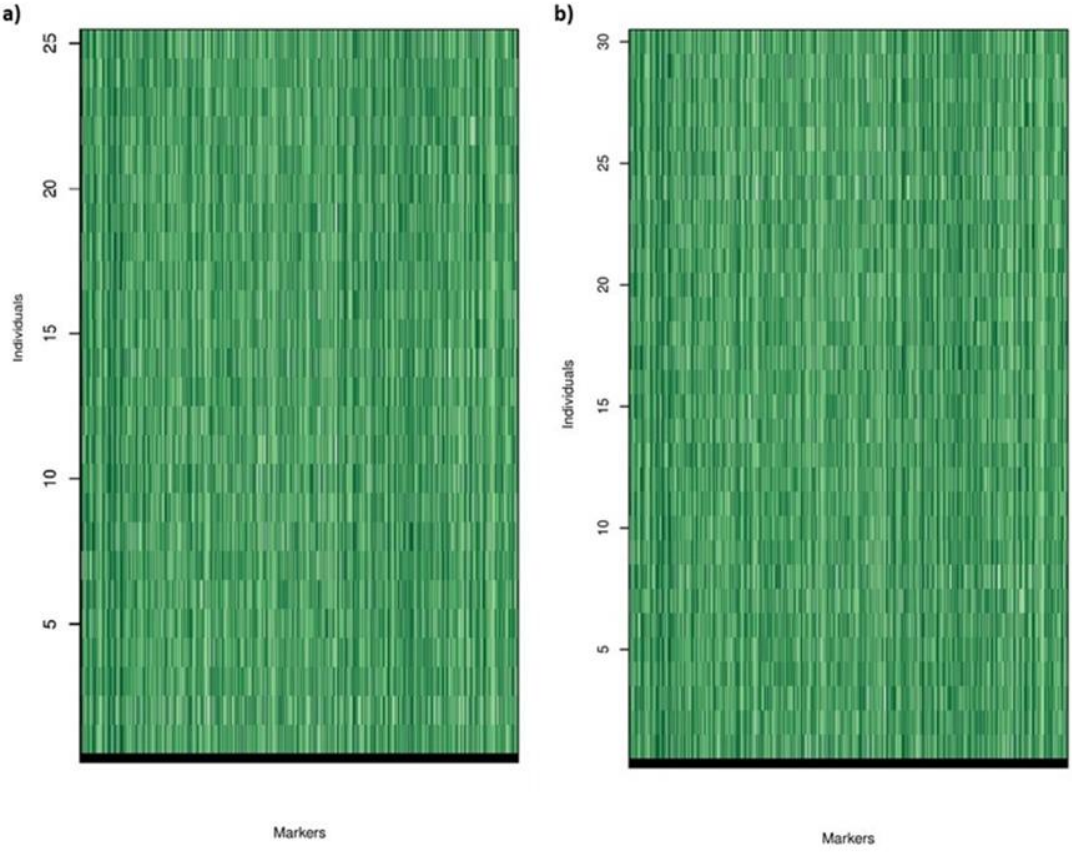


Fig. 3.3 Admixture plots produced by the *introgress* R package. Levels of admixture in the (a) founder and (b) wild population based individual hybrid indices (proportion *E.h.kulan* ancestry). Hybrid indices were estimated using maximum likelihood methods and 5,577 SNPs. Individuals are represented as horizontal bars; SNPs are vertical lines. Genotypes are colour-coded: Dark green=homozygosity for *E.h.onager*, bright green=homozygosity for *E.h.kulan*, intermediate green=heterozygosity, white=missing data.

3.4.1.2 Spatial autocorrelation based on hybrid indices

The spatial analysis revealed no signs of spatial autocorrelation between individuals based on hybrid indices (Moran's *I*=0.042, *z*=1.574, p=0.115, Fig. 3.4).

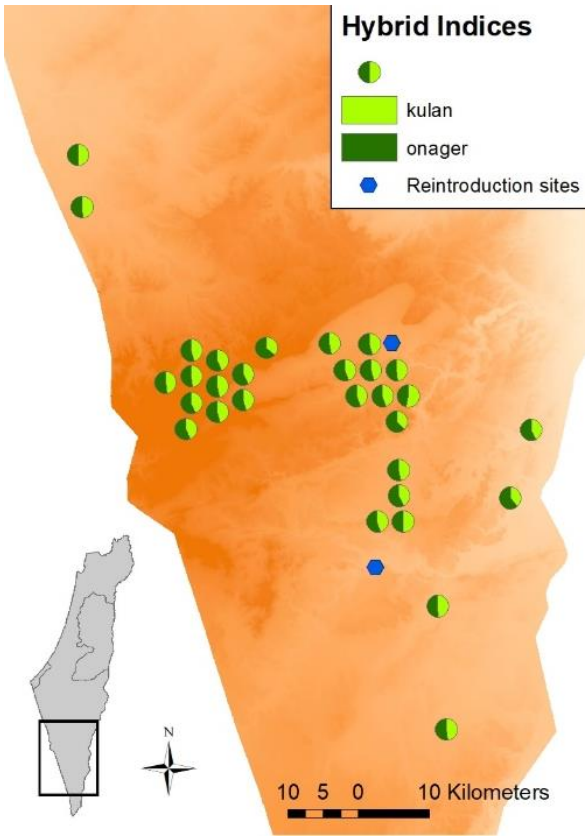


Fig. 3.4 Spatial distribution of individual samples from the wild population. Pie charts represent individual sampling locations and indicate proportion ancestry from each parental subspecies. Background shading displays elevation gradient, with darker colour representing higher elevation.

3.4.1.3 Genetic structure

In the genetic structure analysis, visualization of the first two PCs indicated the differentiation of the data into four clusters, consistent with the original populations. The two zoo populations could be identified distinctly while there was some overlap between the founder and wild population. Specifically, PC1(11.03% of total variance) separated the kulans from the other samples, while there was some overlap between the onagers, founder and wild population along PC2(6.26%). This was resolved along PC3(4.66%) which separated the onager, founder and wild population (Fig. 3.5).

In the DAPC K=3 was identified as the optimal number of clusters, based on the BIC value (Fig. 3.6). The DAPC separated kulans and onagers into distinct clusters, while the founder and wild population made up a third joint cluster (Fig. 3.7). The Israeli cluster was separated from the kulans but not the onagers by the first discriminant function (71.40% of total variance). Only the second discriminant function (28.60%) set apart onagers from the Israeli population. All individuals showed assignment probabilities of 1 (Fig. 3.8). When setting K=4 the DAPC also distinguished between founder and wild population in a third and fourth cluster, however, seven of the Israeli individuals were assigned to the wrong cluster (Fig. 3.9). The founder and wild populations clustered closely together with some overlap (Fig. 3.10). All but two individuals had assignment probabilities of 1 (Fig. 3.9).

In the Bayesian clustering approach using the programme STRUCTURE, the Evanno method identified a clear peak in delta K for K=3 (Delta(K)=1,297.90) and a second smaller peak for K=8 (Delta(K)=312.32; Fig. 3.11). There was very little variation in the mean Ln P(D) between runs, however, a slight increase could be observed until K=3, which is consistent with the results of the Evanno method. Individual assignment to evolutionary clusters inferred by STRUCTURE were compared for different values of K (Fig. 3.12). Kulans were most distinct and separated from the other populations at K=2. In contrast, onagers showed more similarity with the Israeli populations and separated only at K=6. Higher values of K added little information with respect to between group variation but increased individual admixture levels in the founder and wild populations. Furthermore, for K≥4 a distinct cluster of 4 individuals (F4, F5, F6, F7) within the founder population became apparent, which displayed admixture levels

more similar to the wild than the founder population. The same 4 individuals were also falsely assigned to the wild population by the DAPC analysis (Fig. 3.9).

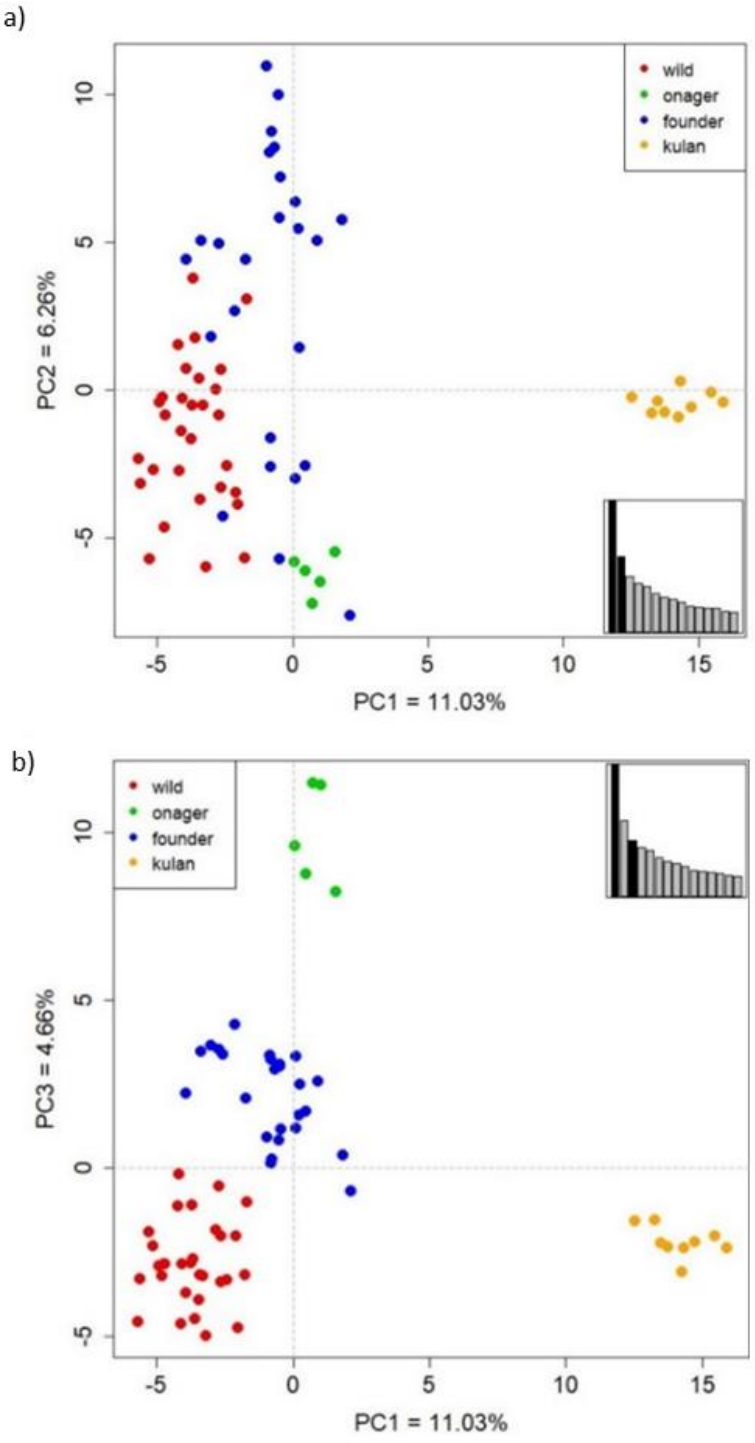


Fig. 3.5 Principal component analysis (PCA) plot displaying variation between the four populations. The analysis was based on the first 15 principal components retaining 52.44 % of the total variance. a) Kulans are separated from the other populations along PC1 (11.03%) on the x-axis. Founder and wild populations overlap along PC2 (6.26%) on the y-axis. b) Founder, wild and onager populations are differentiated by PC3 (4.66%) on the y-axis. Insets display eigenvalues of principal components with the eigenvalues used in the plot marked in black.

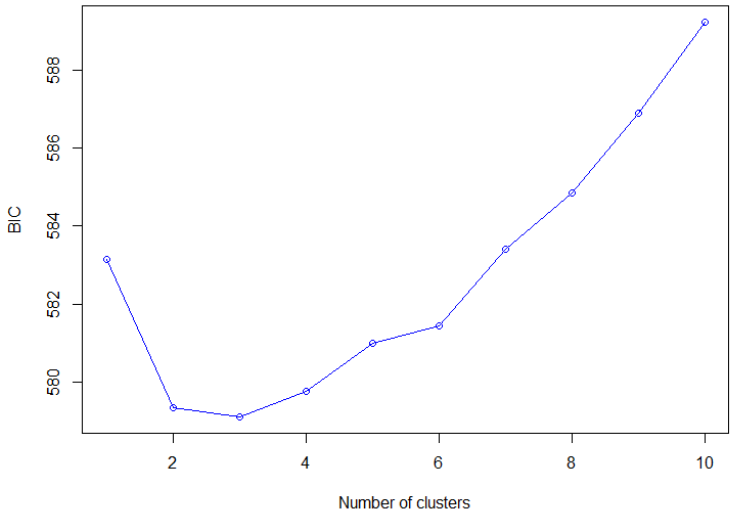


Fig. 3.6 K-means clustering approach of the discriminant analysis of principle components (DAPC). Displayed is the Bayesian Information Criterion (BIC) for different numbers of clusters. A minimum value is reached at three genetic clusters (K=3) indicting the optimal number of clusters.

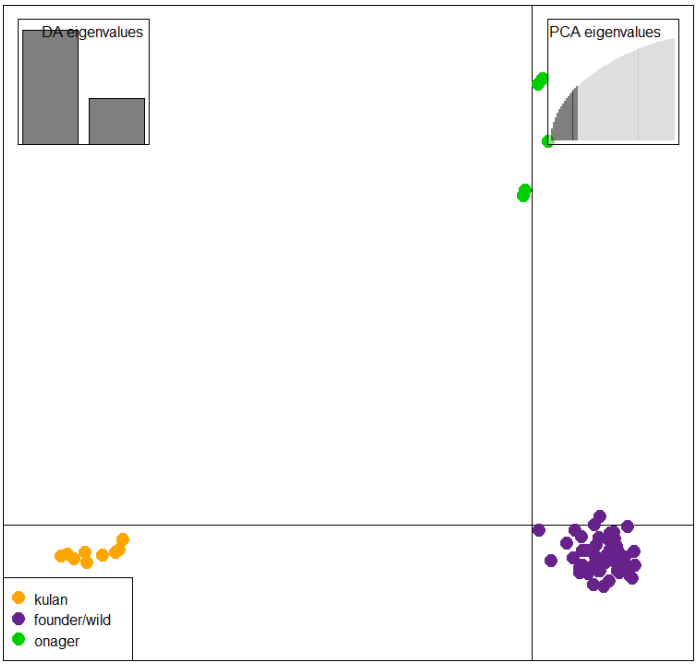


Fig. 3.7 Discriminant analysis of principal components (DAPC) plot for three genetic clusters (K=3). Hybrid individuals (founder and wild) cluster together and are separated from kulans and onagers along the x-axis (DA1) and y-axis (DA2), respectively. The analysis was based on the first 15 principal component analysis (PCA) components which explain 52.40% of the total variance and 2 discriminant analysis (DA) axes retaining all of this variance (DA1=71.40% and DA2=28.60%). Top right inset displays the eigenvalues of principal components of the PCA with dark grey indicating the eigenvalues of the retained principal components. The top left inset displays the eigenvalues of the two retained DAs of the DAPC.

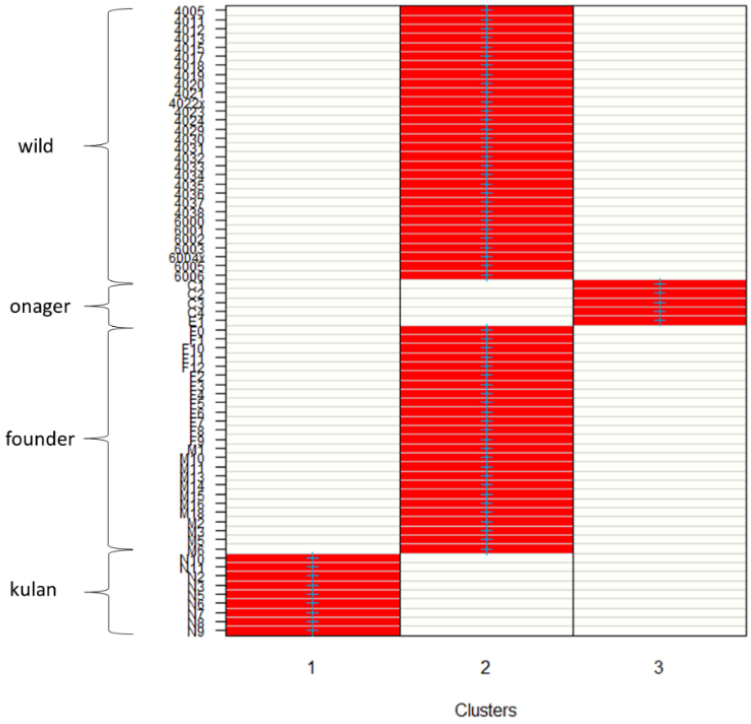


Fig. 3.8 Individual assignment to clusters by the discriminant analysis of principal components (DAPC) for three genetic clusters (K=3). Horizontal bars display individuals arranged by source population. Colours indicate assignment probability to each cluster (red=1, white=0).

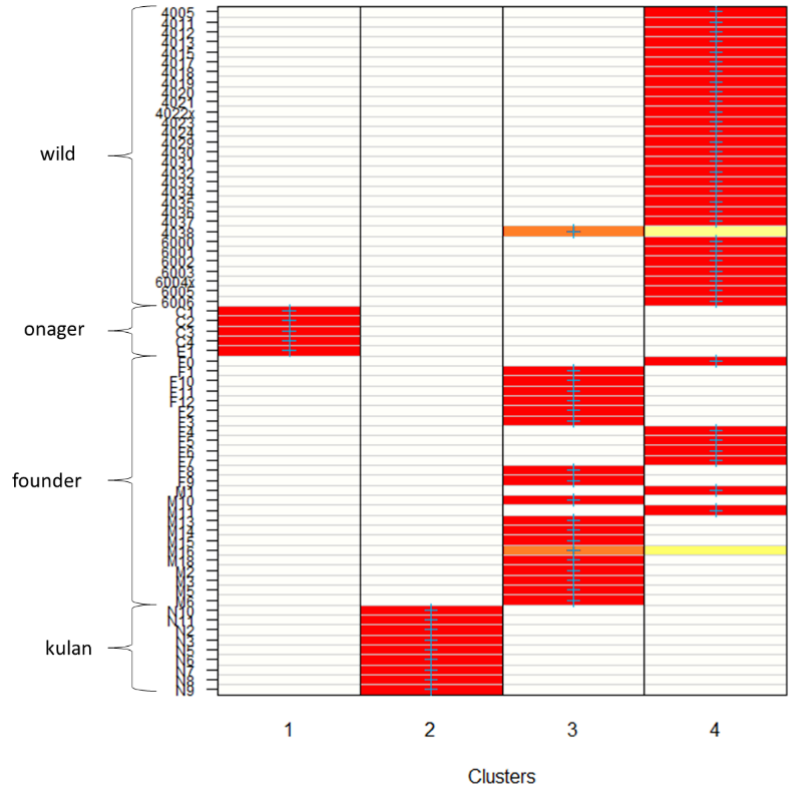


Fig. 3.9 Individual assignment to clusters by the discriminant analysis of principal components (DAPC) for four genetic clusters (K=4). Horizontal bars display individuals arranged by source

population. Colours indicate assignment probability to each cluster (red=1, orange=0.75, yellow=0.25, white=0).

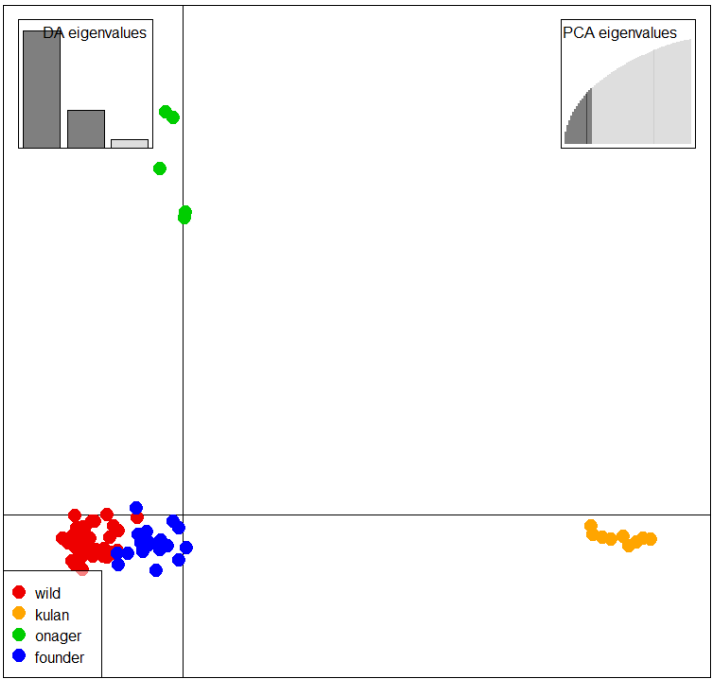


Fig. 3.10 Discriminant analysis of principal components (DAPC) plot for four genetic clusters (K=4). Founders and wild individuals cluster closely together and are separated from kulans and onagers along x-axis (DA1) and y-axis (DA2), respectively. The analysis was based on the first 15 principal component analysis (PCA) components which explain 52.40% of the total variance and 3 discriminant axes (DA) retaining all of this variance (DA1=62.45%, DA2=29.47%, DA3=8.08%). Top right inset displays the eigenvalues of principal components of the PCA with dark grey indicating the eigenvalues of the retained principal components. The top left inset displays the eigenvalues of the retained DAs of the DAPC, with the eigenvalues of the displayed DAs in dark grey.

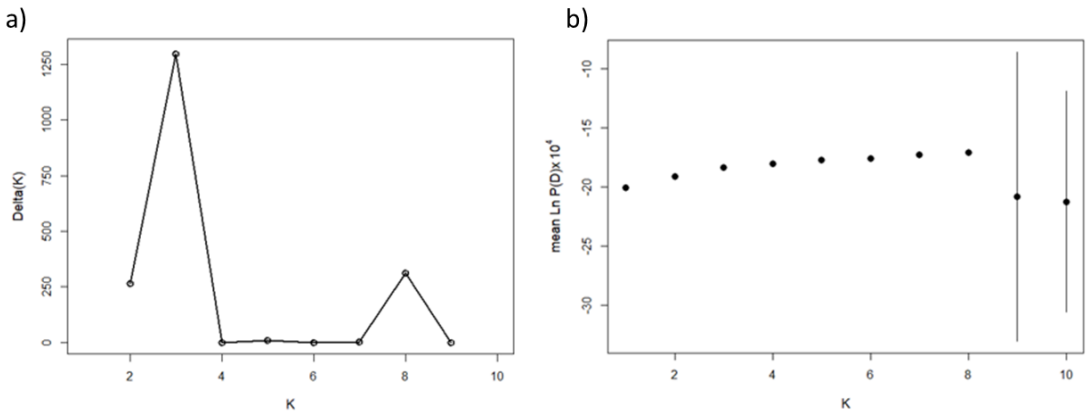


Fig. 3.11 Optimal number of genetic clusters (K) based on STRUCTURE. a) Delta(K) values peak at K=3 and K=8. b) The mean (± SD) log probability as function of K. The value increases slightly until K=3.

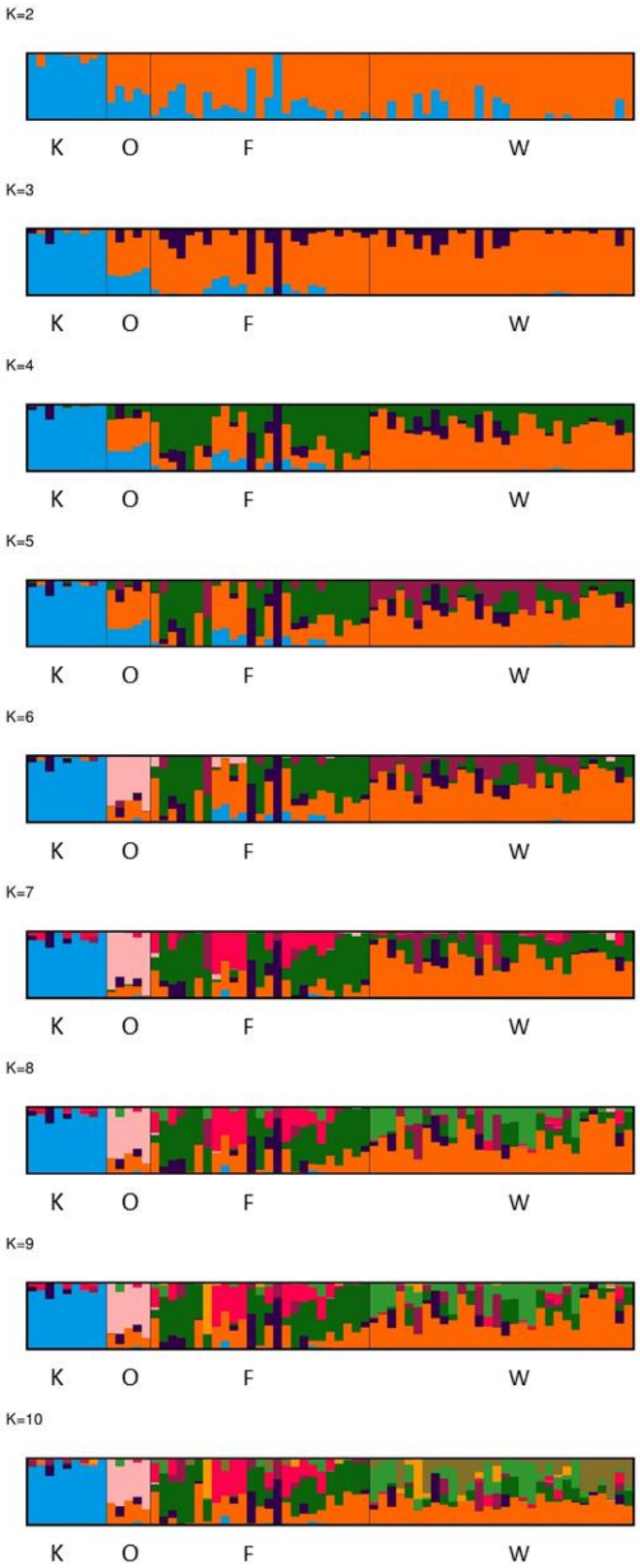


Fig. 3.12 STRUCTURE bar plot for increasing number of genetic clusters (K). Each bar represents and individual and sampled populations are separated by black lines (K=kulan, O=onager, F=founder, W=wild). The colours indicate the inferred clusters.

3.4.2 Heterozygosity and Inbreeding

Expected heterozygosity and observed heterozygosity were similar in all four populations, with *Ho* significantly higher than *He* in all but the wild population, which displayed a significant heterozygote deficit (Table 3.3, but see also Appendix A, Table A3). *Ho* was the highest in the founder population and significantly greater than in the zoo populations of onagers (independent samples Student's t-test: $t(4254.9)=-7.439$, $p<0.001$) and kulans ($t(4360.4)=-4.292$, $p<0.001$; Table 3.3). Furthermore, comparison of the Israeli founder and wild population showed a 10.6% loss of heterozygosity during the reintroduction. This difference was found to be significant ($t(4374.4)=4.569$, $p<0.001$). The founder population also displayed significantly greater variation in individual heterozygosity levels than the wild population (Bartlett's test: K-squared(1)=6.230, $p<0.05$, Fig. 3.13). The inbreeding coefficient was low in all populations, yet significantly smaller in the founder population than in the wild (independent samples Student's t-test: $t(3517.7)=-8.645$, $p<0.001$). Repeating this analysis after removing Hardy-Weinberg equilibrium (HWE) outlier loci, which potentially indicate sequencing error, had only minor effects on the results (see Appendix A, Table A3).

Table 3.3 Estimates of population genetic parameters for the four different populations. Expected heterozygosity (*He*), observed heterozygosity (*Ho*), mean individual inbreeding coefficient (*Fis*). Significant levels (* p<0.01, **p<0.001) are shown for the paired t-test comparison of *He* and *Ho* for all populations and for independent t-test comparison of *Fis* between founder and wild population.

	He	***Ho***	***Fis***
Onager	0.201	0.211**	- 0.057
Kulan	0.227	0.235*	- 0.037
Founder	0.252	0.263**	- 0.033
Wild	0.246	0.235**	0.043**

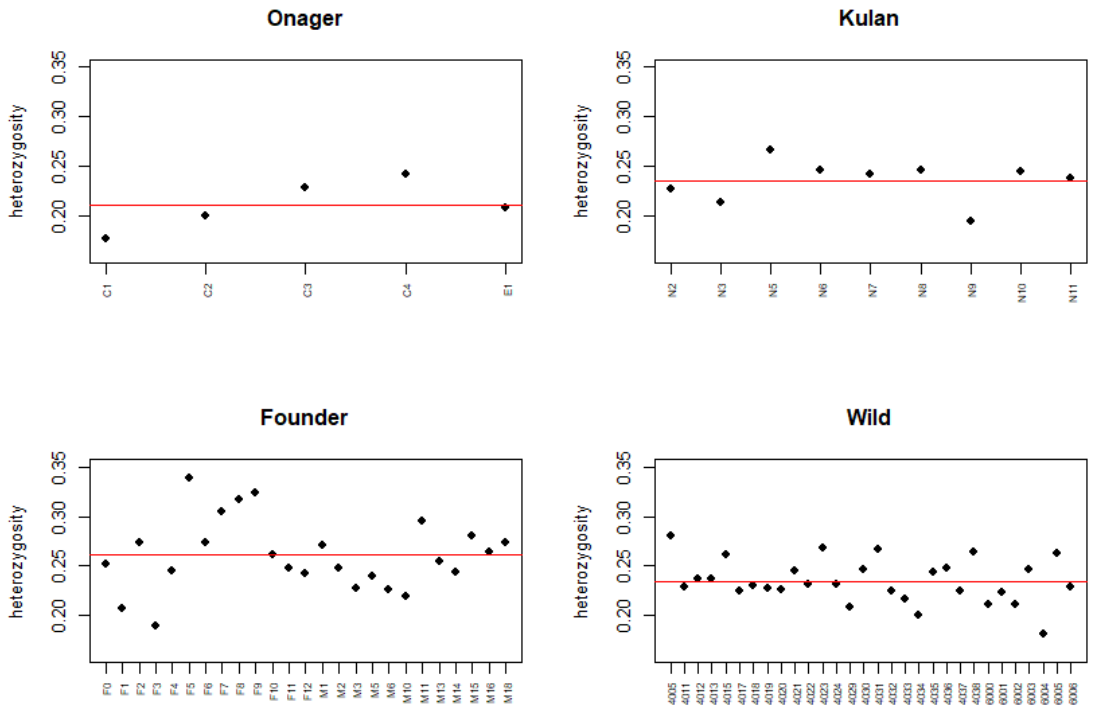


Fig. 3.13 Variation in individual heterozygosity by population. Individual IDs are on the x-axis, proportion of heterozygote markers on the y-axis. Red lines indicate population means.

3.4.3 Effective population sizes

Effective population size estimates for the wild population differed between the two methods used. The temporal method estimated a variance effective population size (95% CIs) of *Nev*=14 (12.9-15.1). The linkage disequilibrium method estimated an inbreeding effective size of *Nef*=26.7 (18.6-42.3) for the wild population. For the founder population the estimated inbreeding effective size was relatively small at *Nef*=7.5 (3.3-13.4).

3.4.4 Heterozygosity simulations under alternative reintroduction scenarios

Despite accurate records of the population's demographic history the simulated SFS did not match the observed SFS. Divergence between the maximum likelihood estimates obtained for the observed SFS and the simulated SFS was high across different demographic scenarios. Interestingly, closest resemblance was achieved by model 2 and model 3, which did not include an admixture event (Table 3.4). Visual assessment highlighted that none of the simulated SFS produced by the different models accurately captured the distribution of the observed SFS. Specifically, the

empirical data were strongly skewed towards intermediate allele frequencies, with a notable increase in alleles with minor count ranging between 13-21 (Fig. 3.14). However, this increase was not displayed by either of the simulated SFSs. In the absence of a model closely resembling the SFS obtained from the empirical data set, it was not possible to test alternative reintroduction scenarios. Therefore, no heterozygosity estimates were obtained, and the analysis was not continued.

Table 3.4 Maximum likelihood estimates for the estimated and the observed site frequency spectra (SFS) for 5 different demographic models. Low Delta value indicates good model fit.

Model	MaxEstLhood	MaxObsLhood	|Delta|
1	-654.717	-626.511	28.206
2	-640.963	-626.511	14.452
3	-641.001	-626.511	14.490
4	-641.662	558.176	1,199.840
5	-643.318	558.176	1,201.490

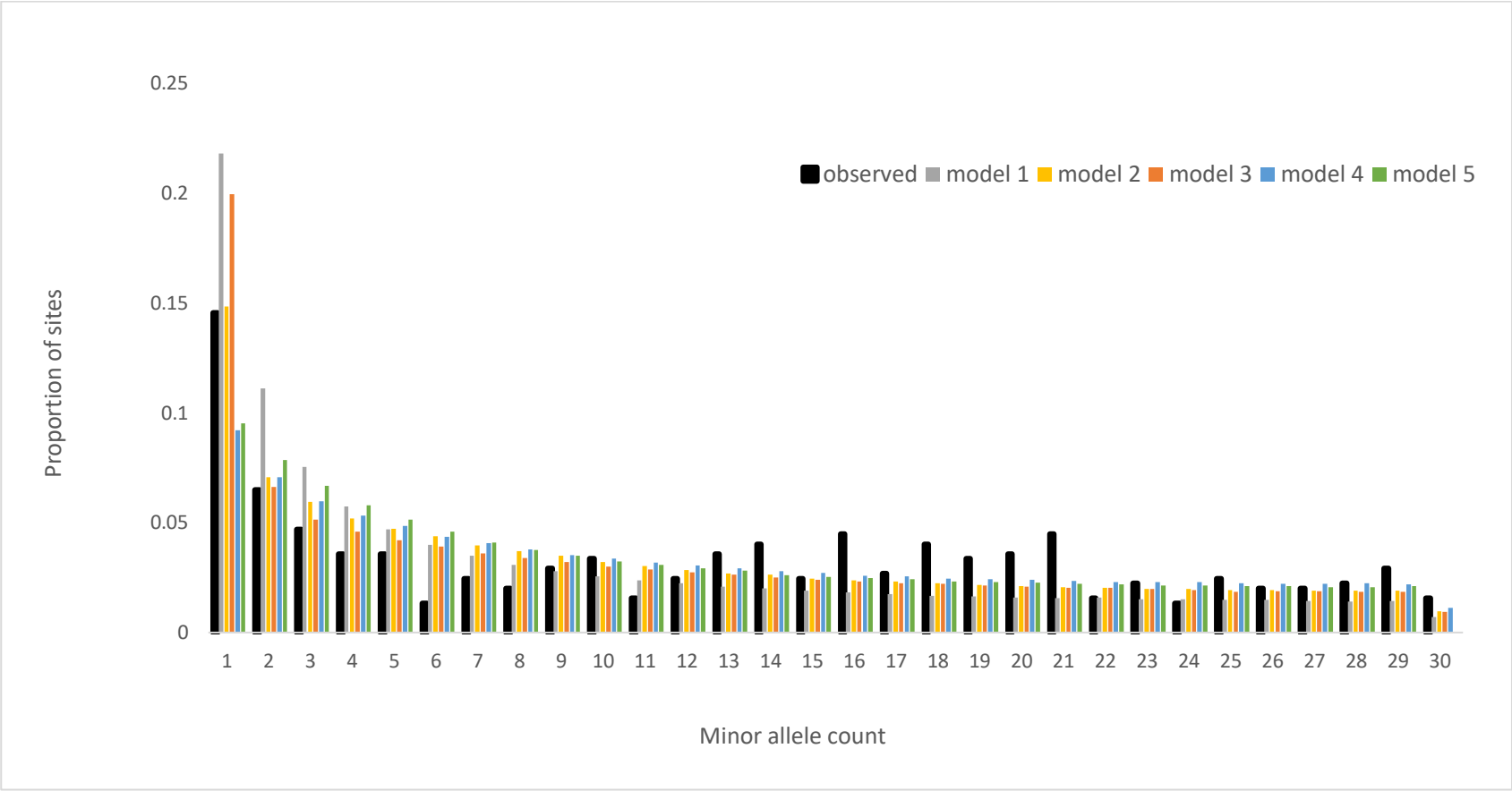


Fig. 3.14 Observed (black bold) and expected site frequency spectra, simulated using fastsimcoal2 for 5 different demographic models for the Israeli *E.h.ssp.* population. Models represent the following demographic histories: model 1=population of constant size, model 2=population after single bottleneck, model 3=population after two bottlenecks, model 4=hybrid population, model 5=hybrid population after two bottlenecks.

3.5 Discussion

3.5.1 Subspecies admixture

The different aspects of the admixture analysis all demonstrated that the two subspecies admixed successfully. Individuals displayed high levels of subspecies admixture in the breeding core, after three generations of unmanaged mating (Gueta et al., 2014). Further, high levels of subspecies admixture were maintained in the wild population. There was no spatial autocorrelation in the wild population based on hybrid indices, indicating no genetic or behavioural admixture barriers between the subspecies.

The genetic clustering analysis in Structure and the DAPC identified four founder individuals which overlapped with the wild population cluster. This could be an indication that these individuals contributed to the gene pool of the wild population (directly through mating or indirectly through mating of a close relative), therefore appearing more closely associated with it. In the PCA kulans were clearly separated from the other populations, while onagers consistently overlapped with the Israeli individuals. This could be an analytical artefact as uneven sample sizes may cause a distorted representation by the PCA (McVean, 2009). Alternatively, the strong differentiation of kulans may be caused by strong genetic drift experienced by the zoo populations causing shifts in allele frequencies which may lead to the observed pattern (Frankham, Briscoe, et al., 2002).

Consistent with the genetic structure analysis, the admixture analysis also indicated that the Israeli populations were genetically closer to the onagers, as the median ancestry proportion was slightly biased towards the onager outgroup. This observed bias could be due to a chance difference in the genetic founders of the population or due to genetic drift. While the captive breeding core was founded with almost equal numbers of onagers (3M,3F) and kulans (2M, 3F), it is likely that not all of these individuals reproduced. Considering that wild ass are strongly polygynous, it is possible that a dominant male onager seized most mating opportunities during the early stages of the breeding programme, causing the observed bias (Greenbaum et al., 2018). The

observed bias could also be the result of strong genetic drift experienced by the captive breeding core population.

The high admixture proportions and apparent lack of a genetic barrier are consistent with the low observed genetic differentiation between the zoo samples. Only 0.05% of the analysed SNPs were fixed for opposite alleles in the two subspecies. Although the samples sizes were small and genetic variation within the subspecies may have been not accurately represented, these results coincide with previous studies demonstrating only low genetic divergence between onagers and kulans (e.g. Bennett *et al.*, 2017, see chapter 1).

3.5.2 Heterozygosity and effective population size

The results of the population genomic analysis were consistent with the population's demographic history. Specifically, the founder, kulan and onager populations all displayed a significant heterozygote excess. There are multiple possible explanations for these results: In the zoo populations this could be due to managed breeding programmes. The captive onager and kulan populations are both part of the European Endangered Species Programmes of the European Association of Zoos and Aquaria, which aims to maximise retained genetic diversity in the captive populations (EAZA, 2019). Managed breeding between the most genetically divergent individuals could results an observed heterozygote excess. Future research could compare gene diversity measures for the zoo populations from this thesis with estimates based on studbook data of onagers and kulans. By analysing the number of founder genome equivalents from the studbooks and tracing the pedigree the likelihood of heterozygote excess due to managed breeding could be tested (Ito, Ogden, Langenhorst, & Inoue-Murayama, 2017).

In the case of the unmanaged Israeli founder population it is more likely that excess heterozygosity is caused by the previous admixture between the two subspecies. This is further supported by the significantly greater levels of observed heterozygosity in the founder population compared to the zoo populations, which suggest that subspecies admixture has resulted in increased genetic diversity. In the wild population observed heterozygosity was significantly lower than expected. This could be due to the mating system and non-random mating in the wild population potentially leading

to increased levels of inbreeding (Frankham, Briscoe, et al., 2002). However, considering the population's history it seems likely that the observed heterozygote deficit is a signature of the genetic bottleneck the population experienced following the release events. An alternative explanation is that heterozygote excess is caused by differing allele frequencies between males and females. In small populations, such as the zoo and Israeli founder populations, stochastic effects can lead to random changes in allele frequencies resulting in an observed heterozygote excess, even if random mating occurs (Templeton, 2018). Finally, genotyping errors introduced during bioinformatic processing could have produced false heterozygotes, resulting in an observed heterozygote excess. However, a recent study suggests that genotyping error, while common in RADseq experiments, is more likely to result in an underestimation of genetic diversity, making this explanation less likely (Bresadola, Link, Buerkle, Lexer, & Wegmann, 2020; see 5.1 for detailed discussion of potential gentopying error).

The 10.6% reduction in the observed heterozygosity from the founder to the wild population is probably due to the bottleneck and founder event during the release into the wild. Reductions in heterozygosity in reintroduced populations have been reported in multiple other species, and particularly in isolated populations (For example: moose *Alces alces,* Broders, Mahoney, Montevecchi, & Davidson, 1999; Merriam's turkey *Meleagris gallopavo merriami,* Mock et al., 2004; European brown bears *Ursus arctos*, De Barba et al., 2010*;* alpine ibex *Capra Ibex,* Grossen et al., 2018). The extent of the genetic diversity loss depends on multiple factors, including gene flow, but also the number of founders (Wright et al., 2014) and post release survival and reproductive rate (Biebach & Keller, 2012). Extended periods of low population growth after release can significantly reduce the genetic diversity maintained, as it has been reported for a reintroduced herd of Pennsylvania elk (*Cervus elaphus*). The population originated from a similar number of founders (N=34) as the Asiatic wild ass in Israel, and displayed a 61% reduction in heterozygosity seven generations after release (Coulson et al., 1998; Williams et al., 2002). The wild ass has been reintroduced in Israel four generations ago and heterozygosity loss, while it may be still ongoing, does not occur at the same rate as in the Pennsylvania elk. Unlike the elk, the wild ass population displayed fast post-release growth, with female net reproductive rate (R_0=1.87)

exceeding those observed in other reintroduced populations of Asiatic wild ass (Saltz & Rubenstein, 1995). This fast population growth would have reduced the duration of the bottleneck and minimised the amount heterozygosity lost due to strong genetic drift in the early stages of population establishment (Templeton, 2006).

Like heterozygosity levels, the effective population size estimates are also consistent with the demographic history of the population. The inbreeding effective size was markedly smaller in the founder (*Nef*=7.5) compared to the wild (*Nef*=26.7) population. This is consistent with the fact that in populations founded by a small number of individuals *Nef* is often closer to the number of original founders (N=11) than the current census size (N>30) (Templeton, 2018). However, this founding effect will disappear over time as the population continues to grow, hence the observed increase in *Nef* in the wild population four generations later.

In contrast, the variance effective populations size is also influenced by the progeny number and expanding populations usually display a larger *Nev* than *Nef* (Braude & Templeton, 2010). However, high variance in reproductive success, as observed in the highly polygynous Asiatic wild ass, can reverse this relationship (Greenbaum et al., 2018; Templeton, 2018). The estimated *Nev* (*Nev*=14) is further supported by a previous study on the wild population based on eight microsatellite markers, which estimated a variance effective size in a similar range of *Nev*=24.3 (Greenbaum et al., 2018).

Both estimates of effective population size for the reintroduced Asiatic wild ass population are very low, which is consistent with the past bottlenecks. This is not unusual, as other studies on reintroduced and isolated populations coming through a bottleneck have also reported extremely small inbreeding effective population sizes (e.g. *Nef*=3 in American marten *Martes americana*, Manlick, Romanski and Pauli, 2018; *Nef*=2-8 in red deer *Cervus elaphus*, Zachos *et al.*, 2016; *Nef*=3 in American black bear *Ursus americanus*, Murphy *et al.*, 2018). However, such small effective population sizes over extended periods can seriously threaten the populations long-term persistence. Franklin (1980) made recommendations for *in situ* management of populations based on the 50/500 rule. This rule of thumb recommends managing populations to achieve a minimum of *Nef*=50 and *Nev*=500 to avoid inbreeding depression and the long-term

loss of genetic diversity due to genetic drift, respectively (Braude & Templeton, 2010; Franklin, 1980; Jamieson & Allendorf, 2012).

While the inbreeding effective size has increased in the wild population it is still significantly below the recommended value. This is a major concern since inbreeding depression has been reported previously for equids. For example, Kaczensky et al. (2018) described a high mean individual inbreeding coefficient together with very low reproductive output in a captive populations of onagers. Furthermore, Sasidharan et al. (2011) compared population genetics between wild populations of mountain zebra (*Equus zebra zebra*), some of which suffering from a virus causing skin tumours. Affected populations displayed low levels of heterozygosity and high levels of internal relatedness compared to healthy populations, suggesting inbreeding depression as a potential cause for poor population health.

The variance effective size is expected to increase as the population continues to grow, however, this may be enhanced by active management aiming to increase the effective number of breeding individuals. Specifically, increasing the number of males contributing to the gene pool has been suggested as an effective measure considering the strong polygynous mating system (Greenbaum et al., 2018; Renan et al., 2018). A more powerful method may be the supplementation of the reintroduced population through additional release of individuals. Other studies have highlighted the need for continued management, including periodical release of individuals, in small and isolated reintroduced populations to avoid long-term genetic erosion (Jamieson, 2010; Saremi et al., 2019; vonHoldt et al., 2007). Individuals could be sourced from the persisting captive breeding core population at the HaiBar Yotvata reserve. Alternatively, the results of this thesis suggest no negative effects of admixture, hence inferring the possibility of introducing kulans or onagers from captive populations to increase genetic diversity in the Israeli wild population. However, a more cautionary approach would be to cross subspecies individuals in a captive setting to avoid any risk not identified by the analysis in this thesis.

The reintroduced population maintained relatively high levels of heterozygosity compared to the breeding core and to other reintroduced populations. However, while the growing census size of the population is crucial to withstand demographic stochasticity, future management must be aimed specifically at increasing effective

population sizes to improve chances of long-term persistence of the population. Furthermore, continued genetic monitoring of the population and its effective population sizes is necessary, as monitoring census size alone may fail to detect potential negative developments (Moraes et al., 2017).

3.5.3 Coalescent simulations

The population's complex demographic history was reflected in the observed SFS, which showed a distinct genomic signature. The SFS displayed an excess of intermediate frequencies, which is typical for a past admixture event caused by an exchange of variants which were previously fixed in the parental populations (Alcala, Jensen, Telenti, & Vuilleumier, 2016). The location and height of these peaks depends on the timing of the event, with peaks fading over time. In addition to this excess, there was a clear deficit in singleton and doubleton alleles, which is characteristic for a recent population bottleneck, during which disproportionally more rare variants are lost (Nei, Maruyama, & Chakraborty, 1975).

It is unclear why the coalescent simulations failed to reproduce the observed SFS. One possible cause could be the short time frame in combination with the high complexity of the simulated history. It is possible that due to potentially imprecise records (e.g. past population sizes) the input parameters did not accurately represent the population's history. However, the fact that that model fit did not improve with increasing model complexity suggests that the underlying problem is not a lack of detailed records. Coalescent simulations are usually applied to scenarios spanning over a much longer time frame (e.g. Sovic, Carstens and Gibbs, 2016; Thomé and Carstens, 2016; Stoffel *et al.*, 2018). It is possible that historic events in the demographic history of the Israeli population occurred over too few generations to leave a strong easily detectable signature in the SFS. Better results might be achieved with forward-time simulations, which tend to perform better for very complex scenarios (Carvajal-Rodríguez, 2008; Yuan, Miller, Zhang, Herrington, & Wang, 2012). Future research could look at a combined approach of coalescent simulations to generate SFSs of the parental subspecies, which are then applied to forward-time simulations based on detailed life history data. This would provide important insights regarding the impact of past subspecies admixture on the current genetic makeup of the reintroduced population. This information would be extremely valuable, specifically for sourcing

individuals for potential future supplementation of the Israeli population, but also for conservation reintroductions in general.

Unfortunately, without the results of the simulations, it is not possible to infer how different reintroduction scenarios would have impacted the genetic makeup of the reintroduced population. Specifically, whether a single source population would have resulted in lower heterozygosity levels than observed today. Nonetheless, at the present stage this reintroduction appears to have been successful with respect to demographic population growth and high levels of retained heterozygosity. This is further supported by comparing the genetic parameters of the Israeli population with those of another reintroduced population of Asiatic wild ass. In the Altyn Emel National Park in Kazakhstan a total of 32 kulans were reintroduced in the early 1980s. The population grew rapidly and is estimated to comprise >3,000 individuals today. However, despite the rapid growth and large census size the population displays a striking heterozygote deficit (estimates based on nine microsattelite markers: *Ho*=0.39, *He*=0.73; Kaczensky, Kovtun, Habibrakhmanov, Reza, *et al.*, 2018). While different genetic methods were applied in this study, comparison of the relative heterozygosity deficit indicate that the Israeli population has retained more genetic diversity following reintroduction than the population in Kazakhstan.

3.5.4 Potential impact of bioinformatic processing and SNP filtering

The bioinformatic processing and SNP filtering are crucial steps in next generation sequencing projects (see Chapter 2). While great care was given to processing of the data set used in this analysis, it cannot be guaranteed that no bias was introduced which may have impacted downstream analysis presented in this chapter. The robustness of the here presented results could be further assessed through additional filtering steps. Physical linkage between sequenced genome fragments or between multiple SNPs on the same fragment could be tested and linked loci removed. Additionally, the data set could be tested for loci under balancing selection using an *Fst* outlier approach (O'Leary et al., 2018). Finally, previous studies have demonstrated how bioinformatic processing and *de novo* assembly of sequencing data can impact downstream population genetics analysis (Rodríguez-Ezpeleta et al., 2016; Shafer et al., 2017). Further analysis could investigate whether a smaller set of SNPs, created by

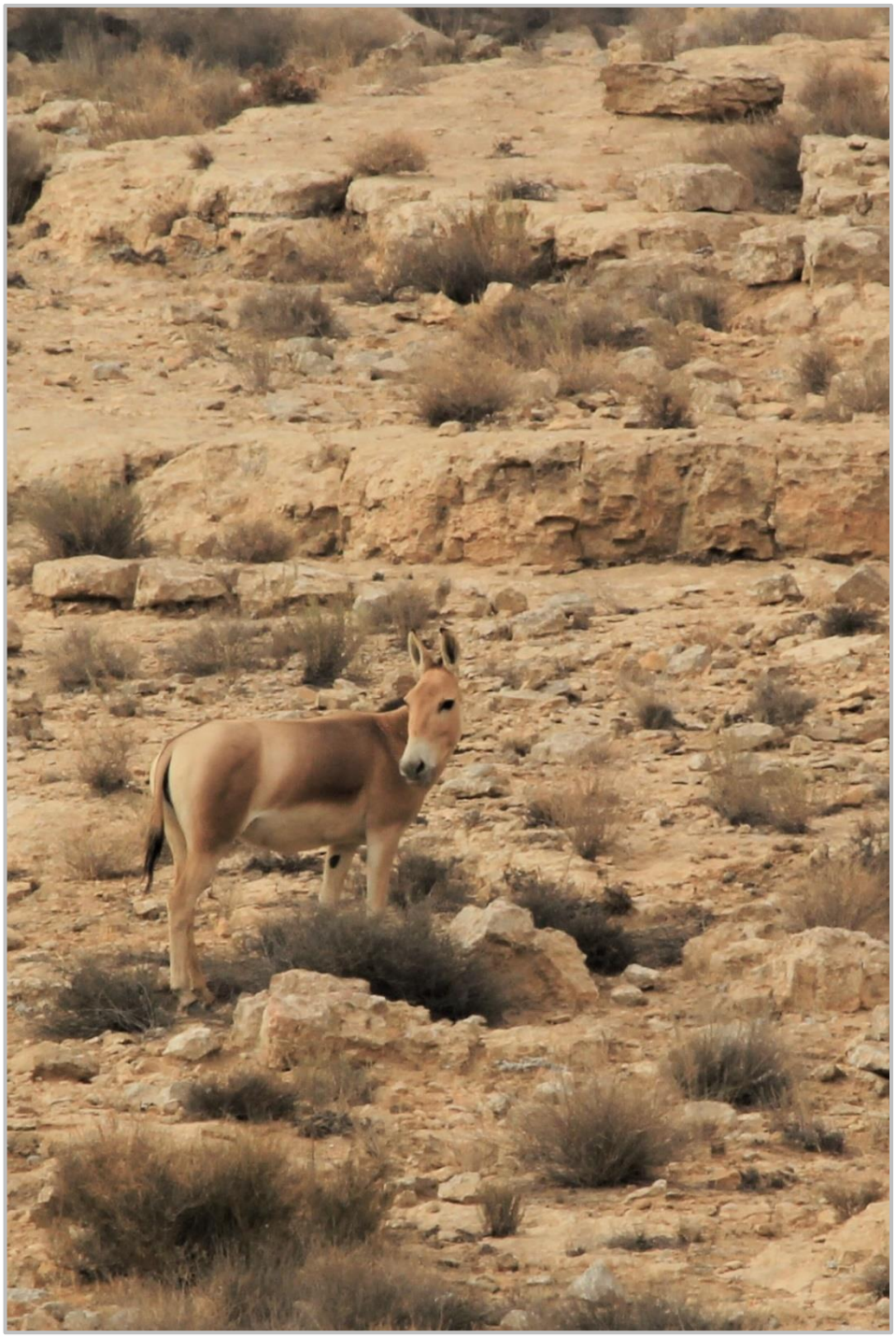
applying more stringent thresholds during the *de novo* assembly (see 2.3.4), has an impact on the results of the population genetic analysis.

3.5.5 Conclusion

There is an ongoing debate about optimal ways to select source populations for reintroductions (Edmands, 2007; Huff et al., 2011; Weeks et al., 2015). While several studies have reported higher levels of heterozygosity in populations of mixed source origin, others have highlighted the risk of mixing too divergent populations (Biebach & Keller, 2012; Huff et al., 2011; Olson, Whittaker, & Rhodes, 2013; White et al., 2018). Especially the admixture of different subspecies is highly controversial. However, it has been argued that the risk of outbreeding depression has been overstated, whereas inbreeding depression is a more imminent threat, especially in small and isolated populations (Frankham et al., 2011). This debate is largely based on a small number of case studies, and long-term genetic data on reintroduced populations are still rare (Seddon & Armstrong, 2016).

The present study provides valuable empirical evidence of a successful and complete admixture between individuals from two different subspecies. Further, the results suggest that the reintroduction has befitted from this admixture by increasing genetic diversity retained in the established population. These results provide important information to the ongoing debate about the use of different source populations for reintroductions specifically, and for the use of admixture for conservation management in general. Intentional admixture may be considered as a conservation tool for systems where genetic differentiation between source populations is low and hence outbreeding depression unlikely, as it is the case with onagers and kulans.

4. Landscape genetics of the Israeli wild ass population



4.1 Abstract

The long-term success of species reintroductions is strongly dependent on the availability of large areas of high-quality habitat. If available habitat is poorly connected this can hinder gene flow between areas and lead to the rise of genetic sub-structuring of the population, potentially increasing its extinction risk. I employed a conservation genomics approach in which I combined analyses of genetic structure with testing for potential landscape effects on habitat selection and gene flow in the reintroduced population of Asiatic wild ass *Equus hemionus ssp* in Israel. First, I investigated genetic structure and pairwise relatedness in the reintroduced population. Then, I tested landscape effects on individual habitat selection using records of global positioning system (GPS) collared individuals. Finally, I built habitat resistance surfaces and used electrical circuit theory to test for landscape effects on genetic relatedness. While genetic structuring (albeit quite weak) was detected, spatial coherence among individuals from the same genetic cluster was low. Landscape variables had a significant impact on individual habitat selection, with wild ass avoiding steep slopes and habitats of low suitability as predicted by a species distribution model. However, the landscape genetic analysis revealed no effect of habitat resistance on genetic relatedness. The results suggest that gene flow in the reintroduced population is not impacted by landscape resistance and I discuss other potential causes for the observed genetic structure. This study highlights the importance of understanding species habitat interactions for the long-term success of species reintroductions.

4.2 Introduction

Identification of suitable habitat is a crucial aspect in species reintroductions and requires careful and thorough consideration (Armstrong & Seddon, 2008; Moorhouse, Gelling, & Macdonald, 2009). However, for the long-term success it is important that considerations extend beyond the reintroduction site to the greater geographical region, termed *release area* by the IUCN (IUCN/SSC, 2013). Ecological considerations should include the extent, composition, and connectivity of the available habitat at a scale that enables population growth and long-term persistence. Especially for large-bodied, highly mobile species habitat connectivity is of great importance (Seddon & Armstrong, 2016). In contrast, habitats with low functional connectivity (whereby the

landscape impedes individual movement) can hinder dispersal and range expansion and prevent reintroduced populations from successfully colonising the available habitat (Ziółkowska, Perzanowski, Bleyhl, Ostapowicz, & Kuemmerle, 2016). Furthermore, low connectivity can also limit gene flow between occupied patches resulting in spatial sub-structuring of the population (Bergl & Vigilant, 2007; Manel et al., 2003). This may explain observations of within-population genetic structure in reintroduced populations with genetic clusters centring around release sites (Grauer et al., 2017; Howell et al., 2016; Moraes et al., 2017). Resulting genetic isolation of subpopulations can make these population fragments more vulnerable to extinction due to inbreeding and stochastic genetic processes (Saccheri et al., 1998). In order to avoid the problem of genetic isolation, individuals must be able to disperse between release sites into new suitable territory at a rate that facilitates sufficient gene flow (Mills & Allendorf, 1996).

Gene flow is limited by factors restricting individual dispersal movements between habitat patches (Benton & Bowler, 2012). In terrestrial mammals, dispersal ability is usually affected by landscape structure, climatic and anthropogenic factors, or specific combinations of these (Howell et al., 2016). Major landscape features (e.g. roads, mountain ridges) may act as physical barriers completely preventing movement across, but areas of less preferred habitat may also impact gene flow (Storfer et al., 2007). For example, in female-philopatric mountain goats gene flow and relatedness across the landscape is best predicted by male habitat selection (Shafer et al., 2012). However, for many reintroduced populations, information on habitat use and preference is limited, especially when the species has been absent from the area for a long time or when it is replaced by a closely-related group (e.g., a different subspecies) which makes prediction of resource use and dispersal more difficult (Seddon & Soorae, 1999). Therefore, directly assessing habitat connectivity and gene flow and the factors impacting it is an important measure to optimise population management and to enhance chances of the population's long-term persistence.

The Asiatic wild ass *Equus hemionus ssp.* reintroduced to Israel presents an ideal opportunity for furthering our understanding of environmental effects on the dispersal and genetic structure in reintroduced populations. After the establishment of a captive breeding core from individuals of two different subspecies (Iranian onager *E.h.onager*

and Turkmen kulan *E.h.kulan*) in 1968, 38 descendants of this breeding core were released into the Negev desert. The reintroduction took place between 1982-1993 and consisted of six release events at two different sites (Fig. 4.1; see Chapter 1) (Saltz & Rubenstein, 1995). The population has since expanded its spatial distribution across the highly heterogenous landscape and is currently estimated at 300 individuals (Renan et al., 2018).

Recent studies have investigated the movement and space-use patterns (Giotto et al., 2015; Ziv, 2016) in the reintroduced population and analysed spatial genetic structuring (Gueta et al., 2014; Renan, 2014). An analysis based on 8 nuclear microsatellite markers and a systematic sampling regime using non-invasively collected DNA from faecal mounds identified a weak spatial genetic structure. A weak yet significant genetic differentiation (pairwise Fst) was discovered between 4 *a priori* defined subpopulation. Wild ass activity centres surrounding permanent water sources, which are separated by areas of low habitat conductivity, were defined as subpopulations (Nezer, Bar-David, Gueta, & Carmel, 2017; Renan, 2014). Gueta et al. (2014) also identified a weak spatial genetic structure of the population based on the analysis of mitochondrial DNA haplotypes. The authors suggested a combined effect of range expansion and low habitat connectivity between colonised areas to be the underlying cause for the observed structure (Gueta, Templeton and Bar-David, 2014). This theory is consistent with previous studies which identified resource distribution and topography as the main predictors for habitat use by the reintroduced wild ass (Davidson, Carmel, & Bar-David, 2013). Nezer et al. (2017) developed a high resolution species distribution model (SDM) based on surveys of wild ass faecal mounds and a data set of nine habitat variables relevant to wild ass biology (Nezer et al., 2017). The model had high predictive power and identified woody vegetation cover and slope to be the strongest predictors of wild ass distribution.

In the Negev, patches of suitable habitat are separated by areas of low resource availability and challenging topography such as steep cliffs and canyons which could act as barriers to wild ass movement, hence limiting gene flow between patches (Gueta et al., 2014; Nezer et al., 2017). Since the recently established population in Israel is geographically isolated with no opportunity for external migrants from neighbouring countries, it is particularly vulnerable to the negative effects of genetic

drift (Frankham, Ballou, et al., 2002). Further spatial subdivision would be a severe threat to this recently established population and could jeopardise the long-term success of the reintroduction.

The aim of the presented analysis was to further our knowledge of wild ass habitat interactions in the reintroduced population. Specifically, I investigate previously suggested landscape effects on the population genetic structure by combining the existing SDM and global positioning system (GPS) movement data with new next generation sequencing genomic data. First, I assessed genetic clustering of the population using high resolution genomic data. Then, I analysed individual GPS collar data and investigated habitat selection with respect to slope and SDM-based habitat suitability. Finally, I created landscape resistance surfaces from habitat selection data and applied electrical circuit theory to test for an effect of habitat resistance on genetic relatedness. Based on wild ass ecology and previous studies of the population, I predict that (i) the population in Israel is genetically structured into spatially distinct clusters. Further, I expect (ii) that individuals avoid areas of low habitat suitability (based on the SDM) and steep slope as reported for wild ass in other populations (e.g. Sharma *et al.*, 2004). (iii) I expect that steep cliffs form a complete barrier to wild ass movement and hence predict a stronger effect of slope-based landscape resistance than suitability-based landscape resistance on genetic relatedness in the population.

4.3 Methods

4.3.1 Study site

The Negev is a hyper-arid desert that extends throughout Southern Israel. The landscape is defined by a steep gradient in elevation ranging from the Negev Highlands in the Northwest (>1,000m above sea level) decreasing towards the Arava valley and the Dead Sea in the East (<300m below sea level) (Stern, Gradus, Meir, Krakover, & Tsoar, 1986). This elevation gradient coincides with a gradient in mean maximum annual temperature and precipitation, ranging from 22.6°C and 150mm in the Negev Highlands to 31.1°C and 30mm in the hotter and more arid Arava (Israel Meteorological Service). This climatic gradient also causes differences in vegetation, with shrub-steppes in the Negev Highlands giving way to sand and desert savannoid vegetation types in the Arava (Danin, 1999). The topography of the Negev is complex

and characterised by steep cliffs and levelled floodplains. Vegetation is mostly limited to ephemeral streambeds and floodplains. Permanent water sources are scarce, however, flash floods occurring after heavy rainfall in the winter fill up rock pools which retain water for several months (Nezer et al., 2017). In addition, there are three artificial water sources which are maintained throughout the year by the INPA to provide wildlife with water, which have also become activity centres of the wild ass population (Gueta et al., 2014; Nezer et al., 2017).

4.3.2 Genetic structure analysis

Between 2011-2017 a total of 33 blood and tissue samples were collected from the reintroduced population. Samples were collected opportunistically from roadkill and individuals captured for fitting of GPS collars. DNA was extracted and sequenced using ddRADseq methods and sequences were processed and filtered using Bioinformatic tools. The final data set contained 30 individuals and 1,645 SNPs (see chapter 2 for details).

Initially, data were explored using Principal component analysis (PCA), which fits orthogonal principal components (PCs) that summarise overall variability between individuals. Subsequently, genetic population structure was investigated in more detail using two different approaches: a discriminant analysis of principal components (DAPC; Jombart, 2008) and a Bayesian clustering analysis implemented in the programme STRUCTURE (Pritchard et al., 2000).

DAPC is a multivariate approach which performs a PCA in a first step and then subjects the PC scores to a discriminant function analysis (DFA). Unlike the PCA, the DFA fits orthogonal discriminant functions that maximise between group relative to within-group variation. Therefore, it is suited to differentiating between genetic groups (Jombart et al., 2010). A K-means clustering approach was applied to assess the numbers and composition of genetic clusters (K) in the data. The best supported model is identified using the Bayesian information criterion (BIC), where the lowest BIC is preferred. The PCA and DAPC were performed in the *adegenet* package in R (Jombart, 2008). In both analyses the first 10PCs were retained, which explained 54.96% of the total variance.

The program STRUCTURE was run with the admixture model and correlated allele frequencies for K=1-10, with 10 repetitions for each K. The runs were performed with $1x10^6$ iterations of the Markov Chain Monte Carlo (MCMC) chain preceded by $1x10^5$ burn-in iterations. STRUCTURE outputs were assessed for the optimal value of K using the log likelihood (Pritchard et al., 2000) and the Evanno method (Evanno et al., 2005) in the web-based version of STRUCTURE HARVESTER (Earl & vonHoldt, 2012). Pritchard et al. (2000) suggest that a value of K which maximizes the model log likelihood Ln(PD) is optimal. However, Ln(PD) often plateaus or continues to increase after reaching the optimal K-value and Evanno et al. (2005) proposed an improved method to estimating optimal K based on the second order rate of change of the likelihood function.

4.3.3 Individual habitat selection

A pre-existing telemetry dataset was used for this analysis. Between 2012 and 2013 five individuals (4M, 1F) of the reintroduced population were fitted with GPS collars (African Wildlife Tracking company) (Giotto et al., 2015) and an additional 2 females were collared in 2015. All individuals were captured in the same area in the Negev highlands near a permanent water source (Fig 4.1). Collars were set to record location every hour and animals wore collars between 10-25 months resulting in a minimum of 7,786 records per individual (Appendix B, Table B1).

To investigate for a potential effect of landscape characteristics on gene flow, habitat resistance surfaces were created. This approach assigns resistance values to each cell in a habitat grid, reflecting the relative cost inflicted on an individual moving through it (Spear, Balkenhol, Fortin, McRae, & Scribner, 2010). To improve the cost assignment, I first tested whether certain landscape factors impact individual movement. I performed a compositional analysis of habitat selection using individual movement data from GPS collar records (Aebischer, Robertson, & Kenward, 1993). The analysis compares the relative abundance of a specific habitat type with its relative use by individuals. This way, habitat types which are avoided by individuals and potentially represent high resistance to movement can be identified.

Relative habitat use was investigated with respect to two variables relevant to wild ass distribution. First, habitat was classified based on a suitability index using the model output of the SDM previously produced by Nezer et al. (2017). The model, which covered most of the area of the present study (Fig.4.1), produced a probabilistic

distribution map which represents the probability of wild ass distribution with values ranging from 0 (low probability) to 99 (high probability). This map was used as an indicator of habitat suitability (habitat suitability index) with low probability values indicating low habitat suitability and high values indicating high habitat suitability.

SDMs are commonly used to inform landscape resistance to movement (e.g. O'Brien *et al.*, 2006; Berkman *et al.*, 2013; Yumnam *et al.*, 2014; Howard *et al.*, 2015). However, some authors have suggested that habitat suitability may have little impact on gene flow (Mateo-Sánchez et al., 2015; Peterman, Conette, Semlitsch, & Eggert, 2014; Wasserman, Cushman, Schwartz, & Wallin, 2010). This is because SDMs often describe individual habitat preference. However, less preferred habitat types may still function as dispersal corridors (Abrahms et al., 2017; Keeley, Beier, Keeley, & Fagan, 2017; Spear, Cushman, & McRae, 2016). While preference may impact individual movement through the landscape, a habitat measure directly linked to movement ability was also included. Previous studies have reported topography as one of the most important constraints to wild ass movement with steep slopes (>30°) being avoided entirely (Davidson et al., 2013; Henley, Ward, & Schmidt, 2007; Nezer et al., 2017; Sharma et al., 2004). Consequently, relative habitat use was analysed with respect to slope as well. The same slope layer from the SDM was used, which was generated from a contour dataset retrieved from the Survey of Israel (MAPI; for further details see Nezer et al. 2017). Slope and habitat suitability raster layers, which had a resolution of 100m, were processed in ArcGIS 10.0.

To categorise habitat types, the grid cell value for habitat suitability index and slope layers were extracted for each GPS collar record using the *extract values to points* function in ArcGIS. Subsequently, data extracted for each variable were divided into categories, to investigate proportional habitat use. For habitat suitability index the range from 0-99 was divided equally, rendering a low (0-33), intermediate (34-66) and high (67-99) suitability category. Since steep slopes (>30°) were rare in the study area, slope was divided into low (0°-15°) and intermediate slope (16°-30°) and a category of steep slope containing all records >30°.

I performed a compositional analysis of habitat selection on the defined habitat categories (Aebischer et al., 1993). The analysis was performed using the *compana* function in the *adehabitatHS* package in R (Calenge, 2006). The analysis first tests for

significance of habitat selection using Wilks lambda and subsequently produces a ranking matrix indicating whether a specific habitat type is used significantly more or less than another. P-values were estimated by randomisation tests (999 permutations of the data). Aebischer et al. (1993) recommend using a minimum of six individuals, therefore, males and females were pooled for the analysis. Habitat use was analysed relative to habitat availability within the entire habitat area. The analysis was deliberately not limited to habitat available within an individual's home range, since gene flow is mediated by long-distance dispersal movements extending beyond home range boundaries. Restricting the analysis to individuals home ranges would have likely resulted in the loss of long-distance movements. Finally, compositional analysis assumes that all individuals select habitat in the same way, and this assumption was tested by eigenanalysis of selection ratios with the *eisra* function.

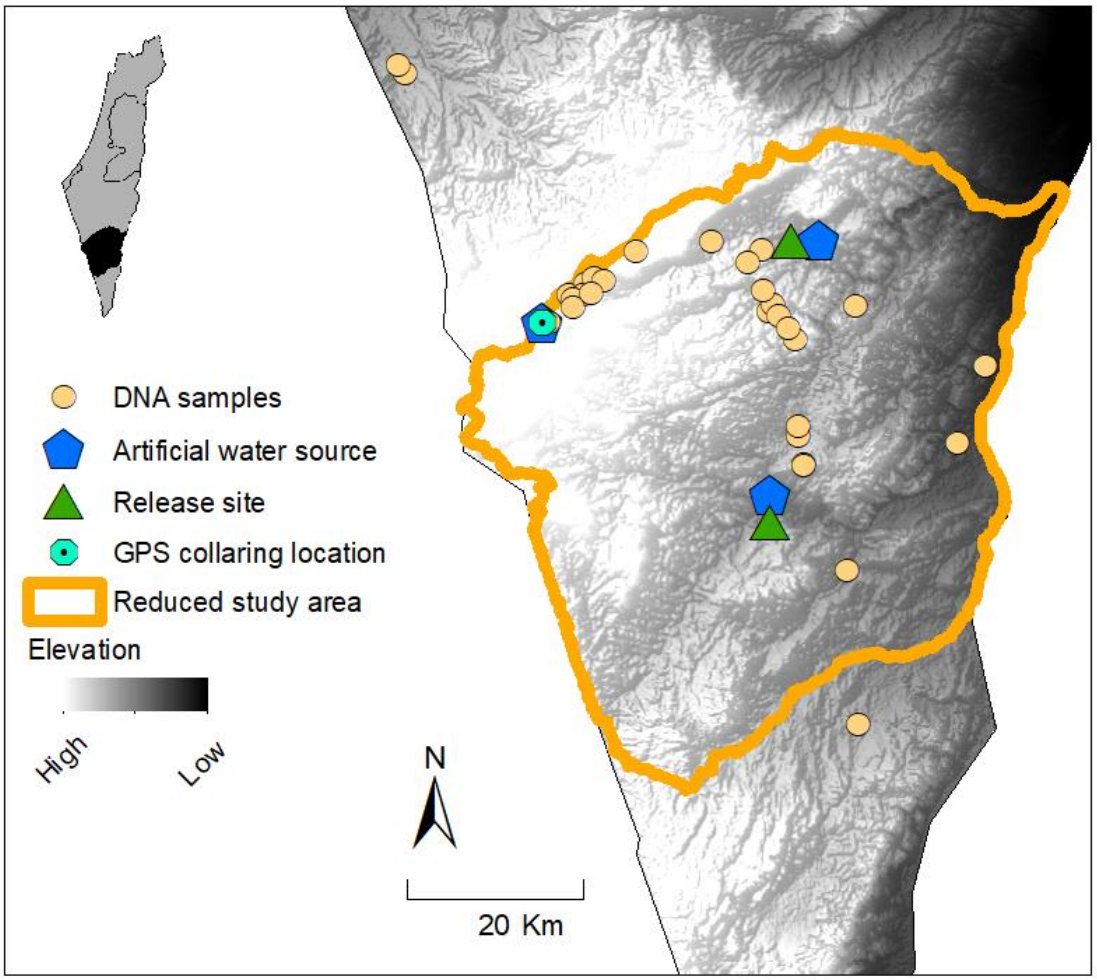


Fig. 4.1 Map of the study area. Depicted are locations of DNA sample collection, release sites of the reintroduction, location of three artificial water sources and capture location of GPS collared individuals. The orange outline indicates the area of the SDM created by Nezer et al. (2017) and the study area for the landscape genetic analysis.

4.3.4 Landscape genetic analysis

4.3.4.1 Resistance surfaces

Three habitat resistance surfaces were created based on habitat suitability index, slope and geographic distance. First, the habitat suitability index values were inverted using the *raster calculator* function in ArcGIS 10.0. Grid cells with a probability of 0 were assigned a marginally positive value of 0.01 to comply with input requirements of downstream analysis. The resulting habitat resistance map based on habitat suitability ranged from 0.01 (low resistance) to 99 (high resistance). The second resistance surface based on slope was parameterised so that grid cells with a slope of 1°-30° were assigned a resistance value of 1-30 respectively. A threshold was set by assigning grid cells with a slope >30° a resistance value of 99 (Table 4.1). Grid cells with a slope of 0° were assigned a resistance of 0.01. Additionally, a control resistance surface, based solely on geographic distance was created by assigning all grid cells a value of 1. All three resistance surfaces had a spatial resolution of 100m and were produced with ArcGIS 10.0 (Table 4.1, Fig. 4.2).

Table 4.1 Parameterisation of three habitat resistance surfaces.

Resistance layer based on	**Parameterisation**
Habitat suitability	Inverted SDM values Range 0.01-99
Slope	Slope =0° (resistance 0.01) Slope >0°≤30° (resistance same as slope) Slope >30° (resistance 99)
Geographic distance (control)	All cells resistance of 1

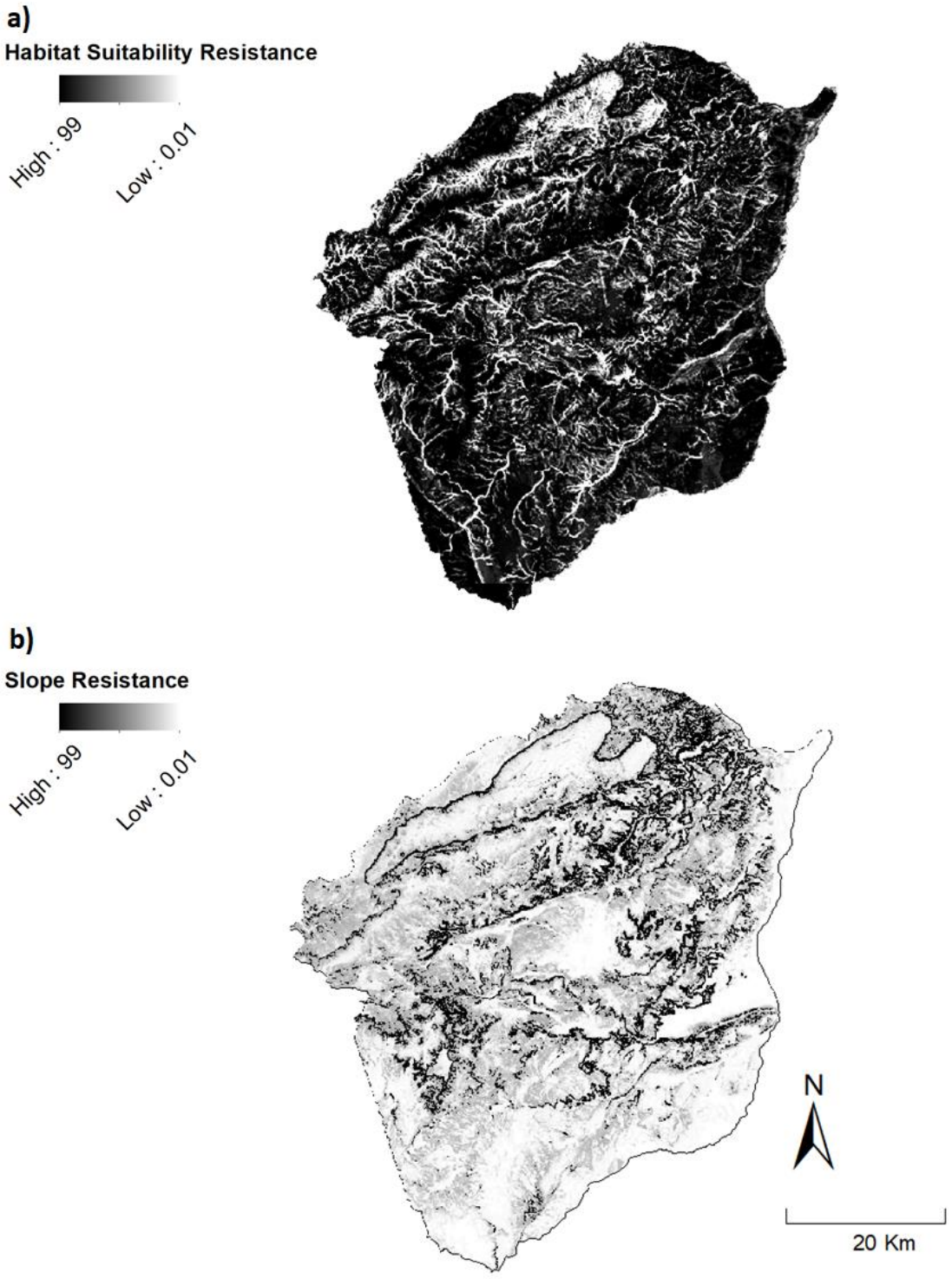


Fig. 4.2 Habitat resistance surfaces based on habitat suitability (a) and slope (b). Shading indicates resistance value.

4.3.4.2 Pairwise distances

Since the SDM from which resistance surfaces were derived did not cover the entire study area, the landscape genetic analysis was restricted to the part of the study area covered by the SDM (Fig. 4.1). As a result, three individuals which fell outside the SDM

area were excluded from the analysis (Fig. 4.1). The programme circuitscape Version 4.0 (McRae, Shah, & Mohapatra, 2013) was used to calculate pairwise resistance distances for the remaining 27 individuals for each of the three resistance surfaces. Circuitscape applies algorithms from electronic circuit theory to estimate resistances to current flow between nodes. The program was run in pairwise mode with individuals set as nodes, connected to all eight neighbouring cells surrounding a node. Pairwise genetic distance was expressed through a relatedness coefficient, which is effectively a measure of the genetic distance between two individuals. Pairwise relatedness coefficients were estimated using the corrected Wang (2002) estimator in the *related* R package (Pew, Muir, Wang, & Frasier, 2015).

4.3.4.3 Distance-based redundancy analysis

To test for a potential relationship between habitat resistance distance and genetic distance a distance-based redundancy analysis (dbRDA) was performed using the *capsscale* function in the *vegan* R package (Oksanen et al., 2010). dbRDA is an extension of multivariate regression which accepts a distance matrix as response variable. The response matrix is transformed into synthetic variables which are then regressed on multiple explanatory variables (Buttigieg & Ramette, 2014; Legendre, Andersson, & Anderson, 1999). First, the pairwise habitat resistance matrices were transformed to generate one-dimensional explanatory variables for the dbRDA. For this purpose, a principal coordinate analysis was performed using the *pcoa* function in the *ape* R package (Paradis & Schliep, 2018) with a Lingoes correction for negative eigenvalues to preserve all variation of the landscape resistance matrices. Subsequently, a broken stick model was used to estimate the number of significant principal coordinates (PCos) (MacArthur, 1957; Appendix B, Fig. B1). For all three resistance variables only the first or first and second PCos explained more variation than expected under the broken stick model. However, since this accounted for only ~35% of variance in each variable, the analysis was repeated with the first four PCos retained which accounted for >50% of variance (Table 4.2).

A total of seven models were tested, once with the first four PCos and once with only the first PCo retained. Three models tested for landscape resistance effects on gene flow, with the pairwise relatedness matrix set as the response variable and one of the three transformed habitat resistance matrices set as explanatory variables.

Additionally, four partial models were tested which controlled for an effect of geographic distance on habitat resistance and the reciprocal. Statistical significance was tested with permutation tests in the *anova.cca* function with 9,999 permutations.

Table 4.2 Percentage of variance explained by the retained principal coordinates (PCos) of different habitat resistance variables.

Variable	**Variance explained by retained principal coordinate**
only first PCo retained	
Habitat suitability resistance	34.92%
Slope resistance	37.07%
Geographic distance	35.18%
first four PCos retained	
Habitat suitability resistance	61.96%
Slope resistance	63.29%
Geographic distance	59.41%

4.4 Results

4.4.1 Genetic structure analysis

The variance explained by the first two principal components of the PCA was low (PC1 9.97%, PC2 7.54%) and no prominent clusters could be identified (Fig. 4.3a). Also, the BIC plot of the DAPC indicated K=1 as optimal (Fig. 4.3c). This suggested no meaningful genetic clustering in the population. In contrast, for the STRUCTURE analysis the Evanno method identified a clear peak in delta(K) for K=4 (delta(K)=57.07; Fig. 4.3d). However, the Evanno method cannot identify an optimum of K=1 and may indicate peaks at higher values of K even in the absence of any genetic structure (Evanno et al. 2005). The mean Ln P(D) across different values of K remains consistent with no distinct maximum value or plateau (Fig. 4.3d), suggesting that there may be only a very weak signal of genetic structure. The STRUCTURE ancestry plot highlights 4 clusters with high admixture levels in some individuals (Fig. 4.4b).

Since the two approaches gave slightly different results, their robustness was assessed by comparing the individual assignments to the four clusters between the multivariate and the Bayesian approach. Based on the results of the Evanno method, the DAPC was repeated with predefined K=4. Three out of the four described clusters were differentiated along the first PC while the fourth cluster was differentiated more strongly by the second PC (Fig. 4.5). Subsequently, individual assignments from the DAPC and STRUCTURE were compared. In the DAPC analysis all individuals had assignment probabilities of 1, whereas in STRUCTURE 12 individuals could not be assigned clearly to a single ancestral population (q-values <0.7) and these individuals were excluded from the comparison. Of the 18 remaining individuals, 16 clustered together in groups consistent between STRUCTURE and DAPC analyses (Figure 4b). However, spatial coherence of these clusters was low. Only one of the four genetic clusters also displayed spatial coherence, with six of the eight individuals assigned to this cluster originated from the same geographical location near an artificial water source. Individuals from the other genetic clusters were dispersed across the study area (Fig. 4.4a).

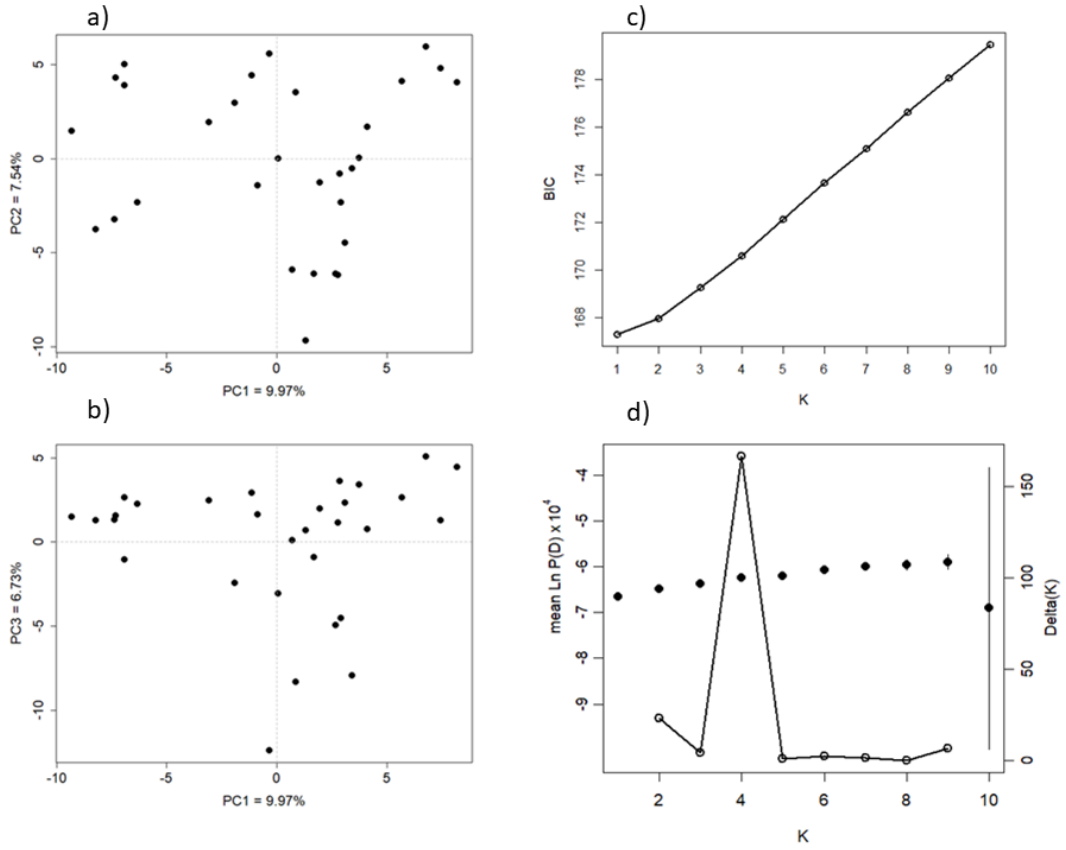


Fig. 4.3 Optimal number of genetic clusters (K) in the wild population. Initial exploration using principal component analysis indicates no distinct clustering along the first and second (a) and along the first and third (b) principal components. The Bayesian information criterion (BIC) is lowest for K=1 indicating no genetic clustering (c). Evanno methods indicates a clear peak for K=4, while the mean Ln P(D) identified no distinct maximum value (d).

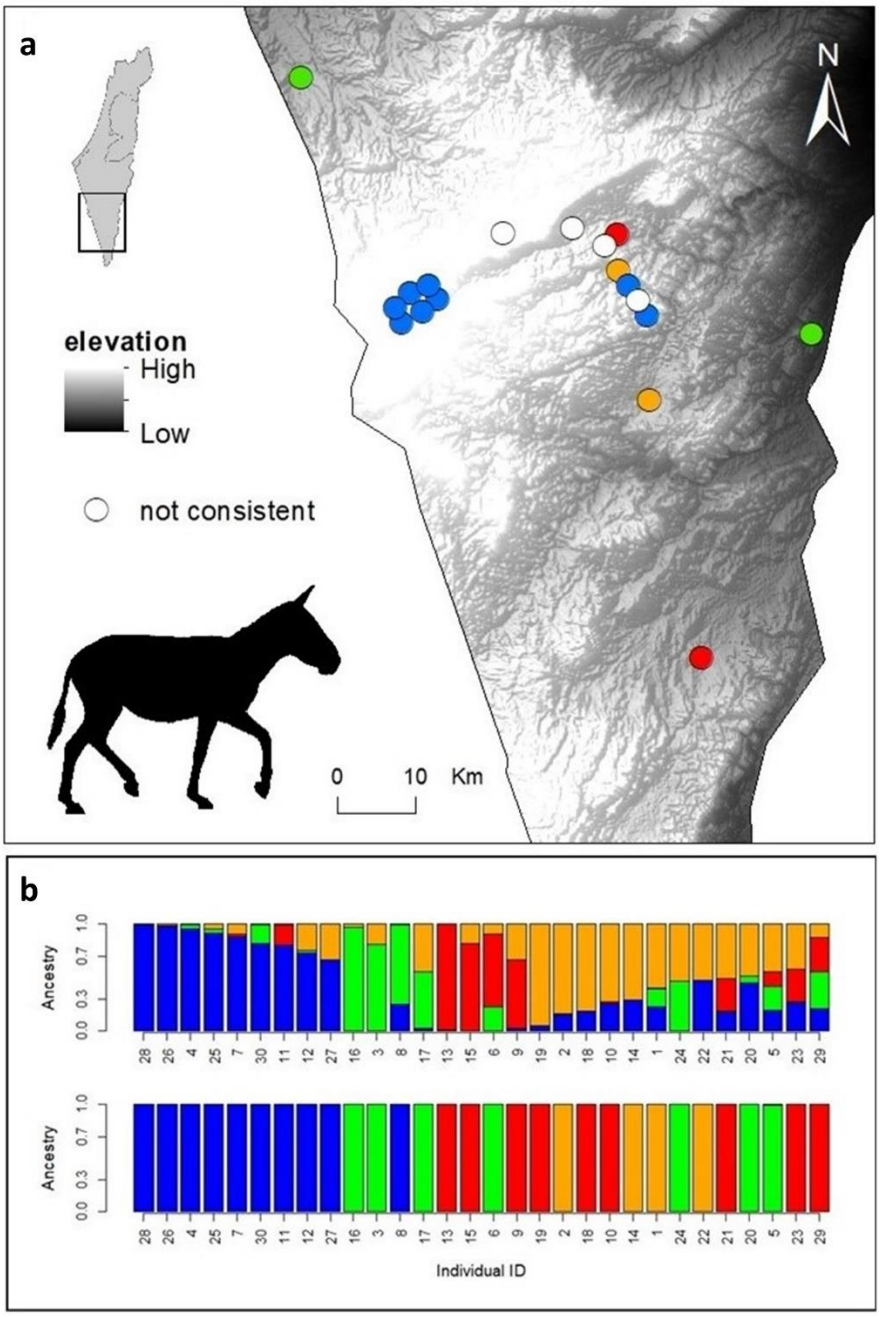


Fig. 4.4 Results of the genetic structure analysis. (a) Spatial distribution of sampling locations for individuals consistently assigned to the same cluster by both STRUCTURE and a discriminant analysis of principal components (DAPC). Only individuals with a high assignment probability (≥ 0.7) to a single genetic cluster are displayed. Colours indicate four genetic clusters (blue, green, red, orange), white points indicate individuals not assigned consistently by the two analyses. (b) Proportional ancestry of all 30 individuals for four genetic clusters (K=4) as estimated by STRUCTURE (top) and DAPC (bottom).

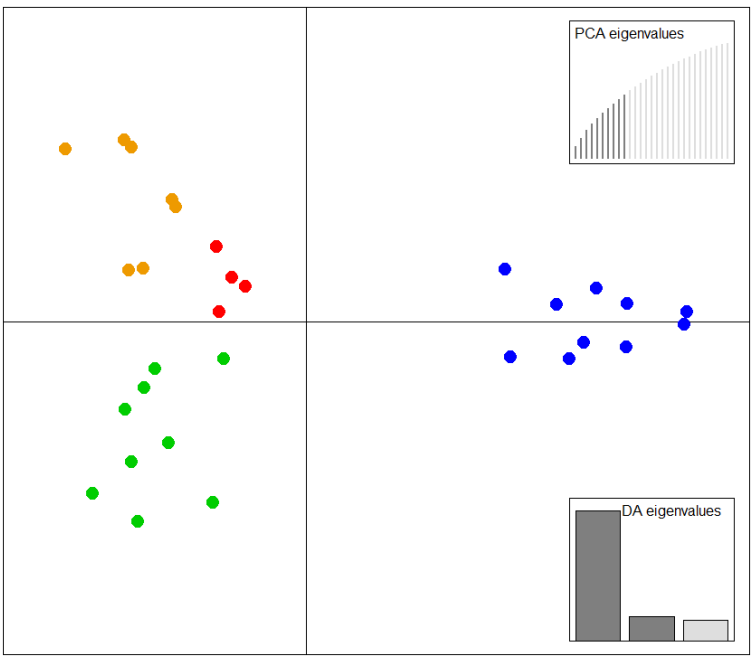


Fig. 4.5 Discriminant analysis of principal components (DAPC) plot of the current wild population for four genetic clusters (K=4). One clusters is clearly separated along the x-axis (DA1), while the other three clusters are differentiated along the y-axis (DA2).The analysis was based on the first 10 principal component analysis (PCA) components which explain 54.96% of the total variance and three discriminant analysis (DA) axes retaining all of this variance (DA1=71.45%, DA2=15.23%, DA3=13.32%). Top right inset displays the eigenvalues of principal components of the PCA with dark grey indicating the eigenvalues of the retained principal components. The bottom right inset displays the eigenvalues of the DAs of the DAPC with dark grey indicating the displayed DAs.

4.4.2 Individual habitat selection

Compositional analysis of habitat categories revealed that individual habitat selection differed significantly from random with respect to habitat suitability index (Lambda=0.013, p=0.012, by randomisation) and slope (Lambda=0.064, p=0.021, by randomisation). The ranking matrix highlighted a clear preference for low slope and high suitability habitats (Appendix B, Table B2). Wild ass used more low slope and more intermediate and high suitability habitat than proportionally available (Fig. 4.6). The analysis using GPS-collar data indicated that both habitat suitability index and slope are relevant variables affecting habitat selection in wild ass. Eigenanalysis of selection ratios indicated no difference in habitat selection between individuals (Appendix B, Fig. B2).

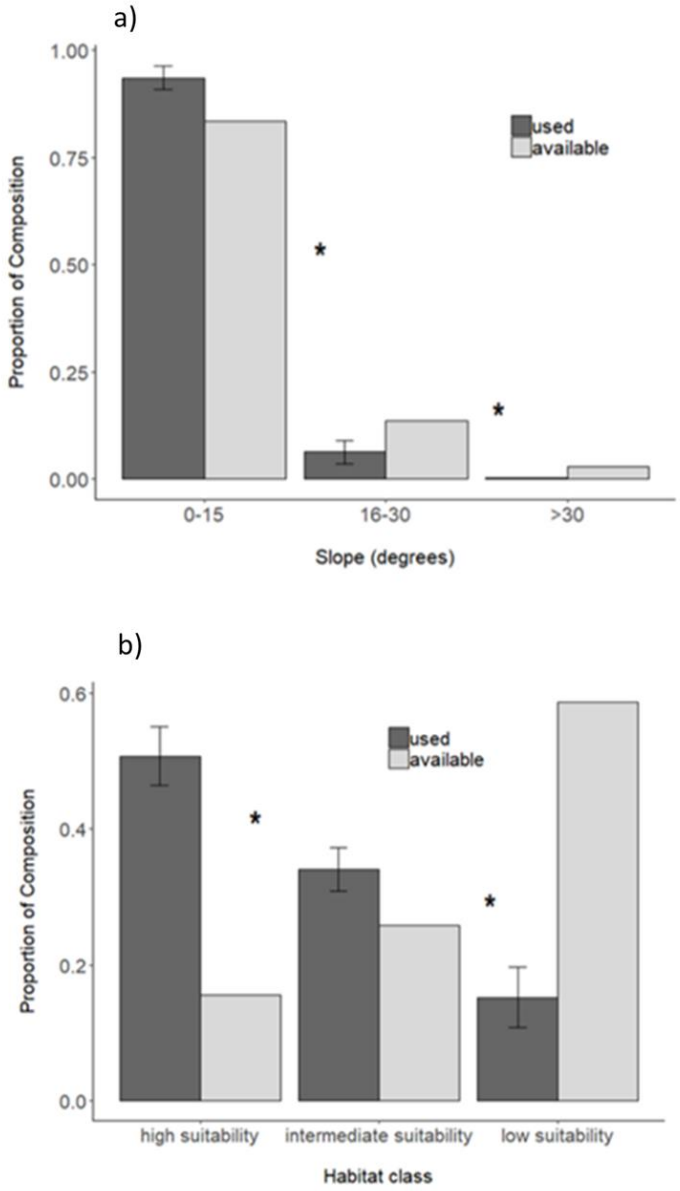


Fig. 4.6 Proportional habitat use by seven individuals in Israel between 2013-2017. Habitat is classified based on a) slope and b) suitability. Dark bars indicate mean (±SD) proportional usage by individuals and light bars indicate proportional availability in the study area of each habitat class. "*" indicates significance by permutation of differences in mean proportional habitat use between categories.

4.4.3 Landscape genetic analysis

None of the tested models returned significant results and the explained variance was close to zero for all predictor variables (Table 4.3). The analysis indicates that neither habitat suitability nor slope or geographic distance affected pairwise genetic distance in the population. This was also true for models controlling for geographic distance and resistance distances, respectively. Therefore, while individual movement is affected by

the habitat categories, it is not expressed in pairwise genetic distance. Habitat suitability and slope explained negligible variation in genetic relatedness between individuals of the population. These relationships were consistent across models that retained only the first PCo or the first four PCos of the explanatory variables, hence, the models appear to be insensitive to these minor variations, indicating robustness of the results.

Table 4.3 Results from the distance-based redundancy analysis. Displayed are the tested models, their total variance (Inertia), the % variance explained (R^2) and adjusted % variance explained (adjusted R^2), the degrees of freedom (df), F-statistic (F) and p-value (p) of the permutation tests (9,999 permutations). Partial models controlling for a third variable are indicated with |.

Variable	Inertia	R^2	Adjusted R^2	df	F	p
only first PCo retained						
SDM resistance	4.53	3.71%	<1%	1	0.964	0.635
Slope resistance	4.53	3.96%	<1%	1	1.031	0.407
Geographic distance	4.53	3.61%	<1%	1	0.936	0.726
SDM resistance | geographic distance	4.53	3.92%	<1%	1	1.018	0.441
Slope resistance | geographic distance	4.53	3.36%	<1%	1	0.867	0.865
Geographic distance | SDM resistance	4.53	3.82%	<1%	1	0.991	0.533
Geographic distance | slope resistance	4.53	3.01%	<1%	1	0.777	0.963
first 4 PCos retained						
SDM resistance	4.53	15.04%	<1%	4	0.974	0.694
Slope resistance	4.53	15.75%	<1%	4	1.028	0.317
Geographic distance	4.53	14.49%	<1%	4	0.932	0.873
SDM resistance | geographic distance	4.53	15.91%	<1%	4	1.028	0.396
Slope resistance | geographic distance	4.53	14.49%	<1%	4	0.918	0.819
Geographic distance | SDM resistance	4.53	15.36%	<1%	4	0.993	0.551
Geographic distance | slope resistance	4.53	13.23%	<1%	4	0.838	0.961

4.5 Discussion

The analysis revealed some genetic structuring in the reintroduced population of wild ass in Israel. However, inconsistencies in the optimal number of clusters and individual assignment between the different methods highlight that the genetic differentiation is weak. These results are consistent with a previous study on the same population using lower resolution genetic data. Renan et al (2014) also showed weak yet significant genetic differentiation, between four *a priori* defined subpopulations using eight microsatellite markers. Hence, there is now strong support for a (weak) genetic structure within the reintroduced population. Nevertheless, my new analysis of landscape resistance to individual movement does not support this as being a cause. The analysis of GPS data showed that landscape resistance affected wild ass habitat selection, with individuals clearly avoiding low suitability habitats and steep slopes. However, the landscape genetic analysis gave no support for an effect of landscape resistance on genetic relatedness.

4.5.1 Landscape effects on gene flow

While wild asses clearly avoid certain habitats, functional connectivity across the study area appears to be maintained. Although large proportions of the habitat have low suitability, these are interwoven by a network of low resistance paths, which likely facilitate individual movement across the landscape (Fig. 4.2). Furthermore, dispersing individuals may be willing to cross even low suitability habitats which are otherwise avoided during movements within the home range (Keeley et al., 2017). However, in contrast to my expectations, also habitat resistance based on slope was found to have no impact on genetic relatedness. Slopes above 30° account for only a very small proportion of the habitat in the Negev, nonetheless, they occur in the form of steep cliffs extending over large geographical areas and are expected to form true barriers to wild ass movement. However, wild ass are large-bodied, highly mobile mammals which have been reported to range long distances, and it is likely that even if individuals are unable to climb these cliffs, they can circumvent them (Nandintsetseg, Kaczensky, Ganbaatar, Leimgruber, & Mueller, 2016). In fact, the GPS data showed some long-distance movements by two females, which support the findings that even high resistance habitat does not prevent movement across the landscape in the Negev

population (Appendix B, Fig. B3). Therefore, despite being important for individual habitat selection, slope and habitat suitability appear to have no effect on gene flow.

I presented a vigorous landscape genetic analysis, using individual GPS collar data to verify the predictive power of a pre-existing SDM prior to parameterisation of resistance surfaces. The SDM was based on systematic surveys of indirect observations (dung piles) and habitat selection analysis demonstrated high predictive power of the model. This suggests that indirect observations are a suitable proxy for direct observations of individuals. Nevertheless, there are some limitations of the analysis, mostly resulting from the sampling regime. For example, all GPS collared individuals were captured in the same geographic area (Fig. 4.1) and consequently individual GPS records were only available for a subset of the study area. Should there be a habitat effect on the predictive power of the SDM then this would, then this may have introduced bias in the parameterisation of the resistance surfaces with potential downstream effects on the analysis.

Another possible limitation is the timing of data collection. GPS collar data, data used for the SDM and genetic samples were collected at different times which could potentially impact the results of the analysis. However, in the present study it is unlikely to have introduced bias in the analysis: The GPS collar data were collected continuously across seasons and for different years thereby accounting for potential seasonal effects of habitat selection. Further, the GPS data demonstrated high predictive power of the SDM, despite the fact that it was based on data collected several years before (aerial photographs taken in 2008; Nezer et al., 2017). This is likely due to the fact, that the main predictor variables identified by the model are permanent landscape features, which display little or no annual variation: woody vegetation cover (shrubs and trees), slope and altitude (Nezer et al., 2017). The genetic data were collected over a period of seven years and hence within the generation time of the species (Ransom et al., 2016).

Finally, in my analysis I used GPS data with hourly location records, which are likely autocorrelated and could have biased the analysis. However, the removal of data points to control for autocorrelation is debated, since it may impact the ability to identify fine-scale individual movements and habitat selection patterns (Cushman, 2009). Rooney et al. (1998) recommended using the shortest possible sampling

intervals and longest possible sampling duration to obtain best estimates of spatial habitat use. Given the relatively long sampling duration of the telemetry data used in this study (10-25 months), the impact of autocorrelation should be minimised. Nevertheless, further analysis could test for a potential effect of autocorrelation of the telemetry data.

The here presented landscape genetic analysis identified no effect of the measured habitat characteristics on gene flow. However, future studies could improve the analysis by increasing the sampling size and applying a systematic sampling design for the telemetry and genetic data.

4.5.2 Genetic structure

The weak genetic structure observed in the wild ass population is likely caused by factors other than landscape resistance. Three potential causes for genetic structuring are related to the population's demographic history. (i) At the onset of the reintroduction, a captive breeding core was created from individuals of two different subspecies (Gueta et al., 2014). Differences in the effective niche of these two subspecies may result in divergent habitat preferences and lead to spatial separation and limited interbreeding, ultimately promoting the rise of genetic substructure (McDonald, Johnson, Henry, & Cunneyworth, 2019). However, an analysis investigating spatial autocorrelation based on individual hybrid indices found no support for spatial segregation based on subspecies ancestry (see chapter 3). Alternatively (ii), the genetic structure could be the signature of the multiple release events during establishment of the wild population. Individuals were released at two reintroduction sites, from which they dispersed across the habitat (Fig. 4.1). Founder effects and genetic drift experienced by the population during early stages of population establishment could have caused the weak genetic differentiation. Other studies have described a genetic signature of release events in translocated populations (Biebach & Keller, 2009; Moraes et al., 2017; Puckett et al., 2014; Williams, Rhodes, & Serfass, 2000). For example, Grauer et al. (2017) reported unique patterns of genetic structure in a reintroduced population of American marten (*Martes americana*), caused by serial release events of individuals from different sources. Finally (iii), a behavioural effect related to the resource-defence-polygyny of the Asiatic wild ass could be the cause for the observed genetic clustering (Renan, 2014). Wild ass males defend territories

around permanent water sources. Increased resource requirements restrict females to the vicinity of these permanent water sources during the foaling and breeding season in the summer (Boyd et al., 2016; Saltz, Rowen, & Rubenstein, 2000; Wallach, Inbar, Scantlebury, Speakman, & Shanas, 2007). The GPS records of collared individuals reflected these behavioural patterns: Males remained close to a water source all year round, while females extended their movement range in the winter, yet returned to the same area of the permanent water source in the summer (Appendix B, Fig. B3). This seasonal range contraction and the resulting highly localised breeding activity could result in a genetic differentiation between individuals from different activity centres (Giotto et al., 2015; Renan, 2014). This could explain the presence of a fine-scale weak genetic structure despite high mobility of the species. A similar effect has been observed in feral horses (*Equus caballus*) in Nevada: during the hot summer, when most of the mating occurred, herds were unable to disperse from the limited water sources, which resulted in a weak genetic differentiation between populations despite overlapping winter ranges (Ashley, 2004).

While the current analysis failed to identify an effect of habitat on gene flow, it is important to consider the short lag time since the initial release of individuals, less than five generations ago (generation time 7.5 years, Ransom *et al.*, 2016). Landscape resistance may have an impact on gene flow, however, not enough time has passed for the signal to become established (Landguth et al., 2010). At this point it is not possible to determine with certainty what is causing the observed weak genetic differentiation. If it is due to range expansion and genetic drift, it is expected to diminish over time due to continued gene flow (Short & Petren, 2011). However, if it is caused by behavioural or a (not yet detectable) landscape effect, then it is likely to persist or even intensify.

4.5.3 Conservation implications

While some restriction to gene flow can increase the potential for retaining genetic diversity and is therefore beneficial (Chesser, 1991; Chesser, Rhodes, Sugg, & Schnabel, 1993), intensification of the genetic structure may lead to population fragmentation and genetic isolation of subpopulations, which could increase the population's extinction risk (Wang et al., 2017; With & King, 1999). In an isolated population of woodland caribou (*Rangifer tarandus caribou)* reduced gene flow has caused the rise of genetic substructure over a short time period (Pelletier et al., 2019). The authors

believe that this rising structure is severely threatening the population's long-term persistence, as a 53% reduction in the population's inbreeding effective size has been recorded over a timespan of only two generations. To avoid the risk of genetic isolation, management of the Asiatic wild ass population should aim to prevent any further reinforcement of the observed structure. Specifically, creating additional permanent water sources is expected to increase the number of activity centres, minimise distances between these, and potentially encourage more dispersal movements. Furthermore, additional permanent water sources provide more high-quality territories for Asiatic wild ass, thereby enabling a greater number of males to contribute to the gene pool (Greenbaum et al., 2018).

Other studies have highlighted the positive impact of artificial water sources on habitat use and dispersal in multiple species in arid environments (Krausman, Rosenstock, & Cain, 2006). A recent study in the Mojave Desert in the United States investigated the effect of artificial water catchments on single-species occupancy models for local wildlife (Rich, Beissinger, Brashares, & Furnas, 2019). The authors identified 18 species of terrestrial mammals, birds and bats whose occupancy was strongly and positively associated with the presence of these water sources. Catchments were particularly important for a large ungulate, the desert bighorn sheep (*Ovis canadensis*), which was seven times more likely to occupy habitats near these artificial water sources.

4.5.4 Conclusion

The presented findings have general implications for species conservation and management in arid landscapes. With globally increasing levels of desertification and simultaneous land conversion for agriculture and development, many species will increasingly struggle for access to water (Sherwood & Fu, 2014). Maintaining artificial water sources may be an important tool to counteract population fragmentation and genetic isolation in arid areas (Drake, Griffis-Kyle, & McIntyre, 2017; Mcintyre, Drake, & Griffis-Kyle, 2016). This highlights the potential impacts of resource availability and distribution on gene flow within populations and hence long-term reintroduction success. While it is standard procedure to assess resource accessibility prior to release of individuals to maximise post-release survival, it must also be continued even after population establishment to ensure that resources are available, and their spatial configuration facilitates dispersal and gene flow between different areas. Finally, the

results highlight the importance for long-term genetic monitoring of reintroduced populations. Genetic structure may develop even after successful establishment of a growing population (Neuwald & Templeton, 2013), and in the absence of obvious landscape barriers. While this may be simply a transient phenomenon caused by a founder effect, it may have other underlying causes. If genetic differentiation persists and intensifies, it can reduce reintroduction success even long after initial release of individuals and thus should be considered in conservation management protocols (Kramer-Schadt, Revilla, Wiegand, & Breitenmoser, 2004).

5. Final discussion and concluding remarks

Within each chapter I have discussed the presented results individually. Here I offer a wider discussion of my results within the broader context of the overall topic. I suggest general and specific directions for future research that build on the findings presented in this thesis.

5.1 Restriction site associated DNA sequencing in conservation genetics

In this thesis, I used double-digest RADseq methods and a *de novo* sequence assembly approach for a non-model species. I successfully recovered a set of >5,000 SNPs, which facilitated the population genomic analysis. Applying a dataset of thousands of genome-wide markers increased the genomic resolution and statistical power of my analysis compared to previous studies on wild ass populations (e.g. Gueta, Templeton and Bar-David, 2014; Renan *et al.*, 2015; Greenbaum *et al.*, 2018; Kaczensky *et al.*, 2018). Furthermore, for accurate hybridisation analysis, the selection of suitable markers is crucial. While neutral SNPs are less informative than microsatellite markers due to their lower variability and diallelic nature (Haasl & Payseur, 2011), they can be advantageous in population specific hybridisation analyses. Previous research has demonstrated that small sets of selected SNPs with high discriminative power outperform sets of microsatellites markers of the same size in hybrid assignment analysis (Väli et al., 2010). Therefore, the use of genomic data enabled accurate estimation of hybrid indices, which is unlikely to have been feasible using the small panel of microsatellites that was previously established for the species (Allendorf, Hohenlohe, & Luikart, 2010).

RADseq is part of a large suite of recent advances in genomics, which have dramatically increased the possibilities for conservation genetic analyses (Narum, Buerkle, Davey, Miller, & Hohenlohe, 2013). Through increased genomic resolution, new methods provide significantly improved statistical and analytical power for traditional population genetics, such as genetic diversity estimates, genetic clustering and gene flow analyses (Allendorf et al., 2010). Whole genome sequencing methods open up even more possibilities to improve conservation genomic studies, including highly accurate inbreeding estimates (Kardos, Taylor, Ellegren, Luikart, & Allendorf, 2016), genomic ancestry and hybridisation analysis (e.g. Mattucci *et al.*, 2019). For example, vonHoldt et al. (2016) used whole genome resequencing to investigate the ancestry of

endemic North American wolf species. The authors analysed 28 wolf genomes and concluded that the red and eastern wolf species were in fact hybrids between grey wolves (*Canis lupus*) and coyotes (*Canis latrans*). These results have important implications for the conservation status and management of grey wolves in North America.

While these advances in the field are exciting, the potential of genomic techniques for conservation has not been fully realised yet. Due to financial and computational limitations, applications of many novel genomic methods are restricted to relatively few case studies. Most of these studies are based on a small number of species, often those of commercial value (e.g., salmonid fish species: Garner *et al.*, 2016; Shafer *et al.*, 2016) or closely related species for which the same genomic resources can be applied, for example bovine SNP chip for bison (*Bison bonasus*, *Bison bison*) (Pertoldi et al., 2010) or domestic dog linkage map for European wolves (*Canis lupus*) (Kardos et al., 2018). Additionally, these new genomic methods are computationally and analytically demanding, making them less suitable for conservation practitioners (Combe et al., 2018). Consequently, there is a mismatch between advanced sequencing methods often applied in scientific studies and more traditional approaches, commonly used in the field of conservation (the "conservation genomics gap"; Shafer et al., 2015).

Compared to whole-genome resequencing methods, reduced representation sequencing (RRS) techniques such as RADseq have several practical advantages which make them more applicable to conservation studies. Lower sequencing costs, available computational pipelines which reduce the level of bioinformatic expertise required and the option of *de novo* sequence assembly without a reference genome, make RADseq methods a more practical alternative for conservation research on non-model species (Peterson et al., 2012). Furthermore, compared with traditional markers, RADseq has several benefits for conservation genetic applications. Discovery and genotyping of new microsatellite markers is labour-intensive and may be less cost-effective than using RADseq protocols, which facilitate simultaneous discovery and genotyping of SNPs (Morin, Luikart, & Wayne, 2004; Senn et al., 2013). Finally, a recent study has highlighted the improved performance of RRS methods for common conservation genetics/genomics analyses. In comparison to microsatellites and

candidate region sequencing, RRS had the lowest error rates and heterozygosity estimates most closely resembled genome heterozygosity (McLennan, Wright, Belov, Hogg, & Grueber, 2019).

While RADseq is already well established in the field of molecular ecology and evolution, it has not yet reached the same ubiquity in conservation genetics (Shafer et al., 2015). Many conservation genetic analyses still rely on traditional markers, such as microsatellites (Combe et al., 2018). This fact has also resulted in a limited comparability of the results from the current study with other population genetic analyses (e.g. Renan, 2014; Kaczensky *et al.*, 2018). The main limitation for the wider application of RADseq studies in conservation is likely to be the requirement of comparatively large amounts of high-quality DNA (500ng-1µg; Peterson *et al.*, 2012) which can often only be obtained via invasive sampling. There are ethical and practical constraints to invasive sampling of endangered species, and especially for highly elusive species in remote habitats, non-invasive sampling may be the only feasible option (Combe et al., 2018). Therefore RADseq methods are frequently used in areas of conservation where high quality DNA samples are more readily available (illegal wildlife trade, *ex situ* breeding programmes, conservation translocations), but in other areas of conservation genetics they are less commonly applied (e.g. Ng *et al.*, 2017; Nash *et al.*, 2018; Zhang *et al.*, 2019).

Future improvements in DNA recovery from non-invasive samples are expected to broaden applications of RADseq in conservation (Chiou & Bergey, 2015; Stowell et al., 2018). Furthermore, advances in sample storage techniques will improve applicability and allow sample transportation from remote field sites (Stowell et al., 2018). Finally, RADseq methods can aid in the development of highly informative, highly variable SNP panels, which in turn can be used in combination with non-invasive sampling techniques (Bourgeois et al., 2018; Janjua et al., 2019). For example, Schmidt *et al.* (2019) used blood samples and RADseq with subsequent *de novo* assembly to identify a set of highly informative SNPs for the threatened Western rattlesnake (*Crotalus oreganus*). Later, the identified SNPs were successfully recovered from non-invasive samples. This study presents an innovative method for applying RADseq to conservation genetics of non-model species, while minimising the need for invasive sampling.

Despite the promising advances for conservation genetics, the application of RADseq techniques includes several drawbacks which originate in technical aspects of this sequencing method. RADseq protocols are susceptible to bias being introduced during the library preparations (Van Dijk, Jaszczyszyn, & Thermes, 2014) and subsequent bioinformatic processing, particularly in the absence of a reference genome (Fountain et al., 2016; O'Leary et al., 2018). This is a special challenge for conservation genetic studies on non-model species with no high-quality references available.

Careful consideration must be given to the Bioinformatic processing steps, as previous studies demonstrated that even the choice of pipeline used for processing of sequencing data can significantly impact downstream population genetic analyses (Shafer et al., 2017). Furthermore, the optimal setting of parameters used during *de novo* sequence assembly, remains a challenging task and the selection of thresholds can directly affect population genetic estimates (Mastretta-Yanes et al., 2015; Rochette & Catchen, 2017). For example, a comparative analysis on population genetics of Atlantic mackerel identified that higher parameter thresholds in the *de novo* assembly produced lower absolute population differentiation estimates (*Fis*, *Fst*). Moreover, threshold selection impacted the results of Bayesian algorithm-based but not multilinear genetic structure analysis (Rodríguez-Ezpeleta et al., 2016). Finally, while SNP filtering is powerful methods commonly applied to correct for errors introduced during bioinformatic processing, filtering settings themselves can create bias in the data. Linck and Battey (2019) demonstrated that the filtering threshold for minor allele frequency can impact downstream analyses such as population genetic clustering. The above studies demonstrate the potential risk of introducing bias in the data at the various stages of a RADseq project. However, the prevalence of these biases in RADseq studies across different systems requires further investigation. Future research should continue to investigate the introduction of bias at the different stages of a RADseq projects and focus on developing guidelines and bioinformatic tools to reduce error rates. This will be particularly important to advance the application of RADseq methods for conservation genetic projects on non-model species without a high-quality reference genome.

To summarise, while academic research drives towards even more advanced sequencing and analysis techniques including whole genome sequencing, applications

in conservation practice are still limited to a small number of species. However, RADseq studies are already well established for non-model organisms and are expected to become more broadly used in conservation, particularly, when compatibility with non-invasive sampling techniques is improved (Narum et al., 2013). As the use of ddRADseq methods increases, the generation of data sets of robust variants and avoidance of introducing bias remains a major task. The work presented in this thesis adds to a growing number of ddRADseq studies on non-model organisms. Future research could build on the discovered SNP data set for Asiatic wild ass to develop an informative SNP panel to be applied to non-invasive sampling and future population genetic monitoring of the population.

5.2 Deliberate admixture in conservation

In this thesis I analysed the level of admixture in the reintroduced Asiatic wild ass population in Israel, which was founded by individuals of two different subspecies (*Equus hemionus onager*, *Equus hemionus kulan*). While this population was not mixed intentionally, it still provides an interesting study system, since only few cases of deliberate admixture for conservation have been reported to this date (see Chan et al. 2019 for a review). My results demonstrate high levels of subspecies admixture and no indication of a genetic or behavioural barrier to the interbreeding of *E.h.onager* and *E.h.kulan*. In chapter 1 I presented previous studies on the taxonomy of Asiatic wild ass and the controversy about the current subspecies status of onagers and kulans. As part of this debate previous studies have referred to the Israeli population as a hybrid population and proof of successful interbreeding between the subspecies (Bennett et al., 2017; Kaczensky et al., 2015). However, subspecies admixture has not been confirmed genetically until now.

The fact that the two subspecies did interbreed is not unexpected. Previous phylogenetic analyses found onagers and kulans clustering together more closely than other subspecies of Asiatic wild ass (Bennett et al., 2017; Vilstrup et al., 2013). Furthermore, hybridisation between subspecies is not uncommon in mammals. A recent review describes species and subspecies hybrids in many wild European ungulate populations, resulting either from translocations or naturally overlapping ranges (Iacolina, Corlatti, Buzan, Safner, & Šprem, 2019). However, while population genetic analyses can document past admixture events, few studies have investigated

the fitness consequences (Huff, Miller, Chizinski, & Vondracek, 2011; Marshall & Spalton, 2000; Sagvik, Uller, & Olsson, 2005). What makes the results presented in this thesis particularly interesting is that the wild ass population in Israel displays no signs of reduced fitness, which would indicate outbreeding depression. In fact, the female net reproductive rate (R0=1.87) exceeded those observed in other reintroduced populations of Asiatic wild ass (Saltz & Rubenstein, 1995).

There are few reports of outbreeding depression (Huff et al., 2011; Marshall & Spalton, 2000; Sagvik, Uller, & Olsson, 2005) and consequently, predicting the outcome of genetic admixture events remains a challenge. Frankham et al. (2011) have reviewed case studies of outbreeding depression. The authors identified four criteria and predict an increased risk of outbreeding depression if one of these is true. If the source populations are of different species, are adapted to different environments, display fixed chromosomal differences, or exchanged no genes in the last 500 years, crossing may result in reduced fitness and should be avoided.

Onagers and kulans inhabit similar habitats with minor differences in temperature extremes and mean annual precipitation (Denzau & Denzau, 1999). No clear estimates of the divergence time between the subspecies exist, however, gene flow between the two mitochondrial clades introduced in Israel could date back as far as 3500 years, based on present records of their cooccurrence in the same geographic area (Bennett et al., 2017; Oakenfull, Lim, & Ryder, 2000). Finally, onagers and kulans display chromosomal differences, resulting in distinct and intermediate karyotypes (E.h.onager 2n=55,56; E.h.kulan 2n=55,54) (Ryder & Chemnick, 1990). Hence, based on the criteria by Frankham et al. (2011) an increased risk of outbreeding depression would be predicted for onagers and kulans.

A reintroduced population of Arabian oryx (Oryx leucoryx) provides an interesting comparison: For this reintroduction individuals were also sourced from different founder stock (Marshall & Spalton, 2000). The desert-adapted species is not differentiated into different subspecies but was formerly distributed across a large geographical area. Prior to its extinction in the wild, captive breeding populations were established with individuals sourced from different countries. Like the onagers and kulans, the Arabian oryx displays a Robertsonian translocation resulting in different karyotypes (2n=56, 57, 58). Consequently, the Arabian oryx fulfils similar criteria to the

Asiatic wild ass in Israel. However, unlike the Asiatic wild ass in Israel, the admixture of Arabian oryx from different source populations resulted in outbreeding depression with the established population displaying reduced in juvenile survival. This example highlights the difficulty of predicting outbreeding depression based on indicators of genetic or evolutionary distance. While literature reviews like the one by Frankham et al. (2011) provide valuable guidelines for population management, reliable predictors are still missing and future research should investigate fitness consequences of genetic admixture to improve our understanding of outbreeding depression and inform future conservation management. The results of this thesis provide novel insights relevant for the management and conservation of Asiatic wild ass populations. While other arguments against a combined management of kulans and onagers remain to be discussed (e.g. genetic purity, dilution and potential loss of unique adaptations), the results presented here demonstrate that genetic or behavioural barriers are unlikely to compromise a mixed stock management approach in *in situ* and *ex situ* conservation.

Historically hybrids were considered to be of little to no conservation value, however, recently there has been a shift in attitude and the discussion about hybrids has become more differentiated (Allendorf et al., 2001; Hamilton & Miller, 2016). There is a general consensus that naturally occurring hybrids deserve protection, however, the ethical and legal status of hybrids created by human actions (intentional or unintentional) remains disputed (Allendorf et al., 2001; Genovart, 2009; Jackiw, Mandil, & Hager, 2015). Recently, this debate has gained another dimension, with a growing number of authors advocating the use of intentional hybridisation for conservation purposes (Chan, Hoffmann, & van Oppen, 2019; van Oppen, Oliver, Putnam, & Gates, 2015; Weeks et al., 2015, 2016). The population genomic analysis I presented in this thesis, suggests that the genetic makeup of the reintroduced population in Israel has been improved by the subspecies hybridisation (heterozygote excess). These results add to a growing number of studies pointing out the potential conservation benefits of intra- or interspecific hybridisation (Bell et al., 2019; Chan et al., 2019; Hamilton & Miller, 2016; vonHoldt, Brzeski, Wilcove, & Rutledge, 2018; Weeks et al., 2016).

Mixing of different genetic lineages can improve the chances of successful population establishment and persistence by increasing genetic diversity and adaptive potential in

the population (Biebach & Keller, 2012; Tordoff & Redig, 2001). However, when the established population is not integrated in an existing metapopulation but remains isolated, it is expected that any positive effects of outbreeding will fade over time. For example, divergent genomic signatures were found in a small and isolated population of Florida panthers (*Puma concolor ssp.*) as a result of historic admixture (Saremi et al., 2019). One individual's genome displayed relatively high levels of heterozygosity caused by past subspecies admixture, however, it also contained long runs of homozygosity, indicative of a recent inbreeding event (Saremi et al., 2019). The authors argue that in small populations the potential benefits of admixture can be eroded quickly by unavoidable inbreeding. Therefore, in the case of isolated populations, continued translocations may be necessary to prevent inbreeding depression despite admixture. In this thesis I demonstrated the high individual admixture levels and relatively high heterozygosity in the reintroduced wild ass in Israel. However, considering the small number of founders (N=11), high individual inbreeding levels are expected. Future research could apply whole genome sequencing methods to investigate individual inbreeding. These results would provide valuable information for the populations' conservation status and also produce insights on the lasting effects of hybridisation in reintroduced populations (Tallmon et al., 2004).

Aside from reintroductions, hybridisation can be used for the recovery of autochthonous populations. If a population has become genetically depleted, genetic rescue, the introduction of individuals from a genetically divergent population, can help to counteract inbreeding depression (Tallmon et al., 2004). This is not a novel technique in conservation, however, increasing rates of habitat fragmentation and population isolation may warrant a more widespread application (Weeks *et al.*, 2015; Weeks, Stoklosa and Hoffmann, 2016; Ralls *et al.*, 2018, but see Tallmon, Luikart and Waples, 2004). Finally, hybridisation may be applied as a precautionary conservation measure against expected future threats to a species (Biebach & Keller, 2012; Tordoff & Redig, 2001). Intentional hybridisation can be used to increase the evolutionary potential of a population (Chan et al., 2019; Hamilton & Miller, 2016). Furthermore, by means of targeted gene flow, key genetic traits that are expected to have conservation benefits can be introduced to populations by moving individuals with desired traits (Kelly & Phillips, 2019). In Australia, northern quolls (*Dasyurus hallucatus*) have

become threatened with extinction due to the introduction of the poisonous cane toad (*Rhinella marina)* (Kelly & Phillips, 2018). Naïve quolls feeding on the toad are killed, however, in some populations, individuals have evolved to avoid them. A common garden experiment has shown that this avoidance behaviour has a genetic component and thus can be inherited (Kelly & Phillips, 2018). The translocation of individuals from 'toad-smart' populations could facilitate the spread of this behaviour. While the authors acknowledge the risk of outbreeding depression, they argue that targeted gene flow may be the best option to ensure the survival of the species.

The growing debate about deliberate admixture for conservation is highly controversial as it challenges some fundamental ideas of conservation practice. Furthermore, the potential risks associated with deliberate admixture complicate decision making. Future research is needed to address the gaps in our knowledge about the likelihood and prevalence of outbreeding depression, to facilitate informed and effective conservation action. Specifically, controlled experiments, conducted in a lab or quasi-*in situ* settings (Shemesh et al., 2018; van Oppen et al., 2015) will help further our knowledge of genetic barriers and incompatibilities. Nevertheless, for many species, especially large-bodied mammals, experiments under controlled conditions are often not feasible. Therefore, case studies, like the one presented in this thesis, are of great value and can provide important information for similar species (Chan et al., 2019).

5.3 Landscape genetics in species reintroductions

Landscape genetics is a fast-developing multidisciplinary field, which combines concepts from spatial ecology and population genetics (Manel et al., 2003). It is still a relatively novel discipline which continues to evolve and expand into new research areas. In this thesis I applied landscape genetic methods to investigate environmental effects on genetic relatedness in the reintroduced Asiatic wild ass in Israel. I combined spatial telemetry data with population genomic analysis and found that neither landscape resistance nor geographic distance had an impact on genetic relatedness.

Until today, only a limited number of studies have applied landscape genetics to reintroduced populations (e.g. Mucci et al., 2010; Wasserman et al., 2010; Williams and Scribner, 2010; Moraes et al., 2018). One constraint is a potential time lag effect

between patterns of genetic association and landscape configuration. This is because the genetic structure of a population is shaped by both past and contemporary patterns of gene flow (Waits, Cushman, & Spear, 2016). For example, in the case of recent landscape change, any observed genetic structure may be more representative of historic rather than present-day landscape connectivity. Indeed, some studies have reported lag times of 20-50 years, with genetic structure correlating more strongly with past environmental conditions (Holzhauer et al., 2006; Spear and Storfer, 2008). Therefore, the genetic configuration in a recently reintroduced population may not yet reflect existing gene flow at the release site but instead may be more strongly influenced by demographic processes prior to the reintroduction and founder effects during spatial range expansion (Excoffier & Ray, 2008).

Despite these potential limitations, landscape genetic analysis may prove to be a vital tool in conservation management of reintroduced populations. Fast post-release population growth is critical to minimise the loss of genetic diversity due to drift (Seddon & Armstrong, 2016). Additionally, the integration of a reintroduced population into a larger metapopulation is an important measure to prevent genetic isolation (Seddon & Armstrong, 2016; Segelbacher, Höglund, & Storch, 2003). Consequently, habitat characteristics limiting range expansion and landscape connectivity are threatening reintroduction success (la Morgia, Malenotti, Badino, & Bona, 2011; Williams et al., 2002). There are several case studies where landscape genetic analysis has been applied successfully to identify factors restricting gene flow in reintroduced populations (Howell et al., 2016; Moraes et al., 2018; Ziółkowska et al., 2016). For example, Neuwald and Templeton (2013) reported that forest areas in the Missouri Ozarks acted as dispersal barriers in translocated populations of collared lizard (*Crotaphytus collaris collaris*). Genetic isolation caused local extinctions and only after the adoption of a woodland burning management regime did the metapopulation recover. The study highlights the importance of landscape genetic monitoring to inform habitat management for reintroductions.

Future studies on reintroduced populations should apply a combined approach of population genetic and landscape genetic monitoring, beginning immediately after release. Furthermore, by using genetic markers, with differing mutation rates (mitochondrial DNA markers, microsatellites, SNPs), inferences about historic and

recent landscape effects on gene flow could be improved (Waits, Cushman and Spear, 2016). Additionally, estimates of landscape effects on relatedness could be enhanced by modelling gene flow more directly. In this thesis I have used generic individual movement data from GPS collars for a small set of adult individuals. However, gene flow may be mediated by rare, sex- and age-specific dispersal events (Templeton, Brazeal, & Neuwald, 2011; Van Dyck & Baguette, 2005). Future research could apply long-term telemetry monitoring of individuals from different age groups. This would allow identification of potentially age- and sex-specific dispersal movements and hence create landscape resistance surfaces which are more informative about gene flow (Killeen et al., 2014). Alternatively, non-invasive genetic sampling techniques in combination with genetic mark-recapture methods could provide a better understanding of individual movements across the landscape. While the latter approach would not provide direct insights into fine-scale individual movement patterns and path selection, it would provide an indirect measure of gene flow across the study area.

While landscape genetics is an important tool for post-release monitoring and management, this field of research is expected to become increasingly valuable during the planning stage of reintroductions (Flanagan, Forester, Latch, Aitken, & Hoban, 2018). In the face of climate change and increasing rates of land conversion, habitats are changing at an unprecedented rate. Modelling future distributions of endangered species is a crucial precursor for their effective conservation (Bar-David, Saltz, & Dayan, 2005; Bar-David, Saltz, Dayan, & Shkedy, 2008). Many species are expected to shift their ranges in response to changing climatic and environmental conditions (IPCC, 2014). Therefore, landscape genetic studies are an important tool to predict future barriers to gene flow and adjust management to promote connectivity and long-term persistence of the population (Razgour, 2015). Such models will be of great importance for future conservation reintroductions to ensure that release sites provide suitable habitat under expected climate change scenarios (Draper, Marques, & Iriondo, 2019; Houde, Garner, & Neff, 2015). Finally, landscape genomics enables the analysis of functional adaptive genetic variation across a species' range (Schwartz, McKelvey, Cushman, & Luikart, 2010). This knowledge can assist the identification of optimal source lineages for reintroductions and select founders which display environmental

adaptations that match current or future conditions at the restoration site (Radinger et al., 2017; Shryock et al., 2017).

5.4 Final conclusions

Species reintroductions have become an increasingly used tool in conservation practice. They are a powerful measure for the recovery of endangered species. However, in the face of the current species extinction crisis and the unprecedented rate of environmental change, traditional approaches might not suffice to reach long-term conservation goals. The reintroduction of Asiatic wild ass in Israel provides a unique case study since the population was founded by individuals from two different subspecies. While the reintroduction has been generally considered successful, until now it was not known whether individuals of the two subspecies interbred. In this thesis I demonstrated high levels of subspecies admixture and my results suggested that the reintroduced population may have benefited from this admixture. My research contributes to a small number of empirical studies on hybridisation for conservation.

The long-term success of reintroductions is strongly dependent on population genetic viability and species-habitat interactions. Previously, little was known about how reintroduced individuals disperse across the new habitat. Using the example of the Asiatic wild ass in Israel, I have shown how landscape features and distribution of key resources affect individual habitat use and movement across the landscape. Finally, I have highlighted the crucial need for continued post-release management and genetic monitoring in reintroduced populations to increase effective population size and prevent population fragmentation.

The here-presented results highlight areas for improvement in species reintroductions. Individual-habitat interactions and long-term genetic monitoring should be integrated in recovery projects from an early stage. The results of my admixture analysis advance our limited knowledge of a potentially powerful yet highly controversial conservation practice. Future research on a wide range of taxa is needed to create an extensive knowledge base and facilitate informed conservation decision making.

6. Bibliography

Abrahms, B., Sawyer, S. C., Jordan, N. R., McNutt, J. W., Wilson, A. M., & Brashares, J. S. (2017). Does wildlife resource selection accurately inform corridor conservation? *Journal of Applied Ecology*, *54*(2), 412–422. https://doi.org/10.1111/1365-2664.12714

Aebischer, N. J., Robertson, P. A., & Kenward, R. E. (1993). Compositional analysis of habitat use from animal radio-tracking data. *Ecology*, *74*(5), 1313–1325.

Alacs, E. A., Georges, A., Fitzsimmons, N. N., & Robertson, J. (2009). DNA detective : a review of molecular approaches to wildlife forensics. *Forensic Science, Medicine, and Pathology*, *6*(3), 180–194. https://doi.org/10.1007/s12024-009-9131-7

Alcala, N., Jensen, J. D., Telenti, A., & Vuilleumier, S. (2016). The genomic signature of population reconnection following isolation: From theory to HIV. *G3 Genes|Genomes|Genetics*, *6*(1), 107–120. https://doi.org/10.1534/g3.115.024208

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, *11*(10), 697–709. https://doi.org/10.1038/nrg2844

Allendorf, F. W., Leary, R. F., Spruell, P., & Wenburg, J. K. (2001). The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution*, *16*(11), 613–622. https://doi.org/10.1016/0305-4403(79)90038-4

Altman, A. (2016). *Female group size in Asiatic wild ass*. Ben-Gurion University of the Negev, Israel.

Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G., & Hohenlohe, P. A. (2016). Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), 81–92. https://doi.org/10.1038/nbt.3301.Mammalian

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Armstrong, D. P., & Seddon, P. J. (2008). Directions in reintroduction biology. *Trends in Ecology and Evolution*, *23*(1), 20–25. https://doi.org/10.1016/j.tree.2007.10.003

Ashley, M. C. (2004). Population genetics of feral horses: Implications of behavioral isolation. *Journal of Mammalogy*, *85*(4), 611–617. https://doi.org/10.1644/brb-123

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., … Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*(10). https://doi.org/10.1371/journal.pone.0003376

Bar-David, S., Saltz, D., & Dayan, T. (2005). Predicting the spatial dynamics of a reintroduced population: The Persian fallow deer. *Ecological Applications*, *15*(5), 1833–1846. https://doi.org/10.1890/04-0798

Bar-David, S., Saltz, D., Dayan, T., & Shkedy, Y. (2008). Using spatially expanding populations as a tool for evaluating landscape planning: The reintroduced Persian fallow deer as a case study. *Journal for Nature Conservation*, *16*(3), 164–174. https://doi.org/10.1016/j.jnc.2008.09.004

Baxter, S. W., Davey, J. W., Johnston, J. S., Shelton, A. M., Heckel, D. G., Jiggins, C. D., & Blaxter, M. L. (2011). Linkage mapping and comparative genomics using next-generation rad sequencing of a non-model organism. *PLoS ONE*, *6*(4). https://doi.org/10.1371/journal.pone.0019315

Bell, D. A., Robinson, Z. L., Funk, W. C., Fitzpatrick, S. W., Allendorf, F. W., Tallmon, D. A., & Whiteley, A. R. (2019). The exciting potential and remaining uncertainties of genetic rescue. *Trends in Ecology & Evolution*, *34*(12), 1070–1079. https://doi.org/10.1016/j.tree.2019.06.006

Bennett, E. A., Champlot, S., Peters, J., Arbuckle, B. S., Guimaraes, S., Pruvost, M., … Geigl, E. M. (2017). Taming the late Quaternary phylogeography of the Eurasiatic wild ass through ancient and modern DNA. *PLoS ONE*, *12*(4), 1–28. https://doi.org/10.1371/journal.pone.0174216

Benton, T. G., & Bowler, D. E. (2012). Dispersal in invertebrates: influences on individual decisions. In J. Colbert, M. Baguette, T. G. Benton, & J. M. Bullock (Eds.), *Dispersal ecology and evolution* (pp. 41–49). Oxford: Oxford University Press.

Bergl, R. A., & Vigilant, L. (2007). Genetic analysis reveals population structure and recent migration within the highly fragmented range of the Cross River gorilla (Gorilla gorilla diehli). *Molecular Ecology*, *16*(3), 501–516. https://doi.org/10.1111/j.1365-294X.2006.03159.x

Berkman, L. K., Nielsen, C. K., Roy, C. L., & Heist, E. J. (2013). Resistance is futile: Effects of landscape features on gene flow of the northern bobwhite. *Conservation Genetics*, *14*(2), 323–332. https://doi.org/10.1007/s10592-013-0471-1

Berry, O., Algar, D., Angus, J., Hamilton, N., Hilmer, S., & Sutherland, D. (2012). Genetic tagging reveals a significant impact of poison baiting on an invasive species. *Journal of Wildlife Management*, *76*(4), 729–739. https://doi.org/10.1002/jwmg.295

Biebach, I., & Keller, L. F. (2009). A strong genetic footprint of the re-introduction history of Alpine ibex (Capra ibex ibex). *Molecular Ecology*, *18*, 5046–5058. https://doi.org/10.1111/j.1365-294X.2009.04420.x

Biebach, I., & Keller, L. F. (2012). Genetic variation depends more on admixture than number of founders in reintroduced Alpine ibex populations. *Biological Conservation*, *147*(1), 197–203. https://doi.org/10.1016/j.biocon.2011.12.034

Bourgeois, S., Senn, H., Kaden, J., Taggart, J. B., Ogden, R., Jeffery, K. J., … McEwing, R. (2018). Single-nucleotide polymorphism discovery and panel characterization in the African forest elephant. *Ecology and Evolution*, *8*(4), 2207–2217. https://doi.org/10.1002/ece3.3854

Boyd, L., Scorolli, A., Nowzari, H., & Bouskila, A. (2016). Social Organization of Wild Equids. In P. Kaczensky & J. I. Ransom (Eds.), *Wild Equids: Ecology, Management and Conservation* (pp. 7–22). Baltimore, MD: Johns Hopkins University Press.

Braude, S., & Templeton, A. R. (2010). Understanding the multiple meanings of ‘ inbreeding ’ and ‘ effective size ’ for genetic management of African rhinoceros populations. *African Journal Ecology*, *47*(4), 546–555.

Brekke, P., Bennett, P. M., Wang, J., Pettorelli, N., & Ewen, J. G. (2010). Sensitive males: Inbreeding depression in an endangered bird. *Proceedings of the Royal Society B: Biological Sciences*, *277*(1700), 3677–3684. https://doi.org/10.1098/rspb.2010.1144

Bresadola, L., Link, V., Buerkle, C. A., Lexer, C., & Wegmann, D. (2020). Estimating and accounting for genotyping errors in RAD-seq experiments. *Molecular Ecology Resources*, (July 2019), 1–15. https://doi.org/10.1111/1755-0998.13153

Broders, H. G., Mahoney, S. P., Montevecchi, W. A., & Davidson, W. S. (1999). Population genetic structure and the effect of founder events on the genetic variability of moose, Alces alces, in Canada. *Molecular Ecology*, *8*, 1309–1315. https://doi.org/10.1007/s10344-005-0087-y

Bull, J. K., Heurich, M., Saveljev, A. P., Schmidt, K., & Förster, D. W. (2016). The effect of reintroductions on the genetic variability in Eurasian lynx populations : the cases of Bohemian – Bavarian and Vosges – Palatinian populations. *Conservation Genetics*, *17*, 1229–1234. https://doi.org/10.1007/s10592-016-0839-0

Buttigieg, P. L., & Ramette, A. (2014). A guide to statistical analysis in microbial ecology: a community-focused, living review of multivariate data analyses. *FEMS Microbiology Ecology*, *90*(3), 543–550. https://doi.org/10.1111/1574-6941.12437

Calenge, C. (2006). The package "adehabitat" for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling*, *197*(3–4), 516–519. https://doi.org/10.1016/j.ecolmodel.2006.03.017

Caro, T. M., & Laurenson, M. K. (1994). Ecological and genetic factors in conservation: A cautionary tale. *Science*, Vol. 263, pp. 485–486. https://doi.org/10.1126/science.8290956

Carvajal-Rodríguez, A. (2008). Simulation of genomes: A review. *Current Genomics*, *9*, 155–159.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3 Genes|Genomes|Genetics*, *1*(3), 171–182. https://doi.org/10.1534/g3.111.000240

Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140. https://doi.org/10.1111/mec.12354

Chan, W. Y., Hoffmann, A. A., & van Oppen, M. J. H. (2019). Hybridization as a conservation management tool. *Conservation Letters*, *12*(5), e12652. https://doi.org/10.1111/conl.12652

Charlesworth, D., & Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics*, *18*(1), 237–268. https://doi.org/10.1146/annurev.es.18.110187.001321

Chesser, R. K. (1991). Influence of gene flow and breeding tactics on gene diversity within populations. *Genetics*, *129*(2), 573–583.

Chesser, R. K., Rhodes, O. E., Sugg, D. W., & Schnabel, A. (1993). Effective sizes for subdivided populations. *Genetics*, *135*(4), 1221–1232.

Chiou, K., & Bergey, C. (2015). FecalSeq: Methylation-based enrichment for noninvasive population genomics from feces. *BioRxiv*. https://doi.org/10.1101/032870

Choudhry, S., Coyle, N. E., Tang, H., Salari, K., Lind, D., Clark, S. L., … Burchard, E. G. (2006). Population stratification confounds genetic association studies among Latinos. *Human Genetics*, *118*(5), 652–664. https://doi.org/10.1007/s00439-005-0071-3

Combe, F. J., Taylor-Cox, E., Fox, G., Sandri, T., Davis, N., Jones, M. J., … Harris, W. E. (2018). Rapid isolation and characterization of microsatellites in the critically endangered mountain bongo (Tragelaphus eurycerus isaaci). *Journal of Genetics*, *97*(2), 549–553. https://doi.org/10.1007/s12041-018-0922-z

Coulson, T. N., Pemberton, J. M., Albon, S. D., Beaumont, M., Marshall, T. C., Slate, J., … Clutton-Brock, T. H. (1998). Microsatellites reveal heterosis in red deer. *Proceedings of the Royal Society B: Biological Sciences*, *265*(1395), 489–495. https://doi.org/10.1098/rspb.1998.0321

Crnokrak, P., & Roff, D. A. (1999). Inbreeding depression in the wild. *Heredity*, *83*(3), 260–270. https://doi.org/10.1038/sj.hdy.6885530

Cushman, S. A. (2009). Animal Movement Data: GPS Telemetry, Autocorrelation and the Need for Path-Level Analysis. In S. A. Cushman & F. Huettmann (Eds.), *Spatial Complexity, Informatics, and Wildlife Conservation* (pp. 131–149). https://doi.org/10.1007/978-4-431-87771-4

Danin, A. (1999). Desert rocks as plant refugia in the Near East. *Botanical Review*, *65*(2), 93–170.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*(7), 499–510. https://doi.org/10.1038/nrg3012

Davidson, A., Carmel, Y., & Bar-David, S. (2013). Characterizing wild ass pathways using a non-invasive approach: Applying least-cost path modelling to guide field surveys and a model selection analysis. *Landscape Ecology*, *28*(8), 1465–1478. https://doi.org/10.1007/s10980-013-9915-8

De Barba, M., Waits, L. P., Garton, E. O., Genovesi, P., Randi, E., Mustoni, A., … Groffs, C. (2010). The power of genetic monitoring for studying demography, ecology and genetics of a reintroduced brown bear population. *Molecular Ecology*, *19*(18), 3938–3951. https://doi.org/10.1111/j.1365-294X.2010.04791.x

Denzau, G., & Denzau, H. (1999). *Wildesel*. Stuttgart: Jan Thorbeck Verlag GmbH & Co.

DeSalle, R. (2005). Genetics at the brink of extinction. *Heredity*, *94*(4), 386–387. https://doi.org/10.1038/sj.hdy.6800641

DeSalle, Rob, & Amato, G. (2004). The expansion of conservation genetics. *Nature Reviews Genetics*, *5*(9), 702–712. https://doi.org/10.1038/nrg1425

Díaz-Arce, N., & Rodríguez-Ezpeleta, N. (2019). Selecting RAD-Seq data analysis parameters for population genetics: The more the better? *Frontiers in Genetics*,

10, 1–10. https://doi.org/10.3389/fgene.2019.00533

Diaz, S., Settele, J., & Brondízio, E. (2019). *Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosyste*. IPBES secretariat, Bonn, Germany.

DiBattista, J. D., Saenz-Agudelo, P., Piatek, M. J., Wang, X., Aranda, M., & Berumen, M. L. (2017). Using a butterflyfish genome as a general tool for RAD-Seq studies in specialized reef fish. *Molecular Ecology Resources*, *17*(6), 1330–1341. https://doi.org/10.1111/1755-0998.12662

Dierickx, E. G., Shultz, A. J., Sato, F., Hiraoka, T., & Edwards, S. V. (2015). Morphological and genomic comparisons of Hawaiian and Japanese Black-footed Albatrosses (Phoebastria nigripes) using double digest RADseq: Implications for conservation. *Evolutionary Applications*, *8*(7), 662–678. https://doi.org/10.1111/eva.12274

Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: Re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, *14*(1), 209–214. https://doi.org/10.1111/1755-0998.12157

Draper, D., Marques, I., & Iriondo, J. M. (2019). Species distribution models with field validation, a key approach for successful selection of receptor sites in conservation translocations. *Global Ecology and Conservation*, *19*. https://doi.org/10.1016/j.gecco.2019.e00653

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Eaton, D. A. R. (2014). PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*(13), 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Edmands, S. (2007). Between a rock and a hard place: Evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology*, *16*(3), 463–475. https://doi.org/10.1111/j.1365-294X.2006.03148.x

Eldridge, F., & Blazak, W. F. (1976). Horse, ass, and mule chromosomes. *Journal of Heredity*, *67*(6), 361–367.

European Association of Zoos and Aquaria. (2019). Specialist Programmes. Retrieved December 30, 2019, from https://www.eaza.net/conservation/programmes/

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, *9*(10), e1003905. https://doi.org/10.1371/journal.pgen.1003905

Excoffier, L., & Foll, M. (2011). fastsimcoal: a continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, *27*(9), 1332–1334. https://doi.org/10.1093/bioinformatics/btr124

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*(3), 564–567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Excoffier, L., & Ray, N. (2008). Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology and Evolution*, *23*(7), 24–28.

Fischer, J., & Lindenmayer, D. B. (2000). An assessment of the published results of animal relocations. *Biological Conservation*, *96*(1), 1–11. https://doi.org/10.1016/S0006-3207(00)00048-3

Flanagan, S. P., Forester, B. R., Latch, E. K., Aitken, S. N., & Hoban, S. (2018). Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evolutionary Applications*, *11*(7), 1035–1052. https://doi.org/10.1111/eva.12569

Fountain, E. D., Pauli, J. N., Reid, B. N., Palsbøll, P. J., & Peery, M. Z. (2016). Finding the right coverage: the impact of coverage and sequence quality on single nucleotide polymorphism genotyping error rates. *Molecular Ecology Resources*, *16*(4), 966–978. https://doi.org/10.1111/1755-0998.12519

Frankel, O. H. (1974). Genetic conservation: our evolutionary responsibility. *Genetics*, *78*(1), 53–65. https://doi.org/10.1093/sysbio/43.3.387

Frankel, O. H., & Soulé, M. E. (1981). *Conservation and evolution*. Cambridge: Cambridge University Press.

Frankham, R. (2005). Genetics and extinction. *Biological Conservation*, *126*(2), 131–140. https://doi.org/10.1016/j.biocon.2005.05.002

Frankham, R., Ballou, J. D., Briscoe, D. A., & McInnes, K. H. (2002). Effects of Population Size Reduction. In D. A. Briscoe, J. D. Ballou, & R. Frankham (Eds.), *Introduction to Conservation Genetics* (pp. 225–226). Cambridge: Cambridge University Press.

Frankham, R., Ballou, J. D., Eldridge, M. D. B., Lacy, R. C., Ralls, K., Dudash, M. R., & Fenster, C. B. (2011). Predicting the probability of outbreeding depression. *Conservation Biology*, *25*(3), 465–475. https://doi.org/10.1111/j.1523-1739.2011.01662.x

Frankham, R., Briscoe, D. A., & Ballou, J. D. (2002). *Introduction to Conservation Genetics*. Cambridge: Cambridge University Press.

Frankham, R., & Ralls, K. (1998). Inbreeding leads to extinction. *Nature*, *392*(6675), 441–442.

Franklin, I. R. (1980). Evolutionary changes in small populations. In *Conservation Biology: An Evolutionary Ecological Perspective*. Sunderland, MA: Sinauer Associates.

Gagnon, M., Yannic, G., Perrier, C., & Côté, S. D. (2019). No evidence of inbreeding

depression in fast declining herds of migratory caribou. *Journal of Evolutionary Biology*, 1–14. https://doi.org/10.1111/jeb.13533

Garamszegi, L. Z., & Møller, A. P. (2010). Effects of sample size and intraspecific variation in phylogenetic comparative studies: A meta-analytic review. *Biological Reviews*, *85*(4), 797–805. https://doi.org/10.1111/j.1469-185X.2010.00126.x

Garner, B. A., Hand, B. K., Amish, S. J., Bernatchez, L., Foster, J. T., Miller, K. M., … Luikart, G. (2016). Genomics in conservation: Case studies and bridging the gap between data and application. *Trends in Ecology and Evolution*, *31*(2), 81–83. https://doi.org/10.1016/j.tree.2015.10.009

Gattepaille, L. M., Jakobsson, M., & Blum, M. G. B. (2013). Inferring population size changes with sequence and SNP data: Lessons from human bottlenecks. *Heredity*, *110*(5), 409–419. https://doi.org/10.1038/hdy.2012.120

Geigl, E.-M., & Grange, T. (2012). Eurasian wild asses in time and space: Morphological versus genetic diversity. *Annals of Anatomy - Anatomischer Anzeiger*, *194*(1), 88–102. https://doi.org/10.1016/J.AANAT.2011.06.002

Genovart, M. (2009). Natural hybridization and conservation. *Biodiversity and Conservation*, *18*(6), 1435–1439. https://doi.org/10.1007/s10531-008-9550-x

Giotto, N., Gerard, J. F., Ziv, A., Bouskila, A., & Bar-David, S. (2015). Space-use patterns of the Asiatic Wild Ass (Equus hemionus): Complementary insights from displacement, recursion movement and habitat selection analyses. *PLoS ONE*, *10*(12), 1–21. https://doi.org/10.1371/journal.pone.0143279

Gitzen, R. A., Keller, B. J., Miller, M. A., Goetz, S. M., Steen, D. A., Jachowski, D. S., … Millspaugh, J. J. (2016). Effective and purposeful monitoring of species reintroductions. In D. S. Jachowski, J. J. Millspaugh, P. L. Angermeier, & R. Slotow (Eds.), *Reintroduction of fish and wildlife populations* (pp. 283–317). Oakland, CA: California University Press.

Gompert, Z., & Buerkle, C. A. (2010). Introgress: A software package for mapping components of isolation in hybrids. *Molecular Ecology Resources*, *10*(2), 378–384. https://doi.org/10.1111/j.1755-0998.2009.02733.x

Gottsberger, B., & Mayer, F. (2019). Dominance effects strengthen premating hybridization barriers between sympatric species of grasshoppers (Acrididae, Orthoptera). *Journal of Evolutionary Biology*, *32*(9), 921–930. https://doi.org/10.1111/jeb.13490

Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, *5*, 184–186. https://doi.org/10.1111/j.1471-8278

Gouin, A., Legeai, F., Nouhaud, P., Whibley, A., Simon, J. C., & Lemaitre, C. (2015). Whole-genome re-sequencing of non-model organisms: Lessons from unmapped reads. *Heredity*, *114*(5), 494–501. https://doi.org/10.1038/hdy.2014.85

Graham, C. F., Glenn, T. C., Mcarthur, A. G., Boreham, D. R., Kieran, T., Lance, S., … Somers, C. M. (2015). Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Molecular Ecology Resources*, *15*(6), 1304–1315.

https://doi.org/10.1111/1755-0998.12404

Grauer, J. A., Gilbert, J. H., Woodford, J. E., Eklund, D., Anderson, S., & Pauli, J. N. (2017). Unexpected genetic composition of a reintroduced carnivore population. *Biological Conservation*, *215*, 246–253. https://doi.org/10.1016/j.biocon.2017.09.016

Greenbaum, G., Renan, S., Templeton, A. R., Bouskila, A., Saltz, D., Rubenstein, D. I., & Bar-David, S. (2018). Revealing life-history traits by contrasting genetic estimations with predictions of effective population size. *Conservation Biology*, *32*(4), 817–827. https://doi.org/10.1111/cobi.13068

Greig, J. C. (1979). Principles of genetic conservation in relation to wildlife management in Southern Africa. *International Symposium on African Wildlife, Research and Management*, *9*, 57–78.

Grossen, C., Biebach, I., Angelone-Alasaad, S., Keller, L. F., & Croll, D. (2018). Population genomics analyses of European ibex species show lower diversity and higher inbreeding in reintroduced populations. *Evolutionary Applications*, *11*(2), 123–139. https://doi.org/10.1111/eva.12490

Groves, C. P., & Mazák, V. (1967). On some taxonomic problems of Asiatic wild asses; with the description of a new subspecies (Perissodactyla; Equidae). *Zeitschrift Für Säugetierkunde*, *32*(6), 321–355.

Gueta, T., Templeton, A. R., & Bar-David, S. (2014). Development of genetic structure in a heterogeneous landscape over a short time frame: the reintroduced Asiatic wild ass. *Conservation Genetics*, *15*(5), 1231–1242. https://doi.org/10.1007/s10592-014-0614-z

Haasl, R. J., & Payseur, B. A. (2011). Multi-locus inference of population structure: A comparison between single nucleotide polymorphisms and microsatellites. *Heredity*, *106*(1), 158–171. https://doi.org/10.1038/hdy.2010.21

Hamilton, J. A., & Miller, J. M. (2016). Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conservation Biology*, *30*(1), 33–41. https://doi.org/10.1111/cobi.12574

Hammer, S. E., Schwammer, H. M., & Suchentrunk, F. (2008). Evidence for introgressive hybridization of captive markhor (Capra falconeri) with domestic goat: Cautions for reintroduction. *Biochemical Genetics*, *46*(3–4), 216–226. https://doi.org/10.1007/s10528-008-9145-y

He, X., Johansson, M. L., & Heath, D. D. (2016). Role of genomics and transcriptomics in selection of reintroduction source populations. *Conservation Biology*, *30*(5), 1010–1018. https://doi.org/10.1111/cobi.12674

Hedrick, P. W., & Kalinowski, S. T. (2000). Inbreeding depression in conservation biology. *Annual Review of Ecology and Systematics*, *31*(1), 139–162.

Henley, S. R., Ward, D., & Schmidt, I. (2007). Habitat selection by two desert-adapted ungulates. *Journal of Arid Environments*, *70*(1), 39–48. https://doi.org/10.1016/j.jaridenv.2006.12.007

Hohenlohe, P. A., Amish, S. J., Catchen, J. M., Allendorf, F. W., & Luikart, G. (2011).

Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular Ecology Resources*, *11*, 117–122. https://doi.org/10.1111/j.1755-0998.2010.02967.x

Houde, A. L. S., Garner, S. R., & Neff, B. D. (2015). Restoring species through reintroductions : Strategies for source population selection. *Restoration Ecology*, *23*(6), 746–753. https://doi.org/10.1111/rec.12280

Howard, A. M., Nibbelink, N. P., Madden, M., Young, L. A., Bernardes, S., & Fragaszy, D. M. (2015). Landscape influences on the natural and artificially manipulated movements of bearded capuchin monkeys. *Animal Behaviour*, *106*, 59–70. https://doi.org/10.1016/j.anbehav.2015.05.003

Howell, P. E., Koen, E. L., Williams, B. W., Roloff, G. J., & Scribner, K. T. (2016). Contiguity of landscape features pose barriers to gene flow among American marten (Martes americana) genetic clusters in the Upper Peninsula of Michigan. *Landscape Ecology*, *31*(5), 1051–1062. https://doi.org/10.1007/s10980-015-0331-0

Huang, J., Zhao, Y., Bai, D., Shiraigol, W., Li, B., Yang, L., … Dugarjaviin, M. (2015). Donkey genome and insight into the imprinting of fast karyotype evolution. *Scientific Reports*, *5*, 1–9. https://doi.org/10.1038/srep14106

Huff, D. D., Miller, L. M., Chizinski, C. J., & Vondracek, B. (2011). Mixed-source reintroductions lead to outbreeding depression in second-generation descendents of a native North American fish. *Molecular Ecology*, *20*(20), 4246–4258. https://doi.org/10.1111/j.1365-294X.2011.05271.x

Illumina. (2020). Read length recommendations. Retrieved from https://www.illumina.com/science/technology/next-generation-sequencing/plan-experiments/read-length.html?langsel=/us/

Ito, H., Ogden, R., Langenhorst, T., & Inoue-Murayama, M. (2017). Contrasting results from molecular and pedigree-based population diversity measures in captive zebra highlight challenges facing genetic management of zoo populations. *Zoo Biology*, *36*(1), 87–94. https://doi.org/10.1002/zoo.21342

IUCN/SSC. (2013). *Guidelines for reintroductions and other conservation translocations*. https://doi.org/10.1016/j.biocon.2015.07.030

Jackiw, R. N., Mandil, G., & Hager, H. A. (2015). A framework to guide the conservation of species hybrids based on ethical and ecological considerations. *Conservation Biology*, *29*(4), 1040–1051. https://doi.org/10.1111/cobi.12526

Jahner, J. P., Matocq, M. D., Malaney, J. L., Cox, M., Wolff, P., Gritts, M. A., & Parchman, T. L. (2019). The genetic legacy of 50 years of desert bighorn sheep translocations. *Evolutionary Applications*, *12*(2), 198–213. https://doi.org/10.1111/eva.12708

Jamieson, I. G. (2010). Founder effects, inbreeding, and loss of genetic diversity in four avian reintroduction programs. *Conservtion Biology*, *25*(1), 115–123. https://doi.org/10.1111/j.1523-1739.2010.01574.x

Jamieson, I. G., & Allendorf, F. W. (2012). How does the 50/500 rule apply to MVPs?

Trends in Ecology and Evolution, *27*(10), 578–584. https://doi.org/10.1016/j.tree.2012.07.001

Janjua, S., Peters, J. L., Weckworth, B., Abbas, F. I., Bahn, V., Johansson, O., & Rooney, T. P. (2019). Improving our conservation genetic toolkit: ddRAD-seq for SNPs in snow leopards. *Conservation Genetics Resources*. https://doi.org/10.1007/s12686-019-01082-2

Jiménez, J. A., Hughes, K. A., Alaks, G., Graham, L., & Lacy, R. C. (1994). An experimental study of inbreeding depression in a natural habitat. *Science*, *266*(5183), 271–273.

Johnson, D. H. (1980). The comparison of usage and availability measurements for evaluating resource preference. *Ecology*, *61*(1), 65–71.

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 94–109.

Jónsson, H., Schubert, M., Seguin-Orlando, A., Ginolhac, A., Petersen, L., Fumagalli, M., … Orlando, L. (2014). Speciation with gene flow in equids despite extensive chromosomal plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(52), 18655–18660. https://doi.org/10.1073/pnas.1412627111

Kaczensky, P., Hrabar, H., Lukarevskiy, V., Zimmermann, W., Usukhjargal, D., Ganbaatar, O., & Bouskila, A. (2016). Reintroductions of wild equids. In J. I. Ransom & P. Kaczensky (Eds.), *Wild equids: Ecology, management and conservation* (pp. 196–214). Baltimore, MD: Johns Hopkins University Press.

Kaczensky, P., Kovtun, E., Habibrakhmanov, R., Hemami, M.-R., Khaleghi, A., Linnell, J. D. C., … Kuehn, R. (2018). Genetic characterization of free-ranging Asiatic wild ass in Central Asia as a basis for future conservation strategies. *Conservation Genetics*, *19*(5), 1169-1184. https://doi.org/10.1007/s10592-018-1086-3

Kaczensky, P., Lkhagvasuren, P Pereladova, O., Hemami, M. R., & Bouskila, A. (2015). Equus hemionus. Retrieved September 16, 2019, from The IUCN Red List of Threatened Species 2015: e.T7951A45171204. website: https://www.iucnredlist.org/species/7951/45171204#bibliography

Kardos, M., Åkesson, M., Fountain, T., Flagstad, Ø., Liberg, O., Olason, P., … Ellegren, H. (2018). Genomic consequences of intensive inbreeding in an isolated Wolf population. *Nature Ecology and Evolution*, *2*(1), 124–131. https://doi.org/10.1038/s41559-017-0375-4

Kardos, M., Taylor, H. R., Ellegren, H., Luikart, G., & Allendorf, F. W. (2016). Genomics advances the study of inbreeding depression in the wild. *Evolutionary Applications*, *9*(10), 1205–1218. https://doi.org/10.1111/eva.12414

Keeley, A. T. H., Beier, P., Keeley, B. W., & Fagan, M. E. (2017). Habitat suitability is a

poor proxy for landscape connectivity during dispersal and mating movements. *Landscape and Urban Planning*, *161*, 90–102. https://doi.org/10.1016/j.landurbplan.2017.01.007

Keller, L. F., & Waller, D. M. (2002). Inbreeding effects in wild populations. *Trends in Ecology and Evolution*, *17*(5), 230–241. https://doi.org/10.1016/S0169-5347(02)02489-8

Kelly, E., & Phillips, B. (2019). How many and when? Optimising targeted gene flow for a step change in the environment. *Ecology Letters*, *22*(3), 447–457. https://doi.org/10.1111/ele.13201

Kelly, E., & Phillips, B. L. (2018). Targeted gene flow and rapid adaptation in an endangered marsupial. *Conservation Biology*, *33*(1), 112–121. https://doi.org/10.1111/cobi.13149

Killeen, J., Thurfjell, H., Ciuti, S., Paton, D., Musiani, M., & Boyce, M. S. (2014). Habitat selection during ungulate dispersal and exploratory movement at broad and fine scale with implications for conservation management. *Movement Ecology*, *2*(1), 1–13. https://doi.org/10.1186/s40462-014-0015-4

Kimura, B., Marshall, F. B., Chen, S., Rosenbom, S., Moehlman, P. D., Tuross, N., … Mulligan, C. J. (2011). Ancient DNA from Nubian and Somali wild ass provides insights into donkey ancestry and domestication. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1702), 50–57. https://doi.org/10.1098/rspb.2010.0708

Krausman, P. R., Rosenstock, S. S., & Cain, J. W. (2006). Developed waters for wildlife: Science, perception, values, and controversy. *Wildlife Society Bulletin*, *34*(3), 563–569.

Krüger, K., Gaillard, C., Stranzinger, G., & Rieder, S. (2005). Phylogenetic analysis and species allocation of individual equids using microsatellite data. *Journal of Animal Breeding and Genetics*, *122*, 78–86. https://doi.org/10.1111/j.1439-0388.2005.00505.x

la Morgia, V., Malenotti, E., Badino, G., & Bona, F. (2011). Where do we go from here? Dispersal simulations shed light on the role of landscape structure in determining animal redistribution after reintroduction. *Landscape Ecology*, *26*(7), 969–981. https://doi.org/10.1007/s10980-011-9621-3

Lande, R. (1988). Genetics and demography in biological conservation. *Science*, *241*(4872), 1455–1460.

Landguth, E. L., Cushman, S. A., Schwartz, M. K., McKelvey, K. S., Murphy, M., & Luikart, G. (2010). Quantifying the lag time to detect barriers in landscape genetics. *Molecular Ecology*, *19*(19), 4179–4191. https://doi.org/10.1111/j.1365-294X.2010.04808.x

Latch, E. K., & Rhodes, O. E. (2005). The effects of gene flow and population isolation on the genetic structure of reintroduced wild turkey populations: Are genetic signatures of source populations retained? *Conservation Genetics*, *6*(6), 981–997. https://doi.org/10.1007/s10592-005-9089-2

Legendre, P. P., Andersson, M. J., & Anderson, M. (1999). Distance-based redundancy

analysis: Testing multispecies responses in multifactorial ecological experiments. *Ecological Monographs*, *69*(1), 1–24. https://doi.org/10.1890/0012-9615(1999)069[0001:DBRATM]2.0.CO;2

Linck, E., & Battey, C. J. (2019). Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Molecular Ecology Resources*, *19*(3), 639–647. https://doi.org/10.1111/1755-0998.12995

López-Rull, I., Lifshitz, N., Macías Garcia, C., Graves, J. A., & Torres, R. (2016). Females of a polymorphic seabird dislike foreign-looking males. *Animal Behaviour*, *113*, 31–38. https://doi.org/10.1016/j.anbehav.2015.12.023

MacArthur, R. H. (1957). On the relative abundance of bird species. *Proceedings of the National Academy of Sciences of the United States of America*, *43*(3), 293–295.

Manel, S., Schwartz, M. K., Luikart, G., Taberlet, P., Manel, Stephanie, Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, *18*(4), 8–9. https://doi.org/10.1016/S0169-5347(03)00008-9

Manlick, P. J., Romanski, M. C., & Pauli, J. N. (2018). Dynamic colonization history in a rediscovered Isle Royale carnivore. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-31130-0

Marshall, T. C., & Spalton, J. A. (2000). Simultaneous inbreeding and outbreeding depression in reintroduced Arabian oryx. *Animal Conservation*, *3*(3), 241–248. https://doi.org/10.1017/S1367943000000962

Martin, C. H., Crawford, J. E., Turner, B. J., & Simons, L. H. (2016). Diabolical survival in Death Valley: recent pupfish colonization, gene flow and genetic assimilation in the smallest species range on earth. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1823), 20152334. https://doi.org/10.1098/rspb.2015.2334

Mastretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T. H., Piñero, D., & Emerson, B. C. (2015). Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Molecular Ecology Resources*, *15*(1), 28–41. https://doi.org/10.1111/1755-0998.12291

Mateo-Sánchez, M. C., Balkenhol, N., Cushman, S., Pérez, T., Domínguez, A., & Saura, S. (2015). A comparative framework to infer landscape effects on population genetic structure: are habitat suitability models effective in explaining gene flow? *Landscape Ecology*, *30*(8), 1405–1420. https://doi.org/10.1007/s10980-015-0194-4

Mattucci, F., Galaverni, M., Lyons, L. A., Alves, P. C., Randi, E., Velli, E., … Caniglia, R. (2019). Genomic approaches to identify hybrids and estimate admixture times in European wildcat populations. *Scientific Reports*, *9*(1), 1–15. https://doi.org/10.1038/s41598-019-48002-w

McCue, M. E., Bannasch, D. L., Petersen, J. L., Gurr, J., Bailey, E., Binns, M. M., … Mickelson, J. R. (2012). A high density SNP array for the domestic horse and extant perissodactyla: Utility for association mapping, genetic diversity, and phylogeny studies. *PLoS Genetics*, *8*(1), e1002451.

https://doi.org/10.1371/journal.pgen.1002451

McDonald, M. M., Johnson, S. M., Henry, E. R., & Cunneyworth, P. M. K. (2019). Differences between ecological niches in northern and southern populations of Angolan black and white colobus monkeys (Colobus angolensis palliatus and Colobus angolensis sharpei) throughout Kenya and Tanzania. *American Journal of Primatology*, e22975. https://doi.org/10.1002/ajp.22975

McLennan, E. A., Wright, B. R., Belov, K., Hogg, C. J., & Grueber, C. E. (2019). Too much of a good thing? Finding the most informative genetic data set to answer conservation questions. *Molecular Ecology Resources*, *19*(3), 659–671. https://doi.org/10.1111/1755-0998.12997

McRae, B., Shah, V., & Mohapatra, T. (2013). CIRCUITSCAPE User Guide. *The Nature Conservancy*, p. 28. Retrieved from http://www.circuitscape.org

McVean, G. (2009). A genealogical interpretation of principal components analysis. *PLoS Genetics*, *5*(10). https://doi.org/10.1371/journal.pgen.1000686

Meng, X. L., & Rubin, D. B. (1993). Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika*, *80*(2), 267–278. https://doi.org/10.1093/biomet/80.2.267

Miller, K. A., Nelson, N. J., Smith, H. G., & Moore, J. A. (2009). How do reproductive skew and founder group size affect genetic diversity in reintroduced populations? *Molecular Ecology*, *18*(18), 3792–3802. https://doi.org/10.1111/j.1365-294X.2009.04315.x

Mills, L. S., & Allendorf, F. W. (1996). The one-migrant-per-generation rule in conservation and management. *Conservation Biology*, *10*(6), 1509–1518. https://doi.org/10.1046/j.1523-1739.1996.10061509.x

Mock, K. E., Latch, E. K., & Rhodes, O. E. (2004). Assessing losses of genetic diversity due to translocation: Long-term case histories in Merriam's turkey (Meleagris gallopavo merriami). *Conservation Genetics*, *5*(5), 631–645. https://doi.org/10.1007/s10592-004-1849-x

Moehlman, P. D., Shah, N., Masseti, M., & Feh, C. (2010). Equus hemionus. Retrieved February 7, 2020, from The IUCN Red List of Threatened Species 2010: e.T7951A12874957 website: https://www.iucnredlist.org/species/7951/12874957

Mondol, S., Sridhar, V., Yadav, P., Gubbi, S., & Ramakrishnan, U. (2014). Tracing the geographic origin of traded leopard body parts in the indian subcontinent with DNA-based assignment tests. *Conservation Biology*, *29*(2), 556–564. https://doi.org/10.1111/cobi.12393

Moodley, Y., & Harley, E. H. (2005). Population structuring in mountain zebras (Equus zebra): The molecular consequences of divergent demographic histories. *Conservation Genetics*, *6*(6), 953–968. https://doi.org/10.1007/s10592-005-9083-8

Moore, D. L., & Vigilant, L. (2014). A population estimate of chimpanzees (Pan troglodytes schweinfurthii) in the Ugalla region using standard and spatially explicit genetic capture – recapture methods. *American Journal of Primatology*,

76(4), 335–346. https://doi.org/10.1002/ajp.22237

Moorhouse, T. P., Gelling, M., & Macdonald, D. W. (2009). Effects of habitat quality upon reintroduction success in water voles: Evidence from a replicated experiment. *Biological Conservation*, *142*(1), 53–60. https://doi.org/10.1016/j.biocon.2008.09.023

Moraes, A. M., Ruiz-Miranda, C. R., Galetti, P. M., Niebuhr, B. B., Alexandre, B. R., Muylaert, R. L., … Ribeiro, M. C. (2018). Landscape resistance influences effective dispersal of endangered golden lion tamarins within the Atlantic Forest. *Biological Conservation*, *224*, 178–187. https://doi.org/10.1016/j.biocon.2018.05.023

Moraes, A. M., Ruiz-Miranda, C. R., Ribeiro, M. C., Grativol, A. D., da S. Carvalho, C., Dietz, J. M., … Galetti, P. M. (2017). Temporal genetic dynamics of reintroduced and translocated populations of the endangered golden lion tamarin (Leontopithecus rosalia). *Conservation Genetics*, *18*(5), 995–1009. https://doi.org/10.1007/s10592-017-0948-4

Morin, P. A., Luikart, G., & Wayne, R. K. (2004). SNPs in ecology, evolution and conservation. *Trends in Ecology and Evolution*, *19*(4), 208–216. https://doi.org/10.1016/j.tree.2004.01.009

Munshi-South, J., Zolnik, C. P., & Harris, S. E. (2016). Population genomics of the Anthropocene: Urbanization is negatively associated with genome-wide variation in white-footed mouse populations. *Evolutionary Applications*, *9*(4), 546–564. https://doi.org/10.1111/eva.12357

Murphy, S. M., Laufenberg, J. S., Clark, J. D., Davidson, M., Belant, J. L., & Garshelis, D. L. (2018). Genetic diversity, effective population size, and structure among black bear populations in the Lower Mississippi Alluvial Valley, USA. *Conservation Genetics*, *19*(5), 1055–1067. https://doi.org/10.1007/s10592-018-1075-6

Nandintsetseg, D., Kaczensky, P., Ganbaatar, O., Leimgruber, P., & Mueller, T. (2016). Spatiotemporal habitat dynamics of ungulates in unpredictable environments: The khulan (Equus hemionus) in the Mongolian Gobi desert as a case study. *Biological Conservation*, *204*, 313–321. https://doi.org/10.1016/j.biocon.2016.10.021

Narum, S. R., Buerkle, C. A., Davey, J. W., Miller, M. R., & Hohenlohe, P. A. (2013). Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, *22*(11), 2841–2847. https://doi.org/10.1111/mec.12350

Nash, H. C., Wirdateti, Low, G. W., Choo, S. W., Chong, J. L., Semiadi, G., … Rheindt, F. E. (2018). Conservation genomics reveals possible illegal trade routes and admixture across pangolin lineages in Southeast Asia. *Conservation Genetics*, *19*(5), 1083–1095. https://doi.org/10.1007/s10592-018-1080-9

Nei, M., Maruyama, T., & Chakraborty, R. (1975). The bottleneck effect and genetic variability in populations. *Evolution*, *29*(1), 1–10.

Nei, M., & Tajima, F. (1981). Genetic drift and estimation of effective population size. *Genetics*, *98*(3), 625–640. https://doi.org/10.1534/genetics.107.075481

Neuwald, J. L., & Templeton, A. R. (2013). Genetic restoration in the eastern collared

lizard under prescribed woodland burning. *Molecular Ecology*, *22*(14), 3666–3679. https://doi.org/10.1111/mec.12306

Nezer, O., Bar-David, S., Gueta, T., & Carmel, Y. (2017). High-resolution species-distribution model based on systematic sampling and indirect observations. *Biodiversity and Conservation*, *26*(2), 421–437. https://doi.org/10.1007/s10531-016-1251-2

Ng, E. Y. X., Garg, K. M., Low, G. W., Chattopadhyay, B., Oh, R. R. Y., Lee, J. G. H., & Rheindt, F. E. (2017). Conservation genomics identifies impact of trade in a threatened songbird. *Biological Conservation*, *214*, 101–108. https://doi.org/10.1016/j.biocon.2017.08.007

Norwegian Institute for Nature Research. (2019). Retrieved July 7, 2019, from Kulan reintroduction to the Steppe of Central Kazakhstan website: https://www.nina.no/english/Fields-of-research/Projects/KulanSteppe

O'Brien, D., Manseau, M., Fall, A., & Fortin, M. J. (2006). Testing the importance of spatial configuration of winter habitat for woodland caribou: An application of graph theory. *Biological Conservation*, *130*(1), 70–83. https://doi.org/10.1016/j.biocon.2005.12.014

O'Leary, S. J., Puritz, J. B., Willis, S. C., Hollenbeck, C. M., & Portnoy, D. S. (2018). These aren't the loci you'e looking for: Principles of effective SNP filtering for molecular ecologists. *Molecular Ecology*, *27*(16), 3193–3206. https://doi.org/10.1111/mec.14792

Oakenfull, E. A., Lim, H. N., & Ryder, O. A. (2000). A survey of equid mitochondrial DNA: Implications for the evolution, genetic diversity and conservation of Equus. *Conservation Genetics*, *1*(4), 341–355. https://doi.org/10.1023/A:1011559200897

Ogden, R., & Linacre, A. (2015). Wildlife forensic science: A review of genetic geographic origin assignment. *Forensic Science International: Genetics*, *18*, 152–159. https://doi.org/10.1016/j.fsigen.2015.02.008

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'hara, R. B., Simpson, G. L., … Wagner, H. (2010). Vegan: community ecology package. R package version 1.17-4. *Http://Cran. r-Project. Org>. Acesso Em*, *23*, 2010.

Olson, Z. H., Whittaker, D. G., & Rhodes, O. E. (2013). Translocation history and genetic diversity in reintroduced bighorn sheep. *Journal of Wildlife Management*, *77*(8), 1553–1563. https://doi.org/10.1002/jwmg.624

Olsson, O. (2007). Genetic origin and success of reintroduced white storks. *Conservation Biology*, *21*(5), 1196–1206. https://doi.org/10.1111/j.1523-1739.2007.00763.x

Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., Stiller, M., … Willerslev, E. (2013). Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature*, *499*(7456), 74–78. https://doi.org/10.1038/nature12323

Paradis, E., & Schliep, K. (2018). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528.

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for stacks. *Methods in Ecology and Evolution*, *8*(10), 1360–1373. https://doi.org/10.1111/2041-210X.12775

Pelletier, F., Turgeon, G., Bourret, A., Garant, D., & St-Laurent, M. H. (2019). Genetic structure and effective size of an endangered population of woodland caribou. *Conservation Genetics*, *20*(2), 203–213. https://doi.org/10.1007/s10592-018-1124-1

Pertoldi, C., Wójcik, J. M., Tokarska, M., Kawałko, A., Kristensen, T. N., Loeschcke, V., … Bendixen, C. (2010). Genome variability in European and American bison detected using the BovineSNP50 BeadChip. *Conservation Genetics*, *11*(2), 627–634. https://doi.org/10.1007/s10592-009-9977-y

Peterman, W. E., Conette, G. M., Semlitsch, R. D., & Eggert, L. S. (2014). Ecological resistance surfaces predict fine-scale genetic differentiation in a terrestrial woodland salamander. *Molecular Ecology*, *23*(10), 2402–2413. https://doi.org/10.1111/mec.12747

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0037135

Pew, J., Muir, P. H., Wang, J., & Frasier, T. R. (2015). related: an R package for analysing pairwise relatedness from codominant molecular markers. *Molecular Ecology Resources*, *15*(3), 557–561.

Pilot, M., Greco, C., VonHoldt, B. M., Jędrzejewska, B., Randi, E., Jędrzejewski, W., … Wayne, R. K. (2014). Genome-wide signatures of population bottlenecks and diversifying selection in European wolves. *Heredity*, *112*(4), 428–442. https://doi.org/10.1038/hdy.2013.122

Pollak, E. (1983). A new method for estimating the effective population size from allele frequency changes. *Genetics*, *104*(3), 531–548.

Porras-Hurtado, L., Ruiz, Y., Santos, C., Phillips, C., Carracedo, A., & Lareu, M. V. (2013). An overview of STRUCTURE: Applications, parameter settings, and supporting software. *Frontiers in Geneticsenetics*, *4*, 98. https://doi.org/10.3389/fgene.2013.00098

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Puckett, E. E., Kristensen, T. V., Wilton, C. M., Lyda, S. B., Noyce, K. V., Holahan, P. M., … Eggert, L. S. (2014). Influence of drift and admixture on population structure of American black bears (Ursus americanus) in the Central Interior Highlands, USA, 50 years after translocation. *Molecular Ecology*, *23*(10), 2414–2427. https://doi.org/10.1111/mec.12748

Puritz, J. B., Hollenbeck, C. M., & Gold, J. R. (2014). dDocent : a RADseq, variant-calling pipeline designed for population genomics of non-model organisms . *PeerJ*, *2*, e431. https://doi.org/10.7717/peerj.431

Radinger, J., Essl, F., Hölker, F., Horký, P., Slavík, O., & Wolter, C. (2017). The future distribution of river fish: The complex interplay of climate and land use changes, species dispersal and movement barriers. *Global Change Biology*, *23*(11), 4970–4986. https://doi.org/10.1111/gcb.13760

Ralls, K., Ballou, J. D., Dudash, M. R., Eldridge, M. D. B. B., Fenster, C. B., Lacy, R. C., … Frankham, R. (2018). Call for a Paradigm Shift in the Genetic Management of Fragmented Populations. *Conservation Letters*, *11*(2), 1–6. https://doi.org/10.1111/conl.12412

Ransom, J. I., Lagos, L., Hrabar, H., Nowzari, H., Usukhjargal, D., & Spasskaya, N. (2016). Wild and feral equid population dynamics. In J. I. Ransom & P. Kaczensky (Eds.), *Wild equids: Ecology, management, and conservation* (pp. 87–104). Baltimore: Johns Hopkins University Press.

Reed, D. H., & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, *17*(1), 230–237. https://doi.org/10.1046/j.1523-1739.2003.01236.x

Renan, S. (2014). *From behavioral patterns to genetic structure: The reintroduced Asiatic Wild Ass (Equus hemionus) in the Negev Desert*. Ben-Gurion University of the Negev, Israel.

Renan, S., Greenbaum, G., Shahar, N., Templeton, A. R., Bouskila, A., & Bar-David, S. (2015). Stochastic modelling of shifts in allele frequencies reveals a strongly polygynous mating system in the re-introduced Asiatic wild ass. *Molecular Ecology*, *24*(7), 1433–1446. https://doi.org/10.1111/mec.13131

Renan, S., Speyer, E., Ben-Nun, T., Ziv, A., Greenbaum, G., Templeton, A. R., … Bouskila, A. (2018). Fission-fusion social structure of a reintroduced ungulate: Implications for conservation. *Biological Conservation*, *222*, 261–267. https://doi.org/10.1016/j.biocon.2018.04.013

Rich, L. N., Beissinger, S. R., Brashares, J. S., & Furnas, B. J. (2019). Artificial water catchments influence wildlife distribution in the Mojave Desert. *Journal of Wildlife Management*, *83*(4), 855–865. https://doi.org/10.1002/jwmg.21654

Rieseberg, L. H., Whitton, J., & Gardner, K. (1999). Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics*, *152*(2), 713–727.

Riley, S. P. D., Pollinger, J. P., Sauvajot, R. M., York, E. C., Bromley, C., Fuller, T. K., & Wayne, R. K. (2006). A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology*, *15*(7), 1733–1741. https://doi.org/10.1111/j.1365-294X.2006.02907.x

Rochette, N. C., & Catchen, J. M. (2017). Deriving genotypes from RAD-seq short-read data using Stacks. *Nature Protocols*, *12*(12), 2640–2659. https://doi.org/10.1038/nprot.2017.123

Rodríguez-Ezpeleta, N., Bradbury, I. R., Mendibil, I., Álvarez, P., Cotano, U., & Irigoien, X. (2016). Population structure of Atlantic mackerel inferred from RAD-seq-derived SNP markers: effects of sequence clustering parameters and hierarchical SNP selection. *Molecular Ecology Resources*, *16*(4), 991–1001.

https://doi.org/10.1111/1755-0998.12518

Rollins, L. A., Woolnough, A. P., Wilton, A. N., Sinclair, R., & Sherwin, W. B. (2009). Invasive species can't cover their tracks : using microsatellites to assist management of starling (Sturnus vulgaris) populations in Western Australia. *Molecular Ecology*, *18*(8), 1560–1573. https://doi.org/10.1111/j.1365-294X.2009.04132.x

Rooney, S. M., Wolfe, A., & Hayden, T. J. (1998). Autocorrelated data in telemetry studies: time to independence and the problem of behavioural effects. *Mammal Review*, *28*(2), 89–98. https://doi.org/10.1046/j.1365-2907.1998.00028.x

Rosenbom, S., Costa, V., Chen, S., Khalatbari, L., Yusefi, G. H., Abdukadir, A., … Beja-Pereira, A. (2015). Reassessing the evolutionary history of ass-like equids: Insights from patterns of genetic variation in contemporary extant populations. *Molecular Phylogenetics and Evolution*, *85*, 88–96. https://doi.org/10.1016/j.ympev.2015.01.005

Rubinoff, D., & Holland, B. S. (2005). Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biologyiology*, *54*(6), 952–961. https://doi.org/10.1080/10635150500234674

Ryder, O. A., & Chemnick, L. G. (1990). Chromosomal and molecular evolution in Asiatic wild asses. *Genetica*, *83*(1), 67–72. https://doi.org/10.1007/BF00774690

Saccheri, I., Kuussaari, M., Kankare, M., Vikman, P., Fortelius, W., & Hanski, I. (1998). Inbreeding and extinction in a butterfly metapopulation. *Nature*, *392*(1), 491–494. https://doi.org/10.1038/33136

Sagvik, J., Uller, T., & Olsson, M. (2005). Outbreeding depression in the common frog, Rana temporaria. *Conservation Genetics*, *6*(2), 205–211. https://doi.org/10.1007/s10592-004-7829-3

Saltz, D., Rowen, M., & Rubenstein, D. I. (2000). The effect of space-use patterns of reintroduced Asiatic wild ass on effective population size. *Conservation Biology*, *14*(6), 1852–1861. https://doi.org/10.1111/j.1523-1739.2000.99227.x

Saltz, D., & Rubenstein, D. I. (1995). Population dynamics of a reintroduced Asiatic wild ass (Equus hemionus) herd. *Ecological Applications*, *5*(2), 327–335.

Sard, N. M., Johnson, M. A., Jacobson, D. P., Hogansen, M. J., O'Malley, K. G., & Banks, M. A. (2016). Genetic monitoring guides adaptive management of a migratory fish reintroduction program. *Animal Conservation*, *19*(6), 570–577. https://doi.org/10.1111/acv.12278

Saremi, N. F., Supple, M. A., Byrne, A., Cahill, J. A., Coutinho, L. L., Dalén, L., … Shapiro, B. (2019). Puma genomes from North and South America provide insights into the genomic consequences of inbreeding. *Nature Communications*, *10*(1), 1–10. https://doi.org/10.1038/s41467-019-12741-1

Sasidharan, S. P., Ludwig, A., Harper, C., Moodley, Y., Bertschinger, H. J., & Guthrie, A. J. (2011). Comparative genetics of sarcoid tumour-affected and non-affected mountain zebra (Equus zebra) populations. *African Journal of Wildlife Research*,

41(1), 36–49. https://doi.org/10.3957/056.041.0117

Saveljev, A., & Milishnikov, A. (2002). Biological and genetic peculiarities of cross-composed and aboriginal beaver populations in Russia. *Acta Zoologica Lituanica*, *12*(4), 397–402. https://doi.org/10.1080/13921657.2002.10512530

Schneider, J. (2011). Review of reintroduction of Atlantic salmon (Salmo salar) in tributaries of the Rhine River in the German Federal States of Rhineland-Palatinate and Hesse. *Journal of Applied Ichthyology*, *27*(3), 24–32. https://doi.org/10.1111/j.1439-0426.2011.01848.x

Schoenecker, K. A., King, S. R. B., Nordquist, M. K., Nandintsetseg, D., & Cao, Q. (2016). Habitat and Diet of Equids. In J. I. Ransom & P. Kaczensky (Eds.), *Wild Equids: Ecology, Management and Conservation* (pp. 41–57). Baltimore, MD: Johns Hopkins University Press.

Schwartz, M. K., Luikart, G., & Waples, R. S. (2007). Genetic monitoring as a promising tool for conservation and management. *Trends in Ecology and Evolution*, *22*(1), 25–33. https://doi.org/10.1016/j.tree.2006.08.009

Schwartz, M. K., McKelvey, K. S., Cushman, S. A., & Luikart, G. (2010). Landscape genomics: A brief perspective. In S.A. Cushman & F. Huettmann (Eds.), *Spatial complexity, informatics, and wildlife conservation* (pp. 165–174). Tokyo: Springer.

Seddon, P. J., & Armstrong, D. P. (2016). Reintroduction and other conservation translocations. In David S Jachowski, J. J. Millspaugh, P. L. Angermeier, & R. Slotow (Eds.), *Reintroduction of fish and wildlife populations* (1st ed., pp. 7–28). Retrieved from http://www.jstor.org/stable/10.1525/j.ctv1xxz9w.6

Seddon, P. J., Armstrong, D. P., & Maloney, R. F. (2007). Developing the science of reintroduction biology. *Conservation Biology*, *21*(2), 303–312. https://doi.org/10.1111/j.1523-1739.2006.00627.x

Seddon, P. J., & Soorae, P. S. (1999). Guidelines for subspecific substitutions in wildlife restoration projects. *Conservation Biology*, *13*(1), 177–184.

Segelbacher, G., Höglund, J., & Storch, I. (2003). From connectivity to isolation: Genetic consequences of population fragmentation in capercaillie across Europe. *Molecular Ecology*, *12*(7), 1773–1780. https://doi.org/10.1046/j.1365-294X.2003.01873.x

Senn, H., Ogden, R., Cezard, T., Gharbi, K., Iqbal, Z., Johnson, E., … McEwing, R. (2013). Reference-free SNP discovery for the Eurasian beaver from restriction site-associated DNA paired-end data. *Molecular Ecology*, *22*(11), 3141–3150. https://doi.org/10.1111/mec.12242

Shafer, A. B. A., Northrup, J. M., White, K. S., Boyce, M. S., Côté, S. D., & Coltman, D. W. (2012). Habitat selection predicts genetic relatedness in an alpine ungulate. *Ecology*, *93*(6), 1317–1329. https://doi.org/10.1890/11-0815.1

Shafer, A. B. A., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B. W. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, *8*(8), 907–917. https://doi.org/10.1111/2041-210X.12700

Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergstro, L., Meester, L. De, Bruford, M. W., … Zielin, P. (2015). Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution*, *30*(2), 78–87. https://doi.org/10.1016/j.tree.2014.11.009

Shafer, A. B. A., Wolf, J. B. W., Alves, P. C., Bergström, L., Colling, G., Dalén, L., … Zieliński, P. (2016). Reply to Garner et al. *Trends in Ecology & Evolution*, *31*(2), 83. https://doi.org/10.1016/j.tree.2015.11.010

Sharma, B. D., Clevers, J., De Graaf, R., & Nawa, R. (2004). Mapping Equus kiang (Tibetan Wild Ass) habitat in Surkhang, Upper Mustang, Nepal. *International Mountain Society*, *24*(2), 149–156.

Shemesh, H., Shani, G., Carmel, Y., Kent, R., & Sapir, Y. (2018). To mix or not to mix the sources of relocated plants? The case of the endangered Iris lortetii. *Journal for Nature Conservation*, *45*, 41–47. https://doi.org/10.1016/j.jnc.2018.08.002

Short, K. H., & Petren, K. (2011). Fine-scale genetic structure arises during range expansion of an invasive gecko. *PLoS ONE*, *6*(10). https://doi.org/10.1371/journal.pone.0026258

Shryock, D. F., Havrilla, C. A., DeFalco, L. A., Esque, T. C., Custer, N. A., & Wood, T. E. (2017). Landscape genetic approaches to guide native plant restoration in the Mojave Desert: *Ecological Applications*, *27*(2), 429–445. https://doi.org/10.1002/eap.1447

Sinai, Y. (1994). *A program for selecting individuals from a breeding core for reintroduction - Equus hemionus in Hai-Bar Yotveta*. The Hebrew University, Jerusalem.

Soland-Reckeweg, G., Heckel, G., Neumann, P., Fluri, P., & Excoffier, L. (2009). Gene flow in admixed populations and implications for the conservation of the Western honeybee, Apis mellifera. *Journal of Insect Conservation*, *13*(3), 317–328. https://doi.org/10.1007/s10841-008-9175-0

Solberg, K. H., Bellemain, E., Drageset, O.-M., Taberlet, P., & Swenson, J. E. (2006). An evaluation of field and non-invasive genetic methods to estimate brown bear (Ursus arctos) population size. *Biological Conservation*, *128*(2), 158–168. https://doi.org/10.1016/j.biocon.2005.09.025

Soulé, M. E. (1985). What is conservation biology? *BioScience*, *35*(11), 727–734.

Sovic, M., Fries, A., Martin, S. A., & Lisle Gibbs, H. (2019). Genetic signatures of small effective population sizes and demographic declines in an endangered rattlesnake, Sistrurus catenatus. *Evolutionary Applications*, *12*(4), 664–678. https://doi.org/10.1111/eva.12731

Sovic, M. G., Carstens, B. C., & Gibbs, H. L. (2016). Genetic diversity in migratory bats: Results from RADseq data for three tree bat species at an Ohio windfarm. *PeerJ*, *4*, e1647. https://doi.org/10.7717/peerj.1647

Sovic, M. G., Fries, A. C., & Gibbs, H. L. (2015). AftrRAD: A pipeline for accurate and efficient de novo assembly of RADseq data. *Molecular Ecology Resources*, *15*(5), 1163–1171. https://doi.org/10.1111/1755-0998.12378

Spear, S. F., Balkenhol, N., Fortin, M. M. J., McRae, B. H., & Scribner, K. (2010). Use of resistance surfaces for landscape genetic studies: Considerations for parameterization and analysis. *Molecular Ecology*, *19*(17), 3576–3591. https://doi.org/10.1111/j.1365-294X.2010.04657.x

Spear, S. F., Cushman, S. A., & McRae, B. H. (2016). Resistance surface modeling in landscape genetics. In N. Balkenhol, S. A. Cushman, A. T. Storfer, & L. P. Waits (Eds.), *Landscape genetics: Concepts, methods, applications* (pp. 129–144). Chichester: John Wiley & Sons, Ltd.

Spielman, D., Brook, B. W., & Frankham, R. (2004). Most species are not driven to extinction before genetic factors impact them. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(42), 15261–15264. https://doi.org/10.1073/pnas.0403809101

Steiner, C. C., Mitelberg, A., Tursi, R., & Ryder, O. A. (2012). Molecular phylogeny of extant equids and effects of ancestral polymorphism in resolving species-level phylogenies. *Molecular Phylogenetics and Evolution*, *65*(2), 573–581. https://doi.org/10.1016/J.YMPEV.2012.07.010

Stern, E., Gradus, Y., Meir, A., Krakover, S., & Tsoar, H. (1986). *Atlas of the Negev*. Jerusalem: Keter Publishing House.

Stoffel, M. A., Humble, E., Paijmans, A. J., Acevedo-Whitehouse, K., Chilvers, B. L., Dickerson, B., … Hoffman, J. I. (2018). Demographic histories and genetic diversity across pinnipeds are shaped by human exploitation, ecology and life-history. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06695-z

Storfer, A., Murphy, M. A., Evans, J. S., Goldberg, C. S., Robinson, S., Spear, S. F., … Waits, L. P. (2007). Putting the "landscape" in landscape genetics. *Heredity*, *98*(3), 128–142. https://doi.org/10.1038/sj.hdy.6800917

Stowell, S. M. L., Bentley, E. G., Gagne, R. B., Gustafson, K. D., Rutledge, L. Y., & Ernest, H. B. (2018). Optimal DNA extractions from blood on preservation paper limits conservation genomic but not conservation genetic applications. *Journal for Nature Conservation*, *46*, 89–96. https://doi.org/10.1016/j.jnc.2018.09.004

Svengren, H., Prettejohn, M., Bunge, D., Fundi, P., & Björklund, M. (2017). Relatedness and genetic variation in wild and captive populations of mountain bongo in Kenya obtained from genome-wide single-nucleotide polymorphism (SNP) data. *Global Ecology and Conservation*, *11*, 196–206. https://doi.org/10.1016/j.gecco.2017.07.001

Tallmon, D. A., Luikart, G., & Waples, R. S. (2004). The alluring simplicity and complex reality of genetic rescue. *Trends in Ecology and Evolution*, *19*(9), 489–496. https://doi.org/10.1016/j.tree.2004.07.003

Taylor, S. S., Jamieson, I. G., & Armstrong, D. P. (2005). Successful island reintroductions of New Zealand robins and saddlebacks with small numbers of founders. *Animal Conservation*, *8*(4), 415–420. https://doi.org/10.1017/S1367943005002337

Templeton, A. R. (2006). Genetic drift in large populations and coalescence. In *Population Genetics and Microevolutionary Theory* (pp. 118–143). Hoboken, NJ:

John Wiley & Sons, Inc.

Templeton, A. R. (2017). Measuring biodiversity and monitoring ecological and evolutionary processes with genetic and genomic tools. In J. Garson, A. Plutynski, & S. Sarkar (Eds.), *The Routledge handbook of philosophy of biodiversity* (pp. 251–265). New York: Routledge.

Templeton, A. R. (2018). *Human population genetics and genomics*. London: Academic Press/Elsevier.

Templeton, A. R., Brazeal, H., & Neuwald, J. L. (2011). The transition from isolated patches to a metapopulation in the eastern collared lizard in response to prescribed fires. *Ecology*, *92*(9), 1736–1747. https://doi.org/10.1890/10-1994.1

Templeton, A. R., Hemmer, H., Mace, G., Seal, U. S., Shields, W. M., & Woodruff, D. S. (1986). Local adaptation, coadaptation, and population boundaries. *Zoo Biology*, *5*(2), 115–125. https://doi.org/10.1002/zoo.1430050206

Thavornkanlapachai, R., Mills, H. R., Ottewell, K., Dunlop, J., Sims, C., Morris, K., … Kennington, W. J. (2019). Mixing genetically and morphologically distinct populations in translocations: Asymmetrical introgression in a newly established population of the boodie (Bettongia lesueur). *Genes*, *10*(9), 729. https://doi.org/10.3390/genes10090729

Thomé, M. T. C., & Carstens, B. C. (2016). Phylogeographic model selection leads to insight into the evolutionary history of four-eyed frogs. *Proceedings of the National Academy of Sciences*, *113*(29), 8010–8017. https://doi.org/10.1073/pnas.1601064113

Thresher, R. E., Hayes, K., Bax, N. J., Teem, J., Benfey, T. J., & Gould, F. (2013). Genetic control of invasive fish: technological options and its role in integrated pest management. *Biological Invasions*, *16*(6), 1201–1216. https://doi.org/10.1007/s10530-013-0477-0

Tollington, S., Jones, C. G., Greenwood, A., Tatayah, V., Raisin, C., Burke, T., … Groombridge, J. J. (2013). Long-term, fine-scale temporal patterns of genetic diversity in the restored Mauritius parakeet reveal genetic impacts of management and associated demographic effects on reintroduction programmes. *Biological Conservation*, *161*, 28–38. https://doi.org/10.1016/j.biocon.2013.02.013

Toonen, R. J., Puritz, J. B., Forsman, Z. H., Whitney, J. L., Fernandez-Silva, I., Andrews, K. R., & Bird, C. E. (2013). ezRAD: A simplified method for genomic genotyping in non-model organisms. *PeerJ*, *1*, 1–15. https://doi.org/10.7717/peerj.203

Tordoff, H. B., & Redig, P. T. (2001). Role of genetic background in the success of reintroduced peregrine falcons. *Conservation Biology*, *15*(2), 528–532. https://doi.org/10.1046/j.1523-1739.2001.015002528.x

Tracy, L. N., Wallis, G. P., Efford, M. G., & Jamieson, I. G. (2011). Preserving genetic diversity in threatened species reintroductions: How many individuals should be released? *Animal Conservation*, *14*(4), 439–446. https://doi.org/10.1111/j.1469-1795.2011.00448.x

Trifonov, V. A., Stanyon, R., Nesterenko, A. I., Fu, B., Perelman, P. L., O'Brien, P. C. M., … Yang, F. (2008). Multidirectional cross-species painting illuminates the history of karyotypic evolution in Perissodactyla. *Chromosome Research*, *16*(1), 89–107. https://doi.org/10.1007/s10577-007-1201-7

Väli, Ü., Saag, P., Dombrovski, V., Meyburg, B. U., Maciorowski, G., Mizera, T., … Fagerberg, S. (2010). Microsatellites and single nucleotide polymorphisms in avian hybrid identification: A comparative case study. *Journal of Avian Biology*, *41*(1), 34–49. https://doi.org/10.1111/j.1600-048X.2009.04730.x

Vamosi, S. M., & Schluter, D. (1999). Sexual selection against hybrids between sympatric stickleback species: Evidence from a field experiment. *Evolution*, *53*(3), 874. https://doi.org/10.2307/2640727

Van Dijk, E. L., Jaszczyszyn, Y., & Thermes, C. (2014). Library preparation methods for next-generation sequencing: Tone down the bias. *Experimental Cell Research*, *322*(1), 12–20. https://doi.org/10.1016/j.yexcr.2014.01.008

Van Dyck, H., & Baguette, M. (2005). Dispersal behaviour in fragmented landscapes: Routine or special movements? *Basic and Applied Ecology*, *6*(6), 535–545. https://doi.org/10.1016/j.baae.2005.03.005

van Oppen, M. J. H., Oliver, J. K., Putnam, H. M., & Gates, R. D. (2015). Building coral reef resilience through assisted evolution. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(8), 2307–2313. https://doi.org/10.1073/pnas.1422301112

Vilstrup, J. T., Seguin-Orlando, A., Stiller, M., Ginolhac, A., Raghavan, M., Nielsen, S. C. A., … Orlando, L. (2013). Mitochondrial phylogenomics of modern and ancient equids. *PLoS ONE*, *8*(2), e55950. https://doi.org/10.1371/journal.pone.0055950

Volf, J. (2010). Sixty years of kulan, Equus hemionus kulan (Groves et Mazák, 1967) breeding at Prague Zoo. *Equus. Prague Zoo, Prague, Czech Republic*, 31–55.

vonHoldt, B. M., Brzeski, K. E., Wilcove, D. S., & Rutledge, L. Y. (2018). Redefining the role of admixture and genomics in species conservation. *Conservation Letters*, *11*(2), 1–6. https://doi.org/10.1111/conl.12371

VonHoldt, B. M., Cahill, J. A., Fan, Z., Gronau, I., Robinson, J., Pollinger, J. P., … Wayne, R. K. (2016). Whole-genome sequence analysis shows that two endemic species of North American wolf are admixtures of the coyote and gray wolf. *Science Advances*, *2*(7), e1501714. https://doi.org/10.1126/sciadv.1501714

vonHoldt, B. M., Stahler, D. R., Smith, D. W., Earl, D. A., Pollinger, J. P., & Wayne, R. K. (2007). The genealogy and genetic viability of reintroduced Yellowstone grey wolves. *Molecular Ecology*, *17*(1), 252–274. https://doi.org/10.1111/j.1365-294X.2007.03468.x

Waits, L. P., Cushman, S. A., & Spear, S. F. (2016). Applications of landscape genetics to connectivity research in terrestrial animals. In N. Balkenhol, S. A. Cushman, A. T. Storfer, & L. P. Waits (Eds.), *Landscape genetics: Concepts, methods, applications* (pp. 199–219). Wiley-Blackwell.

Wallach, A. D., Inbar, M., Scantlebury, M., Speakman, J. R., & Shanas, U. (2007). Water

requirements as a bottleneck in the reintroduction of European roe deer to the southern edge of its range. *Canadian Journal of Zoology*, *85*(11), 1182–1192. https://doi.org/10.1139/Z07-098

Wang, J. (2002). An estimator for pairwise relatedness using molecular markers. *Genetics*, *160*(3), 1203–1215. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10924488

Wang, S., Meyer, E., Mckay, J. K., & Matz, M. V. (2012). 2b-RAD: A simple and flexible method for genome-wide genotyping. *Nature Methods*, *9*(8), 808–810. https://doi.org/10.1038/nmeth.2023

Wang, W., Qiao, Y., Li, S., Pan, W., & Yao, M. (2017). Low genetic diversity and strong population structure shaped by anthropogenic habitat fragmentation in a critically endangered primate, Trachypithecus leucocephalus. *Heredity*, *118*(6), 542–553. https://doi.org/10.1038/hdy.2017.2

Waples, R. S. (1989). A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics*, *121*(2), 379–391.

Waples, R. S., & Do, C. (2008). LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Molecular Ecology Resources*, *8*(4), 753–756. https://doi.org/10.1111/j.1755-0998.2007.02061.x

Wasserman, T. N., Cushman, S. A., Schwartz, M. K., & Wallin, D. O. (2010). Spatial scaling and multi-model inference in landscape genetics: Martes americana in northern Idaho. *Landscape Ecology*, *25*(10), 1601–1612. https://doi.org/10.1007/s10980-010-9525-7

Weeks, A. R., Moro, D., Thavornkanlapachai, R., Taylor, H. R., White, N. E., Weiser, E. L., & Heinze, D. (2015). Conserving and enhancing genetic diversity in translocation programs. *Advances in Reintroduction Biology of Australian and New Zealand Fauna*, 127–140.

Weeks, A. R., Stoklosa, J., & Hoffmann, A. A. (2016). Conservation of genetic uniqueness of populations may increase extinction likelihood of endangered species: The case of Australian mammals. *Frontiers in Zoology*, *13*(1), 1–9. https://doi.org/10.1186/s12983-016-0163-z

Wessels, S., Krause, I., Floren, C., Schütz, E., Beck, J., & Knorr, C. (2017). ddRADseq reveals determinants for temperature-dependent sex reversal in Nile tilapia on LG23. *BMC Genomics*, *18*(1), 1–13. https://doi.org/10.1186/s12864-017-3930-0

White, L. C., Moseby, K. E., Thomson, V. A., Donnellan, S. C., & Austin, J. J. (2018). Long-term genetic consequences of mammal reintroductions into an Australian conservation reserve. *Biological Conservation*, *219*, 1–11. https://doi.org/10.1016/J.BIOCON.2017.12.038

Williams, C. L., Serfass, T. L., Cogan, R., & Rhodes, O. E. (2002). Microsatellite variation in the reintroduced Pennsylvania elk herd. *Molecular Ecology*, *11*(8), 1299–1310. https://doi.org/10.1046/j.1365-294X.2002.01546.x

Williams, R. N., Rhodes, O. E., & Serfass, T. L. (2000). Assessment of genetic variance among source and reintroduced fisher populations. *Journal of Mammalogy*, *81*(3),

895–907. https://doi.org/10.1644/1545-1542(2000)081<0895:aogvas>2.3.co;2

Wilson, G. A., Nishi, J. S., Elkin, B. T., & Strobeck, C. (2005). Effects of a recent founding event and intrinsic population dynamics on genetic diversity in an ungulate population. *Conservation Genetics*, *6*(6), 905–916. https://doi.org/10.1007/s10592-005-9077-6

Wisely, S. M., Mcdonald, D. B., & Buskirk, S. W. (2003). Evaluation of the Genetic Management of the Endangered Black-Footed Ferret (Mustela nigripes). *Zoo Biology*, *22*, 287–298. https://doi.org/10.1002/zoo.10089

Wisely, S. M., Santymire, R. M., Livieri, T. M., Mueting, S. A., & Howard, J. (2008). Genotypic and phenotypic consequences of reintroduction history in the black-footed ferret (Mustela nigripes). *Conservation Genetics*, *9*(2), 389–399. https://doi.org/10.1007/s10592-007-9351-x

With, K. A., & King, A. W. (1999). Extinction thresholds for species in fractal landscapes. *Conservation Biology*, *13*(2), 314–326. https://doi.org/10.1046/j.1523-1739.1999.013002314.x

Witzenberger, K. A., & Hochkirch, A. (2011). Ex situ conservation genetics: A review of molecular studies on the genetic consequences of captive breeding programmes for endangered animal species. *Biodiversity and Conservation*, *20*(9), 1843–1861. https://doi.org/10.1007/s10531-011-0074-4

Wright, D. J., Spurgin, L. G., Collar, N. J., Komdeur, J., Burke, T., & Richardson, D. S. (2014). The impact of translocations on neutral and functional genetic diversity within and among populations of the Seychelles warbler. *Molecular Ecology*, *23*(9), 2165–2177. https://doi.org/10.1111/mec.12740

Wu, T. D., & Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*, *26*(7), 873–881. https://doi.org/10.1093/bioinformatics/btq057

Yoffe, A. (1980). Breeding endangered species in Israel. *International Zoo Yearbook*, *20*(1), 127–137. https://doi.org/10.1111/j.1748-1090.1980.tb00956.x

Yuan, X., Miller, D. J., Zhang, J., Herrington, D., & Wang, Y. (2012). An overview of population genetic data simulation. *Journal of Computational Biology*, *19*(1), 42–54. https://doi.org/10.1089/cmb.2010.0188

Yumnam, B., Jhala, Y. V., Qureshi, Q., Maldonado, J. E., Gopal, R., Saini, S., … Fleischer, R. C. (2014). Prioritizing tiger conservation through landscape genetics and habitat linkages. *PLoS ONE*, *9*(11). https://doi.org/10.1371/journal.pone.0111207

Zachos, F. E., Frantz, A. C., Kuehn, R., Bertouille, S., Colyn, M., Niedziałkowska, M., … Flamand, M. C. (2016). Genetic structure and effective population sizes in European red deer (Cervus elaphus) at a continental scale: Insights from microsatellite DNA. *Journal of Heredity*, *107*(4), 318–326. https://doi.org/10.1093/jhered/esw011

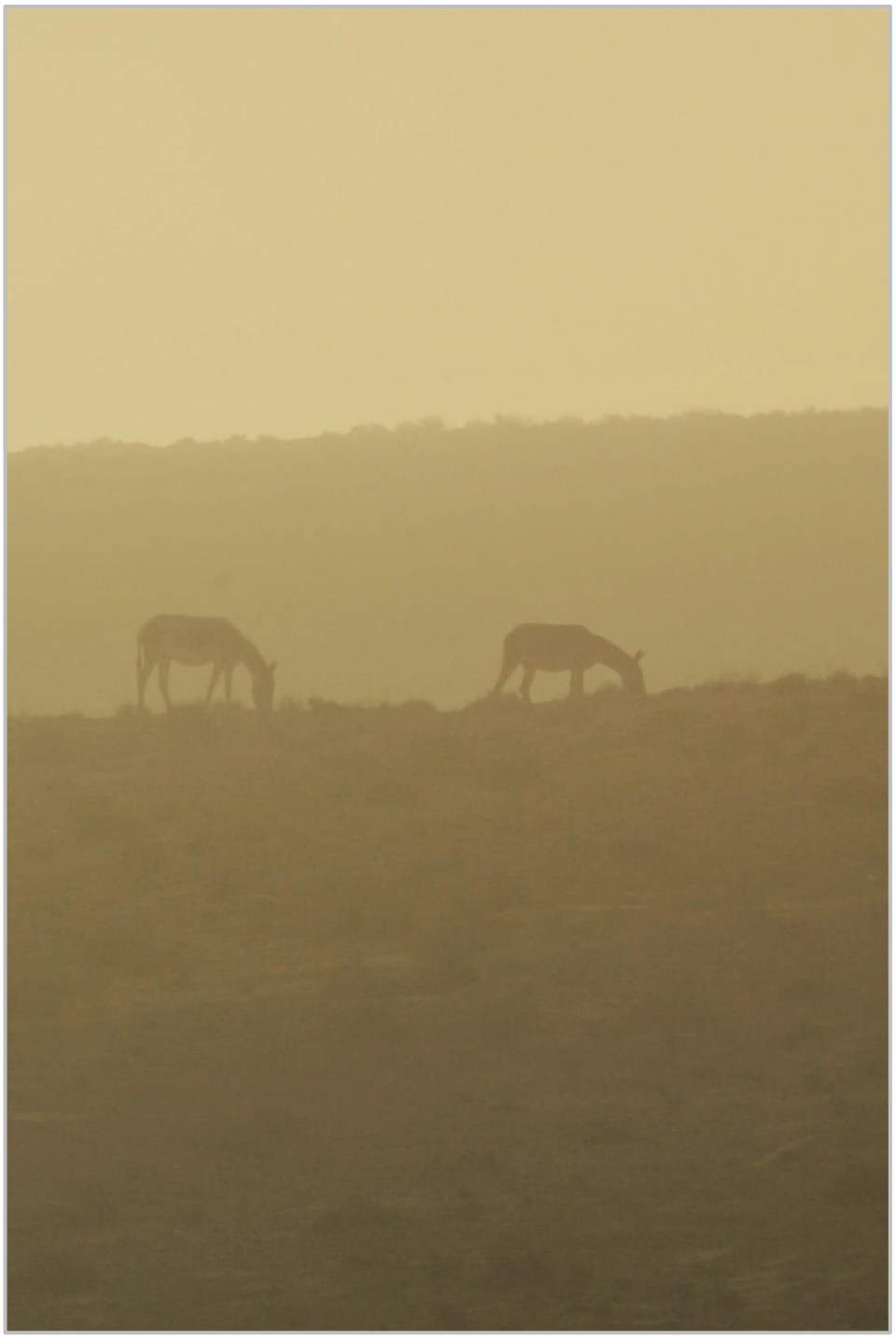
Zhang, Y., Zhong, Y., Hong, Y., Xue, Y., Li, D., Zhou, C., & Liu, S. (2019). RAD-Seq data advance captive-based conservation of wild bactrian camels (Camelus ferus). *Conservation Genetics*, *20*(4), 817–824. https://doi.org/10.1007/s10592-019-

01173-5

Ziółkowska, E., Perzanowski, K., Bleyhl, B., Ostapowicz, K., & Kuemmerle, T. (2016). Understanding unexpected reintroduction outcomes: Why aren't European bison colonizing suitable habitat in the Carpathians? *Biological Conservation*, *195*, 106–117. https://doi.org/10.1016/j.biocon.2015.12.032

Ziv, A. (2016). *Males social behavior and space use patterns in the Asiatic wild ass (Equus hemionus)*. Ben-Gurion University of the Negev.

7. Appendices



Appendix A: Supporting Information for chapter 3

Estimates of hybrid indices after removal of Hardy-Weinberg equilibrium outlier loci

The complete data set contained 5581 SNPs. Hardy-Weinberg equilibrium outlier loci were removed using the –hwe function in the program *vcftools*, which performs an exact test. SNPs with a p-values <0.05 were removed, resulting in a reduced data set containing 4231 SNPs. The removal of HWE outlier loci affected the results of the hybridisation analysis. After removal of HWE outliers the mean admixture levels in the founder and the wild population increased. The mean hybrid index in the wild population was significantly lower than in the founder population (independent samples Student's t-test: $t(37.84)=-3.543$, p=0.001, Table A1). However, these differences have little impact on the biological interpretation of the results. Both data sets display high levels of subspecies admixture in the founder and the wild population, with a shift towards increased onager ancestry in the wild population. This is also consistent with the pairwise fixation indices, which indicate that the combined hybrid population is more strongly differentiated from the kulan than the onager population (Table A2).

Table A1 Comparison of the mean (±SD) hybrid indices of the founder and wild population for the complete data set and after removal of Hardy-Weinberg equilibrium outlier loci. Hybrid indices were calculated calculated using all SNPs and maximum likelihood methods implemented in the R package *introgress* (Gompert & Buerkle, 2010). Statistically significant difference (*p<0.001) is shown for the paired t-test comparison between the founder and the wild population.

	Mean (±SD) hybrid indices	
	Founder	Wild
HWE outliers included	0.454 (±0.056)	0.433 (±0.054)
HWE outliers removed	0.502 (±0.061)	0.453 (±0.037) *

Table A2 Comparison of the Weir & Cockerham's pairwise *Fst* values for the complete data set and after removal of Hardy-Weinberg equilibrium outlier loci. Pairwise *Fst* values were estimated between the parental populations and the combined hybrid population (founder and wild). All values differed significantly from 0 (10,000 bootstrap permutations, p<0.05).

	Pairwise *Fst*	
	HWE outliers included	HWE outliers removed
Kulan - Onager	0.278	0.133
Kulan - Hybrid	0.207	0.248
Onager - Hybrid	0.157	0.154

Heterozygosity estimates after removal of Hardy-Weinberg equilibrium outlier loci

The complete data set used for the population genetic analysis contained 2203 SNPs. Hardy-Weinberg equilibrium outlier loci were removed using the –hwe function in the program *vcftools*, which performs an exact test. SNPs with a p-values <0.05 were removed, resulting in a reduced data set containing 1738 SNPs. Analysis based on the reduced data set produced very similar results, except for the wild population which now displayed a small yet significant heterozygote excess instead of a heterozygote deficit (Table A3). The results for the other populations are consistent between the data sets, indicating robustness.

Table A3 Comparison of heterozygosity estimates for the four different populations with and without prior removal of Hardy-Weinberg equilibrium (HWE) outlier loci. Expected heterozygosity (*He*), observed heterozygosity (*Ho*). Significant levels (* p<0.01, **p<0.001) are shown for the paired t-test comparison of *He* and *Ho* for all populations.

	HWE outliers included		***HWE outliers removed***	
	He	*Ho*	*He*	*Ho*
Onager	0.201	0.211**	0.201	0.226**
Kulan	0.227	0.235*	0.223	0.251**
Founder	0.252	0.263**	0.261	0.289**
Wild	0.246	0.235**	0.253	0.258**

Table A4 Additional information used in the coalescent simulations. a) Timing of demographic events included in the simulations and their references. Generation time=7.5 years (Ransom et al., 2016). b) Search ranges for demographic parameters estimated in fastsimcoal2 simulations. Search ranges are uniform distributions defined by minimum and maximum values. References are given on which these ranges are based. c) Sensitivity analysis indicates minimal impact of search ranges for ancestral population size on model performance. Values in bold were used for the analysis. d) Maximum likelihood estimates for the estimated minor allele frequency site frequency spectra (SFS) and for the observed SFS for the complete model (model 5) with and without population growth rate (included growth rate of 1.87 was based on the female net reproductive rate described by Saltz and & Rubenstein, 1995).

a)

Point in time	Description of event (backwards in time)	Generations in the past	Reference
t_bot2	Start second bottleneck: Change in population size from current wild population to number of individuals released into the wild	3	(Saltz et al., 2000)
t_endbot2	End second bottleneck: Change in population size from number of individuals released into the wild to breeding core	4	(Saltz et al., 2000)
t_bot1	Start first bottleneck: Change in population size from breeding core to individuals imported to Israel	6	(Gueta et al., 2014)
t_admix	Admixture between onager and kulan individuals and end first bottleneck: change in population size from individuals imported to Israel to populations of origin	7	(Saltz et al., 2000)
t_div	Ancestral divergence between *E.h.kulan* and *E.h.onager* subspecies	466	(Bennett *et al.*, 2017)

b)

Parameter	Description	Search range (min-max)	Reference
N_wild	Effective population size (N alleles) of wild population	200-800	(Renan et al., 2018)
N_kulan	Effective population size (N alleles) of kulan population	50-400	(Denzau & Denzau, 1999)
N_onager	Effective population size (N alleles) of onager population	50-400	(Denzau & Denzau, 1999)
N_ancestral	Effective population size (N alleles) of ancestral population	7,000-20,000	See sensitivity analysis section C)
N_bot 2	Effective population size (N alleles) during bottleneck 2 i.e. individuals that were released from the breeding core into the wild population	10-76	(Gueta et al., 2014) Israeli Nature and Parks authority (unpublished)
N_founder	Effective population size (N alleles) of the breeding core population prior to release of individuals	40-200	Bar-David, personal communication
N_bot 1	Effective population size (N alleles) during bottleneck 1 i.e. establishment of the breeding core	4-22	(Saltz et al., 2000)

c)

Parameter	Search range (min-max)	Maximum Estimated Likelihood	Maximum Observed Likelihood	|Delta|
N_ancestral	7,000 – 20,000	-643.318	558.176	1,201.494
	50,000 – 100,000	-643.538	558.176	1,201.714
	100 – 200	-634.322	558.176	1,192.498

d)

Model	Description	Growth rate	Maximum Estimated Likelihood	Maximum Observed Likelihood	|Delta|
Model 5	Hybrid population, created by admixture of onager and kulan population, 2 bottlenecks (release into the wild and creation of the breeding core)	R_0=0	-643.316	558.176	1,201.49
		R_0=1.87	-851.81	558.176	1,968.162

References

Denzau, G., & Denzau, H. (1999). *Wildesel*. Stuttgart: Jan Thorbeck Verlag GmbH & Co.

Gueta, T., Templeton, A. R., & Bar-David, S. (2014). Development of genetic structure in a heterogeneous landscape over a short time frame: the reintroduced Asiatic wild ass. *Conservation Genetics*, *15*(5), 1231–1242. https://doi.org/10.1007/s10592-014-0614-z

Ransom, J. I., Lagos, L., Hrabar, H., Nowzari, H., Usukhjargal, D., & Spasskaya, N. (2016). Wild and feral equid population dynamics. In J. I. Ransom & P. Kaczensky (Eds.), *Wild Equids—ecology, management, and conservation* (pp. 87–104). Baltimore: Johns Hopkins University Press.

Renan, S., Speyer, E., Ben-Nun, T., Ziv, A., Greenbaum, G., Templeton, A. R., … Bouskila, A. (2018). Fission-fusion social structure of a reintroduced ungulate: Implications for conservation. *Biological Conservation*, *222*, 261–267. https://doi.org/10.1016/j.biocon.2018.04.013

Saltz, & & Rubenstein, D. I. (1995). Population Dynamics of a Reintroduced Asiatic Wild Ass (Equus Hemionus) Herd. *Ecological Applications*, *5*(2), 327–335.

Saltz, D., Rowen, M., & Rubenstein, D. I. (2000). The Effect of Space-Use Patterns of Reintroduced Asiatic Wild Ass on Effective Population Size. *Conservation Biology*, *14*(6), 1852–1861. https://doi.org/10.1111/j.1523-1739.2000.99227.x

Appendix B: Supporting Information for chapter 4

Global positioning system (GPS) data collection

Table B1 Location records previously collected for seven individuals equipped with global positioning system (GPS) collars recording at hourly intervals.

ID	Name	Sex	Start date	End date	Total time	Total number of records
594	Nahum Tacum	Male	12.07.2013	31.12.2014	17 months	14,101
595	Idan	Male	08.08.2013	18.01.2015	17 months	14,901
596	Ktsoutsy	Male	16.10.2012	05.12.2014	25 months	15,712
597	Short tail	Male	08.06.2013	18.04.2014	10 months	7,786
598	Gila	Female	07.08.2013	08.02.2015	18 months	14,980
1400	Alona	Female	08.07.2015	22.06.2017	24 months	16,700
1401	Ariela	Female	17.07.2015	18.02.2017	19 months	14,442

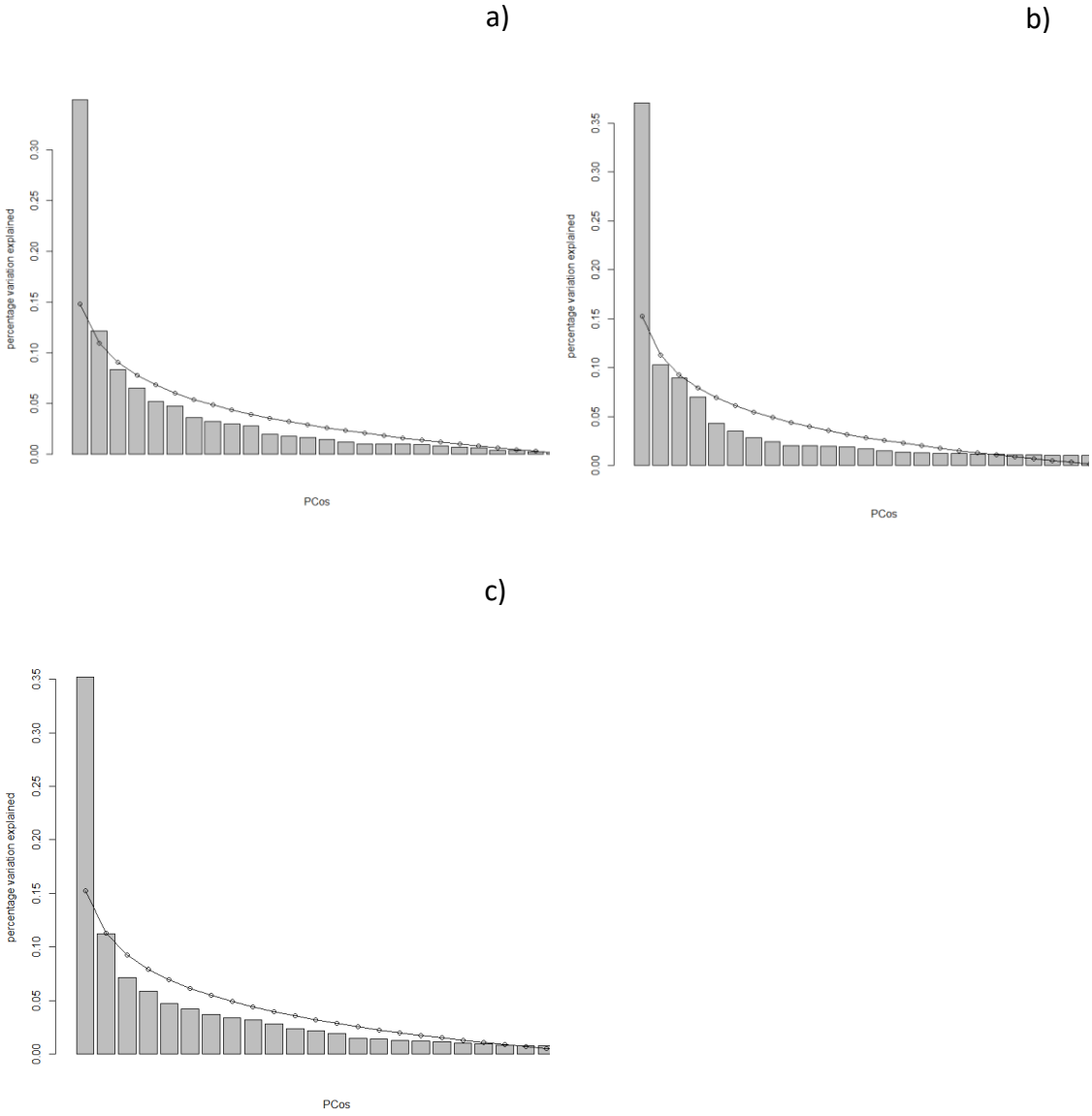


Fig. B1 Broken stick model used in the distance-based redundancy analysis. Percentage variance explained by the principal coordinates of the pairwise resistance matrices based on a) the species distribution model, b) slope and c) geographic distance. Connected dots are indicating the variance explained as expected under a broken stick model. Only the first (b, c) or first and second (a) principal coordinates explain more of the variance than expected.

Table B2 Simplified ranking matrix comparing proportional habitat use with overall habitat availability in the study area. Displayed is proportional habitat use for a) different categories of habitat slope and b) different categories of habitat suitability. "+" indicates the habitat in the row is used more than the habitat in the column, "-" indicates the opposite. "+++" and "---" indicate that the difference is significant at $p<0.05$.

a)	**Habitat slope**			
	0°-15° slope	16° -30° slope	>30° slope	Rank
0°-15° slope		+++	+++	2
16°-30° slope	---		+++	1
>30° slope	---	---		0

b)	**Habitat suitability**			
	High suitability	Intermediate suitability	Low suitability	Rank
High suitability		+++	+++	2
Intermediate suitability	---		+++	1
Low suitability	---	---		0

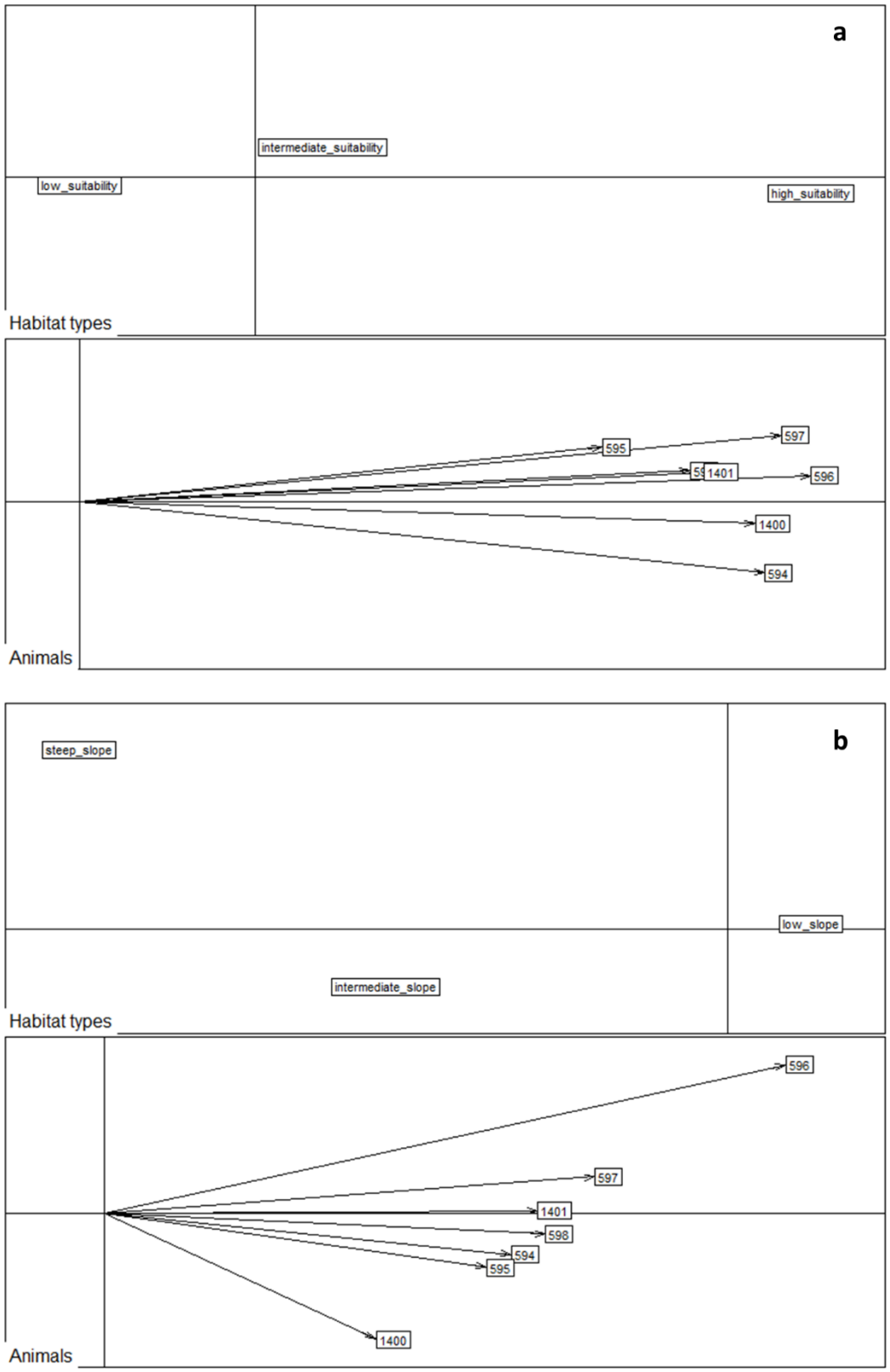


Fig. B2 Results of the eigenanalysis of selection ratios. Habitat selection by seven individuals equipped with global positioning system (GPS) collars was analysed with respect to a) habitat suitability and b) habitat slope. Top figures show the habitat types, bottom figures show habitat preference of each individual.

Individual movement records

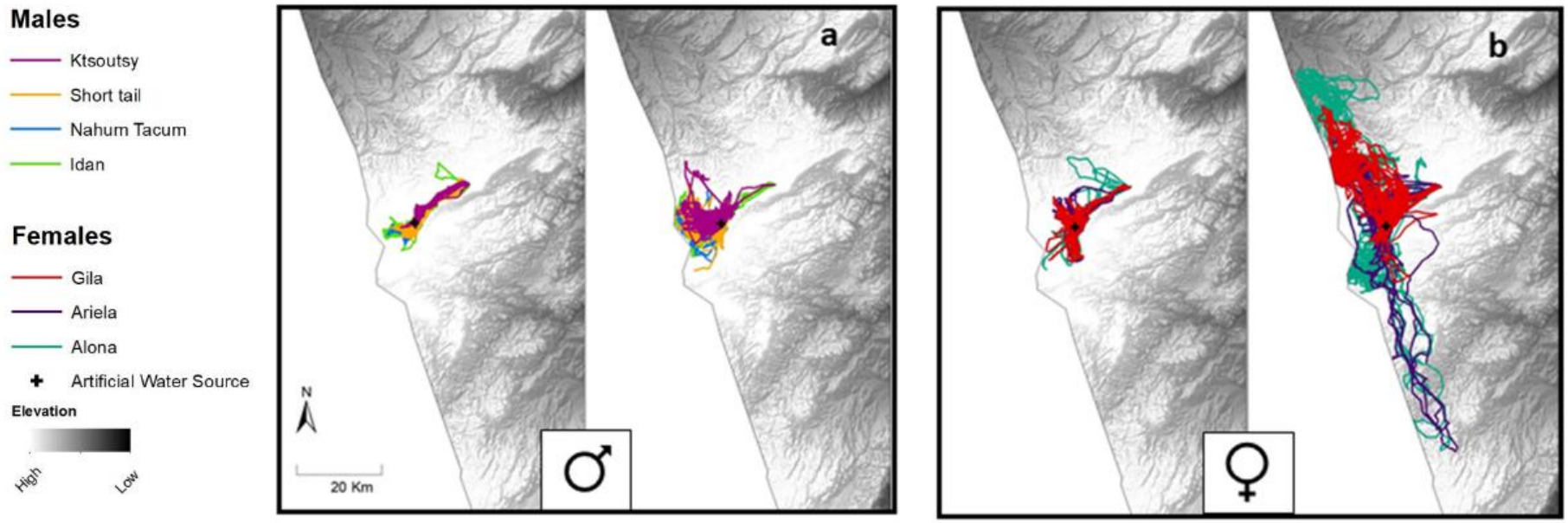


Fig. B3 Individual movement tracks for four males (left) and three females (right). Data represent hourly records obtained from global positioning system (GPS) collars over a minimum period of 10 months. Left panels indicated movements recorded during the breeding season (June-August), right panels represent movements during non-breeding season (October-May). For three individuals (Nahum Tacum, Alona, Ariela) data were obtained for two consecutive breeding seasons. Two females (Alona, Ariela) which displayed long-distance movements during the non-breeding seasons, returned to the area near the permanent water source during breeding season in two consecutive years.