Imperial College London

The ecology and population genetics of introduced deer species

Richard Guy Fautley

Department of Life Sciences, Imperial College London

A thesis submitted for the degree of Doctor of Philosophy

June 2013

Abstract

Deer have been introduced outside their native ranges numerous times worldwide, causing significant economic and environmental impacts. The scale of problems caused by some introduced deer species is expected to increase, while others are relatively harmless and may be important for conservation as their numbers decline in native regions. This thesis examines the factors promoting invasion in non-native deer, and investigates the population genetics of two deer species introduced to Great Britain.

Factors hypothesised to predict invasion success were tested using a comparative analysis of the outcomes of introductions of non-native deer populations. Two modelling approaches were taken to account for confounding effects in species-level and population-level analyses, and different factors were found to explain success at different stages of the invasion process.

Populations of roe deer expanding from refugia and reintroductions were examined, and differences in genetic diversity between core and peripheral populations, alongside a decline in diversity with rate of expansion from the core were found. High levels of differentiation between and among core and peripheral populations suggest that genetic drift is the major factor causing these patterns.

Using a genome scan of British and European roe deer, loci potentially under selection were identified. Association between some of these loci and environmental variables suggests that climatic extremes may have a role in exerting selective pressures on roe populations.

The Chinese water deer is severely declining in range and number in its native habitat, but rapidly expanding after being introduced to Great Britain. Relatively high levels of diversity in native Chinese populations, and significant differentiation between the Chinese and British populations were found. The source population of the British deer is likely to be extinct, and the level of genetic structuring indicates that conserving populations across both ranges is important to maintaining their diversity.

Declaration

I confirm that the work presented in this thesis is my own with the following acknowledgements:

A version of Chapter 2 has been published in Biological Invasions (Fautley, R., Coulson, T. and Savolainen, V. (2012) 'A comparative analysis of the factors promoting deer invasion', Biological Invasions, 14(11), 2271-2281). Vincent Savolainen and Tim Coulson participated in the study design and edited the manuscript. I participated in the study design, compiled datasets, generated DNA sequences, analysed the data and wrote the manuscript.

Some of the data used in Chapters 4 and 5 were contributed by collaborators. Dr Massimo Scandura shared Italian roe deer tissue samples, and Dr Frank Zachos shared German roe deer DNA samples. Dr Min Chen and Prof.Endi Zhang from East China Normal University shared DNA sequence data on native Chinese water deer, and Christian Miquel sequenced samples from one British population.

The use of any materials from other sources is acknowledged throughout this thesis.

Richard Fautley

Copyright Declaration

Acknowledgements

Firstly, I would like to thank my supervisors Vincent Savolainen, Tim Coulson and Rory Putman for all their advice, insight and guidance over the past few years. This work was funded by the BBSRC and the British Deer Society.

I am greatly indebted to the deer stalkers from the British Deer Society membership, the Forestry Commission and elsewhere who sent me roe deer tissue samples. I am also very grateful to Arnold Cooke for his insights into the Chinese water deer populations, and Paul Tinsley-Marshall, Richard Southwood and Nick Linsday who helped with my sample collection forays into the field.

A special mention must go to Martyn Powell, who repeatedly went above and beyond the call of duty, from helping me get set up in the lab to running the sequencer at ungodly hours and much more besides.

I would like to thank the members of the lab past and present: Jan, Alex, Haris, Helen, Silvana, Paul, Luis, Guillaume, Omar, Katya, Hanno, Zuzana, Andy, Adam, Andrew, Simon, Maria and Alan for all their lively discussions, unique senses of humour, valuable advice and support. Thanks also to the Silwood community for providing such a friendly and stimulating atmosphere in which to work, and Silwood Cricket Club for entertaining my attempts at batting.

Many thanks go to Min Chen and Endi Zhang for their kind hospitality during my visit to Shanghai. Thanks also to Betty Lingshan Yu and Xueping Quan for translating documents and helping me set up the collaboration.

Huge thanks to my housemates: Tim, Tariq, Pete, Chris, Coops, Ben and Jen at House 66, and to Tom H in Ealing for the many good times and helping maintain my sanity over the years. I hold Tom G in great esteem for his formatting wizardry. I am also especially grateful to my fellow journeymen Rob, Derek and Dom for sharing in my commiserations and showing me that the end was in sight.

I am eternally grateful to my family for their continued love and support: my mother for encouraging me to follow my interests, and my sister for sharing a childhood which shaped my ambitions.

Finally, I owe a huge debt of gratitude to my girlfriend Shuangzi for her unreserved kindness, keeping me going through the difficult times, and reminding me to focus on the future.

Table of Contents

List of Figures

List of Tables

Chapter 1: General introduction

The introduction of non-native species has resulted in enormous economic and environmental impacts (Pimentel 2002), and has important implications for the conservation of biological diversity (Allendorf 2003). Some introduced species can cause significant damage to their recipient environment and require controlling or eradicating. Other introductions have proved beneficial from a conservation perspective, preserving species and genetic diversity while the populations in the native range have declined. Introduced deer species provide a good study system in which to examine both invasion and conservation scenarios in wild populations.

1.1 Biological invasions

While biotic translocations among continents have occurred repeatedly over geological time (Vermeij 1991), the movement of species between continents happened at incomparably small rates and scales compared to the relatively recent human-mediated biotic interchanges. The creation of a discontinuous distribution of species caused by deliberate or accidental human activity is thought be one of the most important biogeographical processes in recent history (Guo 2006). With the potential to inflict significant environmental and economic damage upon the recipient environment, invasion by non-native species ranks among the greatest threats to global biodiversity (Allendorf 2003).

Changes in the composition and structure of native communities (e.g. Fritts 1998; Suarez 1998) and modification of ecosystem function (e.g. O'Dowd 2003) can occur as a result of environmental damage caused by non-native invasive species (Whitney 2008). The cost of biological invasions is estimated to be over $US394 billion annually for the ~120,000 non-native species introduced into the United States, United Kingdom, Australia, South Africa, India, and Brazil (Pimentel 2001). There has been an understandably large amount of scientific interest in the field of invasion biology, with the ultimate aim of applying the research to predict, control and manage species invasions.However, while some

species invasions can cause major damage to the recipient environment, the effects of many introduced species are neutral or sometimes even beneficial.

1.1.1 Developments in invasion biology

The application of systematic study methods to invasive species was prompted by Elton (1958) in his seminal book *The ecology of invasions by animals and plants.* He set forth the argument that invasions by non-native species *"*...are so frequent nowadays in every continent and island, and even in the oceans, that we need to understand what is causing them and try to arrive at some general viewpoint about the whole business". Research in the field has developed from studies conducted on case histories of individual species into analyses of broad taxonomic groups (e.g. Rejmanek 1996) aiming to identify common attributes associated with invasive success (Reichard & White 2003). The ability to predict the propensity of a species to invade and its impact on the recipient ecosystem are major goals in ecology, driven in large part by the growing realisation of the rising ecological and economic costs of species invasions.

1.1.2 Predicting invasive success

The characteristics that predispose a species to invade are subject to much research, as identifying them could help improve predictions. Initial efforts to define a clear-cut set of predictive traits applicable across species were inconclusive (Williamson 1996). While some generalisations about the invasion process could be made, different characteristics of species were found to be important in different habitats, suggesting that the underlying factors behind successful biological invasions were highly idiosyncratic, with few commonalities between instances of successful invasions. The view of some ecologists became that the ecological behaviour of a non-native species may be nearly impossible to predict (Williamson 1999). Reasons for the failure to find features consistently

associated with invasive success may due to the qualitative nature of the available data, examining only one stage of the invasion, and lack of information on species that failed to become invasive (Kolar & Lodge 2001). Studies now tend to examine traits associated with invasive success using comparisons between and within broad taxonomic groups (Jeschke & Strayer 2006; Jeschke 2008; Sol *et al.* 2008a)

Invasion biology integrates large-scale ecology and evolutionary biology (Hochberg & Gotelli 2005), and is now emerging as an interdisciplinary science, comprising many disparate subject areas from life-history studies to genetics and ecosystem dynamics. There is still little in the way of broad consensus in the field. Areas of agreement are that invasive species are one of the major causes of biodiversity loss, that prevention of introduction is more effective than eradication of established populations, and that more research is required to develop effective predictive and control measures (Allendorf 2003).

1.1.3 Features of biological invasions

Identifying the characteristics of a successful invasion could enable better understanding of the mechanisms responsible, inform efforts to control currently occurring invasions, and may enable the prevention of future impacts. It has been asked whether it is the characteristics of the species which is invading, or the attributes of the ecosystem being invaded that is responsible for invasive success (Guo 2006). The true picture is likely to be both, with a complex interplay between the two sets of factors specific to a particular group of taxa. Factors affecting the success of species introduced to novel environments include the idiosyncrasies of the release event and the characteristics of the recipient environment (Duncan *et al.* 2003).Broadly, the types of characteristics with influence on invasion success can be categorised into 1) Event, 2) Region, and 3) Species specific effects (*sensu* Sol *et al.* 2008a).

1) Event-specific effects relate to the introduction event. The most important event-specific effects relate to 'propagule pressure' or introduction effort (Lockwood *et al.* 2005). Factors include frequency and number of individuals released, whether the release was deliberate or accidental, and to an island or mainland. However, in a meta-analysis Jeschke(2008) found that there is no significant difference in establishment successes between introductions to islands or continents.

2) Region-specific effects relate to features of the recipient environment. Habitats may exhibit differences in their 'invasibility' due to biotic and abiotic factors. For example, biotic factors include the phylogenetic relatedness of the native species to the introduced species (Strauss *et al.* 2006). Abiotic factors include the degree to which features of the novel region are similar to the native region (e.g. climate and latitude).

3) Species-specific effects may be genetic or phenotypic. Genetic factors of the introduced population(s) include having sufficient diversity to enable rapid evolution (Lee 2002; Dlugosch 2008), or the ability to hybridise with native species (Mallet 2005). Phenotypic factors of the species include population characteristics (e.g. range size), life-history, habitat generalism, diet and physiology.

1.1.4 Current trends in mammalian invasion biology

With improved ecological data and the application of rigorous quantitative methods, many large-scale studies examining the characteristics of species introductions into novel environments have been conducted on broad taxonomic groups (Sol *et al.* 2008b). Relative to other animals, mammal species are more likely to successfully establish themselves in novel environments (Clout & Russell 2008). The IUCN Red List classifies 2.6% of extant land mammals as 'successful invaders' – with Artiodactyla containing the highest proportion (14.7%) of successful invaders. The family with the highest proportion of invasive species is Cervidae (deer), with 29.2% of the species having self-sustaining non-native wild populations (Clout & Russell 2008). Of these, the red deer (*Cervuselaphus*)

is one of the few mammal species to have successfully established at more than 30 locations worldwide (Long 2003).

Differences in relative brain size have been shown to be associated with establishment success of introduced mammals (Sol *et al.* 2008a). Species possessing large brains relative to body size are more likely to be successful in establishing themselves in a new environment than those with relatively smaller brain size, even after controlling for differences in propagule pressure and habitat generalism. Similar results were found in a study of introduced bird species correlating larger brain size with higher establishment success (Sol *et al.* 2005), and mechanisms underlying this correlation are better understood (Sol *et al.* 2007). However, in a review of six studies of ecological factors associated with invasive success, only three features were reported to be consistently associated with increased probability of establishment success for non-native mammal species (Clout & Russell 2008): a) Number of individuals released into the habitat, b) Natural range size and c) Climate temperateness in the novel environment.

The only feature consistently shown to be strongly associated with invasive success is propagule pressure (Lockwood *et al.* 2005; Travis *et al.* 2005; Colautti 2006). The studies of other features of biological invasions are constrained by the data available for the ranges of taxa under examination and as such rarely analyse exactly the same sets of features, making broad-scale conclusions on features of biological invasions hard to draw at present. Assessing invasions at a taxon or region-specific level is likely to be more informative about the characteristics associated with the process.

1.2 The genetics of introduced populations

During establishment, genetic differentiation of an introduced population from its source population is expected due to founder effects, drift and natural selection (Nei *et al.* 1975). Strong natural selection is expected to act during the invasion process, due to differences between the native and novel environments (Lee 2002) and also because traits that are advantageous at one stage of the invasion

process may be disadvantageous at another (Kolar & Lodge 2001). In the case of introduced populations undergoing rapid range expansion, and vulnerable populations heading toward extinction in the wild, studies of genetic variation have the potential to contribute to management and conservation plans. In both instances, understanding the capacity of species to adapt to new environments is of great importance.

1.2.1 Colonisation

Sufficient genetic variation is required for adaptive evolution to occur, which may not be present in introduced populations below a certain size (Lee 2002). This corresponds with the demonstrated importance of propagule pressure in establishment success (Forsyth & Duncan 2001; Lockwood *et al.* 2005; Travis *et al.* 2005; Colautti 2006; Jeschke & Strayer 2006), and relates to a paradox in invasion biology: how do introduced species with reduced genetic diversity and small population size succeed relative to native species well-adapted to their surroundings? The initial population dynamics of introduced species can determine levels of genetic diversity retained within and among populations. A population that increases in size rapidly will lose relatively little variation, whereas much variation can be lost when a population remains small over many generations (Nei *et al.* 1975). The history of introduced populations typically involves complex differences in number and size of founding propagules, and in some instances, admixture between populations introduced to different regions (Kolbe *et al.* 2004). Genetic approaches have been applied to identify the sources of introduced populations and routes of invasion (Estoup & Guillemaud 2010), and comparisons of native and introduced populations have attempted to quantify the loss of genetic diversity during the invasion process (Dlugosch & Parker 2008).

1.2.2 Range expansion

Genetic studies have been highly informative about the phylogeography of post-ice age range expansions (e.g. Hewitt 2000). The methods employed in these studies were thought to be applicable to recent invasive events occurring over a much shorter timescale (Sakai 2001) with the potential to improve understanding of the processes involved in invasions. However, applying phylogeographic methodologies to the study of invasive species can be misleading. Analyses that examine concordance between geography and phylogenetic networks or trees require a sufficient number of generations to have elapsed for the effects of drift and mutation to impact upon the distribution genetic variation – a scenario that is highly unlikely to have occurred during the timescale of human-mediated introductions (Fitzpatrick *et al.* 2012).

Population genetic methods can be applied to understanding post-establishment patterns of spread in introduced species. Modelling and simulation studies show that founder events during range expansion can involve genetic differentiation and loss of genetic variation from newly established areas (Austerlitz et al. 1997; Excoffier 2004). Where the population derives from a small number of individuals, the founder event is likely to involve a change in genetic composition (Hartl & Clark 1997). Subdivision of populations generally results in the emergence of genetic structure (Slatkin 1987), and differentiation from founder populations during range expansion may be significant (Le Corre & Kremer 1998). When demographic growth is high, range expansion can be accompanied by large rates of migration, which can act to decrease genetic differentiation between fragmented populations, thereby affecting genetic population structure. By developing a more comprehensive understanding of invasion processes, risk assessments can be better informed and prevention and control measures can be better implemented.

Species that successfully establish populations when introduced into novel environments also provide a ‘natural experiment’ (Sol *et al.* 2008b)that can be used to examine the influence of range expansions on population genetic parameters by analysing the distribution of genetic variation across the range. The central–marginal hypothesis, otherwise known as the ‘abundant centre’ model (Sagarin & Gaines

2002; Eckert *et al.* 2008), assumes that a species is most abundant in the core of the range, and decreases in density towards the edge. Populations at the periphery of the range are expected to have smaller effective population sizes, and be more isolated and fragmented relative to core populations (Vucetich & Waite 2003). This hypothesis predicts the highest levels of genetic diversity and gene flow at core populations, and that these parameters will decrease towards the edge of the range.

1.2.3 Adaptation

Changing environmental conditions and colonisation of different areas can result in rapid adaptive evolution of wild populations to their new environment (Reznick & Ghalambor 2001), and in the case of introduced populations in new environments undergoing rapid range expansion, ecologically based selection could be occurring leading to local adaption (Sakai 2001; Lee 2002). Population expansion of introduced species is expected to be accompanied by local adaptation, where the rapidly increasing population size facilitates a response to strong directional selection (Whitney & Gabler 2008; Sexton *et al.* 2009). While neutral population genetic variation has been assumed as a surrogate for evolutionary potential (Reed & Frankham 2003), natural selection is likely to affect only a small number of loci in the genome (Nielsen 2005). Assessment of adaptive genetic diversity in wild populations using population genomic approaches (Luikart *et al.* 2003) has generated significant insights (Schoville *et al.* 2012). However, accurately identifying loci under selection and understanding which selective pressures are acting on non-model species in the wild is challenging (Nunes *et al.* 2011). Understanding the genetics of adaptation in introduced species is important for identifying candidate genes underlying traits involved in 'invasiveness' (Prentis *et al.* 2008), predicting future distributions (Kawecki 2008) and ultimately understanding how species respond to new and changing environments (Manel *et al.* 2010).

1.2.4 Conservation

Changing climate, coupled with increased habitat destruction and fragmentation pose major threats to the maintenance of biodiversity (Saunders *et al.* 1991). Genetic variation is fundamental to biodiversity, and its loss is expected to have negative effects on the survival of populations (Lande & Shannon 1996). Population genetic theory predicts that an increased frequency of deleterious alleles resulting from genetic drift and inbreeding can reduce short-term viability, while sustained loss of genetic variation is expected to have negative effects on long-term population viability, resulting in decreased adaptive potential (Frankham 1995). Furthermore, loss of genetic variation in small populations may increase extinction risk due to increased disease susceptibility, and decreased reproductive fitness and adaptive flexibility (Allendorf & Luikart 2007). In the context of rapid environmental change, determining the ability of threatened species to survive and adapt is critical, as isolated populations with low levels of dispersal are especially vulnerable to extinction (Lande 1993). Determining levels of genetic diversity in threatened species can therefore help inform suitable conservation strategies (Frankham *et al.* 2002). Non-invasive genetic sampling is very valuable for conservation genetic studies, enabling the generation of genetic data on species in the wild which would not otherwise be possible due to the limitations of conventional sampling techniques(Taberlet *et al.* 1999).

1.3 Deer as a study system

1.3.1 Family Cervidae

Cervidae (deer) are a Family of ruminant mammals within the Order Artiodactyla. They have a global distribution, with native species extant in all continents except Australasia and Antarctica, and consist of a monophyletic group of 51 species within Ruminantia(Grubb 2005). The Cervidae have successfully colonised all continents but Antarctica, and exist in a wide variety of ecosystems (Putman 1988; Geist 1999).

As a result of their aesthetic and economic appeal as exotic species, deer have been introduced numerous times outside their native range, and now comprise the largest proportion of all invasive mammal species (Clout and Russell 2008). The presence of non-native deer species has brought both benefits and costs to the recipient communities, yet there is increasing concern about their impacts (Gill 1992; Fuller & Gill 2001; Cote *et al.* 2004; Moriarty 2004; Ward 2005; Dolman & Waber 2008; Acevedo *et al.* 2010; Gill & Morgan 2010; Nunez *et al.* 2010). Through grazing and browsing, non-native deer may have a significant economic impact on agriculture and forestry (Ratcliffe 1987; Chadwick *et al.* 1996; Putman & Moore 1998) as well as on the structure and dynamics of natural vegetational systems (Pollard & Cooke 1994; Cooke & Lakhani 1996; Cooke & Farrell 2001; Jaksic *et al.* 2002; Relva *et al.* 2010). They may also compete with native ungulate species (Keiper 1985; Chapman *et al.* 1993; Hemami *et al.* 2004, 2005; Feldhamer & Demarais 2009; Feldhamer & Armstrong 1993) or other native species, and in the extreme may impact upon native deer species through hybridisation (Lowe & Gardiner 1975; Simberloff 1996; Goodman *et al.* 1999; Senn & Pemberton 2009).

Comparative studies between closely related taxa exhibiting differences in ability to establish and spread in novel environments holds the promise of identification of traits associated with invasive success and better understanding of the mechanisms underlying a successful invasion (Sol *et al.* 2008b). The family Cervidae contains the highest proportion of invasive species of all mammalian families, and species within the Cervidae exhibit large interspecific variation in ability to become invasive (Clout & Russell 2008). Some are highly localised, whereas others are capable of rapid establishment and spread - invading numerous different habitats. For example, the red deer (*Cervuselaphus*) is one of the few mammal species to have successfully established at more than 30 locations worldwide (Long 2003).

Phylogenetic relationships

The family Cervidae consistently forms a monophyletic group in supertree analyses of Ruminantia(Fernandez & Vrba 2005) and Cetartiodactyla(Price *et al.* 2005), respectively. Views on the phylogenetic relationships within Cervidae have reached a broad consensus among molecular studies (Randi *et al.* 1998; Pitra *et al.* 2004; Gilbert *et al.* 2006). The findings of the most comprehensive molecular phylogenetic analysis of Cervidae to date (Gilbert *et al.* 2006) using DNA sequence data from 25 species for mtDNA and nuclear fragments are congruent with the results of an analysis conducted using only *cytb* sequence data from 11 cervid species (Randi *et al.* 1998). The results of a phylogenetic analysis using 18 characters defined from the vocal behaviour of 11 Cervid species (Cap *et al.* 2008) were congruent with the current molecular phylogenies (Randi *et al.* 1998; Pitra *et al.* 2004; Gilbert *et al.* 2006), having an identical topology to the Cervinae (a la Gilbert et al, 2006) and forming (*Rangifer, Odocoileus*) and (*Capreolus, Hydropotes, Alces*) clades within the Capreolinae (Figure 1.1).

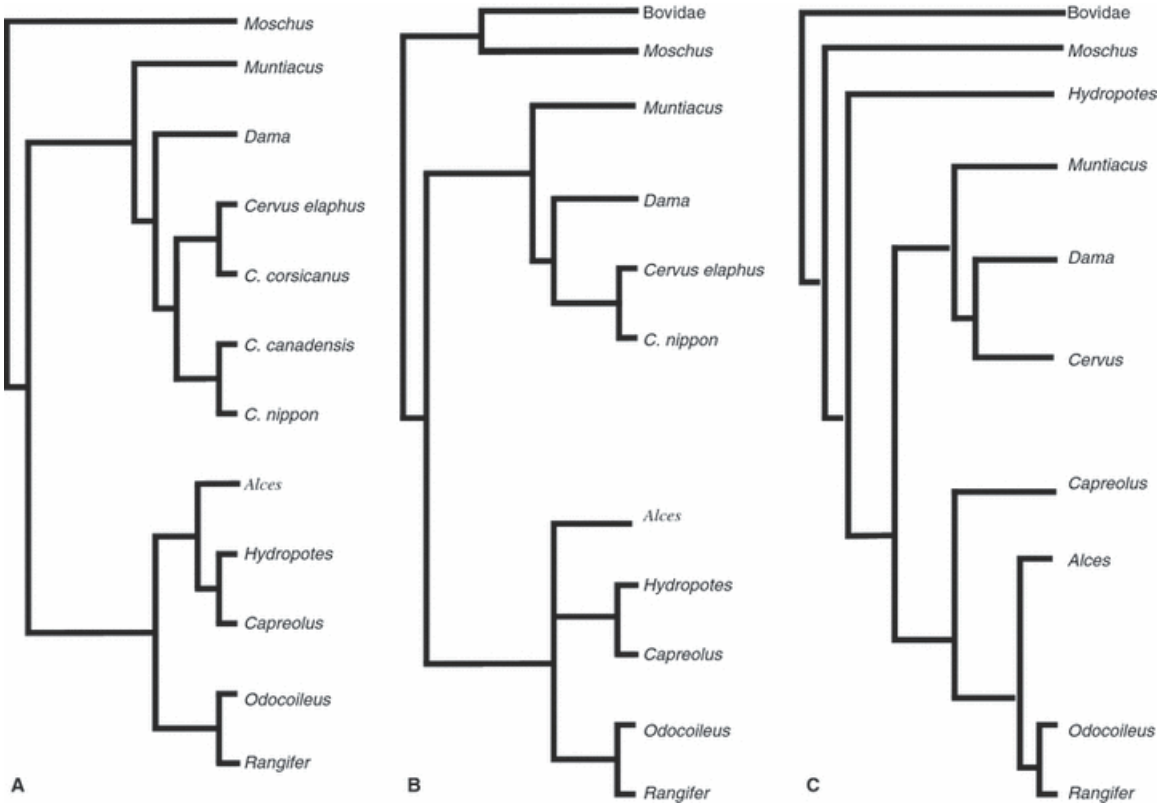


Figure 1.1 Cladograms of Cervid phylogenetic relationships. (A) mtDNA*cytochrome b (cytb)* data (Randi et al., 1998; Pitra et al., 2004); (B) *cytb* and *CO2*mtDNA and nuclear fragments: exon 2 of *alpha-lactalbumin* and intron 1 of the gene encoding *protein kinase C iota* data (Gilbert et al., 2006); (C) morphological data (Groves & Grubb 1987). From Cap et al. (2008).

These studies divide the Cervidae into two subfamilies defined by strongly supported clades: Cervinae and Capreolinae. These two clades correspond to the classical division of the family into Telenmetacarpalia or 'New World Deer' (Capreolinae) and Plesiometacarpalia or 'Old World Deer' (Cervinae). Within the Capreolinae, three tribes are defined: Odocoileni, Capreolini (containing the roe and Chinese water deer) and Alceini. Within the Cervidae, two tribes are defined: Muntiacini (containing the Reeves' Muntjac) and the Cervini (containing the Fallow, Red and Sika deer). These results contradict phylogenies constructed using morphological characters (Groves & Grubb 1987; Janis 1993), regarding the group containing *Alcesalces*and *Capreoluscapreolus*, and the nesting of *Hydropotes inermis* within the Capreolinae rather than sister to the rest of the Cervidae (Fig. 1). Furthermore, following detailed examination, the internal taxonomy of the genus Cervus – which is based on morphological characters which are looking increasingly doubtful (Cap *et al.* 2008), seems to be in need of revision (Ludt *et al.* 2004; Pitra *et al.* 2004; Gilbert *et al.* 2006).

1.3.2 The European roe deer

The European roe deer (*Capreoluscapreolus*L.) is a small ungulate species in the subfamily Capreolinae (Figure 1.2). It is widely distributed across Europe except for Ireland, some Mediterranean islands including Corsica, Sardinia and Sicily, the tundra regions and northern Russia. Range limits are defined by the cold climate and associated snow accumulation in the north (Holand *et al.* 1998) and in the south by water requirements and high temperatures (Wallach *et al.* 2007). Roe have existed in Europe for at least 600,000 years (Lister *et al.* 1998), and are a typical faunal element of the Holocene epoch, from c.10, 000 BP to the present day (Sommer *et al.* 2009). The impacts of glaciations on roe distribution have been determined using phylogeographic analyses (Randi *et al.* 2004; Lorenzini & Lovari 2006), with three mitochondrial clades being defined: Clade East (in Greece and the Balkans), Clade West (in Iberia) and central (across much of Northern, Central, Western and Eastern Europe). Populations in Europe experienced a severe decline due to deforestation and overhunting prior to the nineteenth century, but have greatly increased in range and density in recent

years, aided by numerous translocations (Andersen *et al.* 1998). In post-war Great Britain and Europe, agricultural intensification and woodland fragmentation produced substantial spatial heterogeneity in natural habitats (Robinson & Sutherland 2002; Stoate *et al.* 2009). Roe deer responded well to the changes in landscape features and have spread into a variety of habitats(Andersen *et al.* 1998; Ward 2005; Morellet *et al.* 2011). The behavioural plasticity of the species has been largely credited with underlying this great ecological flexibility (Hewison *et al.* 2001), yet there may also be a role for other factors such as physiological responses and genetics in explaining the adaptation of roe deer to new environments.

Figure 1.2 European roe deer buck

1.3.3 The roe deer of Great Britain

The European roe deer is native to Britain. There was a severe reduction in distribution and abundance of the populations up to the eighteenth century due to habitat destruction and overhunting. The majority of populations in England and Southern Scotland were believed to have gone extinct (Whitehead 1964), with refugial populations surviving mainly in woodlands in Central and North-western Scotland, and potentially in Northern England and lowland Scotland (Whitehead 1964; Prior 1995; Baker 2011). All Southern English populations are thought to be descended from later reintroductions (Whitehead 1964; Hewison 1995, 1997; Ward 2005; Baker & Hoelzel 2013). The reintroductions to England occurred from Austria to Windermere, North-western England, from Germany to East Anglia, and from Scotland and unknown sources to Southern England (Whitehead 1964; Prior 1995; Baker & Hoelzel 2013). Roe currently have the widest distribution of all deer in

Great Britain, and their numbers are estimated to range in the hundreds of thousands (Hewison & Staines 2008). They are now absent only from areas in Kent, the Midlands and Wales (Ward 2005) and most Scottish islands. Roe deer have a high capacity for demographic and range expansion (Andersen & Linnell 2000). Damage caused by roe deer has been recorded in commercial forestry, conservation woodlands, fruit and orchard trees and other horticultural crops, and it is anticipated that more damage will occur as the populations continue to expand (Putman & Moore 1998). Furthermore, roe deer account for approximately 29% of deer-vehicle collisions with reported species detail in England (Langbein 2011).

Integrating genetic, environmental, and phenotypic data is a promising and powerful way of better understanding population demographic history, structure, and evolutionary patterns, and for detecting the selective forces acting upon the genome in wild populations. Roe deer are well suited to such analyses, with their abundance and rapid spread having been the focus of wealth of ecological studies (Andersen *et al.* 1998). Molecular work on the genetics of this charismatic deer has so far yielded insights into the phylogeography of the species (Randi *et al.* 2004; Lorenzini & Lovari 2006; Zachos *et al.* 2006a; Zachos *et al.* 2006b; Gentile *et al.* 2009), interactions with the landscape (Wang & Schreiber 2001; Coulon *et al.* 2004; Coulon *et al.* 2006), the influence of heterozygosity on fitness measures (Zachos *et al.* 2007; Da Silva *et al.* 2009; Baker & Hoelzel In press) and the effects of reintroduction (Hewison 1995; Zachos *et al.* 2006a; Gentile *et al.* 2009; Baker & Hoelzel 2013) on the distribution of genetic variation.

1.3.4 The Chinese water deer

The Chinese water deer (*Hydropotes inermis*Swinhoe 1870) is a small ruminant species in the subfamily Capreolinae and is native to East Asia. Its preferred habitats are reed-beds or tall, damp and undisturbed grasslands (Zhang *et al.* 2006). It is understood to be comparatively primitive (e.g.Randi *et al.* 1998) and like species of the genus Muntiacus, it retains tusks (Figure 1.3) which are used in

combat during the rut (Cooke & Farrell 1998). It is the only extant deer species that does not possess antlers (Haltenorth 1963), with the most likely evolutionary scenario being secondary loss (Randi *et al.* 1998; Gilbert *et al.* 2006). Outside the rutting season its behaviour is almost entirely solitary, quite unlike that of any other deer species and having more in common with the most primitive ruminant families of moschids and tragulids(Dubost *et al.* 2011). While a matriarchal system is often the foundation of social life in deer (Putman 1988) grouping in Chinese water deer is fleeting and unstable. Interactions between individuals are rare and limited to reproduction and rearing of young. After rearing, contact between Chinese water deer mothers and their young is minimal (Dubost *et al.* 2011).

In addition to such features, Chinese water deer also have a number of other 'ancient' chromosomal, anatomical and reproductive characteristics. The species has an ancestral (2n = 70) diploid number (Yang *et al.* 1997), its brain size is intermediate between moschids and other old world cervids (Bützler 1988), and it has the highest reproductive capacity of all cervids with an average gestation length of 175 days, does bearing over three fawns (Cooke & Farrell 1998) which are weaned and grow more quickly than other deer, and take the shortest time (293 days) to reach sexual maturity (Dubost *et al.* 2008).

Figure 1.3 Chinese water deer buck

Two subspecies are recognised: the Chinese water deer *Hydropotes inermisinermis*(Swinhoe 1870) and the Korean water deer *Hydropotes inermisargyropus* (Heude 1884). However, these country-specific designations based on pelage colour are not consistent with molecular relationships, and pelage colour has been observed to change seasonally (Cooke & Farrell 1998). Two sympatric phylogroups based on mtDNA*cytb* and control region haplotypes have been found: a major clade containing individuals from China and Korea and a minor clade containing Korean individuals only (Koh *et al.* 2009).

The historical distribution of the deer once stretched along the eastern part of China from Liaoning to Guangdong out to the lower Yangtze Basin, and the Korean Peninsula (Ohtaishi & Gao 1990). The population in China has reduced considerably in range and numbers recently due to habitat loss and poaching (Hu *et al.* 2006). It is currently restricted to fragmented populations in the eastern Yangtze

Basin and the Zhoushan Islands off the coast of Zhejiang Province. Published estimates of total numbers of the Chinese water deer in China have declined from 10,000-30,000 in 1993 to fewer than 5000 in 2011 (M. Chen unpublished data). The two largest remaining mainland populations are in Jiangsu and Jiangxi Provinces. The estimated numbers in Jiangsu have reduced from 1200-1500 (Ohtaishi & Gao 1990) to fewer than 400 (M. Chen, unpublished data). This is corroborated by studies finding that the populations within the Yancheng Nature Reserve in Jiangsu have reduced in range and become more fragmented (Zhu *et al.* 2004). The population on Jishan Island in Poyang Lake Nature Reserve, Jiangxi has more than halved in number from around 1000 (Ohtaishi & Gao 1990) to fewer than 500, and small, fragmented wild populations remain in eastern Anhui Province where poaching and habitat destruction are on the rise. The Zhoushan Islands are estimated to support 2000-3000 of the remaining deer (M. Chen, unpublished data).

While there are no data on the populations in North Korea, the species is believed to be 'moderately widespread' along the west coast, with 'reasonable numbers' in the lowlands, and abundant in the demilitarised zone (Harris & Duckworth 2008). In South Korea the deer occur in most areas of the country, with greater abundance in rural regions. Their occurrence is negatively correlated with human population size, and density varies with habitat type and altitude (Kim *et al.* 2011).

Chinese water deer were introduced to the UK in the 1870s, initially to London Zoo. Starting in 1896, the deer were sent gradually to Woburn Abbey in Bedfordshire. 19 individuals were imported to the park by 1913 (Chapman 1995). Assuming that no further introductions took place, all of the Chinese water deer in the UK are descended from this one population. 32 individuals were transferred to the nearby Whipsnade Park in 1929 (Cooke and Farrell 2008). During WW2 a number of Chinese water deer escaped from Woburn and established localised feral populations (Cooke 2009). In 1950 a small number of deer were released near Woodwalton Fen National Nature Reserve (Chapman 1995). Wild populations founded from escapes and deliberate releases since the 1940s are now discontinuously distributed across Bedfordshire, Cambridgeshire, Norfolk and Suffolk (Cooke & Farrell 1998), and

areas surrounding the Norfolk Broads contain the highest concentration of favoured habitat (Acevedo *et al.* 2010).

Total estimated numbers in the UK stand at 7000 in the wild (A. Cooke pers. comm.) and >500 in semi-captivity. 300-400 of the semi-captive numbers are in Whipsnade Park (N. Lindsay pers. comm.). Natural rate of spread is estimated to be 1 km/yr (Cooke and Farrell 1998), but may have increased in recent years (Ward *et al.* 2008). Dispersal is likely to be limited by availability of suitable habitat and presence of other deer species, especially muntjac and roe (Acevedo *et al.* 2010). In 2010 Defra included the Chinese water deer on Schedule 9 of the Wildlife and Countryside Act 1981, which bans unauthorised releases in the wild. In its native range the Chinese water deer has been reclassified on the IUCN Red List from LR/NT in 2000 to Vulnerable (Harris & Duckworth 2008). In China, habitat destruction and fragmentation looks set to continue at an accelerating pace. The Lower Yangtze wetlands, especially Dongting and Poyang Lakes are being degraded by pollution and sedimentation (Dudgeon 2010).

1.4 Thesis aim and objectives

In this thesis I aim to determine the drivers of invasion success in introduced deer, and develop a better understanding of the population genetics of two introduced deer species in relation to their history and contemporary processes.

The objectives of this thesis are to investigate:

- The factors underlying establishment and spread in introduced deer species.
- The influence of reintroduction and range expansion on the population genetics of the roe deer in Great Britain.
- The genetic basis of adaptation to environmental conditions in populations of European roe deer.
- The conservation genetics of the native and introduced Chinese water deer.

1.5 Outline of thesis

In Chapter 2 I uncover the determinants of establishment and spread in introduced deer species. Using historical records on the outcomes of non-native deer introductions in conjunction with data on the species, region and introduction event, I employ comparative analyses to determine which factors explain success at two different stages of the invasion process. I also present the largest phylogenetic estimate of the deer, which was used to control for shared ancestry in the statistical tests.

Chapter 3 focuses on range expansion from multiple reintroduced and remnant roe deer populations in Great Britain. I assess population genetic structure in the context of documented population history, compare levels of diversity at the core and periphery of the ranges and test for the effect of range expansion on differentiation between populations. In Chapter 4 I search for signatures of selection in expanding European roe deer populations. Combining genomic data with climatic and land cover variables, I examine the influence the environment has in shaping the distribution of genetic variation in the roe deer genome.

Examining the effects of translocations in a conservation context, in Chapter 5 I study the conservation genetics of the Chinese water deer. Using non-invasive sampling, attempt to determine the source of the British populations, and compare levels of diversity and differentiation between introduced and native Chinese populations with a view to making recommendations for conservation in their native range and management in their introduced range.

Chapter 2: A comparative analysis of the factors promoting deer invasion

A version of this Chapter has been published in Biological Invasions (Fautley, R., Coulson, T. and Savolainen, V. (2012) 'A comparative analysis of the factors promoting deer invasion', Biological Invasions, 14(11), 2271-2281).

Introduction

With the potential to inflict significant economic and environmental damage upon host environments, invasion by non-native species is one of the greatest threats to global biodiversity (Walker & Steffen 1997; Allendorf 2003). Increasing concerns worldwide about invasive non-native species have resulted in international and national commitments to assess the status and impact of non-native species, and to devise action plans for their control. There are initiatives at both the global (Boyle 1996) and national levels aiming to achieve these goals. For example, in England the Non-Native Species Secretariat was established in response to the statutory obligations of signatory member states. It has set out the UK Government's approach to the sustainable management of wild deer in England in the Wild Deer Action Plan (Defra 2004). Policy based on risk assessments of non-native species requires objective and reliable information on the factors predictive of success at different stages of the invasion process.

Invasions characteristically proceed through four stages: transport, introduction, establishment, and spread (Kolar & Lodge 2001). Intervention at the early stages of invasions is crucial for their control and prevention (Courchamp *et al.* 2003; Hulme 2006). Attempts to identify common attributes associated with invasive success have developed from studies conducted on case histories of individual species into analyses of broad taxonomic groups (e.g. Rejmanek 1996). However, lack of historical records on introductions that failed to establish has reduced the scope of many analyses to the later stages of invasion (Kolar & Lodge 2001). Relative to other animals, mammals are more likely to successfully establish in novel environments (Clout & Russell 2008). Invasive mammals are a major problem both from a conservation perspective and in their actions changing local ecosystems (Pimentel 2001; Courchamp *et al.* 2003; White *et al.* 2008; Forsyth *et al.* 2010; Nunez *et al.* 2010). The IUCN Red List classifies 2.6% of extant land mammals as invasive, with Artiodactyla containing the highest proportion of successful invaders (Clout and Russell 2008).

There are various hypotheses to explain successful invasions, which may broadly be grouped into three categories (*sensu* Sol et al. 2008a): (1) species-level, i.e. life-history variables and reproductive

characteristics, (2) region-level, i.e. features of the recipient area, and (3) event-level, i.e. features of the introduction event (Table 2.1). An analysis of mammals introduced to Australia found successfully established species had greater numbers of introductions and numbers of individuals per introduced population, a larger area of suitable habitat available, and larger native ranges (Forsyth *et al.* 2004). In New Zealand, successfully established mammal species tend to have had greater introduction effort (i.e. propagule pressure) and have shorter maximum lifespans (Forsyth & Duncan 2001). Larger relative brain size is also correlated with establishment success in introduced mammals (Sol *et al.* 2008a; Sol *et al.* 2008b). A meta-analysis by Clout and Russell (2008) found three factors consistently associated with increased probability of establishment success for non-native mammal species: a) number of individuals released, b) natural range size and c) climate temperateness in the novel environment.

We attempt to determine which factors predict success at two stages of the invasion process in introduced deer. As a result of their aesthetic and economic appeal as exotic species, deer have been introduced numerous times worldwide. Deer comprise the largest proportion of all invasive mammal species (Clout and Russell 2008). The presence of non-native deer species has brought both benefits and costs to the recipient communities, yet there is increasing concern about their impacts (Gill 1992; Fuller & Gill 2001; Cote *et al.* 2004; Moriarty 2004; Ward 2005; Dolman & Waber 2008; Acevedo *et al.* 2010; Gill & Morgan 2010; Nunez *et al.* 2010). Through grazing and browsing, non-native deer may have a significant economic impact on agriculture and forestry (Ratcliffe 1987; Chadwick *et al.* 1996; Putman & Moore 1998) as well as on the structure and dynamics of natural vegetational systems (Pollard & Cooke 1994; Cooke & Lakhani 1996; Cooke & Farrell 2001; Jaksic *et al.* 2002; Relva *et al.* 2010). They may also compete with native ungulate species (Keiper 1985; Chapman *et al.* 1993; Hemami *et al.* 2004, 2005; Feldhamer & Demarais 2009; Feldhamer & Armstrong 1993) or other native species, and in the extreme may impact upon native deer species through hybridisation (Lowe & Gardiner 1975; Simberloff 1996; Goodman *et al.* 1999; Senn & Pemberton 2009).

Given this potential for non-native deer to have such a significant impact on native systems, better understanding of the factors predicting success at different stages of invasion will help efforts to prevent and manage future invasions. In our analyses we aim to determine: (1) which factors influence the outcome of introductions, that is, lead to established populations; and (2) which factors promote the spread of previously established populations. We expect that propagule pressure, native range size, relative brain mass and introductions within the same biome will be important factors in predicting the establishment of non-native deer. As for spread, we expect that introduction within the same biome and species-level traits associated with population growth rate to predict whether established populations will spread. We compiled a trait dataset for all introduced deer species, assembled a database of deer introductions and their outcomes, and reconstructed a species-level phylogenetic tree of deer for comparative analyses. Since closely related taxa may be more likely to have similar introduction outcomes because of traits shared through common ancestry, a phylogenetic framework is needed to control for evolutionary relatedness in the analyses.

Table 2.1 Event, region and species-level effects hypothesised to influence invasion success in mammals

Category	Variable	Expected correlation with invasion success	References
Event	Number of individuals introduced (propagule size)	Higher (i.e. introduction effort or propagule pressure)	(Forsyth and Duncan 2001; Sol et al. 2008a)
	Number of introductions	Higher (i.e. introduction effort or propagule pressure)	(Forsyth and Duncan 2001; Forsyth et al. 2004; Lockwood et al. 2005; Jeschke and Strayer 2006)
Region	Biome introduced to	Same biome (i.e. climate matching)	(Forsyth and Duncan 2001; Forsyth et al. 2004; Sol et al. 2008a)
	Island or mainland introduction	Islands more likely to be invaded	(Elton 1958; Jeschke 2008; Sol et al. 2008a)
	Biogeographical region	Outside Eurasia (i.e. biotic resistance hypothesis)	(Jeschke and Strayer 2006; Jeschke 2008; Sol et al. 2008b)
Species	Body size (body mass, body length)	Larger	(Ehrlich 1989; Forsyth and Duncan 2001; Forsyth et al. 2004; Jeschke and Strayer 2006; Sol et al. 2008a)
	Relative brain mass	Higher (i.e. brain size environmental change hypothesis)	(Jeschke and Strayer 2006; Sol et al. 2008a)

	Life history (number of litters per year, litter size, neonatal body mass, weaning age, weaning body mass, age at sexual maturity, age at first reproduction, inter-birth interval, gestation length, maximum lifespan)	Higher number of litters per year and litter size, larger neonate mass, shorter maximum lifespan	(Ehrlich 1989; Forsyth and Duncan 2001; Forsyth et al. 2004; Jeschke and Strayer 2006)
	Population features (population density, home range, social group size)	Higher population density; larger home range; gregarious	(Ehrlich 1989; Jeschke and Strayer 2006)
	Native range size; human population density in native range	Larger native area;	(Ehrlich 1989; Sol et al. 2008a)
	Diet breadth	Higher	(Ehrlich 1989; Forsyth et al. 2004; Jeschke and Strayer 2006)
	Habitat breadth	Higher	(Ehrlich 1989; Sol et al. 2008a)

Materials and Methods

Trait dataset

Our dataset of species attributes includes body mass, body length, neonate body mass, weaning body mass, relative brain mass, diet breadth, habitat breadth, maximum lifespan, age at sexual maturity, age at first reproduction, inter-birth interval, gestation length, weaning age, litter size, litters per year, population density, social group size, home range, home range of an individual, native area, human population density in native area and biogeographical region of origin. Data were extracted from extant databases and the literature, and Simon Reader shared data brain masses. Factors linked to the socioeconomics in the recipient area may also be important, but collecting these data was not possible given the nature and scope of our historical dataset. Predation effects are also not considered here due to lack of available data. Details and references are provided in Appendix 2.1. Variables were scaled in order to facilitate comparison, with time scaled to days, lengths to millimetres and areas to square kilometres. We calculated the relative brain mass as the average brain mass-to-body mass ratio for each species. Where the values of a variable were non-normally distributed, they were $\log_{10}$-transformed.

Introduction dataset

We compiled a dataset of all recorded introduction events of deer outside their native ranges (i.e. excluding local translocations). Sven Bacher shared data on mammal introductions (details provided in Appendix 2.2). Data on the numbers of individuals introduced (propagule size), the biome introduced to, whether the introduction was to an island or mainland, and the outcome of the introduction were extracted from a global mammal introduction database (Sol *et al.* 2008a), as well as relevant literature (Chapman *et al.* 1994; Danilkin 1996; Cooke & Farrell 1998; Forsyth & Duncan 2001; Forsyth *et al.* 2004). The resultant dataset was cross-checked with records in Lever (1985) and Long (2003), and the primary literature therein. All introduction events were then scored as either failed (population died

out, but not from eradication by hunting; scored as 0) or successfully established (if a self-sustaining wild population was derived from the initial introduction; scored as 1). Then, each event was further scored as successfully spread where self-sustaining populations had established beyond the area of initial introduction. Establishment and spread success rates of introduced species are often overestimated (Rodriguez-Cabal *et al.* 2009). However, for families of large mammals such as the deer, the under-reporting of introductions failing to establish or spread is likely to be less severe. We took a conservative approach and deleted all entries where the outcome of the introduction was uncertain or where the number of introduced individuals was unavailable. Multiple introductions in the same area at the same time were combined and counted as a single introduction event.

Phylogenetic Analysis

While phylogenetic estimates of certain groups within Cervidae already exist (Randi *et al.* 1998; Pitra *et al.* 2004; Gilbert *et al.* 2006), a comprehensive tree incorporating all existing DNA data for deer species was lacking. We aimed to increase the numbers of taxa sampled in order to: 1) obtain a reliable topology and branch length estimates for phylogenetic correlation tests, 2) gain a better understanding of the relationships within three unresolved genera (*Cervus*, *Muntiacus* and *Mazama*), and 3) determine the relationships of two of the rarest species (*Axis kuhlii* and *Cervusalfredi*) to the rest of Cervidae. Four ruminant species were used as outgroups: *Antilocapraamericana*, *Gazellagranti*, *Moschusmoschiferus*, and *Tragelaphusimberbis*.

Appendix 2.3 provides the DNA accession numbers of all DNA data. We also obtained faecal samples for *A. kuhlii* and *C. alfredi*from Edinburgh and Chester Zoos, respectively. For these two samples, DNA was extracted following Ball *et al.*(2007), and ~400bp of the cytochrome b gene (*cytb*) was sequenced using standard protocols and primers from Hassanin et al. (1998). Complementary strands were sequenced on an ABI 3130*xl* automated DNA sequencer (Applied Biosystems). For the other taxa, a total of six markers were selected for analysis. Four mitochondrial regions: *cytb*, cytochrome

Oxidase II (COII), 12S and 16S rDNA, and two nuclear regions: intron 2 of α-lactalbumin (αLAlb) and intron 1 plus two exonic regions of protein kinase C iota (PRKCι). Sequences were aligned with MacClade and MAFFT (Katoh *et al.* 2002; Katoh *et al.* 2005); the alignment is available from TreeBase accession 10993. In total, our matrix includes 43 species (out of 51 described deer species) and is 5,290 base-pairs long.

MrMODELTEST v2.3 was used to infer the best-fit models of molecular evolution for each gene separately, and for the combined dataset. The models selected by hierarchical log-likelihood ratio tests were GTR+I+G for the mitochondrial DNAs, HKY+G for the nuclear regions, and GTR+I+G for the combined dataset. Maximum Parsimony (MP) and Bayesian analyses using MrBayes v3.1.2 (Ronquist & Huelsenbeck 2003) were conducted on the mitochondrial and nuclear datasets, concatenated separately and combined altogether. For the Bayesian search, we used two million generations and discarded 25% of trees as burn-in.

Statistical analyses

All statistical analyses were conducted in R v2.10.1 (R Development Core Team 2008). Initially, Generalised Linear Models (GLM) were fitted to the data to determine which factors explain establishment and spread success (Appendices 2.4-2.7). However, the outcomes of introduction events are likely to be correlated due to shared evolutionary ancestry of the introduced species, and because species and populations were not randomly introduced to different locations. Therefore, we used two modelling approaches to take into account for the non-independence of these data: Generalised Linear Mixed Modelling and Generalised Estimator Equations (GEE). Explanatory variables were entered into the model as linear terms. Intercept terms were included in the models. Generalised Linear Mixed Models (GLMMs) were fitted to determine factors that predict the establishment and spread of introduced populations, and GEE were used identify factors that explain variation in the establishment and spread success of introduced species. Chi-square tests on change in deviance were used to assess

goodness of fit for GLM. We took a frequentist hypothesis testing approach to inference from the GLMMs and GEE, comparing Wald test statistics to their expected distributions under the null hypothesis and estimating a p value to determine whether to reject the null hypothesis (Bolker *et al.* 2009). For the GLMMs, Z values were compared; for GEE, t values were compared. Plots of predicted probability of establishment and spread were generated for all variables significantly associated with those outcomes (Appendix 2.8).

GLMM

GLMMs were employed to fit hierarchical taxonomic levels (genus and species; Appendix 2.3) and regions of introduction (Appendix 2.2) as random effects (Blackburn & Duncan 2001; Sol *et al.* 2008a). The outcomes of introduction events at two stages of the invasion process (establishment and spread) were the binary response variables, and the remaining characteristics hypothesised to influence success were the explanatory variables (Appendix 2.1). The models were fitted with a binomial error distribution and the logit-link function. Minimal adequate models were fitted using forward selection then dropping and adding terms. Backward selection from a model containing all explanatory variables was not possible due to the large number of explanatory variables (causing the initial full model to fail to converge). The analysis was repeated excluding the three events where over 200 individuals were introduced simultaneously (see Appendix 2.5). Significance levels were assessed using Wald test statistics (Z values).

GEE

GEE are extensions of GLM, which can be used to model correlated observations using a variance-covariance matrix (Paradis & Claude 2002). In our case, we needed to control for the confounding effect of common ancestry on our tests of traits promoting invasiveness: i.e. we may expect closely-related species to have similar invasion success due to traits shared via common ancestry. The variance-covariance matrix was computed from branch lengths in the phylogeny (the smaller the sum of branch lengths, the closer the species are related), with the Analyses of Phylogenetics and Evolution

(APE) (Paradis *et al.* 2004) package in R. Establishment and spread success was calculated as the proportion of times the outcome was positive. The response variables were arcsine-transformed, and models were fitted with a normal distribution. The COMPAR.GEE function in APE was used to fit models incorporating the phylogenetic data.

Results

Deer phylogeny

Although individual genes resolved different topologies, there were no hard incongruences between the mitochondrial and nuclear datasets; therefore we present here phylogenetic relationships depicted from all data combined in the Bayesian search. The phylogeny is presented in Figure 2.1. Cervidae are divided into two subfamilies: Cervinae (Posterior Probability [PP] =0.99) and Capreolinae*sensu*Pocock (1910) (PP=0.99). The relationships within these subfamilies are congruent with previous studies with smaller sampling (Pitra *et al.* 2004; Gilbert *et al.* 2006). The inclusion of four species of *Mazama* that had not been included in a phylogenetic analysis before shows that this genus is polyphyletic. *Cervus* and *Hipocamelus* are also polyphyletic. *Muntiacus* is monophyletic if including *Megamuntiacus*; this latter genus is therefore unjustified. *Axis kuhlii* is found sister to *Axis porcinus*(PP=1). *Cervusalfredi* is found sister to *Cervussensu* Gilbert *et al.* (2006) (PP=0.73).

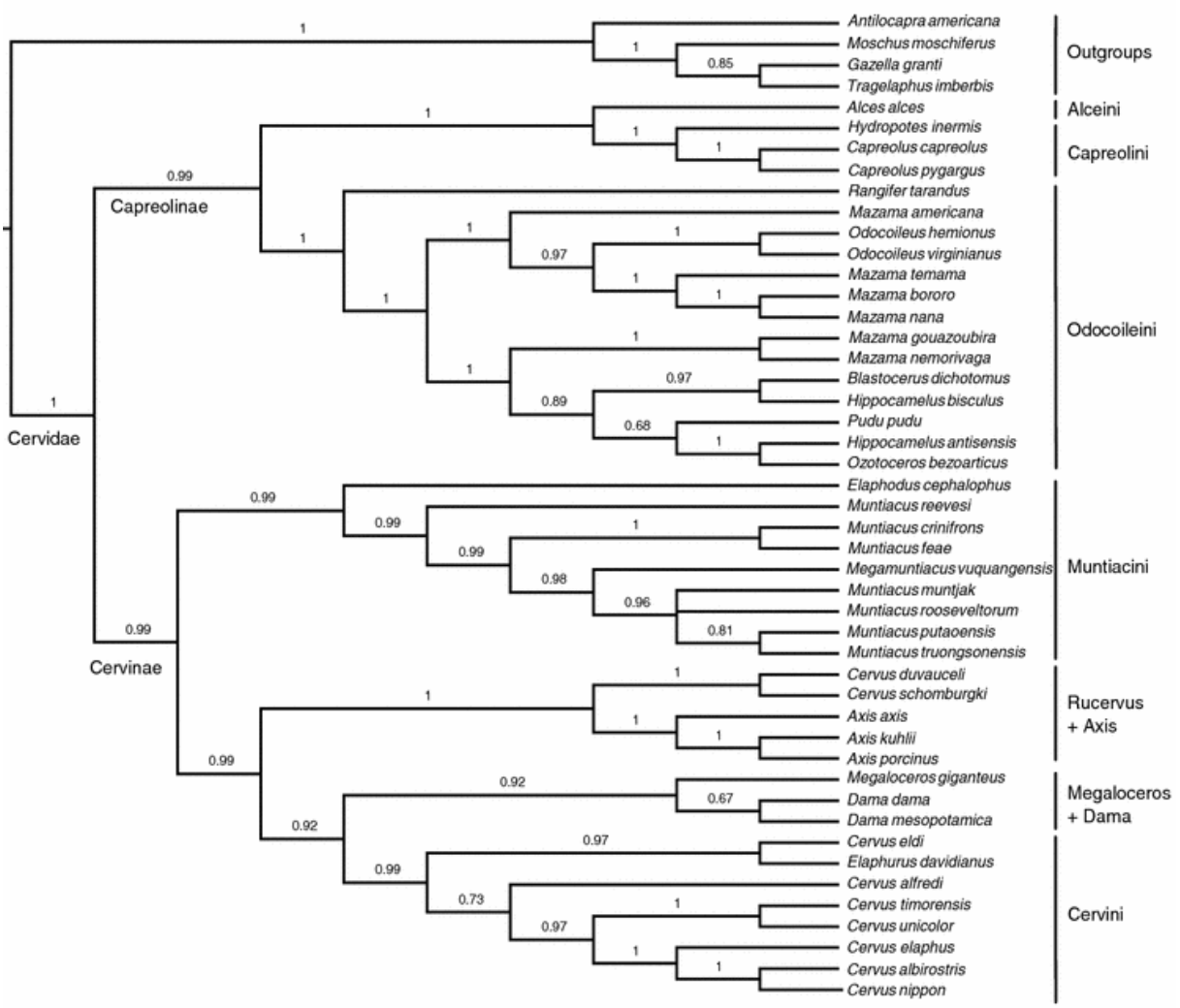


Figure 2.1 Bayesian phylogenetic analysis of deer using six molecular markers.

Values above branches are posterior probabilities.

Establishment

In Appendix 2.2, we document 146 introductions encompassing 4266 individuals from 16 species and at least nine genera. Of these, 100 resulted in successful establishment. Greater propagule size is positively correlated with the establishment success of introduced populations (GLMM $Z = 3.97$, $p < 0.001$; [$Z = 4.33$, $p < 0.001$ excluding three outliers]; Table 2.2 and Appendices 2.4 and 2.8). This association holds with or without controlling for the non-independence of introduction events (Appendix 2.4). At the species level, lower age at sexual maturity is a significant predictor of establishment (GEE $t = -4.25$, $p < 0.01$; Table 2.2 and Appendices 2.5 and 2.8). In addition, lower weaning age (GEE $t = -3.99$, $p < 0.01$) and larger native area (GEE $t = 3.03$, $p < 0.05$) become significant predictive factors of establishment success once phylogenetic effects are taken into account (Table 2.2 and Appendices 2.5 and 2.8).

Spread

Of the 100 successfully established populations, 58 of them spread (Appendix 2.2). Controlling for the non-independence of introduction events using GLMM, habitat breadth (GLMM $Z = 2.93$ to -2.76, $p < 0.01$) and diet breadth (GLMM $Z = 2.06$ to -2.2, $p < 0.05$) were significant when fitted singly (Table 2.2 and Appendices 2.6 and 2.8). As habitat and diet breadth are categorical variables, the test statistics from the models have a range of values. The following explanatory variables predict spread in the minimal adequate model: diet breadth, litter size and biome match, with no significant interactions (not shown). When using GEE to incorporate phylogenetic information and control for common ancestry, species with greater weaning body mass had significantly higher spread success (GEE $t = 1.57$, $p < 0.01$; Table 2.2 and Appendices 2.7 and 2.8).

Table 2.2 Factors associated with establishment and spread success in introduced deer.

	ESTABLISHMENT				SPREAD			
	Introduction events		Species		Introduction events		Species	
	Estimate (±S.E.)	GLMM Z	Estimate (±S.E.)	GEE t	Estimate (±S.E.)	GLMM Z	Estimate (±S.E.)	GEE t
Propagule size	0.84 (0.21)	3.97***	na	na	0.29 (0.24)	1.24	na	na
Habitat breadth[a]	0.66 to 0.16 (0.47 to 0.3)	1.39 to 0.12	1.59 to -0.7 (0.39 to 0.21)	4.94 to -0.81	1.34 to -1.56 (1.48 to 0.46)	2.93** to -2.76	0.88 to -0.18 (0.40 to 0.22)	0.76 to -0.46
Diet breadth[a]	0.64 to -0.4 (0.59 to 0.41)	1.09 to -0.28	1.69 to -0.68 (0.48 to 0.23)	4.83 to -2.87	1.4 to -1.76 (0.8 to 0.68)	2.06* to -2.2	0.94 to -0.66 (0.42 to 0.2)	0.54 to -1.55
Weaning body mass	-0.61 (0.53)	-1.15	-0.39 (0.05)	6.93	0.64 (0.68)	0.92	0.33 (0.21)	1.57**
Weaning age	0.00 (0.00)	0.84	-0.04 (0.01)	-3.99**	0.00 (0.00)	-1.922	-0.00 (0.00)	-0.63
Age at sexual maturity	-0.47 (-0.89)	-0.52	-0.75 (0.17)	-4.25**	-0.252 (1.05)	-0.24	0.378 (0.328)	1.15
Native area	0.00 (0.00)	0.33	3.17 (1.04)	3.03*	0.00 (0.00)	0.447	-1.75 (1.45)	-1.21

Outcomes of introduction were analysed at the population (introduction event) and species levels. Estimates and test statistics are presented for Generalised Linear Mixed Models (GLMMs) and Generalised Estimator Equations (GEE) fitted with outcome as the response variable and the hypothesised predictors of success as explanatory variables. [a]As habitat and breadth diet breadth are categorical variables, we present the maximum and minimum values. Significance codes: $p < 0.001$ = '***'; $p < 0.01$ = '**'; $p < 0.05$ = '*', S.E.= standard error.

Discussion

Significant progress has been made in determining factors predicting successful invasions by quantitatively examining successful and failed introductions of different species (Sol *et al.* 2008b). Studies conducted across broad taxonomic groups attempt to find factors suitable for inclusion in risk assessment models (Bomford *et al.* 2009). However, better understanding is likely to come from examining specific groups of taxa across a broad range of hypotheses. Here we analyse data on the outcomes of introduction events of non-native deer at two stages of the invasion process in order to determine which characteristics are predictive of success. We obtain differing results from analyses at the species and population levels, and show that traits predicting success at one stage of invasion do not necessarily predict success at other stages.

Which characteristics predict the establishment of exotic deer? Deer species natively distributed over larger areas had a higher establishment success rate, a finding which agrees with that of exotic ungulates introduced to New Zealand (Forsyth & Duncan 2001) and mammals introduced in Europe and North America (Jeschke & Strayer 2006). We also found that the life history characteristics of lower weaning age and age at sexual maturity are also associated with greater establishment success at the species level. While this appears to contradict results of other studies in mammals (Forsyth *et al.* 2004; Jeschke & Strayer 2006; Sol *et al.* 2008a), these latter studies were conducted over a broader taxonomic range than our analysis on deer, which may explain these discrepancies. For example, factors affecting establishment and spread in deer may not be the same as for carnivores, and thus may explain instances where our results do not reflect findings from cross-family analyses of non-native mammal introductions.

Below the level of the species, in assessment of the success or failure of an introduction in establishment of a population of that species, our analysis showed that only propagule size was associated with establishment success. This is in keeping with theoretical predictions (Lande 1993) and previous species-level comparative analyses of introduced vertebrates (Lockwood *et al.* 2005), birds (Blackburn & Duncan 2001; Duncan 2001; Sol *et al.* 2005) and mammals (Forsyth & Duncan 2001; Kolar & Lodge 2001; Forsyth *et al.* 2004; Sol *et al.* 2008a). A larger founding population is

likely to be more resilient to extinction, and should contain higher levels of genetic diversity. Higher diversity, in turn, is likely to confer greater adaptive potential and reduce the deleterious effects of inbreeding depression (Reed & Frankham 2003; Theodorou & Couvet 2006).

What about spread? Only higher weaning body mass out of all the species traits examined in the GEE analysis was significantly associated with spread success. In our population analysis, we found greater habitat and diet breadth and introduction within the same biome negatively correlated with spread success. The island invasibility hypothesis (Elton 1958) is not supported here. A potential explanation for this finding is that many of the islands where deer are introduced have high rainfall and therefore nutritionally poor foliage for herbivores e.g. New Zealand (Forsyth *et al.* 2002). This 'neutral resistance' may explain why we do not find any significant differences between the outcomes of deer introductions to island or continental habitats, as many islands may be just as tough for non-native herbivores to invade.

Are there any deer species that are unable to invade outside their native range? Returning to the historical records (Appendix 2.2), there are two species that have failed to establish, however the very small number of introductions makes it difficult to draw any firm conclusions here. The Barasingha (*Cervusduvaucelli*) has failed to establish wild populations from two introductions to Australia. The first was by the Victorian Acclimatisation Society in Gippsland and the second in Port Essington (Long 2003). Assuming that they were not exterminated by humans, neither introduction resulted in a self-sustaining population becoming established in the wild. The Indian muntjac (*Muntiacusmuntjak*) also failed to establish after introduction to England, but this was due to them being deliberately shot out (Long 2003). The small sample size and lack of data on propagule size (three introductions; one with data on numbers introduced) make these records of limited utility. If there were numerous introductions of a species that failed everywhere, this would provide firmer evidence upon which to state that a species is intrinsically non-invasive.

Comparative analyses of detailed species introduction records can predict which factors determine success at different stages of invasion, and generate useful results for quantitatively assessing the risk posed by past and present introductions. An ideal experiment would require random introductions of

species and populations to all climates, across islands and mainland habitats. This would eliminate biases towards successful outcomes in the data, and enable a greater level of confidence in stating that a particular characteristic is genuinely associated with an outcome. The data presented here are sourced from historical records of introduction events, and as such there are many external factors that may influence the outcome of the introductions. For example, successful introductions are likely to reflect the expertise of the people involved in the translocation and initial release (e.g. Beck et al. 2007). This may bias the number of recorded successful establishments at the expense of failed introduction or establishment. The age and sex composition of the introduced propagule may also explain some of the variation in outcomes. Lack of data on these factors precludes the inclusion of them in our analysis. Additionally, socioeconomics and predation effects in the recipient area may also be important, but collecting information on these factors was not possible given the nature and scope of our historical dataset.

This analysis examines many hypothesised determinants of success at different stages of the invasion process in a mammalian family with great economic importance and a key role in ecosystem functioning. A human-influenced characteristic of the invasion process, propagule pressure, predicts largely the successful establishment of a population. In contrast, characteristics of the region of introduction and of the introduced species are key to predicting spread in non-native deer populations. However, at the species level a larger native range size and traits associated with population growth rate predict establishment, and a life history characteristic predicts spread. These findings have applicability to the management of currently established deer populations, the control of established populations, and policy on the introduction of non-native deer species.

Chapter 3:The influence of range expansion on genetic variation and a test of the central-marginal hypothesis in British roe deer

Introduction

The distribution and demographics of species have a large influence on levels of genetic diversity and population structure (Frankham *et al.* 2002). Numerous species have undergone reductions in population size, and then recovered through demographic and range expansion. Geographical isolation and serial founder events during these processes can result in decreased genetic diversity in the new range (Hewitt 1996; Hewitt 2000). The genetic implications of demographic expansions are well studied (Slatkin & Hudson 1991; Beaumont 1999; Zhivotovsky *et al.* 2000), and recently attention has focused on the influence of range expansion on the distribution of genetic variation (Vucetich & Waite 2003; Wegmann *et al.* 2006; Excoffier *et al.* 2009a; Slatkin & Excoffier 2012).

Population abundance, individual density and reproductive output are expected to be highest in core populations, decreasing at the range limits (Sagarin *et al.* 2006), which influences population genetic parameters (Vucetich & Waite 2003; Eckert *et al.* 2008). Populations derived from the same source may become genetically differentiated due to drift and founder effects, with the strength of differentiation varying with migration rates (Austerlitz *et al.* 1997; Excoffier 2004; Arenas *et al.* 2012) and environmental heterogeneity (Wegmann *et al.* 2006). Where range expansion is driven by a small number of dispersing individuals, genetic diversity is expected to be lost due to passing through bottlenecks and founder effects (Broders *et al.* 1999; DeYoung *et al.* 2003; Excoffier *et al.* 2009a).

There is broad support for expecting lower genetic diversity and increased genetic differentiation of populations near range limits (Eckert *et al.* 2008). However, as ranges expand genetic diversity can increase due to admixture if genetically diverged populations merge (Zenger *et al.* 2003; Kolbe *et al.* 2004; Kolbe *et al.* 2008; Sakaguchi *et al.* 2011).

The roe deer of Great Britain, with their complex history of extirpation, reintroduction from multiple sources and rapid recent range expansion make a good study system with which to examine the influence of contemporary processes on the distribution of genetic variation. There was a severe reduction in distribution and abundance of roe deer in Britain up to the eighteenth century, with the majority of populations in England and Southern Scotland believed to have gone extinct (Whitehead 1964). Refuge populations survived mainly in woodlands in Central and North-western Scotland, and potentially in Northern England (Whitehead 1964; Prior 1995). All Southern English populations are thought to be descended from later reintroductions (Whitehead 1964; Hewison 1995, 1997; Ward 2005; Baker & Hoelzel 2013). The reintroductions to England occurred from Austria to Windermere in North-western England, from Germany to East Anglia, and from Scotland and unknown sources to Dorset and West Sussex in Southern England (Whitehead 1964; Prior 1995; Baker & Hoelzel 2013).

Roe are currently widespread throughout mainland Britain (Figure 3.1), with the widest distribution of all deer in Great Britain, and their numbers are estimated to stand in the hundreds of thousands (Hewison & Staines 2008). Roe deer have a high capacity for demographic and range expansion (Andersen & Linnell 2000), and are now absent only from areas in Kent, the Midlands and Wales (Ward 2005) and most Scottish islands. Damage caused by roe deer has been recorded in commercial forestry, conservation woodlands, fruit and orchard trees and other horticultural crops, and it is anticipated that more damage will occur as the populations continue to expand (Putman & Moore 1998).Furthermore, roe deer account for approximately 29% of deer-vehicle collisions with reported species detail in England (Langbein 2011). The impact of range expansions on population genetic parameters is important as information on population genetic parameters can contribute to conservation and management plans (King & Burke 1999; Nussey *et al.* 2006), and knowledge of genetic structuring can be employed in the definition of management areas or units (Zannese *et al.* 2006) for roe deer.

Here, I investigate the influence of range expansion on genetic diversity and population structure in several independent groups of reintroduced and remnant native roe deer populations, and use the results to test the predictions of the central-marginal hypothesis (Eckert *et al.* 2008).

Firstly, I examine how the average diversity of populations in a region varies with the overall recent range expansion rate in that region. I expect that in regions where range expansion is higher, overall levels of genetic diversity will be lower than in regions where range expansion has occurred more slowly. However, where expanding populations from multiple introductions in Southern England have overlapped, admixture may have resulted in an increase in overall diversity. Secondly, I test whether genetic diversity of peripheral populations correlates with their distance from the core population and rate of range expansion from it. I expect that diversity will decline with distance from the core, and that populations which have spread faster will have lower diversity.Thirdly, I investigate the influence that population history and range expansion has on population genetic differentiation. I test for Isolation-By-Distance (IBD) at different spatial scales, contrasting the patterns between regions of natural expansion following contraction, and regions where reintroductions occurred. I also examine the relationship between the rate of range expansion of peripheral populations with their pairwise genetic distance from core populations. Where populations are geographically isolated, differentiation between them is expected to arise under the classical model of IBD. However, in the case of multiple introduction events expanding at different rates into overlapping ranges in heterogeneous environments, patterns of IBD are expected to be disrupted (Wegmann *et al.* 2006; Excoffier *et al.* 2009a) and the rate of expansion is expected to influence levels of differentiation.

Alongside these analyses, I test the predictions made by the central-marginal hypothesis (Eckert *et al.* 2008) of reduced genetic diversity and increased genetic differentiation of peripheral populations when compared with populations at the core of the range. I expect that there will be distinct structuring between peripheral populations that have expanded into new areas. Finally, I examine the underlying genetic structure of the populations using individual-based cluster analysis.

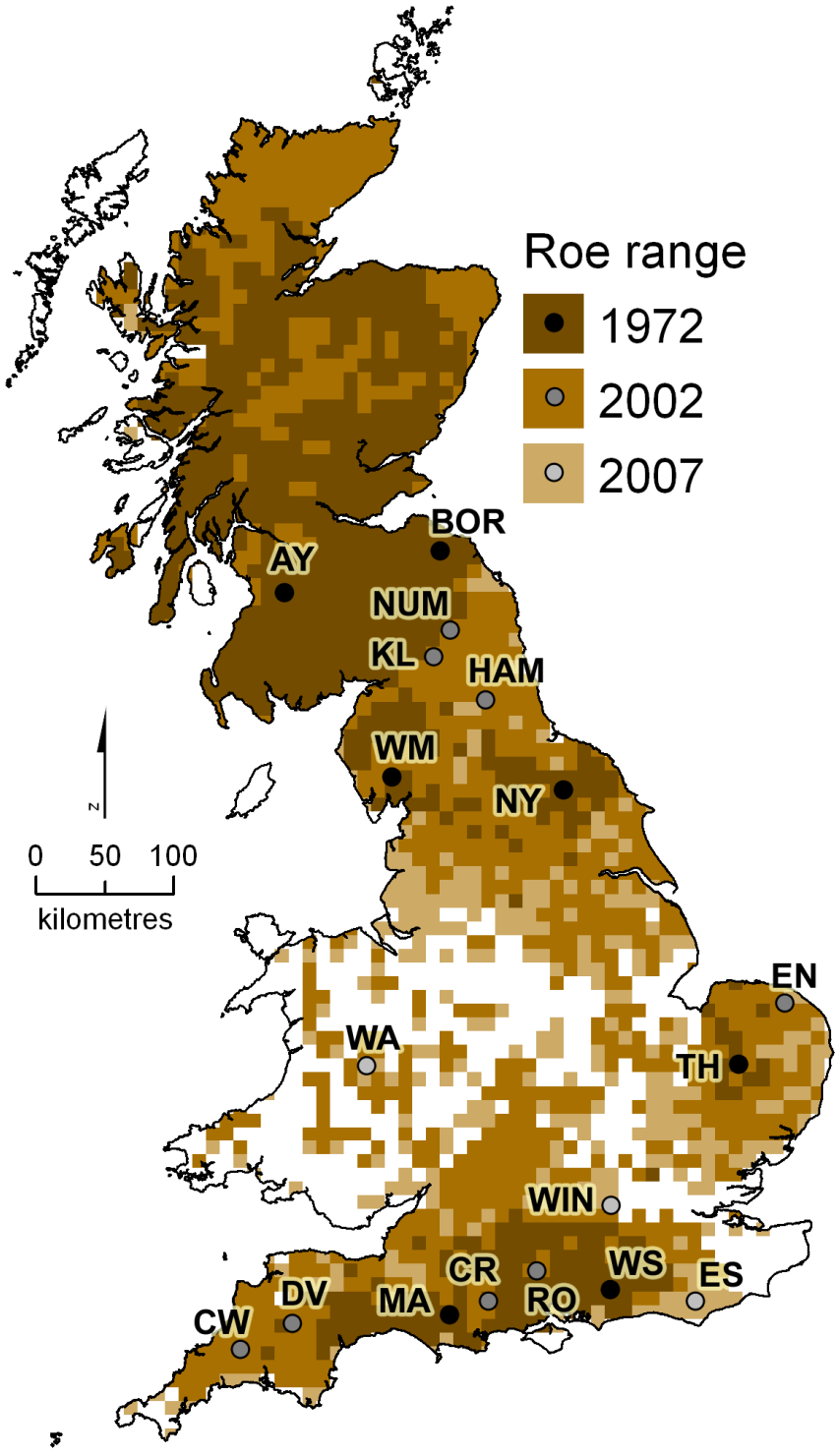


Figure 3.1 Roe deer distribution in Great Britain and sample sites. For site codes see Table 3.1.

Black circles = core populations, grey circles = peripheral populations.

Materials and Methods

Range expansion calculations

At the regional level, the rate of range expansion was defined as the Compound Annual Growth Rate (CAGR) in 10 km^2 cell occupancy between 1972 and 2007. CAGR was calculated from the raw data used to generate the map in Figure 3.1 (A. Ward pers. comm.) for regions where spread has occurred from reintroduction sites in England and south from Scotland. As the two range expansions in Southern England have overlapped, the cell counts were pooled for the CAGR calculation. Factors affecting the accuracy and precision of the data used to produce Figure 3.1 are described by Ward et al. (2008). Broadly, it is not possible to distinguish the difference between genuinely new sightings in 10 km^2 cells in some areas and improvements in coverage or better reporting from survey respondents. This may explain the patchy distribution in some areas, and means that measures of range expansion should be interpreted with some caution. Additionally, data on deer presence in some areas has only recently started being recorded, and therefore there are no data available for the distribution in 1972 to enable comparison with other regions. For these reasons, analysis of the CAGR of the Welsh region is not presented here. For the population level analyses, a population was coded as 'core' if it was recorded in the 1972 distribution data and 'peripheral' if it was recorded after 1972 (Figure 3.1, Table 3.1). The rate of range expansion of was defined as the distance of a population from the core divided by years since the core population was founded. It was calculated using data from geographic distance matrices and historical records on roe deer introductions (Whitehead 1964).

Sample collection and georeferencing

A total of 350 roe deer samples were collected from 18 sites in Great Britain (Figure 3.1). Sites were selected to represent a diverse range of colonisation histories, geographic locations, and to cover the core and periphery of each range expansion. The vast majority of deer sampled (n=348) were shot in season as part of management programmes, with the remaining two samples taken from road-kill. Tongue tissue samples were preserved in 20% w/v DMSO in ddH2O saturated with NaCl, and stored

at -20°C before extraction. Six-figure grid references were recorded for the location of each individual, and converted to latitude/longitude using the UK Grid Reference Finder Batch Convert Tool (http://gridreferencefinder.com/batchConvert/batchConvert.htm).

Table 3.1 Sampled regions and sites of roe deer in Great Britain. Latitude and longitude co-ordinates are midpoints of the sampled sites, C = Core, P = Peripheral locations, n = number of AFLP genotypes individuals, PPL = percentage of polymorphic loci at the 5% level, H_j = expected heterozygosity, H_w = average gene diversity S.E. = standard error.

Region	Site name	Site code	Latitiude	Longitude	Type	n	PPL	Hj	S.E. (Hj)	Hw	S.E. (Hw)
Scotland	Ayrshire	AY	55.2105	-4.4816	C	24	46.2	0.211	0.014		
Scotland	Borders	BOR	55.8118	-2.3487	C	8	72.5	0.166	0.012		
Northern England	Northumberland	NUM	55.6224	-2.2889	P	25	28.6	0.131	0.012	*0.242*	*0.012*
Northern England	Hamsterley	HAM	54.8924	-1.9127	P	22	29.7	0.142	0.011		
Northern England	Kielder	KL	55.4154	-2.1200	P	25	37.9	0.159	0.011		
Northern England	North Yorkshire	NY	54.2935	-0.6449	C	28	71.4	0.251	0.013		
Northern England	Windermere	WM	54.2670	-2.9085	C	6	13.2	0.099	0.010		
East Anglia	Thetford	TH	52.5152	0.6852	C	24	37.4	0.157	0.013	*0.167*	*0.013*
East Anglia	East Norfolk	EN	52.8589	1.3079	P	4	34.7	0.129	0.012		
Southeast England	West Sussex	WS	50.9226	-0.6910	C	19	32.4	0.180	0.014	*0.262*	*0.014*
Southeast England	East Sussex	ES	51.0654	-0.2646	P	18	30.2	0.153	0.013		
Southeast England	Windsor	WIN	51.4702	-0.6916	P	4	58.8	0.180	0.012		
Southeast England	Romsey	RO	50.9807	-1.4973	P	9	18.1	0.126	0.012		
Southwest England	Milton Abbas	MA	50.8382	-2.2957	C	14	34.6	0.176	0.013		
Southwest England	Cranborne	CR	50.9193	-1.9217	P	26	73.1	0.270	0.014		
Southwest England	Devon	DV	50.8594	-3.2503	P	29	24.2	0.111	0.012		
Southwest England	Cornwall	CW	50.8737	-4.4844	P	5	7.1	0.065	0.009		
Wales	Powys	WA	52.2956	-3.1422	P	11	79.1	0.266	0.014		

DNA extraction

Total genomic DNA was extracted from tissue samples using DNeasy Blood & Tissue Kits (Qiagen, UK) following the manufacturer's instructions. DNA was extracted from approximately half of the samples using the 96 Blood&Tissue kit with 96-well plates; the remainder of DNA extractions were conducted with Blood & Tissue Kits in batches of 24 individual extraction tubes. Some samples were coated with hair or soil, so were wiped clean with ethanol before the digestion step. For each sample ~25mg of tissue was cut from the tongue and diced using sterilised forceps and a fresh scalpel blade. Each batch of extractions included one negative to check for contamination between wells/tubes. Subsets of tissue samples from each population were extracted twice independently, in order to allow estimation of genotyping error rates (Bonin *et al.* 2004). 1µl of each DNA extract was added to 4µl loading buffer and electrophoresed on 1.5% agarose gels in 1X TAE buffer for 90 min at ~120V to examine the extract for degradation. Substantially degraded DNA extracts were excluded from further analysis. DNA concentration was then measured using a Nanodrop 2000 (Thermo Scientific).

Genotyping

Amplified Fragment Length Polymorphism (AFLP) profiles were obtained from the samples using a modified version of Protocol 2 described by Papa et al. (2005). The AFLP technique comprises restriction enzyme digestion of total genomic DNA followed by ligation of adaptors, then two rounds of selective PCRs which generate a set of polymorphic fragments (Vos *et al.* 1995). Potential combinations of selective amplification primers were trialled on a subset of samples representing most populations, in order to test the genotyping quality and select the best four with which to genotype the rest of the samples. Primers were selected on the basis of maximising the number of polymorphic peaks per sample, evenness of peak size distribution in the 50-500bp range, and repeatability. 30 combinations of selective amplification primers were initially tested, and four were chosen for

subsequent analyses. Reactions were carried out as follows: ~500 ng of DNA was dried at 60°C in a vacuum oven then resuspended in 5.5 µl nuclease-free water (Sigma-Aldrich). Initial restriction enzyme digestion using 5 U *Taq*I (Thermo Scientific) per sample was carried out at 65°C in a buffer solution (10 mMTris-HCl [pH 8.0 at 37°C], 5 mM $MgCl_2$, 100 mMNaCl, 0.1 mg/mL Bovine Serum Albumin [BSA]) for 90 min. Subsequent digestion with 10 U *Eco*RI (Applied Biosystems) was carried out at 37°C in the buffer solution described above for 2 h.Next, the *Taq*I adaptors (Ajmone-Marsan *et al.* 1997) were prepared in 100 µl batches at 50 pmol/µl by heating an equimolar solution of the two oligonucleotides *Taq*I top strand (5'-GACGATGAGTCCTGAC) and *Taq*1 bottom strand (5'-CGGTCAGGACTCAT) to 95°C for 3 min and cooling slowly to room temperature. *Eco*RI adaptors came from the AFLP Regular Genome Plant Mapping Kit (Applied Biosystems). To ligate adaptors, a solution containing 5 pmol*Eco*RI adaptors and 50 pmol*Taq1* adaptors, 1 U T4 DNA ligase (Promega), 1X T4 DNA Ligase buffer (Promega), 50 ng/µl BSA and 1mM Adenosine 5′-triphosphate (ATP) (Sigma-Aldrich) was added to the digested DNA solution, mixed gently and the combined solution incubated at 16°C for 17 h. The resulting template mixture was diluted fourfold in nuclease-free water (Sigma-Aldrich). All pre-selective and selective amplifications were carried out with Fermentas (Termo Scientific) PCR mastermix, using a Veriti® 96-Well Fast Thermal Cycler (Applied Biosystems).A random subsample of the template fragments was obtained through pre-selective amplification using the primers E01 (5'-GAC TGC GTA CCA ATT CA) and T01 (5′-GAT GAG TCC TGA CCG AA) described by Ajmone-Marsan et al. (1997). 5 µl of the diluted DNA template mixture was added to 75 ng of each primer in a 20 µlPCR mix. PCR conditions were as described in Papa et al. (2005). The reaction products were diluted 20 times with nuclease-free water (Sigma-Aldrich). Selective amplifications used 5 µl of the diluted product with 10 – 15 ng of fluorescently labelled *Eco*RI (5′-GAC TGC GTA CCA ATT CNN N_{DYE} [Applied Biosystems]) selective primer and 30 ng of *Taq*I (5′-GAT GAG TCC TGA CCG ANN N) selective primer in a 20 µlPCR mix. Touchdown PCR conditions were as described in Papa et al. (2005). The four *Eco*RIand *Taq*I primer combinations were: *Eco*RI-ACA_{FAM}/*Taq*I-ACT, *Eco*RI-AAG_{JOE}/ *Taq*1-CAT, *Eco*RI-AAC_{NED}/*Taq*I-CCA, *Eco*RI-ACA_{FAM}/*Taq*I-CCA. 1 – 1.2 µl of each selective PCR product from each sample was mixed with 10 µlHi-Di formamide (Applied Biosystems) and 0.2 µl GeneScan-500 ROX size standard (Applied

Biosystems) in a single well for fragment analysis by capillary electrophoresis on an ABI 3130*xl* genetic analyser (Applied Biosystems). Samples with genotype failures were re-run once. They were re-amplified from genomic DNA and the undiluted pre-selective reaction products were electrophoresed on 1.5% agarose gels at 90V for 2 h to check for the presence of amplified fragments before selective amplification and capillary fragment analysis was re-run.

AFLP electropherograms were visualised using GeneMapper v. 4.0 (Applied Biosystems). Profiles with poor PCR amplification (i.e. few peaks or peak strength declining rapidly with fragment size) or poor sizing were removed from the dataset. Bins within the range of 50 to 500 bp were manually identified. All bins had at least one peak of height ≥100 relative frequency units (*rfu*). In order to minimise the effect of size homoplasy(Caballero *et al.* 2008)on the results, bin widths were set at < 1 bp and had to be non-overlapping. Bins in which peaks were > 0.3 bp apart were removed from the analysis, and different bins had to be separated by > 0.4 bp. Once the maximal bin set was created, tables of 'sum of signal' normalised peak heights for each primer pair were exported from Genemapper. AFLPScore v. 1.4b (Whitlock *et al.* 2008) was used to optimise scoring parameters and create a binary genotype table for each primer pair.Locus and phenotype *rfu* selection thresholds were explored for permutations of scoring methods (i.e. filtered loci/absolute thresholds, unfiltered/absolute, filtered/relative and unfiltered/relative) in AFLPScore, with the aim of maximising the number of loci retained and minimising the mismatch error rate. A binary genotype table for each primer pair was then generated under the optimal threshold settings, and the four tables were concatenated. Mean mismatch error rate was 6.2 % (± 0.6 S.D.), which is slightly higher than the normal range of error rates using AFLPs (Bonin *et al.* 2004). Loci with zero allele frequencies, and those with a fragment present or absent in one individual only were then removed, giving a dataset comprising 168 polymorphic loci for 301 individuals. As the aim here was to infer population genetic parameters based on neutral genetic variation the influence of non-neutral loci should be removed (Bonin *et al.* 2006). Thirty-six loci identified as diverging from neutral expectations by DFDIST (Beaumont & Balding 2004; Antao & Beaumont 2011) and SAM (Joost *et al.* 2007; Joost *et al.* 2008) analysis (see Chapter 4 for details) were excluded from the dataset, giving a final dataset of 132 loci.

Population genetic analyses

AFLP-based estimates of genetic diversity were calculated using AFLP-SURV v. 1.0 (Vekemans *et al.* 2002). Allele frequencies were estimated using a Bayesian method with non-uniform prior distribution (Zhivotovsky 1999) assuming Hardy-Weinberg Equilibrium (HWE) (i.e. $F_{IS} = 0$). Estimates of average gene diversity (H_w) were calculated for each region, and genetic diversity within populations was estimated using the percentage of polymorphic loci at the 5% level (PPL) and expected heterozygosity (H_j). To explore the relationship between range expansion and overall within-region diversity, H_w was plotted against CAGR of each region. To examine the influence of range expansion on within-population genetic diversity, univariate linear regressions of H_j for peripheral populations against distance, and rate of range expansion from core populations were conducted in R v2.15.3 and plotted in Excel. The within-population analyses were repeated excluding the Welsh population (WA) from the dataset (as there is little evidence from the distribution map upon which to assign it to a core population), and the Romsey (RO) population was assumed to have descended from the introductions in West Sussex (WS), on the basis that RO was genetically and geographically closer to WS than the Milton Abbas (MA) population. The central-marginal hypothesis was tested by assessing differences in levels of diversity between the core and peripheral populations for the entire dataset and within regions using a Mann–Whitney U-test (Sokal & Rohlf 1995).

Overall genetic differentiation (F_{ST}) was estimated for the entire dataset from allele frequencies using AFLP-SURV following Lynch and Milligan (1994). Significance was assessed using a permutation test with 10,000 pesudoreplicates. Pairwise population differentiation was estimated by Analysis of MOlecularVAriance (AMOVA) (Excoffier *et al.* 1992) using GeneAlEx v. 6.5 (Peakall & Smouse 2012). AMOVA produces estimates of Φ_{PT} (an analogue of F_{ST} which calculates a squared Euclidean distance matrix between AFLP genotypes). Φ_{PT} is a band-based measure of pairwise population genetic differentiation that is recommended for the analysis of AFLP datasets (Bonin *et al.* 2007), and does not rely on assumptions which may lead to underestimation of genetic variability (Excoffier *et*

*al.*1992). Despite the complex population histories, expanding populations may show a classic pattern of IBD. Testing this hypothesis can be informative on how dispersal occurs at a national level and within regions. A strong pattern indicates that most dispersal events involve movement between geographically adjacent sites, while a very weak pattern can indicate either high levels of gene flow acting to reduce differentiation, or that an appreciable proportion of dispersal events are long distance. The hypothesis of IBD was tested among all populations and subsets of populations corresponding to range expansions within regions using matrices of pairwise Φ_{PT} values and pairwise geographical distances. Pairwise geographical distances were calculated from the midpoint of each population on the WGS84 datum surface using a Mercator Projection in the Geographic Distance Matrix Generator (http://biodiversityinformatics.amnh.org/open_source/gdmg). Geographic distances were used unprocessed and log-transformed. The significance of the relationships between genetic and geographic distances was assessed by Mantel tests with 10,000 randomisations using IBDWS v. 3.23 (Jensen *et al.* 2005). To examine the influence of range expansion on between-population genetic differentiation, univariate linear regressions of pairwise Φ_{PT} between peripheral and core populations against rate of range expansion were conducted in R and plotted in Excel. To test the prediction that population differentiation is greater among peripheral populations than among core populations, pairwise Φ_{PT} values were compared using a Mann–Whitney U-test (Sokal & Rohlf 1995).

Individual-level Bayesian clustering analysis was conducted using STRUCTURE v. 2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2007) in order to analyse population structure, and to ascertain whether the sampled individuals group into distinct genetic lineages. The most plausible number of populations (K), and the probability of assignment for each individual to each population (Q) were assessed with and without admixture-based ancestry models, assuming independent and correlated allele frequencies (Falush *et al.* 2003). For each model, seven runs were conducted for each value of K between 2 and 10, each using 100,000 Markov Chain Monte Carlo (MCMC) generations with 50,000 generations discarded as burn-in. The most probable value of K was approximated using two methods. The first determines the highest log-probability of the data Pr(X|K) as described in the STRUCTURE manual (Pritchard *et al.* 2000) using a plot of Pr(X|K) averaged across runs against K. The smallest value of K

where log-probabilities are similar was determined to be the value of K with the highest hierarchical level of population structure in these data. The second approach assessed the rate of change in log-likelihood between successive values of K (i.e. ΔK) (Evanno *et al.* 2005). ΔK = |mL (K + 1) - 2mL (K) + mL (K - 1)| / sdL (K), where mL is the mean and sdL is the standard deviation of the likelihood value (L) across the seven independent runs. The number of populations described by the modal value in the ΔK distribution was determined to be the correct value of K. Results from the runs were summarised and plotted using STRUCTURE HARVESTER Web v. 0.6.93 (Earl &vonHoldt, 2012).

Results

Rates of range expansion

CAGR between 1972 and 2007 was highest in the East Anglian population at 7.25%, and lowest in Southern England at 3.67%. CAGR in Northern England was 5.10%.

Genetic diversity and range expansion at the regional and population levels

Estimates of H_w per region, and H_j and PPL for each population are presented in Table 3.1. The CAGR of roe deer distributions in the English regions and the average gene diversity in those regions are plotted in Figure 3.2. There was no significant relationship between the within-population diversity of all peripheral populations and a) distance from the nearest core population b) rate of range expansion measured in km/yr. However, when the Welsh population was excluded from the analyses, a significant negative relationship was observed between within-population diversity and both measures (Figure 3.3: a) $r^2 = 0.411$, $P < 0.05$, b) $r^2 = 0.33$, $P < 0.05$). Within each region the average diversity of core populations tended to be higher than peripheral populations (Table 3.1), but Mann–Whitney U tests showed no significant differences in diversity between all core and peripheral populations ($U = 26$, $P > 0.1$). When the analysis was repeated excluding the Welsh population and populations in areas of overlap (CR and RO), the difference was marginally significant ($U = 11$, $P < 0.1$).

Isolation by distance at the national and regional levels

The coefficient of genetic differentiation among populations (F_{ST}) estimated by partitioning the total gene diversity assuming HWE was 0.191 (tested using 10,000 permutations, $P < 0.001$), indicating substantial population structure across Great Britain. Testing for patterns of IBD, the correlations between genetic and geographic distances between populations at the national level was not

statistically significant (Figure 3.4). Additionally, no significant patterns of IBD were observed in the groupings of Southern English and East Anglian populations. However, a marginally significant correlation ($r^2 = 0.506$, $P < 0.1$) in the Northern English group of populations was observed (Figure 3.4).

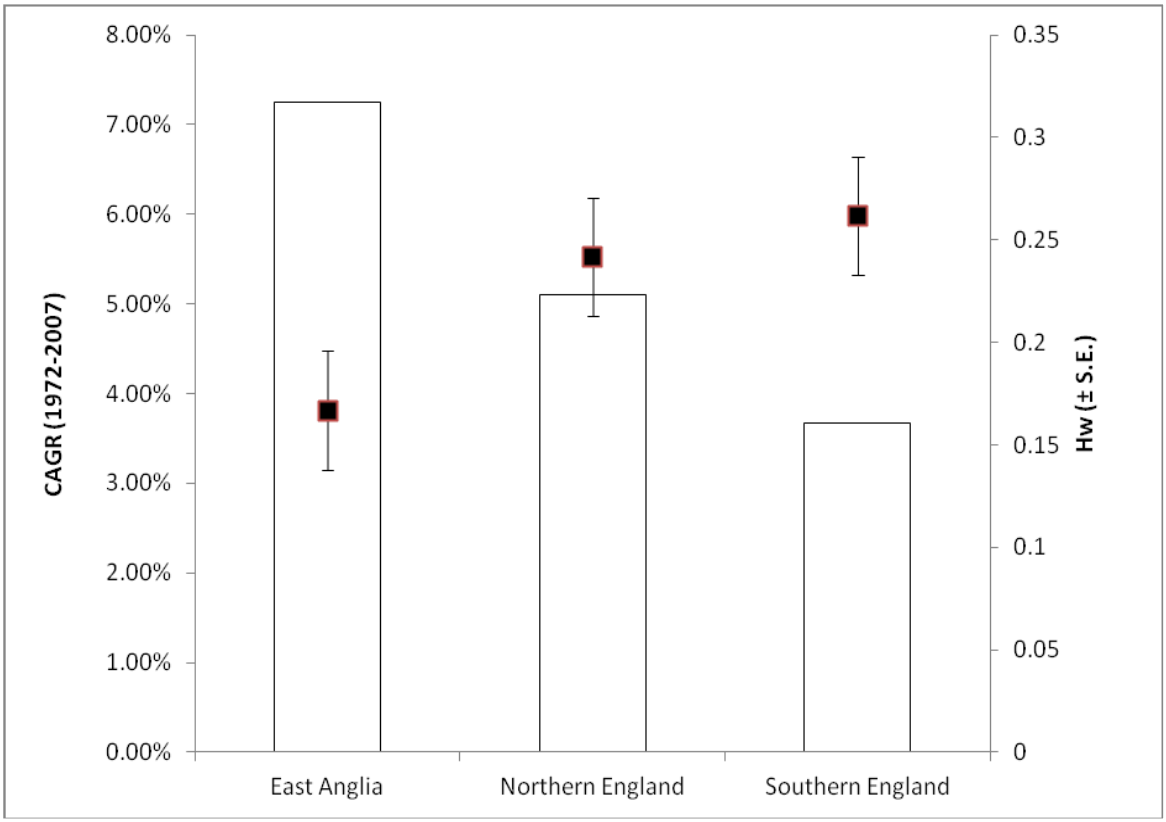


Figure 3.2 Compound Annual Growth Rate (CAGR) and average gene diversity (H_w) in roe deer ranges by region.

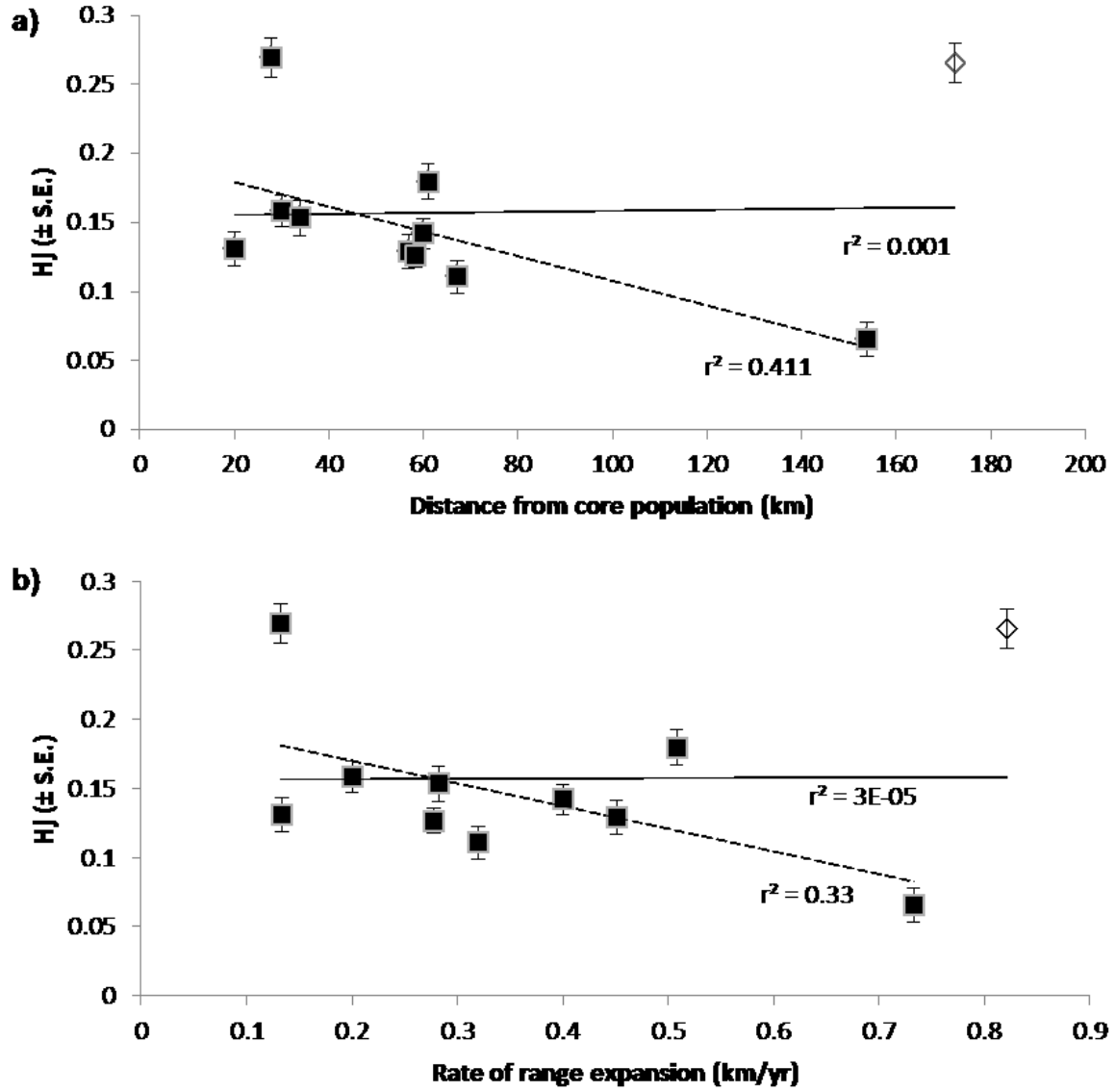


Figure 3.3 Expected heterozygosity (H_j) of peripheral populations regressed against a) distance and b) rate of range expansion from their core population. Regression results are represented by coefficients of determination (r^2). Dashed lines indicate trends when excluding the Welsh population (represented by a white diamond) from the analysis.

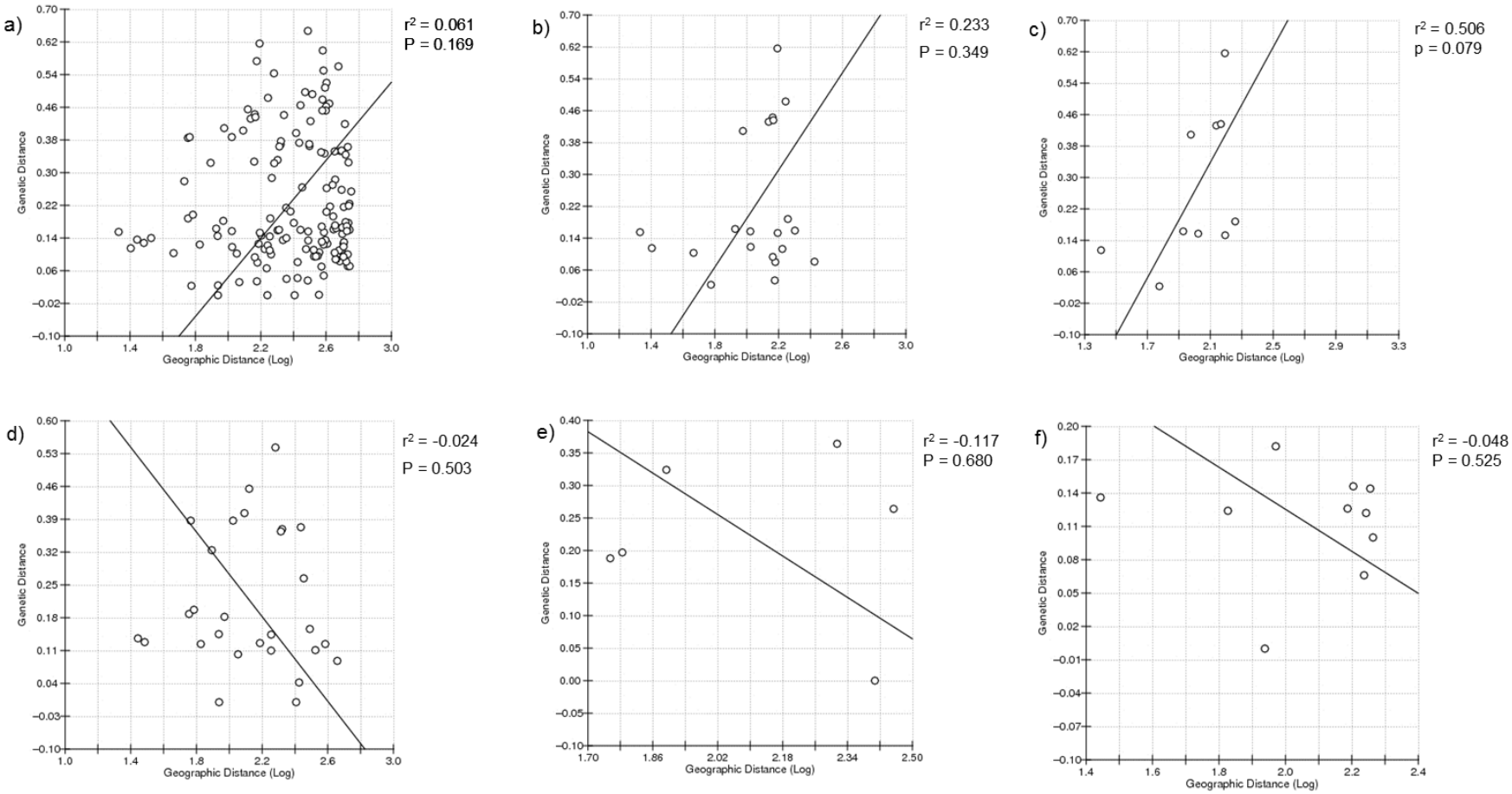


Figure 3.4 IBD relationships between pairwise genetic distance (Φ_{PT}) values and log geographic distance (km). a) Great Britain, b) Southern Scotland and Northern England, c) Northern England, d) Southern England (pooled), e) SE England, f) SW England. Regression results are represented by coefficients of determination (r^2) and their statistical significance according to the Mantel test.

Genetic differentiation and range expansion

Estimates of pairwise genetic distance between populations varied widely, ranging from $\Phi_{PT} = 0$ to 0.647 (Table 3.2). Within regions, the highest between-population Φ_{PT} values were found in the grouping of Northern English and Scottish populations, with the Windermere (WM) population the most differentiated from all other populations in the region. Φ_{PT} values were generally lower between populations in Southern England than between populations in Northern England and Scotland (Table 3.2). High Φ_{PT} values were observed between populations in the known reintroduction sites in the south (e.g. WS-TH: $\Phi_{PT} = 0.331$), suggesting significant historical differentiation and low levels of contemporary gene flow. There was no significant correlation between the rate of expansion of peripheral populations and the pairwise genetic distances from their core population, and the exclusion of Welsh population from the analysis did not alter this finding (Figure 3.5). The prediction of the central-marginal hypothesis that differentiation between pairs of peripheral populations is greater than between peripheral and core populations was not supported in southern populations with ($U = 23$, $P > 0.1$) or without ($U = 10$, $P > 0.1$) the Welsh population included in the analysis.

Genetic structure

The mean posterior probabilities of the STRUCTURE runs suggested an optimal $K = 8$ under both models of admixture and no admixture. Assessing ΔK using the method of Evanno et al. (2005) confirmed that $K = 8$ was the most strongly supported value of K for both models (Figure 3.6). There was a clear division between clusters of individuals sampled from Scotland and Northern England, and those sampled in Southern England in the non-admixture model, while the admixture model highlighted the introduction of Scottish deer into Southern England. Exceptions to this trend are observed where some individuals from the SW English populations (WS and ES) cluster with the Scottish (AY and BOR), and some Northern English deer (from the NY population) cluster with the East Anglian (EN and TH) populations (Figure 3.7).

Table 3.2 Pairwise Φ_{PT} values (below diagonal) and geographic distances in km (above diagonal) between populations. Significant Φ_{PT} values are in bold ($P < 0.01$).

	AY	BOR	CR	CW	DV	EN	ES	HAM	KL	MA	NUM	NY	RO	TH	WA	WIN	WM	WS
AY	-	150	507	483	491	540	460	168	151	508	146	267	511	453	336	486	146	540
BOR	0.034		545	568	555	546	405	106	46.4	554	21.4	201	541	417	395	495	176	555
CR	**0.102**	**0.177**	-	180	93.5	117	310	442	501	27.8	524	385	31	253	175	105	379	86.4
CW	**0.083**	**0.254**	**0.144**	-	86.7	297	455	479	530	154	548	461	210	401	183	273	392	266
DV	**0.109**	**0.160**	**0.182**	-	-	211	384	458	513	67.1	534	421	124	328	160	191	380	180
EN	**0.100**	**0.325**	0.032	**0.497**	**0.377**	-	227	440	500	145	525	360	86.8	174	241	54.0	399	33.9
ES	**0.088**	**0.126**	**0.156**	**0.088**	**0.124**	**0.214**	-	310	362	335	386	205	284	56.8	307	206	320	256
HAM	**0.113**	**0.118**	**0.161**	**0.110**	**0.164**	**0.193**	**0.134**	-	59.7	452	84.7	106	436	315	300	390	94.7	449
KL	**0.080**	**0.103**	**0.150**	**0.082**	**0.130**	**0.166**	**0.105**	0.023	-	510	25.4	157	495	371	354	449	137	509
MA	**0.126**	**0.224**	**0.136**	**0.126**	**0.124**	**0.327**	**0.111**	**0.171**	**0.119**	-	533	401	58.3	278	172	132	384	113
NUM	**0.093**	**0.155**	**0.179**	**0.219**	**0.161**	**0.344**	**0.154**	**0.163**	**0.115**	**0.169**	-	181	519	397	375	474	156	534
NY	**0.081**	**0.159**	**0.048**	**0.166**	**0.217**	0.001	**0.160**	**0.157**	**0.153**	**0.204**	**0.188**	-	373	217	278	314	147	375
RO	**0.216**	**0.363**	**0.128**	**0.369**	**0.403**	0.024	**0.264**	**0.270**	**0.258**	**0.387**	**0.419**	**0.070**	-	228	185	78.3	378	56.9
TH	**0.283**	**0.469**	**0.177**	**0.520**	**0.492**	0.000	**0.385**	**0.365**	**0.350**	**0.465**	**0.452**	**0.135**	**0.140**	-	261	150	308	201
WA	**0.095**	**0.135**	**0.122**	**0.100**	**0.146**	**0.205**	0.036	**0.113**	**0.096**	**0.066**	**0.168**	**0.160**	**0.287**	**0.397**	-	192	220	228
WIN	**0.353**	**0.354**	**0.387**	**0.373**	**0.543**	**0.279**	**0.364**	**0.347**	**0.352**	**0.455**	**0.560**	**0.370**	**0.324**	**0.573**	**0.323**	-	345	61.0
WM	**0.443**	**0.483**	**0.479**	**0.508**	**0.599**	**0.462**	**0.426**	**0.409**	**0.432**	**0.550**	**0.616**	**0.436**	**0.452**	**0.647**	**0.441**	0.095	-	401
WS	**0.071**	0.071	**0.145**	**0.042**	**0.110**	0.140	0.000	**0.104**	**0.094**	**0.102**	**0.157**	**0.131**	**0.188**	**0.331**	0.040	**0.197**	**0.262**	-

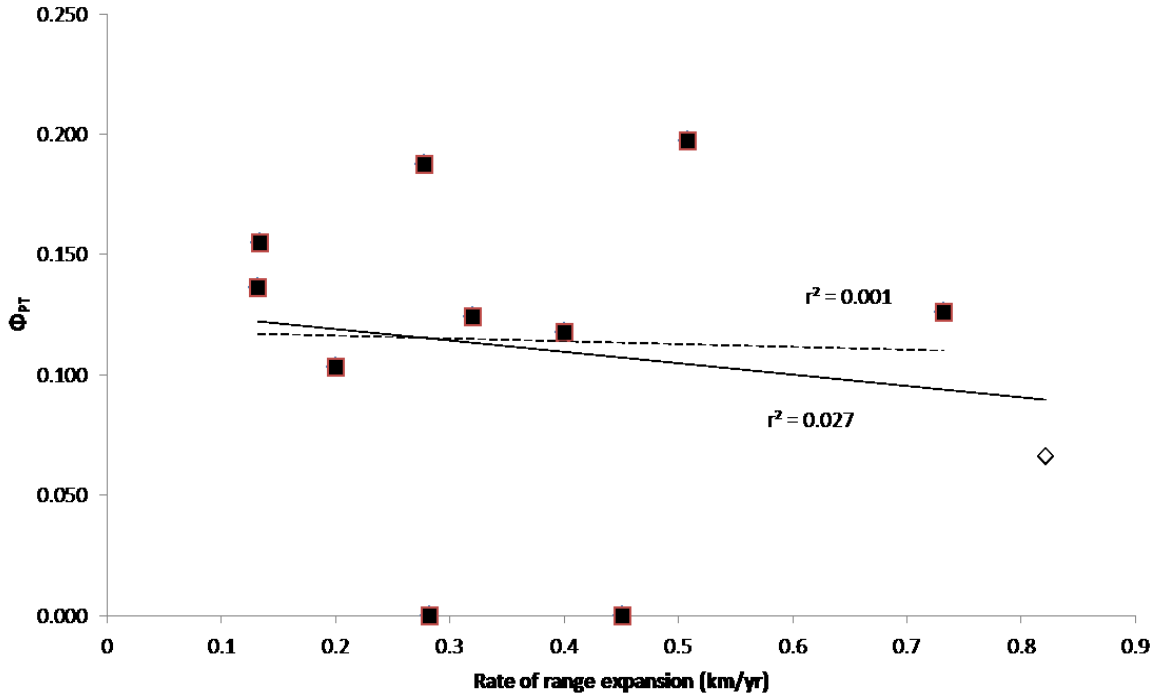


Figure 3.5 Regressions of pairwise genetic distances between core and peripheral populations (Φ_{PT}) against rate of range expansion from the core population. Regression results are represented by coefficients of determination (r^2). Dashed lines indicate trend excluding the Welsh population (represented by a white diamond) from the analysis.

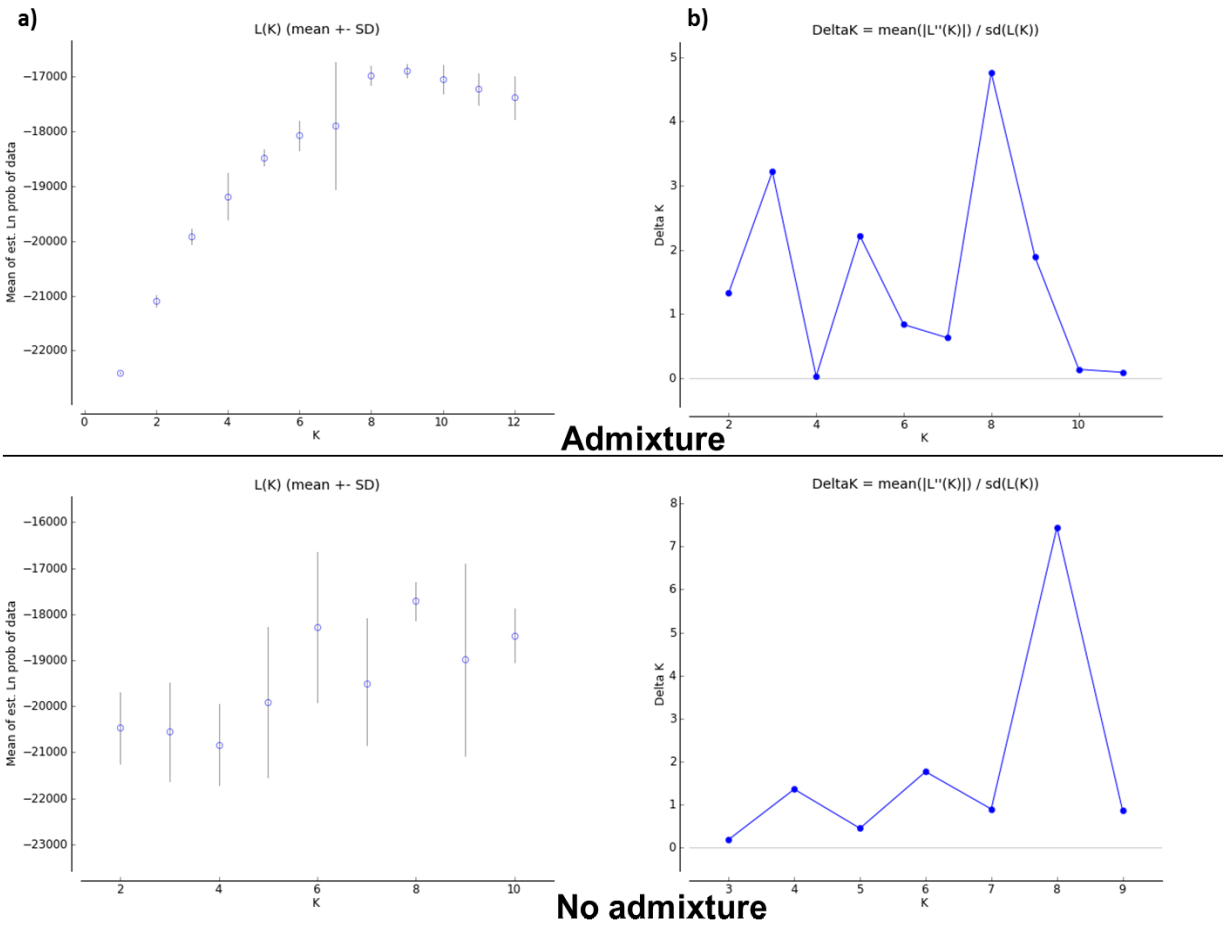


Figure 3.6 Summaries of STRUCTURE output. a) Mean posterior probabilities of the data averaged across runs against *K*. Error bars represent standard deviations after burn-in. b) Plots of Δ*K* against *K*. Upper plots represent models assuming admixture, lower plots represent models assuming no admixture.

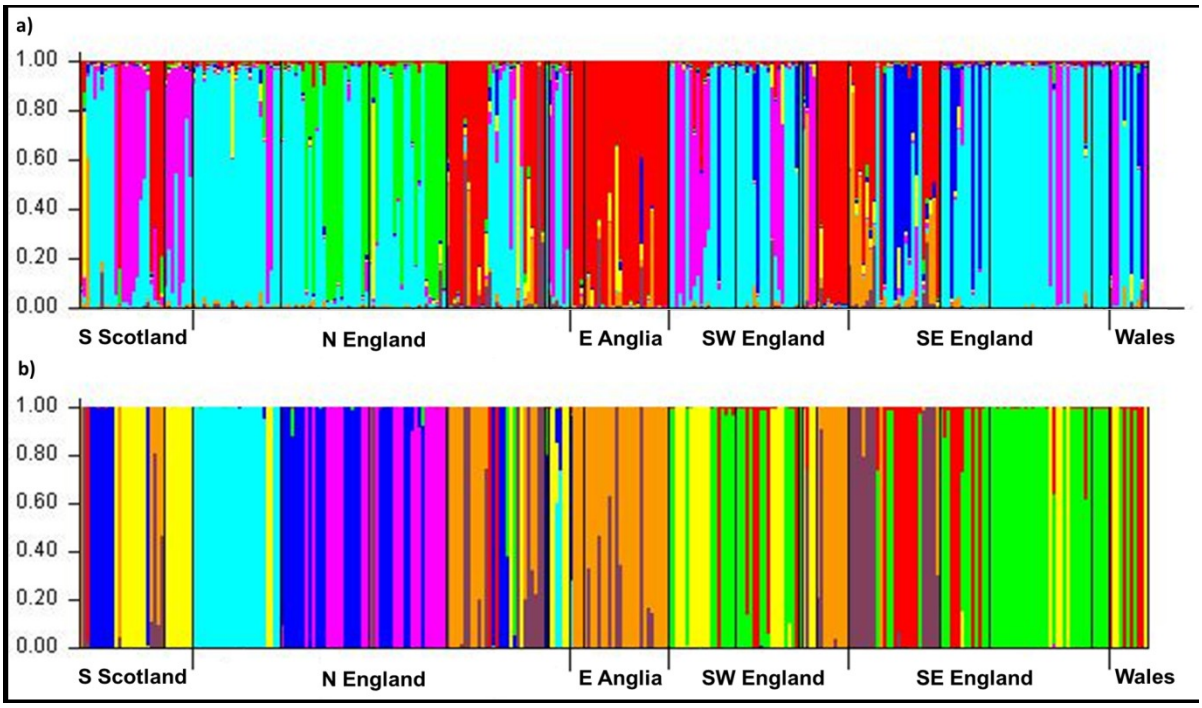


Figure 3.7 Histograms of posterior assignment probabilities determined by STRUCTURE where K = 8. Models assuming a) admixture b) no admixture. Each vertical bar represents an individual and its assignment proportion into one of eight clusters. Individuals are arranged into regions, and subdivided by sample sites.

Discussion

The scenario of recent range expansion from refugia and reintroductions in the roe deer of Great Britain can be thought of as a natural experiment which enables an examination of the influence on population genetic parameters by contemporary processes following documented historical bottlenecks. Populations in Scotland and Northern England were expected to have greater diversity than those derived from reintroductions in Southern England. Across all regions geographically peripheral populations were expected to be less diverse than core populations and more differentiated from each other than with their centre of abundance (Eckert *et al.* 2008). Many studies claim to find associations between genetic parameters and range expansion but do not explicitly test for a relationship (e.g. Pruett *et al.* 2011; Yang *et al.* 2012). Here, the availability of survey data on roe distribution allowed a rough estimation of the recent rate of range expansion in each region with which to test for correlations with population genetic diversity and differentiation.

Range expansion and genetic diversity

Of the three regions examined, the roe distribution in Southern England has the highest average gene diversity but the lowest spatial growth rate (Figure 3.2). Multiple introduction events and potential admixture in regions of overlap (e.g. CR, see Figure 3.1) may have contributed to the relatively high diversity in Southern England, while in Northern England the larger ancestral population size (Baker 2011) may explain the comparable level of diversity. The regional variation in spatial growth rate is likely to be due environmental heterogeneity (Acevedo *et al.* 2010). Differences in habitat availability and latitude have been shown to influence the rate of dispersal into new areas in Swedish and Norwegian roe, with deer in marginal habitats spreading faster than those in more favourable conditions (Andersen *et al.* 2004) under a conditional dispersal strategy (Andersen & Linnell 2000). Roe dispersal in Southern England may also be affected by habitat fragmentation, presence of barriers as a function of greater urbanisation in the south (Baker & Hoelzel 2013), and possibly interspecific

competition with other deer (Hemami *et al.* 2005). In areas of high muntjac density, roe deer numbers are lower than expected (Chapman *et al.* 1993), and there is substantial overlap between the roe and muntjac distributions in the south of England (Acevedo *et al.* 2010).

There was no clear north to south trend in levels of within-population diversity, in contrast to the findings of Baker &Hoelzel (2013). Rather, the peripheral populations within regions tended to have lower diversity than the core populations, with the notable exception of Windermere (Table 3.1). Direct comparisons of within-population diversity with previous genetic studies of the roe in Great Britain using allozymes(Hewison 1995), mtDNA and microsatellites (Baker & Hoelzel 2013) is likely to be misleading due to different rates of molecular evolution of these markers (Zachos *et al.* 2006a). However, when comparing results from microsatellite and AFLP markers, within-population expected heterozygosities estimated from AFLP data were lower than the values from microsatellites, and genetic distance values were larger, consistent with other comparisons (Mariette *et al.* 2001; Mariette *et al.* 2002; Alacs *et al.* 2011). Nevertheless, the hypothesis of Baker &Hoelzel (2013) that the Lancashire population has remained isolated is supported by these findings (using samples from forests in Windermere) of low genetic diversity in the WM population. However, it is also possible that the low sample size (unfortunately due to the majority of the WM samples yielding degraded DNA) may be giving an inaccurate measure of diversity. The higher diversity of the CR population relative to the two core populations in Southern England (MA and WS) suggests that admixture may have occurred in the area (Figure 3.1, Table 3.1). The central-marginal hypothesis of lower genetic diversity in peripheral populations is not supported when pooling both groups and comparing group differences. However, within-population diversity of peripheral populations had a significant negative correlation with both the distance from the site of introduction and the rate of range expansion (Figure 3.3). These patterns are in line with the predictions of Excoffier et al. (2009a) that the most rapidly expanding populations will have the lowest diversities. In roe deer this may be due to founder effects from smaller numbers of dispersers (Andersen *et al.* 1998) colonising new areas. The expansion rates from the core populations are considerably lower than in Scandinavian roe deer (Andersen *et al.* 2004) where the rate varies between 4 and 8km/yr. However, this calculation was based on more accurate

data (hunting records dating back to 1850) and range expansion rates are very unlikely to be constant, and should be interpreted as approximate measures for comparisons.

Range expansion and population structure

The high levels of population differentiation observed at the national and regional levels, with significant pairwise genetic distances between almost all populations (Table 3.2), reflect the complex introduction histories of the expanding populations. During range expansion, differentiation can result from founder effects and limited migration (Austerlitz *et al.* 1997), especially in the case of leptokurtic (long distance) dispersal (Ibrahim *et al.* 1996; Excoffier *et al.* 2009a). In this instance, genetic distances were generally higher between core populationsthanamong core and peripheral populations, demonstrating stronger historical genetic structure in refugial and reintroduction sites than in newly founded peripheral populations. Pairwise genetic distances between the East Anglian populations and most other populations were some of the highest, strongly supporting the documentary evidence of an introduction of 12 individuals from Württemberg, Germany to Thetford Forest (Whitehead 1964). The WM population was also strongly differentiated from almost all other populations, even neighbouring ones in Northern England. Interestingly, the highest genetic distance was observed between TH and WM despite their source populations (Southern Germany and Austria, respectively) being geographically close and likely to belong to the same 'Clade Central' of European roe deer (Randi *et al.* 2004), again suggesting a strong influence of founder effects during establishment.

Finding weak support for IBD in northern populations fits with a history of expansions from long-standing populations in Northern England and potential supplementary introductions into Ayrshire (Whitehead 1964), although it is likely to be disrupted by the presence of the introduced population in Windermere. This pattern is consistent with the findings of Baker &Hoelzel (2013) that northern populations in the area show a pattern of IBD, indicating philopatric behaviour (Nies *et al.* 2005). The majority of roe deer from Northern England in this study were sampled from woodland. IBD has

previously been observed in fragmented woodland habitat (Coulon *et al.* 2004), suggesting that gene flow is affected by landscape connectivity, although they could not rule out the effects of other land use types on gene flow. Additional support comes from a recent study of phenotypic and environmental correlates of roe deer dispersal finding that forest-dwelling populations have a lower dispersal probability than those inhabiting heterogeneous habitats (Debeffe *et al.* 2012). The rapid expansion of roe in East Anglia - where the open farmland habitat is suboptimal (Prior 1995) - lends support to the findings of Andersen et al. (2004)that roe tended to disperse further in unfavourable habitat. This conditional dispersal behaviour (Andersen & Linnell 2000) may underlie the very low levels of differentiation between the core and peripheral populations in the East Anglian area.

The lack of a relationship between genetic differentiation and rate of range expansion is surprising, as populations further from the core are expected to be spatially isolated with small effective population sizes (Vucetich & Waite 2003), and the rate of expansion should intensify the effects of this on gene flow. As roe tend to be highly sedentary in their established ranges (Hewison *et al.* 1998; Pettorelli *et al.* 2003b), the rate of expansion may not influence gene flow significantly as migration between populations is very low. However, in regions where range expansion is happening the most rapidly, the pairwise genetic distances between neighbouring populations are lowest. Φ_{PT} values between populations at the westernmost (CW-DV) and easternmost (ES-WS) areas of the Southern English region, along with East Anglia (EN-TH) were the lowest among all pairwise comparisons. An analysis of factors associated with roe deer presence (Acevedo *et al.* 2010) indicates that these areas have low favourability, lending support to the hypothesis that expansion is most rapid in unfavourable environments (Andersen *et al.* 2004). Rapid expansion may mean that the dispersed populations in these areas have founded so recently that differentiation has yet to occur, or that there are large numbers of migrants (Excoffier 2004). As the more peripheral populations tended to be the most genetically homogenous based on within-population diversity and population cluster analysis, the former explanation is the most likely in this scenario. These findings may reflect the huge variation in dispersal patterns of roe deer, where the majority of the individuals at the edge of the range disperse over much larger distances than those at the core (Wahlstrom & Liberg 1995).

In the Bayesian cluster analysis the 'no admixture' model showed a clear north-south divide between clusters, whereas strong signals of underlying Scottish ancestry were revealed in southern populations under the admixture model. Populations derived from the MA and WS cores exhibited similar patterns to the Scottish populations, comprising individuals from two main clusters. Approximately one third of the CR population appeared to be admixed. Northern populations also fell into two clusters, one local and one shared with the Scottish population, concordant with a southward spread from Scotland and natural expansion from refugia. The East Anglian populations clustered distinctly from the majority of the other regions, reflecting the historical differentiation between native British and introduced German deer. However, some individuals from NY and RO clustered with the East Anglian, mirroring the results of a craniometric study (Hewison 1997) which identified a separate morphotype in those areas[1], indicating that relatively recent secondary translocations may have occurred.

However, STRUCTURE has been found not to perform well when sampled populations deviate from HWE, which is the case in other roe populations (e.g. Coulon *et al.* 2006), and small sample sizes can affect the power to correctly assign individuals to the main genetic clusters (Evanno *et al.* 2005; Kalinowski 2011). Conducting a spatially-explicit analysis that accounts for admixture and does not require defining parental populations such as TESS (Durand *et al.* 2009) may be better able to describe population structure in the dataset(Francois & Durand 2010). These results lend tentative support to the existence of refugial roe populations in Northern England, documentary evidence of European roe being introduced in Windermere and East Anglia, and indicate that populations descended from multiple introductions in Southern England are admixing in areas of overlap.

[1] The sites of Romsey (RO) sampled here and Ringwood (sampled by Hewison(1997)) are less than 30 km apart, connected by the New Forest.

Implications

The principal finding here is that while within-population diversity declines with distance and rate of expansion from core populations of roe deer, there is no corresponding increase in genetic differentiation. Genetic drift in founder populations is likely to be the main factor explaining these patterns. Overall, the findings provide mixed support for the central-marginal hypothesis. Genetic diversities tended to decline with distance from core populations, but there was no significant difference between core and peripheral diversities or genetic distances. Rates of expansion vary widely between regions, and differences in the current distribution of roe introduced from different sources appear in large part to be due to natural expansion, excepting potential secondary translocations from East Anglian stock into localised areas within the southern and northern regions.

Incorporating population genetic information into wildlife management plans is a promising approach, particularly in the case of controlling invasive species (Hampton *et al.* 2004), and understanding population genetic structure is important for both conservation and management programmes (Zannese *et al.* 2006; Palsboll *et al.* 2007). Higher genetic diversity has been associated with establishment success (Gamfeldt *et al.* 2005; Lockwood *et al.* 2005; Crawford & Whitney 2010), and heterozygosity-fitness correlations have been demonstrated in a semi-captive roe deer population (Da Silva *et al.* 2009). As the reintroduced populations are currently significantly differentiated from each other, efforts could focus on preventing connectivity between them, as admixture can increase the invasive potential of populations (Kolbe *et al.* 2007).

Environmental factors associated with rate of spread in British roe deer populations have been investigated (Acevedo *et al.* 2010), calculating the rate of expansion as one figure for the entire country. Factors thatmay influence the rate and direction of past and future expansions of reintroduced or remnant native roe populations include the interaction between environment (such as availability of suitable habitat, barriers to dispersal, and presence of dispersal corridors) and population characteristics (such as population size, fecundity, and productivity). While this and other national-level studies of deer (e.g. Haanes *et al.* 2010; Baker & Hoelzel 2013) can identify broad patterns of

diversity and structure and indirectly infer the influence of landscape features on population genetic parameters, it is desirable to directly determine how gene flow is affected by landscape features (Manel *et al.* 2003; Schoville *et al.* 2012). Small-scale studies have demonstrated higher correlation between genetic distance and geographic distance when taking landscape features into account (Coulon *et al.* 2004; Coulon *et al.* 2006; Perez-Espona *et al.* 2008). To better explain the variation in rates of range expansion at regional and national levels, environmental heterogeneity should be accounted for (Segelbacher *et al.* 2010). Determining how genetic parameters are associated with landscape features and rate of range expansion in roe deer could enable better planning of future management based on predictions of deer movements into new areas (e.g. Finnegan *et al.* 2012). A further extension of this work could focus on developing a predictive model using data on roe deer dispersal behaviour and environmental conditions, comparing the predictions of this model to the spatial genetic patterns described here in order to fully assess the contribution of information on genetic variation to local and landscape-level management plans.

Chapter 4: Detection of adaptive population divergence and associated environmental correlates in European roe deer

Introduction

Understanding the genetic basis of adaptation is a major goal of evolutionary genetics (Feder & Mitchell-Olds 2003). Changing environmental conditions and colonisation of different areas can result in rapid adaptive evolution of wild populations to their new environment (Reznick & Ghalambor 2001), and adaptation has a crucial role in the extension of species ranges to new habitats (Kawecki 2008). In the case of introduced populations in new environments undergoing rapid range expansion, ecologically based selection can lead to local adaption (Sakai 2001; Lee 2002; Schoville *et al.* 2012). Identifying adaptive genetic variation in wild populations is crucial to developing an understanding of species responses to environmental change (Holderegger & Wagner 2008; Manel *et al.* 2010), which has implications for conserving the adaptive potential of populations threatened with extinction (Allendorf *et al.* 2010) and controlling invasions (Allendorf 2003; Lavergne & Molofsky 2007; Dlugosch & Parker 2008).

Identifying loci under selection and understanding which selective pressures are acting on non-model species in the wild is challenging (Manel *et al.* 2010; Nunes *et al.* 2011). Recently, population genomic approaches suitable for detecting loci influenced by selection in non-model species have emerged (Luikart *et al.* 2003). Genome scans can identify loci that appear to be under divergent selection. The underlying assumptions made by the genome scan approach are that population demographic and neutral evolutionary history will broadly affect all neutral loci in the genome similarly, whereas loci influenced by selection (and their neighbouring genomic regions) should exhibit distinct 'outlier' patterns of variation (Luikart *et al.* 2003; Nielsen 2005). The application of AFLP genotyping enables numerous loci distributed randomly across the genome to be surveyed simultaneously in non-model organisms (Bensch & Akesson 2005), although comprehensive representation of the entire genome is not made. This outlier-based approach has been fruitful in

identifying candidate loci associated with adaptation to environmental variables such as altitude (Bonin *et al.* 2006) and temperature (Jump *et al.* 2006). In addition to identifying candidate loci under selection, it is also possible to determine the selective forces acting upon them (Joost *et al.* 2007). Methods of identifying adaptive genetic variants in natural populations have been developed and adapted for use with dominant markers such as AFLPs. Two main approaches have emerged: a) outlier-based locus detection and b) landscape genomic modelling.

Outlier-based locus detection uses estimates of population differentiation to distinguish loci that deviate from expected neutral among-population differentiation relative to the genome-wide average neutral differentiation among populations (Beaumont & Nichols 1996; Beaumont & Balding 2004; Beaumont 2005; Foll & Gaggiotti 2008). Using stringent Confidence Interval (CI) and False Discovery Rate (FDR) (Benjamini & Hochberg 1995) thresholds alongside outlier replication can offer increased confidence that divergent selection is responsible for the result (Bonin *et al.* 2006; Nosil *et al.* 2009), due to the low probability that the F_{ST} of any given locus would repeatedly exceed its neutral expectation by chance. However, loci responding to selection pressures acting at smaller spatial scales than populations are designated at may be missed using a population-based analysis (Bothwell *et al.* 2013). Furthermore, outlier-based analyses are by themselves unable to demonstrate associations between environmental and genetic data. In such instances, landscape genomic modelling can be employed. This individual-based approach assumes that variation in environmental variables generate changes in allele frequencies at loci under selection (Endler 1977; Joost *et al.* 2007; Manel *et al.* 2010). The method is able to identify loci potentially linked to the targets of selection as well as generating new hypotheses regarding the environmental factors driving adaptation (Joost *et al.* 2007). Employing both analytical approaches and comparing the results allows for greater confidence that loci genuinely under selection have been detected (Manel *et al.* 2009; Nunes *et al.* 2011; Schoville *et al.* 2012).

Even when both approaches are combined, the challenge of distinguishing associations between identified loci and environmental variables from associations with other spatially structured processes

remains (Manel *et al.* 2010; Segelbacher *et al.* 2010).Patterns of genetic variation that appear to be due to selection may also result from historical and spatial effects (Kawecki & Ebert 2004; Schmidt *et al.* 2008). Restricted gene flow among populations will cause changes in allele frequency due to genetic drift (Wright 1938), but these effects should be genome-wide rather than restricted to specific loci. Analysing pools of populations from independent regions can help to mitigate the influence of confounding factors derived from historical and spatial processes (Poncet *et al.* 2010). By sampling multiple areas exhibiting a range of environmental variation, it should be possible to test for the generality of landscape genomic inference (Segelbacher *et al.* 2010). Europe is geographically complex, with a broad range of climates and habitat types. Identifying the same candidate loci in independent regions would provide strong support for a genuine association with a genomic region responding to selection from environmental conditions.

The European roe deer *(Capreoluscapreolus)* has been present across most of Europe from c.10, 000 BP to the present day(Sommer *et al.* 2009). Populations in Europe experienced a severe decline due to deforestation and overhunting prior to the nineteenth century, but have greatly increased in range and density in recent years, aided in some instances by active translocation (Andersen *et al.* 1998). Spread of populations has been assisted by an increase in availability of suitable habitat provided initially by woodland fragmentation and agricultural intensification, and subsequently by new woodland planting and agricultural set-aside (Robinson & Sutherland 2002; Stoate *et al.* 2009). Roe deer have responded well to the changes in landscape features and have spread into a variety of habitats(Andersen *et al.* 1998; Ward 2005; Morellet *et al.* 2011). The behavioural plasticity of the species has been largely credited with underlying this great ecological flexibility (Hewison *et al.* 2001), yet there may also be a role for other factors such as physiological responses (e.g. changes in the digestive tract (Guilloteau *et al.* 2012; Serrano Ferron *et al.* 2012)) and genetics in explaining the adaptation of roe deer to new environments. Climate can directly affect behaviour and physiology, and indirectly affect habitat resource availability in roe deer. Variation in life history and population dynamics is observed in areas with mild climate (Gaillard *et al.* 1993; Gaillard *et al.* 1996; Hewison *et al.* 1996). Climate is also expected to play an important role at the range limits (Mysterud & Ostbye 2006), and in fawn survival

(Putman *et al.* 1996; Gaillard *et al.* 1997; Andersen & Linnell 1998) and body mass (Danilkin 1996; Kjellander *et al.* 2006). Roe deer allocate high levels of resources to reproduction, and are therefore very sensitive to variation in the availability of resources in spring (Pettorelli *et al.* 2003a). As income breeders, female roe rely more on current resource intake than fat reserves, to offset the costs of reproduction (Hewison *et al.* 1996; Andersen *et al.* 2000). Future changes in climate are predicted to have considerable effects on the distribution of deer in Europe (Irvine *et al.* 2007), and roe are predicted to increasingly move from forest to open habitat in response to resource availability (Gaillard *et al.* 2013). Recent studies of roe deer in Great Britain have identified both climatic and habitat structure variables associated with distribution and range expansion. Factors most strongly associated with roe deer presence were climate extremes: precipitation of the wettest month and temperature of the warmest month and the following habitat structures: broad-leaved woodland, mountainous/upland terrain, and arable horticultural land (Acevedo *et al.* 2010). Factors associated with range expansion were the landscape features of forest, mountainous/upland terrain and pasture, and biodiversity (Shannon index) (Putman & Ward, 2010).

The questions I attempt to answer in this chapter are: 1) which AFLP loci are under selection in the roe deer genome? 2) are these loci affected by climate or habitat type? 3) if environmental variables are found to be associated with loci under selection, do they concur with the results of ecological models predicting roe distribution? Roe deer from expanding populations in three contrasting environments are examined: the Apennines in Italy, the lowlands in northern Germany and the oceanic wet and cold climate of Great Britain. This study design should help ameliorate the interference of differences in population history on the detection of adaptive genetic variants by isolating locus-specific versus genome-wide patterns of variation, and capture a range of environmental variation. Despite the wide intra-specific variation of European roe (Andersen *et al.* 1998; Putman & Flueck 2011), it is unlikely that these processes affecting the distribution of genome-wide variation would result in the same genetic patterns at a neutral locus in separate environmental conditions (e.g. Wark & Peichel 2010). Population-based analyses were conducted to search for outlier loci in each region, complemented by an individual-based landscape genomic exploratory analysis for loci associated with

climatic, altitudinal and land cover variables. Finally,the results of these analyses were compared with environmental variables found to predict roe deer distribution in order to identify loci potentially underlying the adaptation of roe deer to new environments.

Methods

Sample collection and georeferencing

Roe deer were sampled from 18 locations in Great Britain, five locations in Germany and two in Italy (Figure 4.1, Table 4.1). The regions sampled encompass a range of climatic variation, altitudes (1–535 m above sea level), and habitat types. Tongue tissues from the British deer were preserved in 20% w/v DMSO in ddH_2O saturated with NaCl, and stored at -20°C before extraction. Ear tissues from the Italian deer were dehydrated using 100% EtOH then dried before posting from Italy by Dr Massimo Scandura. DNA aliquots from the German populations were shared by Dr Frank Zachos. Six-figure grid references were recorded for the location of each British individual, and converted to latitude/longitude using the UK Grid Reference Finder Batch Convert Tool (http://gridreferencefinder.com/batchConvert/batchConvert.htm). Longitude and latitude co-ordinates for each European individual were converted from x/y location data using Batch Geocode (http://findlatitudeandlongitude.com/batch-geocode).

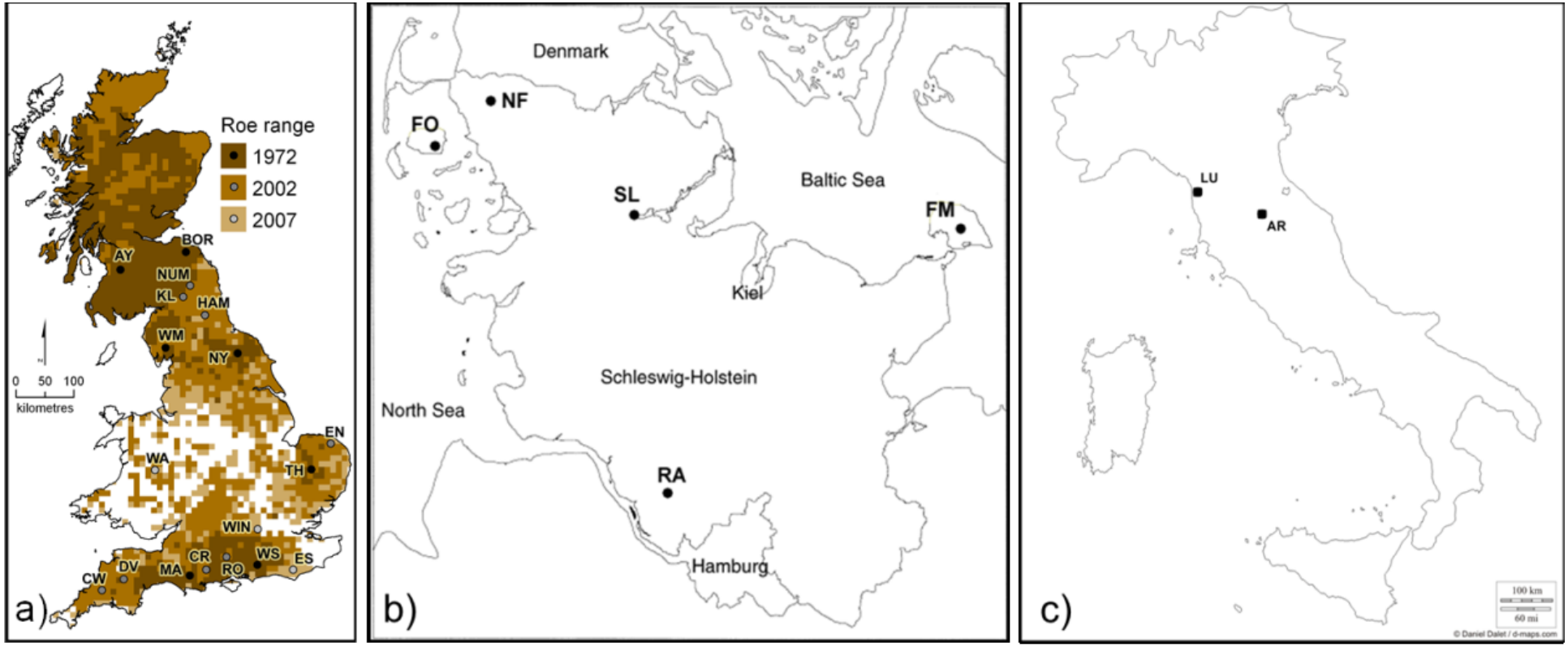


Figure 4.1 Maps of the sampled regions and sites of British and European roe deer. a) Great Britain, b) Northern Germany, c) Italy.Maps not to scale.

Names and regions of sample sites are presented in Table 4.1.

Environmental variables

Variables are presented in Table 4.2. Altitude and 19 GIS-derived bioclimatic variables for current conditions (~1950-2000) at 2.5 arc-minutes resolution were sourced from the WorldClim database (Hijmans *et al.* 2005). These variables are derived from monthly temperature and rainfall values in order to generate more biologically meaningful variables representing annual trends, seasonality, and extreme or limiting environmental factors (http://www.worldclim.org/bioclim), falling into two broad categories – temperature and precipitation. To minimise the redundancy of bioclimatic variables in the analysis, correlation coefficients between each pair of variables were calculated. Where pairs of variables had a correlation coefficient > 0.8, the variable deemed to be less biologically-relevant (e.g. not *a priori* associated with species distribution, derived variables) was excluded from further analyses. Nine bioclimatic variables were retained in the dataset. Land cover data at 30 arc-seconds resolution was sourced from the Global Land Cover 2000 (GLC2000) Project (http://bioval.jrc.ec.europa.eu/products/glc2000/products.php). Three forest (Tree Cover, broadleaved, deciduous, closed; Tree Cover, needle-leaved, evergreen; Tree Cover, mixed phenology or leaf type) and three non-forest (Shrub Cover, closed-open, deciduous; Herbaceous Cover, closed-open; Cultivated and managed areas) GLC2000 Global Classes were represented in the dataset. These were classified into four habitat structure types: broad-leaved woodland (deciduous, mixed, open birch, scrub), coniferous forest (conifers, felled and new plantations), pastoral (grassland), arable/horticultural land (arable cereals, arable horticulture and non-rotational arable and horticulture) and recoded following Acevedo et al. (2010). Variables at the sampling location of each individual were extracted using DIVA-GIS 7.1.6.2 (http://www.diva-gis.org).

Table 4.1 Sampled regions and sites of British and European roe deer. GB = Great Britain, latitude and longitude co-ordinates are midpoints of the sampled sites, n = number of AFLP genotyped individuals.

Region	Site name	Site code	Latitiude	Longitude	n
GB - North	Ayrshire	AY	55.2105	-4.4816	24
	Borders	BOR	55.8118	-2.3487	8
	Northumberland	NUM	55.6224	-2.2889	25
	Hamsterley	HAM	54.8924	-1.9127	22
	Kielder	KL	55.4154	-2.1200	25
	North Yorkshire	NY	54.2935	-0.6449	28
	Windermere	WM	54.2670	-2.9085	6
GB - East Anglia	Thetford	TH	52.5152	0.6852	24
	East Norfolk	EN	52.8589	1.3079	4
GB - South	West Sussex	WS	50.9226	-0.6910	19
	East Sussex	ES	51.0654	-0.2646	18
	Windsor	WIN	51.4702	-0.6916	4
	Romsey	RO	50.9807	-1.4973	9
	Milton Abbas	MA	50.8382	-2.2957	14
	Cranborne	CR	50.9193	-1.9217	26
	Devon	DV	50.8594	-3.2503	29
	Cornwall	CW	50.8737	-4.4844	5
	Powys	WA	52.2956	-3.1422	11
Northern Germany	Fehmarn	FN	54.4601	11.1337	12
	Föhr	FO	54.7182	8.5031	12
	Nordfriesland	NF	54.8788	8.8485	11
	Rantzau	RA	53.7278	9.8497	15
	Schleswig	SL	54.5854	9.5113	13
Italian Apennines	Arezzo	AR	43.5000	11.9570	8
	Lucca	LU	43.9600	10.4783	25

Table 4.2 Bioclimatic, topographic and land cover variables extracted for the location of each individual sampled. Nine bioclimatic variables indicated in bold and all four habitat types were retained for use in the analyses.

	Variable	Abbreviation / Habitat type
Topography	Altitude	Alt
Temperature	**Annual Mean Temperature**	**BIO1**
	Mean Diurnal Range (Mean of monthly (max temp - min temp))	BIO2
	Isothermality (BIO2/BIO7) (* 100)	BIO3
	Temperature Seasonality (standard deviation *100)	**BIO4**
	Max Temperature of Warmest Month	**BIO5**
	Min Temperature of Coldest Month	**BIO6**
	Temperature Annual Range (BIO5-BIO6)	BIO7
	Mean Temperature of Wettest Quarter	**BIO8**
	Mean Temperature of Driest Quarter	BIO9
	Mean Temperature of Warmest Quarter	BIO10
	Mean Temperature of Coldest Quarter	BIO11
Precipitation	**Annual Precipitation**	**BIO12**
	Precipitation of Wettest Month	**BIO13**
	Precipitation of Driest Month	BIO14
	Precipitation Seasonality (Coefficient of Variation)	BIO15
	Precipitation of Wettest Quarter	BIO16
	Precipitation of Driest Quarter	**BIO17**
	Precipitation of Warmest Quarter	BIO18
	Precipitation of Coldest Quarter	**BIO19**
GLC2000 Global Class	1. Tree Cover, broadleaved, deciduous, closed	Broad-leaved woodland
	6. Tree Cover, mixed phenology or leaf type	
	2. Tree Cover, needle-leaved, evergreen	Coniferous forest
	12. Shrub Cover, closed-open, deciduous	Pastoral
	13. Herbaceous Cover, closed-open	
	16. Cultivated and managed areas	Arable/horticultural land

Genetic data

DNA extraction and AFLP genotyping was carried out as described in Chapter 3. Roe deer samples from across Great Britain, Italy and Germany (Table 4.1) were genotyped and scored concurrently, with samples included on the same plates and using the same PCR machine in order to ensure comparability and repeatability among profiles.

DNA extraction

Total genomic DNA was extracted from tissue samples using DNeasy Blood & Tissue Kits (Qiagen, UK) following the manufacturer's instructions. DNA was extracted from approximately half of the samples using the 96Blood&Tissuekit with 96-well plates; the remainder of DNA extractions were conducted with the Blood & Tissue kits in batches of 24 individual extraction tubes. Some samples were coated with hair or soil, so were wiped clean with ethanol before the digestion step. For each sample ~25mg of tissue was cut from the tongue and diced using sterilised forceps and a fresh scalpel blade. Each batch of extractions included one negative to check for contamination between wells/tubes. Subsets of tissue samples from each population were extracted twice independently, in order to allow estimation of genotyping error rates (Bonin *et al.* 2004). 1µl of each DNA extract was added to 4µl loading buffer and electrophoresed on 1.5% agarose gels in 1X TAE buffer for 90 min at ~120V to examine the extract for degradation. Substantially degraded DNA extracts were excluded from further analysis. DNA concentration was then measured using a Nanodrop 2000 (Thermo Scientific).

Genotyping

AFLP profiles were obtained from the samples using a modified version of Protocol 2 described by Papa et al. (2005). The AFLP technique comprises restriction enzyme digestion of total genomic DNA

followed by ligation of adaptors, then two rounds of selective PCRs which generate a set of polymorphic fragments (Vos *et al.* 1995). Potential combinations of selective amplification primers were trialled on a subset of samples representing most populations, in order to test the genotyping quality and select the best four with which to genotype the rest of the samples. Primers were selected on the basis of maximising the number of polymorphic peaks per sample, evenness of peak size distribution in the 50-500bp range, and repeatability. 30 combinations of selective amplification primers were initially tested, and four were chosen for subsequent analyses. Reactions were carried out as follows: ~500 ng of DNA was dried at 60°C in a vacuum oven then resuspended in 5.5 µl nuclease-free water (Sigma-Aldrich). Initial restriction enzyme digestion using 5 U *Taq*I (Thermo Scientific) per sample was carried out at 65°C in a buffer solution (10 mMTris-HCl [pH 8.0 at 37°C], 5 mM $MgCl_2$, 100 mMNaCl, 0.1 mg/mL Bovine Serum Albumin [BSA]) for 90 min. Subsequent digestion with 10 U *Eco*R1 (Applied Biosystems) was carried out at 37°C in the buffer solution described above for 2 h. Next, the *Taq*I adaptors (Ajmone-Marsan *et al.* 1997) were prepared in 100 µl batches at 50 pmol/µl by heating an equimolar solution of the two oligonucleotides *Taq*I top strand (5'-GACGATGAGTCCTGAC) and *Taq*I bottom strand (5'-CGGTCAGGACTCAT) to 95°C for 3 min and cooling slowly to room temperature. *Eco*RI adaptors came from the AFLP Regular Genome Plant Mapping Kit (Applied Biosystems). To ligate adaptors, a solution containing 5 pmol*Eco*RI adaptors and 50 pmol*Taq*I adaptors, 1 U T4 DNA ligase (Promega), 1X T4 DNA Ligase buffer (Promega), 50 ng/µl BSA and 1mM Adenosine 5′-triphosphate (ATP) (Sigma-Aldrich) was added to the digested DNA solution, mixed gently and the combined solution incubated at 16°C for 17 h. The resulting template mixture was diluted fourfold in nuclease-free water (Sigma-Aldrich). All pre-selective and selective amplifications were carried out with Fermentas (Termo Scientific) PCR mastermix, using a Veriti® 96-Well Fast Thermal Cycler (Applied Biosystems). A random subsample of the template fragments was obtained through pre-selective amplification using the primers E01 (5'-GAC TGC GTA CCA ATT CA) and T01 (5′-GAT GAG TCC TGA CCG AA) described by Ajmone-Marsan et al. (1997). 5 µl of the diluted DNA template mixture was added to 75 ng of each primer in a 20 µlPCR mix. PCR conditions were as described in Papa et al. (2005). The reaction products were diluted 20 times with nuclease-free water (Sigma-Aldrich). Selective amplifications used 5 µl of the diluted

product with 10 – 15 ng of fluorescently labelled *Eco*RI (5′-GAC TGC GTA CCA ATT CNN N_{DYE} [Applied Biosystems]) selective primer and 30 ng of *Taq*I (5′-GAT GAG TCC TGA CCG ANN N) selective primer in a 20 µlPCR mix. Touchdown PCR conditions were as described in Papa et al. (2005). The four *Eco*RIand *Taq*I primer combinations were: *Eco*RI-ACA_{FAM}/*Taq*I-ACT, *Eco*RI-AAG_{JOE}/*Taq*I-CAT, *Eco*RI-AAC_{NED}/*Taq*I-CCA, *Eco*RI-ACA_{FAM}/*Taq*I-CCA. 1 – 1.2 µl of each selective PCR product from each sample was mixed with 10 µlHi-Di formamide (Applied Biosystems) and 0.2 µl GeneScan-500 ROX size standard (Applied Biosystems) in a single well for fragment analysis by capillary electrophoresis on an ABI 3130*xl* genetic analyser (Applied Biosystems). Samples with genotype failures were re-run once. They were re-amplified from genomic DNA and the undiluted pre-selective reaction products were electrophoresed on 1.5% agarose gels at 90V for 2 h to check for the presence of amplified fragments before selective amplification and capillary fragment analysis was re-run.

AFLP electropherograms were visualised using GeneMapper v. 4.0 (Applied Biosystems). Profiles with poor PCR amplification (i.e. few peaks or peak strength declining rapidly with fragment size) or poor sizing were removed from the dataset. Bins within the range of 50 to 500 bp were manually identified. All bins had at least one peak of height ≥100 relative frequency units (*rfu*). In order to minimise the effect of size homoplasy (Caballero *et al.* 2008) on the results, bin widths were set at <1 bp and had to be non-overlapping. Bins in which peaks were >0.3 bp apart were removed from the analysis, and different bins had to be separated by >0.4 bp. Once the maximal bin set was created, tables of 'sum of signal' normalised peak heights for each primer pair were exported from Genemapper. AFLPScore v. 1.4b (Whitlock *et al.* 2008) was used to optimise scoring parameters and create a binary genotype table for each primer pair.Locus and phenotype *rfu* selection thresholds were explored for permutations of scoring methods (i.e. filtered loci/absolute thresholds, unfiltered/absolute, filtered/relative and unfiltered/relative) in AFLPScore, with the aim of maximising the number of loci retained and minimising the mismatch error rate. A binary genotype table for each primer pair was then generated under the optimal threshold settings, and the four tables were concatenated. Mean error rate was 7.3% (± 0.4 S.D.) which is slightly higher than the normal range of error rates using AFLPs

(Bonin *et al.* 2004). Loci with zero allele frequencies, and those with a fragment present or absent in one individual only were then removed, giving a final dataset comprising 168 polymorphic locifor 397 individuals.

Population structure and historical effects

Support for considering regional groupings of populations as independent regions with no recent admixture between them was assessed by individual-level Bayesian cluster analysis conducted in STRUCTURE v. 2.3.4 (Pritchard *et al.* 2000; Falush *et al.* 2007). The most plausible number of populations (K), and the probability of assignment for each individual to each population (Q) was assessed with and without admixture-based ancestry models, assuming independent and correlated allele frequencies (Falush *et al.* 2003). For each model, seven runs were conducted for each value of K between 1 and 20, each using 100,000 Markov Chain Monte Carlo (MCMC) generations with 50,000 generations discarded as burn-in. The most probable value of K was approximated using two methods. The first determines the highest log-probability of the data Pr(X|K) as described in the STRUCTURE manual (Pritchard *et al.* 2000) using a plot of Pr(X|K) averaged across runs against K. The smallest value of K where log-probabilities are similar was determined to be the value of K with the highest hierarchical level of population structure in these data. The second approach assessed the rate of change in log-likelihood between successive values of K (i.e. ΔK) (Evanno *et al.* 2005). $\Delta K = |mL(K + 1) - 2mL(K) + mL(K - 1)| / sdL(K)$, where mL is the mean and sdL is the standard deviation of the likelihood value (L) across the seven independent runs. The number of populations described by the modal value in the ΔK distribution was determined to be the correct value of K. Results from the runs were summarised and plotted using STRUCTURE HARVESTER Web v. 0.6.93(Earl & vonHoldt 2012). Potentially admixed individuals can introduce uncertainty to landscape genomic analyses (Bothwell *et al.* 2013). In order to minimise the incidence of spurious associations between AFLP loci and environmental variables, all individuals with an average population membership probability <0.8 were excluded from the dataset, resulting in a final dataset of 385 individuals. These were grouped

into regions on the basis of the Structure analysis and known population history in Great Britain. Pairwise genetic differentiation (Φ_{PT}) among the five regions was estimated using a dataset with outlier loci removed in GeneAlEx v. 6.5 (Peakall & Smouse 2012).

Outlier-based analysis

To identify outlier loci, the regions were analysed separately taking the approach of Beaumont & Nichols (1996) and Beaumont & Balding (2004) using a modification of the FDIST2 algorithm (Beaumont & Nichols 1996) for dominant markers (DFDIST) implemented in Mcheza (Antao& Beaumont 2011). DFDIST estimates F_{ST} values conditional on expected heterozygosity. Allele frequencies were estimated under a Bayesian method (Zhivotovsky 1999), and unbiased F_{ST} values with respect to sample size (Wier & Cockerham 1984) were estimated for each locus. The mean "neutral" F_{ST} was calculated and used in the simulations (Caballero *et al.* 2008). Coalescent simulations of 50,000 loci were performed to generate an F_{ST} null distribution. Loci with unusually high F_{ST} values relative to the null distribution are identified as outliers (potentially under divergent selection). Only outliers under divergent selection were considered, as the power of this method to detect balancing selection is low (Beaumont & Nichols 1996; Beaumont & Balding 2004). Analyses were conducted under the most stringent settings with a 99.5% CI and a 1% FDR to minimise the occurrence of false positives. All other parameters were kept at the default values with $\theta = 0.1$, β-a andβ-b = 0.25 and a critical frequency = 0.99.

Landscape genomic analysis

The Spatial Analysis Method (SAM) is individual-based, and does not rely on assumptions regarding the structure of the sampled populations (Joost *et al.* 2008). The null hypothesis of the SAM - that there is no association between a locus and an environmental variable - is rejected when the model

containing the environmental variable better explains the allele frequency distribution better than a model containing only a constant (Joost *et al.* 2007). Multiple univariate logistic regressions between environmental variables at sample locations and allele frequencies at each locus were carried out using MatSAM v. 2 (Joost *et al.* 2008). Wald tests were used to estimate statistical significance of coefficients in the logistic regression models but not likelihood ratio (G) tests, as these do not allow for missing data. AFLP loci were considered to be significantly associated with environmental variables at a 99% CI after Bonferroni correction of the significance level for multiple comparisons (set to 2.59×10^{-6}).

Results

Outlier-based analyses

A total of 168 polymorphic AFLPs were generated using four primer pairs from 397 individuals in 25 populations (Table 4.2). 30 individuals retained in the dataset had < 50% missing genotype data. The mean posterior probabilities of the STRUCTURE runs suggested an optimal $K = 9$ under the admixture model (Figure 4.2). There was a clear division between clusters of individuals sampled from Germany, Italy and Great Britain (Figure 4.3). Within the British samples, the admixture model highlighted the introduction of Scottish deer into Southern England, while the East Anglian deer clustered distinctly. Based on the population demographic history of roe deer in Great Britain, with southern populations having been founded from multiple translocations while northern populations have expanded from refugia(Whitehead 1964; Baker & Hoelzel 2013) populations from these areas were pooled into separate regions for subsequent analyses. The Welsh population was assigned to the southern region on the basis of cluster membership from the Structure results. All five groupings of regional populations were significantly genetically differentiated from each other (Table 4.3). Out of 168 AFLP loci analysed, 16 (9.52%) were identified as outliers by DFDIST analysis (99.5% CI; FDR = 0.01) in total across all regions. Four of these loci were identified in two regions (Table 4.4).

Table 4.3 Pairwise Φ_{PT} values between the five regions below diagonal, and probabilities based a permutation test with 10,000 pseudoreplicates are above the diagonal.

	GB - North	GB - East Anglia	GB - South	Northern Germany
GB - North	-	0.002	0.000	0.000
GB - East Anglia	0.055	-	0.000	0.000
GB - South	0.148	0.160	-	0.001
Northern Germany	0.102	0.106	0.303	-
Italian Apennines	0.126	0.098	0.321	0.144

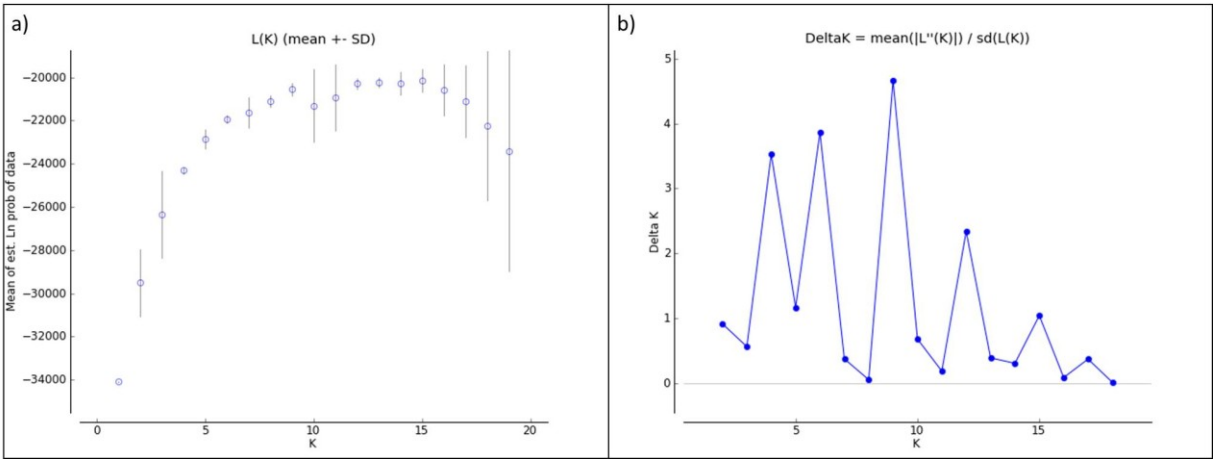


Figure 4.2 Summaries of STRUCTURE output.a) Mean posterior probabilities of the data averaged across seven runs against K for K = 1 to K = 20. Error bars represent standard deviations after burn-in. b) Plots of ΔK against K.

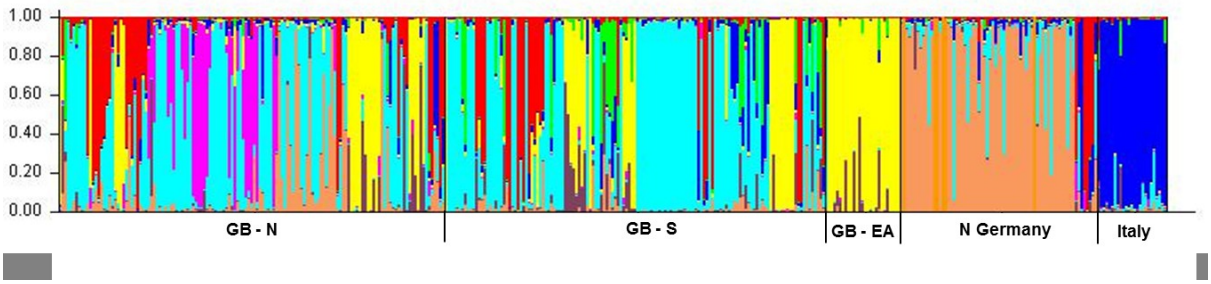


Figure 4.3 Histogram of posterior assignment probabilities determined by STRUCTURE where K = 9 assuming an admixture ancestry model. Each vertical bar represents an individual and its assignment proportion into one of nine clusters. Individuals are arranged into regions, and subdivided by sample sites. GB-N = Northern Great Britain, GB-S = Southern Great Britain, GB-EA = East Anglia in Great Britain, N Germany = Northern Germany, Italy = Italian Apennines.

Table 4.4 AFLP loci detected as outliers potentially under divergent selection by DFDIST, and loci with significant associations with environmental variables using SAM. An "x" indicates the locus was detected as an outlier under divergent selection at the 99.5% CI and a 1% FDR, and "xx"indicates that theoutlier status was detected in two regions. For SAM, environmental variables significantly correlated with a locus at a 99% CI after Bonferronicorrection are listed. Loci detected by both approaches are in bold.

Locus	DFDIST	SAM (associated variables)				Environmental assocation
3	xx					
4	x	BIO13	BIO17			Precipitation
5		BIO4	BIO8			Temperature
7	x					
10	x					
11		Alt	BIO1	BIO6		Altitude+Temperature
14		BIO1				Temperature
15		BIO6				Temperature
16	x					
19		BIO6				Temperature
32		BIO1	BIO4	BIO5	BIO17	Temperature+Precipitation
41		BIO1	BIO5			Temperature
42	xx	BIO12	BIO13			Precipitation
43		BIO13				Precipitation
44	x					
45		BIO6				Temperature
46	x	BIO12	BIO13			Precipitation
74	x	BIO4	BIO6	BIO8		Temperature
77		BIO4	BIO6	BIO8		Temperature
82	x	BIO13				Precipitation
83		BIO6	BIO13			Temperature+Precipitation
88	x	Alt	BIO13	BIO17		Altitude+Precipitation
89	xx					
98	x	BIO1	BIO4	BIO5		Temperature
99		BIO1	BIO4	BIO5		Temperature
119		BIO12				Precipitation
123	x	BIO13				Precipitation
124		BIO13				Precipitation
127		Alt				Altitude
136	xx					
139		Alt	BIO1	BIO4	BIO5	Altitude+Temperature
149		BIO1	BIO5	BIO6		Temperature
159		BIO4				Temperature
161		BIO4				Temperature
162		BIO4	BIO5			Temperature
165	x					

Loci-environment associations

In the SAM analysis, 28 loci (16.7%) were identified as being associated with one or more environmental variables (Table 4.4). Eight bioclimatic variables and altitude were associated with at least one locus; however no associations with habitat type were detected. 11 loci were associated with temperature, 18 with precipitation and four with altitude. Two loci were associated with both temperature and precipitation, two had an association with both altitude and temperature, and one was associated with altitude and precipitation. One locus was associated with altitude only.

Concordance with environmental variables predicting roe distribution

Seven of the 28 loci associated with environmental variables were also detected as outliers by DFDIST. Of those seven, one locus was detected as an outlier in two regions and associated with two precipitation variables: Annual precipitation (BIO12) and precipitation in the wettest month (BIO13).

Of the eight climatic variables associated with loci putatively under section, two have been found to predict roe deer presence in Great Britain (Acevedo *et al.* 2010). The variable most strongly associated with roe presence (BIO13) is linked to 9 loci, of which 6 were detected as outliers. The second most strongly associated variable: Maximum temperature of the warmest month (BIO5) is linked to 7 loci, of which only one was detected as an outlier.

Discussion

Connection with environmental predictors of roe deer distribution

These findings provide evidence for the association of loci under divergent selection with climatic variables. The association of variation in extremes of precipitation (BIO13) and temperature (BIO5) with roe deer presence in Great Britain and with loci under selection suggests that climatic extremes may be acting as selective pressures on roe deer populations. Spring rainfall is known to affect the summer survival of both male and female fawns (Gaillard *et al.* 1997). The effect may be acting directly through lower mortality due to overheating during droughts (or higher mortality in moist areas such as Scotland), or indirectly through effects on resource availability.

The relationship between temperature and selective pressure is less clear. Severe winters result in higher mortality in roe deer (Gaillard *et al.* 1993). The minimum temperature of the coldest month (BIO6) was negatively associated with roe presence in Great Britain (Acevedo *et al.* 2010), and six loci (one identified as being under divergent selection in the Italian population) are associated here. Resource availability and nutritional quality are seasonally variable (Moser *et al.* 2006) and higher annual and maximum temperatures should also correlate with greater resource availability. A recent study found that higher spring temperatures around the birth of a female fawn have a negative effect on their adult body mass later in life, potentially due to environmental stress (Hamel *et al.* 2009a), and this impacts upon annual reproductive success of females (Hamel *et al.* 2009b).

The lack of any genetic association with habitat type is likely to reflect the considerable behavioural plasticity of the roe deer in response to landscape structure (Hewison *et al.* 2001; Jepsen & Topping 2004). Open-land roe deer populations show distinct differences in spatial and social behaviour, including larger group sizes (Putman 1988), compared with forest-dwelling roe deer populations (Jepsen & Topping 2004). Given the ability of roe to change rapidly from ‘solitary’ forest-dwelling behaviour to grouping behaviour in open habitats, phenotypic plasticity or behavioural adaptation is

more likely to explain roe presence in these environments than natural selection (Putman & Flueck 2011).

Evidence for adaptation

The proportion of outliers detected by DFDIST (9.52%) is within the range of most AFLP genome scans using the program across all species (5 – 10%) (Nosil *et al.* 2009) and for plants (0.2 – 17.1%; mean of ~5.5%) (Strasburg *et al.* 2012). However, direct comparisons can be misleading as the sampling, study designs, and stringency criteria intervals used by other studies vary widely (see Nosil et al. (2009) for a review). Number of outlier loci detected by DFDIST per region varied between seven in Italy and only one in Germany. Interestingly, the locus identified as an outlier in Germany (136) was also detected as an outlier in Italy. Applying the landscape genomic method (SAM analysis) to the data set detected a considerably higher proportion of outlier loci (16.7%) than the outlier analysis. Taken together, these results suggest that annual precipitation (associated with 5 variables in the SAM analysis, 4 of which are identified as candidate loci under divergent selection by the outlier analysis) and precipitation of the wettest month (associated with 4 variables in the SAM analysis, 3 of which are identified as candidate loci under divergent selection by the outlier analysis) are exerting a selective pressure on roe deer populations.

Differing results from genome scans using different methods are common (Nosil *et al.* 2009; Nunes *et al.* 2011). The differences in number and identity of loci detected by DFDIST and SAM analysis may be explained by the different focus on tempo and mode of selection between the methods (Nielsen 2005; Sabeti *et al.* 2006; Nielsen *et al.* 2007). DFDIST implemented in Mcheza(Antao & Beaumont 2011) is population-based, while SAM is an individual-based method which performs analyses without reference to the population with which the individuals were sampled from (Joost *et al.* 2007; Joost *et al.* 2008). Only loci identified as being under divergent selection were reported from the DFDIST analysis due to its low power to detect balancing selection, while loci detected by the SAM

analysis may be under divergent or balancing selection (Joost *et al.* 2007). Additionally, a variety of additional biological and technical factors can have a role in causing false-positive and false-negative results in genome scans (Strasburg *et al.* 2012). For example, the Italian and German study sites have lower breadth of environmental variation than the British, which may have influenced the power to detect significant environmental associations in those regions.

The pooling of populations with varying patterns of neutral differentiation from geographically distinct sites may violate the assumptions of the finite-island model used by the DFDIST outlier-based analyses, however simulations have shown that the method is robust to heterogeneous levels of gene flow between populations (Beaumont & Nichols 1996). Studies adopting a similar approach to pooling distinct populations and making comparisons between them for repeated outlier detection have yielded convincing results (Bonin *et al.* 2006; Nosil *et al.* 2008; Nosil *et al.* 2009; Buckley *et al.* 2012), and Mcheza implements an improved method for accurately approximating the simulated distribution of mean F_{ST} by applying a correction factor (Antao & Beaumont 2011). Nevertheless, these results ought to be cautiously interpreted as significant deviations from the model can result in a higher false-positive rate (Excoffier *et al.* 2009b).

While the outlier-based analysis should be robust to heterogeneous levels of gene flow between the pooled populations (Beaumont & Nichols 1996; Beaumont 2005), and the most stringent CI and FDR criteria were applied to the DFDIST analysis, the application of a different outlier-based method could help to confirm which outliers are most likely to be genuinely under divergent selection. Of particular concern here is the confounding factor of population structure within the regions analysed. Isolated populations and hierarchical structuring can lead to an increased rate of false-positive results from outlier-based approaches assessing allele frequency divergence (Excoffier *et al.* 2009b). The overall F_{ST} across sites varied between the regions analysed. The British regions had relatively high between population differentiation relative to the Italian and German regions studied. STRUCTURE has been found not to perform well when sampled populations deviate from HWE, which is the case in other roe populations (e.g. Coulon *et al.* 2006), and small sample sizes can affect the power to correctly

assign individuals to the main genetic clusters (Evanno *et al.* 2005; Kalinowski 2011). Conducting a spatially-explicit analysis which can account for admixture and does not require defining parental populations, such as TESS (Durand *et al.* 2009) may be better able to describe underlying population structure in the dataset.

A population genetic study of the north German roe deer based on microsatellite and mtDNA control region data found that all populations were significantly differentiated from one another (Zachos *et al.* 2006a), while population-level analysis of the AFLP data generated from a subset of the same DNA samples used in those studies did not find significant levels of between-population genetic differentiation (Appendix 4.2). The two island populations FO and FN, separated by the mainland have the highest level of differentiation ($\Phi_{PT} = 0.031$, $P > 0.1$). These island populations were founded by translocations and subsequently supplemented by additional introductions (Zachos *et al.* 2006a; Zachos *et al.* 2006b), which is likely to have reduced genetic differentiation between them and mainland populations. Zachos et al. (2006b) used highly polymorphic markers with many alleles to analyse differentiation. Discrepancies between AFLP and microsatellite markers are common, with AFLPs often finding different levels of differentiation between populations (Mariette *et al.* 2001; Mariette *et al.* 2002; Alacs *et al.* 2011). Bayescan(Foll & Gaggiotti 2008) allows for variable within-population F_{ST} values, may have a lower false-positive rate than other outlier-detection methods (Narum & Hess 2011) and is the most conservative outlier-based method in a simulation study (Perez-Figueroa *et al.* 2010). However, both methods are still susceptible to finding false-positives where allele frequencies are correlated between pooled populations (Robertson 1975; Excoffier *et al.* 2009b). Comparing results between different regions is therefore unlikely to be fruitful using this additional method. Given the recent expansion of the populations within the regions analysed here, some false-positives are expected due to not taking population structuring into account. A new methodology which estimates the null F_{ST} distribution while accounting for population structure has recently been developed (Excoffier *et al.* 2009b), but unfortunately an implementation for dominant marker data is not yet available.

Limitations and future work

This study is the first to apply population genomic methods to identifying loci under selection in European roe deer. The aim was to find candidate loci underlying adaptation to environmental conditions, some of which have been demonstrated *a priori*via ecological modelling to be associated with roe distribution (Acevedo *et al.* 2010) and expansion (Putman & Ward, 2010). Examining more roe populations from different areas representing a broad range of environmental variation could give greater confidence in the results. Accounting for population structure could further reduce the false-positive rate (Zhang *et al.* 2008) in the outlier based analyses. Finding a significant correlation when linearly regressing the allele frequencies of the candidate loci in each population analysed against the associated environmental variables would provide further support for their having a role in local adaptation(Narum *et al.* 2010). However, in some of the regions the populations are rapidly expanding their range, and under these conditions rare alleles can suddenly increase in frequency in peripheral populations at the leading-edge of the expansion (Excoffier 2004; Excoffier & Ray 2008; Excoffier *et al.* 2009a; Arenas *et al.* 2012). As a result, the false-positive rate may increase as loci swept up by a wave of increasing allele frequency are detected alongside loci under selection.

This study is also limited by the influence of population structure on the study regions, the influence of spatial autocorrelation on patterns of genetic variation, the marker system used, and the statistical methods used. AFLPs are a reasonably cost-effective tool for characterising the genome of non-model species (Meudt & Clarke 2007), but are not capable of genome-wide coverage and are hampered by their dominant nature and difficulties with reliable scoring (Herrmann *et al.* 2010). Attention has recently turned towards developing statistical methods for next-generation population genomics, with strong advances being made in reducing the influence of underlying assumptions on the false positive rate in identifying loci under selection, such as incorporating genetic distances in measures of genetic differentiation (Gompert & Buerkle 2011).

The ideal markers for studying adaptation will be directly involved in the genetic mechanisms of adaptive traits, have a sequence of known function, and exhibit quantifiable variation. Characterising the candidate loci in roe deer is therefore the next step in determining their contribution to adaptation. Further studies could isolate and sequence the genomic regions identified as being under selection, then compare the sequences with the annotated bovid genomes. This approach can help to determine whether the identified genomic region has a functional or regulatory role (Nielsen 2005; Stinchcombe & Hoekstra 2008).

Chapter 5: Genetic structure and diversity of native and introduced Chinese water deer: implications for conservation

Introduction

Sufficient genetic variation is vital for the long-term survival of a population (Lande 1988), particularly in the case of rapid environmental change (Lande & Shannon 1996). The adaptive potential and reproductive fitness of a population is generally enhanced by greater levels of genetic diversity, while inbreeding is likely to occur in small populations with low levels of genetic diversity (Frankham 1995). Furthermore, loss of genetic variation in small populations may increase extinction risk due to increased disease susceptibility, and decreased reproductive fitness and adaptive flexibility (Allendorf & Luikart 2007). Determining levels of genetic diversity in threatened species can therefore help inform suitable conservation strategies (Frankham *et al.* 2002). The aim of this study is to investigate levels of genetic diversity and partitioning of genetic variation in native and introduced populations of Chinese water deer (*Hydropotes inermis*), in order to provide recommendations for the international conservation of the species.

The Chinese water deer is a small ruminant species in the family Cervidae, and the only cervid species that does not possess antlers. Its preferred habitats are reed-beds or tall, damp and undisturbed grasslands (Zhang *et al.* 2006). The historical distribution of the species once stretched along the eastern part of China from Liaoning to Guangdong out to the lower Yangtze Basin, and the Korean Peninsula (Ohtaishi & Gao 1990). The population in China has reduced considerably in range and numbers recently due to habitat loss and poaching (Hu *et al.* 2006). It is currently restricted to fragmented populations in the eastern Yangtze Basin and the Zhoushan Islands off the coast of Zhejiang Province. Published estimates of total Chinese water deer numbers in China have declined from ~20,000 in 1993 to fewer than 10,000 in 2009 (Cooke 2009). The species is classified as Vulnerable on the IUCN Red List. In China, habitat destruction and fragmentation looks set to

continue at an accelerating pace. The Lower Yangtze wetlands, especially Dongting and Poyang Lakes are being degraded by pollution and sedimentation (Dudgeon 2010).

Chinese water deer were introduced to Great Britain in the 1870s, initially to London Zoo. Starting in 1896, individuals were sent gradually to Woburn Abbey in Bedfordshire. In 1929, 32 individuals were transferred to the nearby Whipsnade Zoo (Cooke & Farrell 1998). During World War II a number of Chinese water deer escaped from Woburn and established localised feral populations (Cooke 2009). In 1950 a small number of deer were released near Woodwalton Fen National Nature Reserve (Chapman 1995). Wild populations founded from escapes and deliberate releases since the 1940s are now discontinuously distributed across Bedfordshire, Cambridgeshire, Norfolk and Suffolk (Cooke & Farrell 2008). The most recent distribution map from the National Deer Survey conducted by the British Deer Society (Figure 5.1) shows continued range expansion of the Chinese water deer in Great Britain. Total estimated numbers in Great Britain stand at 7,000 in the wild (A. Cooke pers. comm.) and >500 in semi-captivity. This represents a significant proportion (over 40%) of the global numbers of this subspecies of Chinese water deer. Various zoological parks across Europe have been stocked with Chinese water deer from Whipsnade Zoo (N. Lindsay pers. comm.).

Based on analysis of the mtDNA control region of the deer in China, Hu *et al.* (2006) showed significant differentiation between Zhoushan Islands and mainland populations and relatively high levels of genetic diversity in the mainland Chinese populations. These analyses however pooled very low numbers of samples per mainland population ($1<n<6$) collected up to 47 years apart across four provinces (some over 500km apart), and compared them with a sample from the Zhoushan Islands of which over one third were sourced from a captive breeding centre. The findings are unlikely to represent true levels of genetic diversity within mainland Chinese populations, nor to give an accurate estimate of differentiation of these populations from the Zhoushan Islands populations. A subsequent study using seven microsatellite loci to compare a mainland Zoo population (again containing individuals from different regions) with one wild and one captive population from the Zhoushan Islands also found high levels of genetic diversity in the mainland population, significant

differentiation between the mainland and Zhoushan Islands, and described the Zhoushan Islands population as "severely inbred" (Hu *et al.* 2007). Both studies recommend the establishment of a captive breeding programme for mainland populations. A programme aiming to reintroduce the deer to the suburbs and wetlands surrounding Shanghai was already underway at the same time. This programme has now established five captive breeding populations sourced from the Zhoushan Islands, and recently released two of these into the wild (M. Chen pers. comm.).

This investigation of the population genetics of the Chinese water deer focuses on two main questions: 1) how do levels of diversity and differentiation compare between native and introduced populations? 2) what is the source of the British population? If the British populations are not significantly differentiated from the mainland Chinese populations, they may be able to supplement declining populations. Alternatively, if the British populations are significantly differentiated, or their source population has gone extinct they may have value as a 'reservoir' of lost genetic variants which could be sourced for reintroduction to uninhabited areas of the historical range.

In this study, mtDNA sequence data from the entire non-coding control region (926 bp) and protein-coding *cytb* gene (1140 bp) were analysed for over 106 individuals from 10 locations in the native and introduced ranges. This is the first genetic study of Chinese water deer to examine populations in the introduced range. Both molecular markers show high levels of diversity in native Chinese populations, significant differentiation between and within the Chinese and British populations, and that very few haplotypes are shared between deer in the native and introduced ranges.

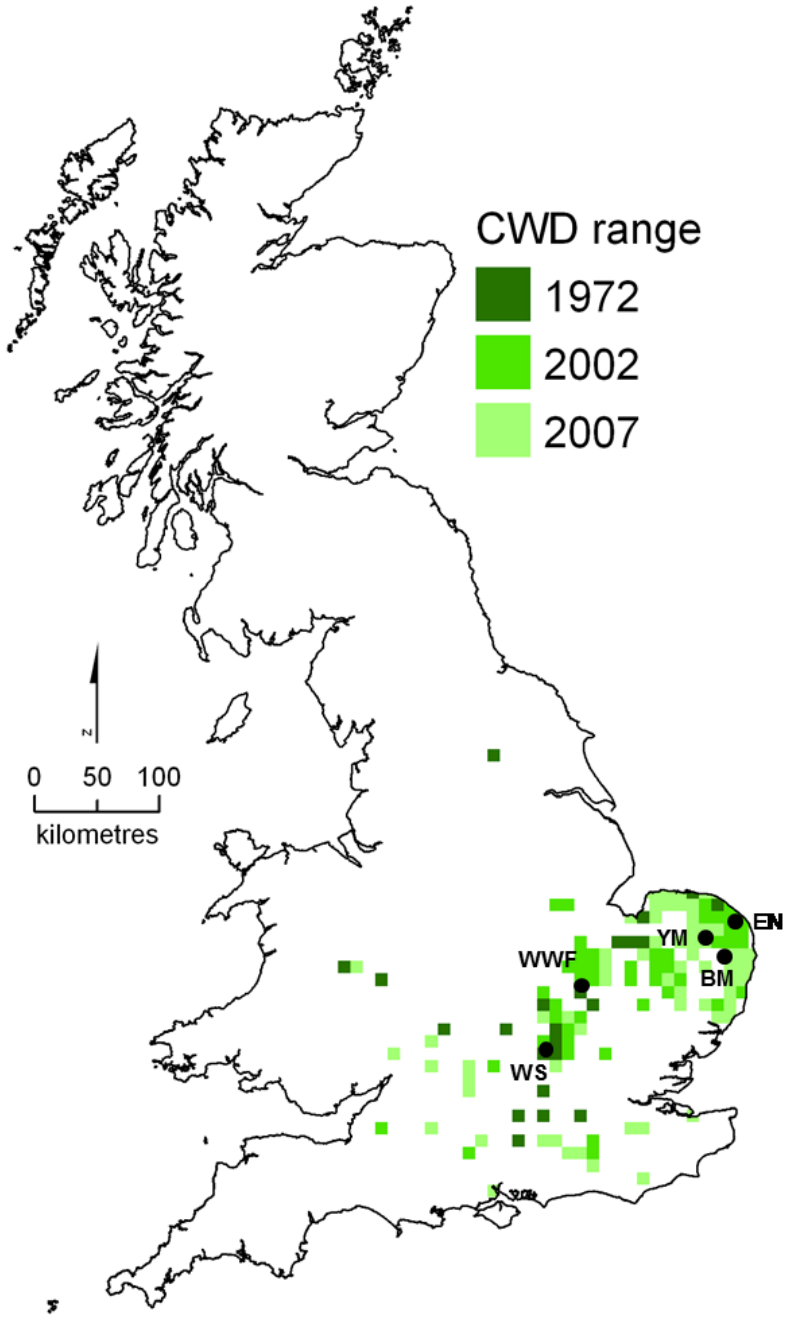


Figure 5.1 Chinese water deer distribution in Great Britain and sample sites. The occupied grid square from 1972 in North Yorkshire is now unoccupied. EN = East Norfolk, BM = Bure Marshes, YM = Yare Marshes, WS = Whipsnade Zoo, WWF = Woodwalton Fen.

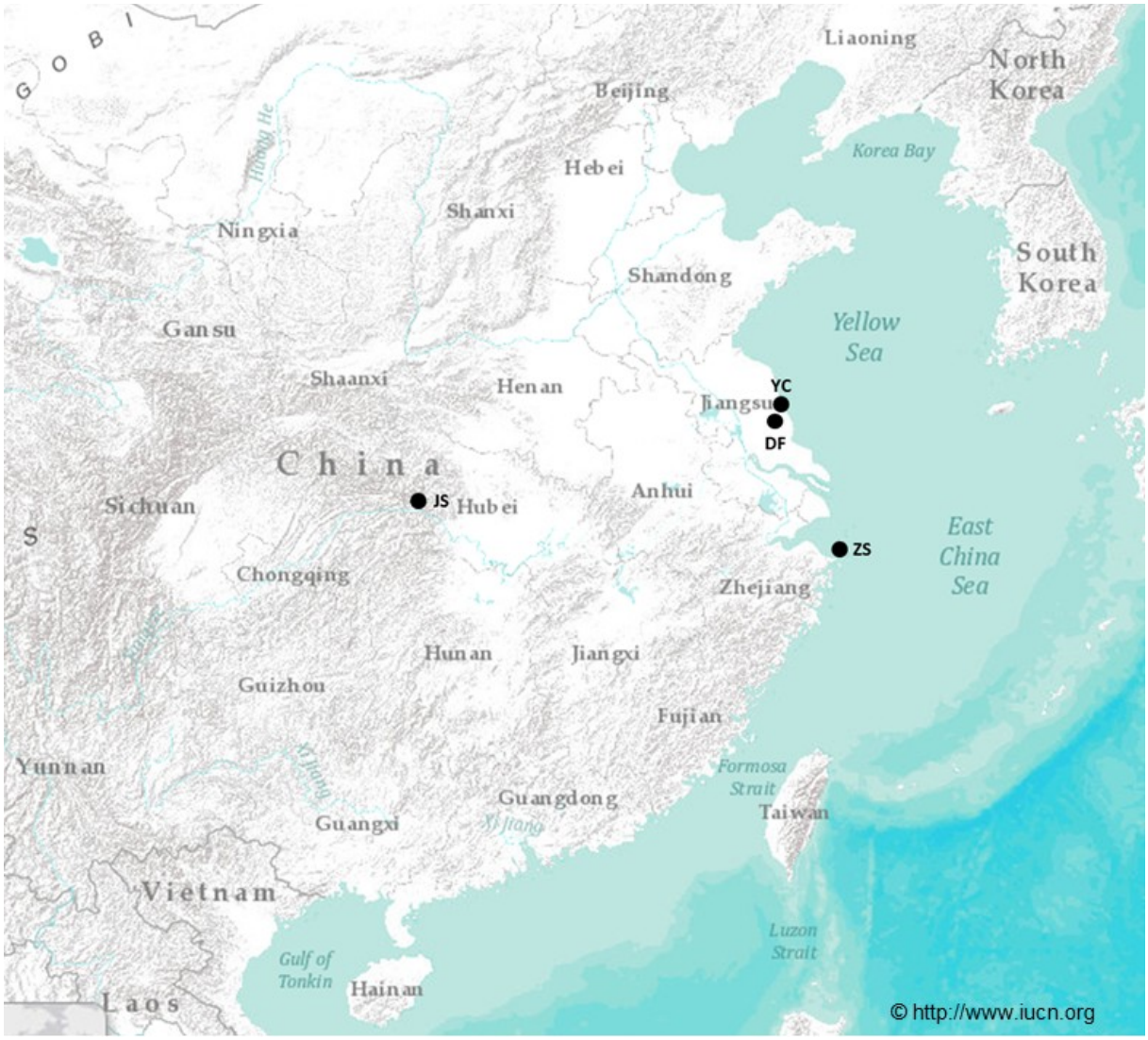


Figure 5.2 Sample sites of Chinese water deer in the native range. JS = Jishan Island, Dongting Lake, YC = Yancheng Coastal Nature Reserve, DF = Dafeng Nature Reserve, ZS = Zhoushan Islands.

Materials and Methods

In order to compare genetic variation within and between populations in China and those in Great Britain and France, a population genetic study of the Chinese water deer encompassing their native and introduced ranges was conducted. Samples collected from the introduced range were genotyped in this study, and combined with existing sequence data from the native range for the analyses. mtDNA control region and *cytb* sequences fromChinese populations (Figure 5.2) were obtained by Dr Min Chen from populations in three mainland sites: Jishan Island on Poyang Lake in Jiangxi Province, Yancheng and Dafeng coastal wetlands in Jiangsu Province, and an archipelago: Zhoushan Islands in Zhejiang Province using a combination of faecal and hair samples.

Sample collection

Samples from Great Britain were obtained from wild populations in Cambridgeshire and Norfolk, and from a semi-captive population in Bedfordshire (Figure 5.1). Samples from France were obtained from a semi-captive population in Rhodes (St. Croix). Sample collection locations and number are presented in Table 5.1. Muscle tissue was collected from deer shot in season (at East Norfolk and three from Woodwalton Fen) and post-mortems (St. Croix). Tissue samples were cut from tongue and stored in 20% w/v DMSO in water saturated with NaCl. In all other sites, samples were faecal material (5-10 freshly deposited faecal pellets from each dropping encountered, collected using sterile toothpicks and stored in 95% ethanol). All faecal samples were collected between November and February from frozen ground, as the genotyping error rate has been demonstrated to be significantly lower from ungulate faeces collected in winter (Maudet *et al.* 2004). All samples were kept frozen at -20°C prior to analysis.

DNA extraction, PCR amplification and sequencing

The mucosal layer containing intestinal cells was washed off the faecal pellets with lysis buffer (0.1 m Tris–HCl, 0.1 m EDTA, 0.01 m NaCl, 1% N-lauroylsarcosine, pH 7.5–8). One pellet from each sample was placed in a 15mL tube and washed in 2 mL of lysis buffer at room temperature for 20 minutes on a rotary agitator. 250 µL was aliquoted into a 2 mL tube, 250 µL buffer AL and 55 µL protease were added and the mixture, which was then incubated at 55°C for two hours on a rotary agitator. Next, 250 µL of 100% EtOH was added and DNeasy Mini Spin Columns (Qiagen) were used for the wash and elution steps. Total DNA was eluted using 150 µL AE buffer heated to 70°C left to stand for 10 minutes in the column before elution. Tissue samples were finely chopped and digested for 16 hours with proteinase-K at 37°C before total DNA extraction using the DNeasy tissue kit (Qiagen, UK) following the manufacturer's instructions. DNA concentration was measured using a Nanodrop 2000 (Thermo Scientific).

All amplification and sequencing reactions were conducted using a Veriti® 96-Well Fast Thermal Cycler (Applied Biosystems). The entire 926bp mtDNA control region was amplified using Lpro and Hphe primers (Table 5.2) under the PCR conditions described by Royo et al. (2007) from 100-200ng DNA extracted from tissue samples. For the DNA extracted from faecal samples the overlapping internal primers H493 and L362 (Table 5.2) were used in combination with Lpro and Hphe respectively to generate two shorter overlapping fragments. The PCR protocol for these reactions started with an initial activation step at 95°C for 5 min, followed by 35–40 cycles, with denaturation at 94°C for 30 s, annealing at 50°C for 30 s, extension at 72°C for 60 s and final extension at 72°C for 7 min. In some samples, amplification of the control region using the four published primers failed. Primer 3 (Rozen and Skaletsy 1998) was then used to design modified Chinese water deer-specific versions of the primers to amplify two shorter overlapping fragments of the control region (Table 5.2) using the PCR protocol mentioned previously.

The entire 1140bp mtDNA*cytb* gene was amplified using two primer pairs (*cytb*_L14723F + *cytb*_H15350R, and *cytb*_L15171F* + *cytb*_H15915R) to amplify four overlapping fragments under

the PCR conditions described by Pitra et al. (2004) from 100-200 ng DNA extracted from tissue samples. For the DNA extracted from faecal samples all four *cytb* primer pairs in Table 5.2 were used to generate four overlapping fragments under the same PCR conditions. 5 µL of each PCR product was run on a 1% agarose gel stained with GelRed™ (Invitrogen) and the quantity and quality of the product was estimated. If there was a single, strong band signifying successful amplification, 2 µL of ExoSAP-IT® (Affymetrix) was used to clean the PCR product following the manufacturer's instructions. Where there was a single, weak band the PCR reaction was repeated for that sample, the products were pooled then cleaned using a Qiaquick PCR cleanup kit (Qiagen, UK) following the manufacturer's instructions. Where there were multiple bands, the entire PCR product was run separately on a 0.5% agarose gel for 2 hours, then the portion containing the fragment of interest was excised with a scalpel blade on a UV lightbox and purified using the Qiaquick gel extraction kit (Qiagen, UK) following the manufacturer's instructions but including an extra wash step before elution, then purifying the eluate using ethanol precipitation to remove excess salts.

Sequencing reactions were carried out using the BigDye^{TM} Terminator Cycle Sequencing Ready Reaction Kit (v 3.1, Applied Biosystems). Both forward and reverse strands were sequenced for amplified fragments on an ABI 3130*xl* automated DNA sequencer (Applied Biosystems). DNA sequences were analysed in Geneious (Biomatters) and revised manually. For the poor quality samples, PCR amplification and sequencing were repeated several times. A negative control was included as standard to test for contamination and quality test controls were conducted by repeating PCRs and sequencing 10% of samples. Samples that did not show identical results in consecutive analyses were excluded from the dataset. Samples from the Whipsnade population were sequenced by Christian Miquel at the Laboratoired'Ecologie Alpine, Université Joseph Fourier.

All available complete sequences from Genbank for the control region of individuals from France (Branféré Park – JN632649), the USA (San Diego Zoo – Y08208), China (Yancheng - EU315254; NC_011821) and South Korea (Gyeonggi-do – JF802125), and for *cytb* from France (Branféré Park – AJ000028; JN632649), China (Yancheng - EU315254; NC_011821), and four South Korean regions

(Gangwon-do - EF139148; EF139149, Gyeongsangbuk-do - EF139150; EF139151; EF139152, Jeollanam-do - EF139153 EF139154; Gyeonggi-do - EF139155; JF802125) were included in the haplotype network dataset. The two Genbank sequences from Yancheng were pooled with the newly generated data in the population-level analyses.

Table 5.1 Sample sites and total number of samples collected, PCR amplified and sequenced for mtDNA control region and *cytb*.

Country	County	Site	# samples collected	# sequenced at control region	control region PCR success rate (%)	# sequenced at *cytb*	*cytb* PCR success rate (%)
Great Britain	Norfolk	Bure Marshes	33	8	24.2	20	60.6
	Norfolk	East Norfolk	4	4	100	4	100
	Bedfordshire	Whipsnade	44	43	97.7	42	95.5
	Cambridgeshire	Woodwalton Fen	20	15	75	20	100
	Norfolk	Yare Marshes	4	3	75	4	100
France	Rhodes	St. Croix	5	4	80	5	100
Total			*110*	*77*	-	*95*	-
Average			-	-	*75.3*	-	*92.7*

Table 5.2 Oligonucleotide primers used to amplify and sequence mtDNA*cytb* and control region sequences. # = Primers designed using Chinese water deer mtDNA genome reference sequence NC_011821.

mtDNA sequence	Name	Section	Sequence	Reference
control region	L-pro	-	CGTCAGTCTCACCATCAACCCCCAAAGC	Jäger, F., Hecht, W. and Herzog, A. (1992). Untersuchungen an mitochondrialer DNS (mtDNS) von hessischem Rehwild (*C. capreolus*). *Z Jagdwiss*, 38: 26–33.
	Lpro_CWD21#	-	CCCACCATCAACACCCAAAGC	Own design
	H-493	-	TGAGATGGCCCTGAAGAAAGAACC	Randi E., Mucci N., Pierpaoli M. & Douzery E. (1998) New phylogenetic perspectives on the Cervidae (Artiodactyla) are provided by the mitochondrial cytochrome)b gene. Proceedings of the Royal Society of London, series B 265 : 793-801
	L-362	-	AATCACCATGCCGCGTGAAACC	Randi E., Mucci N., Pierpaoli M. & Douzery E. (1998) New phylogenetic perspectives on the Cervidae (Artiodactyla) are provided by the mitochondrial cytochrome)b gene. Proceedings of the Royal Society of London, series B 265 : 793-802
	H-phe	-	GGGAGACTCATCTAGGCATTTTCAGTG	Jäger, F., Hecht, W. and Herzog, A. (1992). Untersuchungen an mitochondrialer DNS (mtDNS) von hessischem Rehwild (*C. capreolus*). *Z Jagdwiss*, 38: 26–33.
	Hphe_CWD22#	-	ACTCATCTAGGCATTTTCAGTG	Own design
cytb	cytb_L14723F	***1***	ACCAATGACATGAAAAATCATCGTT	Hassanin, A., Pasquet, E. & Vigne, J.-D. (1998) Molecular Systematics of the Subfamily Caprinae (Artiodactyla, Bovidae) As Determined From Cytochrome B Sequences J. Mamm. Evol. 5, 217–326.
	cytb_H15152R	***1***	CCTCAGAATGATATTTGTCC	Hassanin, A., Pasquet, E. & Vigne, J.-D. (1998) Molecular Systematics of the Subfamily Caprinae (Artiodactyla, Bovidae) As Determined From Cytochrome B Sequences J. Mamm. Evol. 5, 217–326.
	cytb_L14908F	***2***	GCCTATTCCTAGCAATACAC	Hassanin, A., Pasquet, E. & Vigne, J.-D. (1998) Molecular Systematics of the Subfamily Caprinae (Artiodactyla, Bovidae) As Determined From Cytochrome B Sequences J. Mamm. Evol. 5, 217–326.
	cytb_H15350R	***2***	CCTGTDGGGTTGTTDGANCCTGTTTC	Hassanin, A., Pasquet, E. & Vigne, J.-D. (1998) Molecular Systematics of the Subfamily Caprinae (Artiodactyla, Bovidae) As Determined From Cytochrome B Sequences J. Mamm. Evol. 5, 217–326.
	cytb_L15171F_CWD#	***3***	ATGAGGACAAATATCATTCTGAGG	Own design
	cytb_H15672R_CWD#	***3***	TATGCTGCGTTGTTTAGATG	Own design
	cytb_L15612F	***4***	CGATCAATYCCYAAYAAACTAGG	Hassanin, A., Pasquet, E. & Vigne, J.-D. (1998) Molecular Systematics of the Subfamily Caprinae (Artiodactyla, Bovidae) As Determined From Cytochrome B Sequences J. Mamm. Evol. 5, 217–326.
	cytb_H15915R	***4***	TCTCCATTTCTGGTTTACAAGAC	Hassanin, A., Pasquet, E. & Vigne, J.-D. (1998) Molecular Systematics of the Subfamily Caprinae (Artiodactyla, Bovidae) As Determined From Cytochrome B Sequences J. Mamm. Evol. 5, 217–326.

Population genetic analyses

Sequences downloaded from Genbank were aligned with the newly generated sequences using the G-INS-I strategy in MAFFT v.7 (Katoh *et al.* 2005; Katoh & Toh 2008). The best fitting models of nucleotide substitution were determined using MEGA 5.0 (Tamura *et al.* 2011). The Bayesian Information Criterion (BIC) found that the T92 model with $\gamma = 0.05$ fitted the control region data best (BIC = 6450.7), and the HKY model with $\gamma = 0.12$ fitted the *cytb* data best (BIC = 9179.3). The T92 model and rate heterogeneity correction were used in all population genetic analyses of the control region data. As ARLEQUIN v3.5 (Excoffier & Lischer 2010) does not feature the option to use the HKY model, the next best fitting model – T92 with $\gamma = 0.10$ (BIC = 9202.5) was used for analyses of*cytb*data.

Estimates of haplotype diversity (*h*; the probability that two randomly selected haplotypes are different), mean pairwise difference (*k*) and nucleotide diversity (π; the average number of nucleotide differences per site in pairwise sequence comparisons) were calculated using ARLEQUIN. To determine the partitioning of genetic variation among countries, among populations within countries, and within populations an Analysis of MOlecularVAriance (AMOVA) was conducted in ARLEQUIN using a framework of predefined geographical structure. AMOVAs were performed using both *Φ*-statistics (which consider haplotype relationships using a model of genetic distance and a gamma parameter estimate), and conventional F-statistics (using haplotype frequencies). Significance values were obtained from a minimum of 10,000 permutations. Pairwise F_{ST} and Φ_{ST} values were estimated using ARLEQUIN and their significance was tested using a minimum of 1000 permutations.

Haplotype relationships were calculated and depicted in NETWORK v4.2.0.1 (Fluxus Technology Ltd.) using a median-joining network approach (Bandelt *et al.* 1999). NETWORK calculates Minimum Spanning Trees (MST) from a haplotype distance matrix, and then combines the MST to form a single network with missing node haplotypes inferred using the parsimony criterion added to the network to reduce its total length. For the analysis of intraspecific genetic variation, haplotype networks give a more complete representation of relationships than tree algorithms as they explicitly

allow for the co-existence of ancestral and descendant alleles in a sample, instead of treating all sequences as terminal taxa (see Posada and Crandall (2001) for a review). DNAsp v.5 (Librado & Rozas 2009) was used to analyse haplotype distributions, and to prepare and construct files for analyses in ARLEQUIN and NETWORK. With the exception of the haplotype network analysis, population-level analyses were only performed for regions where four or more individuals were sampled.

Results

Out of 110 Chinese water deer samples that were collected in Great Britain and France, 77 yielded reliable sequence data for the mtDNA control region, and 95 for *cytb*. Chinese samples sequenced by Dr Min Chen yielded 24 control region and 34 cytb sequences, and a further five control region and 13 *cytb* sequences were available from Genbank.

mtDNA control region analyses

Genetic variation

Details on polymorphism in the 926 bp sequence in each country, and levels of genetic diversity in each sampling site are presented in Table 5.3. Excluding sample sites where n<4, the Chinese populations contained higher levels of diversity than those in Great Britain and France. In Chinese populations *h* ranges from 0.905 (±0.054) in the Zhoushan Islands to 1 (±0.177) in Dafeng and π x 10^3ranges from 6.95 (±3.91) in the Zhoushan Islands to 28.4 (±16.8) in Yancheng, while in Great Britain *h* ranges from 0.391 (±0.084) in Whipsnade Park to 0.893 (±0.111) in Bure Marshes and π x 10^3ranges from 2.4 (±1.68) in Bure Marshes to 3.88 (±2.22) in Whipsnade Park.

Population structure of native and introduced Chinese water deer populations

There is continued debate regarding the accuracy and reliability of comparing haplotype frequencies using F-statistics versus comparing genetic distances using Φ-statistics (Excoffier *et al.* 1992; Formia *et al.* 2006). AMOVA F-statistic results are presented in Table 5.4 for comparison with Φ-statistics, and pairwise F_{ST} values are presented in Appendix 5.1. As Φ_{ST} is likely to be more resilient to the effects of small sample size than F_{ST}, and Φ-statistics should represent more accurate estimates of

differentiation where the populations are distinguished by geographic separation rather than haplotype frequencies (O'Corry-Crowe *et al.* 1997), only Φ_{ST} values are presented in Table 5.5.

AMOVA on five British and four Chinese populations showed that a significant amount of the total variation (~42%) occurs within populations (Φ_{ST} = 0.579, P < 0.001), and that a significant proportion (~11%) is also explained by differences among populations within countries (Φ_{SC} = 0.208, P < 0.001). AMOVAs performed on British and Chinese populations grouped by country confirmed that the majority of the total variation in control region sequences occurs within populations. Variation within Chinese populations explained ~78% of the total variance (Φ_{ST} = 0.211, P = 0.009), compared to ~82% within British populations (Φ_{ST} = 0.184, P < 0.001).

Genetic differentiation among populations

Pairwise genetic differentiation between Chinese, French and British populations was estimated using F_{ST} and Φ_{ST} values. Pairwise comparisons between the Chinese Zhoushan Islands population and all other populations were all highly significant. Differentiation from the two mainland Chinese populations (Φ_{ST} ranging from 0.239 to 0.253) was considerably lower than that of the French and British populations (Φ_{ST}= 0.660 for Whipsnade Park and Φ_{ST}= 0.764 for Woodwalton Fen). The Dafeng and Yancheng populations are also significantly differentiated from almost all others, but not from each other. The St. Croix Park population in France is significantly differentiated from Bure Marshes and Woodwalton Fen but not Whipsnade Park, with which it shares the lowest Φ_{ST}value (0.202). Within Great Britain, Whipsnade Park is also significantly differentiated from Bure Marshes and Woodwalton Fen, but not from East Norfolk or Yare Marshes.

Haplotype relationships

The haplotype network (Figure 5.2) shows a clear split between introduced and native populations of Chinese water deer, separated by a number of missing haplotypes. No haplotypes are shared between any British and Chinese populations, and all populations are characterised by unique haplotypes. Only one haplotype (CR-H_16) is shared between the French and British populations overall. Interestingly, it is shared between the Woodwalton Fen population rather than the documented source population in Whipsnade.

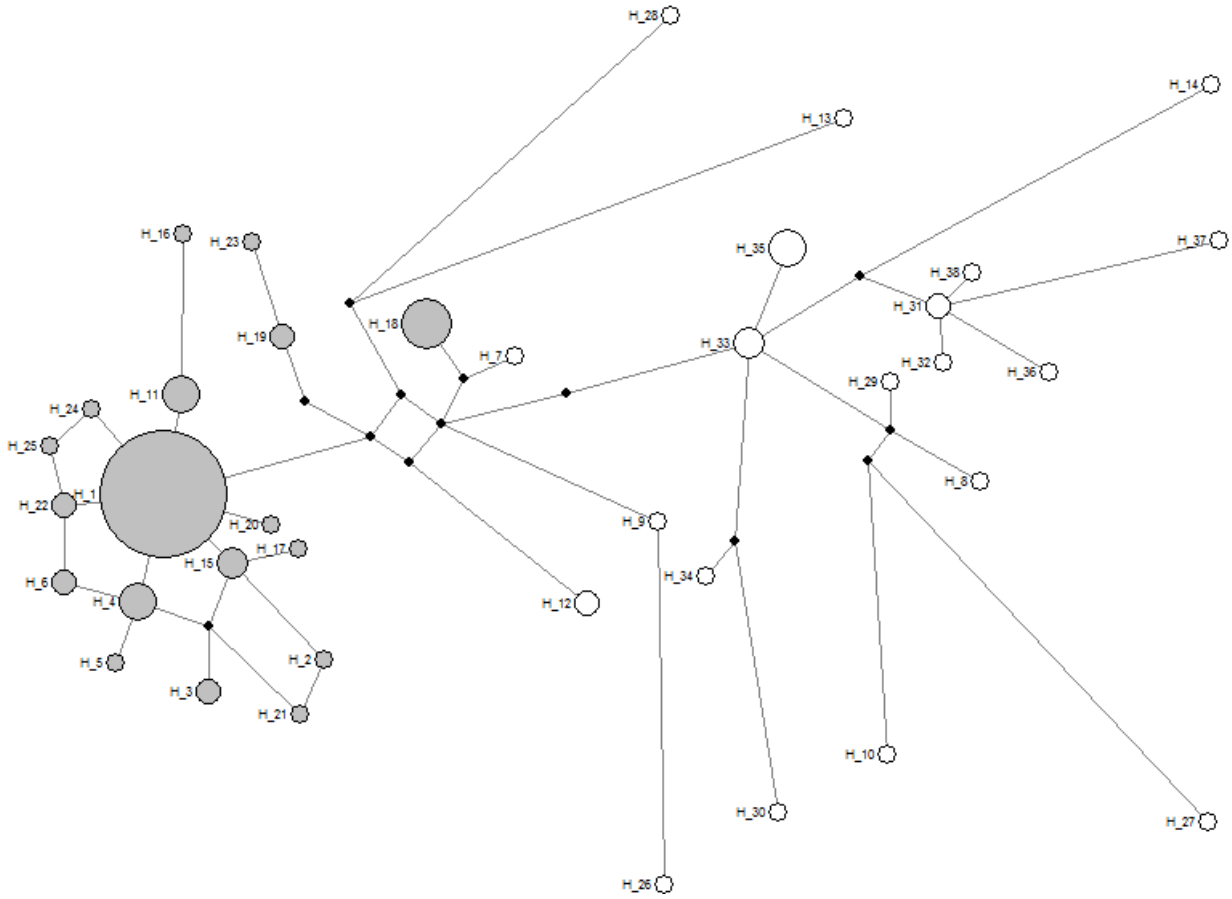


Figure 5.3 Median-joining network analysis of mtDNA control region haplotypes. Missing haplotypes are represented as black dots. Branch length is proportional to the number of mutational steps occurring between haplotypes. Grey filled circles are haplotypes present in introduced populations, white filled circles are haplotypes present in native populations. The size of circles is proportional to haplotype frequencies.

mtDNAcytb analyses

Genetic variation

Details on polymorphism in the 1140 bp sequence in each country, and levels of genetic diversity in each sampling site are presented in Table 5.3. Again, the Chinese populations contained higher levels of diversity than those in Great Britain and France, with *h* ranging from 0.891 (±0.063) in Yancheng to 1 (±0.177) in Dafeng and π x 10^3 ranging from9.17 (±4.97) in the Zhoushan Islands to 14.9 (±8.98) in Jishan, while in Great Britain *h* ranges from 0.361 (±0.083) in Whipsnade Park to 0.584 (±0.127) in Bure Marshes and π x 10^3 ranges from 0.74 (±0.6) in Whipsnade Park to 1.13 (±0.82) in Bure Marshes.

Population structure of native and introduced Chinese water deer populations

AMOVA on the British and Chinese populations (Table 5.4) showed while that the majority of the variation (~ 65-67%) occurs within populations (F_{ST} = 0.348, P < 0.001; Φ_{ST} = 0.332, P < 0.001), a significant proportions are also explained by differences among populations within countries (~ 17-21%), and among countries (14-16%). AMOVAs performed on British and Chinese populations grouped by country confirmed that the majority of the total variation in *cytb* occurs within populations, and that British populations are more differentiated than those that remain in China. Variation within Chinese populations explained ~87% of the total variance (i.e. Φ_{ST} = 0.13), compared to ~64% within British populations (Φ_{ST} = 0.360).

Genetic differentiation among populations

Pairwise comparisons of 10 populations using Φ_{ST} estimates are presented in Table 5.5, and F_{ST} estimates are presented in Appendix 5.1. The *cytb* dataset enables the inclusion of the Chinese

JishanIsland population in this analysis, as there is only one control region sequence for this population. For reasons mentioned above, only Φ_{ST} results are summarised here. Pairwise comparisons between the Chinese Jishan Island population and all other populations were all highly significant. Differentiation from the two mainland Chinese populations (Φ_{ST} ranging from 0.167 to 0.188) was lower than from the Zhoushan Islands, French and British populations (Φ_{ST} = 0.228 with ZS, Φ_{ST} = 0.264 with St. Croix Park and Φ_{ST} ranges between 0.223 with Yare Marshes and 0.644 with Woodwalton Fen). In contrast to the control region results, the Chinese Zhoushan Islands population is not significantly differentiated from the Dafeng or Yancheng mainland populations. However, it is differentiated from all French and British populations except East Norfolk and Yare. The Dafeng population is differentiated from all other populations except for Yancheng and Zhoushan Islands. However the Φ_{ST} value with Zhoushan Islands is higher than with Yancheng (Φ_{ST} 0.083 compared to 0.052). The St. Croix Park population in France is significantly differentiated from all Chinese populations except Yancheng, and all British populations except East Norfolk and Yare Marshes. For British populations, Whipsnade Park is also significantly differentiated from all other populations except East Norfolk or Yare. The deer in Bure Marshes are strongly differentiated from all of the Chinese, French and British populations. The small sample size of the East Norfolk and Yare populations regions may explain some of these unexpected results, especially as it tends to decrease the power of randomisation tests (Excoffier et al. 1992). A repeated analysis pooling these two neighbouring populations shows significant differentiation from all Chinese and French populations (not shown).

Haplotype relationships

The haplotype network (Figure 5.3) again shows a clear split between introduced and native populations of Chinese water deer, separated by a number of missing haplotypes. The most common haplotype (CYTB-H_6) is shared between the Whipsnade population and the Chinese Yancheng population. All populations are again characterised by unique haplotypes. There are a larger number of

mutational steps between haplotypes common in Chinese populations, and between Chinese and Korean haplotypes than there are between the British haplotypes. Only one haplotype (CYTB-H_1) is shared between the French and British populations, however on this occasion it is shared between St. Croix and the documented source population in Whipsnade.

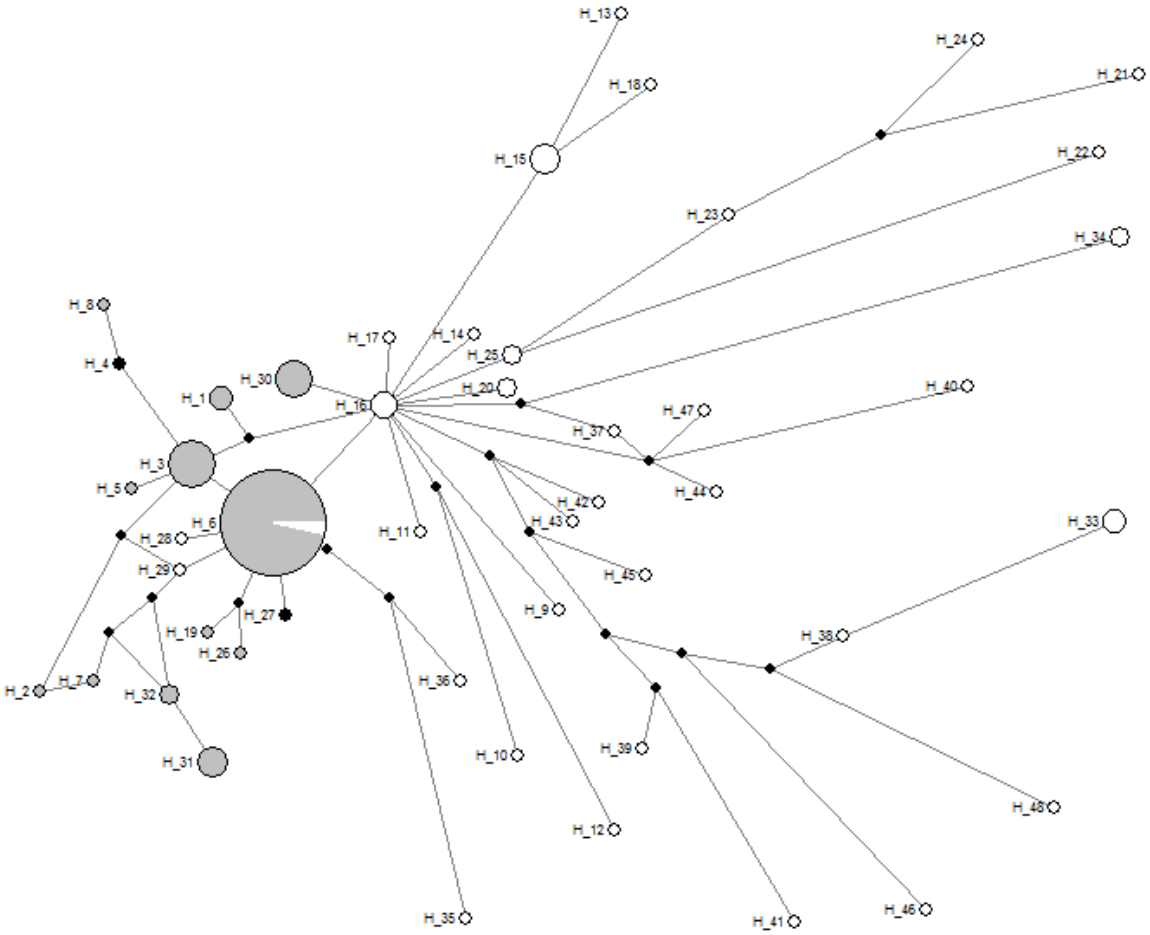


Figure 5.4 Median-joining network analysis of mtDNA*cytb* haplotypes. Missing haplotypes are represented as black dots. Branch length is proportional to the number of mutational steps occurring between haplotypes. Grey filled circles are haplotypes present in introduced populations, white filled circles are haplotypes present in native populations. The size of circles is proportional to haplotype frequencies. Haplotype 'CYTB-H_6' is present in both native and introduced populations, the circle is filled proportionally.

Table 5.3 Sample sites of native and introduced populations. Numbers of individuals examined (n), and measures of mtDNA diversity: haplotype diversity (*h*), mean number of pairwise differences (*k*), nucleotide diversity (π). GB = Great Britain.

mtDNA sequence	Country	Location	n	# Haplotypes	*h*	*k*	$\pi \times 10^3$
control region	China	Dafeng	4	4	1 (±0.177)	15.8 (±9.00)	17 (±11.5)
		Jishan	1	1	na	na	na
		Yancheng	6	5	0.933 (±0.122)	26.5 (±13.6)	28.4 (±16.8)
		Zhoushan Islands	15	9	0.905 (±0.054)	6.46 (±3.24)	6.95 (±3.91)
	France	St. Croix	4	3	0.833 (±0.222)	5.5 (±3.34)	5.93 (±4.3)
	GB	Bure Marshes	8	6	0.893 (±0.111)	2.36 (±1.43)	2.4 (±1.68)
		East Norfolk	4	1	na	na	na
		Whipsnade	43	4	0.391 (±0.084)	3.6 (±1.86)	3.88 (±2.22)
		Woodwalton Fen	15	7	0.724 (±0.121)	2.76 (±1.55)	2.98 (±1.87)
		Yare Marshes	3	3	1 (±0.272)	2 (±1.51)	2.16 (±2.03)
cytb	China	Dafeng	4	4	1 (±0.177)	12.1 (±6.97)	10.6 (±7.29)
		Jishan	6	5	0.933 (±0.122)	17.1 (±8.87)	14.9 (±8.98)
		Yancheng	11	6	0.891 (±0.063)	16.7 (±8.06)	14.6 (±7.98)
		Zhoushan Islands	15	13	0.971 (±0.039)	10.4 (±5.05)	9.17 (±4.97)
	France	St. Croix	5	5	1 (±0.126)	2.88 (±1.81)	2.53 (±1.86)
	GB	Bure Marshes	20	7	0.584 (±0.127)	1.28 (±0.839)	1.13 (±0.82)
		East Norfolk	4	2	0.5 (±0.265)	1.54 (±1.14)	1.35 (±1.19)
		Whipsnade	42	3	0.361 (±0.083)	0.844 (±0.611)	0.74 (±0.6)
		Woodwalton Fen	20	3	0.532 (±0.1)	1.16 (±0.78)	1.02 (±0.77)
		Yare Marshes	4	1	na	na	na

Table 5.4 Hierarchical AMOVA results for Chinese water deer populations in China and Great Britain.

mtDNA sequence	Source of variation	Total variance (%)	Fixation indices based on haplotype frequencies	*P* value	Total variance (%)	Fixation indices based on genetic distance	*P* value
control region	Among countries	13.7	F_{CT} = 0.137	0.123	46.8	Φ_{CT} = 0.468	0.127
	Among populations within countries	20.0	F_{SC} = 0.231	< 0.001	11.0	Φ_{SC} = 0.208	< 0.001
	Within populations	66.4	F_{ST} = 0.336	< 0.001	42.1	Φ_{ST} = 0.579	< 0.001
	Among populations within GB	29.7	F_{ST} = 0.297	< 0.001	18.4	Φ_{ST} = 0.184	0.009
	Within GB populations	70.3			81.6		
	Among populations within China	6.8	F_{ST} = 0.06	0.024	21.2	Φ_{ST} = 0.211	< 0.001
	Within Chinese populations	93.2			78.8		
cytb	Among countries	13.5	F_{CT} = 0.135	0.086	16.0	Φ_{CT} = 0.160	0.014
	Among populations within countries	21.3	F_{SC} = 0.246	< 0.001	17.3	Φ_{SC} = 0.206	< 0.001
	Within populations	65.2	F_{ST} = 0.348	< 0.001	66.8	Φ_{ST} = 0.332	< 0.001
	Among populations within GB	34.6	F_{ST} = 0.346	< 0.001	36.0	Φ_{ST} = 0.360	< 0.001
	Within GB populations	65.4			64.0		
	Among populations within China	5.8	F_{ST} = 0.058	< 0.001	13.1	Φ_{ST} = 0.130	< 0.001
	Within Chinese populations	94.2			87.0		

Table 5.5 Pairwise Φ_{ST} values between populations from control region (below diagonal) and *cytb* data (above diagonal). P values: * = significant at P < 0.05, ** = significant at P < 0.01, *** = significant at P < 0.001.

Population		1	2	3	4	5	6	7	8	9	10
1	China - Dafeng	-	0.167 *	0.052	0.083	0.176 **	0.535 ***	0.143 *	0.486 ***	0.556 ***	0.149 *
2	China - Jishan	na	-	0.188 *	0.228 **	0.264 *	0.561 ***	0.222 *	0.538 ***	0.644 ***	0.223 *
3	China - Yancheng	-0.070	na	-	0.064	0.048	0.292 ***	0.010	0.233 ***	0.297 ***	-0.010
4	China - Zhoushan Islands	0.239 **	na	0.253 ***	-	0.141 *	0.362 ***	0.106	0.303 ***	0.339 ***	0.087
5	France - St. Croix	0.453 *	na	0.167 **	0.702 ***	-	0.378 ***	-0.066	0.182 *	0.199 *	-0.060
6	GB - Bure Marshes	0.583 **	na	0.301 ***	0.735 ***	0.239 *	-	0.403 **	0.449 ***	0.463 ***	0.392 ***
7	GB - East Norfolk	0.519 *	na	0.161 *	0.730 ***	0.348	0.491 ***	-	0.185	0.151	0.000
8	GB - Whipsnade	0.526 **	na	0.354 **	0.660 ***	0.202	0.213 **	0.222	-	0.248 ***	0.097
9	GB - Woodwalton Fen	0.673 ***	na	0.393 ***	0.764 ***	0.212 *	0.053	0.438 **	0.131 *	-	-0.035
10	GB - Yare Marshes	0.411 *	na	0.080	0.695 ***	0.214	0.278 *	0.727 *	0.096	0.168	-

Discussion

Genetic diversity of native and introduced populations

The Chinese populations contain higher levels of intra-population genetic diversity than the British populations for both mtDNA markers, whereas the Zhoushan Islands population tends to have higher haplotype but lower nucleotide diversity. Levels of haplotype and nucleotide diversity at the mtDNA control region are also comparatively higher in Chinese water deer than other rare deer in China e.g. the Chinese sika (*Cervusnippon*) (Wu *et al.* 2004; Lu *et al.* 2006), black muntjac (*Muntiacuscrinifrons*) (Wu & Fang 2005), and Siberian roe deer (*Capreoluspygargus)* in Northeastern China (Xiao *et al.* 2007). The total current number of native Chinese water deer is estimated at below 5,000, while in the 1980s Sheng & Lu (1985) estimated the total number killed by hunting to be in the thousands. Variability is rapidly lost in small populations after a reduction in size, while in larger populations it takes a severe bottleneck to be maintained over a large number of generations to significantly reduce levels of diversity (Nei *et al.* 1975). The high levels of diversity despite relatively low numbers of present-day Chinese water deer suggest that there was a very recent sharp decline in population size in China.

In Great Britain, lower levels of diversity are likely to be the result of a bottleneck during introductions. Chinese water deer in Great Britain were largely confined to two semi-captive sites until the 1940s. The fact that the Whipsnade population appears to have the lowest level of diversity in Great Britain is puzzling, as this park has been the source of translocations to other sites which now have higher levels of diversity (e.g. Woodwalton Fen). This may be explained by the smaller sample taken from Woodwalton Fen giving an inflated measure of diversity. Assuming that the sampling is representative, the low levels of diversity in Whipsnade Zoo may be explained by the finding that an isolated population of white-tailed deer (*Odocoileusvirginianus*) established from small translocations shows signatures of founder effects (DeYoung *et al.* 2003), in contrast to rapidly expanding mammal populations retaining higher levels of genetic diversity (Zenger *et al.* 2003). Additionally, Chinese water deer population sizes tend to increase rapidly then plateau in semi-captivity (Dubost *et al.* 2008).

These observations may help to explain the differing levels of diversity between the bevy of semi-captive deer at Whipsnade and other wild British populations.

Haplotype distribution

No control region haplotypes are shared between any British or Chinese individuals (Figure 5.2). However, the major *cytb* haplotype (Figure 5.3) is central in the network, is distributed across the Chinese mainland, British and French populations, and is the most common haplotype in the Whipsnade population. This suggests that there has been an expansion from a single source, consistent with the records of introduction to Whipsnade followed by escapes from Woburn and a translocation to Woodwalton Fen. The lack of shared haplotypes and presence of inferred mutational steps in the haplotype network suggests that the source population of the Chinese water deer imported to Great Britain is unlikely to be directly descended from any of the locations sampled in China. Since these Chinese populations sampled represent the vast majority of the remaining mainland distribution, and according to the haplotype network British deer are more closely related to them than to Zhoushan Islands populations, it suggests the source of British populations is from mainland China, and the ancestral population is now extinct. Given the historical context of the introduction it is most likely that the deer were sourced from around Shanghai.

Population structure and gene flow

In China the connectivity between populations has been reduced either by urbanisation and habitat destruction on the mainland (Xu et al. 1998), or long-standing geographical separation in the case of the Zhoushan archipelago, which has been separated from the mainland for thousands of years. In Great Britain, the populations' foundation via secondary translocations means that the distance between most populations has been determined by human intervention rather than natural dispersal. As

such, it would be misleading to attempt to explain the degree of contemporary differentiation between populations with geographic distance.

The control region results are consistent with the findings of Hu *et al.*(2006), showing significant differentiation between Zhoushan Islands and mainland populations. In contrast, the *cytb* results show no differentiation. This is unsurprising as the control region evolves much more rapidly than the protein-coding *cytb* gene. The lack of any shared haplotypes between the Zhoushan Islands and the mainland, combined with the highly significant Φ_{ST} value and the geographic isolation of the archipelago from the mainland strongly suggests that there is currently no gene flow between the mainland and island populations. The current reintroduction programme is releasing captive-bred CWD from the Zhoushan Islands in two areas near Shanghai. The results presented here, and from studies using polymorphic mtDNA and nuclear markers (Hu *et al.* 2006; Hu *et al.* 2007) indicate that mainland deer are significantly differentiated from Zhoushan Islands deer. The overall patterns of genetic diversity and differentiation in native and introduced ranges suggest a number of discrete populations that are isolated and could therefore be considered demographically independent. The habitat fragmentation and population decline are likely to be the key factors contributing to these patterns in China, while the secondary translocations and availability of suitable habitat are important in contributing to the observed patterns.

Limitations

Patterns of genetic structure observed from mtDNA data may be less strongly correlated with the true distribution if sex-biased dispersal occurs (Nussey *et al.* 2006). Examining nuclear loci should give a more accurate assessment of population structure, reveal a more nuanced picture of the partitioning of genetic variation within and between populations, and allow for a test of genetic bottlenecking in concordance with historical records. The recent discoveries of a large number of nuclear microsatellite

loci by two independent studies (Lee *et al.* 2011; Yu *et al.* 2011) make this is a feasible prospect provided that sufficient quantities of nuclear DNA could be extracted from non-invasive samples.

Additionally, small sample sizes for some populations also meant that the power to estimate accurate measures of differentiation was low for some comparisons. For example, the Yare site samples did not differ significantly from any Chinese populations (Φ_{ST} = 0.080), yet were found to differ significantly from the neighbouring East Norfolk site (Φ_{ST} = 0.727*).

Despite their limitations, these mtDNA sequence data reveal lower levels of genetic diversity in the British populations, significant differentiation between Chinese and British populations and among British populations, and that the source population of British deer is likely to be extinct. These findings are congruent with a rapid decline of population size and distribution in China, and reflect the expected outcome of a bottleneck during introduction to Great Britain and serial founder effects due to subsequent translocations.

Recommendations

On account of the high levels of genetic differentiation among the geographically separated Chinese populations and their sets of unique haplotypes, the mainland and Zhoushan Islands should be considered and managed as evolutionary significant units. However, caution should be exercised when defining conservation units using a neutral genetic marker, as genetic diversity at adaptive loci may not necessarily be predicted by variation at neutral loci (Fraser & Bernatchez 2001).

It has been recommended that a special mainland breeding centre be established (Hu *et al.* 2007). This is likely to be impractical due to the remaining mainland populations being protected (Dafeng and Yancheng), or so rare in the wild as to have not been sighted during extensive population surveys (M. Chen unpublished data). In Great Britain, the deer are abundant in semi-captivity and expanding their range rapidly in the wild (Ward *et al.* 2008). Again, these populations are significantly differentiated

from mainland populations, but according to the haplotype networks more closely related to them than to Zhoushan Islands populations. This suggests that the Chinese water deer residing in Great Britain may be the descendants of a now extinct Chinese mainland population - a reservoir of genetic variation that has been lost in the native range. In addition to focusing conservation efforts on the remaining mainland populations, and the ongoing release of Zhoushan Islands-descended deer, the feasibility of introducing Chinese water deer from British stocks to supplement levels of adaptive genetic variation in mainland populations should be examined.

Aside from the evidence that Chinese water deer have negligible commercial or environmental impact (Cooke 2011), calls for a cull of British populations should also be rejected on the basis that British populations are important for the international conservation of Chinese water deer. Populations are significantly differentiated from the rapidly declining native populations in China, could provide a viable source to supplement reintroduction efforts in the mainland. Additionally, the relatively low genetic diversity of British populations (many of which have undergone a series of recent bottlenecks) means that they should be given the time and space to recover without interference. A long term study of wolves recently demonstrated that population growth alone may not be sufficient to retain high levels of genetic diversity, especially when combined with low levels of gene flow (Jansson *et al.*2012). This is the scenario facing British populations, and highlights the importance of managing deer populations at a landscape level by promoting connectivity between isolated populations while minimising the occurrence of deer vehicle collisions (Corlatti *et al.* 2009).

The Chinese water deer is an exception among cervid species, possessing numerous ancestral traits and behaviours (Dubost *et al.* 2011). The rapidly expanding British populations are less genetically diverse than the declining populations in China, and are significantly differentiated from them. All of the populations examined here are characterised by unique haplotypes. Given that the abundance of Chinese water deer is strongly determined by habitat preference (Cooke & Farrell 1998) and that habitat destruction and fragmentation is occurring at an alarming rate in China, maximising adaptive variation in the remaining Chinese populations must be a priority for their conservation. This can be

achieved by increasing connectivity between the Dafeng and Yancheng populations, protecting the Jishan and Zhoushan Islands populations, and exploring the feasibility of reintroducing genetically distinct individuals from Great Britain to separate regions.

The reintroduction and maintenance of genetically diverse populations in the Chinese mainland, combined with appropriate management and stakeholder engagement, could help this primitive (Dubost *et al.* 2011), 'Vulnerable' (Harris & Duckworth 2008) and seemingly harmless (Cooke 2009) species to flourish once more in its native range.

Chapter 6: General discussion

Overview

In this thesis I have investigated the factors driving invasion in non-native deer, and applied molecular methods to better understand the population genetics of two deer species introduced to Great Britain. Building and analysing a dataset of global introduction outcomes of non-native deer introductions, I discovered factors promoting invasion success. I then examined the relationship between range expansion and population genetic parameters using the British roe deer as a study system. Next, I took a genome scan approach to search for signatures of selection in British and European populations. Using the Chinese water deer as a second study system, I conducted a conservation genetic analysis investigating the distribution of genetic variation in the introduced and native ranges with a view to making recommendations for the conservation of the species in its native range.

Invasion success of non-native deer species

In Chapter 2, I determined that different factors are important at different stages of the invasion process in non-native deer. Examining the outcomes of introduction events at the species and population levels while controlling for the confounding effects of common ancestry gave differing results.At the species level, reproductive characteristics and native range size were predictive of establishment, while a reproductive characteristic predicted spread. At the population level establishment success is primarily determined by the number of introduced individuals whereas breadth of habitat and diet predicted spread. With the increasing global impact of deer invasions, these findings can help inform international-level policy on preventing and controlling the spread of non-native deer into new areas. The factors examined here may not be the only ones associated with invasion, and future studies could look to include data on predation levels (Colautti *et al.* 2004) and socioeconomic factors in the analysis.

This approach of using historical records on introduction outcomes to determine which factors are predictive of success at different stages of the invasion process could be fruitfully applied to other species (Sol *et al.* 2008b). However, concerns about under-reporting of failed introductions (Rodriguez-Cabal *et al.* 2009) may limit the application to groups of taxa for which there are reliable records (e.g. large mammals).

Range expansion of roe deer in Great Britain

In Chapter 3, I investigated genetic structure and rates of range expansion in roe deer in Great Britain derived from refugia and reintroductions, and used these 'natural experiments' to assess empirical support for hypotheses regarding the influence of spatial expansion on the distribution of genetic variation. I found that rates of expansion varied widely between regions, and contrasting patterns of genetic diversity and structure within the ranges of separate expansions. Genetic diversity tended to decline with rate of range expansion, suggesting that peripheral populations are founded by low numbers of dispersers (Austerlitz *et al.* 1997) although there is very limited evidence for dispersal between existing roe populations (Baker & Hoelzel 2013). Given this, it appears that the majority of roe dispersal is occurring into new areas rather than between established populations.However, in an area of overlap between populations expanding from different sources, I found evidence for admixture. While admixture has been inferred to drive invasion success, this hypothesis is rarely tested directly (Handley *et al.* 2011). Demonstrating that admixed individuals differ from their parental populations in life history traits underlying invasive success (e.g. dispersal, fecundity) is required to meet the criteria for validating the hypothesis (Wolfe *et al.* 2007). Outbreeding depression is generally observed only when highly divergent populations interbreed (Edmands 2007), and previous studies have suggested that admixture between native and introduced stocks has resulted in higher population growth in red deer (Haanes *et al.* 2013) and white-tailed deer (DeYoung *et al.* 2003). Should this be the case in roe deer, the implication for management would be to prevent merging of populations derived from separate introductions, focusing on areas in the overlap zones between the expanding

ranges in the south of England. I also found evidence for limited gene flow, even between geographically close populations, indicating that strong population structuring has occurred during range expansion (Excoffier 2004) and suggesting that low levels of dispersal occur between newly colonised areas, which over time may result in local adaptation (Dlugosch & Parker 2008). Adaptation can occur rapidly in invasive species in response to environmental change (Prentis *et al.* 2008) and selection for dispersal is expected to act during range expansions (Travis & Dytham 2002).

This study forms a foundation for examining the genetic characteristics associated with range expansion of roe deer. I found a negative relationship between genetic diversity and rate of range expansion, although the explanatory power of the findings is limited by only being able to assess contemporary levels of genetic variation without taking into account how population genetic processes are affected by spatial and temporal environmental heterogeneity. Determining how the rate of environmental change influences range expansion and how rapidly adaptation can occur under such scenarios is important for predicting future colonisation patterns (Thomas *et al.* 2001; Hill *et al.* 2011). An extension to this work could identify barriers causing discontinuities and landscape features promoting connectivity between populations (Coulon *et al.* 2006; Storfer *et al.* 2010), but at a national level. Taking a landscape genetic approach could enable a more accurate assessment of the relationship between genetic variation and rate of range expansion, and better inform roe deer management plans (Zannese *et al.* 2006).

Evidence for adaptation to climate in European roe deer

In Chapter 4 I conducted a genome scan in European roe deer, finding loci potentially under selection in separate regions. Using multiple study sites and combining the use of different approaches (population and landscape genomics) I attempted to limit the proportion of false-positives detected from the data. Using outlier-based loci detection alongside regression analyses of AFLP allele frequency with environmental variables led to the identification of i) loci potentially under divergent

selection and ii) the selection pressures putatively acting upon this study system. The results provide evidence that adaptation to climate has occurred in the European roe deer, and identifies candidate loci which may be involved in the response to selective pressure. The lack of association of any loci with habitat types supports long-standing ecological evidence for behavioural plasticity (Hewison 2001), rather than adaptation or different genotypes, in explaining the ability of roe to rapidly colonise new habitats. However, there may be other biotic or abiotic factors contributing to the observed patterns (Manel *et al.* 2009; Segelbacher *et al.* 2010).

This study serves as a basis on which to build research to identify functional variants underlying adaptation to novel environments by roe deer. Further work in this area could focus on isolating, cloning and sequencing the most strongly supported candidate loci identified in Chapter 4. However, the cost of applying next-generation sequencing (NGS) technologies to non-model organisms is falling rapidly (Ekblom & Galindo 2011). Using NGS to generate a much more dense and informative picture of genome-wide variation , and applying new statistical methods capable of further reducing the false-positive rate may be a more valuable future direction to take in order to bridge the gap between detecting outlier loci identifying their functional significance in wild populations. The next step is to link genetic, phenotypic and environmental characteristics. This can be done by determining how the function of loci under selection affects the capacity of populations to respond to changing environmental conditions. With the ultimate aim of being able to predict how various populations will spread and respond to climate change, finding the genetic basis of traits associated with these traits is a major goal. Further investigation into adaptations in introduced and peripheral populations compared with ancestral and core populations, focused on characterising the most ecologically relevant loci (Stinchcombe & Hoekstra 2008; Poncet *et al.* 2010) will help improve predictions of future population growth and distribution, with important implications for management and conservation plans.

Conservation genetics of the Chinese water deer

In Chapter 5, I conducted a conservation genetic analysis of the Chinese water deer, comparing levels of diversity and partitioning of variation in native and introduced populations, in order to provide recommendations for the international conservation of the species. Using mtDNA data from non-invasively collected samples, I conducted the first population genetic study of Chinese water deer in their introduced and native ranges. I found high levels of diversity in native Chinese populations, significant differentiation between and within the Chinese and British populations, and that very few haplotypes are shared between deer in the native and introduced ranges. Taken together the results strongly suggest a severe and recent decline in numbers in the native range, and that there was a single source of introduction to Great Britain which is likely to be extinct. Given the context and timing of the introduction (in the 1870s from China to Britain), it is most likely that the deer were sourced from around Shanghai. Finding that the source population of the Chinese water deer in Britain is likely to be extinct should raise the profile of the semi-captive and wild populations in Britain as a valuable conservation resource. A reintroduction program aiming to establish wild populations in the areas around Shanghai is currently underway, using deer sourced from the Zhoushan Islands. I found that both introduced and Zhoushan Island populations have relatively low diversity and are differentiated from mainland populations. However, these results are based on maternally-inherited markers. The recent development of Chinese water deer-specific microsatellite markers (Lee *et al.* 2011) means that further studies with a more informative marker system are now viable. The next steps for this project are to assess whether the British or Korean populations are suitable for reintroduction, and if so whether they could supplement the existing reintroduction program or be used to found new populations in other parts of the ancestral range where the native deer are now extinct.

Conclusion

An integrated knowledge of the history, ecology and genetics of introduced species is important for predicting invasive success, and determining the influence of introduction, range expansion and adaptation on local populations. These studies have identified factors associated with invasion by non-native deer, and examined the genetics of introduced and expanding populations of two species in contrasting scenarios. The results of these investigations represent a basis upon which to advance our understanding of the roles of environmental and genetic variation in the establishment and spread of introduced deer species, and how this information can be used for the management of invasions and conservation of vulnerable species in a rapidly changing world.

References

Acevedo, P., Ward, A.I., Real, R. & Smith, G.C. (2010). Assessing biogeographical relationships of ecologically related species using favourability functions: a case study on British deer. *Diversity and Distributions*, 16, 515-528.

Ajmone-Marsan, P., Valentini, A., Cassandro, M., Vecchiotti-Antaldi, G., Bertoni, G. & Kuiper, M. (1997). AFLP™ markers for DNA fingerprinting in cattle. *Animal Genetics*, 28, 418-426.

Alacs, E.A., Spencer, P.B.S., de Tores, P.J. & Krauss, S.L. (2011). Population genetic structure of island and mainland populations of the quokka, Setonix brachyurus (Macropodidae): a comparison of AFLP and microsatellite markers. *Conservation Genetics*, 12, 297-309.

Allendorf, F. & Luikart, G. (2007). *Conservation and the genetics of populations.* Blackwell Publishing, Malden, MA.

Allendorf, F.W. (2003). Introduction: population biology, evolution, and control of invasive species. *Conservation Biology*, 17, 24–30.

Allendorf, F.W., Hohenlohe, P.A. & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics*, 11, 697-709.

Andersen, R. & Linnell, J.D.C. (1998). Ecological correlates of mortality of roe deer fawns in a predator-free environment. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, 76, 1217-1225.

Andersen, R. & Linnell, J.D.C. (2000). Irruptive potential in roe deer: density-dependent effects on body mass and fertility. *Journal of Wildlife Management*, 64, 698–706.

Andersen, R., Duncan, P. & Linnell, J. (1998). *The European roe deer: the biology of success*. Scandinavian University Press, Oslo.

Andersen, R., Gaillard, J.M., Linnell, J.D.C. & Duncan, P. (2000). Factors affecting maternal care in an income breeder, the European roe deer. *Journal of Animal Ecology*, 69, 672-682.

Andersen, R., Herfindal, I., Saether, B.E., Linnell, J.D.C., Odden, J. & Liberg, O. (2004). When range expansion rate is faster in marginal habitats. *Oikos*, 107, 210-214.

Antao, T. & Beaumont, M.A. (2011). Mcheza: a workbench to detect selection using dominant markers. *Bioinformatics*, 27, 1717-1718.

Arenas, M., Ray, N., Currat, M. & Excoffier, L. (2012). Consequences of Range Contractions and Range Shifts on Molecular Diversity. *Molecular Biology and Evolution*, 29, 207-218.

Austerlitz, F., JungMuller, B., Godelle, B. & Gouyon, P.H. (1997). Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology*, 51, 148-164.

Baker, K.H. & Hoelzel, A.R. (2013). Evolution of population genetic structure of the British roe deer by natural and anthropogenic processes (*Capreolus capreolus*). *Ecology and Evolution*, 3, 89-102.

Baker, K.H. & Hoelzel, A.R. (In press). Fluctuating asymmetry in populations of British roe deer

(*Capreolus capreolus*) following historical bottlenecks and founder events. *Mammalian Biology - Zeitschrift für Säugetierkunde*.

Baker, K.H. (2011). Population genetic history of the British roe deer (Capreolus capreolus) and its implications for diversity and fitness. In: *School of Biological and Biomedical Sciences*. Durham University.

Ball, M.C., Pither, R., Manseau, M., Clark, J., Petersen, S.D., Kingston, S. *et al.* (2007). Characterization of target nuclear DNA from faeces reduces technical issues associated with the assumptions of low-quality and quantity template. *Conservation Genetics*, 8, 577-586.

Bandelt, H.J., Forster, P. & Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16, 37-48.

Beaumont, M.A. & Balding, D.J. (2004). Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, 13, 969-980.

Beaumont, M.A. & Nichols, R.A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B-Biological Sciences*, 263, 1619-1626.

Beaumont, M.A. (1999). Detecting population expansion and decline using microsatellites. *Genetics*, 153, 2013-2029.

Beaumont, M.A. (2005). Adaptation and speciation: what can F-st tell us? *Trends in Ecology & Evolution*, 20, 435-440.

Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological*, 57, 289-300.

Bensch, S. & Akesson, M. (2005). Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology*, 14, 2899-2914.

Blackburn, T.M. & Duncan, R.P. (2001). Determinants of establishment success in introduced birds. *Nature*, 414, 195-197.

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H. *et al.* (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24, 127-135.

Bomford, M., Darbyshire, R.O. & Randall, L. (2009). Determinants of establishment success for introduced exotic mammals. *Wildlife Research*, 36, 192-202.

Bonin, A., Bellemain, E., Eidesen, P.B., Pompanon, F., Brochmann, C. & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, 13, 3261-3273.

Bonin, A., Ehrich, D. & Manel, S. (2007). Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. *Molecular Ecology*, 16, 3737-3758.

Bonin, A., Taberlet, P., Miaud, C. & Pompanon, F. (2006). Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (Rana temporaria). *Molecular Biology and Evolution*, 23, 773-783.

Bothwell, H., Bisbing, S., Therkildsen, N.O., Crawford, L., Alvarez, N., Holderegger, R. *et al.* (2013). Identifying genetic signatures of selection in a non-model species, alpine gentian (Gentiana nivalis L.), using a landscape genetic approach. *Conservation Genetics*, 14, 467-481.

Boyle AE, (1996) 'The Rio Convention on Biological Diversity', in M. Bowman and C. Redgwell (eds.) International Law and the Conservation of Biological Diversity, London, The Hague and Boston: Kluwer Law International, pp 33–49

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.

Buckley, J., Butlin, R.K. & Bridle, J.R. (2012). Evidence for evolutionary change associated with the recent range expansion of the British butterfly, Aricia agestis, in response to climate change. *Molecular Ecology*, 21, 267-280.

Bützler, W. (1988) Wasserhirsche. In: Grzimeks Enzyklopädie, Säugetiere, Band 5. Kindler Verlag, München, pp 198–199.

Caballero, A., Quesada, H. & Rolan-Alvarez, E. (2008). Impact of amplified fragment length polymorphism size homoplasy on the estimation of population genetic diversity and the detection of selective loci. *Genetics*, 179, 539-554.

Cap, H., Deleporte, P., Joachim, J. & Reby, D. (2008). Male vocal behavior and phylogeny in deer. *Cladistics*, 24, 917-931.

Chadwick, A.H., Ratcliffe, P.R. & Abernethy, K. (1996). Sika deer in Scotland: density, population size, habitat use and fertility – some comparisons with red deer. *Scottish Forestry*, 50, 8-16.

Chapman, N. (1995). Our neglected species. *Deer*, 9, 360-362.

Chapman, N., Harris, S. & Stanford, A. (1994). Reeves' Muntjac *Muntiacus reevesi* in Britain: their history, spread, habitat selection, and the role of human intervention in accelerating their dispersal. *Mammal Review*, 24, 113-160.

Chapman, N.G., Claydon, K., Claydon, M., Forde, P.G. & Harris, S. (1993). Sympatric populations of muntjac (*Muntiacus reevesi*) and roe deer (*Capreolus capreolus*) - a comparative analysis of their ranging behavior, social organization and activity. *Journal of Zoology*, 229, 623-640.

Clout, M.N. & Russell, J.C. (2008). The invasion ecology of mammals: a global perspective. *Wildlife Research*, 35, 180-184.

Colautti, R.I. (2006). Propagule pressure: a null model for biological invasions. *Biological Invasions*, 8, 1023-1037.

Colautti, R.I., Ricciardi, A., Grigorovich, I.A. & MacIsaac, H.J. (2004). Is invasion success explained by

the enemy release hypothesis? *Ecology Letters*, 7, 721-733.

Cooke, A. & Farrell, L. (1998). *Chinese Water Deer*. Mammal Society and British Deer Society, London.

Cooke, A. &Farrell,L. (2008) Chinese Water Deer. In S. Harris and D.W. Yalden (eds.) Mammals of the British Isles: Handbook; 4th edition. The Mammal Society, pp 617-622.

Cooke, A. (2011) 'GB Non-native Organism Risk Assessment for *Hydropotes inermis'*, www.nonnativespecies.org

Cooke, A.S. & Farrell, L. (2001). Impact of muntjac deer (Muntiacus reevesi) at Monks Wood National Nature Reserve, Cambridgeshire, eastern England. *Forestry*, 74, 241-250.

Cooke, A.S. & Lakhani, K.H. (1996). Damage to coppice regrowth by muntjac deer Muntiacus reevesi and protectin with electric fencing. *Biological Conservation*, 75, 231-238.

Cooke, A.S. (2009). Chinese water deer *Hydropotes inermis* in Great Britain. *International Urban Ecology Review*, 4, 32-43.

Corlatti, L., Hacklander, K. & Frey-Roos, F. (2009). Ability of Wildlife Overpasses to Provide Connectivity and Prevent Genetic Isolation. *Conservation Biology*, 23, 548-556.

Cote, S.D., Rooney, T.P., Tremblay, J.P., Dussault, C. & Waller, D.M. (2004). Ecological impacts of deer overabundance. *Annual Review of Ecology Evolution and Systematics*, 35, 113-147.

Coulon, A., Cosson, J.F., Angibault, J.M., Cargnelutti, B., Galan, M., Morellet, N. *et al.*(2004). Landscape connectivity influences gene flow in a roe deer population inhabiting a fragmented landscape: an individual-based approach. *Molecular Ecology*, 13, 2841-2850.

Coulon, A., Guillot, G., Cosson, J.F., Angibault, J.M.A., Aulagnier, S., Cargnelutti, B. *et al.*(2006). Genetic structure is influenced by landscape features: empirical evidence from a roe deer population. *Molecular Ecology*, 15, 1669-1679.

Courchamp, F., Chapuis, J.L. & Pascal, M. (2003). Mammal invaders on islands: impact, control and control impact. *Biological Reviews*, 78, 347-383.

Crawford, K.M. & Whitney, K.D. (2010). Population genetic diversity influences colonization success. *Molecular Ecology*, 19, 1253-1263.

Da Silva, A., Gaillard, J.-M., Yoccoz, N.G., Hewison, A.J.M., Galan, M., Coulson, T. *et al.* (2009). Heterozygosity-fitness correlations revealed by neutral and candidate gene markers in roe deer from a long-term study. *Evolution*, 63, 403-417.

Danilkin, A. (1996). *Behavioural Ecology of Siberian and European Roe Deer*. Chapman and Hall, London.

Debeffe, L., Morellet, N., Cargnelutti, B., Lourtet, B., Bon, R., Gaillard, J.-M. *et al.*(2012). Condition-dependent natal dispersal in a large herbivore: heavier animals show a greater propensity to disperse and travel further. *Journal of Animal Ecology*, 81, 1327-1337.

Defra (2004) The sustainable management of wild deer populations in England: An action plan.

Defraand Forestry Commission, England

DeYoung, R.W., Demarais, S., Honeycutt, R.L., Rooney, A.P., Gonzales, R.A. & Gee, K.L. (2003). Genetic consequences of white-tailed deer (Odocoileus virginianus) restoration in Mississippi. *Molecular Ecology*, 12, 3237-3252.

Dlugosch, K.M. & Parker, I.M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, 17, 431-449.

Dlugosch, K.M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, 17, 431-449.

Dolman, P.M. & Waber, K. (2008). Ecosystem and competition impacts of introduced deer. *Wildlife Research*, 35, 202-214.

Dubost, G., Charron, F., Courcoul, A. & Rodier, A. (2008). Population characteristics of a semi-free-ranging polytocous cervid, *Hydropotes inermis*. *Mammalia*, 72, 333-343.

Dubost, G., Charron, F., Courcoul, A. & Rodier, A. (2011). Social organization in the Chinese water deer, *Hydropotes inermis*. *Acta Theriologica*, 56, 189-198.

Dudgeon, D. (2010). Requiem for a river: extinctions, climate change and the last of the Yangtze. *Aquatic Conservation-Marine and Freshwater Ecosystems*, 20, 127-131.

Duncan, R.P. (2001). High predictability in introduction outcomes and the geographical range size of introduced Australian birds: a role for climate. *The Journal of Animal Ecology*, 70, 621-623.

Duncan, R.P., Blackburn, T.M. & Sol, D. (2003). The ecology of bird introductions. *Annual Review of Ecology Evolution and Systematics*, 34, 71-98.

Durand, E., Jay, F., Gaggiotti, O.E. & Francois, O. (2009). Spatial Inference of Admixture Proportions and Secondary Contact Zones. *Molecular Biology and Evolution*, 26, 1963-1973.

Earl, D.A. & vonHoldt, B.M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4, 359-361.

Eckert, C.G., Samis, K.E. & Lougheed, S.C. (2008). Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Molecular Ecology*, 17, 1170-1188.

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, 463-475.

Ekblom, R. & Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, 107, 1-15.

Elton, C.S. (1958). *The ecology of invasions by animals and plants*. University Chicago Press, Chicago.

Endler, J.A. (1977). Geographic variation, speciation, and clines. *Monographs in population biology*, 10, 1-246.

Estoup, A. & Guillemaud, T. (2010). Reconstructing routes of invasion using genetic data: why, how and so what? *Molecular Ecology*, 19, 4113-4130.

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14, 2611-2620.

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, 564-567.

Excoffier, L. & Ray, N. (2008). Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution*, 23, 347-351.

Excoffier, L. (2004). Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology*, 13, 853-864.

Excoffier, L., Foll, M. & Petit, R.J. (2009a). Genetic Consequences of Range Expansions. In: *Annual Review of Ecology Evolution and Systematics*, pp. 481-501.

Excoffier, L., Hofer, T. & Foll, M. (2009b). Detecting loci under selection in a hierarchically structured population. *Heredity*, 103, 285-298.

Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotypes - application to human mitochondrial-DNA restriction data. *Genetics*, 131, 479-491.

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

Falush, D., Stephens, M. & Pritchard, J.K. (2007). Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes*, 7, 574-578.

Feder, M.E. & Mitchell-Olds, T. (2003). Evolutionary and ecological functional genomics. *Nature Reviews Genetics*, 4, 651-657.

Feldhamer, G. & Armstrong, W. (1993). Interspecific competition between four exotic species and native artiodactyls in the United States. In: *Transaction of the North American Wildlife and Natural Resources Conference 58*, pp. 468-478.

Feldhamer, G.A. & Demarais, S. (2009). Free ranging and confined sika deer in North America: current status, biology and management. In: *Sika deer biology and management of native and introduced populations* (eds. McCullough, DR, Takatsuki, S & Kaji, K). Springer New York, USA, pp. 615-642.

Fernandez, M.H. & Vrba, E.S. (2005). A complete estimate of the phylogenetic relationships in Ruminantia: a dated species-level supertree of the extant ruminants. *Biological Reviews*, 80, 269-302.

Finnegan, L.A., Wilson, P.J., Price, G.N., Lowe, S.J., Patterson, B.R., Fortin, M.J. *et al.* (2012). The complimentary role of genetic and ecological data in understanding population structure: a case study using moose (Alces alces). *European Journal of Wildlife Research*, 58, 415-423.

Fitzpatrick, B.M., Fordyce, J.A., Niemiller, M.L. & Reynolds, R.G. (2012). What can DNA tell us about

biological invasions? *Biological Invasions*, 14, 245-253.

Foll, M. & Gaggiotti, O. (2008). A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics*, 180, 977-993.

Formia, A., Godley, B.J., Dontaine, J.F. & Bruford, M.W. (2006). Mitochondrial DNA diversity and phylogeography of endangered green turtle (Chelonia mydas) populations in Africa. *Conservation Genetics*, 7, 353-369.

Forsyth, D. & Duncan, R. (2001). Propagule size and the relative success of exotic ungulate and bird introductions to New Zealand. *American Naturalist*, 157, 583-595.

Forsyth, D.M., Coomes, D.A., Nugent, G. & Hall, G.M.J. (2002). Diet and diet preferences of introduced ungulates (Order : Artiodactyla) in New Zealand. *New Zealand Journal of Zoology*, 29, 323-343.

Forsyth, D.M., Duncan, R.P., Bomford, M. & Moore, G. (2004). Climatic suitability, life-history traits, introduction effort, and the establishment and spread of introduced mammals in Australia. *Conservation Biology*, 18, 557-569.

Forsyth, D.M., Wilmshurst, J.M., Allen, R.B. & Coomes, D.A. (2010). Impacts of introduced deer and extinct moa on New Zealand ecosystems. *New Zealand Journal of Ecology*, 34, 48-65.

Francois, O. & Durand, E. (2010). Spatially explicit Bayesian clustering models in population genetics. *Molecular Ecology Resources*, 10, 773-784.

Frankham, R. (1995). Conservation genetics. *Annual Review of Genetics*, 29, 305-327.

Frankham, R., Ballou, J.D. & Briscoe, D.A. (2002). *Introduction to conservation genetics*. Cambridge University Press.

Fraser, D.J. & Bernatchez, L. (2001). Adaptive evolutionary conservation: towards a unified concept for defining conservation units. *Molecular Ecology*, 10, 2741-2752.

Fritts, T.H. (1998). The role of introduced species in the degradation of island ecosystems: A case history of Guam. *Annual Review of Ecology and Systematics*, 29, 113-140.

Fuller, R.J. & Gill, R.M.A. (2001). Ecological impacts of increasing numbers of deer in British woodland. *Forestry*, 74, 193-199.

Gaillard, J.M., Boutin, J.M., Delorme, D., VanLaere, G., Duncan, P. & Lebreton, J.D. (1997). Early survival in roe deer: causes and consequences of cohort variation in two contrasted populations. *Oecologia*, 112, 502-513.

Gaillard, J.M., Delorme, D., Boutin, J.M., VanLaere, G. & Boisaubert, B. (1996). Body mass of roe deer fawns during winter in 2 contrasting populations. *Journal of Wildlife Management*, 60, 29-36.

Gaillard, J.M., Delorme, D., Boutin, J.M., Vanlaere, G., Boisaubert, B. & Pradel, R. (1993). Roe deer survival patterns - a comparative analysis of contrasting populations. *Journal of Animal Ecology*, 62, 778-791.

Gaillard, J.-M., Hewison, A.J.M., Bonenfant, C., Plard, F., Douhard, M. & Davison, R. (2013). How does climate change influence demographic processes of widespread species? Lessons from the

comparative analysis of contrasted populations of roe deer. *Ecology Letters*, 16, 48–57.

Gamfeldt, L., Wallén, J., Jonsson, P.R., Berntsson, K.M. & Jon, N.H. (2005). Increasing Intraspecific Diversity Enhances Settling Success in a Marine Invertebrate. *Ecology*, 86, 3219-3224.

Geist, V. (1999). *Deer of the World. Their Evolution, Behaviour, and Ecology*. Stackpole Books, Mechanisburg.

Gentile, G., Vernesi, C., Vicario, S., Pecchioli, E., Caccone, A., Bertorelle, G. *et al.*(2009). Mitochondrial DNA variation in roe deer (Capreolus capreolus) from Italy: Evidence of admixture in one of the last C. c. italicus pure populations from central-southern Italy. *Italian Journal of Zoology*, 76, 16-27.

Gilbert, C., Ropiquet, A. & Hassanin, A. (2006). Mitochondrial and nuclear phylogenies of Cervidae (Mammalia, Ruminantia): Systematics, morphology, and biogeography. *Molecular Phylogenetics and Evolution*, 40, 101-117.

Gill, R.M.A. & Morgan, G. (2010). The effects of varying deer density on natural regeneration in woodlands in lowland Britain. *Forestry*, 83, 53-63.

Gill, R.M.A. (1992). A Review of Damage by Mammals in North Temperate Forests: 1.Deer. *Forestry*, 65, 145-169.

Gompert, Z. & Buerkle, C.A. (2011). A Hierarchical Bayesian Model for Next-Generation Population Genomics. *Genetics*, 187, 903-917.

Goodman, S.J., Barton, N.H., Swanson, G., Abernethy, K. & Pemberton, J.M. (1999). Introgression through rare hybridization: A genetic study of a hybrid zone between red and sika deer (genus Cervus) in Argyll, Scotland. *Genetics*, 152, 355-371.

Groves, C.P. & Grubb, P. (1987). Relationships of living deer. In: *Biology and management of the Cervidae*. Washington, DC: Smithsonian Institution Press Washington, DC: Smithsonian Institution Press, pp. 21-59.

Grubb, P. (2005). *Mammal Species of the World. A Taxonomic and Geographic Reference*. John Hopkins University Press, Baltimore.

Guilloteau, P., Vitari, F., Meuth, V.M.-L., Le Normand, L., Rome, V., Savary, G. *et al.* (2012). Is there adaptation of the exocrine pancreas in wild animal? The case of the Roe Deer. *Bmc Veterinary Research*, 8.

Guo, Q.F. (2006). Intercontinental biotic invasions: what can we learn from native populations and habitats? *Biological Invasions*, 8, 1451-1459.

Haanes, H., Roed, K.H., Flagstad, O. & Rosef, O. (2010). Genetic structure in an expanding cervid population after population reduction. *Conservation Genetics*, 11, 11-20.

Haanes, H., Rosvold, J. & Roed, K.H. (2013). Non-indigenous introgression into the Norwegian red deer population. *Conservation Genetics*, 14, 237-242.

Haltenorth, T. (1963). *Klassifikation der Säugetiere: Artiodactyla*.

Hamel, S., Cote, S.D., Gaillard, J.-M. & Festa-Bianchet, M. (2009a). Individual variation in reproductive costs of reproduction: high-quality females always do better. *Journal of Animal Ecology*, 78, 143-151.

Hamel, S., Gaillard, J.M., Festa-Bianchet, M. & Cote, S.D. (2009b). Individual quality, early-life conditions, and reproductive success in contrasted populations of large herbivores. *Ecology*, 90, 1981-1995.

Hampton, J.O., Spencer, P.B.S., Alpers, D.L., Twigg, L.E., Woolnough, A.P., Doust, J. *et al.*(2004). Molecular techniques, wildlife management and the importance of genetic population structure and dispersal: a case study with feral pigs. *Journal of Applied Ecology*, 41, 735-743.

Handley, L.J.L., Estoup, A., Evans, D.M., Thomas, C.E., Lombaert, E., Facon, B. *et al.* (2011). Ecological genetics of invasive alien species. *Biocontrol*, 56, 409-428.

Harris, R.B. and Duckworth, J.W. (2008) Hydropotes inermis. In: IUCN 2012. IUCN Red List of Threatened Species. Version 2012.2. <www.iucnredlist.org> [Accessed 07 March 2013].

Hartl, D.L. & Clark, A.G. (1997). Principles of population genetics. Third edition. *Principles of population genetics. Third edition.*, i-xiii, 1-542.

Hassanin, A., Pasquet, E. & Vigne, J.-D. (1998). Molecular systematics of the subfamily caprinae (Artiodactyla, Bovidae) as determined from cytochrome b sequences. *Journal of Mammalian Evolution*, 5, 217-236.

Hemami, M.R., Watkinson, A.R. & Dolman, P.M. (2004). Habitat selection by sympatric muntjac (Muntiacus reevesi) and roe deer (Capreolus capreolus) in a lowland commercial pine forest. *Forest Ecology and Management*, 194, 49-60.

Hemami, M.R., Watkinson, A.R. & Dolman, P.M. (2005). Population densities and habitat associations of introduced muntjac Muntiacus reevesi and native roe deer Capreolus capreolus in a lowland pine forest. *Forest Ecology and Management*, 215, 224-238.

Herrmann, D., Poncet, B.N., Manel, S., Rioux, D., Gielly, L., Taberlet, P. *et al.* (2010). Selection criteria for scoring amplified fragment length polymorphisms (AFLPs) positively affect the reliability of population genetic parameter estimates. *Genome*, 53, 302-310.

Hewison, A.J.M. & Staines, B.W. (2008). Roe deer. In: *Mammals of the British Isles: Handbook; 4th edition* (eds. S., H & D.W., Y). The Mammal Society, pp. 605- 617.

Hewison, A.J.M. (1995). Isozyme variation in roe deer in relation to their population history in Britain. *Journal of Zoology*, 235, 279-288.

Hewison, A.J.M. (1997). Evidence for a genetic component of female fecundity in British roe deer from studies of cranial morphometrics. *Functional Ecology*, 11, 508-517.

Hewison, A.J.M., Angibault, J.M., Boutin, J., Bideau, E., Vincent, J.P. & Sempere, A. (1996). Annual variation in body composition of roe deer (Capreolus capreolus) in moderate environmental conditions. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, 74, 245-253.

Hewison, A.J.M., Vincent, J.P. & Reby, D. (1998). Social organisation of European roe deer. *European Roe Deer: the Biology of Success*, 189-219.

Hewison, A.J.M., Vincent, J.P., Joachim, J., Angibault, J.M., Cargnelutti, B. & Cibien, C. (2001). The effects of woodland fragmentation and human activity on roe deer distribution in agricultural landscapes. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, 79, 679-689.

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907-913.

Hewitt, G.M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, 58, 247-276.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965-1978.

Hill, J.K., Griffiths, H.M. & Thomas, C.D. (2011). Climate Change and Evolutionary Adaptations at Species' Range Margins. *Annual Review of Entomology, Vol 56*, 56, 143-159.

Hochberg, M.E. & Gotelli, N.J. (2005). An Invasions Special Issue. *Trends in Ecology & Evolution*, 20, 211-211.

Holand, O., Mysterud, A., Wannag, A. & Linnell, J.D.C. (1998). Roe deer in northern environments: Physiology and behaviour. *European Roe Deer: the Biology of Success*, 117-137.

Holderegger, R. & Wagner, H.H. (2008). Landscape genetics. *Bioscience*, 58, 199-207.

Hu, J., Fang, S.G. & Wan, Q.H. (2006). Genetic diversity of Chinese water deer (Hydropotes inermis inermis): Implications for conservation. *Biochemical Genetics*, 44, 161-172.

Hu, J., Pan, H.J., Wan, Q.H. & Fang, S.G. (2007). Nuclear DNA microsatellite analysis of genetic diversity in captive populations of Chinese water deer. *Small Ruminant Research*, 67, 252-256.

Hulme, P.E. (2006). Beyond control: wider implications for the management of biological invasions. *Journal of Applied Ecology*, 43, 835-847.

Ibrahim, K.M., Nichols, R.A. & Hewitt, G.M. (1996). Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, 77, 282-291.

Irvine, R.J., Broadmeadow, M., Gill, R.M.A. & Albon, S.D. (2007). Deer and global warming: how will climate change influence deer populations? *Deer*, 14, 34-37.

Jaksic, F.M., Iriarte, J.A., Jiménez, J.E. & Martínez, D.R. (2002). Invaders Without Frontiers: Cross-border Invasions of Exotic Mammals. *Biological Invasions*, 4, 157-173.

Janis, C.M. (1993). Tertiary mammal evolution in the context of changing climates, vegetation, and tectonic events. *Annual Review of Ecology and Systematics*, 24, 467-500.

Jansson, E., Ruokonen, M., Kojola, I. & Aspi, J. (2012). Rise and fall of a wolf population: genetic diversity and structure during recovery, rapid expansion and drastic decline. *Molecular Ecology*, 21, 5178-5193.

Jensen, J.L., Bohonak, A.J. & Kelley, S.T. (2005). Isolation by distance, web service. *Bmc Genetics*, 6.

Jepsen, J.U. & Topping, C.J. (2004). Modelling roe deer (Capreolus capreolus) in a gradient of forest fragmentation: behavioural plasticity and choice of cover. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, 82, 1528-1541.

Jeschke, J.M. & Strayer, D.L. (2006). Determinants of vertebrate invasion success in Europe and North America. *Global Change Biology*, 12, 1608-1619.

Jeschke, J.M. (2008). Across islands and continents, mammals are more successful invaders than birds. *Diversity and Distributions*, 14, 913-916.

Joost, S., Bonin, A., Bruford, M.W., DesprÉS, L., Conord, C., Erhardt, G. *et al.* (2007). A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, 16, 3955-3969.

Joost, S., Kalbermatten, M. & Bonin, A. (2008). Spatial analysis method(SAM): a software tool combining molecular and environmental data to identify candidate loci for selection. *Molecular Ecology Resources*, 8, 957-960.

Jump, A.S., Hunt, J.M., Martinez-Izquierdo, J.A. & Penuelas, J. (2006). Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in Fagus sylvatica. *Molecular Ecology*, 15, 3469-3480.

Kalinowski, S.T. (2011). The computer program STRUCTURE does not reliably identify the main genetic clusters within species: simulations and implications for human population structure. *Heredity*, 106, 625-632.

Katoh, K. & Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, 9, 286-298.

Katoh, K., Kuma, K., Toh, H. & Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, 33, 511-518.

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30, 3059-3066.

Kawecki, T.J. & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7, 1225-1241.

Kawecki, T.J. (2008). Adaptation to Marginal Habitats. *Annual Review of Ecology Evolution and Systematics*, 39, 321-342.

Keiper, R.R. (1985). Are sika deer responsible for the decline of white-tailed deer on Assateague Island, Maryland? *Wildlife Society Bulletin*, 13, 144-146.

Kim, B.J., Oh, D.H., Chun, S.H. & Lee, S.D. (2011). Distribution, density, and habitat use of the Korean water deer (Hydropotes inermis argyropus) in Korea. *Landscape and Ecological Engineering*, 7, 291-297.

King, T.L. & Burke, T. (1999). Special issue on gene conservation: Identification and management of genetic diversity. *Molecular Ecology*, 8, S1-S3.

Kjellander, P., Gaillard, J.M. & Hewison, A.J.M. (2006). Density-dependent responses of fawn cohort

body mass in two contrasting roe deer populations. *Oecologia*, 146, 521-530.

Koh, H.S., Lee, B.K., Wang, J.X., Heo, S.W. & Jang, K.H. (2009). Two Sympatric Phylogroups of the Chinese Water Deer (Hydropotes inermis) Identified by Mitochondrial DNA Control Region and Cytochrome b Gene Analyses. *Biochemical Genetics*, 47, 860-867.

Kolar, C.S. & Lodge, D.M. (2001). Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution*, 16, 199-204.

Kolbe, J., Glor, R., Rodríguez Schettino, L., Lara, A., Larson, A. & Losos, J. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, 431, 177-181.

Kolbe, J.J., Glor, R.E., Schettino, L.R., Lara, A.C., Larson, A. & Losos, J.B. (2007). Multiple sources, admixture, and genetic variation in introduced Anolis lizard populations. *Conservation Biology*, 21, 1612-1625.

Kolbe, J.J., Larson, A., Losos, J.B. & de Queiroz, K. (2008). Admixture determines genetic diversity and population differentiation in the biological invasion of a lizard species. *Biology Letters*, 4, 434-437.

Lande, R. & Shannon, S. (1996). The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, 50, 434-437.

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

Lande, R. (1993). Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *American Naturalist*, 142, 911-927.

Langbein, J. (2011) Monitoring reported deer road casualties and related accidents in England to 2010, Deer Initiative contract report for The Highways Agency

Lavergne, S. & Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 3883-3888.

Le Corre, V. & Kremer, A. (1998). Cumulative effects of founding events during colonisation on genetic diversity and differentiation in an island and stepping-stone model. *Journal of Evolutionary Biology*, 11, 495-512.

Lee, C.E. (2002). Evolutionary genetics of invasive species. *Trends in Ecology & Evolution*, 17, 386-391.

Lee, Y.S., Choi, S.K., An, J., Park, H.C., Kim, S.I., Min, M.S. *et al.* (2011). Isolation and characterization of 12 microsatellite loci from Korean water deer (Hydropotes inermis argyropus) for population structure analysis in South Korea. *Genes & Genomics*, 33, 535-540.

Lever, C. (1985). *Naturalized Mammals of the World*. Longman, Essex, UK.

Librado, P. & Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, 25, 1451-1452.

Lister, A.M., Grubb, P. & Sumner, S.R.M. (1998). Taxonomy, morphology and evolution of European

roe deer. *European Roe Deer: the Biology of Success*, 23-46.

Lockwood, J.L., Cassey, P. & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*, 20, 223-228.

Long, J.L. (2003). *Introduced Mammals of the World*. CSIRO, Victoria.

Lorenzini, R. & Lovari, S. (2006). Genetic diversity and phylogeography of the European roe deer: the refuge area theory revisited. *Biological Journal of the Linnean Society*, 88, 85-100.

Lowe, V.P.W. & Gardiner, A.S. (1975). Hybridization between Red deer (*Cervus elaphus*) and Sika deer (*Cervus nippon*) with particular reference to stocks in N.W. England. *Journal of Zoology*, 177, 553-566.

Lu, X.P., Wei, F.W., Li, M., Yang, G. & Liu, H. (2006). Genetic diversity among Chinese sika deer (Cervus nippon) populations and relationships between Chinese and Japanese sika deer. *Chinese Science Bulletin*, 51, 433-440.

Ludt, C.J., Schroeder, W., Rottmann, O. & Kuehn, R. (2004). Mitochondrial DNA phylogeography of red deer (Cervus elaphus). *Molecular Phylogenetics and Evolution*, 31, 1064-1083.

Luikart, G., England, P.R., Tallmon, D., Jordan, S. & Taberlet, P. (2003). The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*, 4, 981-994.

Lynch, M. & Milligan, B.G. (1994). Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, 3, 91-99.

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

Manel, S., Conord, C. & Despres, L. (2009). Genome scan to assess the respective role of host-plant and environmental constraints on the adaptation of a widespread insect. *Bmc Evolutionary Biology*, 9.

Manel, S., Joost, S., Epperson, B.K., Holderegger, R., Storfer, A., Rosenberg, M.S. *et al.* (2010). Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*, 19, 3760-3772.

Manel, S., Schwartz, M.K., Luikart, G. & Taberlet, P. (2003). Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, 18, 189-197.

Mariette, S., Chagne, D., Lezier, C., Pastuszka, P., Baffin, A., Plomion, C. *et al.* (2001). Genetic diversity within and among Pinus pinaster populations: comparison between AFLP and microsatellite markers. *Heredity*, 86, 469-479.

Mariette, S., Cottrell, J., Csaikl, U.M., Goikoechea, P., Konig, A., Lowe, A.J. *et al.* (2002). Comparison of levels of genetic diversity detected with AFLP and microsatellite markers within and among mixed Q-petraea (MATT.) LIEBL. and Q-robur L. stands. *Silvae Genetica*, 51, 72-79.

Maudet, C., Luikart, G., Dubray, D., Von Hardenberg, A. & Taberlet, P. (2004). Low genotyping error rates in wild ungulate faeces sampled in winter. *Molecular Ecology Notes*, 4, 772-775.

Meudt, H.M. & Clarke, A.C. (2007). Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends in Plant Science*, 12, 106-117.

Morellet, N., Klein, F., Solberg, E. & Andersen, R. (2011). The census and management of populations of ungulates in Europe. *Ungulate management in Europe: problems and practices.*, 106-143.

Moriarty, A. (2004). The liberation, distribution, abundance and management of wild deer in Australia. *Wildlife Research*, 31, 291-299.

Moser, B., Schutz, M. & Hindenlang, K.E. (2006). Importance of alternative food resources for browsing by roe deer on deciduous trees: The role of food availability and species quality. *Forest Ecology and Management*, 226, 248-255.

Mysterud, A. & Ostbye, E. (2006). Effect of climate and density on individual and population growth of roe deer Capreolus capreolus at northern latitudes: the Lier valley, Norway. *Wildlife Biology*, 12, 321-329.

Narum, S.R. & Hess, J.E. (2011). Comparison of F-ST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, 11, 184-194.

Narum, S.R., Campbell, N.R., Kozfkay, C.C. & Meyer, K.A. (2010). Adaptation of redband trout in desert and montane environments. *Molecular Ecology*, 19, 4622-4637.

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

Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, 39, 197-218.

Nielsen, R., Hellmann, I., Hubisz, M., Bustamante, C. & Clark, A.G. (2007). Recent and ongoing selection in the human genome. *Nature Reviews Genetics*, 8, 857-868.

Nies, G., Zachos, F.E. & Hartl, G.B. (2005). The impact of female philopatry on population differentiation in the European roe deer (Capreolus capreolus) as revealed by mitochondrial, DNA and allozymes. *Mammalian Biology*, 70, 130-134.

Nosil, P., Egan, S.P. & Funk, D.J. (2008). Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. *Evolution*, 62, 316-336.

Nosil, P., Funk, D.J. & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, 18, 375-402.

Nunes, V.L., Beaumont, M.A., Butlin, R.K. & Paulo, O.S. (2011). Multiple approaches to detect outliers in a genome scan for selection in ocellated lizards (Lacerta lepida) along an environmental gradient. *Molecular Ecology*, 20, 193-205.

Nunez, M.A., Bailey, J.K. & Schweitzer, J.A. (2010). Population, community and ecosystem effects of exotic herbivores: A growing global concern. *Biological Invasions*, 12, 297-301.

Nussey, D.H., Pemberton, J., Donald, A. & Kruuk, L.E.B. (2006). Genetic consequences of human management in an introduced island population of red deer (Cervus elaphus). *Heredity*, 97, 56-

65.

O'Corry-Crowe, G.M., Suydam, R.S., Rosenberg, A., Frost, K.J. & Dizon, A.E. (1997). Phylogeography, population structure and dispersal patterns of the beluga whale Delphinapterus leucas in the western Nearctic revealed by mitochondrial DNA. *Molecular Ecology*, 6, 955-970.

O'Dowd, D.J. (2003). Invasional 'meltdown' on an oceanic island. *Ecology Letters*, 6, 812–817.

Ohtaishi, N. & Gao, Y. (1990). A review of the distribution of all species of deer (Tragulidae, Moschidae and Cervidae) in China. *Mammal Review*, 20, 125-144.

Palsboll, P.J., Berube, M. & Allendorf, F.W. (2007). Identification of management units using population genetic data. *Trends in Ecology & Evolution*, 22, 11-16.

Papa, R., Troggio, M., Ajmone-Marsan, P. & Marzano, F.N. (2005). An improved protocol for the production of AFLP (TM) markers in complex genomes by means of capillary electrophoresis. *Journal of Animal Breeding and Genetics*, 122, 62-68.

Paradis, E. & Claude, J. (2002). Analysis of comparative data using generalized estimating equations. *Journal of Theoretical Biology*, 218, 175-185.

Paradis, E., Claude, J. & Strimmer, K. (2004). APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, 20, 289-290.

Peakall, R. & Smouse, P.E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, 28, 2537-2539.

Perez-Espona, S., Perez-Barberia, F.J., McLeod, J.E., Jiggins, C.D., Gordon, I.J. & Pemberton, J.M. (2008). Landscape features affect gene flow of Scottish Highland red deer (Cervus elaphus). *Molecular Ecology*, 17, 981-996.

Perez-Figueroa, A., Garcia-Pereira, M.J., Saura, M., Rolan-Alvarez, E. & Caballero, A. (2010). Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, 23, 2267-2276.

Pettorelli, N., Dray, S., Gaillard, J.M., Chessel, D., Duncan, P., Illius, A. *et al.* (2003a). Spatial variation in springtime food resources influences the winter body mass of roe deer fawns. *Oecologia*, 137, 363-369.

Pettorelli, N., Gaillard, J.M., Duncan, P., Maillard, D., Van Laere, G. & Delorme, D. (2003b). Age and density modify the effects of habitat quality on survival and movements of roe deer. *Ecology*, 84, 3307-3316.

Pimentel, D. (2001). Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems & Environment*, 84, 1-20.

Pimentel, D. (2002). *Biological invasions: economic and environmental costs of alien plant, animal, and microbe species*. CRC Press Inc., Boca Raton; USA.

Pitra, C., Fickel, J., Meijaard, E. & Groves, P.C. (2004). Evolution and phylogeny of old world deer. *Molecular Phylogenetics and Evolution*, 33, 880-895.

Pocock RI (1910) On the Specialized Cutaneous Glands of Ruminants. Proc Zoo Soc 80:840-845

Pollard, E. & Cooke, A.S. (1994). Impact of muntjac deer (*Muntiacus reevesi*) on egg-laying sites of the white admiral butterfly (*Ladoga camilla*) in a Cambridgeshire wood. *Biological Conservation*, 70, 189-191.

Poncet, B.N., Herrmann, D., Gugerli, F., Taberlet, P., Holderegger, R., Gielly, L. *et al.*(2010). Tracking genes of ecological relevance using a genome scan in two independent regional population samples of Arabis alpina. *Molecular Ecology*, 19, 2896-2907.

Posada, D. & Crandall, K.A. (2001). Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology & Evolution*, 16, 37-45.

Prentis, P.J., Wilson, J.R.U., Dormontt, E.E., Richardson, D.M. & Lowe, A.J. (2008). Adaptive evolution in invasive species. *Trends in Plant Science*, 13, 288-294.

Price, S.A., Bininda-Emonds, O.R.P. & Gittleman, A.L. (2005). A complete phylogeny of the whales, dolphins and even-toed hoofed mammals (Cetartiodactyla). *Biological Reviews*, 80, 445-473.

Prior, R. (1995). *The roe deer: conservation of a native species*. Swan Hill Press.

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

Pruett, C.L., Tanksley, S.M., Small, M.F. & Taylor, J.F. (2011). The Effects of Range Expansion on the Population Genetics of White-Winged Doves in Texas. *American Midland Naturalist*, 166, 415-425.

Putman, R.J. & Flueck, W.T. (2011). Intraspecific variation in biology and ecology of deer: magnitude and causation. *Animal Production Science*, 51, 277-291.

Putman, R.J. & Moore, N.P. (1998). Impact of deer in lowland Britain on agriculture, forestry and conservation habitats. *Mammal Review*, 28, 141-163.

Putman, R.J. & Ward, A. (2010) *Predicting abundance and distribution of wild deer in England*. Deer Initiative contract report for Natural England

Putman, R.J. (1988). *The Natural History of Deer*. Christopher Helm, London.

Putman, R.J., Langbein, J., Hewison, A.J.M. & Sharma, S.K. (1996). Relative roles of density-dependent and density-independent factors in population dynamics of British deer. *Mammal Review*, 26, 81-101.

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Randi, E., Alves, P.C., Carranza, J., Milosevic-Zlatanovic, S., Sfougaris, A. & Mucci, N. (2004). Phylogeography of roe deer (Capreolus capreolus) populations: the effects of historical genetic subdivisions and recent nonequilibrium dynamics. *Molecular Ecology*, 13, 3071-3083.

Randi, E., Mucci, N., Pierpaoli, M. & Douzery, E. (1998). New phylogenetic perspectives on the Cervidae (Artiodactyla) are provided by the mitochondrial cytochrome b gene. *Proceedings of*

the Royal Society of London Series B-Biological Sciences, 265, 793-801.

Ratcliffe, P.R. (1987). Distribution and current status of sika deer, *Cervus nippon*, in Great Britain. *Mammal Review*, 17, 39.

Reed, D.H. & Frankham, R. (2003). Correlation between fitness and genetic diversity. *Conservation Biology*, 17, 230-237.

Reichard, S.H. & White, P.S. (2003). Invasion biology: An emerging field of study. *Annals of the Missouri Botanical Garden*, 90, 64-66.

Rejmanek, M. (1996). What Attributes Make Some Plant Species More Invasive? *Ecology*, 77, 1655–1661.

Relva, M.A., Nunez, M.A. & Simberloff, D. (2010). Introduced deer reduce native plant cover and facilitate invasion of non-native tree species: evidence for invasional meltdown. *Biological Invasions*, 12, 303-311.

Reznick, D.N. & Ghalambor, C.K. (2001). The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica*, 112, 183-198.

Robertson, A. (1975). Gene frequency distributions as a test of selective neutrality. *Genetics*, 81, 775-785.

Robinson, R.A. & Sutherland, W.J. (2002). Post-war changes in arable farming and biodiversity in Great Britain. *Journal of Applied Ecology*, 39, 157-176.

Rodriguez-Cabal, M.A., Barrios-Garcia, M.N. & Simberloff, D. (2009). Across island and continents, mammals are more successful invaders than birds (Reply). *Diversity and Distributions*, 15, 911-912.

Ronquist, F. & Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, 19, 1572-1574.

Royo, L.J., Pajares, G., Alvarez, I., Fernandez, I. & Goyache, F. (2007). Genetic variability and differentiation in Spanish roe deer (Capreolus capreolus): A phylogeographic reassessment within the European framework. *Molecular Phylogenetics and Evolution*, 42, 47-61.

Sabeti, P.C., Schaffner, S.F., Fry, B., Lohmueller, J., Varilly, P., Shamovsky, O. *et al.* (2006). Positive natural selection in the human lineage. *Science*, 312, 1614-1620.

Sagarin, R.D. & Gaines, S.D. (2002). The 'abundant centre' distribution: to what extent is it a biogeographical rule? *Ecology Letters*, 5, 137-147.

Sagarin, R.D., Gaines, S.D. & Gaylord, B. (2006). Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology & Evolution*, 21, 524-530.

Sakaguchi, S., Takeuchi, Y., Yamasaki, M., Sakurai, S. & Isagi, Y. (2011). Lineage admixture during postglacial range expansion is responsible for the increased gene diversity of Kalopanax

septemlobus in a recently colonised territory. *Heredity*, 107, 338-348.

Sakai, A.K. (2001). The population biology of invasive species. *Annual Review of Ecology, Evolution, and Systematics*, 32, 305-332.

Saunders, D.A., Hobbs, R.J. & Margules, C.R. (1991). Biological consequences of ecosystem fragmentation: a review. *Conservation Biology*, 5, 18-32.

Schmidt, P.S., Serrao, E.A., Pearson, G.A., Riginos, C., Rawson, P.D., Hilbish, T.J. *et al.* (2008). Ecological genetics in the North Atlantic: environmental gradients and adaptation at specific loci. *Ecology*, 89, S91-S107.

Schoville, S.D., Bonin, A., Francois, O., Lobreaux, S., Melodelima, C. & Manel, S. (2012). Adaptive Genetic Variation on the Landscape: Methods and Cases. In: *Annual Review of Ecology, Evolution, and Systematics, Vol 43* (ed. Futuyma, DJ), pp. 23-43.

Segelbacher, G., Cushman, S.A., Epperson, B.K., Fortin, M.-J., Francois, O., Hardy, O.J. *et al.* (2010). Applications of landscape genetics in conservation biology: concepts and challenges. *Conservation Genetics*, 11, 375-385.

Senn, H.V. & Pemberton, J.M. (2009). Variable extent of hybridization between invasive sika (Cervus nippon) and native red deer (C-elaphus) in a small geographical area. *Molecular Ecology*, 18, 862-876.

Serrano Ferron, E., Verheyden, H., Hummel, J., Cargnelutti, B., Lourtet, B., Merlet, J. *et al.*(2012). Digestive plasticity as a response to woodland fragmentation in roe deer. *Ecological Research*, 27, 77-82.

Sexton, J.P., McIntyre, P.J., Angert, A.L. & Rice, K.J. (2009). Evolution and Ecology of Species Range Limits. In: *Annual Review of Ecology Evolution and Systematics*, pp. 415-436.

Sheng, H. L., and Lu, H. J. (1985). Cervid resources in subtropical and tropical areas of China. *J. East China Norm. Univ., Sci. Ed.* 1, 96–104 (in Chinese).

Simberloff, D. (1996). Hybridization between native and introduced wildlife species: Importance for conservation. *Wildlife Biology*, 33, 143-150.

Slatkin, M. & Excoffier, L. (2012). Serial Founder Effects During Range Expansion: A Spatial Analog of Genetic Drift. *Genetics*, 191, 171-181.

Slatkin, M. & Hudson, R.R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, 129, 555-562.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. *Science*, 236, 787-792.

Sokal, R.R. & Rohlf, F.J. (1995). *Biometry: the principles and practice of statistics in biological research. Third edition*.

Sol, D., Bacher, S., Reader, S.M. & Lefebvre, L. (2008a). Brain size predicts the success of mammal species introduced into novel environments. In: *Symposium of the American Society of*

Naturalists Christchurch, New Zealand, pp. S63-S71.

Sol, D., Duncan, R.P., Blackburn, T.M., Cassey, P. & Lefebvre, L. (2005). Big brains, enhanced cognition, and response of birds to novel environments. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 5460-5465.

Sol, D., Szekely, T., Liker, A. & Lefebvre, L. (2007). Big-brained birds survive better in nature. *Proceedings of the Royal Society B-Biological Sciences*, 274, 763-769.

Sol, D., Vila, M. & Kuhn, I. (2008b). The comparative analysis of historical alien introductions. *Biological Invasions*, 10, 1119-1129.

Sommer, R.S., Fahlke, J.M., Schmolcke, U., Benecke, N. & Zachos, F.E. (2009). Quaternary history of the European roe deer Capreolus capreolus. *Mammal Review*, 39, 1-16.

Stinchcombe, J.R. & Hoekstra, H.E. (2008). Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, 100, 158-170.

Stoate, C., Baldi, A., Beja, P., Boatman, N.D., Herzon, I., van Doorn, A. *et al.* (2009). Ecological impacts of early 21st century agricultural change in Europe - A review. *Journal of Environmental Management*, 91, 22-46.

Storfer, A., Murphy, M.A., Spear, S.F., Holderegger, R. & Waits, L.P. (2010). Landscape genetics: where are we now? *Molecular Ecology*, 19, 3496-3514.

Strasburg, J.L., Sherman, N.A., Wright, K.M., Moyle, L.C., Willis, J.H. & Rieseberg, L.H. (2012). What can patterns of differentiation across plant genomes tell us about adaptation and speciation? *Philosophical Transactions of the Royal Society B-Biological Sciences*, 367, 364-373.

Strauss, S.Y., Webb, C.O. & Salamin, N. (2006). Exotic taxa less related to native species are more invasive. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 5841-5845.

Suarez, A.V. (1998). Effects of fragmentation and invasion on native ant communities in coastal southern California. *Ecology*, 79, 2041-2056.

Taberlet, P., Waits, L.P. & Luikart, G. (1999). Noninvasive genetic sampling: look before you leap. *Trends in Ecology & Evolution*, 14, 323-327.

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

Theodorou, K. & Couvet, D. (2006). On the expected relationship between inbreeding, fitness, and extinction. *Genetics Selection Evolution*, 38, 371-387.

Thomas, C.D., Bodsworth, E.J., Wilson, R.J., Simmons, A.D., Davies, Z.G., Musche, M. *et al.* (2001). Ecological and evolutionary processes at expanding range margins. *Nature*, 411, 577-581.

Travis, J.M.J. & Dytham, C. (2002). Dispersal evolution during invasions. *Evolutionary Ecology Research*, 4, 1119-1129.

Travis, J.M.J., Hammershoj, M. & Stephenson, C. (2005). Adaptation and propagule pressure determine invasion dynamics: insights from a spatially explicit model for sexually reproducing species. *Evolutionary Ecology Research*, 7, 37-51.

Vekemans, X., Beauwens, T., Lemaire, M. & Roldan-Ruiz, I. (2002). Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology*, 11, 139-151.

Vermeij, G.J. (1991). When Biotas Meet: Understanding Biotic Interchange. *Science*, 253, 1099-1104.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M. *et al.* (1995). AFLP - a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23, 4407-4414.

Vucetich, J.A. & Waite, T.A. (2003). Spatial patterns of demography and genetic processes across the species' range: Null hypotheses for landscape conservation genetics. *Conservation Genetics*, 4, 639-645.

Wahlstrom, L.K. & Liberg, O. (1995). Patterns of dispersal and seasonal migration in roe deer (*Capreolus capreolus*). *Journal of Zoology*, 235, 455-467.

Walker, B. & Steffen, W. (1997). An overview of the implications of global change for natural and managed terrestrial ecosystems. Conservation Ecology, 1, http://www.consecol.org/vol1/iss2/art2.

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-Revue Canadienne De Zoologie*, 85, 1182-1192.

Wang, M. & Schreiber, A. (2001). The impact of habitat fragmentation and social structure on the population genetics of roe deer (Capreolus capreolus L.) in Central Europe. *Heredity*, 86, 703-715.

Ward, A.I., Etherington, T. & Ewald, J. (2008). Five years of change. *Deer*, 14, 17-20.

Ward, A.I. (2005). Expanding ranges of wild and feral deer in Great Britain. *Mammal Review*, 35, 165-173.

Wark, A.R. & Peichel, C.L. (2010). Lateral line diversity among ecologically divergent threespine stickleback populations. *Journal of Experimental Biology*, 213, 108-117.

Wegmann, D., Currat, M. & Excoffier, L. (2006). Molecular diversity after a range expansion in heterogeneous environments. *Genetics*, 174, 2009-2020.

White, P.C.L., Ford, A.E.S., Clout, M.N., Engeman, R.M., Roy, S. & Saunders, G. (2008). Alien invasive vertebrates in ecosystems: pattern, process and the social dimension. *Wildlife Research*, 35, 171-179.

Whitehead, G.K. (1964). *The Deer of Great Britain and Ireland*. Routledge & Kegan Paul, London.

Whitlock, R., Hipperson, H., Mannarelli, M., Butlin, R.K. & Burke, T. (2008). An objective, rapid and reproducible method for scoring AFLP peak-height data that minimizes genotyping error.

Molecular Ecology Resources, 8, 725-735.

Whitney, K.D. & Gabler, C.A. (2008). Rapid evolution in introduced species, 'invasive traits' and recipient communities: challenges for predicting invasive potential. *Diversity and Distributions*, 14, 569-580.

Whitney, K.D. (2008). Rapid evolution in introduced species,'invasive traits' and recipient communities: challenges for predicting invasive potential. *Diversity & Distributions*, 14, 569-580.

Wier, B.S. & Cockerham, C.C. (1984). Estimating F statistics for the analysis of population structure. *Evolution*, 38, 1358-1370.

Williamson, M. (1996). The characters of successful invaders. *Biological Conservation*, 78, 163-170.

Williamson, M. (1999). Invasions. *Ecography*, 22, 5-12.

Wolfe, L.M., Blair, A.C. & Penna, B.M. (2007). Does intraspecific hybridization contribute to the evolution of invasiveness?: an experimental test. *Biological Invasions*, 9, 515-521.

Wright, S. (1938). Size of population and breeding structure in relation to evolution. *Science*, 87, 430-431.

Wu, H., Wan, Q.H. & Fang, S.G. (2004). Two genetically distinct units of the Chinese sika deer (Cervus nippon): analyses of mitochondrial DNA variation. *Biological Conservation*, 119, 183-190.

Wu, H.L. & Fang, S.G. (2005). Mitochondrial DNA genetic diversity of black muntjac (Muntiacus crinifrons), an endangered species endemic to China. *Biochemical Genetics*, 43, 407-416.

Xiao, C.-T., Zhang, M.-H., Fu, Y. & Koh, H.-S. (2007). Mitochondrial DNA distinction of northeastern China roe deer, Siberian roe deer, and European roe deer, to clarify the taxonomic status of northeastern China roe deer. *Biochemical Genetics*, 45, 93-102.

Xu, H. F., Zheng, X. Z., and Lu, H. J. (1998). Impact of human activities and habitat changes on distribution of Chinese water deer along the coast area in northern Jiangsu. *ActaTheriologicaSinica.* 18, 161–167 (in Chinese with English abstract).

Yang, C.-C., Ascunce, M.S., Luo, L.-Z., Shao, J.-G., Shih, C.-J. & Shoemaker, D. (2012). Propagule pressure and colony social organization are associated with the successful invasion and rapid range expansion of fire ants in China. *Molecular Ecology*, 21, 817-833.

Yang, F., O'Brien, P.C.M., Wienberg, J., Neitzel, H., Lin, C.C. & Ferguson-Smith, M.A. (1997). Chromosomal evolution of the Chinese muntjac (*Muntiacus reevesi*). *Chromosoma*, 106, 37-43.

Yu, J.-N., Won, C., Jun, J., Lim, Y. & Kwak, M. (2011). Fast and Cost-Effective Mining of Microsatellite Markers Using NGS Technology: An Example of a Korean Water Deer Hydropotes inermis argyropus. *Plos One*, 6.

Zachos, F.E., Hartl, G.B. & Suchentrunk, F. (2007). Fluctuating asymmetry and genetic variability in the roe deer (Capreolus capreolus): a test of the developmental stability hypothesis in mammals using neutral molecular markers. *Heredity*, 98, 392-400.

Zachos, F.E., Hmwe, S.S. & Hartl, G.B. (2006a). Biochemical and DNA markers yield strikingly

different results regarding variability and differentiation of roe deer (Capreolus capreolus, Artiodactyla : Cervidae) populations from northern Germany. *Journal of Zoological Systematics and Evolutionary Research*, 44, 167-174.

Zachos, F.E., Otto, M., Hmwe, S.S. & Hartl, G.B. (2006b). Population genetic investigations for identification of the roe deer (Capreolus capreolus Linnaeus, 1758; Artiodactyla: Cervidae) from Schleswig-Holstein. *Bonner Zoologische Beitraege*, 55, 89-94.

Zannese, A., Morellet, N., Targhetta, C., Coulon, A., Fuser, S., Hewison, A.J.M. *et al.*(2006). Spatial structure of roe deer populations: towards defining management units at a landscape scale. *Journal of Applied Ecology*, 43, 1087-1097.

Zenger, K.R., Richardson, B.J. & Vachot-Griffin, A.M. (2003). A rapid population expansion retains genetic diversity within European rabbits in Australia. *Molecular Ecology*, 12, 789-794.

Zhang, E., Teng, L. & Wu, Y. (2006). Habitat selection of the Chinese water deer (*Hydropotes inermis*) in Yancheng Reserve, Jiangsu Province. *Acta Theriologica Sinica*, 26, 49-53.

Zhang, F., Wang, Y.P. & Deng, H.W. (2008). Comparison of Population-Based Association Study Methods Correcting for Population Stratification. *Plos One*, 3, 7.

Zhivotovsky, L.A. (1999). Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, 8, 907-913.

Zhivotovsky, L.A., Bennett, L., Bowcock, A.M. & Feldman, M.W. (2000). Human population expansion and microsatellite variation. *Molecular Biology and Evolution*, 17, 757-767.

Zhu, H.G., Qin, P. & Wang, H. (2004). Functional group classification and target species selection for Yancheng Nature Reserve, China. *Biodiversity and Conservation*, 13, 1335-1353.

Appendices

Appendix 2.1 Trait data for introduced deer species

Genus	Species	Adult Body Mass (g)	Adult Length (mm)	Age at First Reproduction (d)	Diet Breadth*	Gestation Length (d)	Habitat Breadth**	Home Range (km^2)***
Alces	*alces*	461901	2930	1217	1	235	2[e]	71.75
Axis	*axis*	69500	1850[c]	450[c]	4[c]	227	4[e]	0.67
Axis	*porcinus*	37448	1225[c]	300[c]	3	220	4[e]	0.52
Capreolus	*capreolus*	22502	1230[c]	730	4[c]	196	4[e]	0.52
Capreolus	*pygargus*	41371	1320	-	-	290	4[e]	0.99
Cervus	*duvaucelii*	171224	1500	1095	2	245	4[e]	14.11
Cervus	*elaphus*	240867	2137	994	5	236	5[e]	58.42
Cervus	*nippon*	53000	1201	1095	2	224	4[e]	1.43
Cervus	*timorensis*	66376	1635	630[f]	2	250	4[e]	-
Cervus	*unicolor*	177523	2040	788	3	246	4[e]	-
Dama	*dama*	57225	1530[d]	1095	2	230	3[e]	0.71
Hydropotes	*inermis*	12760	890[d]	210[c]	2[c]	175	3[e]	0.32[a]
Muntiacus	*reevesi*	13500	1070[c]	168[f]	2	214	2[e]	0.12
Odocoileus	*hemionus*	84561	1517	548	2	203	5[e]	2.01
Odocoileus	*virginianus*	75901	1514	463	4	201	5[e]	1.99
Rangifer	*tarandus*	109089	2224	855[c]	3	223	3[e]	2771.22[b]

*Diet Breath: Number of categories eaten by a species. Categories defined as: Vertebrate, invertebrate, fruit, flowers/nectar/pollen, leaves/branches/bark, seeds, grass and roots/tubers.
**Habitat Breath: Number of types of habitat occupied by a species. Categories defined as: Grassland/scrubland; dense forest; desert; rocky; tundra; swamp.
***Home Range: Size of the area within which everyday activities of individuals or groups (of any type) are typically restricted.

[a]Sheng, H. & Ohtaishi, N. (1993) The status of deer in China. In: Deer of China: biology and management (eds. N. Ohtaishi & H. Sheng) pp. 1-11. Elsevier, Amsterdam.
[b]Ernest S.K.M. (2003) Life history characteristics of placental non-volant mammals. Ecology, 84:3402
[c]Long, J. L. (2003) Introduced Mammals Of The World: Their History, Distribution And Abundance. CSIRO Publishing, Collingwood, Victoria, Australia
[d]Jeschke, J.M. & Strayer, D.L. (2006) Determinants of vertebrate invasion success in Europe and North America. Glob Change Biol, 12:1608-1619
[e]Price, S. A. & Gittleman, J. L. (2007) Hunting to extinction: biology and regional economy influence extinction risk and the impact of hunting in artiodactyls. Proc R Soc B, 274:1845-1851
[f]Animal Diversity Web http://animaldiversity.ummz.umich.edu

Genus	Species	Home Range of Individual (km^2)†	Interbirth Interval (d)††	Litter Size [Fecundity]	Litters Per Year	Maximum Lifespan (m)	Neonate Body Mass (g)	Neonate Length (mm)
Alces	*alces*	73.26	365	1.25	1.7	324	13000	-
Axis	*axis*	2.67	426	1.01	1[c]	250	3157	-
Axis	*porcinus*	-	-	1	1[c]	240	2483	-
Capreolus	*capreolus*	0.49	365	1.79	1[b]	204	1209	-
Capreolus	*pygargus*	1.06	-	2	1[b]	206	1924	-
Cervus	*duvaucelii*	20	-	1	1[c]	276[b]	-	-
Cervus	*elaphus*	54.81	-	1.09	0.8[b]	322	8256	-
Cervus	*nippon*	0.62	-	1	1[c]	305	4277	570
Cervus	*timorensis*	-	366	1	1[f]	253[b]	2532[f]	-
Cervus	*unicolor*	-	-	1	1[c]	317	10270	-
Dama	*dama*	0.61	365	1	1[b]	300	4698	621
Hydropotes	*inermis*	-	365[a]	3	1[c]	144[b]	1016	-
Muntiacus	*reevesi*	0.11	215	0.98	1.5[b]	236[b]	1050	-
Odocoileus	*hemionus*	2.85	365	1.61	0.9[b]	264	3007	-
Odocoileus	*virginianus*	2.29	304	1.57	1[b]	276	2950	-
Rangifer	*tarandus*	2599.89	365	2	1[b]	242	5491	-

†Home Range of Individual: Size of the area within which everyday activities of individuals are typically restricted.
††Interbirth Interval: The length of time between successive births of the same female(s) after a successful or unspecified litter.

[a]Sheng, H. & Ohtaishi, N. (1993) The status of deer in China. In: Deer of China: biology and management (eds. N. Ohtaishi & H. Sheng) pp. 1-11. Elsevier, Amsterdam
[b]Ernest S.K.M. (2003) Life history characteristics of placental non-volant mammals. Ecology, 84:3402
[c]Long, J. L. (2003) Introduced Mammals Of The World: Their History, Distribution And Abundance. CSIRO Publishing, Collingwood, Victoria, Australia
[f]Animal Diversity Web http://animaldiversity.ummz.umich.edu

Genus	Species	Population Density (n/km^2)	Population Group Size‡	Sexual Maturity Age (d)	Social Group Size‡‡	Weaning Age (d)‡‡‡	Relative Brain Mass	Weaning Body Mass (g)
Alces	*alces*	0.4	-	668	1	99	0.00137[h]	86300
Axis	*axis*	6.69	-	544	16[c]	122	0.00205[h]	-
Axis	*porcinus*	16.5	-	464	1.8	180[b]	0.00473[h]	-
Capreolus	*capreolus*	7.61	-	401	1/ family groups (2-10)[c]	80	0.00669[h]	8692.5[b]
Capreolus	*pygargus*	0.09[g]	-	909	-	136	-	28750[b]
Cervus	*duvaucelii*	11.24	401	828	64.3	210[f]	0.00107[h]	-
Cervus	*elaphus*	2.9	-	660	17.5	105	0.00454[h]	43977
Cervus	*nippon*	0.75		518	7.5	198	0.00124[h]	28590[b]
Cervus	*timorensis*	-	-	666	6-100[c]	227	0.00256[h]	-
Cervus	*unicolor*	4.89	-	621	1[c]	210[b]	0.00214[h]	-
Dama	*dama*	21.99	-	543	10.5	178	0.00518[h]	16720
Hydropotes	*inermis*	-	1	293	1[c]	60[c]	0.00496[h]	-
Muntiacus	*reevesi*	0.67	-	295	1[c]	60[f]	0.00432[h]	-
Odocoileus	*hemionus*	16	-	527	up to 50[c]	73	0.00394[h]	23900
Odocoileus	*virginianus*	17.32	-	365	2.3	80	0.00806[h]	29986
Rangifer	*tarandus*	2.55	100000	759	-	137	0.00346[h]	24907

‡Population Group Size: Number of individuals, adults or definition unspecified in a group that spends the majority of their time in a 24 hour cycle together, measured over any duration of time, using non-captive populations.

‡‡Social Group Size: Number of individuals, adults or definition unspecified in a group that spends the majority of their time in a 24 hour cycle together where there is some indication that these individuals form a social cohesive unit, measured over any duration of time, using non-captive populations.

‡‡‡Weaning Age: Age when primary nutritional dependency on the mother ends and independent foraging begins.

[g]Danilkin, A. (1996) Behavioural Ecology of Siberian and European Roe Deer. Chapman and Hall, London.

[b]Ernest S.K.M. (2003) Life history characteristics of placental non-volant mammals. Ecology, 84:3402

[c]Long, J. L. (2003) Introduced Mammals Of The World: Their History, Distribution And Abundance. CSIRO Publishing, Collingwood, Victoria, Australia

[h]Pérez-Barbería, F. J. & Gordon, I. J. (2005) Gregariousness increases brain size in ungulates. Oecologia, 145:41–52

[f]Animal Diversity Web http://animaldiversity.ummz.umich.edu

Genus	Species	Native Grid Referenced Area (million km^2)[§]	Mean Human Population Density (n/km^2)[§§]	Biogeographic region[§§§]
Alces	*alces*	18.75[i]	-	1,2[c]
Axis	*axis*	2.04	361.5	3[c]
Axis	*porcinus*	2.02	144.4	3[c]
Capreolus	*capreolus*	6.78	101.8	1[c]
Capreolus	*pygargus*	9.82	26.6	1[c]
Cervus	*duvaucelii*	0.26	428.7	3[c]
Cervus	*elaphus*	14.17	60.8	1,2[c]
Cervus	*nippon*	2.71	331.6	1[c]
Cervus	*timorensis*	0.13	804.2	3[c]
Cervus	*unicolor*	7.05	218.9	3[c]
Dama	*dama*	3.08	128.4	1[c]
Hydropotes	*inermis*	1.14	373.9	1[c]
Muntiacus	*reevesi*	1.77	356.7	3[c]
Odocoileus	*hemionus*	6.31	14.4	2[c]
Odocoileus	*virginianus*	14.45	34.1	2,4[c]
Rangifer	*tarandus*	18.38	1.1	1,2[c]

[§]Native Grid Referenced Area: Total extent of a species range with a global equal-area projection (Mollweide).

[§§]Mean Human Population Density: Mean human population density (persons per km^2) using the Gridded Population of the World (GPW).

[§§§]Biogeographic Region: 1 - Palearctic; 2 - Nearctic; 3 - Indo-malayan; 4 - Neotropical; 5 - Holarctic; 6 - Nearctic and Neotropical.

[c]Long, J. L. (2003) Introduced Mammals Of The World: Their History, Distribution And Abundance. CSIRO Publishing, Collingwood, Victoria, Australia

[i]Morrison, J., Sechrest, W., Dinerstein, E., Wilcove, D. & Lamoreux, J. (2007) Persistence of large mammal faunas as indicators of global human impacts. J Mammalogy, 88:1363–1380

Base dataset compiled from Jones et al. (2009)[#].

[#]Jones et al. (2009) PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. Ecology 90:2648

Appendix 2.2 Deer introduction data

Genus	Species	Recipient Biome	Within same biome?	Island Introduction?	Propagule size*	Established?	Spread?
Alces	alces	Australian	no	Yes	4	No	NA
Alces	alces	Australian	no	Yes	10	Yes	No
Axis	axis	Australian	no	Yes	5	No	NA
Axis	axis	Australian	no	Yes	5	No	NA
Axis	axis	Australian	no	No	4	No	NA
Axis	axis	Australian	no	No	3	No	NA
Axis	axis	Australian	no	Yes	3	No	NA
Axis	axis	Australian	no	Yes	2	No	NA
Axis	axis	Australian	no	Yes	7	No	NA
Axis	axis	Oceanian	no	Yes	2	No	NA
Axis	axis	Palaearctic	no	No	20	No	NA
Axis	axis	Australian	no	Yes	7	Yes	No
Axis	axis	Australian	no	Yes	7	Yes	No
Axis	axis	Palaearctic	no	No	24	Yes	No
Axis	axis	Palaearctic	no	No	21	Yes	No
Axis	axis	Palaearctic	no	No	39	Yes	No
Axis	axis	Palaearctic	no	No	24	Yes	No
Axis	axis	Palaearctic	no	No	47	Yes	No
Axis	axis	Palaearctic	no	No	19	Yes	No
Axis	axis	Australian	no	No	22	Yes	Yes
Axis	axis	Nearctic	no	No	8	Yes	Yes
Axis	axis	Oceanian	no	Yes	12	Yes	Yes
Axis	axis	Oceanian	no	Yes	9	Yes	Yes
Axis	axis	Oceanian	no	Yes	8	Yes	Yes
Axis	porcinus	Australian	no	No	4	No	NA
Axis	porcinus	Australian	no	No	12	No	NA
Axis	porcinus	Australian	no	No	2	No	NA
Axis	porcinus	Australian	no	No	33	Yes	No
Axis	porcinus	Australian	no	No	7	Yes	No
Capreolus	capreolus	Nearctic	no	No	18	Yes	No
Capreolus	pygargus	Palaearctic	Yes	No	60	No	NA
Capreolus	pygargus	Palaearctic	Yes	No	59	Yes	Yes
Capreolus	pygargus	Palaearctic	Yes	No	15	Yes	Yes
Cervus	duvaucelii	Australian	no	No	2	No	NA
Cervus	elaphus	Australian	no	No	2	No	NA
Cervus	elaphus	Australian	no	Yes	3	No	NA
Cervus	elaphus	Australian	no	Yes	2	No	NA
Cervus	elaphus	Australian	no	Yes	2	No	NA
Cervus	elaphus	Australian	no	Yes	2	No	NA
Cervus	elaphus	Nearctic	Yes	Yes	8	No	NA
Cervus	elaphus	Nearctic	Yes	Yes	8	No	NA
Cervus	elaphus	Nearctic	Yes	Yes	3	No	NA
Cervus	elaphus	Nearctic	Yes	Yes	24	No	NA
Cervus	elaphus	Nearctic	Yes	No	18	No	NA
Cervus	elaphus	Australian	no	Yes	3	Yes	No
Cervus	elaphus	Nearctic	Yes	Yes	4	Yes	No
Cervus	elaphus	Palaearctic	Yes	Yes	10	Yes	No
Cervus	elaphus	Australian	no	No	9	Yes	Yes
Cervus	elaphus	Australian	no	No	7	Yes	Yes
Cervus	elaphus	Australian	no	No	4	Yes	Yes
Cervus	elaphus	Australian	no	Yes	4	Yes	Yes
Cervus	elaphus	Australian	no	Yes	17	Yes	Yes
Cervus	elaphus	Australian	no	Yes	8	Yes	Yes
Cervus	elaphus	Australian	no	Yes	9	Yes	Yes
Cervus	elaphus	Australian	no	Yes	17	Yes	Yes
Cervus	elaphus	Australian	no	Yes	18	Yes	Yes
Cervus	elaphus	Australian	no	Yes	6	Yes	Yes
Cervus	elaphus	Nearctic	Yes	Yes	8	Yes	Yes
Cervus	elaphus	Nearctic	Yes	No	67	Yes	Yes
Cervus	nippon	Nearctic	no	No	20	No	NA
Cervus	nippon	Australian	no	Yes	3	No	NA
Cervus	nippon	Palaearctic	Yes	No	6	No	NA
Cervus	nippon	Australian	no	Yes	3	Yes	No
Cervus	nippon	Palaearctic	Yes	No	40	Yes	No
Cervus	nippon	Palaearctic	Yes	No	4	Yes	No
Cervus	nippon	Palaearctic	Yes	Yes	7	Yes	No
Cervus	nippon	Palaearctic	Yes	No	20	Yes	No
Cervus	nippon	Australian	no	Yes	6	Yes	Yes
Cervus	nippon	Nearctic	no	Yes	5	Yes	Yes
Cervus	nippon	Palaearctic	Yes	No	18	Yes	Yes
Cervus	nippon	Palaearctic	Yes	No	25	Yes	Yes

Cervus	timorensis	Australian	no	No	2	No	NA
Cervus	timorensis	Australian	no	No	400	Yes	No
Cervus	timorensis	Australian	no	Yes	4	Yes	Yes
Cervus	timorensis	Australian	no	Yes	8	Yes	Yes
Cervus	timorensis	Australian	no	Yes	12	Yes	Yes
Cervus	timorensis	Australian	no	Yes	14	Yes	Yes
Cervus	unicolor	Australian	no	Yes	7	Yes	No
Cervus	unicolor	Australian	no	Yes	2	Yes	No
Cervus	unicolor	Australian	no	Yes	2	Yes	No
Cervus	unicolor	Australian	no	No	7	Yes	Yes
Dama	dama	Australian	no	No	14	No	NA
Dama	dama	Australian	no	No	2	No	NA
Dama	dama	Nearctic	no	No	12	No	NA
Dama	dama	Nearctic	no	No	85	No	NA
Dama	dama	Nearctic	no	No	13	No	NA
Dama	dama	Palaearctic	Yes	Yes	3	No	NA
Dama	dama	Australian	no	No	6	Yes	No
Dama	dama	Australian	no	No	6	Yes	No
Dama	dama	Nearctic	no	No	20	Yes	No
Dama	dama	Nearctic	no	No	51	Yes	No
Dama	dama	Palaearctic	Yes	No	10	Yes	No
Dama	dama	Palaearctic	Yes	No	20	Yes	No
Dama	dama	Palaearctic	Yes	No	100	Yes	No
Dama	dama	Palaearctic	Yes	No	67	Yes	No
Dama	dama	Australian	no	No	20	Yes	Yes
Dama	dama	Australian	no	Yes	18	Yes	Yes
Dama	dama	Australian	no	Yes	18	Yes	Yes
Dama	dama	Australian	no	No	3	Yes	Yes
Dama	dama	Nearctic	no	No	40	Yes	Yes
Dama	dama	Nearctic	no	No	60	Yes	Yes
Dama	dama	Nearctic	no	No	73	Yes	Yes
Dama	dama	Palaearctic	Yes	No	54	Yes	Yes
Dama	dama	Palaearctic	Yes	Yes	90	Yes	Yes
Hydropotes	inermis	Palaearctic	Yes	Yes	32	Yes	Yes
Hydropotes	inermis	Palaearctic	Yes	Yes	19	Yes	Yes
Muntiacus	reevesi	Palaearctic	no	Yes	28	Yes	Yes
Odocoileus	hemionus	Australian	no	Yes	5	No	NA
Odocoileus	hemionus	Australian	no	Yes	9	No	NA
Odocoileus	hemionus	Palaearctic	no	Yes	4	No	NA
Odocoileus	hemionus	Nearctic	Yes	Yes	7	No	NA
Odocoileus	hemionus	Nearctic	Yes	No	2	No	NA
Odocoileus	hemionus	Nearctic	Yes	No	50	No	NA
Odocoileus	hemionus	Nearctic	Yes	Yes	13	Yes	No
Odocoileus	hemionus	Nearctic	Yes	No	13	Yes	No
Odocoileus	hemionus	Australian	no	Yes	2	Yes	Yes
Odocoileus	hemionus	Australian	no	Yes	8	Yes	Yes
Odocoileus	hemionus	Nearctic	Yes	Yes	12	Yes	Yes
Odocoileus	hemionus	Nearctic	Yes	Yes	24	Yes	Yes
Odocoileus	hemionus	Nearctic	Yes	Yes	11	Yes	Yes
Odocoileus	hemionus	Nearctic	Yes	Yes	16	Yes	Yes
Odocoileus	hemionus	Nearctic	Yes	Yes	7	Yes	Yes
Odocoileus	hemionus	Oceanian	Yes	Yes	40	Yes	Yes
Odocoileus	virginianus	Australian	no	Yes	5	No	NA
Odocoileus	virginianus	Australian	no	Yes	9	Yes	No
Odocoileus	virginianus	Australian	no	Yes	9	Yes	Yes
Odocoileus	virginianus	Neotropical	no	Yes	5	Yes	Yes
Odocoileus	virginianus	Palaearctic	no	No	5	Yes	Yes
Rangifer	tarandus	Palaearctic	Yes	Yes	14	No	NA
Rangifer	tarandus	Palaearctic	Yes	Yes	5	No	NA
Rangifer	tarandus	Palaearctic	Yes	Yes	3	No	NA
Rangifer	tarandus	Palaearctic	Yes	Yes	200	No	NA
Rangifer	tarandus	Afrotropical	no	Yes	3	Yes	No
Rangifer	tarandus	Nearctic	Yes	Yes	15	Yes	No
Rangifer	tarandus	Nearctic	Yes	Yes	25	Yes	No
Rangifer	tarandus	Neotropical	Yes	Yes	7	Yes	No
Rangifer	tarandus	Palaearctic	Yes	Yes	23	Yes	No
Rangifer	tarandus	Palaearctic	Yes	Yes	15	Yes	No
Rangifer	tarandus	Palaearctic	Yes	Yes	10	Yes	No
Rangifer	tarandus	Afrotropical	no	Yes	10	Yes	Yes
Rangifer	tarandus	Nearctic	Yes	No	1280	Yes	Yes
Rangifer	tarandus	Nearctic	Yes	Yes	24	Yes	Yes
Rangifer	tarandus	Nearctic	Yes	Yes	29	Yes	Yes
Rangifer	tarandus	Neotropical	Yes	Yes	10	Yes	Yes
Rangifer	tarandus	Palaearctic	Yes	Yes	35	Yes	Yes
Rangifer	tarandus	Palaearctic	Yes	Yes	35	Yes	Yes

Appendix 2.3 Sources of DNA sequence data used in the phylogenetic analysis

Family	Subfamily	Tribe	Genus	Species	cytb	COII	12S rRNA	16S rRNA	PRKCI	αLAlb
Antilocapridae	Antilocaprinae	Antilocaprini	*Antilocapra*	*americana*	AF091629	U62571	U86969	U87019	AF165669	AY122014
Bovidae	Antilopinae	Antilopini	*Gazella*	*granti*	AF034723	U18824	AY670652	EF033136	AF165749	AY122029
Bovidae	Bovinae	Strepsicerotini	*Tragelaphus*	*imberbis*	AF036279	U18815	AF091697	M86493	AF165733	AY122025
Cervidae	Capreolinae	Alcinae	*Alces*	*alces*	AJ000026	DQ379322	AY184437	DQ318382	DQ379338	DQ379360
Cervidae	Capreolinae	Capreolini	*Capreolus*	*capreolus*	AJ000024	DQ365690	AY184439	AY122048	DQ365692	AY122021
Cervidae	Capreolinae	Capreolini	*Capreolus*	*pygargus*	AJ000025	-	-	-	-	-
Cervidae	Capreolinae	Capreolini	*Hydropotes*	*inermis*	AJ000028	DQ379323	EU315254	EU315254	DQ379340	AY122020
Cervidae	Capreolinae	Odocoileinae	*Blastocerus*	*dichotomus*	DQ379306	DQ379324	-	-	DQ379341	DQ379361
Cervidae	Capreolinae	Odocoileinae	*Hippocamelus*	*antisensis*	DQ379307	DQ379325	-	-	DQ379342	DQ379362
Cervidae	Capreolinae	Odocoileinae	*Hippocamelus*	*bisculus*	DQ789178	-	-	-	-	-
Cervidae	Capreolinae	Odocoileinae	*Mazama*	*americana*	DQ789222	DQ379326	Y08209	-	DQ379343	DQ379363
Cervidae	Capreolinae	Odocoileinae	*Mazama*	*bororo*	DQ789228	-	-	-	-	-
Cervidae	Capreolinae	Odocoileinae	*Mazama*	*gouazoubira*	DQ379308	DQ379368	Y08570	-	DQ379344	DQ379364
Cervidae	Capreolinae	Odocoileinae	*Mazama*	*nana*	DQ789210	-	-	-	-	-
Cervidae	Capreolinae	Odocoileinae	*Mazama*	*nemorivaga*	DQ789213	-	-	-	-	-
Cervidae	Capreolinae	Odocoileinae	*Mazama*	*temama*	AJ000027	-	AJ000030	-	-	-
Cervidae	Capreolinae	Odocoileinae	*Odocoileus*	*hemionus*	AF091630	DQ379369	AF091708	DQ318369	DQ379345	AY122022
Cervidae	Capreolinae	Odocoileinae	*Odocoileus*	*virginianus*	DQ379370	U18816	M35874	M35874	DQ379346	DQ379365
Cervidae	Capreolinae	Odocoileinae	*Ozotoceros*	*bezoarticus*	DQ789198	-	-	-	-	-
Cervidae	Capreolinae	Odocoileinae	*Pudu*	*puda*	DQ379309	DQ379327	-	-	DQ379347	DQ379366
Cervidae	Capreolinae	Odocoileinae	*Rangifer*	*tarandus*	AJ000029	DQ379328	AY184438	DQ318374	AF165693	AY122019
Cervidae	Cervinae	Cervini	*Axis*	*axis*	AY607040	DQ379310	DQ017832	AY391766	DQ379329	DQ379348
Cervidae	Cervinae	Cervini	*Axis*	*kuhlii*	This study	-	-	-	-	-
Cervidae	Cervinae	Cervini	*Axis*	*porcinus*	DQ379301	DQ379311	AY775785	AY391768	DQ379367	DQ379349
Cervidae	Cervinae	Cervini	*Cervus*	*albirostris*	AF423202	DQ379312	AY184429	-	DQ379330	DQ379350
Cervidae	Cervinae	Cervini	*Cervus*	*alfredi*	This study	-	-	-	-	-
Cervidae	Cervinae	Cervini	*Cervus*	*duvauceli*	AY607041	DQ379313	EU908275	EU084668	DQ379331	DQ379351
Cervidae	Cervinae	Cervini	*Cervus*	*elaphus*	AY244490	DQ365689	DQ153244	EU144036	AY846793	AY122017
Cervidae	Cervinae	Cervini	*Cervus*	*eldi*	AY157735	DQ379314	AY184432	AF108041	-	DQ379353
Cervidae	Cervinae	Cervini	*Cervus*	*nippon*	AY035876	DQ379315	AY184433	GU457433	DQ379332	DQ379352
Cervidae	Cervinae	Cervini	*Cervus*	*schomburgki*	AY607036	-	-	-	-	-
Cervidae	Cervinae	Cervini	*Cervus*	*timorensis*	AF423200	DQ379316	-	-	DQ379333	DQ379354
Cervidae	Cervinae	Cervini	*Cervus*	*unicolor*	AF423201	DQ379317	AY184434	M35875	DQ379334	DQ379355
Cervidae	Cervinae	Cervini	*Dama*	*dama*	AJ000022	DQ379318	AJ885203	DQ922639	DQ379335	DQ379356
Cervidae	Cervinae	Cervini	*Dama*	*mesopotamica*	DQ379304	DQ379319	-	-	DQ379336	DQ379357
Cervidae	Cervinae	Cervini	*Elaphurus*	*davidianus*	AF423194	DQ379320	AY397660	-	DQ379337	DQ379358
Cervidae	Cervinae	Cervini	*Megaloceros*	*giganteus*	AM072745	-	-	-	-	-
Cervidae	Cervinae	Muntiacini	*Elaphodus*	*cephalophus*	DQ379305	DQ379321	AY184436	AF108040	DQ379339	DQ379359
Cervidae	Cervinae	Muntiacini	*Megamuntiacus*	*vuquangensis*	AF042720	-	-	AF108034	-	-
Cervidae	Cervinae	Muntiacini	*Muntiacus*	*crinifrons*	DQ445735	AY239042	AY239042	AY239042	-	-
Cervidae	Cervinae	Muntiacini	*Muntiacus*	*feae*	AF042721	-	-	AF108036	-	-
Cervidae	Cervinae	Muntiacini	*Muntiacus*	*muntjak*	AF042718	NC_004563	AM778453	EF523639	-	-
Cervidae	Cervinae	Muntiacini	*Muntiacus*	*putaoensis*	EF523669	-	-	EF523642	-	-
Cervidae	Cervinae	Muntiacini	*Muntiacus*	*reevesi*	AF042719	NC_004069	M35877	M35877	AF165677	AY122018
Cervidae	Cervinae	Muntiacini	*Muntiacus*	*rooseveltorum*	-	-	-	AF108031	-	-
Cervidae	Cervinae	Muntiacini	*Muntiacus*	*truongsonensis*	-	-	-	AF108033	-	-
Moschidae			*Moschus*	*moschiferus*	AY121995	DQ365691	AY184428	AY122045	DQ365693	AY122033

Appendix 2.4 Results of GLM and GLMM analyses of establishment at the population level

	Variable	n	GLM* Δ deviance	GLMM Z Value
	Species	146	5.74	-
	Genus	146	14.6	-
Introduction Effects	Biome match	146	0.96	0.85
	Propagule size (Log10)	146	19.9*** (23.4***)	3.97*** (4.33***)
	Island introduction	146	0.07	0.31
Regional Effects	Region of origin	146	2	0.91 to -0.3
	Native area	146	0	0.33
	Home Range (Log10)	140	0.06	0.19
	Population density	146	0.01	-0.09
	Human population density (Log10)	144	0.14	-0.38
Species Characteristics	Habitat breadth	146	2.13	1.39 to 0.12
	Diet breadth	146	2.1	1.09 to -0.28
	Relative brain mass (Log10)	143	0.11	0.36
	Social group size	146	3.53	0.69 to -1.07
Life History Traits	Weaning body mass (Log10)	133	1.34	-1.15
	Body mass (Log10)	146	1.02	-1.01
	Weaning age	146	0.68	0.84
	Maximum lifespan (Log10)	146	0.27	-0.52
Reproductive Traits	Age at sexual maturity (Log10)	146	0.27	-0.52
	Litters per year (Log10)	146	0.63	0.78
	Litter size (Log10)	146	0.93	0.95
	Interbirth interval (Log10)	137	1.86	-1.31
	Gestation length (Log10)	146	0.13	-0.38

*GLM were fitted to the data, with the outcome of the introduction events as the binary response variable, and the characteristics hypothesised to influence success as the explanatory variables.

The models were fitted with a binomial error distribution and the logit-link function. Models were fitted using forward selection then dropping and adding terms. The effect of adding each term alone to a null model was calculated, then the variable which caused the largest drop in deviance was added to the model until the addition of new variables explained no significant ($p<0.05$) additional variation. Chi-square tests on the change in deviance resulting from the addition of a new variable to the model were used to determine significance. The GLM analysis was repeated excluding the three events where over 200 individuals were introduced simultaneously.

The effects of variables in explaining establishment success at the population level. n = Number of introduction events.
Values in brackets obtained from analysis excluding three outlying introduction events (>200 individuals introduced). Significance code: p < 0.001 = '*'**

Appendix 2.5 Results of GLM and GEE analyses of establishment at the species level

	Variable	n	GLM Δ deviance	GEE t value
Regional Effects	Region of origin	16	0.15	-0.56 to -1.92
	Native area	16	0.03	3.03*
	Home Range (Log10)	15	0.31	-1.65
	Population density	16	0.07	-3.8
	Human population density (Log10)	15	0.00	0.46
Species Characteristics	Habitat breadth	16	0.30	4.94 to -0.81
	Diet breadth	16	0.10	4.83 to -2.87
	Relative brain mass (Log10)	15	0.6*	2.3
	Social group size	16	0.28	1.39 to -1.09
Life History Traits	Weaning body mass (Log10)	11	0.28***	1.57
	Body mass (Log10)	16	0.23*	-0.56
	Weaning age	16	0.08	-3.99**
	Maximum lifespan (Log10)	16	0.16	-0.15
Reproductive Traits	Age at sexual maturity (Log10)	16	0.3*	-4.25**
	Litters per year (Log10)	16	0.29	-1.35
	Litter size (Log10)	16	0.08	2.13
	Interbirth interval (Log10)	14	0.32	-1.47
	Gestation length (Log10)	16	0.16	-2.05

The effects of variables in explaining establishment success at the species level. n = Number of species included in analysis.
Significance codes: p < 0.001 = '*'; p < 0.01 = '**'; p <0.05 = '*'**

Appendix 2.6 Results of GLM and GLMM analyses of spread at the population level

	Variable	n	GLM^ Δ deviance	GLMM Z Value
	Species	100	24.4*	-
	Genus	100	12.4	-
Introduction Effects	Biome match	100	1.28	-0.60
	Propagule size (Log10)	100	0.30	1.24
	Island introduction	100	3.88*	1.90 .
Regional Effects	Region of origin	100	5.14	-0.91 to -2.76
	Native area	100	0.51	0.45
	Home Range (Log10)	95	0.11	0.27
	Population density	100	0.00	-0.01
	Human population density (Log10)	99	0.26	-0.35
Species Characteristics	Habitat breadth	100	8.78*	2.93** to -2.76
	Diet breadth	100	8.72.	2.06* to -2.2
	Relative brain mass (Log10)	98	3.54 .	1.71 .
	Social group size	100	4.31	1.64 to 0.00
Life History Traits	Weaning body mass (Log10)	91	1.82	0.92
	Body mass (Log10)	100	0.10	-0.49
	Weaning age	100	3.69.	-1.92 .
	Maximum lifespan (Log10)	100	0.15	-0.97
Reproductive Traits	Age at sexual maturity (Log10)	100	0.01	-0.24
	Litters per year (Log10)	100	4.50*	-1.90 .
	Litter size (Log10)	100	1.66	1.48
	Interbirth interval (Log10)	94	3.89*	-1.80 .
	Gestation length (Log10)	100	0.09	-0.29

^GLM were fitted to the data with the success rate of the species as the response variable, and the characteristics hypothesised to influence success as the explanatory variables.

The effects of introduction event, regional in explaining spread success at the population level. n = Number of introduction events.

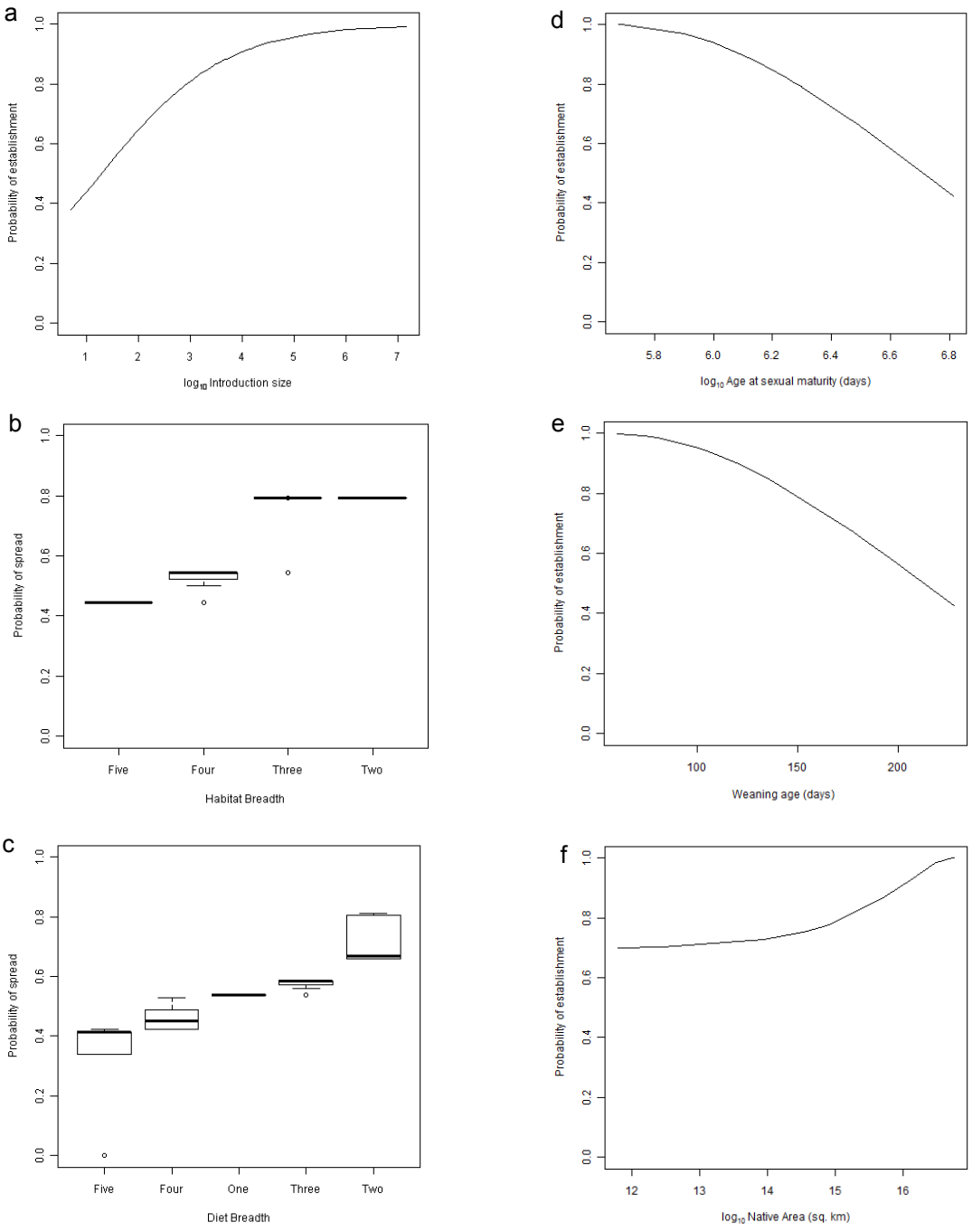
Significance codes: p < 0.01 = ''; p <0.05 = '*'; p< 0.1 = '.'**

Appendix 2.7 Results of GLM and GEE analyses of spread at the species level

	Variable	n	GLM Δ deviance	GEE t value
Regional Effects	Region of origin	15	0.31	0.26 to 1.46
	Native area	15	0.15	-1.21
	Home Range (Log10)	14	0.17	-1.37
	Population density	15	0.04	-0.33
	Human population density (Log10)	14	0.00	-0.12
Species Characteristics	Habitat breadth	15	0.56	0.76 to -0.46
	Diet breadth	15	1.84	0.54 to -1.55
	Relative brain mass (Log10)	14	0.21	1.57
	Social group size	15	0.32	0.72 to 0.29
Life History Traits	Weaning body mass (Log10)	11	0.42	1.57**
	Body mass (Log10)	15	0.59	-1.87
	Weaning age	15	0.27	-0.63
	Maximum lifespan (Log10)	15	0.54	-1.91
Reproductive Traits	Age at sexual maturity (Log10)	15	0.08	1.15
	Litters per year (Log10)	15	0.12	-0.68
	Litter size (Log10)	15	0.38	1.61
	Interbirth interval (Log10)	13	0.51	-1.01
	Gestation length (Log10)	15	0.00	2.02

The effects of variables in explaining spread success at the species level. n = Number of species included in analysis.
Significance code: p< 0.01 = ''**

Appendix 2.8 Plots of predicted probability of establishment and spread



Plots of predicted probability of success at two stages of the invasion process: establishment and spread.

a-c Plots of fitted values of GLMM models against variables found to significantly predict success at the population level.
d-g Plots of fitted values of GEE models against variables found to significantly predict success at the species level (see Table 2).

Appendix 4.1 Pairwise Φ_{PT} values between North German roe deer populations

	FN	FO	NF	RA	SL
FN	-	P > 0.1	P > 0.1	P > 0.1	P > 0.1
FO	0.031	-	P > 0.1	P > 0.1	P > 0.1
NF	0.012	0.024	-	P > 0.1	P > 0.1
RA	0.020	0.016	0.022	-	P > 0.1
SL	0.015	0.026	0.008	0.008	-

Appendix 5.1 Population pairwise F_{ST} values for Chinese water deer control region (below diagonal) and cytb sequences (above diagonal).

P values: * = significant at P < 0.05, ** = significant at P < 0.01, *** = significant at P < 0.001

Population		1	2	3	4	5	6	7	8	9	10
1	China - Dafeng	-	0.037	0.065	0.017	0.000	0.282 *	0.250	0.322 *	0.493 ***	0.5 *
2	China - Jishan	na	-	0.0901 *	0.045 *	0.035	0.285 ***	0.254 *	0.322 ***	0.482 ***	0.456 **
3	China - Yancheng	0.037	na	-	0.067 ***	0.026	0.266 **	0.144	0.214 **	0.349 ***	0.302 *
4	China - Zhoushan Islands	0.058	na	0.083 *	-	0.016	0.232 ***	0.197 **	0.260 ***	0.402 ***	0.344 ***
5	France - St. Croix	0.083	na	0.111	0.122 *	-	0.251 *	0.098	0.187 *	0.345 *	0.322 *
6	UK - Bure Marshes	0.062	na	0.088	0.101 **	0.132	-	0.395 **	0.403 ***	0.511 ***	0.520 **
7	UK - East Norfolk	0.5 *	na	0.456 **	0.382 ***	0.583 *	0.441 **	-	-0.013	-0.010	0.000
8	UK - Whipsnade	0.465 **	na	0.457 ***	0.412 ***	0.511 ***	0.384 ***	0.675 ***	-	0.098 *	0.059
9	UK - Woodwalton Fen	0.176	na	0.190 **	0.186 ***	0.211 *	0.102	0.488 ***	0.079 *	-	-0.025
10	UK - Yare Marshes	0.000	na	0.040	0.063	0.092	0.026	0.579 *	0.257	-0.001	-

Appendix: Summary of permission for third party copyright works

Springer licence for reproduction of the paper:

Fautley, R., Coulson, T. and Savolainen, V. (2012) 'A comparative analysis of the factors promoting deer invasion', Biological Invasions, 14(11), 2271-2281

SPRINGER LICENSE

License Number	3150831384987
License date	May 16, 2013
Licensed content publisher	Springer
Licensed content publication	Biological Invasions
Licensed content title	A comparative analysis of the factors promoting deer invasion
Licensed content author	Richard Fautley
Licensed content date	Jan 1, 2012
Volume number	14
Issue number	11
Type of Use	Thesis/Dissertation
Portion	Full text