THE UNIVERSITY of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Hybridisation and introgression of exotic *Cervus* (*nippon* and *canadensis*) with red deer (*Cervus elaphus*) in the British Isles.

Stephanie Lindsay Smith

(BSc. Hons. University of Aberdeen)

This thesis is submitted for the degree of Doctor of Philosophy to the College of Science and Engineering, School of Biological Sciences.

The University of Edinburgh

March 2013

Declaration

I declare that the work described in this thesis has been carried out by myself unless otherwise acknowledged. It is entirely of my own composition and has not, in whole or part, been submitted for any other degree.

Stephanie Smith

Edinburgh

March 2013

Abstract

Europe's largest population of wild red deer (*Cervus elaphus)* resides in the British Isles and has been present since the end of the last ice age, c. 11,000BP. Since the mid-19th century, multiple introductions of Japanese sika (*Cervus nippon*) and wapiti (*Cervus canadensis*) have taken place across the British Isles. While wapiti introductions have generally gone extinct, sika have thrived and expanded and now often live in sympatry with red deer. Hybridisation between these species has been demonstrated in captivity and in the wild. This study sought to determine the extent of hybridisation and introgression between red and sika across large parts of the British Isles and elucidate some of its potential consequences.

Chapter 2 addresses the extent of hybridisation and introgression across Scotland and NW England. A total of 2984 samples from the North Highlands, the central Highlands, the Hebrides, Kintyre and the English Lake District were genotyped at 22 microsatellite loci, which are highly diagnostic for red and sika and strongly diagnostic for red and wapiti and a mitochondrial marker that is diagnostic for red and sika, alongside 49 wapiti samples from Canada. Microsatellite data was analysed using the Bayesian clustering program Structure 2.3 to determine the extent of admixture between species. There was some evidence for very low-level introgression by wapiti into a small number of Scottish red deer (<0.2% of total). Only two areas (both in Kintyre, Argyll) showed extensive introgression with collapse of assortative mating between red and sika (50.4% and 61.8% of sampled individuals were hybrid in West Loch Awe and South Kintyre, respectively). However, rare and widely scattered individuals with low-level sika introgression or cytonuclear disequilibrium suggest hybridisation has occurred in several other places in mainland Scotland and Cumbria in the past without subsequent loss of assortative mating.

Chapter 3 addresses the extent of hybridisation in Ireland. There are now an estimated 4,000 red deer in Ireland and their numbers are increasing. It has recently been determined that the red deer in Killarney, County Kerry are descended from an ancient (c. 5,000BP) introduction and therefore merit genetic conservation. Introduction of exotic species, including Japanese sika and North America wapiti, since the 19th century have primarily occurred via the now defunct Powerscourt Park, County Wicklow, which was the source of many translocations to the rest of Ireland as well as to the UK. 374 deer samples from across Ireland were analysed as in Chapter 2. Wapiti introgression was again very low, with trace amounts of introgression detected in a small proportion of samples (0.53%), whilst 41% of 197 deer sampled in Co. Wicklow and 47% of 15

deer sampled in Co. Cork were red-sika hybrids according to either their nuclear genome or mitochondrial haplotype. No pure red deer were detected in Co. Wicklow, suggesting that in this region the red deer has disappeared following hybridisation. Whilst no hybrids were detected among 37 red samples and 77 sika samples in Co. Kerry, the Co. Cork hybrids pose a threat to the Killarney populations due to their proximity.

Chapter 4 investigates population genetic structure *within* red and sika populations across the British Isles and investigates whether low-level introgression by the other species influences the resolved population structure. Structure analysis was conducted separately using 2307 'pure' red deer individuals and 752 'pure' sika animals from the British Isles (defined as $Q > 0.95$ for red and $Q < 0.05$ for sika) and then on reduced sample sizes using more stringent purity criteria ($Q \geq 0.99$ and $Q \leq 0.01$). As might be predicted, the more stringent criteria removed individuals in areas known to contain advanced backcrosses. In red deer, there was some evidence for a loss of within-species population structure under the more stringent criteria, while for sika there was not. Datasets were also analysed using Discriminate Analysis of Principal Components; a multivariate method designed to infer and describe genetic population structure. In red deer, both analytical approaches confirmed the strong separation of the deer on Harris and Lewis from others, and there is support for clusters typified by the other Hebridean islands, Kintyre, central and North Scotland and the English sites. Among sika, both approaches supported the likelihood of three clusters which are presumably the result of bottleneck events as each introduction was made.

Chapter 5 investigates the phenotypic consequences of hybridisation by three approaches. Firstly, carcass weight was regressed against genetically-determined hybrid scores (at two stringency levels, see Chapter 4) and heterozygosity (in terms of red and sika alleles). Among hybrids, carcass weight is linearly related to hybrid score (Q) and there is some evidence for a positive relationship with heterozygosity. This suggests that additive genetic variation explains variation in carcass weight to a greater extent than heterosis. Secondly, analysis of five case studies representing individual putative hybrids submitted by stalkers from areas without known hybridisation, two proved to be hybrids, while the other three were pure sika. Lastly, in regions known to contain hybrids, the accuracy of ranger-assigned phenotype averaged 78% and revealed that in Scotland accuracy tends to decline as an individual becomes more genetically intermediate; whilst in Co. Wicklow it is the identification of pure parental animals that is more challenging.

In conclusion, the existence of rare and widely scattered advanced red-sika backcrosses with low-level nuclear introgression and/or mitochondrial introgression (e.g. in North of Scotland, Cumbria) highlight that some hybridisation events are followed by extensive backcrossing without the breakdown of assortative mating, while others are followed by the generation of a hybrid swarm (e.g. in South Kintyre, West Loch Awe, Co. Wicklow, Co. Cork). Phenotypic traits can become intermediate due to hybridisation and this may facilitate further gene flow and hybridisation. New molecular tools including next generation sequencing (NGS) will enable better understanding the hybridisation process and its phenotypic consequences in this and other systems.

Acknowledgements

I am writing my acknowledgements not on the conclusion of my project but at a time when I really need to acknowledge the people that have enabled me to do this. It is clear, that without the help of my parents, I would not have been able to tackle challenges beyond that of a thesis. I know it is the belief of many a child, but I don't understand how I could have had the most supportive and understanding and loving parents in the world. Also to some of the most incredible people I have ever had the fortune of calling friends, old and new; people that seem to go beyond the call of friendship when you didn't even realise you may have needed it. With special thanks to Jessica Hedge, Andrew Grimmer and Claire Webster. To a wonderful supervisor, who went out of her way to help me and has taught me so much.

I must also thank Forestry Commission Scotland rangers, including Kevin McKillop, Willie Lamont, Derick Macaskill, David Bain, staff of the Deer Commission Scotland (now Scottish Natural Heritage), Will Boyd-Wallis and the Cairngorm National Park Authority and the private estate owners and stalkers who cooperated within this project and provided samples. Thanks to Eleni Socratous and supervisors from the University of Leicester and Mike Thornley (Forestry Commission Chief Ranger) and the rangers of the Forestry Commission Cumbria for the samples they provided from the Lake District. Thanks to Prof. David Coltman (University of Alberta) for the provision of wapiti samples. Thanks also to Elizabeth Heap, Sheena Morrissey, Elizabeth Mittell and Sarah MacDonald for genotyping samples. Big thanks also to Phil Ellis for his help genotyping and haplotyping samples and for being great to work alongside. Lastly, thanks to the GenePool for their excellent service and NERC for funding.

So above all, apart from exploring the discipline of molecular ecology that I find so exceptionally fascinating, I would hope to acknowledge the people and experiences this project had enabled me to have.

Table of contents

Chapter 1: General Introduction……………………………………………………………..1

1.1 Defining Hybridisation and Introgression……………......…………………………..1
1.2 The Study of Hybridisation..…………………………..…………………………3
1.3 Types of Hybridisation and their Outcomes…………….…………………………..4
1.3.1 Natural hybridisation……………………………………………………………….4
1.3.2 Anthropogenic hybridisation…………………………………………………………..5
1.4 Red deer and Sika in the British Isles……………………………………………..6
1.4.1 Red deer (*Cervus elaphus*)………………………....……………………………8
1.4.2 Sika (*Cervus nippon*)……………………………………………...………………10
1.5 Hybridisation between red and sika in the British Isles…...………………………....12
1.5.1 Scotland…………………………………………………………………………...12
1.5.2 England…………………………………………………………………………...14
1.5.3 Ireland……………………………………………………………………………14
1.6 Studying hybridisation………………………………………………………………15
1.6.1 Genetic markers……………………………………………………………………..15
1.6.2 Statistical approaches to inferring population genetic structure………………………16
1.7 This study…………………………………………………………………………18
1.7.1 Objectives of this thesis.………………………………………………………………..19
1.8 **Acknowledgements**…….……………………………………………………………19

Chapter 2: Introgression of exotic *Cervus (nippon* and *canadensis)* into red deer (*Cervus elaphus*) in Scotland and northwest England……………….…………………………21

2.1 Abstract………………………………………………………………………………..22
2.2 Introduction………………………………………………………………………..23
2.2.1 Hybridisation……………………………………………………………………...23
2.2.2 *Cervus* in the British Isles…………………………………………………………..23
2.2.3 This study…………………………………………………………………………26
2.3 Materials and Methods……………………………………………………………27
2.3.1 Study area and sampling………………………………………………………………..27
2.3.2 DNA analysis……………………………………………………………………...29
2.3.3 Objective 1: Extent of hybridisation ………………………………………………...30
2.3.4 Objective 2: Direction of hybridisation ……………………………………………...31

2.4 Results……32
2.4.1 Genotypes……32
2.4.2 Objective 1: Extent of hybridisation ……34
2.4.3 Objective 2: Direction of hybridisation ……45
2.5 Discussion……47
2.5.1 Objective 1: Extent of hybridisation ……47
2.5.2 Objective 2: Direction of hybridisation ……50
2.5.3 Objective 3: Context and Causation ……51
2.5.4 Objective 4: Management suggestions ……53
2.6 Acknowledgements……54
2.7 Appendices……54

Chapter 3: A survey of the hybridisation status of *Cervus* deer species on the island of Ireland……69

3.1 Abstract……70
3.2 Introduction……70
3.2.1 Hybridisation……70
3.2.2 *Cervus* in Ireland……71
3.2.3 This study……74
3.3 Materials and Methods……75
3.3.1 Sampling Sites……75
3.3.2 DNA analysis……76
3.3.3 Objective 1: Extent of hybridisation……76
3.3.4 Objective 2: Direction of hybridisation……78
3.3.5 Objective 3: Source of hybrids outwith Co. Wicklow……78
3.3.6 Objective 4: Accuracy of Ranger-assigned phenotypes……79
3.4 Results……79
3.4.1 Objective 1: Extent of hybridisation……81
3.4.2 Objective 2: Direction of hybridisation……93
3.4.3 Objective 3: Source of hybrids outwith Co. Wicklow……93
3.4.4 Objective 4 Accuracy of Ranger-assigned phenotypes……95
3.5 Discussion……96
3.5.1 Objective 1: Extent of hybridisation……96
3.5.2 Objective 2: Direction of hybridisation ……98
3.5.3 Objective 3: Source of hybrids outwith Co. Wicklow……98
3.5.4 Objective 4 Accuracy of Ranger-assigned phenotypes……99

3.5.5 Objective 5: Management suggestions……………………………………………….99
3.6 Acknowledgements..…………………………………………………………………...100
3.7 Appendices……………………………………………………………………………..100

Chapter 4: Population structure and genetic diversity within two *Cervus* species in the British Isles……………………………………………………………………………...**111**

4.1 Abstract…………………………………………………………………………………112
4.2 Introduction……………………………………………………………………………112
4.2.1 Factors which can influence population structure…………………………...113
4.2.2 Red deer in the British Isles……………………………………………………....115
4.2.3 Sika in the British Isles…………………………………………………………...117
4.2.4 This study…………………………………………………………………………..118
4.3 Materials and Methods………………………………………………………………...120
4.3.1 Sampling Sites………………………………………………………………………120
4.3.2 DNA Analysis………………………………………………………………………121
4.3.3 Objective 1 and 2: Genetic population structure in 'pure' red and 'pure' sika ..……...122
4.3.4 Objective 3: Individuals removed by more stringent purity criteria………………….123
4.4 Results.…………………………………………………………………………………123
4.4.1 Objective 1: Genetic population structure in 'pure' red……………………………..125
4.4.2 Objective 2: Genetic population structure in 'pure' sika…………………………….132
4.4.3 Objective 3: Individuals removed by more stringent purity criteria………………….138
4.5 Discussion.……………………………………………………………………………..142
4.5.1 Objective 1: Genetic population structure in 'pure' red……………………………..142
4.5.2 Objective 2: Genetic population structure in 'pure' sika…………………………….145
4.5.3 Objective 3: Individuals removed by more stringent purity criteria………………….146
4.5.4 Objective 4: Management suggestions ..……………………………………………147
4.6 Acknowledgements……………………………………………………………………148
4.7 Appendices……………………………………………………………………………..150

Chapter 5: Phenotypic consequences of introgression…..………………………..………**163**

5.1 Abstract………………………………………………………………………………....164
5.2 Introduction…………………………………………………………...……….165
5.2.1 Phenotypic consequences of hybridisation……………………………………...165
5.2.2 Red-sika hybridisation and its potential phenotypic consequences………………..166
5.2.3 This study…………………………………………………………………………..168

5.3 Materials, Methods and Results ...169
5.3.1 Objective 1: Carcass weight...169
5.3.1.1 Data selection...169
5.3.1.2 Statistical analysis...172
5.3.1.3 Results...173
5.3.2 Objective 2: Case studies...177
5.3.2.1 Description of case studies...177
5.3.2.2 Genetic analysis...182
5.3.2.3 Results...182
5.3.3 Objective 3: Stalker-assigned phenotype...183
5.3.3.1 Data selection...183
5.3.3.2 Statistical analysis...184
5.3.3.3 Results...184
5.4 Discussion...190
5.4.1 Objective 1: Carcass weight...190
5.4.2 Objective 2: Case studies...191
5.4.3 Objective 3: Stalker-assigned phenotype...191
5.4.4 Objective 4: Management implications...193
5.5 Acknowledgements...194
5.6 Appendices...194

Chapter 6: General Conclusions ...**195**

6.1 Summary of main findings...195
6.2 Implications for management...199
6.2.1 Hybrid swarm...199
6.2.2 Low-level introgression...200
6.2.3 No hybridisation...201
6.2.4 Further aspects to consider...201
6.3 Future research...202

Appendix 1: The taxonomic and hybridisation status of *Cervus* deer species in various populations in the south of England and France...207

7.1 Abstract...208
7.2 Introduction...208
7.3 Materials and Methods...209
7.3.1 Study area and sampling...209

7.3.2 DNA analysis……………………………………………………………….211
7.3.3 Extent of hybridisation………………………………………………………...211
7.4 Results…………………………………………………………………………212
7.4.1 Genotypes…………………………………………………………………..212
7.4.2 Extent of hybridisation………………………………………………………...212
7.5 Discussion……………………………………………………………………...216
7.5.1 Extent of hybridisation………………………………………………………...216
7.6 Acknowledgements…………………………………………...……………….217
7.7 Appendices……………………………………………………………………..218
8.0 References……………………………………………………………………...219

Chapter 1: General Introduction

1.1 Defining Hybridisation and Introgression

The first obstacle in studying hybridisation is its definition and that of a species (Allendorf *et al.* 2001). There are over 24 different definitions of a species that have been developed and discussed since Darwin proposed the concept of evolving groups (Hey 2001). One of the longest-established and best known is the Biological Species Concept (BSC) (Mayr 1942), which assigns species as groups of individuals that interbreed and produce fertile offspring. This, however, does not account for cases where hybridisation takes place between phenotypically distinct taxa (Beebee & Rowe 2004; Fitzpatrick *et al.* 2010; Mallet 2005). Further difficulties arise when trying to determine to what extent populations which are not in contact, animals which reproduce asexually or are now extinct, fit the criteria for the different definitions when their adherence can't be tested directly (Agapow *et al.* 2004; Arnold *et al.* 2008).

Following the acquisition of more molecular data on populations, the Phylogenetic Species Concept assigns species based on the smallest 'aggregation' of either sexual or asexual individuals which share a unique set of 'characters' (diagnostic approach) or share a similar 'history' (coalescent approach) and is more flexible in allowing for monophyletic lineages from a hybridisation event to be included (Baum & Donoghue 1995; Donoghue 1985; Nixon & Wheeler 1990). This concept, however, can be subjective in its choice of traits used to distinguish populations, whilst the differential coalescent rates of genes can confuse history-based concepts (Baum & Donoghue 1995; Kraaijeveld 2000).

A more recent concept, that of the 'evolutionary significant unit', is a dynamic definition suggesting that a unit has evolved over time, but ultimately it requires species to be reciprocally monophyletic (all individuals within each species share a more recent common ancestor than individuals between species) for their mtDNA and show significant divergence in the frequency of nuclear alleles (Crandall *et al.* 2000). Despite the fact this concept better integrates the evolutionary lineage of a taxon, such genetic criteria can be difficult to demonstrate between functionally different units, especially in the face of high gene flow (such as that generated by hybridisation). Whilst they are still largely used, therefore, species concepts based on biological (e.g. reproductive isolation) or phylogenetic (e.g. fixed genetic attributes) constraints are limited. By not recognising hybrid lineages or their frequency, heritable genetic variation or 'adaptive potential' worthy of conservation may be ignored. Modern approaches are attempting to combine the strong commonalities amongst concepts with a more holistic perspective to describe

'species' as evolutionary and demographically dynamic entities, rather than static, discrete taxa, incorporating the evolutionary processes that create them (e.g. gene exchange) and may impact them (e.g. hybridisation) (Hey *et al.* 2003).

Whether species are seen as discrete entities or incompletely differentiated points on a continuum of biodiversity also rests on the markers and the units used to define them (Fitzpatrick *et al.* 2010; Mallet 2005). Technological progress, notably the advent of PCR and large-scale laboratory automation has brought genetic resolution down to DNA nucleotides, increasing the genetic detail of individuals and populations and questioning some species allocations made by different concepts. The exact definition of a species and the popularity of hybridisation as an evolutionary force has, therefore, fluctuated over the years (Ellstrand 2003; Mallet 2005). As the study of hybridisation continues to increase and technology improves, the system of taxonomy will have to adapt its concepts in light of new information.

In this thesis, hybridisation is defined as the interbreeding of genetically distinct taxa (Allendorf *et al.* 2001). The significance of the process was noted by Darwin in the *Origin of Species* (1859), who observed that hybrid animals are "in several respects the most important to us". It is a widespread phenomenon; around 25% of plant species and 10% of animal species are suspected to be involved in hybridisation (Mallet 2005). Particular 'phylogenetic hotspots' of hybridisation exist in the wild, for example, amongst duck species (almost one in two species of Anseriformes shown to hybridise; (Grant & Grant 1992)), game birds and Heliconiine butterflies; (Fitzpatrick *et al.* 2010; Mallet 2005). Hybridisation has been described as an evolutionary catalyst generating novel genotypes over much shorter time scales than by mutation and recombination (Arnold *et al.* 2008; Schwenk *et al.* 2008). As mentioned, hybridisation challenges species concepts and that of the biological 'tree of life', as they don't necessarily allow for lateral or horizontal gene transfer; life could perhaps be better represented by a 'web of life' (Arnold *et al.* 2008).

Introgression is the horizontal gene flow between populations that may result after initial hybridisation events, which can dramatically influence the evolutionary trajectory of a species (Allendorf *et al.* 2001). It can allow advantageous alleles to introgress from one species into another, thus a species to gain adaptive traits (Arnold *et al.* 2008). Examples include the introgression of genetic regions underlying cold tolerance from *Rhododendrum catawbiense* into *R. ponticum,* allowing the latter species to colonise more northern, colder parts of its range (Facon *et al.* 2006) and introgression of sexually-selected male plumage traits between manakin species (*Manacus*) (Parsons *et al.* 1993).

Following hybridisation, the rate of introgression can be exceptionally fast, especially if it is favoured by selection. For example, the selective introgression of three particular markers 90km into the California Tiger Salamander (*Ambystoma* spp.) range occurred within 60 years, whilst the discovery of herbicide-resistant alleles in *Brassica* species 550m from the transgenic crops which sourced them occurred within 17 months (Ellstrand 2003; Fitzpatrick *et al.* 2010; Hall *et al.* 2000). The uncertainty around the impact of global climate change on species and the consequences of introgression between genetically modified species and wild congeners means that the study of hybridisation remains paramount (Fitzpatrick *et al.* 2010; Hails & Morley 2005; Schwenk *et al.* 2008).

1.2 The Study of Hybridisation

By studying hybridisation events, we can better understand the mechanics of outcrossing, reproductive isolation and natural selection in a way analogous to the study of mutated genes in order to better understand their function (Schwenk *et al.* 2008). Hewitt (1988) described hybrid systems as providing "natural laboratories" for the study of evolutionary processes. Massive improvements in the tools and approach to this field of biology have revealed that hybridisation and introgression may have been largely underestimated in previous studies (Schwenk *et al.* 2008). Before the advent of molecular methods, detecting hybridisation by observational data appears to have given delayed indication of the extent of introgression; in a particular time frame we could be observing the conspicuous phenotypic tip of a genotypically introgressed iceberg. Since the 1960s the means of detecting hybridisation has evolved from the use of morphological, physiological and biochemical markers, through the use of protein polymorphisms, towards one of robust molecular markers such as microsatellite markers, single nucleotide polymorphisms (SNPs) and most recently, whole genome sequencing approaches (e.g. Roach *et al.,* (2010)). Whilst the use of neutral markers has prevailed over the last 50 years, we have also been reminded the importance of incorporating information on functional, adaptive genes (Jiggins *et al.* 2008).

1.3 Types of Hybridisation and their Outcomes

1.3.1 Natural hybridisation

Hybridisation can occur naturally, with the result that the two species can form sterile or unfit offspring, they may form a hybrid zone, regions of their genome can introgress, or, alternatively, a new hybrid species may be formed. The first outcome, the production of hybrid offspring which are sterile or unfit (e.g. between *Drosophila melanogaster* and *D. simulans,* (Kaneshiro & Val 1977) and species of deer mice, *Peromyscus* spp. (Gray 1971)) means that there are unlikely to be hybrids beyond those formed in the first generation and such matings effectively constitute a waste of reproductive effort (Allendorf *et al.* 2001).

Hybrid zones generally occur at species boundaries, often on secondary contact of two independently evolving species, creating an area over which parallel clines in introgressing allele frequency can form (Barton & Hewitt 1989). Such hybrid zones are thought to be maintained by a balance between dispersal and selection against hybrids, whether this be exogenous (ecological selection) as modelled by the 'geographical selection gradient' algorithm or endogenous (selection against maladapted gene complexes) leading to a 'tension zone' (Barton & Hewitt 1989; Moore & Price 1993). A hybrid zone between field crickets *Gryllus pennslyvanicus* and *G. firmus* was found to be maintained by selection for soil type and is an example of exogenous selection in action (Rand & Harrison 1989). It was concluded that endogenous selection was likely to be maintaining a hybrid zone between species of dwarf shrub (*Phyllodoce* spp.) in Japan, accounting for the absence of F2s and backcrossed individuals (Kameyama & Kudo 2011). A third way a hybrid zone may be maintained is by frequency-dependent selection, as exemplified by the hybrid zones between Mullerian morphs of *Heliconius* spp. in the Neotropics; the hybrid zone looks like a tension zone, but the selection is induced by frequency-dependent predation of specific morphs by birds (Mallet 2008).

Within a hybrid zone, the barrier across which introgression occurs can have different degrees of permeability (Rand & Harrison 1989). In the *Bombina* toad hybrid system, for example, across 6km of the central hybrid zone in southern Poland diagnostic enzyme frequencies change by up to 0.9; however, introgressed alleles have also been detected up to 38km from the zone centre (Barton & Hewitt 1989). Such permeable species boundaries can lead to a mosaic-like hybrid zone (such as that between species of field cricket, *Gryllus* spp.), rather than a clinal pattern, due to biotic variation and choice of

genetic markers (Rand & Harrison 1989). Fitzpatrick *et al.* (2010) emphasised that the number of informative markers can determine the ability to identify selectively advantageous introgressed alleles in the genetic background of the other species; using 68 markers they were able to identify three markers that were selectively advantageous in *Ambystoma* spp. between-species introgression can preferentially occur in specific parts of the genome, whilst other parts of the genome are more resilient if, for example, under the control of divergent selection within the parental species (e.g. different parapatric inversions remain differentiated in hybridising *Drosophila* spp. due to their adaptive role in their parental species (Mallet 2005; Noor *et al.* 2001)).

Natural hybridisation can also lead to hybrid speciation, whereby introgressive hybridisation generates a novel species (Jiggins *et al.* 2008). This occurs when the recombinant genetic constitution of hybrid offspring produced between two species is such that it confers reproductive isolation (usually by rapid karyotypic change). Hybrid speciation can occur by polyploidy (for example *Primula* spp.; (Ramsey & Schemske 1998)) or homoploidy (for example, between *Heliconius melpomene* and *H. cydbo* (Jiggins *et al.* 2008)). It is normally a founder event, such that the newly formed hybrids exploit a new niche, locality or resource (Ungerer *et al.* 1998). In the language of Wright (1932) adaptive landscape this can involve travelling across an 'adaptive valley' (e.g. slightly reduced fertility (Ungerer *et al.* 1998)) in order to reach another, higher adaptive peak.

1.3.2 Anthropogenic hybridisation

Hybridisation can also be induced anthropogenically. Human intervention and disruption can bring a non-native species into contact with a native species and threaten its persistence (Allendorf *et al.* 2001). With international trade and transport of fish stocks, cultivated crops, livestock and animal products, distinguishing anthropogenic hybridisation from that which occurs naturally is essential in order to apply appropriate conservation measures. For example, if hybridisation between a non-native and a native has not progressed too far there is the potential that the genetic integrity of the parental species could be recovered (Allendorf *et al.* 2001). The various anthropogenic practices which may induce hybridisation between two species include selective breeding, habitat modification and degradation, genetic engineering, introduction of exotic species and even indirectly through captive rearing and management (Karl *et al.* 1995; Leary *et al.* 1993; Rhymer & Braun 1994).

The introduction of alien species by humans into the range of a native conspecific can be devastating. If the species are genetically similar enough that hybridisation can occur,

there are various possible consequences. If hybrids are sterile or unfit, hybridisation leads to wasted reproductive efforts and can displace the native species. For example, the hybrid offspring between the introduced brook trout (*Salvelinus fontinalis)* and native bull trout (*S. confluentus*) are sterile, as are those formed between the introduced minnow species, *Pseudorasbora parva* and the endangered *P. pumila* in eastern Japan (Konishi & Takata 2004; Leary *et al.* 1993). If hybrids are fertile and introgression occurs, this can break up locally adapted suites of genes, leaving the native population vulnerable to demographic fluctuations and extinctions or introduce novel alleles for adaptive traits. If introgression leads to an intermediate phenotype this could cause assortative mating to collapse and species-specific mate recognition to decline (Jiggins *et al.* 2008). In this case the system can collapse into a 'hybrid swarm', in which many individuals have hybrid ancestry and individuals of the parental species are lost (Allendorf *et al.* 2001). Hybrid swarms have been reported between New Zealand mallards and grey ducks (*Anas* spp.) in New Zealand (Rhymer & Braun 1994); between the pecos pupfish and sheepshead minnow (*Cyprinodon* spp.) in Texas (Childs *et al.* 1996).

1.4 Red deer and Sika in the British Isles

Over the last few years an increasing number of studies have detected interspecific hybridisation of European mammals with introduced exotic congeners. The system studied in this project is that between the native red deer (*Cervus elaphus)* and the introduced Japanese sika deer (*Cervus nippon*). These species are markedly different in many aspects, most notably in body size and pelage (Table 1.1; (Senn & Pemberton 2009)). Both species will now be introduced in more detail.

Table 1.1. Some of the most conspicuous phenotypic differences between red and sika deer, based on animals found in Scotland. Reproduced from Senn & Pemberton (2009).

Character	Western European red deer (*Cervus elaphus*)	Japanese sika deer (*Cervus nippon*)
Shoulder height	♂102–112 cm ♀91–102 cm	♂ 81–86 cm ♀76–81 cm
Adult pelage winter:	Grey-brown to dark-brown	Grey-brown, spots invisible
summer:	Chestnut red	Red-brown, large white spots
Rump patch	Off-white to light brown Short tail of same colour	Prominent white, bordered with black Long white tail with black stripe to tip
Ear shape	Pointed	Rounded
Antlers (males only)	Up to 12 points or more	Up to eight points
Rutting call (males only)	Deep roar	Whistle, sometimes high pitched scream

Numerous attempts have been made to resolve the phylogenetic history of the genus Cervus across its Holarctic distribution (Cook *et al.* 1999; Ludt *et al.* 2004). Using mitochondrial sequence information, one of the most widely accepted suggestions is that the progenitor of red and sika deer originated from around Kyrgyzstan, approximately 4.5 MYA (Ludt, 2004; Pitra, 2004). Following this, deer forming a western-migrating clade became the medium-sized red deer (*C. elaphus*) whilst those moving east bifurcated into the larger wapiti (*C. canadensis*) and smaller sika which itself diversified throughout south-eastern Asia including in Japan (*C. n. nippon*), China (*C. n. mantchuricus*) and Taiwan (*C. n. taiouanus*) (Cook *et al.* 1999; Kuwayama & Ozawa 2000; Ludt *et al.* 2004). Despite red and sika sitting at opposite ends of the *Cervus* continuum and showing substantial mitochondrial genetic divergence, a difference of two Robertsonian translocations between their karyotypes and large morphological differences (Table 1.1), they still appear to be able to hybridise in captivity and in the wild (Goodman *et al.* 1999; Harrington 1973; Huang *et al.* 2006; Lowe & Gardiner 1975; Senn & Pemberton 2009). The position of sika as a sister taxon to wapiti suggests they are more genetically similar to each other than either to red, (possibly connected at one point by a land bridge between north-eastern Asian and North America) such that the large differences in their morphology may be due to phenotypic plasticity (Kuwayama & Ozawa 2000; Mahmut *et al.* 2002). Within the sika clade itself there are up to an estimated 13 subspecies, including the Japanese sika (*C. n. nippon)* and Manchurian sika (*C. n. mantchuricus*) (Cook *et al.* 1999). In addition to morphological differences between the two subspecies, Cook *et al.,* (1999) demonstrated a difference in the number of 39bp repeat motifs within a region of the mitochondrial control region, for which Japanese sika has three and Manchurian sika has seven. Despite this, Manchurian and Japanese sika are able to hybridise freely and produce fertile offspring (Gray 1971).

1.4.1 Red deer (*Cervus elaphus*)

Figure 1.1. Red deer. Clockwise from top-left: red stag in winter pelage; two red hinds moulting into winter pelage (photos taken by Clare Andrews); a red hind in summer; two red stags in summer coats during the rut (photos taken by Megan Wyman).

Approximately 30% of Europe's red deer reside in the British Isles. In wild land such as the Scottish uplands, they are a pivotal species as they impact the flora through grazing, browsing, trampling and seed dispersal and they are an economic resource in terms of stalking, tourism and venison production (Figure 1.1, (DCS 2008)). They have been present in Europe since the middle of the Pleistocene, constricted to particular refugia during the Last Glacial Maximum which ended ~10-11,000 BP, after which they recolonized northern and western parts of Europe and spread into areas of forest (Sommer *et al.* 2008). Subsequently, Mesolithic hunting and Neolithic farming reduced suitable habitat (~5000 BP) and triggered a population decline concluding at the end of the 1700s (Pérez-Espona *et al.* 2009a). Numbers increased in the British Isles (especially in the Scottish highlands) to high densities during the 19th century and have been

maintained to date by a combination of conservation, forestry protection and sporting incentives (Clutton-Brock *et al.* 1989; Pérez-Espona *et al.* 2009a).

Within Scotland there are currently an estimated 450,000 red deer primarily across the mainland and the Hebridean islands, but largely absent from the Central Belt and the south east region of the mainland (Clutton-Brock *et al.* 2004; Pérez-Espona *et al.* 2009a). In England, their distribution is far less continuous with isolated populations totalling around 16,000-20,000 animals (Pérez-Espona *et al.* 2009a; Ward 2007). In addition to the major demographic changes following the last glaciation, the resident red deer population in Europe and the British Isles has also been subject to a long history of introductions, translocations and management by humans. During the 19th century the popularity of deer hunting increased, as did the introduction of red deer from various European countries, primarily into deer parks but also into the wild ((Pérez-Espona *et al.* 2009a; Whitehead 1964); Further details in Chapter 2)).

Within Ireland there are thought to be around 4,000 red deer (Pérez-Espona *et al.* 2009a), which are descended from both ancient and recent postglacial introductions by man (Carden *et al.* 2012). The red deer centred in Killarney, Co. Kerry, are descended from a human introduction from Britain during the Neolithic period (Carden *et al.* 2012), whilst other populations are descended from more recent introductions from Britain and continental Europe (Carden *et al.* 2012; McDevitt *et al.* 2009a).

1.4.2 Sika (Cervus nippon)

Figure 1.2. Sika deer. Clockwise from top-left: Sika stag in winter pelage (photo taken by Josephine Pemberton); sika hinds in winter coat; a sika hind (left) next to a sika stag (right), both in summer coat (photos by Kevin McKillop).

Since the mid-19th century, numerous introductions of sika deer (*Cervus nippon*) from Japan have taken place across sites throughout the British Isles, mainly because they make attractive park deer (Ratcliffe 1987) (Figure 1.2). They were first introduced in 1860 by Viscount Powerscourt, to Powerscourt Estate, Co. Wicklow, and by the Zoological Society of London which acquired various specimens (named as *Cervus nippon nippon* and *C. n. hortulorum*) around the same time (Lowe & Gardiner 1975; Powerscourt 1884). Unfortunately, park sika have repeatedly escaped or been deliberately introduced to the wild. In the wild, sika are elusive and nocturnal, preferring forest habitats, where

they can be particularly damaging to tree plantations through browsing and bark stripping (Pérez-Espona *et al.* 2009a; Ratcliffe 1987).

Based on the mitochondrial haplotype they carry, the populations introduced to the UK probably came from the island of Kyushu. Nagasaki, on Kyushu, was the only Japanese port open to international trade around the time of export (Goodman *et al.* 2001). In 1860, one male and three female sika were introduced from Japan to Powerscourt where they successfully established and expanded their range (by 1884 there were over 100 in the park) and become the main source of sika which directly or indirectly stocked other sites throughout the British Isles (Powerscourt 1884; Ratcliffe 1987).

There are thought to be approximately 15,000-20,000 sika in Scotland (Pérez-Espona *et al.* 2009a), the distribution of which is attributed to twelve separate introductions from which animals escaped and became established (Ratcliffe 1987) and now they occupy around 40% (~ 14,000km^2) of Scotland (Lowe & Gardiner 1975; Pérez-Espona *et al.* 2009a; Ward 2005). Within England, there are an estimated 1,500-2,000 sika in the wild in more discrete populations (Díaz *et al.* 2006; Pérez-Espona *et al.* 2009a). In addition, up to 1,000 sika are held in parks throughout the UK, highlighting the continued risk of further escapes (Swanson & Putman 2009). Since their introduction to Ireland over 150 years ago, sika have been translocated around the country and have successfully established in the wild at several sites (Carden *et al.* 2010; Whitehead 1964).

Sika, therefore, exemplify a highly successful bioinvasion (Facon *et al.* 2006). In Kintyre, Scotland it is estimated that their population has expanded at a rate of 9.2% per year in terms of numbers with a dispersal rate of approximately 3.7km per year, whilst in Ireland sika have expanded the area of their range at around 5% per annum over the last 30 years (Carden *et al.* 2010; Goodman *et al.* 2001). In Japan sika deer can reach densities of up to around 40 deer/ km^2 (Ratcliffe 1987), similar to estimates of 14 – 44 deer/ km^2 in County Wicklow and up to 42 – 45 deer/ km^2 in parts of Scotland (Swanson & Putman 2009).

The successful establishment and expansion of Japanese sika populations in the British Isles has led to inevitable overlap with the range of the native red deer and has provided ample opportunity for hybridisation and introgression. Despite being separated by ~5–7 MY of independent evolution (Ludt *et al.* 2004) and with major phenotypic differences (Table 1.1), red and sika have been shown to hybridise in captivity (Harrington 1973) and in the wild in Scotland, Ireland and an area of the Lake District, by observation

(Ratcliffe 1987), skull morphometrics (Lowe & Gardiner 1975) and molecular approaches (Abernethy 1994; Goodman *et al.* 1999; Harrington 1973; Senn & Pemberton 2009).

1.5 Hybridisation between red and sika in the British Isles

This study of red-sika hybridisation may be restricted to that within the British Isles; however, it is by no means a problem exclusive to this set of countries. Sika have been introduced to many other countries. Hybridisation has also been demonstrated in the former Czechoslovakia (now Czech Republic and Slovakia) by craniological analyses (Bartos *et al.* 1981), phenotypic observation (Bartos & Zirovnicky 1981) and behavioural studies (Bartos & Zirovnicky 1982). Red-sika hybrids have also been confirmed by genetic analyses in many countries throughout Eastern Europe (Poland, Russia and Lithuania) and have also been observed in the wild in New Zealand (Biedrzycka *et al.* 2012; Davidson 1973). However, hybridisation between red and sika animals has perhaps been best documented in regions throughout the British Isles by various molecular methods. The situation in each region will now be described in turn.

1.5.1 Scotland

One of the best studied red-sika systems in the British Isles is on the Kintyre peninsula, Argyll (Abernethy 1994; Goodman *et al.* 1999; Senn *et al.* 2010a; Senn & Pemberton 2009; Senn *et al.* 2010b). A small founding population of nine female and two male sika deer were introduced to Carradale, Kintyre from Japan in 1893 at a time when red deer were absent or rare on the peninsula (Ratcliffe 1987). After escaping from their enclosure the wild sika population expanded and are thought to have made permanent contact with red deer expanding south down the peninsula from the mainland around the 1960s (Senn & Pemberton 2009). Since then hybridisation between these two species was first identified using four nuclear microsatellite loci on a sample of deer shot in 1990-1 (Abernethy 1994) and then by a more reliable panel of 11 microsatellite markers and more robust statistical approaches on deer shot in 1995-1996 (Goodman *et al.* 1999; Slate *et al.* 1998). Whilst the majority of animals analysed in this more recent study were found to be either red-like or sika-like based on their nuclear markers, up to 40% of deer were found to be introgressed at the site where the two species had overlapped for longest (Knapdale; <30km from West Loch Awe, 50km from Carradale), highlighting the uneven spatial distribution of hybrid activity (Goodman *et al.* 1999). The multiply-introgressed genotypes identified at this site and significant linkage disequilibrium amongst both red and sika populations was interpreted as evidence of

recent hybridisation. This is because the time elapsed since older hybridisation events allows segregation to break up the associations between introgressed alleles and produces a system closer toward linkage equilibrium (Goodman *et al.* 1999).

Most recently, Senn & Pemberton (2009) analysed 735 individuals shot in 2005-2006 from throughout Kintyre using a panel of 22 diagnostic microsatellites, a single mitochondrial marker and the analytical software Structure (Pritchard *et al.* 2000), providing the highest resolution analysis of the extent of hybridisation and introgression in this system to date. While hybrid animals were generally rare, at one site, West Loch Awe (WLA), 43% of individuals were hybrid. At WLA the number of hybrids appears to have exceeded a threshold and led to a breakdown in normal assortative mating patterns (red with red and sika with sika) and collapse into a hybrid swarm. This situation is proposed to have occurred very rapidly from as few as five hybridisation events since contact was established around the 1960s and no change in the extent of introgression was found recently over a 15 year period (Senn *et al.* 2010a). The pregnancy rates of hybrid individuals at this site do not differ significantly from either of parental species, suggesting the species are relatively compatible and fertile (Senn *et al.* 2010b).

The use of a mitochondrial DNA marker allows the maternal inheritance to be traced and the mtDNA type of F1 hybrids would allow the direction of initial hybridisation events to be determined. Senn & Pemberton (2009) identified no F1 hybrids in Kintyre (n = 735); however, 60 out of 61 examples of mitochondrial discordance were in sika-like animals with red deer mtDNA, lending support to the hypothesis that hybridisation events occur mainly between sika stags and red hinds. However, this cannot be confirmed due to the absence of F1 hybrids.

Regarding other parts of Scotland, previous genetic work with a small number of microsatellite markers (5-10) suggested that there may be extensive, but very low-level, introgression in parts of the North Highlands, particularly of red alleles into sika (Swanson 2000), but these results are unpublished and large areas of mainland Scotland remain unscreened. A previous genetic study of seven of the Hebridean islands (1998-1999; prior to those sampled 10 years later in 2009-2010, Chapter 2) also remain unpublished, but showed an absence of recent hybridisation with sika amongst 317 red deer analysed (Pemberton 2000).

1.5.2 England

The population of red deer in Grizedale in the Lake District is thought to be one of very few remaining of native English descent (*Cervus elaphus scoticus*) (Pérez-Espona *et al.* 2009a). Sika deer are reported to have been introduced to a site 50km south east of here, Rigmaden, later released into the Bowland area and during the early 1900's hybrids were observed (the first shot hybrid stag in 1944; (Lowe & Gardiner 1975)). Increasing hybrid activity in this region may have been directed westward by (biased) culling restrictions, afforestation and habitat regeneration efforts, inadvertently providing suitable corridors for sika expansion (Lowe & Gardiner 1975). Similar to the site at County Wicklow, this site in the south Lake District is thought to have experienced complete introgression and consist of hybrids only (Ratcliffe 1987). However, inferences to date have been based on the use of polymorphism in the blood protein transferrin and multivariate approaches on skull parameters (Lowe & Gardiner 1975; McDougal & Lowe 1968), highlighting the need for revision of this area using more robust and objective genetic methods.

The only other molecular study of hybridisation in England was conducted by Diaz *et al.* (2006), who reported low-level introgression of red nuclear DNA in a sika population in New Forest and Purbeck, refuting recent hybridisation and concluding that the species have remained genetically distinct at these sites. However, this study only used eight diagnostic microsatellite markers and sample sizes were small. The absence of apparent hybrids was attributed to stronger assortative mating or less sympatry compared to other populations (Díaz *et al.* 2006).

1.5.3 Ireland

Red and sika deer live sympatrically throughout many counties in Ireland, with the major strongholds being Co. Wicklow in the East, Co. Kerry in the south west and Co. Galway north to Co. Donegal in the north west.

County Wicklow contains the longest-standing example of a putative hybrid swarm between red and sika. In the early 20th century, a dwindling wild red deer population faced numerous escaped Powerscourt sika, which may have already been hybrids due to hybridisation in the park (Delap 1936; Harrington 1973; Whitehead 1964). Harrington (1973) devised a diagnostic rocket immunoelectrophoresis method for red and sika, and during a study initiated early in 1972, found no pure red deer in Co. Wicklow in over 200 animals sampled. A study using nine non-diagnostic microsatellite markers and Structure analysis on samples from Co. Wicklow suggested that the majority of

phenotypically red deer were actually hybrid (McDevitt *et al.* 2009a). Even though a substantial number of 'pure' sika were identified in the region using this panel of microsatellite markers, introgressive hybridisation is extensive and there are unlikely any remaining 'pure' red deer (Harrington 1979).

Sika were introduced from Powerscourt to Killarney, Co Kerry, the home of the oldest lineage of Irish red deer, in 1864, before documented hybridisation in Powerscourt (McDevitt *et al.* 2009a; Powerscourt 1884). Even with density ratios of 1:5 red to sika individuals interacting since the 1800's, hybrids have not been observed in County Kerry and both species appear reproductively isolated at this site, including when tested by the rocket immunoelectrophoresis and microsatellite methods applied to Co. Wicklow deer (Harrington 1973; McDevitt *et al.* 2009a). Historic reports have suggested the presence of hybrids across several counties other than these mentioned (McDevitt *et al.* 2009a; Whitehead 1964). Hybrid swarms, such as that documented in Co. Wicklow, may exist undetected elsewhere on the island of Ireland where these species overlap and could be expanding at a rate which threatens the genetic and phenotypic integrity of the ancient-origin red deer in Co. Kerry.

In summary, work to date in the British Isles suggests that red-sika hybridisation occurs infrequently and is typically followed by backcrossing into one or other parental species. However, where hybridisation does occur, introgression can sometimes be extensive and rapid and can breakdown into a hybrid swarm. It has been observed in numerous other systems that the frequency of hybrid formation is highly variable in nature (Hails & Morley 2005).

1.6 Studying hybridisation

1.6.1 Genetic markers

Documenting hybridisation by observational data appears to provide a delayed indication; in a particular time frame we could be observing the conspicuous phenotypic tip of a genotypically introgressed iceberg. The use of molecular techniques, however, has greatly facilitated the detection of hybridisation and introgression. Since the 1960s the means of detecting hybridisation has evolved from the use of morphological, physiological and biochemical markers, through the use of protein polymorphisms, towards one of robust molecular markers such as microsatellite markers and single nucleotide polymorphisms (SNPs).

This study adopts the use of 22 microsatellites and a single mitochondrial marker developed to discriminate between red deer and sika. Despite our panel of microsatellites being robust and effective, such genetic markers can have relatively low coverage and resolution, show variable rates of mutation, null alleles or can exhibit homoplasy (identical character states due to multiple mutation to the same allele size) at particular loci, which can lead to population structure being underestimated (Coates *et al.* 2009; Morin *et al.* 2004). Further, standardising allele sizes for comparison between laboratories (e.g. electrophoresis methods and specific standards used) can be difficult (Coates *et al.* 2009; Morin *et al.* 2004). Therefore, microsatellites are informative for population-level questions; however, analysis would be improved by more markers (Morin *et al.* 2004). The first improvement in marker panel could, therefore, be the detection and application of further diagnostic loci between red and sika deer either as single sequence repeats or single nucleotide polymorphisms.

The marker panel in this study also uses a mitochondrial DNA marker for inferred matrilineal phylogeny and introgression. In addition, our use of a single mitochondrial marker has helped identify past hybridisation and directionality; however, it represents a single, maternally-inherited marker with limited molecular resolution and may be experiencing different selection pressures to the nuclear genome (Hurst & Jiggins 2005; Twyford & Ennos 2012). An improvement to our panel could also be the compliment of a Y-chromosome marker, such as that looking for wapiti introgression in Scottish red deer (Pérez-Espona *et al.* 2010b). This would allow the paternal line to be traced and would complement the use of nuclear and mitochondrial markers (Isoda *et al.* 2000). However, Y-chromosome markers provide relatively low molecular resolution as they are a single non-recombining locus, and their inference power is compromised by high variability in male reproductive success, lowering the probability that they would persist as introgressed material in the opposite species (Twyford & Ennos 2012). A diagnostic Y-chromosome marker between *Cervus* species was not developed or used in this study due to weighting time and labour costs against the information it could provide, however, could be worth investing in with future work.

1.6.2 Statistical approaches to inferring population genetic structure

There are numerous population genetic methods for analysing hybrid populations based on genotypic data generated by the markers mentioned above; some of the most appropriate to the analysis of data in this study are described subsequently. It can be beneficial to try numerous different statistical approaches and compare the outcome in

order to better understand underlying structure (Marie *et al.* 2011). Our situation is made slightly more difficult by not being able to sample the allele frequencies for the parental sika population due to their uncertain purity and bottlenecked diversity in the British Isles. Also, population structure amongst the red deer across the sample sites may confound patterns found. It is lastly important to note that the quality of the data collected restrict the extent of biological inferences that can be made (Anderson *et al.,* 2002).

One of the most widespread approaches used to inferring population structure based on genetic data is the Bayesian clustering software Structure 2.3.3 (Falush *et al.* 2003, 2007; Pritchard *et al.* 2000). It calculates the number of inferred, genetically distinct populations (K) that maximises the likelihood (Ln Pr(X|K)) of the dataset, assuming they are in Hardy-Weinberg equilibrium and linkage equilibrium. This software provides a robust approach for inferring population structure, can incorporate prior information and account for null alleles, however, estimating the most likely number of populations (K) remains slightly subjective and suboptimal (Falush *et al.* 2007; Pritchard *et al.* 2000).

A second Bayesian approach is New Hybrid (Anderson & Thompson 2002). Whilst Structure uses an inheritance model to calculate the probability that an individual has recent ancestry in two or more populations, New Hybrid uses an inheritance model to estimate the probability of an individual belonging to a set of pre-defined parental species and hybrid categories. New Hybrid draws individual samples randomly from a system undergoing recent hybridisation and calculates the posterior probability that they belong to each of the different hybrid classes based on their genotypic data (Anderson & Thompson 2002). When both approaches mentioned so far were applied to various stocking scenarios of wild brook charr (*Salvelinus fontinalis*), Structure exhibited a higher efficiency in assigning individuals (a greater number of individuals assigned, less prone to fluctuations in hybrid number), whilst New Hybrid exhibited better assignment accuracy (Marie *et al.* 2011). It suggested an optimal approach was using both types of software in combination (Marie *et al.* 2011).

A third approach, Geneland (Guillot 2008; Guillot *et al.* 2005a; Guillot *et al.* 2005b; Guillot & Santos 2010; Guillot *et al.* 2008), has been developed to make much more use of the spatial coordinates of sampled animals, as well as their genotypic data to infer the underlying population genetic and gene flow structure using a spatial model. The Poisson – Voronoi tessellation approach is used to model the genetic and geographical information to approximate the spatial domain of each of K assumed populations by

the union of a few polygonal domains. This software provides an efficient method in the field of landscape genetics, by locating genetic discontinuities without prior knowledge and can detect migrants and uncertainty as illustrated by analysis of on wolverine populations in the northwest US (Guillot *et al.* 2005a).

Alternatively, discriminant analysis of principal components (DAPC) is a multivariate approach which uses the programming language R. 2.15 to identify the most likely population clusters by maximising between-population variation and minimising within-population variation and avoids any assumptions of an explicitly evolutionary model (Jombart *et al.* 2010). The graphical output from this approach can be used to generate a scatter plot, in which individuals are located according to coordinates determined by the principal component analysis. This software is deemed faster and more applicable than other approaches at inferring population subdivision, however, some of the underlying algorithms are rather simple and selecting the number of principal components is also slightly subjective (Jombart *et al.* 2010).

Overall, selecting the best software to use is likely to depend on the objective of the study. Whilst Structure may be better for assessing hybridisation in the wild, New Hybrid may be better for analysing the extent of hybridisation in managed systems (e.g. fisheries, stocking game birds) where hybridisation is known to occur. Ultimately, the different software packages calculate almost analogous parameters for proportion population membership; Structure generates Q for each individual, as a vector for admixture proportions, New Hybrid also has Q, the "genetic heritage proportion" and Geneland, the vector *p,* for parameterising the population memberships (Anderson & Thompson 2002; Guillot 2008; Pritchard *et al.* 2000). Numerous studies concluded that the resolution and efficiency of an inference is higher with the proportion of hybridisation in the system being analysed (Marie *et al.* 2011).

1.7 This study

This project seeks to apply the highly diagnostic panel of 22 microsatellites and a single mitochondrial DNA marker, developed by Senn & Pemberton (2009), to a much wider geographical area, in order to give a more extensive and uniform account of the extent of hybridisation between these species in the British Isles. The Bayesian clustering software Structure 2.3.3 (Pritchard *et al.* 2000) was used to analyse the extent of individual and population admixture using the microsatellite genotype data in datasets. It calculates the number of inferred, genetically distinct populations (K) that maximises the

likelihood (Ln Pr(X|K)) of the dataset, assuming they are in Hardy-Weinberg equilibrium and linkage equilibrium. Running the nuclear genotype data in Structure when two populations are assumed, red and sika populations differentiate and the proportion membership of each species assigned to red ancestry (Q) was used as a hybrid score. This scale was such that $Q = 1$ represents a ‘pure’ red and $Q = 0$, ‘pure’ sika and a hybrid defined as an individual with a Q value of $0.05 \leq Q \leq 0.95$ (Senn & Pemberton 2009). However, this project also explores population structure at a much more stringent definition of purity ($Q>0.99$ and $Q<0.01$) to investigate whether this removes further low-level introgression.

1.5.4 Objectives of this thesis

1. Determine the extent of hybridisation and introgression between exotic *Cervus* (*C. nippon* and *C. canadensis*) and red deer (*Cervus elaphus*) across Scotland and northwest England.
2. Determine the extent of hybridisation and introgression between exotic *Cervus* (*C. nippon* and *C. canadensis*) and red deer (*Cervus elaphus*) in Ireland.
3. Describe the genetic population structure of pure red and pure sika populations.
4. Explore the phenotypic consequences of introgression in hybrid deer.
5. Comment on how the information obtained might be enhanced and used to improve management of red and sika populations in the British Isles.

1.6 Acknowledgments

With thanks to Helen Senn for some of her data and to Clare Andrews, Josephine Pemberton, Megan Wyman and Kevin McKillop for photographs.

Chapter 2: Introgression of exotic Cervus (*nippon* and *canadensis*) into red deer (*Cervus elaphus*) in Scotland and northwest England.

Author's contributions:

Of the 2887 red and sika individuals analysed in this project, 737 were genotyped by Helen Senn from the Kintyre peninsula who also genotyped a further 121 from central Scotland. 726 samples were genotyped from the Outer Hebrides and 49 wapiti from Canada by the combined effort of Elizabeth Heap and Sheena Morrison. SS organised the sampling of and completed the genotyping of 568 samples from across the North highlands, a further 315 samples from particular sites in Kintyre (101 of which were provided by Megan Wyman who shared the process of their genotyping with SS) and 50 samples from the Cairngorm National Park. A further 233 samples were provided by Silvia Perez-Espona from central Scotland and were genotyped by SS. Lastly 137 tissue samples were provided by Eleni Socratous from the Lake District were genotyped by the combined efforts of Elizabeth Mittel, Sarah MacDonald and SS. Assistance in the laboratory was also provided by Philip Ellis. Statistical analysis of all samples was performed by SS. SS wrote the MS. JMP guided the study and edited and commented on the MS.

2.1 Abstract

Since the mid-19th century, multiple introductions of Japanese sika deer (*Cervus nippon*) and wapiti (*C. canadensis*) have taken place across Scotland and North West England. While wapiti introductions have generally gone extinct, sika have established and now occupy over 40% of the range of native red deer (*C. elaphus*) in Scotland. Hybridisation between these species has been demonstrated in captivity and in the wild. Using a panel of 22 microsatellite loci that are highly diagnostic for red-sika and strongly diagnostic for red-wapiti and a diagnostic mitochondrial marker for red-sika, we analysed 2943 deer using the Bayesian clustering software Structure 2.3.3 to investigate the extent of introgression between these species. There was some evidence for very low level introgression by wapiti into a small number Scottish red deer (but not Cumbrian red deer or Scottish sika). Despite large areas of sympatry, only two areas (both in Kintyre, Argyll) show extensive introgression with collapse of assortative mating between red and sika. However, rare and widely scattered individuals with low-level sika introgression in Cumbria and mainland Scotland and occasional individuals with cytonuclear disequilibrium (sika nuclear genetics with red mitochondrial genome) in Scotland suggest hybridisation has occurred in several places in the past without subsequent loss of assortative mating. The Hebridean red deer refuge appears free of sika introgression.

Key words: Cervus, microsatellite, mitochondrial, hybridisation, introgression, sika, wapiti, red.

2.2 Introduction

2.2.1 *Hybridisation*

Hybridization is the interbreeding of genetically distinct taxa and is widespread amongst eukaryotes (Allendorf *et al.* 2001). Introgression is the resultant gene flow (described as 'horizontal') between populations whose members are hybridising and can dramatically influence the evolutionary trajectory of a species (Allendorf *et al.* 2001). Whilst hybridisation can occur naturally, it may also be induced by human activity, which can cause taxa which were previously not in contact to become sympatric. This can be detrimental to native or endemic species. The generation of hybrids between an invasive and a native without introgression (e.g. due to genetic incompatibility or strong negative selection against F1 hybrids) can result in substantial wasted reproductive costs, whilst if introgression does occur, it can be highly destructive and has led to the extinction of numerous species, races and locally adapted ecotypes (Allendorf *et al.* 2001). Examples of anthropogenically-induced hybridisation include that between Antarctic fur seals (*Arctocephalus gazelle, A. tropicalis*) and New Zealand fur seals (*A. forsterri*) threatening population homogenisation through the disturbances caused by seal harvesting, between endemic mouse lemur species from Southern Madagascar (*Microcebus* spp.) where deforestation has facilitated asymmetric gene flow, the introgression of maladaptive gene complexes into wild American mink (*Neovison vison*) from escaped domestic farmed American mink causing population decline and the generation of sterile hybrids between the native bull trout (*Salvelinus confluetus*) and the introduced brook trout (*S. fontinalis)* in North America, ultimately leading to the displacement of the former species (Gligor *et al.* 2009; Kidd *et al.* 2009; Lancaster *et al.* 2006; Leary *et al.* 1993).

2.2.2 Cervus in the British Isles

Within Scotland there are an estimated 450,000 red deer (*Cervus elaphus*) (Clutton-Brock *et al.* 2004). Whilst this population is far from threatened in terms of population size, introgression from an invasive species can break up locally adapted suites of genes or introduce novel traits, leaving the native population vulnerable to demographic fluctuations. Since the mid-19th century, a series of introductions of exotic deer including North America wapiti (*Cervus elaphus canadensis* or *C. canadensis,* not to be confused with the European elk, *Alces alces)* and Japanese sika (*C. nippon nippon*) into the British Isles has created many opportunities for hybridisation with the red deer.

North American elk or wapiti have been introduced to a few widely-spaced sites in Scotland in attempts to supplement red deer stocks and improve trophy quality

(Polziehn & Strobeck 1998; Whitehead 1964). The first individual introduced to Scotland is suspected to have been to Dunkeld, in the early 1800s, by a former Duke of Atholl (Whitehead 1964). During the 1890s wapiti were introduced and cross bred with red deer at Monymusk, Aberdeenshire, until the remaining herd (c. 30 animals) were translocated to Mamore forest, Inverness-shire in the early 1900s (Whitehead 1964). In England, wapiti were introduced to Derby around the 1790s, and herds were kept in Woburn (Bedfordshire), Buckinghamshire, Kent, Sussex and Northamptonshire. Around the turn of the 20th century wapiti were introduced to Rigmaden Park, Cumbria (Whitehead 1964). Hybridisation between red and wapiti has occurred both in the wild in Scotland (Whitehead 1964) and in captivity (Moore & Littlejohn 1989; Shackell *et al.* 2003). Introgression with wapiti causes changes from the conventional phenotype of the red stag (Whitehead 1964). Overall, however, the impact of wapiti has been limited; the wet British climate renders them highly susceptible to lung disease and foot malformation and they show delayed female maturity and lower levels of stag aggression than red deer in the rut (Asher *et al.* 2005; Pérez-Espona *et al.* 2010a). This poor acclimatisation is corroborated by the absence of wapiti Y chromosome haplotypes in a recent survey of red deer from Mamore and adjacent areas (Pérez-Espona *et al.* 2010b).

Numerous introductions of sika deer from Japan have also taken place across sites throughout the British Isles, including Scotland, as an ornamental species for deer parks (Ratcliffe 1987). It likely there are now more than the 15,000 – 20,000 sika in Scotland, the distribution of which is attributed to twelve separate episodes of introduction, release or escape (Pérez-Espona *et al.* 2009a; Ratcliffe 1987). It is estimated that they occupy around 40% (~ 14,000km^2) of Scotland; however, they have a patchier distribution in England (Díaz *et al.* 2006; Pérez-Espona *et al.* 2009a; Ratcliffe 1987; Ward 2005). This has led to inevitable overlap with the range of the native red deer and has provided an opportunity for hybridisation and introgression. Despite being separated by ~5 – 7 MY of independent evolution (Ludt *et al.* 2004) and with major phenotypic differences (summarised in Table 1.1, Chapter 1 originally from Senn & Pemberton, 2009), red and sika have been shown to hybridise in captivity (Harrington 1973) and the wild in Scotland (Goodman *et al.* 1999; Senn & Pemberton 2009) and Cumbria (Lowe & Gardiner 1975). We note here that a further sika subspecies, Manchurian sika (*C. n. mantchuricus)* has been introduced to England and Ireland and hybrids subsequently identified, however there are no concrete reports of their introduction to Scotland or Cumbria (Powerscourt 1884; Ratcliffe 1987; Whitehead 1964) and they are not addressed further in this chapter.

One of the best-studied examples of red-sika hybridisation in the wild in Scotland is that which has occurred in Kintyre, Argyll. A small founding population of sika deer was introduced to this peninsula in 1893, which expanded in sympatry with red, and hybrid animals were subsequently observed (Goodman *et al.* 1999; Ratcliffe 1987; Whitehead 1964). More recently, Senn & Pemberton (2009) analysed 735 red deer and sika sampled from throughout Kintyre and using 22 strongly diagnostic microsatellite markers and a diagnostic mitochondrial marker, found that while hybrids were generally rare, at one site, West Loch Awe (WLA), 43% of the samples were from hybrids. Here, the number of hybrids appears to have exceeded a threshold, leading to a breakdown in normal mating patterns (red with red and sika with sika) and a collapse into a so-called 'hybrid swarm'. The pregnancy rates of hybrid individuals at this site do not differ significantly from either of parental species, suggesting they are relatively compatible and fertile (Senn *et al.* 2010b). Regarding other parts of Scotland, previous genetic work suggests that there may be extensive, but very low-level, introgression in parts of the North Highlands, particularly of red alleles into sika (Swanson 2000), but large areas remain unscreened.

Sika have also been introduced to England and red-sika hybrids subsequently reported, however, studies have been less extensive and lacked power. Following the introduction of Japanese sika during the late 1800s to the Lake District, Cumbria, hybridisation with red deer was reported in and around Rigmaden Park during the 1920s (Lowe & Gardiner 1975; Ratcliffe 1987). The adjacent red deer population in Grizedale, are thought to be one of the few remaining of native English descent (*Cervus elaphus scoticus*), and its genetic integrity would be at risk from nearby hybrids (Pérez-Espona *et al.* 2009a). However, existing evidence of hybridisation comes from multivariate analyses on a suite of skull parameters collected after the 1950s (Lowe & Gardiner 1975), which may not reliably indicate the extent of hybridisation, such that introgression in this area has yet to be assessed with genetic approaches amongst current populations.

Since evidence for red-sika hybridisation exists, it is possible that many ecological, fitness-related traits could introgress between species, altering their ecology and ultimately, management (Senn *et al.* 2010b). Sika have a longer rutting season than red deer, can live on a much poorer quality and more fibrous diet and exhibit a greater resistance to lungworm, *Elaphostrongylus* spp (Bohm *et al.* 2006; Chadwick *et al.* 1996). On the other hand, the introgression of red deer genes could improve the compatibility of sika animals to Scottish conditions and facilitate their already successful colonisation (Harrington 1973). Animals with an intermediate appearance may, in turn, promote further hybridisation and cause assortative mating to break down (Senn *et al.* 2010b).

Whilst based on low numbers of samples and crude analyses, Lowe & Gardiner (1975) reported intermediate craniological morphology for putative hybrids in the Lake District, whilst in a well-documented hybrid swarm in County Wicklow, Ireland, Harrington (1973) observed a general "coalescence" of form and colour, with a bias toward a sika-like phenotype amongst hybrids. In a wild red-sika study system in Kintrye, Argyll, Senn (2010b) performed regression of phenotypic trait values against genetically-determined hybrid scores to quantify the impact of hybridisation on phenotype. Carcass weight was greater in sika-like hybrids than in 'pure' sika and lower in red-like hybrid females than in 'pure' red females. Within sika-like females, hybrids had increased jaw length and incisor arcade breadth (IAB) compared with 'pure' sika, whilst IAB was low in red-like hybrid females compared to 'pure' red (see below for definition of 'pure'). Overall, phenotypic modifications such as these highlight the (additive) genetic variation for quantitative traits in hybrid deer and the substantial potential for change under selection. This can greatly exacerbate effective management of these populations.

2.2.3 *This study*

This study samples red and sika deer from four main regions of Scotland and one in the North West of England; namely Kintyre, the Central highlands, the North highlands, the Hebrides and within the Lake District. It then seeks to investigate the extent of hybridisation between the introduced *Cervus* species, wapiti and Japanese sika, with native red deer at these sites using the most powerful marker panel for these heterospecifics to date. The specific aims of this study were:

1. *To assess the current extent of hybridisation and introgression between introduced* Cervus *deer and native red deer across Scotland and a region of the Lake District.*
2. *To determine, if possible, the initial direction of hybridisation.*
3. *To compare the outcome of this study with other forms of anthropogenically-induced hybridisation and try to identify causative factors.*
4. *To indicate what future management actions may be required to protect putatively pure populations from hybridisation.*

2.3 Materials and Methods

2.3.1 *Study area and sampling*

The study area consisted of four major regions throughout Scotland and the Lake District (Figure 2.1), covering many areas of sympatry between these two species (Ward 2005).

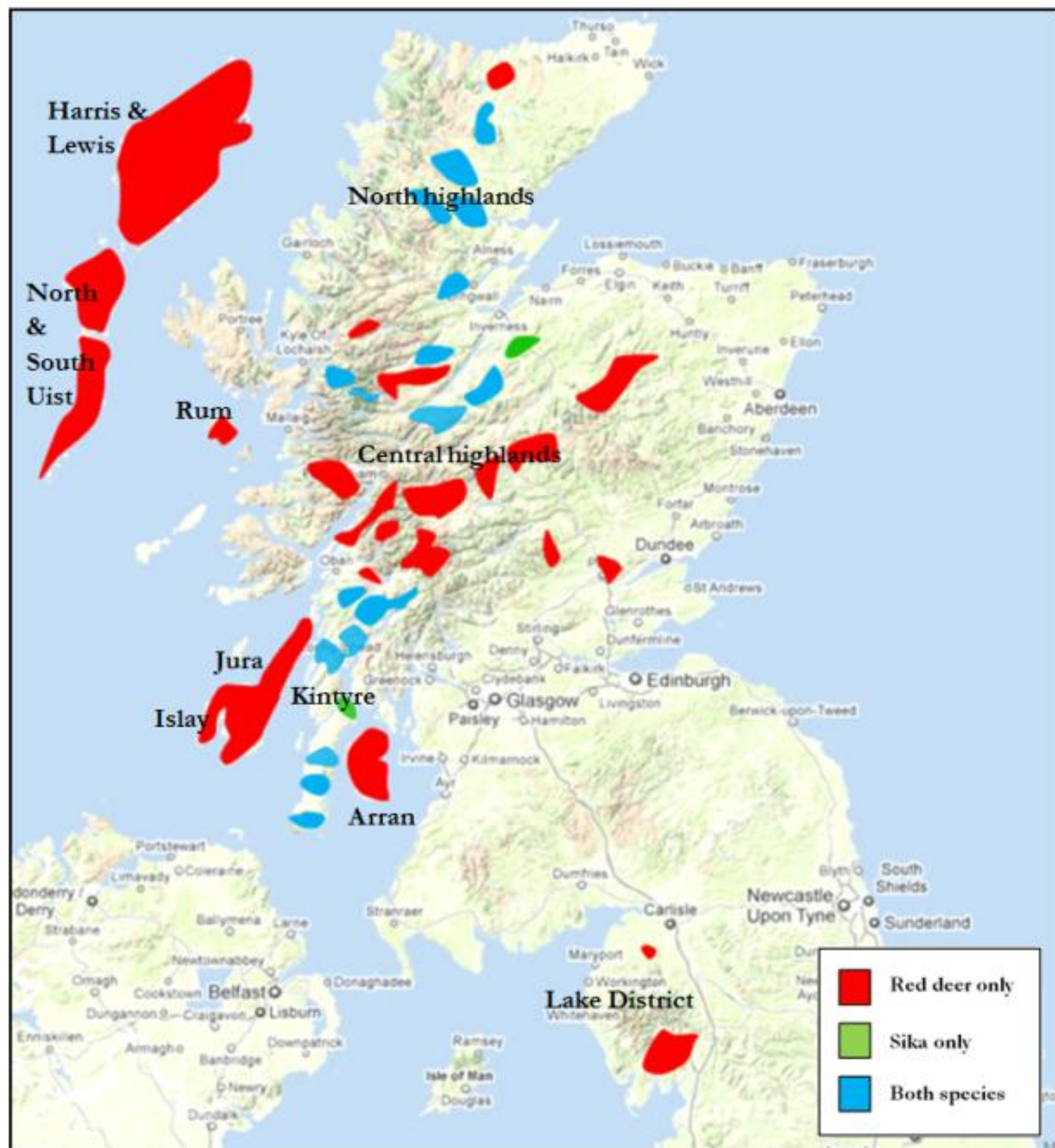


Figure 2.1. Map showing the sites from which samples were obtained; red shows those from which phenotypically red deer only were sampled, green from which phenotypically sika only were sampled and blue from which both species were sampled.

Further to the study conducted by Senn & Pemberton (2009) throughout the Kintyre peninsula, additional samples were collected from this region for the current study. A set of 735 previously studied animals (Senn & Pemberton 2009) were supplemented with 69 more individuals from South Kintyre, 163 more from WLA and adjacent sites and individuals from previously unsampled sites, including Ardchonnel (n = 36) and West Carron (n = 3). This gave a total of 1054 animals sampled from this region.

Tissue samples obtained from the central highlands included those from in and around the Cairngorm National Park and the Loch Lomond and the Trossachs National Park (n = 121, June 2008-May 2012) and from around the central highlands previously collected by Silvia Perez-Espona (n = 235, (Pérez-Espona *et al.* 2009b)). Those collected from the National Parks included animals sampled from nine estates within the parks (Abernethy, Ralia, Rothiemurchus, Inshariach, Breadalbane collected 2008-2009 and genotyped by Helen Senn and from Glen Spean, Craig Dhu, Glenbanchor, Kinveachy, collected June 2011-May 2012). Those from the Pérez-Espona *et al.* (2009b) study were obtained from ten open hill estates. This gave a total of 356 animals from the central highlands. Samples were also collected from a large area of the North Highlands (n = 514, June 2009-May 2010, further n = 58, May 2011), covering 18 Forestry Commission Scotland deer units.

Many of the islands in the Hebrides were designated as refugia for native red deer in 1999 to protect them from introgression from other *Cervus* species. This designation requires routine assessment of populations for sika introgression in order to confirm the efficiency of the protection, its management and extent. This study includes genotype information on 727 animals from across nine of the major islands of the Hebrides as part of this assessment.

All samples described thus far were collected by Forestry Commission Scotland rangers and estate stalkers from deer shot as part of their standard culling operations, during the appropriate season. Culling operations are assumed to be unbiased and where red and sika co-occur should, therefore, reflect their relative proportions (Goodman *et al.* 1999). This excludes those samples provided by Silvia Perez-Espona from the central highlands of Scotland for which red deer only were targeted. These samples all consisted of ear tips preserved in 100% ethanol.

A total of 137 red deer were obtained from the Lake District, Cumbria, provided by Eleni Socrates (PhD student), University of Leicester. These samples consisted of frozen tissue kept on dry ice.

Lastly, 49 extracted DNA samples from wapiti were obtained from Prof. D.W. Coltman, University of Alberta, Canada.

2.3.2 DNA analysis

Genomic DNA was extracted from samples with the DNeasy Blood and Tissue Kit (Qiagen) whilst some samples obtained from Pérez-Espona *et al.* (2009b) were extracted using the DNAace Spin Tissue Mini kit (Bioline), both according to the manufacturer's instructions. Individual samples were genotyped at a panel of 22 diagnostic microsatellite markers following previously-published protocols (Senn & Pemberton 2009), the details of which are given in Appendix Table 2.A1. Originally derived from cattle (*Bos taurus*) and sheep (*Ovis aries*), these markers have been selected to discriminate between red deer and Japanese sika because when used to genotype 44 red deer and 44 sika from diverse geographical locations, they shared no common alleles (Goodman *et al.* 1999; Slate *et al.* 1998). In addition, they also have some discriminatory power between red and wapiti (10/22 strongly diagnostic loci; J. Pemberton, pers. comm). PCR products were run on an ABI3730 capillary sequencer (Applied Biosystems), using the internal standard Genescan LIZ 500 (Applied Biosystems). Fragment analysis was carried out using Genemapper version 4.0 (Applied Biosystems).

Individuals were also screened for their haplotype in the mitochondrial control region which in some *Cervus* species includes a diagnostic number of 39bp tandem repeats: red deer have a single repeat while Japanese sika have three repeats and Manchurian sika have seven (Cook *et al.* 1999). Amplification followed a published protocol (Cook *et al.* 1999) and repeat number was determined by assay on 4% agarose gels stained with ethidium bromide (Goodman *et al.* 1999) where red deer have a 350bp band, and Japanese sika a 430bp band (80bp difference due to two extra 39bp repeats and variation in the length of flanking regions).

2.3.3 *To assess the current extent of hybridisation and introgression between introduced* Cervus *deer and native red deer across Scotland and a region of the Lake District (objective 1)*

The Bayesian clustering software Structure 2.3.3 (Falush *et al.* 2003, 2007; Pritchard *et al.* 2000) was used to analyse the extent of individual and population admixture using the microsatellite genotype data in both datasets. The number of inferred, genetically distinct populations (K) that maximises the likelihood (Ln Pr(X|K)) of the dataset, assuming they are in Hardy-Weinberg equilibrium and linkage equilibrium, was estimated by running five independent replicates at different values of K (1-8) and selecting the smallest value of K with the highest log likelihood (Ln Pr(X|K)), prior to it plateauing (Pritchard *et al.* 2000). A more objective approach for estimating the best value of K, ΔK, was also used. This parameter is related to the second order rate of change of the log likelihood and is estimated as the maximum rate of change in (Ln Pr(X|K)) between consecutive values of K (Evanno *et al.* 2005). Datasets were all run with the same parameters as in Senn & Pemberton (2009), namely the standard model of admixed ancestry (with the parameter α inferred from the data, using a uniform prior) and the model of correlated allele frequency ($\lambda = 1$), a burnin of 5×10^4 and a run length of 10^6 Markov chain Monte Carlo steps. Null alleles can cause deviation from Hardy-Weinberg equilibrium by causing a systematic pattern of missing genotype data and can jeopardise rates of hybridisation observed (Falush *et al.* 2003; Senn 2009). The frequency of null alleles were therefore estimated concurrently by incorporating a row of "999" values into the second line of the data set and activating the option RECESSIVE ALLELES = 1. This function enables Structure to 'suspect' particular alleles as null alleles if, for example, they exhibit allele-specific PCR failure. It will then then treat these suspected null alleles as recessive instead of missing data and estimate their frequency at each and every locus (Falush *et al.* 2007; Senn 2009). Structure output data were manipulated using the software Distruct (Rosenberg *et al.* 2002), for illustrative purposes.

Analysis by Structure 2.3.3 generated a Q value for each individual, which represents the estimated proportion of ancestry to each of K groups. When simulations are run at K = 2 (as is typical for hybridisation between two taxa), the Q values for membership to one of the two ancestral populations can be used as an index of the hybrid status of an individual; here Q = 0 represents a sika and Q = 1, a red. Delimiting the proportion of admixture that qualifies as a hybrid is difficult, principally due to the possibility that at some loci there may be ancestral allele sharing in the taxa under consideration. Here a hybrid was defined on the basis of nuclear markers as an individual returning a Q value

of $0.05 \leq Q \leq 0.95$ between two taxa, following previous practice (Senn & Pemberton 2009). A red-sika hybrid was also defined if the mtDNA haplotype was discordant with a 'pure' nuclear genotype (i.e. red mtDNA in an animal with $Q<0.05$ or sika mtDNA in animals with $Q>0.95$). This latter type of hybrid indicates introgression beyond the resolution of the nuclear markers.

Various datasets were analysed sequentially using Structure 2.3.3, addressing each of the study aims. Initially, all red and sika samples from Scotland and Cumbria together with the 49 wapiti were analysed to resolve the most likely population structure that recognises these three species (analysis 1, n = 2943). Secondly, the wapiti controls and any individuals showing evidence of wapiti introgression were excluded, in order to assess the extent of red-sika hybridisation only across Scotland and Cumbria (analysis 2, n = 2887). In the third dataset all 'pure' sika and red-sika hybrids were removed such that only red deer and wapiti individuals remained, to identify the presence of any introgression between these two species (analysis 3, n = 2230). Lastly, all 'pure' sika and wapiti individuals were analysed to investigate the remote possibility of wapiti-sika introgression (analysis 4, n = 571).

The average number of alleles and genetic diversity indices for each of the three species at each of the 22 microsatellite loci and within each population, respectively, were also calculated using CERVUS 3.0 (Kalinowski *et al.* 2007).

2.3.4 *To determine, if possible, the initial direction of hybridisation (objective 2)*

The direction of initial hybridisation events (i.e. which taxon was the female parent) can only be assessed from cytonuclear data in F1 hybrids. An F1 individual should have a Q close to 0.5 in a K=2 Structure analyses and it should be heterozygous for red and sika alleles at all loci. In order to determine whether we had sampled any F1 hybrids we examined the posterior allele frequencies for the parental taxa generated by Structure following analysis 2 and assigned these as red-specific, sika-specific or inconclusive, according to conservative criteria (Appendix Table 2.A2). The genotypes of hybrids were recoded according to the origin of each allele at each locus to determine the proportion of loci that were red-sika heterozygous relative to all loci genotyped in that individual.

2.4 Results

2.4.1 Genotypes

In total, 2943 individuals were successfully genotyped for at least 20 out of 22 of the nuclear loci (Table 2.1). Genetic diversity indices are given for each locus (Table 2.2) and within each population (Table 2.3), for red deer, sika and wapiti. Almost 9% of all individuals failed to amplify at the single locus TGLA337 (predominantly animals from the Hebrides), accounting for almost 40% of the missing data.

Table 2.1. Sample sizes, stalker-assigned phenotypes and genetic data set completeness for the 2943 individuals successfully genotyped (at least 20 out of the 22 markers genotyped), shown for the five regions sampled and the wapiti controls.

Site, County	Total number of individuals	No. of phenotypic red	No. of phenotypic sika	No. of phenotypic wapiti	No. of phenotypic hybrid animals	No. animals not assigned phenotype	Nuclear dataset (% complete)	MtDNA dataset (% complete)
Kintyre, Scotland	1054	671	309	0	32	42	99.71	98.96
Central Highlands, Scotland	406	366	1	0	0	39	98.9	99.75
Refugia, Scotland	727	727	0	0	0	0	97.49	100
North Highlands , Scotland	570	256	206	0	35	73	99.78	99.65
Cumbria, England	137	137	0	0	0	0	98.77	100
Canada	49	0	0	49	0	0	99.63	NA

Table 2.2. Genetic diversity indices for each of the 22 loci in our microsatellite marker panel in phenotypic red deer (n = 2164), sika (n = 521) and wapiti (n = 49) calculated in Cervus 3.0. Subscripts $_{r, s, w}$ represent parameters calculated in red, sika and wapiti datasets independently. Parameters are k, the number of alleles at each locus in each species, N, number of samples typed at each locus, H_o, observed heterozygosity, H_e, expected heterozygosity and Null, the frequency of null alleles at each locus, after Table 3 in Senn & Pemberton (2009).

Locus	k_r	N_r	$H_{O\,r}$	$H_{E\,r}$	$Null_r$	k_s	N_s	$H_{O\,s}$	$H_{E\,s}$	$Null_s$	k_w	N_w	$H_{O\,w}$	$H_{E\,w}$	$Null_w$
AGLA293	3	2122	0.194	0.281	0.1884	3	520	0.094	0.118	0.1237	3	49	0.163	0.19	0.0612
BM4006	4	2153	0.336	0.405	0.1054	4	521	0.098	0.126	0.1294	1	49	0	0	ND
BM6438	7	2134	0.497	0.596	0.0798	6	518	0.411	0.559	0.1517	3	49	0.224	0.27	0.0779
BM757	16	2163	0.612	0.668	0.0486	11	519	0.145	0.185	0.1544	7	49	0.857	0.819	-0.0295
BOVIRBP	10	2153	0.66	0.762	0.0729	7	521	0.088	0.118	0.1638	5	48	0.563	0.714	0.1216
FCB193	20	2110	0.768	0.87	0.0616	13	519	0.191	0.213	0.0746	9	49	0.592	0.552	-0.0578
FSHB	28	2146	0.831	0.9	0.0402	15	519	0.277	0.382	0.1934	5	49	0.633	0.552	-0.0815
IDVGA29	3	2124	0.432	0.451	0.0209	3	521	0.136	0.175	0.129	1	49	0	0	ND
IDVGA55	11	2109	0.731	0.801	0.045	9	520	0.242	0.273	0.0574	2	49	0.469	0.504	0.0307
INRA5	2	2160	0.01	0.014	0.1193	5	521	0.175	0.19	0.0386	1	49	0	0	ND
INRA6	6	2162	0.414	0.454	0.0461	4	520	0.102	0.151	0.2514	2	49	0.245	0.217	-0.0603
INRA131	9	2164	0.536	0.575	0.0367	8	521	0.238	0.301	0.1204	3	49	0.51	0.532	0.0101
MM012	5	2163	0.327	0.363	0.0523	3	520	0.181	0.245	0.1606	3	49	0.653	0.535	-0.1232
RM12	12	2149	0.755	0.862	0.0653	7	521	0.069	0.078	0.0797	4	49	0.286	0.345	0.0764
RM188	16	2144	0.633	0.748	0.0866	14	518	0.595	0.647	0.0386	4	49	0.224	0.316	0.1769
RM95	14	2153	0.759	0.829	0.044	12	519	0.189	0.287	0.2294	7	49	0.735	0.796	0.0362
RME025	9	2155	0.326	0.357	0.052	3	518	0.087	0.114	0.1521	3	49	0.347	0.408	0.082
TGLA40	10	2156	0.519	0.634	0.1012	6	520	0.238	0.343	0.1786	2	49	0.449	0.444	-0.0103
TGLA126	7	2164	0.021	0.045	0.3283	4	521	0.493	0.571	0.0686	2	49	0.265	0.34	0.1186
TGLA127	14	2160	0.707	0.8	0.0621	8	520	0.34	0.489	0.1823	4	49	0.408	0.497	0.088
TGLA337	12	1930	0.648	0.794	0.1004	9	516	0.314	0.509	0.2661	3	47	0.277	0.33	0.08
UWCA47	4	2150	0.14	0.178	0.1115	3	521	0.115	0.172	0.2019	1	49	0	0	ND

Table 2.3. Genetic diversity indices within each population for phenotypic red deer, sika and wapiti calculated in Cervus 3.0. Parameters H_o and H_e represent observed heterozygosity and H_e, expected. heterozygosity respectively.

Species	*Population*	*Sample Size*	*Mean No. alleles per locus*	H_E	H_O
Red	Kintyre, Scotland	677	7.91	0.5688	0.529
	Central highlands, Scotland	368	7.27	0.5507	0.524
	Arran, Scotland	61	4.36	0.5071	0.429
	Islay, Scotland	162	5.77	0.4921	0.484
	Jura, Scotland	198	6.05	0.4842	0.466
	Rum, Scotland	20	4.27	0.4542	0.436
	Scarba, Scotland	7	3.32	0.489	0.482
	North and South Uist	88	5.41	0.5081	0.442
	Harris and Lewis, Scotland	190	3.77	0.3272	0.324
	North highlands, Scotland	256	8.5	0.5898	0.544
	Lake District, England	137	5.36	0.4686	0.469
Sika	Argyll, Scotland	314	5.5	0.256	0.172
	Central highlands, Scotland	1	1.18	0.1818	0.182
	North highlands, Scotland	206	5.77	0.3456	0.292
Wapiti	Canada	49	3.5	0.383	0.358

2.4.2 *To assess the current extent of hybridisation and introgression between introduced* Cervus *deer and native red deer across Scotland and a region of the Lake District (objective 1)*

Analysis 1: All individuals including Canadian wapiti controls (n = 2943)

The logarithmic likelihoods calculated in Structure revealed K = 2 was the smallest number of genetic clusters that was optimal to describe the majority of the population structure, with an average Ln Pr(X|K) (natural logarithm of the probability of data X, conditional on K) of -14603.8 (s.d. 10.7) and a rate of change of 2506.9 (Figure 2.2). At this value of K, as might be predicted from the choice of marker, red and Japanese sika are differentiated, but not wapiti, which cluster with red (see Appendix, Figure 2.A1). Due to strong population differentiation between some of the Hebridean red deer from mainland red deer, only at K=4 (-140655.48, s.d. 661.36; rate of change of 7.7) do the wapiti individuals become differentiated from red and sika (Figure 2.3). Therefore, whilst analysis suggests K=2 most likely describes the structure, for our purposes K=4 is appropriate. However, for presentation in Table 2.4 and Figure 2.4 and the allele frequencies shown in Appendix Table 2.A3, the two red deer population clusters were combined. The variation in the log likelihood generated during replicated simulations at the same value of K may be attributed to slight variation in the sampling (or "mixing") of the Markov chain, as part of the Bayesian analysis, when converging on the posterior distribution of each of the required parameters (Pritchard *et al.* 2000).

From these analyses, it is initially apparent that there are no three-way hybrids between red, sika and wapiti, according to our marker panel (Table 2.4, categories 6, 9 or 12). Only 7 individuals in the dataset showed signs of recent wapiti introgression and these were red-like individuals from Kintyre (n = 2), the Central highlands (n = 2), the Hebrides (n = 1) and North Highlands (n = 2) (Figure 2.4c). No sika individuals sampled gave an indication of recent wapiti introgression. This infrequent wapiti introgression is in contrast to the identification of 98 individuals classified as 'red-like hybrid with recent sika ancestry' (category 4) and 78 'sika-like hybrids with recent red ancestry' (Table 2.4, category 7). Whilst 6.1% of individuals provide evidence for red-sika hybridisation, only 0.24% of individuals show evidence of red-wapiti hybridisation. Since it is possible that the inclusion of wapiti genotypes could confound the analysis of red-sika hybridisation, in analysis 2 we repeated the analysis after removing the 49 wapiti control samples and the seven deer with evidence of wapiti introgression.

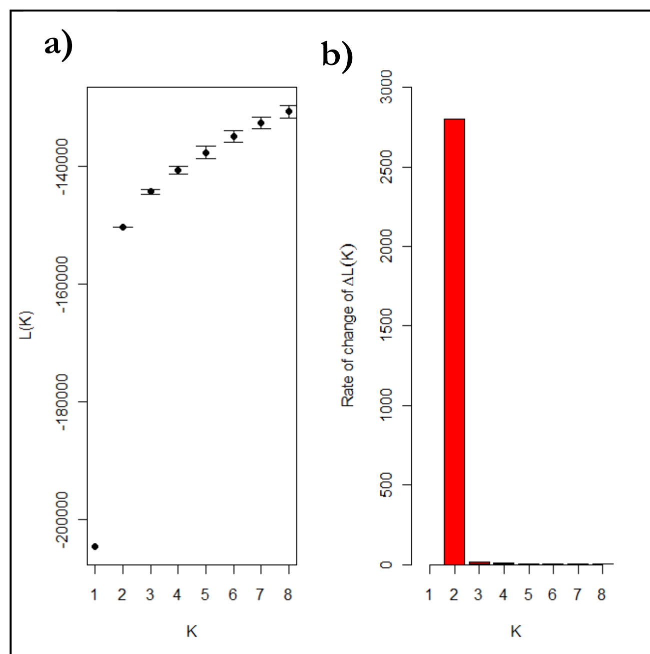


Figure 2.2. Assessment of the most likely number of populations using Structure 2.3.3 analysis 1 of dataset containing all red deer, sika and wapiti (n=2,943) at K = 1 – 8. Two likelihood parameters are assessed; of which the results for a) the log-likelihood (with standard error) of the each value of K (number of populations) given the dataset and b) the rate of change in log likelihood between values of K. Both provide evidence that K = 2 are the most likely.

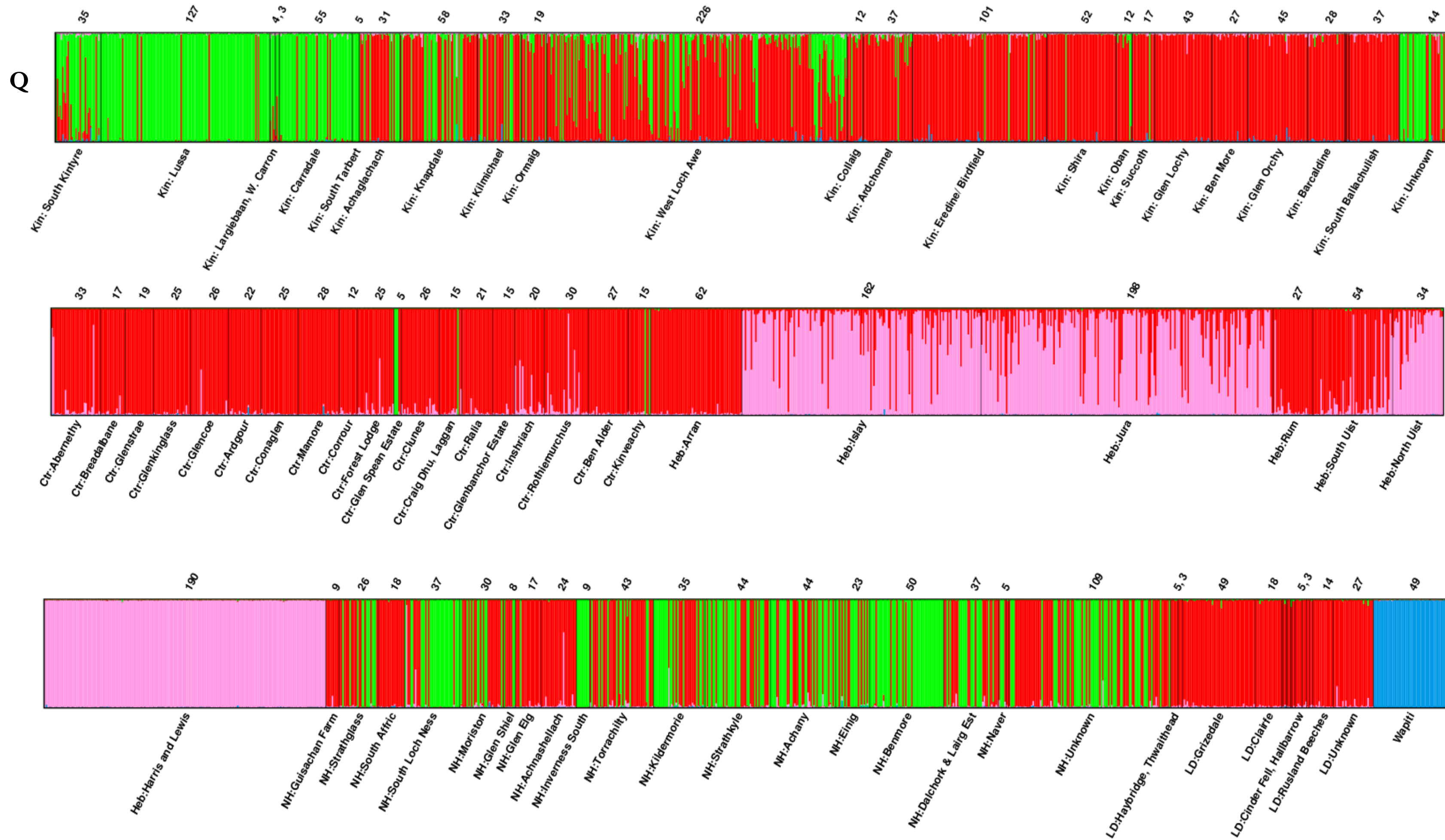


Figure 2.3. Bar chart showing the results of *analysis 1* in STRUCTURE at K = 4 for each individual in the dataset consisting of red deer, sika and wapiti animals (n = 2943). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (shown in red and pink) and the proportion attributable to sika ancestry (green) and to wapiti ancestry (blue). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis. Scottish sites are plotted in an approximately south to north order, followed by the sample sites in the Lake District, Cumbria and lastly the wapiti controls. Abbreviations represent; Kin= Kintyre, Ctr= Central highlands, Heb= Hebrides, NH= North Highlands and LD= Lake District, Cumbria.

Table 2.4. Admixture classification of all individuals, based on Q value from analysis 1 in Structure at K = 4 (allele frequency in red cluster I and II combined) based on a breakdown of the classification approach of Senn & Pemberton (2009).

Estimated membership to sika	Estimated membership to wapiti	Category	Scotland				England	Total
			No. (& %) of animals from Kintyre	No. (& %) of animals from Central Highlands	No. (& %) of animals from Hebrides	No. (& %) of animals from North Highalnds	No. (& %) of animals from Cumbria	No. (& %) of animals from all sites
$0 \le Q < 0.05$	$0 \le Q < 0.05$	'pure' red	618 (58.63)	398 (98.03)	726 (99.86)	299 (52.46)	134 (97.81)	2175 (73.90)
$0.90 < Q \le 1$	$0 \le Q < 0.05$	'pure' sika	265 (25.14)	6 (1.48)	0 (0)	265 (46.49)	0 (0)	536 (18.21)
$0 \le Q < 0.05$	$0.90 < Q \le 1$	'pure' wapiti	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	49 (1.66)
$0.05 \le Q < 0.50$	$0 \le Q < 0.05$	red-like hybrid with recent sika ancestry	94 (8.92)	0 (0)	0 (0)	1 (0.18)	3 (2.19)	98 (3.33)
$0 \le Q < 0.05$	$0.05 \le Q \le 0.50$	red-like hybrid with recent wapiti	2 (0.19)	2 (0.49)	1 (0.14)	2 (0.35)	0 (0)	7 (0.24)
$0.05 \le Q \le 0.95$	$0.05 \le Q \le 0.95$	red-like hybrid with recent sika and recent wapiti ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
$0.50 \le Q < 0.95$	$0 \le Q < 0.05$	sika-like hybrid with recent red ancestry	75 (7.12)	0 (0)	0 (0)	3 (0.53)	0 (0)	78 (2.65)
$0.50 \le Q < 0.95$	$0.05 \le Q < 0.50$	sika-like hybrid with recent wapiti	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
$0.50 \le Q < 0.95$	$0.05 \le Q \le 0.95$	sika-like hybrid with recent red and recent wapiti ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
$0 < = Q < 0.05$	$0.50 \le Q < 0.95$	wapiti-like hybrid with recent red ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
$0.05 \le Q < 0.50$	$0.50 \le Q < 0.95$	wapiti-like hybrid with recent sika ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
$0.05 \le Q < 0.95$	$0.50 \le Q < 0.95$	wapiti-like hybrid with recent red and recent sika ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

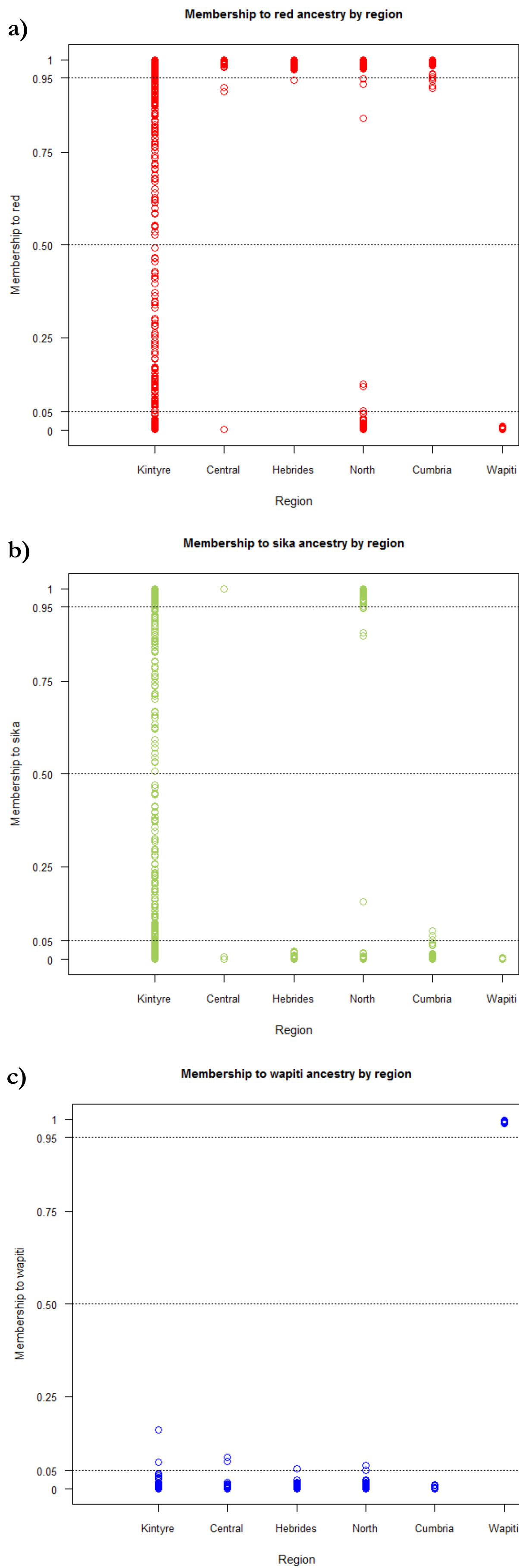


Figure 2.4. The proportion of inferred a) red, b) sika and c) wapiti ancestry determined by Q value generated in analysis 1 in Structure at K=4 (however red cluster I and II were combined for this figure) for the four regions in Scotland, in Cumbria and the wapiti control samples.

Analysis 2: All animals excluding Canadian wapiti controls and wapiti-introgressed animals (n = 2887)

The logarithmic likelihoods calculated in Structure support K = 2 as the smallest number of genetic clusters that describes the majority of the population structure, with an average Ln Pr(X|K) of 143907.42 (s.d. = 17.63) and rate of change of 3007.69 (Figure 2.5). Plots of Q values are given in figure 2.6 and 2.7, with allele frequencies for population clusters at K = 2 given in Appendix Table 2.A2. In practice there was little detectable difference in red-sika Q values compared with those in analysis 1. Considering first Kintyre, in total, 617 (58.7%) of the deer from this area were 'pure' red, 270 (25.6%) were 'pure' sika and 165 (15.7%) were hybrid. The main hotspots of hybrid activity are in South Kintyre and at WLA although low numbers of nuclear or mitochondrial hybrids were also detected in the north of Kintyre, the North highlands and the Lake District (Figures 2.6 and 2.7).

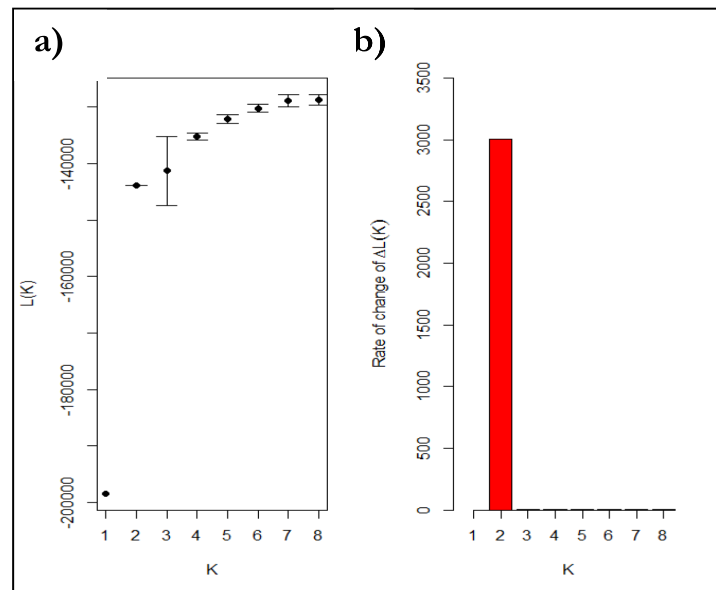


Figure 2.5. Assessment of the most likely number of populations using Structure 2.3.3 analysis 2 of dataset containing red deer and sika individuals only (n = 2,894) at K = 1 – 8. Two likelihood parameters are assessed; of which the results for a) the log-likelihood (with standard error) of the each value of K (number of populations) given the dataset and b) the rate of change in log likelihood between values of K. Both provide evidence that K = 2 are the most likely.

Samples obtained from in and around the Cairngorm National park (n = 171) appeared free from hybridisation. A subset of these samples (n = 50) were obtained in the most recent cull of the western borders of the park, where interaction with pioneering sika was considered most likely. Whilst the older samples (n = 121) we all 'pure' red

individuals (Q > 0.99), the additional 50 individuals were made up of six 'pure' sika (Q < 0.01), sampled from three of the estates and 44 'pure' red (again Q>0.99), confirming the lack of hybrid activity (Figures 2.6 and 2.7). Throughout the more western regions of the central highlands, a further 233 individuals were sampled (provided by Silvia Perez-Espona), all of which were identified as 'pure' red (Q≥0.99, Figures 2.6 and 2.7).

Similarly, all the 726 individuals sampled from the Hebrides were identified as 'pure' red individuals (Q>0.95), such that there was no evidence for recent introgressive hybridisation with sika in the islands forming the red deer refugia off Scotland's west coast (Figures 2.6 and 2.7).

The 568 individuals from throughout the North Highlands could be broken down into 299 'pure' red deer, 266 'pure' sika and three individuals showing signs of recent hybridisation (Torrachility, n = 2, Benmore, n = 1; Figures 2.6 and 2.7), according to our markers. Therefore, considering the sample size, we found little evidence of recent hybridisation (0.53% of individuals) in these samples, however, the Q values assigned to two of the three hybrids had 90% confidence intervals that did not overlap with zero or 1 (Torrachility, Q=0.124 (0.0350 -0.2350) and Q=0.841 (0.7330 – 0.9290)) vindictive of genuine hybrid status.

Similarly, three out of the 137 individuals from Cumbria returned Q values which fell within our definition of a hybrid (2.2%, Grizedale, Brigsteere and an unknown region in Cumbria; Figures 2.6 and 2.7), two of which had 90% confidence intervals that did not overlap with zero or 1(Grizedale, Q = 0.925 (0.840 – 0.988), Brigsteere, Q = 0.933 (0.855 – 0.987)) whilst all the remaining animals were 'pure' red (Q>0.95), according to our markers. The Q values assigned to the three hybrid animals suggested they were red-like with low-level sika introgression (0.925, 0.933 and 0.949 respectively).

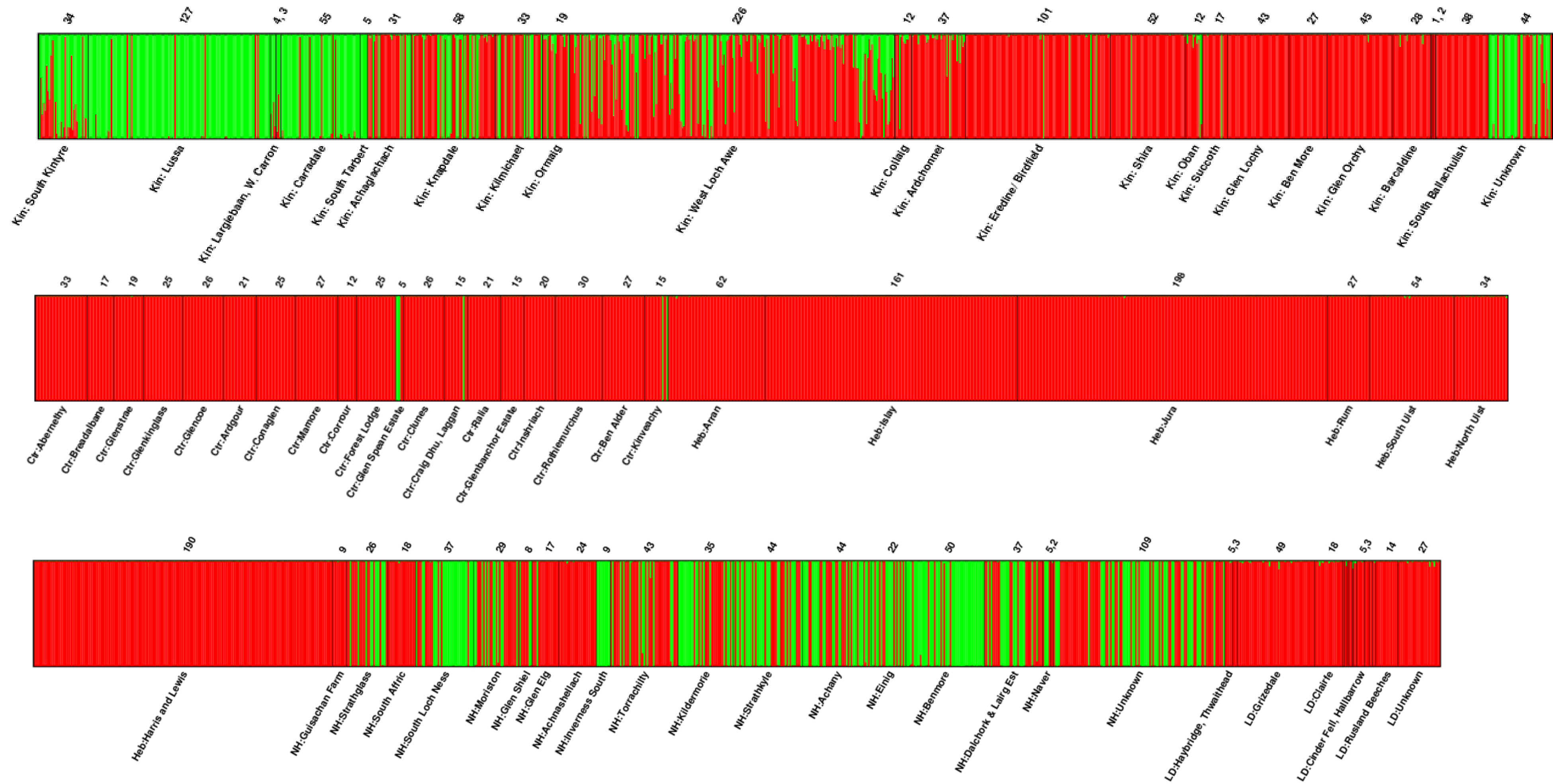


Figure 2.6. Bar chart showing the results of *analysis 2* in STRUCTURE at K = 2 for each individual in the dataset consisting of red and sika animals only (n = 2887). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (shown in red) and the proportion attributable to sika ancestry (shown in green). A hybrid is defined as an animal with membership ancestry of $0.05 \leq Q \leq 0.95$ to both red and sika. Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis. Scottish sites are plotted in an approximately south to north order, followed by the sample sites in the Lake District, Cumbria. Abbreviations represent; Kin= Kintyre, Ctr= Central highlands, Heb= Hebrides, NH= North Highlands and LD= Lake District, Cumbria.

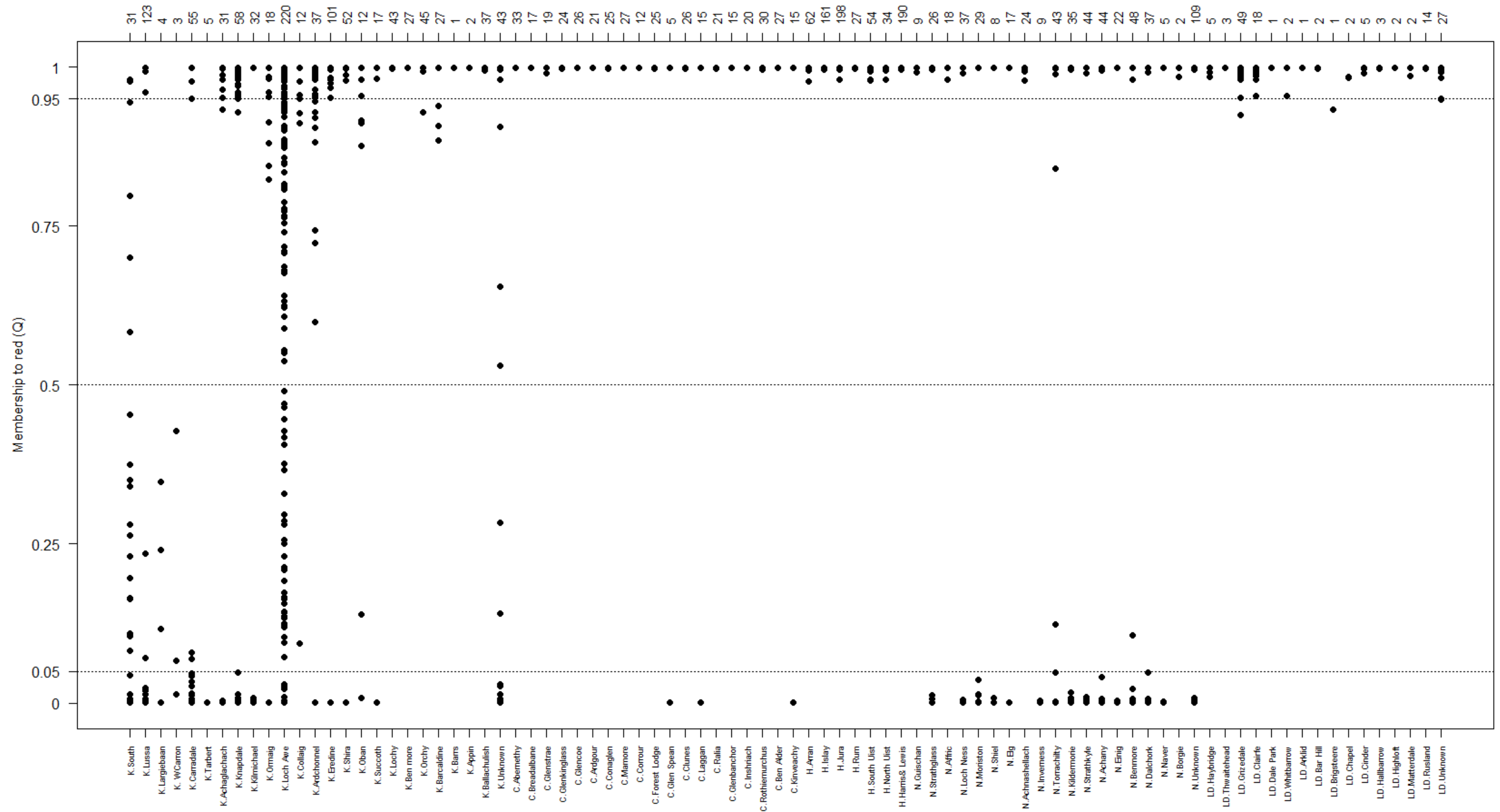


Figure 2.7. The membership to red (Q) from analysis 2 at K = 2, as calculated by Structure 2.3.3 plotted against the site from which the individual was obtained. Scottish sites are plotted in an approximately south to north order, followed by the sample sites in the Lake District, Cumbria. Abbreviations represent main areas of sampling; K= Kintyre, C= Central highlands, H= Hebrides, N= North Highlands and LD= Lake District, Cumbria.

Mitochondrial DNA analysis added further resolution to the Structure analyses (Figure 2.8). First, we note that the mitochondrial marker is not diagnostic for wapiti. Amongst the red and sika animals in analysis 2, 12.7% of sika-like animals (Q<0.5) sampled showed mtDNA discordance and 0.04% of red-like animals (Q>0.5) showed mtDNA discordance of which 92% came from Kintyre with the remaining 8% from sites in the North Highlands. Interestingly within nuclear hybrids from Kintyre, the haplotype prevalence varied between the southern site in South Kintyre and WLA; 17 out of 27 nuclear hybrids had the sika mtDNA haplotype in South Kintyre, whilst all of the 102 nuclear hybrids in WLA carried the red haplotype.

Of the animals showing mitochondrial discordance almost half of these would be considered mitochondrial hybrids (5.74% sika-like animals); that is they carried the haplotype of the opposite species against the 'pure' nuclear background of their own. This represents a level of introgression beyond the detection of our nuclear marker panel and was only apparent as 'pure' sika animals with a red mitochondrial haplotype. As with the nuclear hybrids, nearly all evidence for mtDNA introgression from animals sampled was spatially clustered; 17 of the cases came from the South Kintyre below Carradale and 11 came from in and around WLA. Six mitochondrial hybrids were also identified in the North Highlands; Kildermorie (n=4), Benmore (n=1) and an unknown site in this region (n=1). Kildermorie and Benmore sit approximately 50 miles apart across the central region of the North Highlands. Despite 11.4% of the animals coming from Kildermorie showing mitochondrial introgression, there were no nuclear hybrids identified at this site, however, one nuclear and one mitochondrial hybrid were sampled from the 50 animals obtained from Benmore. This suggests a low frequency of hybridisation events sometime in the past, the offspring of which have largely backcrossed into their parental species.

No mitochondrial hybrids were identified in 'pure' red animals and only one individual had a sika mtDNA against a red-like nuclear background (Q>0.5); this animal was sampled from the South of Kintyre and genotyped previously by Senn & Pemberton (2009). Even though we increased the sample size from South Kintyre from two to 34 compared with Senn & Pemberton (2009), we found no further evidence for this extent of introgression in this direction, or in fact any of the sites sampled in this study. No cytonuclear disequilibria in either direction were observed in the Central highlands, the Hebrides or in Cumbria.

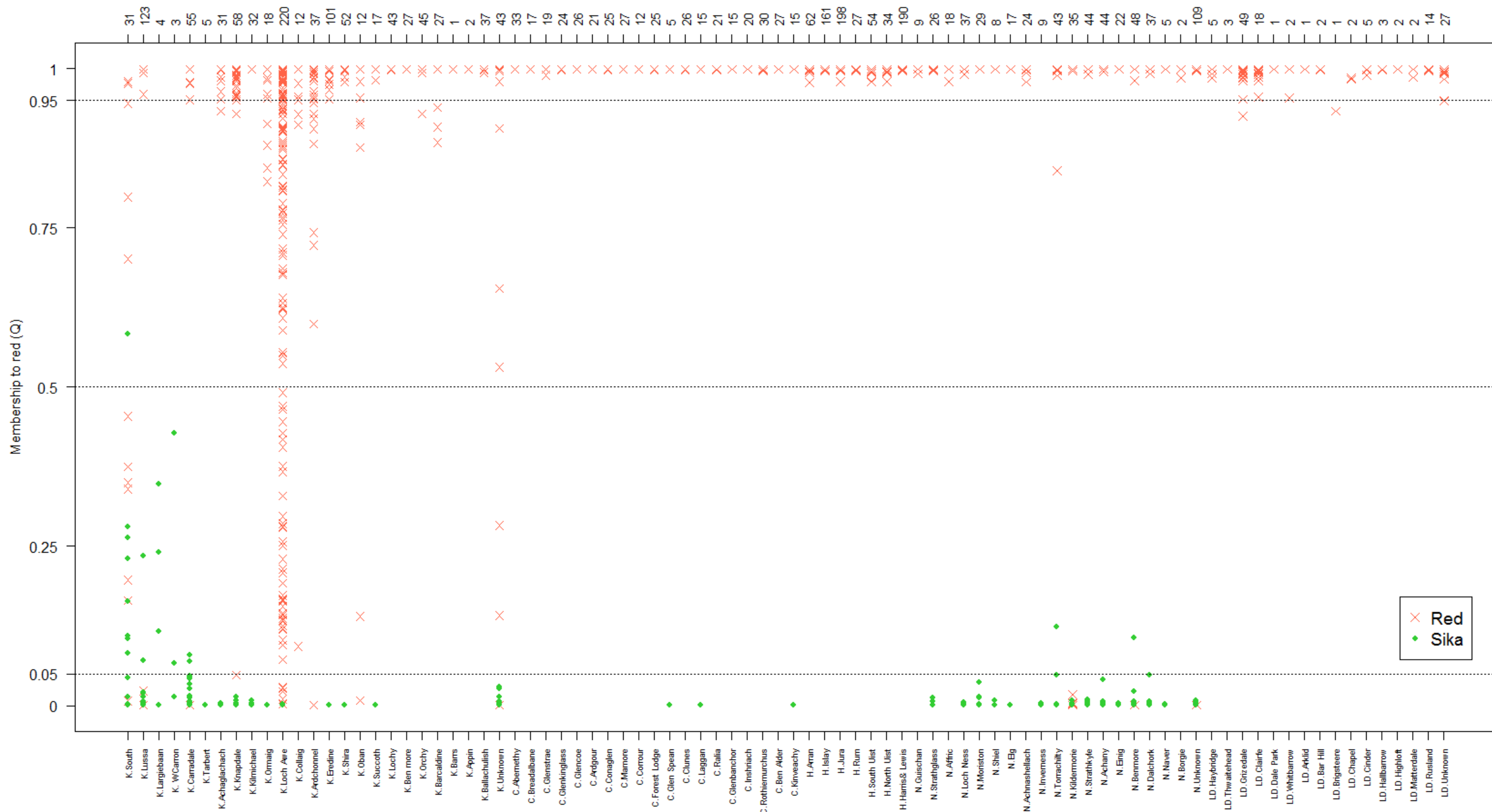


Figure 2.8. The membership to red (Q) from analysis 2, as calculated by Structure 2.3.3 at K = 2, plotted against the site from which the individual was obtained. Scottish sites are plotted in an approximately south to north order, followed by the sample sites in the Lake District, Cumbria. Abbreviations represent main areas of sampling; K= Kintyre, C= Central highlands, H= Hebrides, N= North Highlands and LD= Lake District, Cumbria. The mtDNA haplotype the individual carried is also indicated.

Analysis 3: All animals excluding sika and sika-red hybrids (n = 2230)

After removing all sika and red-sika hybrid individuals, a dataset comprising 2230 red and wapiti animals was analysed in Structure. The logarithmic likelihoods calculated in Structure revealed that, although there is relatively greater variation around the different values of K in this analysis, the most likely population structure amongst this dataset is at K = 7, with an average Ln Pr (X|K) of -109223.16 (s.d. = 16.34) and rate of change of 31.09 (see Appendix, Figure 2.A2). Whilst K=7 is the most likely population structure, the objective of this analyses was to assess introgression between red and wapiti individuals. Wapiti became differentiated from red deer in this analysis at K=4 and whilst not the most likely, for our purposes, it is most appropriate (Appendix Figure 2.A4). This analysis identifies six of the seven red-wapiti hybrids previously identified in analysis 1 (n = 2943), namely the two from Kintyre, two from the central highlands and the two from the north highlands. Based on their nuclear genome, the total proportion of these animals assigned red ancestry showed that those from the North highlands (Q = 0.946, 0.943) and central Scotland (Q = 0.936, 0.915) had relatively low-levels of wapiti introgression, with slightly more in one of those obtained from Kintyre (Q = 0.931, 0.832). The remaining animal (from the Hebrides), had only just qualified as hybrid in the context of the more inclusive analysis 1, whereas in analysis 3 its Q values placed it in the category of 'pure' red. Allele frequencies for the population clusters at K = 4 are shown in Appendix Table 2.A4.

Analysis 4: All animals excluding red and sika-red hybrids (n = 591)

Lastly, whilst a dataset consisting of all putatively pure sika and wapiti individuals only (n=591) was subject to analyses in Structure, the population structure was best explained by K = 2 and the integrity of both species was complete (see Appendix Figures 2.A3 and 2.A5).

2.4.3 *To determine, if possible, the initial direction of hybridisation (objective 2)*

Based on our best estimates of species-specific allele frequencies generated from analysis 2 (Appendix Table 2.A2) the majority of red-sika hybrid animals had low heterozygosity, however, a single individual (WYM080) returned a heterozygosity index which may be consistent with an F1 (Figure 2.9). This potential F1 hybrid animal, shot at WLA, has a Q value of 0.465, is heterozygous for a red and a sika allele at 21/ 22 markers with a single locus (RM95) homozygous for two sika alleles (frequency of red null alleles at this locus is 0.035). This animal also has a red mitochondrial haplotype suggesting that, if this was an F1, the sire was a sika and the dam was a red deer.

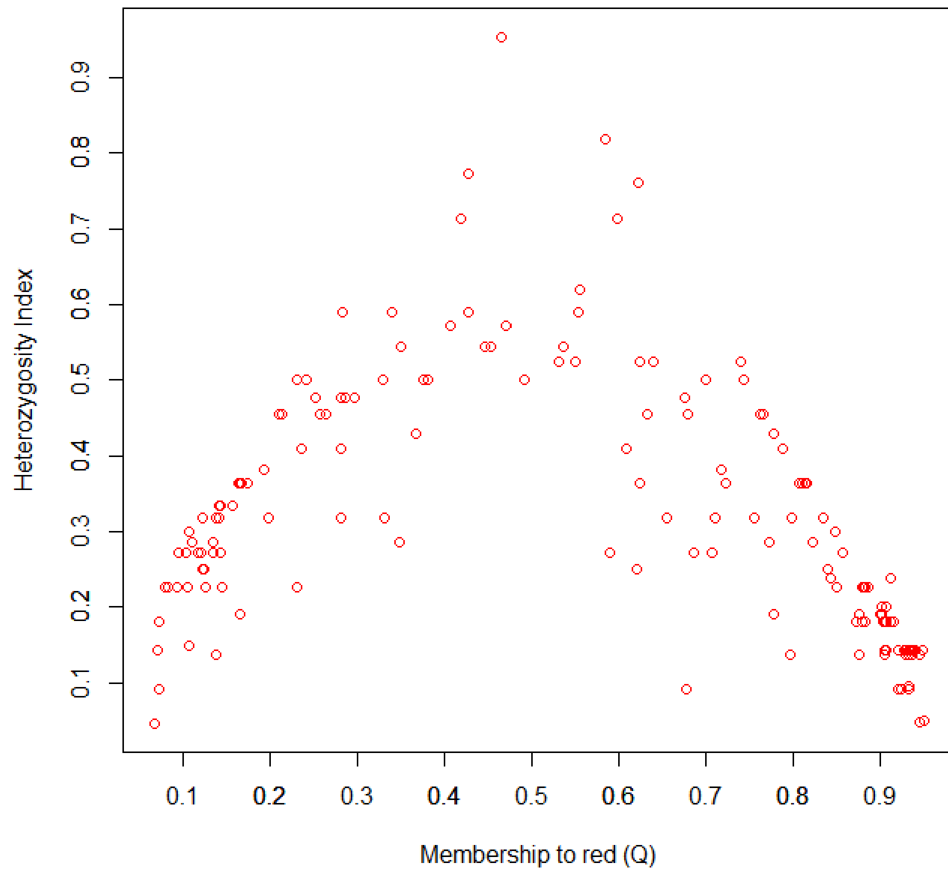


Figure 2.9. A Heterozygosity index (calculated based on the number of loci in an individual's genotype which are heterozygous for red and sika alleles, divided by the total scored loci in the genotype) by the membership to red (Q) of all hybrid individuals.

2.5 Discussion

In our analysis we genotyped 2894 individuals from Scotland and Cumbria and 49 wapiti from Canada across a panel of 22 highly diagnostic (between sika and red deer) microsatellite markers and a single mitochondrial marker in order to determine the extent of hybridisation between these species.

2.5.1. *To assess the current extent of hybridisation and introgression between introduced* Cervus *deer species, the Japanese sika deer and North American wapiti, with native red deer across Scotland and a region of the Lake District (objective 1)*

Initially it is interesting to note that the wapiti population only appeared to differentiate from the red and sika populations at K = 4 in the Structure analysis, rather than at K = 3. This is surprising given the genetic distinctness of wapiti (Kuwayama & Ozawa 2000; Ludt *et al.* 2004) and the fact they originated in North America. The pattern of differentiation would suggest that their existed greater genetic differentiation between red deer populations on the Hebridean islands compared to red deer on the Scottish mainland, than between red and sika with the wapiti species. The less diverse and more distinct genetics of the red deer on particular islands on the Hebrides (notably Harris and Lewis) is likely due to their long standing isolation and a consequence of an extreme genetic bottleneck on introduction to the island, followed by subsequent inbreeding. Population structure within the red deer population only in Britain and Ireland is discussed further in Chapter 4.

Evidence for low-level wapiti introgression was found in seven red-like animals when analysed in the context of all samples (analysis 1, n = 2943) and once the data was reanalysed without sika and red-sika hybrid animals (analysis 3, n = 2230), six of these animals retained Q values that fell within our definition of a hybrid. These individuals were from Kintyre (n = 2), the central region (n = 2) and the north highlands (n = 2) and either represent genuine wapiti introgression from previous hybrid events or are an artefact of ancestral polymorphism, the limited resolution of this marker panel or imprecision in our definition of a 'hybrid' animal. Combined with their largely unsuccessful establishment in Britain, the impact of wapiti as an introduced species across Scotland and the Lake District appears negligible (Whitehead 1964).

Further to that discussed above, it is interesting to note that even after the removal of sika animals, Structure analysis showed that it still took till K = 4 before the wapiti became differentiated as a separate cluster. The genetic sub-structuring within the red

deer from the Hebrides compared to the mainland is clearly apparent and could warrant subspecies to be designated if genuine. However such patterns may also be the consequences of genetic bottlenecks occurring on the island or an artefact of the marker panel.

The trivial impact of wapiti on Scottish red deer is consistent with a previous study (Pérez-Espona *et al.* 2013) with which this study shares 235 animals. The study by Pérez-Espona adopted a more conservative definition of a hybrid (admixture by Q≥0.1) than used in this study (admixture by Q≥0.05), such that the two wapiti-introgressed animals we identified here were recognised as putatively pure red animals, rather than previously undetected hybrids. Out of interest these animals from the Central highlands were from Ardgour and Mamore; the latter of which historically held a large wapiti herd (Whitehead 1964). The Q values representing membership to wapiti ancestry were 0.075 (0.001-0.197) and 0.085 (0.000-0.248) respectively. Despite this slight discrepancy, both studies did agree that the impact of wapiti introgression on red deer in Scotland and the British Isles was negligible.

In contrast, it is evident from analysis of all samples (analysis 1, n = 2943) and red and sika animals only (analysis 2, n=2887) that there is substantial gene flow and introgression between red deer and sika in some parts of Scotland. Specifically, within two regions in Kintyre, around WLA and the southern site in South Kintyre, 50.4% and 61.8% of individuals were hybrid respectively, based on either their nuclear genotype or their mitochondrial haplotype. While the hybrid swarm at WLA has been documented previously (Senn & Pemberton 2009), the south Kintyre swarm has only become apparent with the additional samples collected for this study. Further to this, sites either side of WLA contained substantial numbers of nuclear or mitochondrial hybrids (Ormaig, 21.1% of samples from this site, Collaig, 25% and Ardchonnel, 24.3%) as did Lussa (11.8%) and West Carrom, (66.7%), just north of the southern tip of Kintyre.

A major difference between these two regions of hybrid activity in Kintyre was the absence of the sika haplotype amongst hybrids in WLA and its prevalence in the southern sites. A possible explanation for this contrast may be that the relative species densities in the two areas were different when hybridisation started. At WLA, sika were invading an area containing red deer; backcrossing into red was, therefore, more likely. In contrast in the south of Kintyre red deer were rare relative to sika which were locally abundant (Senn & Pemberton 2009), such that hybrids generated here were more likely to backcross with sika. In south Kintyre it is possible hybridisation occurred after an escape of farmed red deer (K. McKillop pers. comm.)

Regarding other parts of Scotland, only 3/568 (0.53%) nuclear red-sika hybrids were identified in the North highlands, whilst twice this number had a mitochondrial haplotype discordant with their nuclear genotype, increasing the number of hybrid animals from this area to 9/568 or 1.6%. This is consistent with evidence for some hybridisation in this region by Swanson (Swanson 2000). In addition, the unexpected appearance of animals showing mitochondrial discordance from sites otherwise relatively free of nuclear introgression (e.g. Kildermorie) is extremely interesting and highlights the potential for many more hybridisation events in the past, outside Kintyre, which failed to break down assortative mating between species and progress to a hybrid swarm.

On the other hand, there was no evidence for hybrid animals from across the central highlands, even given the addition of a further 50 animals collected from the western borders of the Cairngorm National Park into which sika are encroaching. Similarly, no evidence for red-sika hybrids were found in the Hebrides, suggesting the integrity of the refugia is still effective.

Within the Lake District, only 3/137 (2.2%) of animals were considered hybrid; these were red-like animals with enough apparent sika introgression to qualify them as hybrid. There was no evidence for mitochondrial introgression for animals from this region. Whilst Whitehead (1964) and Lowe & Gardiner (1975) document a history of hybrid sightings in this region, these may have been eliminated through increased culling pressure on conspicuous hybrids by the Lake District Deer Control Society around the 1970s.

With regard to the two introduced species, there was no evidence for hybridisation and introgression between wapiti and sika across Scotland and the Lake District in the larger analysis (analysis 1, n = 2943) or the more exclusive analysis (analysis 4, n = 591). Their genetic integrity is consistent with the fact their probability of meeting would have been low and they have large morphological differences, likely to make them incompatible.

Overall, this leads us to conclude that whilst introgression from wapiti is not considered a threat to native red deer, hybridisation with sika evidently is, however it is stochastic as to where it occurs. Further, the presence of low-level nuclear and mitochondrial introgression amongst our samples from Kintyre, the North highlands and the Lake District highlights that the downstream outcome of an initial hybridisation event can

vary. Whilst in some populations hybridisation is followed by a breakdown of assortative mating and the generation of a hybrid swarm (e.g. South Kintyre and WLA), in others it is followed by repeated backcrossing to sika (e.g. the North Highlands) or into red (e.g. the Lake District) and assortative mating amongst parental taxa largely continues.

2.5.2 *To determine, if possible, the initial direction of hybridisation (objective 2)*

In terms of the initial direction of hybridisation, the animal most likely to be an F1 hybrid in our entire dataset had a mother who carried the red mitochondrial haplotype. This is consistent with that expected for the red hind-sika stag direction.

The substantial number of mitochondrial hybrids in the form of 'pure' sika based on their nuclear genotype but carrying the red mtDNA haplotype highlights that mitochondrial introgression to this extent is largely unidirectional and that matings are primarily occurring (at least those immediately preceding sampling) between red-like hinds and sika-like stags. This is consistent with the literature (Pérez-Espona *et al.* 2009b; Senn & Pemberton 2009). The only example for a red-like hybrid (Q=0.584) carrying a sika mitochondrial haplotype was obtained from the South of Kintyre; however, not only was there an unusually high prevalence of the sika haplotype in this region but the genotype of the animal was not entirely consistent with that of an F1 hybrid. This is such that it was unlikely to be the offspring of a 'pure' red stag and a 'pure' sika hind, but may in fact have been generated by mating between a red-like stag and a sika hind or a red stag and a sika-like hind. The hybrid nature of either one of the parental species in this case may have facilitated the compatibility and success of this mating. Such cytonuclear disequilibria could be driven by assortative mating behaviour in terms of size compatibility; here a red hind and sika stag are more comparable in size than a red stag and sika hind (Senn & Pemberton 2009). This suggests a pattern of sex-biased dispersal, whereby sika stags disperse into areas with more sedentary red hinds and initiate matings. Such a behavioural explanation for unidirectional mitochondrial introgression has also been suggested in tree frogs from Alabama in response to anthropogenic-induced habitat disruption (Lamb & Avise 1986).

2.5.3 *To compare the outcome of this study with other forms of anthropogenically-induced hybridisation and try to identify causative factors (objective 3)*

In the context of other studies of anthropogenically-induced hybridisation, we see there is not a uniform route for gene flow but numerous trajectories and scenarios that allele frequencies within populations can converge on (Allendorf *et al.* 2001). There can be hybridisation without introgression, the main disadvtange of which is wasted reproductive costs, as between trout species (*Salvelinus*) in America (Leary *et al.* 1993) and stilt species (*Himantopus*) in New Zealand. There can be situations in which hybridisation may occur and introgression become widespread; however, putatively pure populations of the parental species remain. This may describe the situation between gull species (*Larus* spp.) in Iceland (Vigfusdottir *et al.* 2008), European and domestic wildcats (*Felis* spp.) and between wolves and domestic dogs (*Canis* spp.) (Randi *et al.* 2001), within wolf species themselves (Fain *et al.* 2010) and between native caribou and semi-domesticated reindeer (*Rangifer* spp.) (Jepsen *et al.* 2002). Lastly, anthropogenic hybridisation may result in complete admixture, whereby assortative mating is entirely broken down, the system collapses into a hybrid swarm and remaining populations of 'pure' animals are few if not lost entirely. This has been a concern with up to 14 species of cutthroat trout species (*Oncorhynchus*) in North America (Trotter & Behnke 2008), between grey ducks and mallards (*Anas* spp.) in New Zealand (Rhymer & Simberloff 1996) and between the American black duck and mallards in North America (Mank *et al.* 2004). The introduction of wapiti appears to have had negligible impact on the genetic integrity of native red deer in Scotland as we detected very few hybrids and they had very low levels of introgression. The impact of sika deer, however, appears to sit somewhere on the spectrum between the second and the last scenario described above. Hybridisation events appear infrequent, so 'pure' populations of red deer still remain in some areas (e.g. Hebridean refugia, areas of the Cairngorm National Park). When hybridisation does occur there can be at least two outcomes: repeated backcrossing leading to hybrids which can only sometimes be detected by nuclear markers or by discordant mitochondrial haplotypes (e.g. some parts of Kintyre, North Highlands, Cumbria). Alternatively there can be a breakdown of assortative mating leading to a hybrid swarm (e.g. South Kintyre, WLA). Overall, the pattern and extent of red-sika hybridisation and introgression throughout Scotland may be described as "mottled" (Hauffe & Searle 1993); where some areas are entirely free from hybridisation and genetic integrity is strong, whilst others have collapsed. In Cumbria, it seems that past hybrid activity may have occurred; however, the offspring of such activities have

backcrossed into red to the extent that the population appears largely composed of 'pure' individuals.

The question of hybridisation causation appears hierarchal. Firstly, what causes the initial hybridisation event between these two species? Secondly, what is the fate of this hybrid individual; does it mate with other intermediate animals or backcross into the parental populations? Lastly, if the system does proceed to a hybrid swarm, what determines the direction of nuclear and mitochondrial introgression?

Identifying the factors that cause hybridisation events to occur is difficult due to their infrequency, such that we can only postulate. The initial densities of both species would be important in determining species and sex ratios and hence mating dynamics. If sika stags, which migrate ahead of sika hinds, enter an area with high density of red hinds this may heighten the chance of hybridisation. The expansion of sika populations has been proposed to follow the planting of commercial forestry, which facilitate their spread and access to red populations (Pérez-Espona *et al.* 2009a; Swanson & Putman 2009). Sika have also not reached all parts of Scotland, including most of the Hebrides and much of the East of Scotland, so the absence of hybridisation in some areas could simply be through lack of opportunity. The duration of a time both species have been interacting may also govern the likelihood of hybridisation. Despite red and sika being present in Kintyre for approximately 115 years (~38 generations), reports suggest it wasn't until the 1960's that permanent contact was established between the two species, hence the system is very young (Senn & Pemberton 2009). Stochastic forces, therefore, may still govern the likelihood of species interacting and hybridising.

Following a hybridisation event the ambient density of both parental species may determine the trajectory of subsequent matings. The general absence of extensive hybridisation across a large region of the North highlands may be attributable to relatively equal densities of red and sika deer, such that assortative mating remains strong and access to a species own parental population is facilitated. On the other hand, the hybrid swarms in the south of Kintyre and around WLA may have been due to a small number of one species in the presence of a high density of the other; in South Kintyre, individual red deer that escaped from a deer farm in the area would have found themselves in sika-dominated territory, whilst at WLA pioneering sika stags that migrated from the south would have met the higher red deer densities at this central region in Kintyre. These individuals would have been more likely to mate with the more abundant parental species, initiating hybridisation, the offspring of which either

continued to backcross into the parental population or mated amongst themselves (potentially leading to a hybrid swarm).

If a hybrid swarm develops the direction with which nuclear and mitochondrial genetic material introgresses may be determined by contemporary selection pressures, the presence of cytonuclear disequilibrium and, again, the relative density of both species. The greater frequency of nuclear hybrid animals in the three most southern sites in Kintyre (60%) carrying the sika mitochondrial haplotype compared to its absence amongst 100 nuclear hybrids from WLA suggests the source and densities of both species, their previous interaction with the other species and site-specific adaptations may all influence the consequential trajectory of gene flow. In addition, of over 2300 animals tested, no genetically pure red deer carried a sika mitochondrial haplotype, whilst 6% of pure sika carried the red haplotype. This unidirectional occurrence of mitochondrial hybrids could indicate selection for the red haplotype, cytonuclear disequilibria between the red haplotype and nuclear alleles likely to introgress or an incompatibility between a red-like nuclear genome and the sika mitochondrial haplotype (Arnold 1993).

2.5.4 *To indicate what future management actions may be required to protect putatively pure populations from hybridisation (objective 4)*

The red-sika system, therefore, needs to be addressed if the integrity of both species is to be preserved. Areas around reported hybrid swarms should be monitored in Scotland, in case hybrid animals disperse, whilst in other areas where both species are sympatric management should strive to discourage the initiation or accumulation of hybrid animals.

In regions within which red and sika populatons were sympatric in this study, stalkers accurately identified the species status of animals ('red', 'sika' and 'hybrid') in approximately 88% of cases, however misidentifications (12%) highlight the difficulty in identifying introgressed individuals based on phenotype in the field. This suggests that attempting to selectively cull hybrids will not be totally effective and introgressed animals are likely to escape undetected (Senn & Pemberton, 2009, Chapter 5). Managers should remain vigilant as rare sika amongst large populations of red are potentially conducive to hybridisation and they should target culling toward conspicuous hybrid animals and pioneering sika stags. In situations where red deer are massively outnumbered, efforts could be focused on increasing their numbers or rather addressing

the imbalance with sika numbers and, thereby, lowering their susceptibility to hybridisation (Vilà *et al.* 2003). Not only management of deer populations, but management of the land could help ameliorate hybridisation. Just as landscape features have been shown to have significant impact on red deer gene flow in Scotland (Pérez-Espona *et al.* 2008), similarly, hybridisation patterns may be influenced by patterns of increasing forestry cover (Carden *et al.* 2010). Working together with foresters could allow deer managers to play a role in shaping the layout of future forests in a way that reduces access and suitable corridors for dispersal of the invasive sika.

2.6 Acknowledgements

We thank the Forestry Commission Scotland rangers, including Kevin McKillop, Willie Lamont, Derick Macaskill, David Bain, staff of the Deer Commission Scotland (now Scottish Natural Heritage), Will Boyd-Wallis and the Cairngorm National Park Authority and the private estate owners and stalkers who cooperated within this project and provided samples. Thanks to Eleni Socratous, Eleanor A. M. Graham and Guy N. Rutty from the University of Leicester and Mike Thornley (Forestry Commission Chief Ranger) and the rangers of the Forestry Commission Cumbria for the samples they provided from the Lake District. Thanks to Prof. David Coltman (University of Alberta) for the provision of wapiti samples. Thanks also to Elizabeth Heap, Sheena Morrissey Phil Ellis, Elizabeth Mittell and Sarah MacDonald for genotyping and haplotyping some samples. Lastly, thanks to the GenePool for their excellent service and NERC for funding.

2.7 Appendices

Table 2.A1. Details of the molecular markers used in this study from Senn & Pemberton (2009) (next page).

Loading panel	PCR plex	Locus name	Annealing temperature (° C)	Primer concentration (µM)	Label	Size Range	Deer linkage group	Primers (5'-3')
A	1	AGLA293*	58	0.06	PET	128-147	3	GTCTGAAATTGGAGGCAATGAGGC CCCAAGACAACTCAAGTCAAAGGACC
	1	RM12§		0.06	VIC	116-151	9	CTGAGCTCAGGGGTTTTTGCT ACTGGGAACCAAGGACTGTCA
	1	INRA6¶		0.1	NED	128-138	20	AGGAATATCTGTATCAACCTCAGTC CTGAGCTGGGGTGGGAGCTATAAATA
	1	TGLA126*		0.12	6-FAM	100-105	-	CTAATTTAGAATGAGAGAGGCTTCT TTGGTCTCTATTCTCTGAATATTCC
	2	IDVGA55¥	59	0.12	NED	191-221	4	GTGACTGTATTTGTGAACACCTA TCTAAAACGGAGGCAGAGATG
	2	BM6438†		0.5	6-FAM	249-275	31	TTGAGCACAGACACAGACTGG ACTGAATGCCTCCTTTGTGC
B	3	FSHB‡	56	0.5	6-FAM	180-210	1	CAGTTTCTAAGGCTACATGGT TGGGATATAGACTTAGTGGC
	3	BOVIRBP**		0.25	NED	140-159	-	TGTATGATCACCTTCTATGCTTC GCTTTAGGTAATCATCAGATAGC
	3	INRA131§§		0.12	PET	92-105	11	GGTAAAATCCTGCAAAACACAG TGACTGTATAGACTGAAGCAAC
	3	BM4006†		0.06	VIC	85-95	-	CAATGTGCATTATTTCCAAAGTG AGAAATAACTCTTTCTCCTTGGAGG
	Solo	RM188§	61	0.35	VIC	115-182	18	GGGTTCACAAAGAGCTGGAC GCACTATTGGGCTGGTGATT
C	4	MM12¶¶	60	0.12	NED	89-104	26	CAAGACAGGTGTTTCAATCT ATCGACTCTGGGGATGATGT
	4	BM757†		0.25	6-FAM	160-202	28	TGGAAACAATGTAAACCTGGG TTGAGCCACCAAGGAACC
	4	OarFCB193¥¥		0.5	PET	103-143	5	TTCATCTCAGACTGGGATTCAGAAAGGC GCTTGGAAATAACCCTCCTGCATCCC
	Solo	TGLA40*	56	0.25	6-FAM	91-108	10	GCTTCTCTGCCAACTAATATTATCC CACCAGGTAAGCCCCTTATATATGT
	Solo	RM095†	54	0.12	VIC	118-147	31	TCCATGGGGTCGCAAACAGTGG ATCCCTCCATTTGTTGTGGAGTT
D	5	TGLA127*	53	0.08	NED	161-192	20	CAATTGTGTGGTAGTTTGGACATTC ACACTATTGCAAAAGGACCTCCAATT
	5	UWCA47††		0.5	6-FAM	225-240	29	GGAAAGTCCTTAGATGGAGGATTGT TTGAGAACTTGTCCCGAGAGAA
	5	INRA5¶		0.25	VIC	129-143	30	CAATCTGCATGAAGTATAAATAT CTTCAGGCATACCCTACACC
E	6	IDVGA29¥	54	0.25	VIC	136-156	-	CCCACAAGGTTATCTATCTCCAG CCAAGAAGGTCCAAAGCATCCAC
	6	TGLA337*		0.5	PET	126-147	13	TTTGTTAAGGATAGTAGGCTACT GCTCTTCCCTTGGTTTCCTTG
	Solo	RME25‡‡	54	0.5	6-FAM	151-207	12	AGTGGGTAAAGGAGCCTGGT TTATTGATCCCAGCCTGTGC
Mt DNA marker	-	SikaL3*** H16498§§§	52	0.5	-	430(sika) 350(red)	-	TTAAACTATTCCCTGACGCTT CCTGAAGTAGGAACCAGATG

Table 2.A2. Posterior allele frequencies from analysis 2 (n = 2887) at K = 2 and species specific allele assignment. An allele was not assigned to a species if its frequency was less than 1% (0.01) for both species. Alleles were assigned to a species (red = red, green = sika) if its frequency in the other species was 0 or if its frequency was five-fold larger than the other species.

Locus	% Missing data	Allele Size	Estimated allele frequency in Red	Estimated allele frequency in Sika	Allele species - specific assignment
ALGA293	1.50%	128	0.083	0.004	R
		144	0.773	0.021	R
		147	0.054	0.969	S
		Null	0.091	0.005	R
BM4006	0.40%	85	0.001	0.974	S
		87	0.096	0.000	R
		93	0.731	0.024	R
		95	0.124	0.000	R
		Null	0.047	0.002	R
BM6438	1.20%	249	0.545	0.001	R
		251	0.201	0.002	R
		253	0.098	0.000	R
		257	0.005	0.000	NA
		259	0.000	0.080	S
		261	0.072	0.000	R
		265	0.000	0.273	S
		275	0.001	0.551	S
		Null	0.078	0.092	NA
BM757	0.10%	160	0.068	0.010	R
		162	0.543	0.003	R
		164	0.007	0.000	NA
		172	0.000	0.916	S
		174	0.003	0.053	S
		179	0.053	0.000	R
		183	0.075	0.000	R
		185	0.044	0.001	R
		187	0.037	0.000	R
		189	0.002	0.000	NA
		196	0.000	0.000	NA
		197	0.000	0.000	NA
		198	0.056	0.003	R
		200	0.069	0.000	R
		202	0.010	0.000	R
		210	0.004	0.000	NA
		Null	0.029	0.013	NA
BOVIRP	0.40%	140	0.001	0.959	S
		142	0.000	0.016	S
		147	0.061	0.000	R
		149	0.056	0.000	R
		151	0.184	0.000	R
		153	0.365	0.008	R
		155	0.055	0.002	R
		157	0.202	0.000	R
		159	0.021	0.000	R
		163	0.000	0.000	NA
		Null	0.055	0.015	NA
FCB193	2.00%	101	0.002	0.000	NA
		103	0.039	0.008	NA
		105	0.000	0.000	NA
		107	0.090	0.000	R
		109	0.161	0.000	R
		111	0.018	0.000	R
		113	0.236	0.000	R
		115	0.008	0.000	NA
		118	0.039	0.000	R
		120	0.101	0.000	R
		122	0.101	0.001	R
		124	0.049	0.000	R
		126	0.010	0.028	NA
		128	0.011	0.051	NA
		130	0.060	0.000	R
		132	0.006	0.901	S
		134	0.004	0.002	NA
		140	0.003	0.000	NA
		141	0.000	0.000	NA
		143	0.007	0.000	NA
		Null	0.055	0.008	R
FSHB	0.70%	179	0.000	0.009	NA
		180	0.004	0.744	S
		181	0.000	0.111	S
		182	0.000	0.029	S
		184	0.048	0.000	R
		185	0.189	0.000	R
		186	0.001	0.000	NA
		187	0.003	0.000	NA
		188	0.126	0.000	R
		189	0.127	0.002	R
		190	0.002	0.028	S
		191	0.088	0.000	R
		192	0.018	0.000	R

		193	0.000	0.000	NA
		194	0.026	0.002	R
		195	0.000	0.000	NA
		196	0.009	0.000	NA
		197	0.005	0.000	NA
		198	0.072	0.000	R
		199	0.021	0.000	R
		200	0.000	0.000	NA
		201	0.006	0.000	NA
		202	0.026	0.000	R
		203	0.021	0.007	NA
		204	0.011	0.000	R
		205	0.079	0.003	R
		206	0.034	0.000	R
		207	0.038	0.000	R
		210	0.010	0.000	R
		211	0.002	0.000	NA
		Null	0.032	0.064	NA
IDVGA29	1.50%	136	0.668	0.018	R
		143	0.318	0.044	R
		156	0.002	0.911	S
		Null	0.011	0.028	NA
IDVGA55	1.90%	191	0.037	0.000	R
		193	0.087	0.000	R
		195	0.226	0.001	R
		197	0.284	0.000	R
		199	0.209	0.005	R
		202	0.023	0.000	R
		204	0.042	0.000	R
		208	0.000	0.001	NA
		210	0.001	0.868	S
		212	0.000	0.064	S
		214	0.000	0.051	S
		217	0.037	0.000	R
		219	0.015	0.000	R
		221	0.000	0.000	NA
		Null	0.038	0.010	NA
INRA005	0.10%	124	0.000	0.036	S
		126	0.993	0.045	R
		129	0.000	0.002	NA
		136	0.000	0.002	NA
		143	0.000	0.913	S
		Null	0.006	0.001	NA
INRA006	0.10%	128	0.000	0.001	NA
		130	0.000	0.947	S
		132	0.040	0.000	R
		134	0.683	0.042	R
		136	0.244	0.004	R
		138	0.010	0.000	R
		Null	0.022	0.007	NA
INRA131	0.00%	87	0.000	0.000	NA
		92	0.042	0.000	R
		94	0.008	0.086	S
		98	0.590	0.002	R
		100	0.233	0.000	R
		102	0.071	0.000	R
		104	0.037	0.000	R
		106	0.000	0.779	S
		113	0.000	0.048	S
		115	0.000	0.011	S
		Null	0.018	0.073	NA
MM012	0.10%	89	0.742	0.104	R
		91	0.230	0.004	R
		93	0.000	0.835	S
		97	0.001	0.000	NA
		104	0.000	0.000	NA
		Null	0.026	0.056	NA
RM012	0.60%	116	0.003	0.991	S
		120	0.005	0.000	NA
		125	0.165	0.001	R
		127	0.051	0.000	R
		129	0.072	0.000	R
		131	0.083	0.000	R
		133	0.244	0.000	R
		137	0.016	0.000	R
		139	0.085	0.005	R
		141	0.085	0.000	R
		144	0.095	0.000	R
		151	0.039	0.000	R
		Null	0.057	0.002	R
RM188	0.80%	115	0.020	0.000	R
		117	0.037	0.000	R
		121	0.000	0.000	NA
		123	0.042	0.000	R
		125	0.074	0.000	R
		127	0.411	0.003	R
		129	0.207	0.009	R
		131	0.033	0.000	R
		132	0.030	0.000	R
		133	0.001	0.000	NA
		134	0.041	0.000	R
		137	0.041	0.000	R
		139	0.003	0.037	S
		141	0.000	0.009	NA
		143	0.000	0.550	S

		161	0.000	0.205	S
		163	0.000	0.002	NA
		176	0.000	0.027	S
		182	0.000	0.143	S
		Null	0.057	0.015	NA
RM95	0.50%	116	0.000	0.117	S
		118	0.054	0.000	R
		120	0.002	0.000	NA
		122	0.011	0.796	S
		124	0.086	0.000	R
		126	0.041	0.000	R
		128	0.179	0.001	R
		130	0.302	0.011	R
		132	0.099	0.000	R
		134	0.006	0.000	NA
		136	0.077	0.000	R
		138	0.088	0.000	R
		140	0.019	0.000	R
		142	0.002	0.000	NA
		147	0.000	0.001	NA
		Null	0.035	0.074	NA
RME025	0.50%	151	0.020	0.000	R
		155	0.064	0.000	R
		157	0.001	0.000	NA
		159	0.003	0.000	NA
		168	0.764	0.005	R
		170	0.106	0.007	R
		183	0.001	0.000	NA
		193	0.000	0.976	S
		207	0.010	0.000	R
		Null	0.031	0.011	NA
TGLA40	0.30%	91	0.195	0.000	R
		96	0.003	0.000	NA
		97	0.492	0.010	R
		98	0.000	0.000	NA
		99	0.041	0.000	R
		101	0.192	0.000	R
		102	0.002	0.000	NA
		104	0.001	0.758	S
		106	0.000	0.155	S
		108	0.001	0.001	NA
		Null	0.073	0.076	NA
TGLA126	0.00%	100	0.001	0.350	S
		101	0.000	0.548	S
		105	0.935	0.057	R
		130	0.000	0.003	NA
		132	0.002	0.000	NA
		134	0.006	0.000	NA
		136	0.003	0.000	NA
		138	0.000	0.000	NA
		Null	0.052	0.042	NA
TGLA127		161	0.000	0.603	S
		167	0.014	0.000	R
		169	0.321	0.006	R
		171	0.000	0.000	NA
		172	0.003	0.003	NA
		174	0.027	0.290	S
		176	0.018	0.000	R
		178	0.241	0.012	R
		180	0.053	0.000	R
		184	0.100	0.000	R
		186	0.072	0.000	R
		188	0.002	0.000	NA
		190	0.065	0.000	R
		192	0.039	0.000	R
		Null	0.047	0.085	NA
TGLA337	8.50%	126	0.005	0.592	S
		128	0.000	0.039	S
		130	0.201	0.000	R
		132	0.112	0.000	R
		134	0.002	0.000	R
		136	0.255	0.006	R
		138	0.042	0.193	NA
		142	0.001	0.000	NA
		145	0.234	0.001	R
		147	0.065	0.022	NA
		153	0.001	0.000	NA
		155	0.003	0.016	S
		Null	0.079	0.131	NA
UWCA47	0.50%	225	0.031	0.000	R
		229	0.049	0.000	R
		231	0.867	0.082	R
		240	0.000	0.878	S
		Null	0.053	0.040	NA

Table 2.A3. Posterior allele frequencies from analysis 1 (n = 2943) at K = 4.

Locus	% Missing data	Allele Size	Estimated allele frequency in Wapiti	Estimated allele frequency in Sika	Estimated allele frequency in Red I	Estimated allele frequency in Red II
AGLA293	1.50%	128	0.002	0.004	0.105	0.000
		144	0.921	0.018	0.738	0.993
		147	0.033	0.974	0.064	0.001
		149	0.011	0.000	0.000	0.000
		Null	0.032	0.003	0.093	0.006
BM4006	0.40%	85	0.001	0.977	0.001	0.000
		87	0.057	0.000	0.070	0.358
		93	0.827	0.021	0.839	0.138
		95	0.077	0.000	0.086	0.502
		Null	0.038	0.001	0.004	0.002
BM6438	1.20%	249	0.403	0.001	0.583	0.426
		251	0.176	0.001	0.233	0.000
		253	0.028	0.000	0.120	0.001
		257	0.033	0.000	0.001	0.000
		259	0.000	0.080	0.000	0.000
		261	0.007	0.000	0.024	0.561
		263	0.102	0.000	0.000	0.000
		265	0.021	0.273	0.000	0.000
		275	0.008	0.551	0.000	0.000
		Null	0.222	0.094	0.038	0.011
BM757	0.10%	160	0.107	0.010	0.066	0.033
		162	0.354	0.001	0.510	0.958
		164	0.000	0.000	0.009	0.000
		172	0.000	0.919	0.000	0.000
		173	0.039	0.000	0.000	0.000
		174	0.001	0.053	0.003	0.000
		175	0.009	0.000	0.000	0.000
		177	0.034	0.000	0.000	0.000
		179	0.083	0.000	0.055	0.000
		183	0.021	0.000	0.092	0.000
		185	0.001	0.000	0.056	0.000
		187	0.199	0.000	0.017	0.003
		189	0.000	0.000	0.003	0.000
		192	0.011	0.000	0.000	0.000
		196	0.000	0.000	0.001	0.000
		197	0.000	0.000	0.000	0.000
		198	0.039	0.003	0.071	0.000
		200	0.039	0.000	0.087	0.000
		202	0.000	0.000	0.013	0.000
		210	0.000	0.000	0.005	0.000
		Null	0.061	0.013	0.012	0.005
BOVIRP	0.40%	140	0.008	0.961	0.000	0.000
		142	0.000	0.016	0.000	0.000
		145	0.052	0.000	0.000	0.000
		147	0.055	0.000	0.077	0.000
		149	0.265	0.000	0.030	0.016
		151	0.070	0.000	0.162	0.481
		153	0.263	0.007	0.421	0.027
		155	0.002	0.002	0.060	0.071
		157	0.098	0.000	0.192	0.401
		159	0.075	0.000	0.015	0.000
		161	0.028	0.000	0.000	0.000
		163	0.000	0.000	0.000	0.000
		Null	0.084	0.013	0.043	0.004
FCB193	1.90%	101	0.005	0.000	0.001	0.000
		103	0.002	0.008	0.049	0.000
		105	0.000	0.000	0.001	0.001
		107	0.023	0.000	0.111	0.000
		109	0.014	0.000	0.203	0.000
		111	0.139	0.000	0.002	0.004
		113	0.151	0.000	0.270	0.038
		115	0.000	0.000	0.010	0.000
		118	0.011	0.000	0.047	0.003
		120	0.250	0.000	0.054	0.303
		122	0.058	0.001	0.102	0.127
		124	0.001	0.000	0.061	0.000
		126	0.104	0.027	0.003	0.077
		128	0.041	0.051	0.003	0.045
		130	0.001	0.000	0.032	0.364
		132	0.012	0.903	0.001	0.034
		134	0.007	0.002	0.005	0.000
		140	0.029	0.000	0.000	0.000
		141	0.000	0.000	0.000	0.000
		143	0.008	0.000	0.009	0.000
		145	0.006	0.000	0.000	0.000
		150	0.023	0.000	0.000	0.000
		Null	0.115	0.009	0.035	0.003
FSHB	0.70%	179	0.000	0.009	0.000	0.000
		180	0.025	0.746	0.001	0.000
		181	0.000	0.111	0.000	0.000
		182	0.102	0.029	0.000	0.000
		184	0.023	0.000	0.061	0.000
		185	0.074	0.000	0.232	0.000
		186	0.000	0.000	0.001	0.000
		187	0.002	0.000	0.004	0.000
		188	0.135	0.000	0.104	0.294
		189	0.175	0.002	0.126	0.054
		190	0.015	0.028	0.001	0.002
		191	0.055	0.000	0.077	0.195
		192	0.000	0.000	0.022	0.000

		193	0.000	0.000	0.000	0.000
		194	0.000	0.002	0.033	0.000
		195	0.000	0.000	0.000	0.000
		196	0.001	0.000	0.011	0.000
		197	0.002	0.000	0.006	0.006
		198	0.041	0.000	0.071	0.106
		199	0.008	0.000	0.024	0.006
		200	0.004	0.000	0.000	0.000
		201	0.008	0.000	0.006	0.000
		202	0.000	0.000	0.033	0.000
		203	0.001	0.007	0.026	0.000
		204	0.068	0.000	0.000	0.026
		205	0.097	0.001	0.049	0.299
		206	0.005	0.000	0.041	0.000
		207	0.081	0.000	0.036	0.000
		210	0.000	0.000	0.012	0.000
		211	0.000	0.000	0.003	0.000
		Null	0.075	0.064	0.020	0.012
IDVGA29	1.40%	134	0.112	0.000	0.000	0.000
		136	0.524	0.016	0.697	0.400
		143	0.150	0.043	0.298	0.597
		156	0.025	0.914	0.000	0.000
		Null	0.189	0.026	0.006	0.002
IDVGA55	1.90%	191	0.106	0.000	0.038	0.034
		193	0.012	0.000	0.109	0.000
		195	0.221	0.001	0.254	0.006
		197	0.252	0.000	0.298	0.294
		199	0.112	0.003	0.166	0.649
		202	0.000	0.000	0.029	0.000
		204	0.001	0.000	0.054	0.000
		208	0.001	0.000	0.000	0.000
		210	0.003	0.871	0.001	0.000
		212	0.000	0.064	0.000	0.000
		214	0.000	0.051	0.000	0.000
		217	0.246	0.000	0.008	0.008
		219	0.000	0.000	0.019	0.000
		221	0.000	0.000	0.000	0.000
		Null	0.044	0.008	0.023	0.007
INRA005	0.10%	124	0.000	0.036	0.000	0.000
		126	0.979	0.042	0.994	0.995
		129	0.000	0.002	0.000	0.000
		136	0.000	0.002	0.000	0.000
		143	0.000	0.917	0.000	0.000
		Null	0.020	0.001	0.005	0.005
INRA006	0.10%	128	0.003	0.000	0.000	0.000
		130	0.001	0.953	0.000	0.000
		132	0.001	0.000	0.051	0.000
		134	0.638	0.039	0.638	0.892
		136	0.180	0.003	0.288	0.098
		138	0.001	0.000	0.012	0.000
		Null	0.177	0.005	0.010	0.010
INRA131	0.00%	87	0.000	0.000	0.000	0.000
		92	0.111	0.000	0.051	0.000
		94	0.006	0.086	0.009	0.000
		98	0.446	0.001	0.567	0.890
		100	0.359	0.000	0.243	0.003
		102	0.006	0.000	0.090	0.000
		104	0.001	0.000	0.034	0.099
		106	0.000	0.780	0.000	0.000
		113	0.000	0.048	0.000	0.000
		115	0.000	0.011	0.000	0.000
		Null	0.069	0.073	0.008	0.008
MM012	0.10%	89	0.624	0.102	0.715	0.991
		91	0.262	0.003	0.270	0.004
		93	0.048	0.839	0.000	0.000
		97	0.000	0.000	0.001	0.000
		104	0.000	0.000	0.000	0.000
		Null	0.065	0.056	0.013	0.005
RM012	0.50%	116	0.001	0.992	0.004	0.000
		144	0.002	0.000	0.042	0.665
		129	0.047	0.000	0.084	0.001
		141	0.097	0.000	0.094	0.002
		133	0.125	0.000	0.293	0.000
		127	0.126	0.000	0.060	0.027
		139	0.049	0.005	0.081	0.180
		125	0.169	0.001	0.186	0.000
		137	0.008	0.000	0.020	0.000
		131	0.039	0.000	0.091	0.069
		151	0.208	0.000	0.018	0.012
		120	0.016	0.000	0.001	0.017
		Null	0.112	0.002	0.025	0.026
RM188	0.80%	115	0.005	0.000	0.025	0.002
		117	0.003	0.000	0.046	0.000
		121	0.000	0.000	0.001	0.000
		123	0.035	0.000	0.048	0.000
		125	0.092	0.000	0.078	0.018
		127	0.173	0.001	0.386	0.940
		129	0.051	0.009	0.259	0.001
		131	0.001	0.000	0.042	0.000
		132	0.270	0.000	0.001	0.009
		133	0.000	0.000	0.001	0.000
		134	0.109	0.000	0.051	0.000
		137	0.091	0.000	0.039	0.001
		139	0.001	0.037	0.004	0.000
		141	0.000	0.009	0.000	0.000
		143	0.001	0.551	0.000	0.000

		161	0.000	0.205	0.000	0.000
		163	0.000	0.002	0.000	0.000
		176	0.000	0.027	0.000	0.000
		182	0.000	0.143	0.000	0.000
		Null	0.166	0.015	0.020	0.029
RM95	0.40%	122	0.069	0.795	0.004	0.049
		132	0.281	0.000	0.082	0.011
		128	0.082	0.000	0.174	0.318
		136	0.103	0.000	0.044	0.355
		138	0.077	0.000	0.106	0.000
		130	0.239	0.010	0.350	0.004
		124	0.006	0.000	0.110	0.000
		147	0.000	0.001	0.000	0.000
		118	0.001	0.000	0.037	0.257
		140	0.032	0.000	0.019	0.000
		126	0.001	0.000	0.053	0.000
		134	0.000	0.000	0.007	0.000
		142	0.028	0.000	0.002	0.000
		116	0.000	0.117	0.000	0.000
		120	0.000	0.000	0.002	0.000
		144	0.013	0.000	0.000	0.000
		153	0.002	0.000	0.000	0.000
		Null	0.065	0.077	0.010	0.007
RME025	0.50%	132	0.090	0.000	0.000	0.000
		134	0.035	0.000	0.000	0.000
		136	0.006	0.000	0.000	0.000
		151	0.005	0.000	0.025	0.000
		155	0.001	0.000	0.081	0.000
		157	0.000	0.000	0.000	0.006
		159	0.000	0.000	0.004	0.000
		168	0.640	0.003	0.722	0.986
		170	0.024	0.009	0.130	0.000
		183	0.000	0.000	0.001	0.000
		193	0.001	0.980	0.000	0.000
		207	0.002	0.000	0.012	0.000
		Null	0.196	0.009	0.025	0.007
TGLA40	0.30%	91	0.160	0.000	0.225	0.000
		96	0.024	0.000	0.000	0.000
		97	0.155	0.008	0.531	0.722
		98	0.002	0.000	0.000	0.000
		99	0.060	0.000	0.052	0.001
		101	0.484	0.000	0.140	0.242
		102	0.000	0.000	0.003	0.000
		104	0.001	0.760	0.001	0.000
		106	0.000	0.155	0.000	0.000
		108	0.005	0.001	0.000	0.000
		Null	0.109	0.076	0.048	0.034
TGLA126	0.00%	100	0.001	0.351	0.001	0.000
		101	0.000	0.550	0.000	0.000
		104	0.032	0.000	0.000	0.000
		105	0.899	0.053	0.926	0.992
		130	0.000	0.003	0.000	0.000
		132	0.000	0.000	0.002	0.000
		134	0.000	0.000	0.008	0.000
		136	0.000	0.000	0.004	0.000
		138	0.000	0.000	0.001	0.000
		Null	0.066	0.041	0.058	0.007
TGLA127	0.20%	161	0.000	0.604	0.001	0.000
		167	0.001	0.000	0.018	0.000
		169	0.074	0.006	0.373	0.191
		171	0.001	0.000	0.000	0.000
		172	0.001	0.003	0.004	0.000
		174	0.041	0.290	0.027	0.007
		176	0.148	0.000	0.002	0.001
		178	0.188	0.010	0.205	0.710
		180	0.211	0.000	0.043	0.001
		182	0.004	0.000	0.000	0.000
		184	0.016	0.000	0.115	0.078
		186	0.063	0.000	0.082	0.000
		188	0.000	0.000	0.002	0.000
		190	0.162	0.000	0.056	0.006
		192	0.042	0.000	0.043	0.000
		Null	0.048	0.086	0.030	0.006
TGLA337	8.40%	111	0.037	0.000	0.000	0.000
		118	0.130	0.000	0.000	0.000
		126	0.026	0.594	0.002	0.001
		128	0.001	0.039	0.000	0.000
		130	0.301	0.000	0.153	0.471
		132	0.001	0.000	0.137	0.000
		134	0.000	0.000	0.002	0.000
		136	0.094	0.005	0.256	0.375
		138	0.004	0.193	0.051	0.000
		142	0.000	0.000	0.001	0.000
		145	0.191	0.000	0.247	0.124
		147	0.079	0.021	0.069	0.008
		153	0.001	0.000	0.001	0.000
		155	0.001	0.016	0.003	0.000
		Null	0.133	0.131	0.077	0.021
UWCA47	0.50%	225	0.229	0.000	0.006	0.000
		229	0.014	0.000	0.060	0.002
		231	0.726	0.079	0.895	0.967
		240	0.000	0.882	0.000	0.000
		Null	0.031	0.038	0.039	0.031

Table 2.A4. Posterior allele frequencies from analysis 3 (n = 2230) at K = 4.

Locus	% Missing data	Allele Size	Estimated allele frequency in Red Cluster III	Estimated allele frequency in Red Cluster II	Estimated allele frequency in Wapiti	Estimated allele frequency in Red Cluster I
AGLA293	1.70%	128	0.000	0.002	0.001	0.107
		144	0.991	0.943	0.883	0.735
		147	0.001	0.032	0.040	0.065
		149	0.000	0.000	0.055	0.000
		Null	0.007	0.022	0.021	0.093
BM4006	0.50%	85	0.000	0.000	0.000	0.004
		87	0.360	0.079	0.001	0.072
		93	0.134	0.806	0.992	0.834
		95	0.505	0.096	0.002	0.086
		Null	0.001	0.018	0.005	0.004
BM6438	1.30%	249	0.424	0.548	0.005	0.579
		251	0.001	0.216	0.002	0.236
		253	0.002	0.039	0.001	0.119
		257	0.000	0.041	0.000	0.001
		261	0.565	0.001	0.031	0.026
		263	0.000	0.000	0.795	0.000
		265	0.000	0.001	0.121	0.000
		275	0.000	0.008	0.000	0.000
		Null	0.009	0.145	0.043	0.039
BM757	0.00%	160	0.031	0.133	0.001	0.066
		162	0.959	0.464	0.003	0.505
		164	0.000	0.000	0.000	0.009
		172	0.000	0.000	0.000	0.003
		173	0.000	0.000	0.232	0.000
		174	0.000	0.000	0.000	0.004
		175	0.000	0.000	0.049	0.000
		177	0.000	0.000	0.203	0.000
		179	0.000	0.100	0.001	0.053
		183	0.000	0.024	0.001	0.092
		185	0.000	0.000	0.000	0.059
		187	0.004	0.236	0.043	0.014
		189	0.000	0.000	0.000	0.003
		192	0.000	0.000	0.058	0.000
		196	0.000	0.000	0.000	0.001
		197	0.000	0.000	0.000	0.000
		198	0.000	0.001	0.201	0.071
		200	0.000	0.001	0.203	0.086
		202	0.000	0.000	0.000	0.014
		210	0.000	0.000	0.000	0.006
		Null	0.004	0.038	0.003	0.014
BOVIRP	0.60%	140	0.000	0.012	0.000	0.002
		142	0.000	0.000	0.000	0.000
		145	0.000	0.000	0.337	0.000
		147	0.000	0.001	0.337	0.071
		149	0.015	0.329	0.001	0.030
		151	0.480	0.078	0.050	0.165
		153	0.025	0.357	0.003	0.423
		155	0.071	0.002	0.001	0.063
		157	0.406	0.126	0.068	0.186
		159	0.000	0.087	0.001	0.015
		161	0.000	0.000	0.155	0.000
		163	0.000	0.000	0.000	0.000
		Null	0.003	0.007	0.047	0.045
FCB193	2.50%	101	0.000	0.006	0.000	0.001
		103	0.000	0.001	0.000	0.048
		105	0.000	0.000	0.000	0.001
		107	0.000	0.034	0.001	0.113
		109	0.000	0.016	0.001	0.200
		111	0.004	0.168	0.000	0.002
		113	0.037	0.206	0.001	0.272
		115	0.000	0.000	0.000	0.011
		118	0.004	0.016	0.001	0.046
		120	0.304	0.291	0.071	0.055
		122	0.126	0.072	0.021	0.105
		124	0.000	0.001	0.000	0.055
		126	0.078	0.001	0.657	0.002
		128	0.045	0.045	0.008	0.003
		130	0.366	0.001	0.000	0.034
		132	0.032	0.011	0.020	0.003
		134	0.000	0.002	0.025	0.005
		140	0.000	0.035	0.000	0.000
		141	0.000	0.000	0.000	0.000
		143	0.000	0.000	0.041	0.009
		145	0.000	0.000	0.028	0.000
		150	0.000	0.000	0.121	0.000
		Null	0.003	0.094	0.002	0.033
FSHB	0.80%	180	0.000	0.027	0.000	0.003
		182	0.000	0.000	0.619	0.000
		184	0.000	0.000	0.114	0.064
		185	0.000	0.070	0.109	0.230
		186	0.000	0.000	0.000	0.001
		187	0.000	0.000	0.010	0.004
		188	0.296	0.137	0.138	0.104
		189	0.052	0.221	0.001	0.129
		190	0.002	0.018	0.000	0.001
		191	0.196	0.078	0.001	0.074
		192	0.000	0.000	0.000	0.023

		193	0.000	0.000	0.000	0.000
		194	0.000	0.000	0.000	0.028
		195	0.000	0.000	0.000	0.000
		196	0.000	0.004	0.000	0.012
		197	0.006	0.002	0.000	0.006
		198	0.106	0.059	0.001	0.069
		199	0.006	0.011	0.001	0.024
		200	0.000	0.002	0.000	0.000
		201	0.000	0.010	0.000	0.006
		202	0.000	0.000	0.000	0.033
		203	0.000	0.002	0.000	0.024
		204	0.026	0.082	0.000	0.000
		205	0.297	0.116	0.001	0.049
		206	0.000	0.009	0.000	0.042
		207	0.000	0.099	0.000	0.037
		210	0.000	0.000	0.000	0.013
		211	0.000	0.000	0.000	0.003
		Null	0.011	0.050	0.001	0.020
IDVGA29	1.80%	134	0.000	0.000	0.971	0.000
		136	0.398	0.760	0.008	0.696
		143	0.601	0.203	0.004	0.299
		156	0.000	0.030	0.000	0.001
		Null	0.001	0.006	0.017	0.004
IDGVA55	2.50%	191	0.036	0.028	0.474	0.036
		193	0.000	0.022	0.001	0.103
		195	0.007	0.285	0.002	0.247
		197	0.296	0.205	0.514	0.303
		199	0.650	0.140	0.002	0.169
		202	0.000	0.000	0.000	0.030
		204	0.000	0.000	0.000	0.056
		210	0.000	0.007	0.000	0.003
		217	0.006	0.306	0.001	0.007
		219	0.000	0.000	0.000	0.020
		221	0.000	0.000	0.000	0.000
		Null	0.005	0.006	0.005	0.024
INRA005	0.20%	126	0.997	0.992	0.995	0.992
		143	0.000	0.000	0.000	0.002
		Null	0.003	0.008	0.004	0.006
INRA006	0.00%	128	0.000	0.002	0.000	0.000
		130	0.000	0.000	0.000	0.002
		132	0.000	0.001	0.001	0.054
		134	0.893	0.896	0.123	0.637
		136	0.100	0.068	0.872	0.285
		138	0.000	0.000	0.000	0.013
		Null	0.006	0.032	0.003	0.009
INRA131	0.00%	87	0.000	0.000	0.000	0.000
		92	0.000	0.009	0.613	0.051
		94	0.000	0.008	0.000	0.009
		98	0.893	0.530	0.279	0.575
		100	0.003	0.438	0.103	0.234
		102	0.000	0.011	0.001	0.089
		104	0.100	0.001	0.001	0.036
		106	0.000	0.000	0.000	0.001
		Null	0.003	0.002	0.003	0.005
MM012	0.00%	89	0.993	0.801	0.113	0.718
		91	0.004	0.185	0.622	0.265
		93	0.000	0.000	0.263	0.001
		97	0.000	0.000	0.000	0.002
		104	0.000	0.000	0.000	0.000
		Null	0.003	0.014	0.001	0.014
RM012	0.70%	116	0.000	0.000	0.000	0.007
		120	0.017	0.020	0.000	0.001
		125	0.000	0.210	0.001	0.177
		127	0.027	0.003	0.779	0.062
		129	0.001	0.065	0.001	0.086
		131	0.069	0.054	0.001	0.091
		133	0.001	0.171	0.009	0.295
		137	0.000	0.001	0.039	0.020
		139	0.181	0.028	0.154	0.081
		141	0.002	0.139	0.001	0.093
		144	0.669	0.002	0.001	0.045
		151	0.011	0.265	0.001	0.018
		Null	0.022	0.042	0.013	0.025
RM188	0.90%	115	0.002	0.007	0.001	0.024
		117	0.000	0.007	0.001	0.046
		121	0.000	0.000	0.000	0.001
		123	0.000	0.045	0.001	0.049
		125	0.017	0.115	0.001	0.080
		127	0.946	0.199	0.137	0.383
		129	0.001	0.080	0.001	0.261
		131	0.000	0.001	0.000	0.041
		132	0.008	0.332	0.019	0.001
		133	0.000	0.000	0.000	0.001
		134	0.000	0.001	0.792	0.049
		137	0.001	0.106	0.021	0.034
		139	0.000	0.000	0.000	0.004
		143	0.000	0.002	0.000	0.001

		161	0.000	0.000	0.000	0.000
		182	0.000	0.000	0.000	0.000
		Null	0.024	0.104	0.025	0.024
RM95	0.50%	118	0.259	0.001	0.001	0.039
		120	0.000	0.000	0.000	0.002
		122	0.049	0.021	0.300	0.005
		124	0.000	0.001	0.031	0.115
		126	0.000	0.001	0.000	0.049
		128	0.316	0.105	0.001	0.168
		130	0.004	0.305	0.002	0.357
		132	0.010	0.358	0.001	0.074
		134	0.000	0.000	0.000	0.007
		136	0.354	0.074	0.240	0.045
		138	0.001	0.063	0.184	0.108
		140	0.000	0.042	0.001	0.019
		142	0.000	0.000	0.151	0.002
		144	0.000	0.000	0.073	0.000
		153	0.000	0.000	0.009	0.000
		Null	0.005	0.031	0.006	0.009
RME025	0.40%	132	0.000	0.000	0.709	0.000
		134	0.000	0.000	0.227	0.000
		136	0.000	0.000	0.030	0.000
		151	0.000	0.006	0.001	0.026
		155	0.000	0.001	0.001	0.086
		157	0.006	0.000	0.000	0.000
		159	0.000	0.000	0.000	0.004
		168	0.989	0.953	0.007	0.724
		170	0.000	0.026	0.001	0.121
		183	0.000	0.000	0.000	0.002
		193	0.000	0.003	0.000	0.003
		207	0.000	0.003	0.001	0.013
		Null	0.004	0.008	0.023	0.022
TGLA40	0.40%	91	0.001	0.213	0.002	0.223
		96	0.000	0.029	0.000	0.000
		97	0.732	0.059	0.661	0.526
		98	0.000	0.002	0.000	0.000
		99	0.001	0.002	0.326	0.054
		101	0.238	0.650	0.002	0.140
		102	0.000	0.000	0.000	0.003
		104	0.000	0.002	0.000	0.003
		108	0.000	0.004	0.000	0.000
		Null	0.028	0.038	0.007	0.050
TGLA126	0.00%	100	0.000	0.000	0.000	0.002
		101	0.000	0.000	0.000	0.002
		104	0.000	0.000	0.190	0.000
		105	0.994	0.983	0.786	0.921
		132	0.000	0.000	0.000	0.002
		134	0.000	0.000	0.000	0.008
		136	0.000	0.000	0.000	0.004
		138	0.000	0.000	0.000	0.001
		Null	0.006	0.016	0.022	0.061
TGLA127	0.20%	161	0.000	0.000	0.000	0.002
		167	0.000	0.002	0.000	0.018
		169	0.191	0.104	0.002	0.368
		171	0.000	0.000	0.000	0.000
		172	0.000	0.000	0.000	0.004
		174	0.007	0.053	0.021	0.028
		176	0.000	0.184	0.000	0.001
		178	0.712	0.088	0.625	0.203
		180	0.001	0.191	0.320	0.043
		182	0.000	0.000	0.018	0.000
		184	0.078	0.018	0.001	0.115
		186	0.000	0.087	0.001	0.083
		188	0.000	0.000	0.000	0.002
		190	0.006	0.203	0.001	0.058
		192	0.000	0.053	0.001	0.042
		Null	0.004	0.015	0.010	0.031
TGLA337	0.20%	111	0.000	0.000	0.183	0.000
		118	0.000	0.000	0.760	0.000
		126	0.002	0.033	0.001	0.004
		128	0.000	0.000	0.000	0.000
		130	0.470	0.428	0.010	0.151
		132	0.000	0.001	0.001	0.137
		134	0.000	0.000	0.000	0.003
		136	0.367	0.131	0.002	0.242
		138	0.000	0.007	0.001	0.054
		142	0.000	0.000	0.000	0.001
		145	0.119	0.274	0.002	0.251
		147	0.008	0.107	0.001	0.071
		153	0.000	0.000	0.000	0.001
		155	0.000	0.000	0.000	0.003
		Null	0.033	0.017	0.038	0.083
UWCA47	0.60%	225	0.000	0.279	0.001	0.006
		229	0.002	0.018	0.001	0.062
		231	0.969	0.691	0.988	0.896
		240	0.000	0.000	0.000	0.000
		Null	0.029	0.012	0.010	0.036

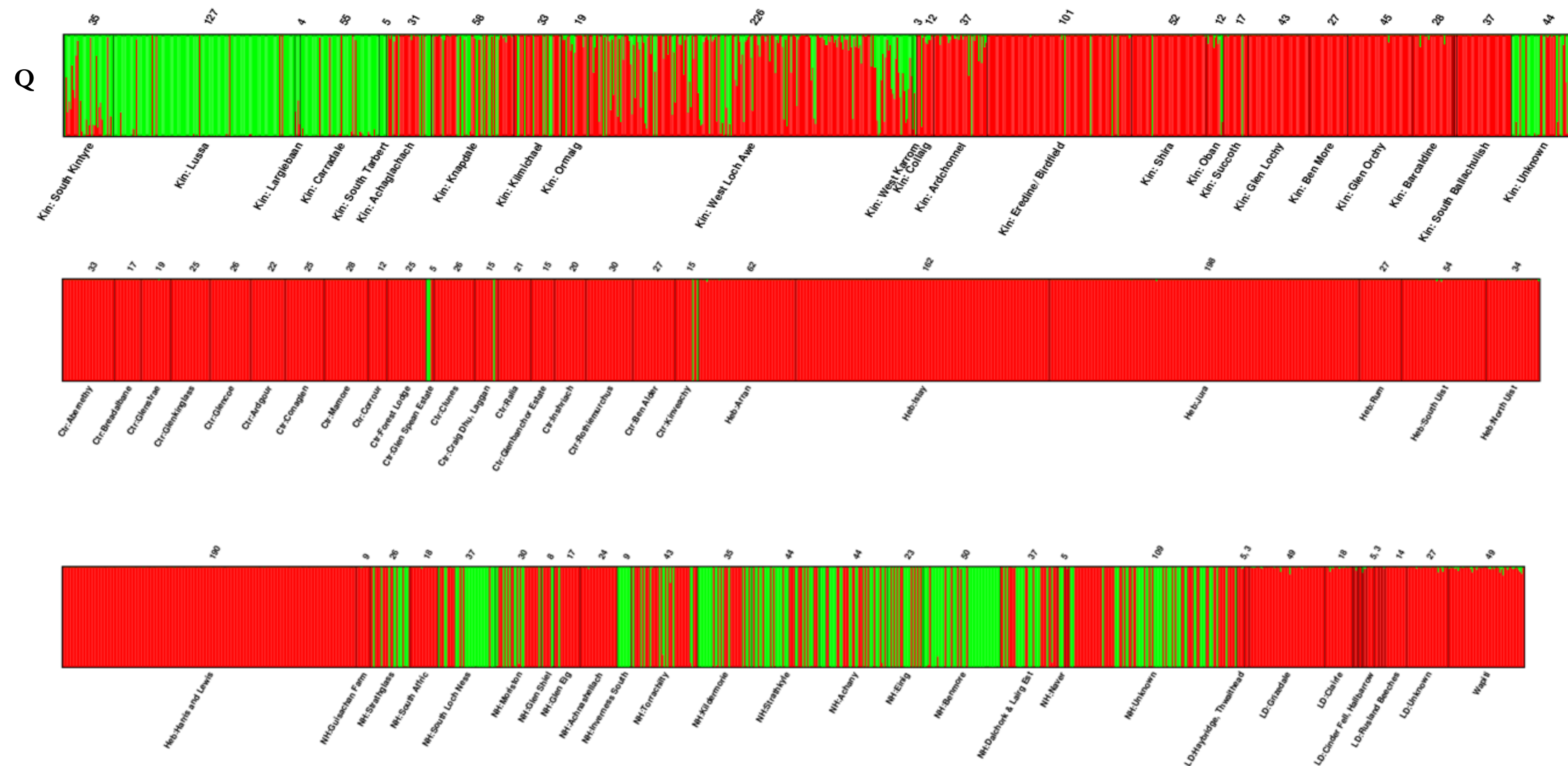


Figure 2A1. Bar chart showing the results of *analysis 1* at in STRUCTURE at K = 2 for each individual in the dataset consisting of red, sika and wapiti animals (n = 2943). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (shown in red) and the proportion attributable to sika ancestry (green). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis. Scottish sites are plotted in an approximately south to north order, followed by the sample sites in the Lake District, Cumbria and lastly the wapiti controls. Abbreviations represent; Kin= Kintyre, Ctr= Central highlands, Heb= Hebrides, NH= North Highlands and LD= Lake District, Cumbria.

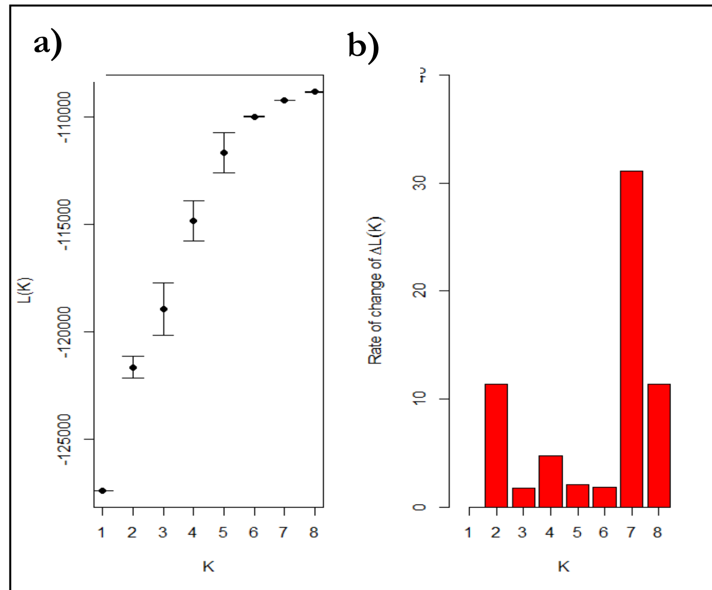


Figure 2.A2. Assessment of the most likely number of populations using Structure 2.3.3 analysis 3 of dataset containing red deer and wapiti individuals only (n = 2,230) at K = 1 – 8. Two likelihood parameters are assessed; of which the results for a) the log-likelihood (with standard error) of the each value of K (number of populations) given the dataset and b) the rate of change in log likelihood between values of K. Whilst K = 7 is most likely, K = 4 is used to meet the objective of this analysis.

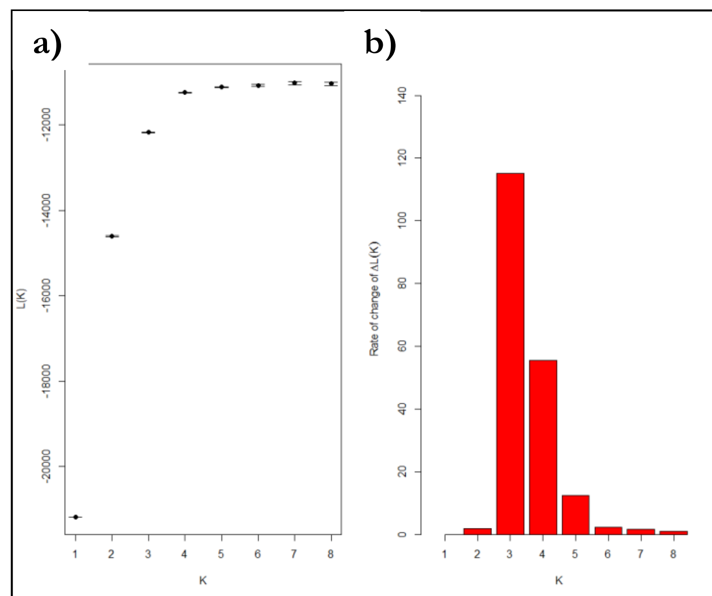


Figure 2.A3. LnPr (X | K) for analysis 4. Assessment of the most likely number of populations using Structure 2.3.3 analysis 4 of dataset containing sika and wapiti individuals only (n = 591) at K = 1 – 8. Two likelihood parameters are assessed; of which the results for a) the log-likelihood (with standard error) of the each value of K (number of populations) given the dataset and b) the rate of change in log likelihood between values of K. Both provide evidence that K = 3 are the most likely.

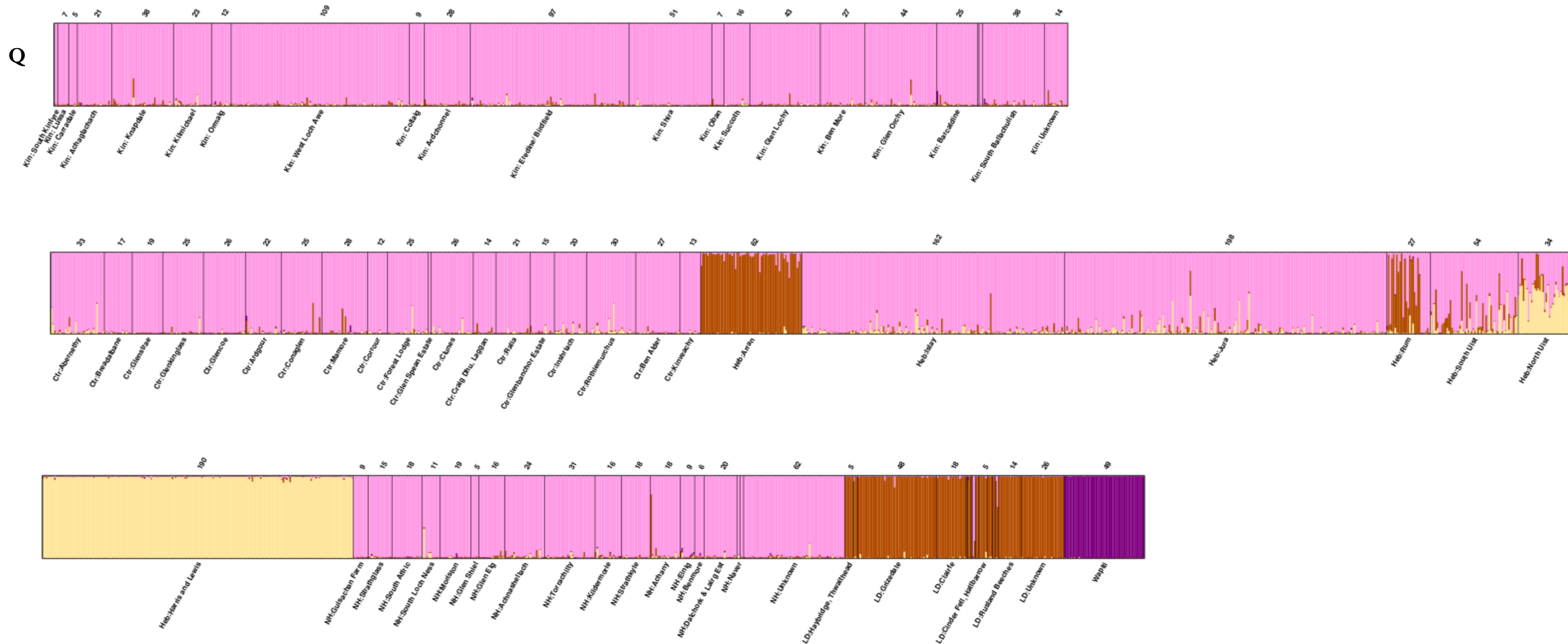


Figure 2.A4. Bar chart showing the results of *analysis 3* in STRUCTURE at K = 4 for each individual in the dataset consisting of red deer and wapiti animals only (n = 2230). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to three red clusters (pink, cream and brown) and the proportion attributable to wapiti ancestry (purple). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis. Scottish sites are plotted in an approximately south to north order, followed by the sample sites in the Lake District, Cumbria and lastly the wapiti controls. Abbreviations represent; Kin= Kintyre, Ctr= Central highlands, Heb= Hebrides, NH= North Highlands and LD= Lake District, Cumbria.

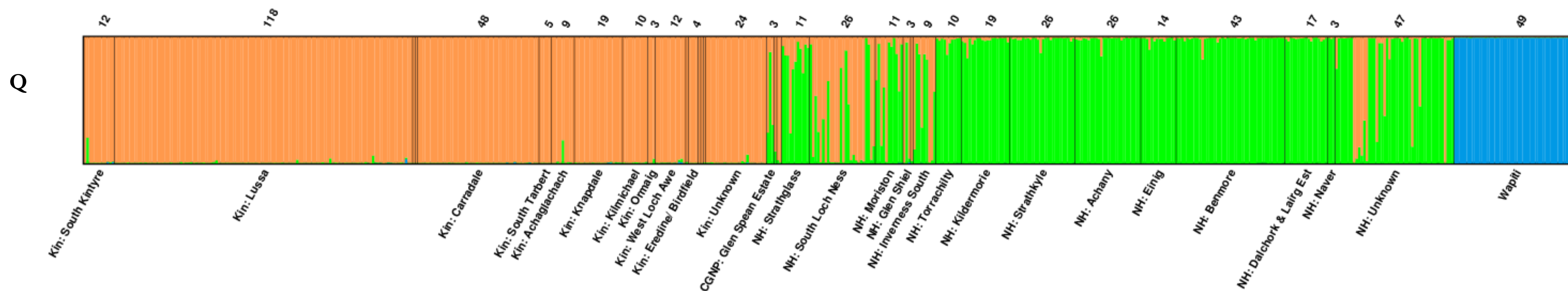


Figure 2.A5. Bar chart showing the results of *analysis 4* in STRUCTURE at K = 3 for each individual in the dataset consisting of sika and wapiti animals only (n = 591). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to sika ancestry (shown in green and orange) and the proportion attributable to wapiti ancestry (blue). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis. Scottish sites are plotted in an approximately south to north order, followed by the sample sites in the Lake District, Cumbria and lastly the wapiti controls. Abbreviations represent; Kin= Kintyre, CGNP = Ctr= Central highlands and NH= North Highlands.

Chapter 3: A survey of the hybridisation status of *Cervus* deer species on the island of Ireland

Author's contributions:

SS organised the sampling and completed the genotyping of 350 samples provided through collaboration with Timothy Burkitt, Ruth Carden and Barry Coad. A further 24 individuals from Ireland were genotyped by Helen Senn. 98 red and sika individuals sampled from Scotland were used as control animals, 91 of which were genotyped by Helen Senn and 8 of which were genotyped by SS. Statistical analysis was performed by SS. SS wrote the MS. JMP guided the study and edited and commented on the MS.

3.1 Abstract

There are now an estimated 4,000 red deer (*Cervus elaphus*) in Ireland and their numbers are increasing. It has recently been confirmed that red deer in County Kerry are descended from an ancient (c. 5,000 BP) introduction and therefore merit genetic conservation. During the mid-19th century exotic species including North American wapiti (*C. canadensis)* and Japanese sika deer (*C. nippon*) were introduced to Ireland via Powerscourt Estate, County Wicklow. Although wapiti struggled to establish themselves, sika thrived and have since dispersed within Co. Wicklow and been translocated to other sites throughout Ireland. Red and sika deer are known to have hybridised in Ireland, but to date there has been no survey of hybridisation and introgression between these species using a panel of highly diagnostic genetic markers. In this study 374 individuals were genotyped at a highly diagnostic set of 22 microsatellite loci and a mtDNA marker. A Bayesian clustering approach and cytonuclear disequilibria were used to assess the extent of hybridisation. Wapiti introgression was very low (trace evidence in 0.53% of individuals), suggesting hybridisation and introgression by this species is negligible. However, 80/197 (41%) deer sampled in Co. Wicklow and 7/15 (47%) deer sampled in Co. Cork were red-sika hybrids according to either their nuclear genome or mitochondrial haplotype. No pure red deer were detected in Co. Wicklow, suggesting that in this region the red deer has disappeared following hybridisation. In contrast, no hybrids were detected in Co. Kerry despite the extensive sympatry of the two species in this area. However, the Co. Cork hybrids pose a threat to the Co. Kerry populations due to their proximity.

Key words: Cervus, microsatellite, hybridisation, introgression, mtDNA, sika, red deer.

3.2 Introduction

3.2.1 *Hybridisation*

Hybridization is the interbreeding of genetically distinct taxa and is widespread amongst eukaryotes (Allendorf *et al.* 2001). Introgression is the resultant gene flow (described as 'horizontal') between populations whose members are hybridising and can dramatically influence the evolutionary trajectory of a species (Allendorf *et al.* 2001). Hybridisation can occur naturally (e.g. between *Partula* spp. of land snails (Clarke *et al.* 1998)); however habitat degradation and species transfer by humans can also bring non-native species into contact with native heterospecifics. The detrimental impact of such anthropogenically-induced hybridisation can vary; the generation of hybrids between an

invasive and a native without introgression (e.g. due to strong negative selection against F1 hybrids) can result in substantial wasted reproductive costs, whilst if introgression does occur, it can be highly destructive to the integrity of the locally adapted species, race or ecotype and can lead to extinction (Allendorf *et al.* 2001). Examples of anthropogenically-induced hybridisation include that between Antarctic fur seals (*Arctocephalus* spp.) and New Zealand fur seals (*A. forsterri*) threatening population homogenisation due to the disturbances caused by seal harvesting; between endemic mouse lemur species from Southern Madagascar (*Microcebus* spp.) where deforestation has facilitated asymmetric gene flow; the introgression of maladaptive gene complexes into wild American mink (*Neovison vison*) from escaped domestic farmed American mink, causing population decline; and the generation of sterile hybrids between the native Bull trout (*Salvelinus confluetus*) and the introduced brook trout (*S. fontinalis)* in North America, ultimately leading to the displacement of the former species (Gligor *et al.* 2009; Kidd *et al.* 2009; Lancaster *et al.* 2006; Leary *et al.* 1993).

3.2.2 Cervus *in Ireland*

Since the mid-19th century, a series of introductions of exotic deer including North American wapiti (*Cervus canadensis)* and Japanese sika (*C. nippon nippon*) into the British Isles has created many opportunities for hybridisation with the red deer (*C. elaphus*), which is native to Britain and was introduced to Ireland as long ago as 5,000 BP (Carden *et al.* 2012). Hybridisation between red deer and the physically larger wapiti has occurred in Scotland (Whitehead 1964) and following the introduction of both species to Fiordland, New Zealand, a heavily introgressed population exists there now (Shackell *et al.* 2003). They can be hybridised with relative ease in captivity and hybrids are now common on New Zealand deer farms (Moore & Littlejohn 1989; Shackell *et al.* 2003). However, in the British Isles the unfavourable climate has prevented wapiti becoming visibly established in the wild, perhaps because it is highly susceptible to lung disease and foot malformation, delayed female maturity and lower levels of stag aggression than red deer in the rut (Asher *et al.* 2005; Pérez-Espona *et al.* 2010a).

Compared with wapiti, introductions of the diminutive Japanese sika deer have proved more successful and hybridisation and introgression appears to occur readily. Based on phenotype, hybridisation has been documented in captivity (Harrington 1973; Powerscourt 1884) and in the wild in Britain, Ireland and the former Czechoslovakia (Bartos *et al.* 1981; Harrington 1973; Lowe & Gardiner 1975). Hybridisation has also been documented by means of biochemical and genetic approaches in both Ireland and

Britain (Goodman *et al.* 1999; Harrington 1973; McDevitt *et al.* 2009a; Senn & Pemberton 2009). A study in Argyll, Scotland found that pregnancy rates of genetically-confirmed female hybrids do not differ significantly from the parental species, indicating little selection against hybrids in this fitness component (Senn *et al.* 2010b).

Hybridisation between *Cervus* deer has substantial phenotypic consequences. On New Zealand deer farms, red deer have been deliberately and successfully hybridised with wapiti to increase carcass and antler size (Moore & Littlejohn 1989). However, it is thought that earlier attempts to achieve this in British deer parks were generally unsuccessful (Whitehead 1964). In a wild red-sika study system in Kintyre, Argyll, Senn *et al.* (2010) regressed phenotypic trait values against genetically-determined hybrid scores to quantify the impact of red-sika hybridisation on phenotype. Carcass weight was greater in sika-like hybrids than in 'pure' sika and lower in red-like hybrid females than in 'pure' red females. Within sika-like females, hybrids had increased jaw length and incisor arcade breadth (IAB) compared with 'pure' sika, whilst IAB was low in red-like hybrid females compared to 'pure' red (see below for definition of 'pure'). Overall, phenotypic modifications such as these highlight the (additive) genetic variation for quantitative traits in hybrid deer and the substantial potential for change under selection. This can greatly exacerbate effective management of these populations. Harrington (1973) noted that the predominance of red-like characters amongst hybrid individuals in Co. Wicklow made the identification of hybrids or introgressed animals there very difficult.

In Ireland, the modern red deer population is descended from ancient and recent postglacial introductions by man (Carden *et al.* 2012). There are currently thought to be around 4,000 phenotypically red deer in Ireland (Pérez-Espona *et al.* 2009a). They are present in the East (Co. Wicklow), the South West (Co. Kerry) and the North West (Co. Galway north to Co. Donegal) and have shown a 7% range expansion from these sites over the last 30 years (Carden *et al.* 2010; Pérez-Espona *et al.* 2009a; Whitehead 1964). Recent work has established that the red deer centred on Killarney, Co. Kerry, are descended from a human introduction from Britain during the Neolithic period (Carden *et al.* 2012). The other populations are descended from more recent introductions from Britain and continental Europe, in several cases indirectly through deer parks, primarily Powerscourt Park, Co. Wicklow, where they may have interacted with other *Cervus* species (Carden *et al.* 2012; McDevitt *et al.* 2009a). Co. Kerry has itself received more recent red deer introductions from Co. Roscommon and possibly from Scotland and from Windsor Great Park (Whitehead 1964).

At least four exotic subspecies or species of the genus *Cervus* have been introduced to Ireland, largely through the activities of Viscount Powerscourt at his deer park in Co. Wicklow (Powerscourt 1884). In 1865 two male and one female North American wapiti were introduced to Powerscourt and towards the end of the 1800s a single wapiti female was introduced to Co. Tyrone (Whitehead 1964). Whilst reports suggest there were up to five wapiti individuals at Powerscourt, around 1880 these animals had been disposed of (Powerscourt 1884; Whitehead 1964) and we are not aware of any other reports of hybrids between red deer and wapiti in Ireland. Japanese sika were also introduced to Powerscourt in 1860. Although this introduction involved only one male and three females, they successfully established and by 1884 there were over 100, despite culling and translocation to other counties (Powerscourt 1884). A further sika subspecies, Manchurian sika (*C. n. mantchuricus*) was introduced and is now free living in Co. Mayo and may have been supplemented by further illegitimate translocations (McDevitt *et al.* 2009a). Despite little information surrounding their introduction to the park Powerscourt Park, there is mention of hybrids between Manchurian sika and the red deer in Powerscourt, but the fate of these animals is unknown (McDevitt *et al.* 2009a; Powerscourt 1884; Ratcliffe 1987). Red-Manchurian sika hybrids were also suspected in Co. Fermanagh between 1885 and 1891 (Whitehead 1964). Lastly, sambar deer (*C. unicolor*) were also introduced around the mid-19th century to Powerscourt (Powerscourt 1884). Hybrids between red deer and sambar were also reported in Powerscourt; however, these are believed to have died out (Powerscourt 1884). The deer park at Powerscourt was disbanded by 1960, when it is believed poachers broke down the perimeter walls and the deer escaped (Powerscourt Estate staff, pers. comm).

Overwhelmingly, it is the Japanese sika that appears to have most frequently hybridised with red deer both in captivity and the wild in Ireland. Viscount Powerscourt reported three or four animals that were "certainly hybrids" between red and Japanese sika in his park, with the red hind in each case being the dam (Powerscourt 1884). Harrington (1973) reported around 250 hybrid animals across Co. Wicklow within a wild population of around 3,000 sika-like individuals. Based on phenotype, it has been apparent for many years that the Co. Wicklow deer population contains substantial numbers of red-sika hybrids and these likely originated from Powerscourt estate, during disturbances in 1922 (Harrington 1973; Whitehead 1964). These escapees are believed to have thrived outside the overcrowded conditions of Powerscourt estate and excelled when they returned for the rut, which probably resulted in further mixing (Delap 1936; Whitehead 1964). From Powerscourt individuals were translocated elsewhere, including Co. Kerry

in 1864 (McDevitt *et al.* 2009a) and various sites in the UK (Ratcliffe 1987). Over the last 30 years, sika have expanded their range at around 5% per annum from populations in the East, South West and North West of Ireland (Carden *et al.* 2010). The extent of hybridisation in the North West has been less well studied, however, the presence of red-sika hybrids in this region has been suspected (Carden *et al.* 2010; Harrington 1973; Pérez-Espona *et al.* 2009a). In contrast to this, there has been no evidence to date to suggest hybridisation has occurred in Co. Kerry. However, estimated expansion rates highlight the risk that sika and hybrid animals from the East may spread into apparently hybrid-free zones of the South West and hence threaten the genetic integrity of the ancient red deer in this area (Carden *et al.* 2010; Pérez-Espona *et al.* 2008).

Two previous studies have investigated the genetic interactions between red deer and Japanese sika in Ireland. Using rocket immunoelectrophoresis, Harrington (1973) found no pure red deer in Co. Wicklow, which matched his conclusions from phenotypic observations (above). Consistent with this finding, a survey using eight non-diagnostic microsatellite markers and analysis using the software package Structure (Pritchard *et al.* 2000) in Co. Wicklow, suggested that the majority of phenotypically red deer were actually hybrid, whilst the majority of sika were putatively pure (McDevitt *et al.* 2009a). As for the rest of Ireland historic reports have suggested the presence of hybrids across several counties (Whitehead 1964).

3.2.3 *This study*

Taken overall, concern is growing over the extent and consequences of hybridisation and introgression among deer, particularly between red and Japanese sika, in Ireland. Hybrid swarms, such as that previously documented in Co. Wicklow, may exist undetected elsewhere on the island of Ireland where these species overlap and could be expanding at a rate that threatens the genetic and phenotypic integrity of the ancient-origin red deer in Co. Kerry. This study builds on the preliminary work of McDevitt *et al.* (2009) by genotyping a large sample of individuals at a set of 22 microsatellites, which are highly diagnostic for red and Japanese sika and moderately diagnostic for red and wapiti and a single mtDNA marker that is diagnostic for red deer, Japanese sika and Manchurian sika. The specific objectives are:

1. *To assess the current extent of hybridisation and introgression between* Cervus *deer on the island of Ireland and whether it threatens either parental taxon.*
2. *To determine, if possible, the initial direction of hybridisation.*

3. To investigate whether any hybrids outside Co. Wicklow derive from the Wicklow hybrids or new hybridisation events.
4. To investigate the accuracy with which hybrids are identified from stalker-assigned phenotype.
5. To indicate what management actions may be required to protect putatively pure populations from hybridisation.

3.3 Materials and Methods

3.3.1 *Sampling Sites*

Samples were obtained from seven counties in the Republic of Ireland and a single county in Northern Ireland covering the major red deer and Japanese sika populations (Carden *et al.* 2010). Across the whole island, 392 individuals were collected for genotyping (details of the final dataset in Table 3.1; Figure 3.1). Most samples obtained from Co. Kerry, Co. Cork and Co. Wicklow were shot during the 2011-2012 season; the remaining samples, collected from all eight counties, were sampled between 2006 and 2012.

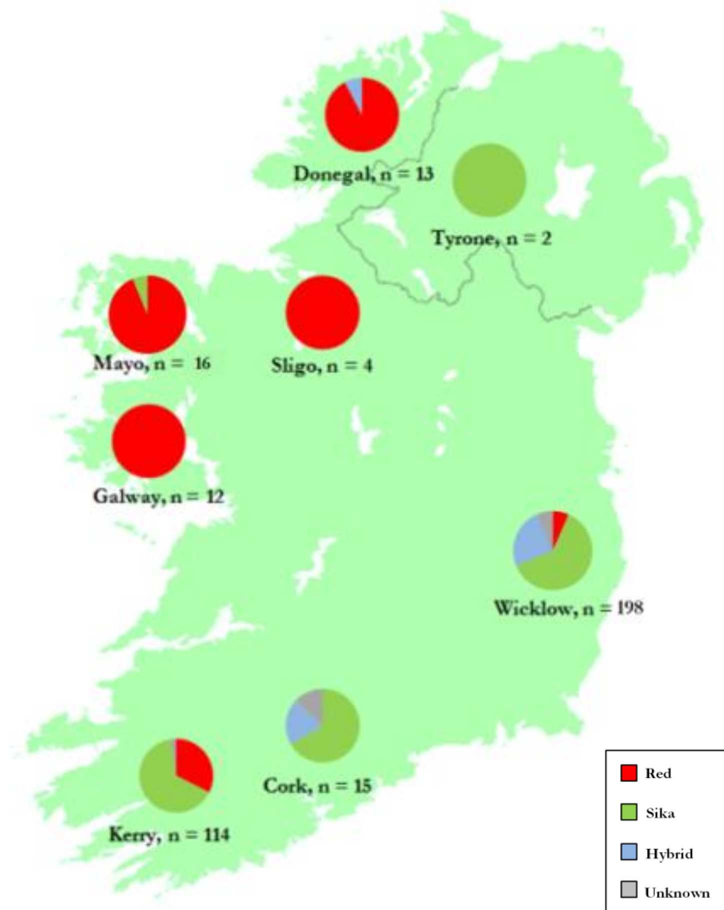


Figure 3.1. Counties from which samples were sourced in this study, showing the number of individuals sampled (n) and the proportion of stalker-assigned phenotypes at each site.

3.3.2 *DNA analysis*

Genomic DNA was extracted from samples with the DNeasy Blood and Tissue Kit (Qiagen) according to the manufacturer's instructions. Individual samples were genotyped at a panel of 22 diagnostic microsatellite markers following previously-published protocols (Senn & Pemberton 2009), the details of which are given in Appendix Table 2.A1. Originally derived from cattle (*Bos taurus*) and sheep (*Ovis aries*), these markers have been selected to discriminate between red deer and Japanese sika because when used to genotype 44 red deer and 44 sika from diverse geographical locations, they shared no common alleles (Goodman, 1999; Slate 1998). In addition, they also have some discriminatory power between red and wapiti (10/22 strongly diagnostic loci; J. Pemberton pers. comm). The marker panel is not diagnostic for Manchurian sika. PCR products were run on an ABI3730 capillary sequencer (Applied Biosystems), using the internal standard Genescan LIZ 500 (Applied Biosystems). Fragment analysis was carried out using Genemapper version 4.0 (Applied Biosystems).

Individuals were also screened for their haplotype in the mitochondrial control region, which in deer includes a diagnostic number of 39bp tandem repeats: red deer have a single repeat, Japanese sika have three and Manchurian sika have seven (Cook *et al.* 1999). Amplification followed a published protocol (Cook *et al.* 1999) and repeat number was determined by assay on 4% agarose gels stained with ethidium bromide (Goodman *et al.* 1999) where red deer have a 350bp band, Japanese sika a 430bp band and Manchurian sika a larger band.

3.3.3 *To assess the current extent of hybridisation and introgression between* Cervus *deer on the island of Ireland and whether it threatens either parental taxon (objective 1).*

The Bayesian clustering software Structure 2.3.3 (Falush *et al.* 2003, 2007; Pritchard *et al.* 2000) was used to analyse the extent of individual and population admixture using the microsatellite genotype data in a number of separate datasets. In the first analysis (analysis 1) the dataset of all Irish deer was supplemented with genotypes for 50 putatively pure red deer from central Scotland, 50 putatively pure Japanese sika from Kintyre, Scotland and 49 Canadian wapiti as control samples. For analysis 2 the wapiti samples and any Irish deer showing signs of wapiti introgression were deleted from the analysis 1 data set. The number of inferred, genetically distinct populations (K) that maximises the likelihood (Ln Pr (X|K)) of the dataset, assuming they are in Hardy-

Weinberg equilibrium and linkage equilibrium, was estimated by running five independent replicates at different values of K (1-8) and selecting the smallest value of K with the highest log likelihood (Ln Pr (X|K)), prior to it plateauing (Pritchard *et al.* 2000). A more objective approach for estimating the best value of K, estimating the maximum rate of change in the log probability of the data between consecutive values of K (ΔK), was also used to indicate the appropriate value of K (Evanno *et al.* 2005). Analyses 1 and 2 were run with the same parameters as in previous studies ((Senn & Pemberton 2009); Chapter 2), namely the standard model of admixed ancestry (with the parameter α inferred from the data, using a uniform prior) and the model of correlated allele frequency ($\lambda = 1$), a burnin of 5×10^4 and a run length of 10^6 Markov chain Monte Carlo steps. Null alleles can cause deviation from Hardy-Weinberg equilibrium by causing a systematic pattern of missing genotype data and can jeopardise rates of hybridisation observed (Falush *et al.* 2003; Senn 2009). The frequency of null alleles were therefore estimated concurrently by incorporating a row of "999" values into the second line of the data set and activating the option RECESSIVE ALLELES = 1. This function enables Structure to 'suspect' particular alleles as null alleles if, for example, they exhibit allele-specific PCR failure. It will then then treat these suspected null alleles as recessive instead of missing data and estimate their frequency at each and every locus (Falush *et al.* 2007; Senn 2009). Structure output data were manipulated using the software Distruct (Rosenberg *et al.* 2002), for illustrative purposes.

Analysis by Structure 2.3.3 generated a Q value for each individual, which represents the estimated proportion of ancestry to each of K groups. When simulations are run at K = 2 (as is typical for hybridisation between two taxa), the Q values for membership to one of the two ancestral populations can be used as an index of the hybrid status of an individual; here Q = 0 represents a sika and Q = 1, a red. Delimiting the proportion of admixture that qualifies as a hybrid is difficult, principally due to the possibility that at some loci there may be ancestral allele sharing in the taxa under consideration. Here a hybrid was defined on the basis of nuclear markers as an individual returning a Q value of $0.05 \leq Q \leq 0.95$, following previous practice (Senn & Pemberton 2009). Individuals outside these boundaries were defined as 'pure', although may still contain introgressed alleles beyond the detection limit of the markers. In analysis 2 a hybrid was also defined if the mtDNA haplotype was discordant with a 'pure' nuclear genotype (i.e. red mtDNA in an animal with $Q < 0.05$ or sika mtDNA in animals with $Q > 0.95$). This latter type of hybrid indicates introgression beyond the resolution of the nuclear markers.

The average number of alleles and genetic diversity indices for red and sika at each of the 22 microsatellite loci and within each population, respectively, were also calculated using CERVUS 3.0 (Kalinowski *et al.* 2007). The diversity indices for wapiti were previously calculated and shown in Table 2.2.

3.3.4 *To determine, if possible, the initial direction of hybridisation (objective 2)*

The direction of initial hybridisation events (i.e. which taxon was the female parent) can only be assessed from cytonuclear data in F1 hybrids. An F1 individual should have a Q close to 0.5 in a K=2 Structure analyses and it should be heterozygous for red and sika alleles at all loci. In order to determine whether we had sampled any F1 hybrids we examined the posterior allele frequencies for the parental taxa generated by Structure following analysis 2 and assigned these as red-specific, sika-specific or inconclusive, according to conservative criteria (see Appendix Table 3.A1). The genotypes of hybrids were recoded according to the origin of each allele at each locus to determine the proportion of loci that were red-sika heterozygous relative to all loci genotyped in that individual.

3.3.5 *To investigate whether any hybrids outwith Co. Wicklow derive from the Wicklow hybrids or new hybridisation events (objective 3)*

Since this study revealed a previously undocumented hybrid population in Co. Cork, we sought to determine the origin of the Japanese sika and red deer contributing to the Co. Cork population and specifically whether they could have been translocated as hybrids from Co. Wicklow. For sika, this was achieved by running analysis 3, a Structure analysis using all individuals that appeared to be pure Japanese sika in analysis 2, i.e. they returned a $Q < 0.05$. Structure run parameters were as described above. Since no apparently pure red deer ($Q > 0.95$ in analysis 2) were sampled in Co. Cork, the origin of the red alleles in the Co. Cork population cannot be determined in the same way. Instead, using the posterior allele frequencies from analysis 2 (Appendix Table 3.A1), we identified the red alleles found among Co. Cork hybrids and asked whether or not they occurred in other Irish deer populations, especially Co. Wicklow.

3.3.6 *To investigate the accuracy with which hybrids are identified from stalker-assigned phenotype (objective 4).*

To address the accuracy with which stalkers identified hybrids between red and Japanese sika the Q value derived from analysis 2 and mitochondrial haplotype for animals from Co. Wicklow and Co. Cork was compared with the phenotype assigned by the stalker when the deer was sampled.

3.4 Results

Across the 392 individual samples, genotypes were obtained for at least 20 of the 22 nuclear markers for 374 individuals and the mitochondrial haplotype was determined for all individuals (Table 3.1). Genetic diversity indices are given for each locus (Table 3.2) and within each population (Table 3.3), for red deer, sika and wapiti.

Table 3.1. Sample sizes, stalker-assigned phenotypes and genetic data set completeness for the 374 individuals successfully genotyped (at least 20 out of the 22 markers genotyped), shown for the eight counties sampled.

County	Total number of individuals	No. of phenotypic red	No. of phenotypic sika	No. of phenotypic hybrid animals	No. animals not assigned phenotype	Nuclear dataset (% complete)	MtDNA dataset (% complete)
Donegal	13	12	0	1	0	95.45	100
Tyrone	2	0	2	0	0	100.00	100
Sligo	4	4	0	0	0	95.45	100
Mayo	16	15	1	0	0	95.45	100
Galway	12	12	0	0	0	96.59	100
Kerry	114	37	74	0	3	98.88	100
Cork	15	0	10	3	2	99.70	100
Wicklow	198	13	123	49	13	98.90	100

Table 3.2. Genetic diversity indices for each of the 22 loci in our microsatellite marker panel in phenotypic red deer (n = 93) and sika (n = 209) calculated in Cervus 3.0. Subscripts $_{r,\ s}$ represent parameters calculated in red and sika datasets independently. Parameters are k, the number of alleles at each locus in each species, N, number of samples typed at each locus, H_o, observed heterozygosity, H_e, expected heterozygosity and Null, the frequency of null alleles at each locus, after Table 3 in Senn & Pemberton (2009). Diversity indices for wapiti are given in Table 2.2.

Locus	k_r	N_r	$H_{O\,r}$	$H_{E\,r}$	$Null_r$	k_s	N_s	$H_{O\,s}$	$H_{E\,s}$	$Null_s$
AGLA293	3	90	0.189	0.28	0.1852	3	209	0.043	0.052	0.0852
BM4006	4	93	0.462	0.504	0.0264	4	209	0.062	0.079	0.1631
BM6438	6	88	0.159	0.303	0.3221	5	196	0.408	0.478	0.0756
BM757	11	91	0.659	0.711	0.0405	5	209	0.258	0.256	-0.0084
BOVIRBP	8	93	0.645	0.704	0.0498	4	209	0.033	0.038	0.1078
FCB193	14	89	0.64	0.797	0.1221	9	209	0.306	0.377	0.1232
FSHB	18	91	0.736	0.891	0.0954	14	208	0.337	0.364	0.0613
IDVGA29	4	87	0.437	0.519	0.0814	5	202	0.277	0.36	0.1229
IDVGA55	9	93	0.484	0.712	0.2073	8	209	0.373	0.414	0.0365
INRA5	3	92	0.065	0.104	0.2185	3	208	0.269	0.342	0.1166
INRA6	5	92	0.359	0.453	0.1295	3	209	0.038	0.052	0.2321
INRA131	6	93	0.505	0.55	0.0579	5	209	0.278	0.282	-0.0009
MM012	4	93	0.226	0.267	0.1119	3	209	0.057	0.079	0.2199
RM12	11	93	0.677	0.814	0.0861	5	209	0.038	0.052	0.2327
RM188	15	92	0.783	0.866	0.0494	11	208	0.596	0.681	0.0626
RM95	9	93	0.699	0.82	0.0786	6	208	0.361	0.371	0.0105
RME025	5	93	0.28	0.545	0.3335	5	207	0.266	0.304	0.0547
TGLA40	7	92	0.5	0.749	0.1975	5	208	0.221	0.493	0.3846
TGLA126	3	93	0.086	0.152	0.3427	4	207	0.469	0.513	0.042
TGLA127	12	91	0.725	0.822	0.0562	7	209	0.416	0.491	0.0826
TGLA337	7	68	0.662	0.745	0.0557	8	200	0.65	0.768	0.0824
UWCA47	4	91	0.143	0.176	0.0898	3	206	0.175	0.215	0.1003

Table 3.3. Genetic diversity indices within each population for phenotypic red and sika calculated in Cervus 3.0. Parameters H_o and H_e represent observed heterozygosity and H_e, expected heterozygosity respectively. Diversity indices for wapiti are given in Table

Species	*Population*	*Sample Size*	*Mean No. alleles per locus*	H_E	H_O
Red	Co. Donegal	12	3.91	0.5425	0.470
	Co. Sligo	4	2.68	0.4984	0.489
	Co. Mayo	15	4.86	0.5435	0.516
	Co. Galway	12	3.86	0.5065	0.499
	Co. Kerry	37	3.27	0.388	0.347
	Co. Wicklow	13	4.77	0.6667	0.653
Sika	Co. Tyrone	2	1.18	0.0985	0.136
	Co. Mayo	1	1.05	0.0909	0.091
	Co. Kerry	73	1.91	0.1642	0.154
	Co. Cork	10	2.36	0.2694	0.218
	Co. Wicklow	123	5.18	0.3707	0.346

3.4.1. *To assess the current extent of hybridisation and introgression between* Cervus *deer on the island of Ireland and whether it threatens either parental taxon (objective 1)*

Analysis 1: Red, Japanese sika and wapiti genotypes (n = 523)

The log likelihoods calculated in Structure revealed K = 2 was the smallest number of genetic clusters that was optimal to describe the population structure, with an average Ln Pr (X|K) (natural logarithm of the probability of data X, conditional on K) of -23028.94 (s.d., 5.57) and a rate of change of 2232.68 (Figure 3.2). At this value of K, as might be predicted from the choice of markers, red and Japanese sika were differentiated, but not wapiti, which clustered with red (see Appendix Figure 3.A1). The next most likely structure was K = 3 with a likelihood of -21009.2 (s.d. 795.49) and a rate of change of 2.54 and at this K wapiti were differentiated from red and Japanese sika (Figure 3.3). Allele frequencies for the three taxa generated at K = 3 are shown in Appendix Table 3.A2. There is some support from the likelihoods for a larger number of populations (K) but for our purposes K = 3 is appropriate since we are interested in hybridising taxa. The variation in the log likelihood generated during replicated simulations at the same value of K may be attributed to slight variation in the sampling (or "mixing") of the Markov chain, as part of the Bayesian analysis, when converging on the posterior distribution of each of the required parameters (Pritchard *et al.* 2000).

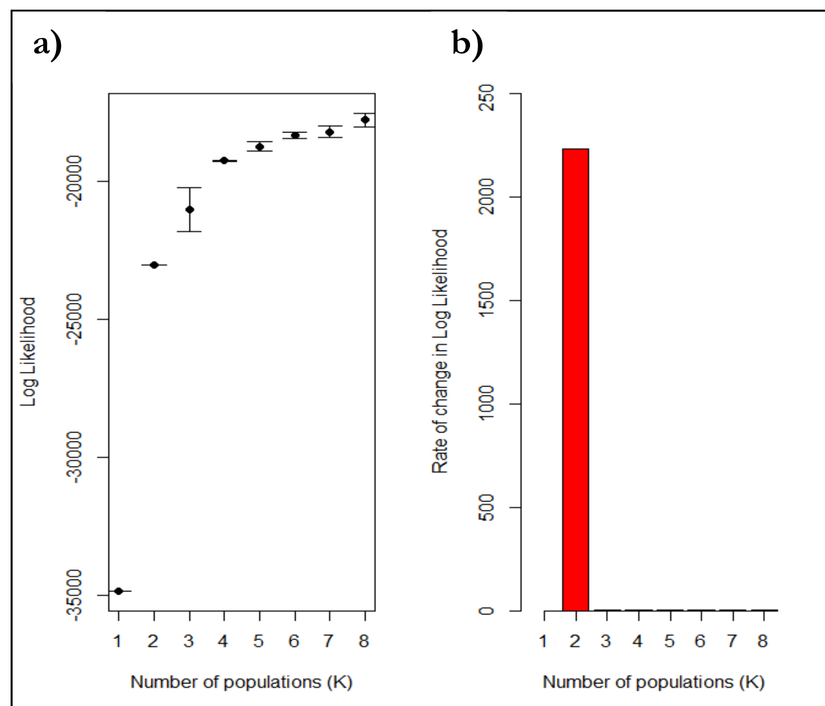


Figure 3.2. Assessment of the most likely number of populations using Structure 2.3.3 Results of Analysis 1 which a) shows the log-likelihood (with standard error) of the value of K (number of populations) given the dataset and b) the rate of change in log likelihood between values of K. Both provide evidence that K = 2 are the most likely.

The analysis including wapiti revealed little evidence for introgression of Irish deer populations by wapiti (Figures 3.3 and 3.4 and Table 3.4). Using the criterion of Q > 0.05 membership to wapiti, in total one “red-like hybrid with recent wapiti introgression” individual was sampled from Co. Mayo (Category 5, Table 3.4) and one “red-like individual with recent sika and recent wapiti ancestry” from Co. Wicklow (Category 6, Table 3.4). Given that 374 Irish individuals were studied, this suggests a very low rate of introgression (0.53%). A single individual amongst the red control animals from central Scotland (RAL09) also showed wapiti introgression.

On the other hand this analysis revealed a spectrum of red-sika hybrids in Co. Wicklow and Co. Cork (Figures 3.3 and 3.4 and Table 3.4). Since it is possible that the inclusion of wapiti genotypes could confound the analysis of red-sika hybridisation, in analysis 2 we repeated the analysis after removing the 49 wapiti control samples, RAL09 (the control Scottish red deer with wapiti introgression) and the two Irish deer with evidence of wapiti introgression.

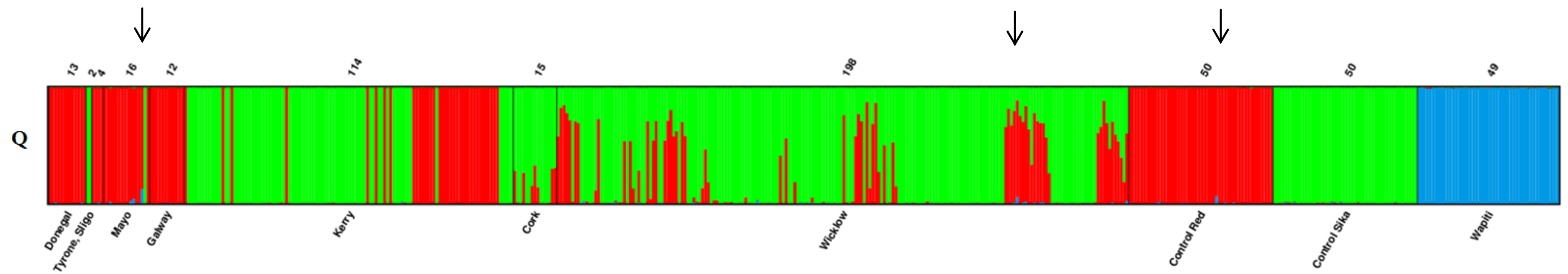
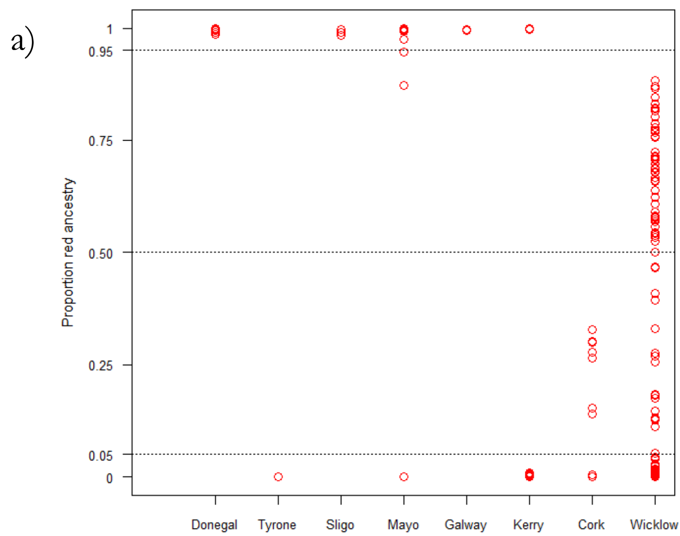
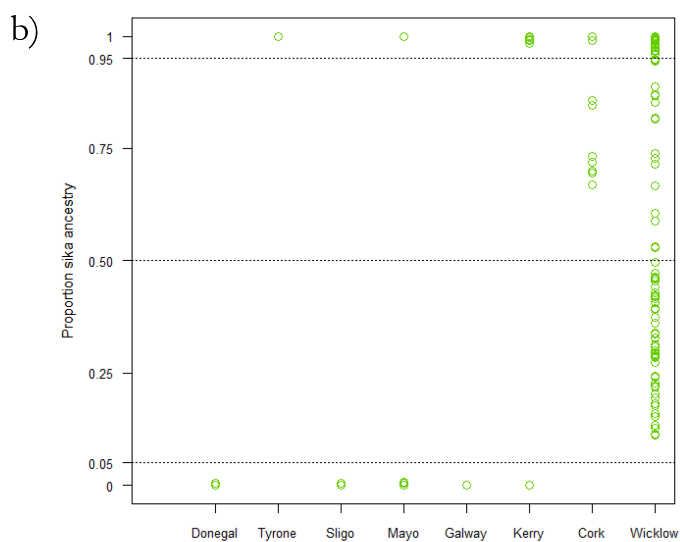


Figure 3.3. Bar chart showing the results of *analysis 1* in STRUCTURE at K = 3 for each individual in the dataset consisting of red deer, sika and wapiti animals (n = 523). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (shown in red), the proportion attributable to Japanese sika ancestry (green) and that attributed to wapiti ancestry (blue). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis. Counties are plotted in an approximately North West to South East order. Arrows indicate the three deer with wapiti introgression at ≥ 5%.

Table 3.4. Admixture classification of all individuals in analysis 1, based on Q values from Structure 2.3.3 with K = 3 following a classification approach expanded from Senn & Pemberton (2009).

Category	Estimated membership to red	Estimated membership to sika	Estimated membership to wapiti	Category	County								Total
					No. (& %) of animals from Donegal	No. (& %) of animals from Tyrone	No. (& %) of animals from Sligo	No. (& %) of animals from Mayo	No. (& %) of animals from Galway	No. (& %) of animals from Kerry	No. (& %) of animals from Cork	No. (& %) of animals from Wicklow	No. (& %) of animals from all sites
1	$0.90 < Q \leq 1$	$0 \leq Q < 0.05$	$0 \leq Q < 0.05$	'pure' red ancestry	13 (100)	0 (0)	4 (100)	14 (87.5)	12 (100)	37 (32.5)	0 (0)	0 (0)	80 (21.4)
2	$0 \leq Q < 0.05$	$0.90 < Q \leq 1$	$0 \leq Q < 0.05$	'pure' sika ancestry	0 (0)	2 (100)	0 (0)	1 (6.25)	0 (0)	77 (67.5)	8 (53.3)	129 (65.1)	217 (58.0)
3	$0 \leq Q < 0.05$	$0 \leq Q < 0.05$	$0.90 < Q \leq 1$	'pure' wapiti ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
4	$0.50 \leq Q < 0.95$	$0.05 \leq Q < 0.50$	$0 \leq Q < 0.05$	red-like hybrid with recent sika ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	51 (25.8)	51 (13.6)
5	$0.50 \leq Q < 0.95$	$0 \leq Q < 0.05$	$0.05 \leq Q \leq 0.50$	red-like hybrid with recent wapiti ancestry	0 (0)	0 (0)	0 (0)	1 (6.25)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.3)
6	$0.50 \leq Q < 0.95$	$0.05 \leq Q \leq 0.95$	$0.05 \leq Q \leq 0.95$	red-like hybrid with recent sika and recent wapiti ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (0.50)	1 (0.3)
7	$0.05 \leq Q < 0.50$	$0.50 \leq Q < 0.95$	$0 \leq Q < 0.05$	sika-like hybrid with recent red ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	7 (46.6)	17 (8.6)	24 (6.4)
8	$0 \leq Q < 0.05$	$0.50 \leq Q < 0.95$	$0.05 \leq Q < 0.50$	sika-like hybrid with recent wapiti ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
9	$0.05 \leq Q < 0.95$	$0.50 \leq Q < 0.95$	$0.05 \leq Q \leq 0.95$	sika-like hybrid with recent red and recent wapiti ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
10	$0.05 \leq Q < 0.50$	$0 \leq Q < 0.05$	$0.50 \leq Q < 0.95$	wapiti-like hybrid with recent red ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
11	$0 \leq Q < 0.05$	$0.05 \leq Q < 0.50$	$0.50 \leq Q < 0.95$	wapiti-like hybrid with recent sika ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
12	$0.05 \leq Q < 0.95$	$0.05 \leq Q < 0.95$	$0.50 \leq Q < 0.95$	wapiti-like hybrid with recent red and recent sika ancestry	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)





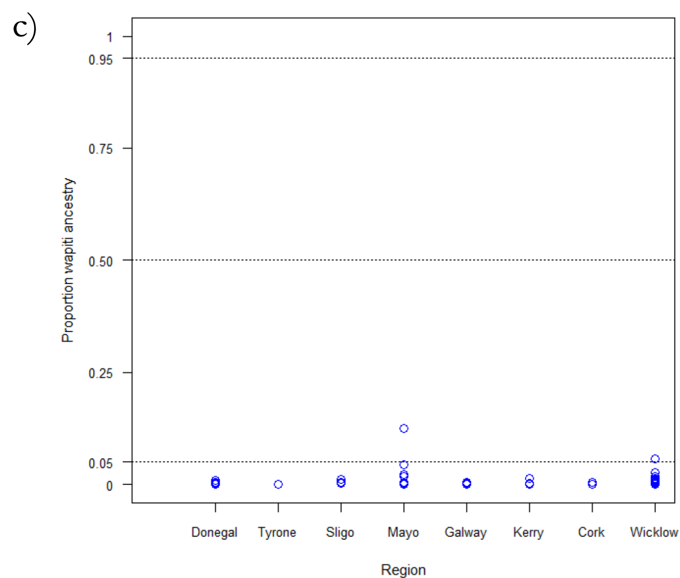


Figure 3.4. The proportion of inferred a) red, b) sika and c) wapiti ancestry determined from Structure analysis 1 when K = 3, plotted for the eight counties sampled and within the wapiti control samples.

Analysis 2: Red and Japanese sika genotypes (n = 471)

As might be expected from analysis 2, the log likelihoods calculated in Structure showed K = 2 was the smallest number of genetic clusters that was optimal to describe the population structure, with an average Ln Pr (X|K) of -18585.8 (s.d. 3.25) and a rate of change of 5264.3 (Figure 3.5). Allele frequencies for the population clusters at K = 2 are shown in Appendix Table 3.A1.

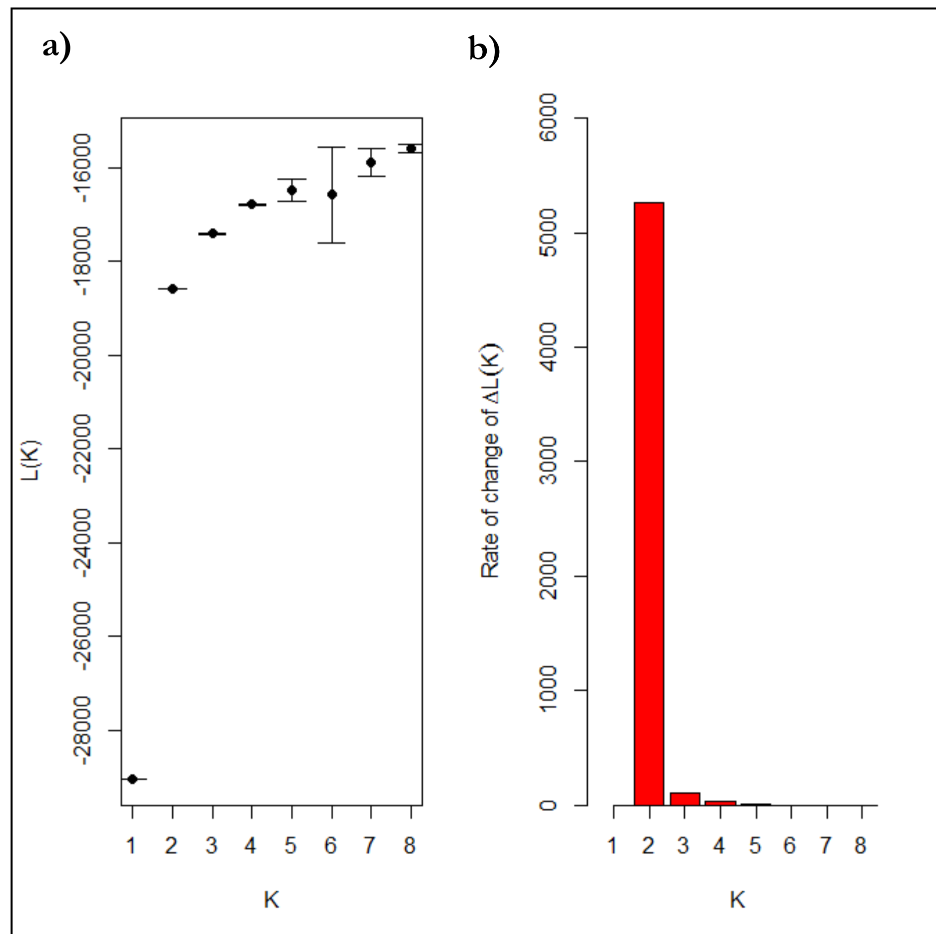


Figure 3.5. Assessment of the most likely number of populations using Structure 2.3.3 analysis 2 of which a) shows the log likelihood (with error bars) of the value of K (number of populations) given the dataset and b) gives the rate of change in log likelihood between values of K. Both provide evidence that K = 2 are the most likely.

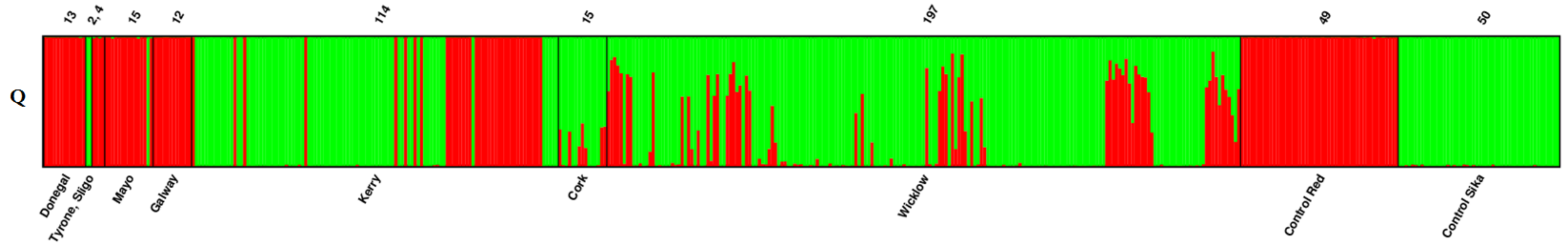


Figure 3.6. Bar chart showing the results of *analysis 2* in STRUCTURE at K = 2 for each individual in the dataset consisting of red deer and sika animals only (n = 471). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (shown in red) and the proportion attributable to sika ancestry (green). A hybrid is defined as an animal with membership ancestry of $0.05 \leq Q \leq 0.95$ to both red and sika. Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis. Counties are plotted in an approximately north west to south-east order.

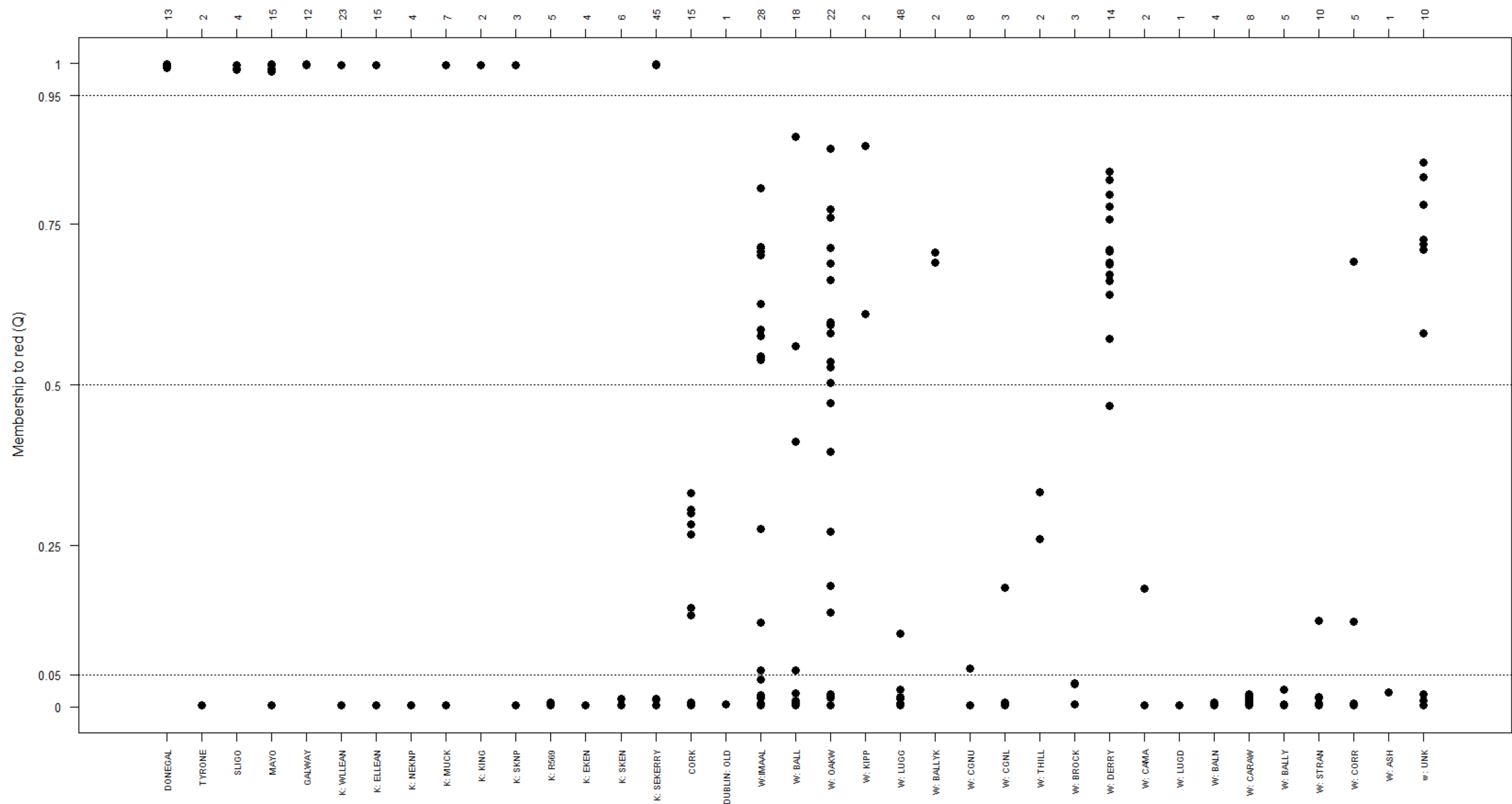


Figure 3.7. The membership to red (Q), as calculated by Structure 2.3.3 analysis 2 against the site from which the individual was obtained. Abbreviated population codes are as follows: "K:" relates to Co. Kerry sites for which WLLEAN= West Lough Leane, ELLEAN= East Lough Leane, NEKNP= North East region of Killarney National Park, MUCK= South Muckross Lake, KING= Kingsboro, Upper Lake, SKNP= Southern border of the Killarney National Park, R569= Inside the R560 road, EKEN= East Kenmare, SKEN= South Kenmare, SEKERRY= South East Kerry, CORK = Co. Cork, then "W:" refers to Co. Wicklow sites for which IMAAL= Glen of Imaal, BALL= Ballinagee, OAKW= Oakwood, KIPP= Kippure, LUGG= Luggala, BALLYK= Ballyknockan, CGNU= Carrigeenduff Upper, CGNL= Carrigeenduff Lower, THILL= Turlough Hill, BROCK= Brockagh, DERRY= Derrybawn, CAMA= Camaderry, LUD= Lugduff, BALN= Ballinacor, CARAW= Carawaystick, BALLY= Ballyward, STRAN= Stranahely, CORR= Corragh, ASH= Ashford and UNK= Co. Wicklow, unknown location.

Analysis 2 supports the results of analysis 1 in showing that a substantial proportion of deer sampled in Co. Wicklow and Co. Cork are introgressed hybrids (with very similar individual estimates of Q to those estimated in analysis 1) while remaining individuals sampled in these counties were 'pure' sika (Figures 3.6 and 3.7). In contrast, no individuals sampled from the North West or Co. Kerry were hybrid. Across all Irish samples, a total of 215 'pure' sika, 80 'pure' red and 77 hybrids were sampled based on their nuclear genotype. Of the hybrid animals, 91% were from Co. Wicklow and 9% from Co. Cork. Results from each of the three main sampling areas will now be described in more detail.

Across the five counties sampled in the North and West, genetic analysis indicated that we sampled 43 'pure' red and three 'pure' sika individuals (Figures 3.6, 3.7 and 3.8a). These putatively pure red were sampled from Co. Donegal (n = 13), Co. Sligo (n = 4), Co. Mayo (n = 14) and Co. Galway (n = 12) and putatively pure sika were sampled from Co. Tyrone (n = 2) and Co. Mayo (n = 1). Since we found no red-sika hybrid animals, there is no genetic evidence of hybridisation in this region; however, since sample sizes per site were generally very low, this is a tentative inference.

In Co. Wicklow, we sampled 127 'pure' sika individuals from 16 of the 20 sites and 70 hybrid individuals from 13 of the sites, (Figures 3.6, 3.7, 3.8b). No 'pure' red deer were sampled from this region. Whilst hybridisation appears to be extensive at particular sites within Co. Wicklow (e.g. 100% hybrids sampled from Kippure, Ballyknockan, Turlough Hill, Derrybawn), it is almost absent from others (e.g. Luggala; Figure 3.8b). Among the hybrids, there were over twice as many genetically red-like individuals (n = 51; $0.5<Q\leq0.95$) compared to sika-like individuals (n = 19; $0.05\leq Q<0.5$).

In Co. Kerry, genetic analysis indicates that we sampled 77 'pure' sika individuals and 37 'pure' red individuals (Figures 3.6, 3.7, 3.8c). The putatively pure sika animals were sampled from nine of the ten sample sites (all except Kingsboro, from which only two samples were obtained), while the putatively pure red were sampled from six sites. However, in neighbouring Co. Cork, seven of the 15 individuals sampled were hybrid based on genetic analysis, while the remainder were putatively pure sika (Figures 3.6, 3.7, 3.8c).

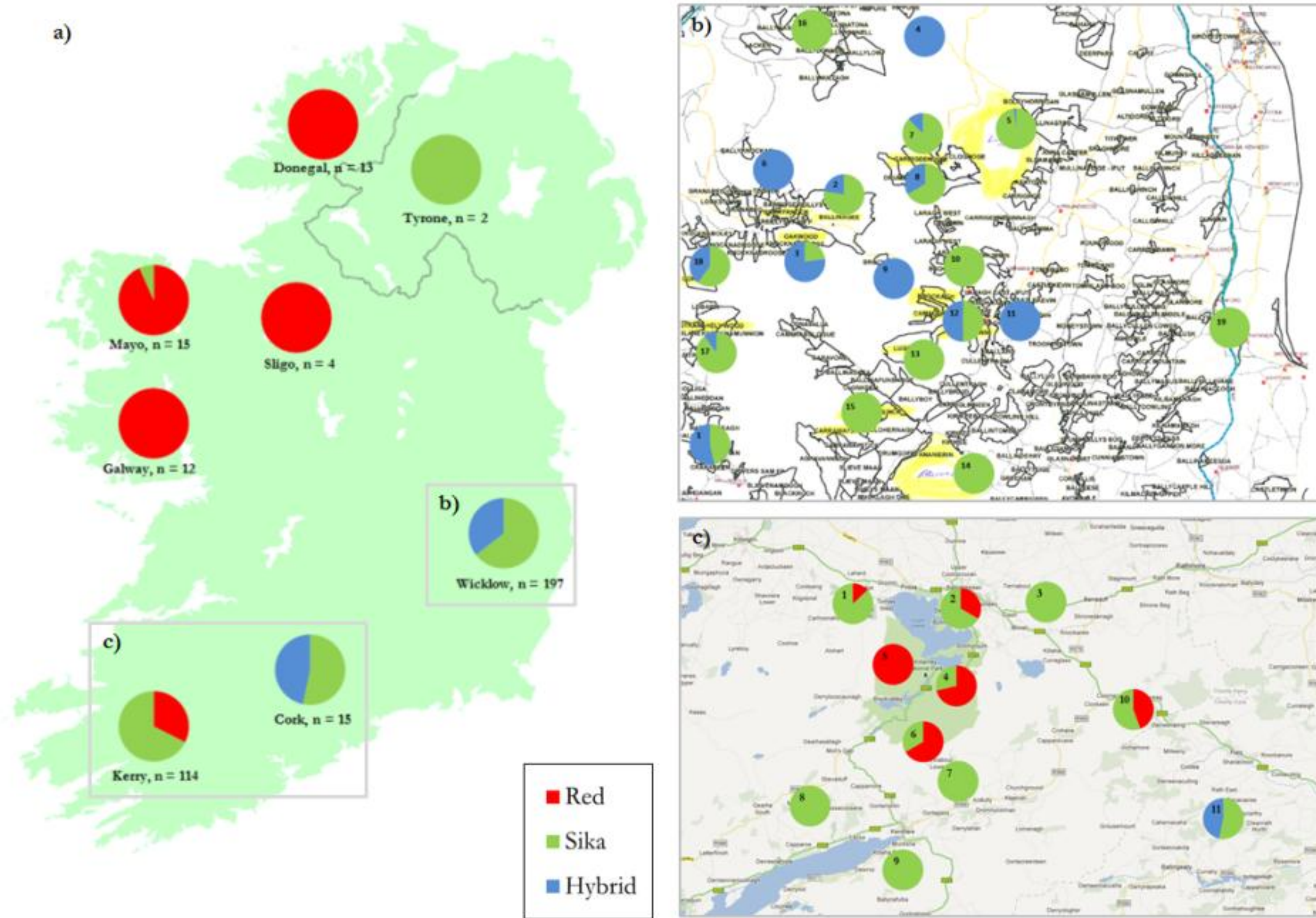


Figure 3.8. a) Overview of proportion of pure red, pure sika and intermediate animals based on their nuclear genotype around Ireland and in further detail from b) Co. Wicklow, where 1= Glen of Imaal, 2= Ballinagee, 3= Oakwood, 4= Kippure, 5= Luggala, 6= Ballyknockan, 7= Carrigeenduff Upper, 8= Carrigeenduff Lower, 9= Turlough Hill, 10 = Brockagh, 11= Derrybawn, 12= Camaderry, 13= Lugduff, 14= Ballinacor, 15= Carawaystick, 16= Ballyward, 17= Stranahely, 18= Corragh, 19= Ashford (background markings in yellow are irrelevant) and from c) Co. Kerry and Co. Cork, where 1= West Lough Leane, 2= East Lough Leane, 3= North East region of Killarney National Park, 4= South Muckross Lake, 5= Kingsboro, Upper Lake 6= Southern border of the Killarney National Park, 7= Inside the R560 road, 8= East Kenmare, 9= South Kenmare, 10= South East Kerry, 11= Cork.

Mitochondrial DNA analysis added further resolution to the Structure analyses. First, it is important to note that the mitochondrial marker is not diagnostic for wapiti and in this study there is no evidence of any Manchurian sika haplotypes in the Irish samples. No cytonuclear disequilibria were noted in counties of the North West or Co. Kerry. However, among the red-sika hybrids in Co. Wicklow and Co. Cork, the red mitochondrial haplotype predominates (Figure 3.9). None of the 131 deer with a Q>0.5 (i.e. red-like) carried a Japanese sika mtDNA haplotype but 31/214 (14.5%) individuals with a Q<0.5 (i.e. sika-like) carried a red deer haplotype. Amongst the 31 sika-like individuals with Q<0.5 and a red haplotype, ten were animals considered 'pure' sika (Q < 0.05) from their nuclear markers (Figure 3.9) and these were all sampled from Co. Wicklow. The inclusion of these mitochondrial hybrids increased the total number of hybrids found, based on either nuclear genotype or mitochondrial haplotype, by 13%.

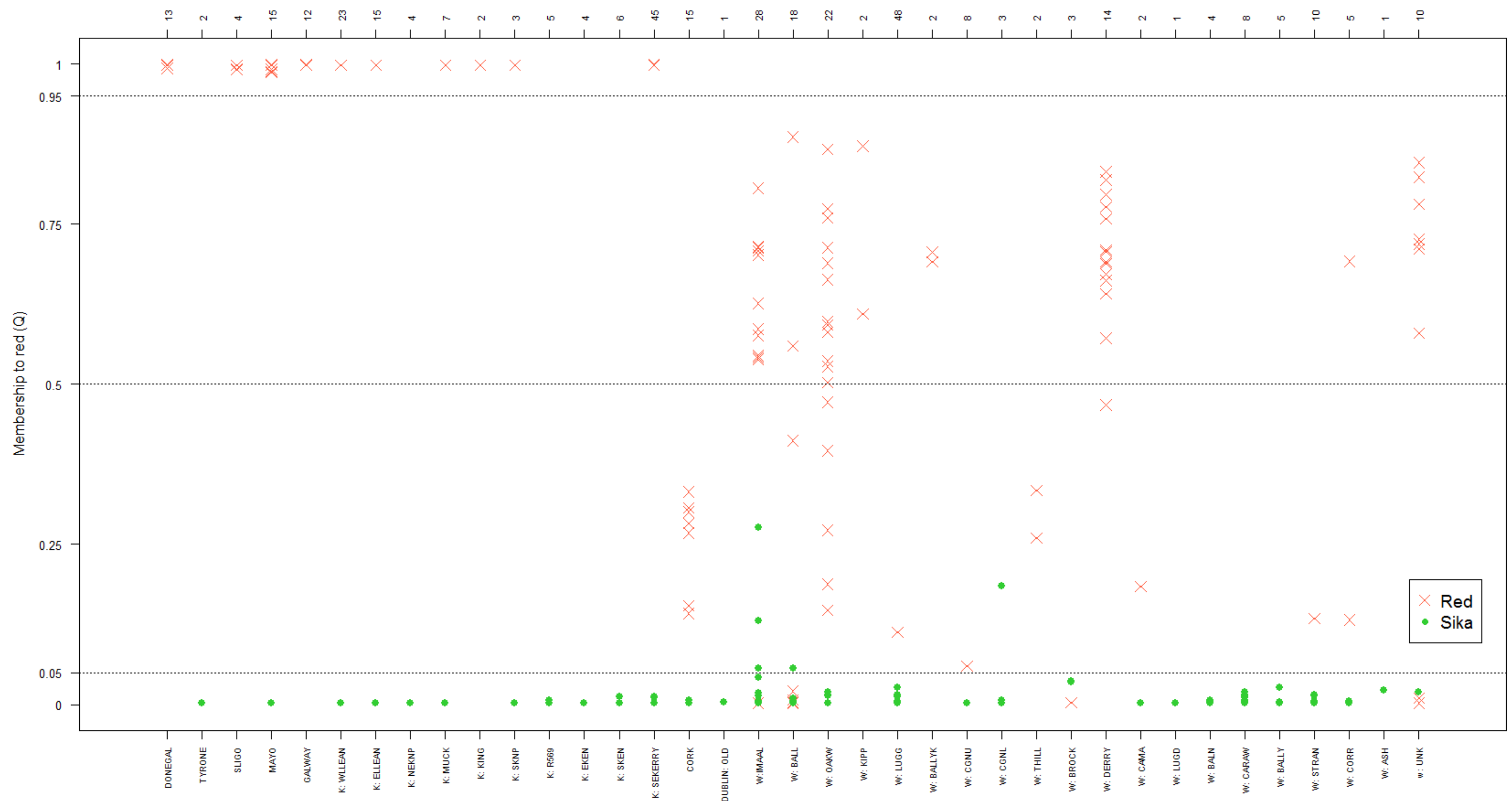


Figure 3.9. The membership to red (Q), as calculated by Structure 2.3.3 plotted against the site from which the individual was obtained. The mtDNA haplotype the individual carried is also indicated. Abbreviated population codes are as follows: "K:" relates to Co. Kerry sites for which WLLEAN= West Lough Leane, ELLEAN= East Lough Leane, NEKNP= North East region of Killarney National Park, MUCK= South Muckross Lake, KING= Kingsboro, Upper Lake, SKNP= Southern border of the Killarney National Park, R569= Inside the R560 road, EKEN= East Kenmare, SKEN= South Kenmare, SEKERRY= South East Kerry, CORK = Co. Cork, then "W" refers to Co. Wicklow sites for which IMMAL= Glen of Imaal, BALL= Ballinagee, OAKW= Oakwood, KIPP= Kippure, LUGG= Luggala, BALLYK= Ballyknockan, CGNU= Carrigeenduff Upper, CGNL= Carrigeenduff Lower, THILL= Turlough Hill, BROCK= Brockagh, DERRY= Derrybawn, CAMA= Camaderry, LUD= Lugduff, BALN= Ballinacor, CARAW= Carawaystick, BALLY= Ballyward, STRAN= Stranahely, CORR= Corragh, ASH= Ashford and UNK= Co. Wicklow, unknown location.

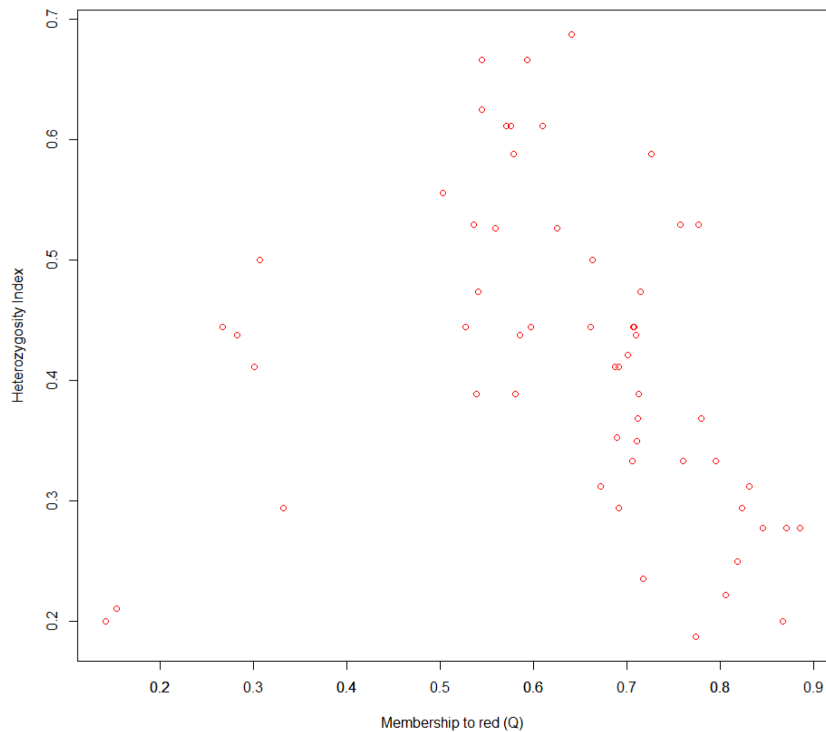


Figure 3.10. A heterozygosity index (calculated as the number of loci in an individual's genotype which are heterozygous for red and sika alleles, divided by the total loci scored) plotted against the membership to red (Q), for all red-sika hybrids at K=2. Since no individuals have Q≈0.5 and heterozygosity index ≈1, we sampled no F1 hybrids.

3.4.2 *To determine, if possible, the initial direction of hybridisation (objective 2)*

No F1 individuals were detected, since no individuals were even close to 100% heterozygous for red and sika alleles (Figure 3.10). This observation is consistent with the lack of 'pure' red samples found in either Co. Wicklow or Co. Cork (above).

3.4.3 *To investigate whether any hybrids outwith Co. Wicklow derive from the Wicklow hybrids or new hybridisation events (objective 3)*

In Structure analysis 3 of 'pure' Japanese sika samples from Ireland, two genetically distinct populations were identified: Co. Wicklow sika and a small number of sika sampled from the North West clustered separately from the Co. Kerry sika (Appendix Figure 3.A3). The Co. Cork sika clustered with the Co. Kerry sika, suggesting that they derived from the long-standing Co. Kerry population. This suggests that the Co. Cork hybrids may be derived independently of the Co. Wicklow hybrids.

Two lines of evidence support the idea that the red deer that founded the Co. Cork hybrids were also not the same as the red deer contributing to the Co. Wicklow hybrids. First, up to six alleles per locus were identified as having introgressed from red into sika-like hybrids in the Co. Wicklow samples, whilst only a maximum of two alleles had introgressed from red into the hybrids sampled in Co. Cork (Appendix Figure 3.A2). This suggests there were multiple ancestral red deer for the Co. Wicklow hybrids, but perhaps only one for the Co. Cork hybrids, although it would also be consistent with an introduction of a small founder population (just 1-2 animals carrying red alleles) from Co. Wicklow to Co. Cork.

Second, at four of the 22 loci, the Co. Cork hybrids had a single private red deer allele which was not present in either the Co. Kerry red deer or the Co. Wicklow hybrids (Figure 3.11). Whilst two of these were found in red deer from the North West counties sampled, the remaining two were absent from all other red deer sampled from Ireland in this project but have been observed in Scottish red deer (S. Smith, pers. obs). Both the sika and the red deer genetic evidence, therefore, suggest that the Co. Cork hybrids have arisen from hybridisation that was independent from Co. Wicklow.

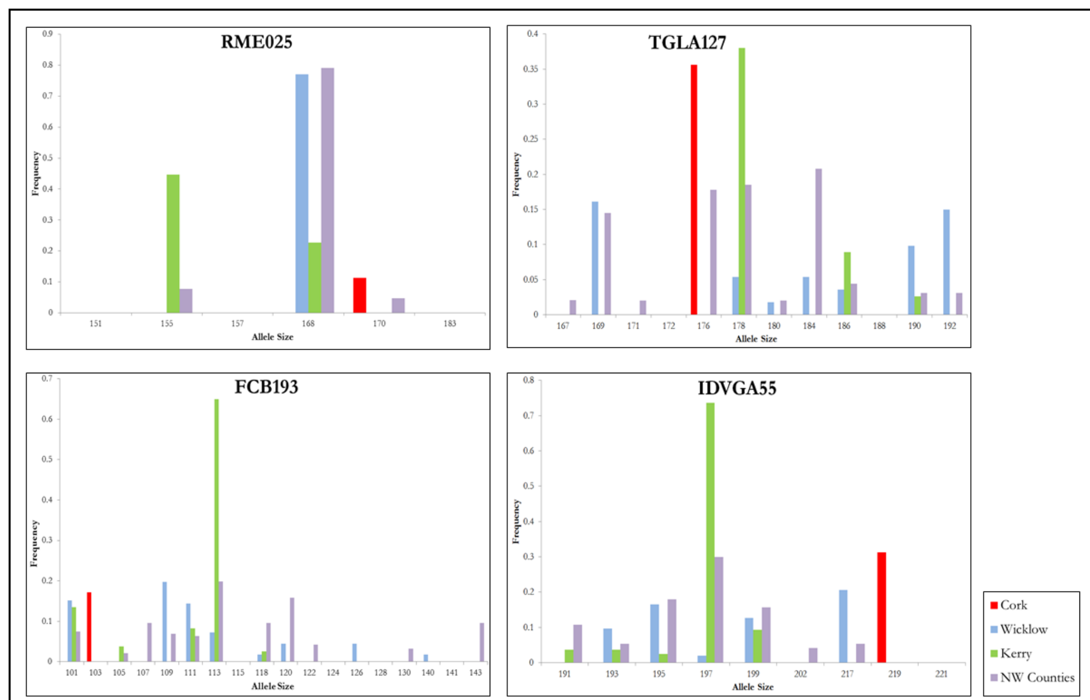


Figure 3.11. The frequency of red deer alleles at each of four loci in animals sampled from Co. Wicklow (blue), Co. Kerry (red) and Co. Cork (green). At these four loci Co. Cork hybrids are fixed for a private allele that is not found in the Co. Wicklow hybrids or the Co. Kerry red deer. This indicates that a genetically distinct red deer population was associated with the origin of the Co. Cork hybrids. Note the scale of the y-axis is different for each locus to highlight the highest relative frequencies.

3.4.4 *To investigate the accuracy with which hybrids are identified from stalker-assigned phenotype (objective 4)*

Amongst all deer assigned a phenotype from Co. Wicklow and Co. Cork, 79% were identified correctly according to their Q value based on nuclear genotype and mitochondrial haplotype. Of the 42 (21%) of animals that were misidentified, 25 were identified as sika but were actually hybrids (including 9 that were mitochondrial hybrids) and 4 were identified as hybrids but genotyped as 'pure' sika (Figure 3.12). Of the thirteen animals which were identified as red deer, twelve were actually hybrid and one was a pure sika; with no 'pure' red deer sampled at these sites, all deer identified as such were incorrect.

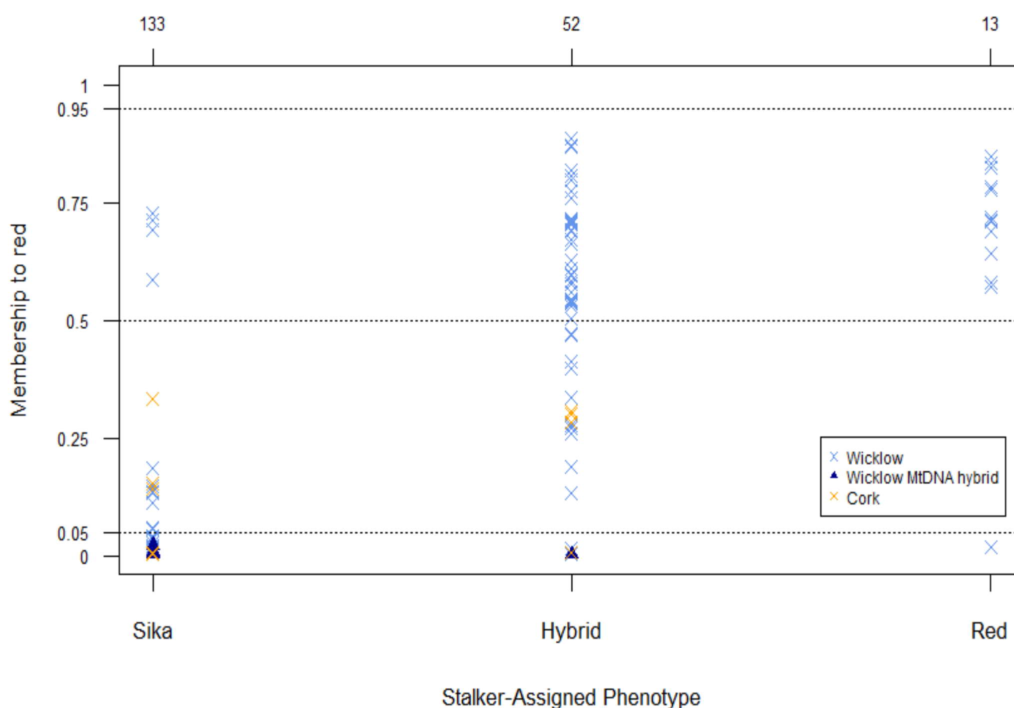


Figure 3.12. The estimated proportion of ancestry (Q) for all animals from Co. Wicklow and Co. Cork (n = 198) plotted against their stalker-assigned phenotype. Mitochondrial hybrids, which are beyond the detection of the nuclear markers, were only found in Wicklow and are represented by triangles.

3.5 Discussion

3.5.1 *To assess the current extent of hybridisation and introgression between* Cervus *deer on the island of Ireland and whether it threatens either parental taxon (objective 1)*

Hybridisation and introgression between Irish red deer and the exotic North American wapiti, subsequent to its introduction over 100 years ago, appears to be negligible. The single individual showing wapiti introgression in Co. Mayo could be attributable to shared ancestral polymorphism between the two species rather than evidence for hybridisation, or the ability of this marker panel, designed to distinguish between red deer and Japanese sika, to distinguish between red deer and wapiti. As the sample sizes from some sites are small, further individuals should be sampled from the North West to test this. However, even following the introduction of two wapiti males and a single female to Powerscourt estate in 1859, Whitehead (1964) regarded it "extremely unlikely" that wapiti material would have persisted this long, and a recent study found a similar negligible genetic impact of wapiti on red deer in Scotland, using a diagnostic Y chromosome marker (Pérez-Espona *et al.* 2010a). Finally, with the only tool available to us, the mtDNA, we found no evidence for Manchurian sika haplotypes. Overall, the ability of the marker panel to distinguish effectively between species was corroborated by the Structure analysis which found three population clusters (one for each species) the most likely underlying genetic structure.

Since the introduction of the second exotic species studied, Japanese sika deer, in 1860, the genetic consequences for Irish red deer have been far greater. Almost 41% of the deer sampled from Co. Wicklow in this study were hybrid based on either their nuclear genotype or mitochondrial haplotype, whilst 47% of those sampled from Co. Cork were also hybrids. On the other hand, there was no evidence for nuclear or mitochondrial introgression from Japanese sika into red deer in samples obtained from the North West and Co. Kerry. These results will now be discussed in more detail for each region.

There was no detectable nuclear or mitochondrial introgression amongst Co. Kerry deer. Using eight nuclear markers, McDevitt *et al.* (2009) assigned Co. Kerry red deer to their own genetic cluster and concluded mtDNA nucleotide and haplotype diversity in this region was up to ten times lower than in other parts of Ireland. Recent research has shown that Co. Kerry reds are likely to be descended from ancestors introduced during

the Irish Neolithic, since when they have experienced bottleneck events and currently exhibit levels of nuclear diversity similar to that amongst threatened deer populations in Tunisia (Carden *et al.* 2012; Hajji *et al.* 2007; McDevitt *et al.* 2009a). Therefore, their longstanding isolation, restricted genetic diversity and the process of genetic drift may have caused the Co. Kerry reds to diverge from other red deer populations to the extent that they have become less genetically and phenotypically compatible with the sika they are now in sympatry with, compared to those that resided in Co. Wicklow. This process may also be paralleled in the Co. Kerry sika; only three sika animals were initially translocated to Co. Kerry in 1864, very soon after their introduction to Powerscourt. Contemporary sika deer in Co. Kerry tend to be physically smaller than in Co. Wicklow and, therefore, may be less compatible with the large Co. Kerry reds, reinforcing their assortative mating to date (McDevitt *et al.* 2009a; Whitehead 1964). Overall, while we found no hybrid animals in Co. Kerry itself, the integrity of Co. Kerry red and sika is threatened by hybrids present in Cork, around 20km away from the Kerry-Cork border, regarded as “no great distance for a travelling stag” (Whitehead 1964). The dispersal of these hybrid animals into Co. Kerry and successful reproduction with deer there could reduce the prevalent interspecific dimorphism and disrupt assortative mating.

It has long been established that Co. Wicklow contains a hybrid swarm; in this study over 37% of all individuals were nuclear hybrids, over 5% were mitochondrial hybrids (beyond the detection of the nuclear markers) and no ‘pure’ red were identified. The mitochondrial hybrids from Co. Wicklow (n = 10) were all putatively pure sika animals carrying the red mitochondrial haplotype. This result is concordant with the results of McDevitt *et al.* (2009) in Ireland and Senn & Pemberton (2009) in Kintyre, Scotland and suggests that at the time of sampling and in generations immediately preceding them, hybrid animals were either descended from a red hind-sika stag cross or there is a tendency for male hybrid animals to backcross with red deer. As in Chapter 2, this cytonuclear disequilibria could be driven by sex-biased dispersal and the fact a red hind and sika stag are a more compatible pairing in terms of size (Senn & Pemberton 2009).

The absence of ‘pure’ red deer from Co. Wicklow, a similar finding to McDevitt *et al.* (2009), suggests this taxon have been lost through hybridisation with sika. Originally there may have been relatively few red deer in Co. Wicklow; in Powerscourt Park, for example, Delap (1936) reports 60-65 red deer inhabitants amongst 500-600 sika deer, the latter of which had ‘over-run the entire park’. Whilst a much smaller sample of animals was obtained from Co. Cork, similarly no ‘pure’ red animals were obtained from

there. Together with the presence of substantial numbers of genetically 'pure' sika obtained from Co. Wicklow (64%) and Co. Cork (53%), this suggested the persistence of the red deer is threatened to a greater extent by sika than *vice versa*. The existence of such skewed species densities may act to facilitate further hybridisation, as the less-common species would have more opportunity to hybridise and become admixed with the more-common species. This is similar to the situation between the dwindling European wolf populations (*Canis lupus*) which increasingly hybridise with domestic dogs (*Canis familiaris*) (Vilà *et al.* 2003).

In the North West, conclusions regarding the extent of red-sika hybridisation are tentative due to small sample sizes. Red deer resident here have a long history of intra-red admixture through various translocations and sika populations still reside in these counties (Carden *et al.* 2010; Whitehead 1964). Whilst our data shows that the samples from Co. Donegal, Co. Tyrone, Co. Sligo, Co. Mayo and Co. Galway were 'pure' sika and 'pure' red, these counties should continue to be monitored for hybridisation.

3.5.2 *To determine, if possible, the initial direction of hybridisation (objective 2)*

No F1 individuals were detected in our dataset in that they fulfilled the criteria of $Q \approx 0.5$ and a genotype heterozygous for red and sika at all loci. The absence of F1 hybrid animals amongst our samples from Co. Wicklow is perhaps unsurprising given the duration of sympatry between red and sika deer and the disappearance of putatively pure red deer from this heavily-hybridised area (above). The absence of F1s in Co. Cork similarly suggests that the hybridisation event leading to this population of hybrids may have been some generations ago or has been missed from our small sample size from this site (n=15). In the absence of F1 individuals, it is not possible to resolve the direction of the initial hybridisation event(s).

3.5.3 *To investigate whether any hybrids outwith Co. Wicklow derive from the Wicklow hybrids or new hybridisation events (objective 3)*

The sika from Co. Cork clustered with those from Co. Kerry in a Structure analysis based on all 'pure' sika sampled across Ireland, which is consistent with the a historical record of the establishment of sika from Co. Kerry to sites in and around West Cork (Whitehead 1964). This suggests that, given the correct combination of circumstances (as evidently manifested in Co. Wicklow), the sika in Co. Kerry may be more susceptible to hybridisation than previously thought (see above).

The absence of 'pure' red deer from Co. Wicklow or Co. Cork prevented us from including them in a parallel analysis of red deer population structure and forced us to look at the red alleles present in the hybrids in these two counties and other Irish populations in order to infer the genetic affinity of the red population involved in the Co. Cork hybrids. The presence of four private red alleles in Co. Cork, at substantial frequencies, suggests an ancestral red population independent from either Co. Wicklow or Co. Kerry was involved in the hybridisation events leading to the Co. Cork hybrids. Red deer were introduced to a deer park at Doneraile in 1895 some 20km away and still reside there alongside two to three other deer species (P. Sleeman, pers. comm). Despite their origin being unknown, red deer sampled in this vicinity were found to carry a single species of lice, *Damalinia concavifrons,* a parasite specific to mainland European red deer and North American wapiti and not found on deer sampled elsewhere in the country (P. Sleeman, pers. comm), suggesting the red deer in question may have been introduced from mainland Europe. Red deer farms in this region are also considered a likely source (T. Burkitt, pers. comm). Further investigation into these farms and comparison with European populations may shed light on the source of the reds involved in the hybrid activity in Co. Cork.

3.5.4 *To investigate the accuracy with which hybrids are identified from stalker-assigned phenotype (objective 4)*

In regions containing hybrid animals, stalkers accurately identified red, sika and hybrids in the field in 79% of cases, however misidentifications (21%) highlight the difficulty in identifying introgressed individuals based on phenotype in the field. This suggests that attempting to selectively cull hybrids will not be totally effective and introgressed animals are likely to escape undetected (see Chapter 5).

3.5.5 *To indicate where management actions may be required to protect putatively pure populations from hybridisation (objective 5)*

The Co. Kerry red deer population is of high conservation value and hitherto no hybridisation with Japanese sika appears to have taken place ((Carden *et al.* 2012; McDevitt *et al.* 2009a), this study). However, the hybrids detected in Co. Cork pose a serious threat to both the red and sika in Co. Kerry since they are nearby and presumably less likely to mate assortatively in the presence of red or sika. A management plan to selectively shoot phenotypic hybrids in Co. Cork is unlikely to be

totally effective (above). An alternative approach, therefore, may be to try and eliminate the deer population in Co. Cork entirely, in order to remove the threat they pose to the Co. Kerry red deer. However, both these options would involve a great investment of resources and labour and would not be simple tasks. They also both carry the risk that heavy culling in an area may displace survivors further afield, with the result that Co. Cork deer might disperse faster toward Co. Kerry than otherwise. A third option might be to try to maintain it by culling a deer-free zone between the two areas.

The situation in Co. Wicklow is advanced and, whilst there is no evidence that any putatively pure red deer remain, management could be directed toward conspicuous hybrids in the county in attempts to preserve and maximise the purity of the remaining Japanese sika at sites in the south, such as Lugduff, Ballinacor and Carawaystick and those resident in Luggala. Elsewhere, managers in the North West counties should remain vigilant as rare sika amongst the large populations of red is potentially conducive to hybridisation (Ratcliffe 1987). In situations where red deer are massively outnumbered, efforts could be focused on addressing the imbalance with sika numbers and, thereby, lowering their susceptibility to hybridisation (Vilà *et al.* 2003).

Not only management of deer populations, but management of the land could help ameliorate hybridisation. As landscape features have been shown to have significant impact on red deer gene flow in Scotland (Pérez-Espona *et al.* 2008), similarly, hybridisation patterns may be influenced by patterns of increasing forestry cover in Ireland (Carden *et al.* 2010). Sika prefer forest habitats, such as commercial conifer forestry and their expansion can parallel that of its planting (Pérez-Espona *et al.* 2009a). Collaborating with foresters could allow deer management to play a role in shaping the layout of future forests in a way that reduces access and suitable corridors for dispersal of the invasive sika and with the leverage that this may also address the economically significant damage that sika deer are having on Irish forestry.

3.6 Acknowledgements

We would like to thank all the stalkers who provided samples from deer. This work was supported by a Natural Environment Research Council PhD studentship.

3.7 Appendices

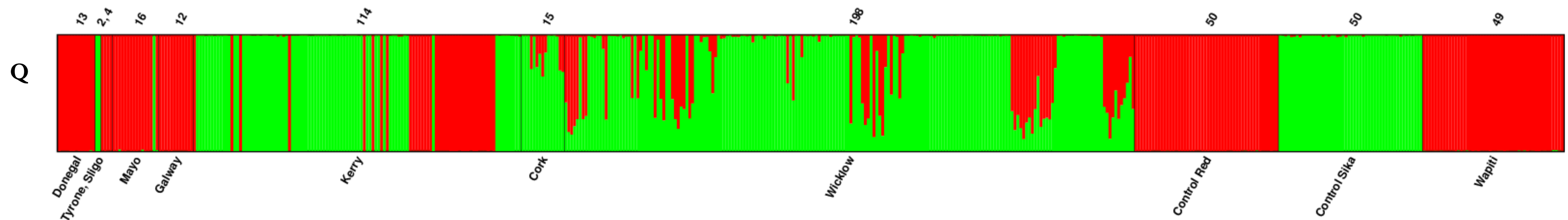


Figure 3.A1. Bar chart showing the results of *analysis 1* in STRUCTURE at K = 2 for each individual in the dataset consisting of red deer, sika and wapiti animals (n = 523). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (shown in red) and the proportion attributable to sika ancestry (green). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis. Counties are plotted in an approximately north west to south-east order.

Table 3.A1. Posterior allele frequencies from analysis 2 (n = 471) at K = 2 and species specific allele assignment. An allele was not assigned to a species if its frequency was less than 1% (0.01) for both species. Alleles were assigned to a species (red = red, green = sika) if its frequency in the other species was 0 or if its frequency was five-fold larger than the other species.

Locus	% Missing data	Allele Size	Estimated allele frequency in Red	Estimated allele frequency in Sika	Allele species - specific assignment
AGLA293	0.60%	128	0.086	0.002	R
		144	0.772	0.002	R
		147	0.056	0.993	S
		Null	0.086	0.004	R
BM4006	0.00%	85	0.002	0.979	S
		87	0.084	0.001	R
		93	0.779	0.005	R
		95	0.113	0.013	R
		Null	0.023	0.002	R
BM6438	4.20%	249	0.712	0.014	R
		251	0.100	0.000	R
		253	0.064	0.000	R
		261	0.023	0.000	R
		263	0.004	0.008	NA
		265	0.002	0.596	S
		273	0.003	0.066	S
		275	0.001	0.243	S
		Null	0.091	0.073	NA
BM757	0.40%	160	0.092	0.000	R
		162	0.483	0.000	R
		172	0.001	0.878	S
		174	0.010	0.116	S
		179	0.061	0.000	R
		183	0.126	0.003	R
		185	0.039	0.000	R
		187	0.072	0.000	R
		198	0.034	0.001	R
		200	0.067	0.000	R
		202	0.007	0.000	NA
		Null	0.007	0.001	R
BOVIRP	0.00%	140	0.002	0.991	S
		144	0.001	0.002	NA
		145	0.008	0.005	NA
		147	0.081	0.000	R
		149	0.074	0.000	R
		151	0.160	0.000	R
		153	0.504	0.000	R
		155	0.053	0.000	R
		157	0.089	0.000	R
		159	0.023	0.000	R
		Null	0.006	0.001	R
FCB193	0.80%	101	0.093	0.000	R
		103	0.018	0.002	R
		105	0.010	0.000	R
		107	0.039	0.000	R
		109	0.160	0.000	R
		111	0.085	0.000	R
		113	0.332	0.000	R
		118	0.054	0.000	R
		120	0.059	0.000	R
		122	0.057	0.000	R
		124	0.007	0.000	NA
		126	0.004	0.022	S
		128	0.004	0.021	S
		130	0.007	0.000	NA
		132	0.003	0.814	S
		134	0.002	0.121	S
		140	0.004	0.000	NA
		143	0.027	0.000	R
		Null	0.035	0.020	NA
FSHB	0.60%	180	0.001	0.831	S
		182	0.001	0.017	S
		183	0.004	0.000	R
		184	0.010	0.000	NA
		185	0.186	0.000	R
		186	0.013	0.000	R
		188	0.076	0.001	R
		189	0.128	0.045	NA
		190	0.061	0.000	R
		191	0.095	0.000	R
		192	0.007	0.000	NA

		193	0.004	0.000	NA
		194	0.004	0.000	NA
		196	0.004	0.000	NA
		197	0.007	0.000	NA
		198	0.125	0.000	R
		199	0.025	0.025	NA
		200	0.018	0.039	NA
		201	0.015	0.000	NA
		202	0.035	0.003	R
		203	0.051	0.000	R
		204	0.024	0.000	R
		205	0.021	0.000	R
		207	0.026	0.034	NA
		208	0.006	0.004	NA
		209	0.001	0.002	NA
		Null	0.055	0.001	R
IDVGA29	3.60%	136	0.645	0.004	R
		143	0.267	0.001	R
		145	0.002	0.078	S
		146	0.003	0.076	S
		156	0.002	0.777	S
		Null	0.082	0.065	NA
IDVGA55	0.20%	191	0.032	0.000	R
		193	0.079	0.000	R
		195	0.150	0.000	R
		197	0.317	0.000	R
		199	0.180	0.000	R
		202	0.042	0.000	R
		204	0.006	0.125	S
		210	0.001	0.729	S
		212	0.001	0.088	S
		214	0.001	0.002	NA
		215	0.001	0.002	NA
		217	0.082	0.006	R
		219	0.021	0.000	R
		Null	0.087	0.048	NA
INRA005	0.80%	124	0.006	0.007	NA
		126	0.941	0.151	R
		136	0.004	0.000	NA
		143	0.010	0.815	S
		Null	0.039	0.026	NA
INRA006	0.60%	130	0.002	0.991	S
		132	0.081	0.002	R
		134	0.673	0.001	R
		136	0.115	0.000	R
		138	0.076	0.000	R
		Null	0.053	0.005	R
INRA131	0.00%	92	0.015	0.000	R
		94	0.028	0.073	NA
		98	0.514	0.002	R
		100	0.300	0.001	R
		102	0.085	0.000	R
		104	0.010	0.000	R
		106	0.002	0.886	S
		113	0.001	0.035	S
		Null	0.044	0.002	R
MM012	0.00%	89	0.751	0.010	R
		91	0.207	0.006	R
		93	0.002	0.981	S
		95	0.004	0.000	NA
		Null	0.035	0.002	R
RM012	0.00%	116	0.001	0.993	S
		120	0.013	0.000	R
		125	0.119	0.000	R
		127	0.016	0.000	R
		129	0.235	0.002	R
		131	0.058	0.000	R
		133	0.168	0.000	R
		137	0.016	0.000	R
		139	0.056	0.001	R
		141	0.101	0.000	R
		144	0.021	0.000	R
		151	0.140	0.000	R
		Null	0.056	0.003	R
RM188	0.60%	113	0.084	0.000	R
		115	0.007	0.000	NA
		117	0.012	0.000	R
		123	0.169	0.000	R
		125	0.067	0.000	R
		127	0.214	0.000	R
		129	0.170	0.000	R
		131	0.021	0.000	R
		132	0.055	0.000	R
		133	0.007	0.000	NA
		134	0.010	0.000	R
		137	0.045	0.000	R
		139	0.070	0.000	R
		143	0.001	0.519	S

		144	0.004	0.000	NA
		145	0.002	0.046	S
		153	0.002	0.084	S
		161	0.001	0.193	S
		163	0.001	0.002	NA
		176	0.001	0.009	S
		182	0.001	0.113	S
		Null	0.059	0.034	NA
RM95	0.20%	116	0.001	0.151	S
		118	0.028	0.000	R
		120	0.004	0.000	NA
		122	0.008	0.842	S
		124	0.121	0.000	R
		126	0.018	0.000	R
		128	0.203	0.000	R
		130	0.124	0.001	R
		132	0.234	0.000	R
		136	0.054	0.000	R
		138	0.047	0.000	R
		140	0.091	0.000	R
		Null	0.065	0.006	R
RME025	0.80%	151	0.012	0.000	R
		155	0.130	0.000	R
		159	0.001	0.002	NA
		168	0.654	0.004	R
		170	0.048	0.001	R
		183	0.004	0.000	NA
		193	0.001	0.867	S
		195	0.001	0.027	S
		207	0.040	0.082	NA
		Null	0.109	0.016	R
TGLA40	0.60%	91	0.167	0.000	R
		95	0.085	0.000	R
		97	0.405	0.005	R
		99	0.069	0.000	R
		101	0.172	0.000	R
		104	0.002	0.563	S
		106	0.002	0.269	S
		Null	0.098	0.162	S
TGLA126	0.40%	99	0.001	0.002	NA
		100	0.007	0.568	S
		101	0.004	0.425	S
		105	0.939	0.003	R
		Null	0.049	0.002	R
TGLA127	0.40%	161	0.001	0.459	S
		167	0.010	0.000	R
		169	0.150	0.001	R
		171	0.004	0.000	NA
		174	0.178	0.499	NA
		176	0.062	0.000	R
		178	0.301	0.002	R
		180	0.010	0.000	R
		184	0.083	0.000	R
		186	0.056	0.000	R
		190	0.049	0.000	R
		192	0.069	0.000	R
		Null	0.027	0.040	NA
TGLA337	8.30%	126	0.001	0.155	S
		128	0.001	0.188	S
		130	0.138	0.000	R
		132	0.064	0.000	R
		134	0.047	0.002	R
		136	0.332	0.000	R
		138	0.068	0.216	NA
		145	0.140	0.003	R
		147	0.151	0.219	NA
		155	0.001	0.086	S
		Null	0.056	0.130	NA
UWCA47	1.50%	225	0.028	0.000	R
		229	0.038	0.000	R
		231	0.898	0.079	R
		240	0.002	0.904	S
		Null	0.034	0.016	NA

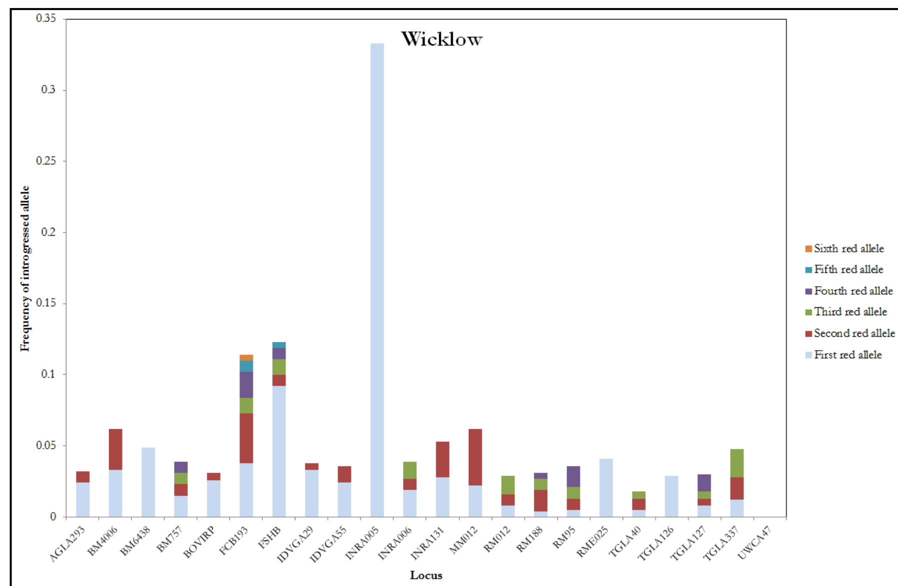
Table 3.A2. Posterior allele frequencies from analysis 1 (n = 523) at K = 3

Locus	% Missing data	Allele Size	Estimated allele frequency in Wapiti	Estimated allele frequency in Red	Estimated allele frequency in Sika
AGLA293	0.60%	128	0.002	0.084	0.002
		144	0.870	0.777	0.002
		147	0.038	0.054	0.993
		149	0.060	0.001	0.000
		Null	0.030	0.084	0.004
BM4006	0.00%	85	0.001	0.001	0.979
		87	0.002	0.085	0.001
		93	0.982	0.784	0.005
		95	0.002	0.111	0.013
		Null	0.013	0.019	0.002
BM6438	3.80%	249	0.003	0.714	0.015
		251	0.001	0.102	0.000
		253	0.001	0.063	0.000
		261	0.031	0.023	0.000
		263	0.819	0.005	0.005
		265	0.116	0.003	0.597
		273	0.001	0.003	0.066
		275	0.001	0.001	0.244
		Null	0.028	0.087	0.073
BM757	0.40%	160	0.001	0.091	0.000
		162	0.001	0.490	0.000
		172	0.001	0.001	0.878
		173	0.238	0.001	0.000
		174	0.001	0.008	0.116
		175	0.050	0.001	0.000
		177	0.208	0.001	0.000
		179	0.001	0.060	0.000
		183	0.001	0.123	0.004
		185	0.001	0.038	0.000
		187	0.040	0.072	0.000
		192	0.060	0.001	0.000
		198	0.197	0.032	0.001
		200	0.197	0.070	0.000
		202	0.000	0.006	0.000
		Null	0.002	0.005	0.001
BOVIRP	0.20%	140	0.001	0.001	0.993
		144	0.001	0.001	0.002
		145	0.363	0.007	0.004
		147	0.330	0.084	0.000
		149	0.001	0.076	0.000
		151	0.049	0.158	0.000
		153	0.001	0.503	0.000
		155	0.001	0.052	0.000
		157	0.063	0.091	0.000
		159	0.001	0.022	0.000
		161	0.159	0.001	0.000
		Null	0.031	0.005	0.001
FCB193	0.80%	101	0.001	0.095	0.000
		103	0.001	0.017	0.002
		105	0.000	0.009	0.000
		107	0.000	0.039	0.000
		109	0.001	0.161	0.000
		111	0.001	0.087	0.000
		113	0.001	0.336	0.000
		118	0.000	0.053	0.000
		120	0.071	0.058	0.000
		122	0.021	0.057	0.000
		124	0.000	0.006	0.000
		126	0.650	0.003	0.020
		128	0.010	0.004	0.021
		130	0.000	0.006	0.000
		132	0.019	0.004	0.818
		134	0.020	0.002	0.121
		140	0.000	0.003	0.000
		143	0.042	0.029	0.000
		145	0.030	0.001	0.000
		150	0.129	0.001	0.000
		Null	0.002	0.028	0.017
FSHB	0.60%	180	0.000	0.001	0.831
		182	0.629	0.001	0.016
		183	0.000	0.003	0.000
		184	0.110	0.009	0.000
		185	0.102	0.188	0.000
		186	0.000	0.012	0.000
		187	0.010	0.000	0.000
		188	0.139	0.073	0.001
		189	0.001	0.127	0.045
		190	0.000	0.060	0.000
		191	0.000	0.098	0.000
		192	0.000	0.012	0.000

		193	0.000	0.003	0.000
		194	0.000	0.003	0.000
		196	0.000	0.003	0.000
		197	0.000	0.006	0.000
		198	0.000	0.130	0.000
		199	0.001	0.024	0.025
		200	0.001	0.017	0.039
		201	0.000	0.015	0.000
		202	0.001	0.034	0.003
		203	0.000	0.051	0.000
		204	0.000	0.024	0.000
		205	0.000	0.021	0.000
		207	0.001	0.025	0.034
		208	0.000	0.005	0.004
		209	0.000	0.001	0.002
		Null	0.001	0.051	0.001
IDVGA29	3.30%	134	0.961	0.001	0.000
		136	0.004	0.646	0.005
		143	0.002	0.270	0.001
		145	0.001	0.003	0.079
		146	0.001	0.003	0.075
		156	0.001	0.001	0.774
		Null	0.031	0.075	0.065
IDVGA55	0.20%	191	0.467	0.031	0.000
		193	0.001	0.085	0.000
		195	0.001	0.148	0.000
		197	0.512	0.315	0.000
		199	0.001	0.180	0.000
		202	0.001	0.041	0.000
		204	0.001	0.005	0.125
		210	0.001	0.001	0.729
		212	0.001	0.001	0.088
		214	0.001	0.001	0.002
		215	0.001	0.001	0.002
		217	0.001	0.083	0.006
		219	0.001	0.021	0.000
		Null	0.013	0.087	0.047
INRA005	0.80%	124	0.001	0.005	0.008
		126	0.983	0.946	0.151
		136	0.001	0.004	0.000
		143	0.001	0.010	0.815
		Null	0.015	0.035	0.026
INRA006	0.60%	130	0.001	0.001	0.992
		132	0.002	0.079	0.003
		134	0.123	0.680	0.001
		136	0.866	0.116	0.000
		138	0.001	0.078	0.000
		Null	0.007	0.045	0.004
INRA131	0.00%	92	0.609	0.015	0.000
		94	0.001	0.027	0.073
		98	0.274	0.517	0.002
		100	0.104	0.300	0.002
		102	0.001	0.090	0.000
		104	0.001	0.010	0.000
		106	0.001	0.001	0.885
		113	0.001	0.001	0.035
		Null	0.009	0.038	0.002
MM012	0.00%	89	0.116	0.761	0.011
		91	0.620	0.207	0.006
		93	0.261	0.002	0.981
		95	0.001	0.004	0.000
		Null	0.002	0.026	0.002
RM012	0.00%	116	0.001	0.001	0.992
		120	0.001	0.012	0.000
		125	0.001	0.121	0.000
		127	0.775	0.014	0.000
		129	0.001	0.238	0.003
		131	0.001	0.057	0.000
		133	0.009	0.167	0.000
		137	0.041	0.016	0.000
		139	0.150	0.055	0.001
		141	0.001	0.103	0.000
		144	0.001	0.021	0.000
		151	0.001	0.139	0.000
		Null	0.020	0.057	0.003
RM188	0.60%	113	0.001	0.083	0.000
		115	0.000	0.006	0.000
		117	0.000	0.012	0.000
		123	0.001	0.170	0.000
		125	0.000	0.067	0.000
		127	0.134	0.213	0.000
		129	0.001	0.175	0.000
		131	0.000	0.021	0.000
		132	0.021	0.054	0.000
		133	0.000	0.006	0.000
		134	0.790	0.009	0.000
		137	0.021	0.047	0.000
		139	0.001	0.069	0.000
		143	0.000	0.001	0.518

		144	0.000	0.003	0.000
		145	0.000	0.002	0.046
		153	0.000	0.001	0.084
		161	0.000	0.001	0.193
		163	0.000	0.001	0.002
		176	0.000	0.001	0.009
		182	0.000	0.001	0.113
		Null	0.027	0.057	0.034
RM95	0.20%	116	0.000	0.001	0.151
		118	0.001	0.028	0.000
		120	0.000	0.003	0.000
		122	0.300	0.010	0.842
		124	0.031	0.126	0.000
		126	0.001	0.018	0.000
		128	0.001	0.202	0.000
		130	0.001	0.123	0.001
		132	0.001	0.235	0.000
		136	0.239	0.053	0.000
		138	0.180	0.049	0.000
		140	0.001	0.090	0.000
		142	0.149	0.001	0.000
		144	0.078	0.001	0.000
		153	0.010	0.001	0.000
		Null	0.007	0.062	0.004
RME025	0.80%	132	0.711	0.001	0.000
		134	0.227	0.001	0.000
		136	0.031	0.001	0.000
		151	0.001	0.012	0.000
		155	0.001	0.127	0.000
		159	0.001	0.001	0.002
		168	0.002	0.659	0.005
		170	0.001	0.047	0.001
		183	0.001	0.003	0.000
		193	0.001	0.001	0.867
		195	0.001	0.001	0.027
		207	0.001	0.039	0.082
		Null	0.023	0.108	0.015
TGLA40	0.60%	91	0.001	0.165	0.000
		95	0.001	0.084	0.000
		97	0.665	0.412	0.004
		99	0.318	0.067	0.000
		101	0.001	0.173	0.000
		104	0.001	0.002	0.563
		106	0.001	0.002	0.270
		Null	0.012	0.097	0.162
TGLA126	0.40%	99	0.001	0.001	0.002
		100	0.001	0.007	0.569
		101	0.001	0.003	0.425
		104	0.203	0.001	0.000
		105	0.767	0.945	0.003
		Null	0.027	0.043	0.002
TGLA127	0.40%	161	0.001	0.001	0.458
		167	0.000	0.009	0.000
		169	0.001	0.151	0.001
		171	0.000	0.003	0.000
		174	0.020	0.178	0.499
		176	0.001	0.061	0.000
		178	0.620	0.298	0.002
		180	0.317	0.011	0.000
		182	0.020	0.001	0.000
		184	0.001	0.087	0.000
		186	0.001	0.056	0.000
		190	0.001	0.048	0.000
		192	0.001	0.068	0.000
		Null	0.017	0.029	0.039
TGLA337	8.00%	111	0.185	0.001	0.000
		118	0.772	0.001	0.000
		126	0.001	0.001	0.155
		128	0.001	0.001	0.187
		130	0.010	0.138	0.000
		132	0.001	0.063	0.000
		134	0.001	0.049	0.002
		136	0.001	0.332	0.000
		138	0.001	0.067	0.216
		145	0.001	0.143	0.004
		147	0.001	0.147	0.220
		155	0.001	0.001	0.086
		Null	0.024	0.057	0.130
UWCA47	1.30%	225	0.001	0.030	0.000
		229	0.001	0.037	0.000
		231	0.983	0.903	0.079
		240	0.001	0.001	0.906
		Null	0.014	0.028	0.015

a)



b)

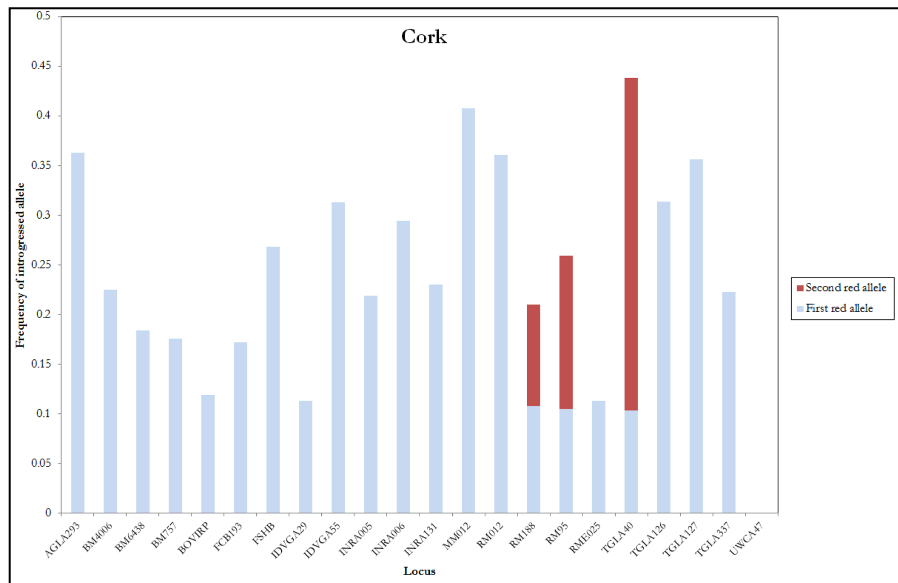


Figure 3.A2. Frequency of introgressed alleles from red deer into Japanese sika-like animals (Q< 0.5) at each locus in a) Co. Wicklow and b) Co. Cork. Note the scale of the y-axis is different for the different regions to highlight the variation in the relative frequencies.

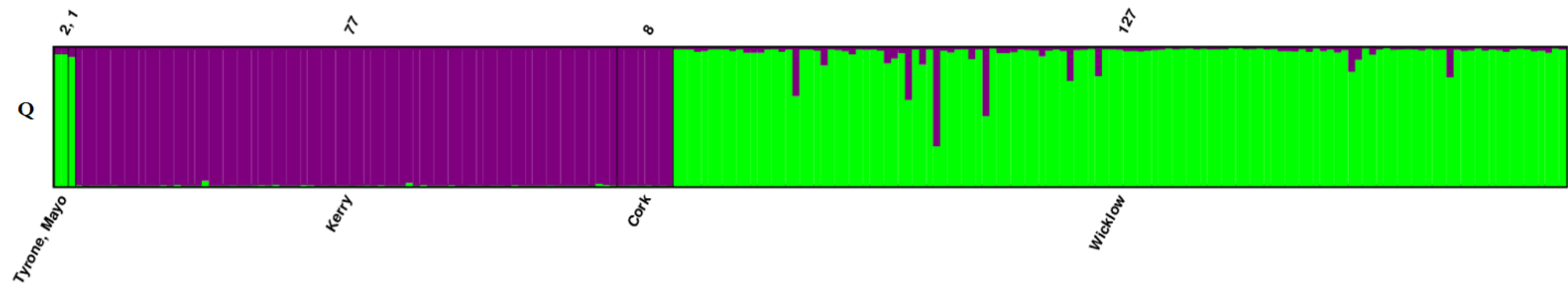


Figure 3.A3. Bar chart showing the results of *analysis 3* in STRUCTURE at K = 2 for each individual in the dataset consisting of all sika individuals sampled across 5 counties in Ireland (n = 215). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to one of the two sika clusters (purple and green). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis. Counties are plotted in an approximately north west to south-east order.

Chapter 4: Population structure and genetic diversity within two *Cervus* species in the British Isles

Author's contributions:

Of the 3059 red and sika individuals analysed in this study, 777 were genotyped by Helen Senn, 725 by Elizabeth Heap and Sheena Morrison, 1261 by SS, 132 by Elizabeth Mittel and Sarah MacDonald under the supervision of SS and 164 by Megan Wyman and SS. Statistical analysis was performed by SS with guidance when using DAPC from Erwan Quemere (INRA, France). SS wrote the MS. JMP guided the study and edited and commented on the MS.

4.1 Abstract

Europe's largest population of wild red deer (*Cervus elaphus)* resides in the British Isles and has been present since the end of the last ice age, c. 11,000BP. 150 years ago, Japanese sika deer (*C. nippon*) were introduced to the British Isles and have since established and expanded their range. Where sika are sympatric with red deer, they sometimes hybridise with them. This study investigates the population genetic structure within red and sika populations independently, using samples from across the British Isles genotyped at 22 microsatellite markers and typed for a diagnostic mtDNA marker. We analysed a microsatellite genotype dataset consisting of 2307 red deer and another with 752 sika from the British Isles defined as 'pure' according to criteria derived from previous analysis using the Bayesian genetic clustering and Discriminant Analysis of Principal Components (DAPC). We then tightened the genetic criteria for a 'pure' animal of each taxon and analysed the remaining red and sika which met this more stringent definition. Comparison of analyses at the two definitions of purity show that under stricter purity criteria the estimated within-taxon population structure changes for red deer. Within red deer, both approaches suggest that at the higher purity criterion the primary differentiation is that between red deer on the Hebridean islands of Harris and Lewis and all other red deer whilst at the less stringent criteria five population clusters are supported, which reflect known differentiation and translocation in the species. Amongst sika deer at both definitions of purity, two to three population clusters are supported, which does not match geography and likely reflects the history of introductions and bottlenecks of this introduced species. Individuals removed by tightening the purity criteria were mainly sampled in areas of known introgression and carried introgressed alleles at higher frequency than in the remaining purer animals. We conclude that in areas of known introgression, there is more introgression than our criteria used hitherto suggest.

Key words: Cervus*, microsatellite, population structure, sika, red.*

4.2 Introduction

The raison d'etre of current conservation genetic efforts is to maintain populations in as natural state as possible (Zachos & Hartl 2011). In more detail, conservation genetics aims to preserve the adaptive diversity and evolutionary lineage of a species, through natural gene flow, across its range (Crandall *et al.* 2000; Hobbs *et al.* 2011). This process

is facilitated by the identification of population units across an area, which it is not always possible to infer from geographic proximity as there may be inconspicuous gene flow barriers or corridors or hybridisation with congeneric species. Combining ecological data, demographic history and population structure inferred from genetic methods is optimal for establishing the most accurate population units (Crandall *et al.* 2000; Goodman *et al.* 2001). This approach enables management to be prioritised on and tailored toward particular genetic units and the gene flow between them. Amongst wild populations, use of genetic markers for guiding management has recently been exemplified in wild boars in the Balkans (*Sus scorfa*), brown bears in Eurasia (*Ursus spp.)*, elephants *(Loxodonta africana*) in the Kenyan savannah and otters (*Lutra lutra*) in the UK to name but a few (Hobbs *et al.* 2011; Okello *et al.* 2008; Tammeleht *et al.* 2010; Velickovic *et al.* 2012).

The same genetic tools can also be used to explore the population genetic structure of exotic, introduced species, in order to manage them to ameliorate their impact and monitor the effectiveness of any control measures implemented. Examples include studies on an invasive weed (*Ageratina spp.)* in China, the introduced fire ant in Taiwan, non-native molluscs (*Cyclope neritea*) in Iberia and feral pigs (*S. scrofa*) in south-western Australia (Couceiro *et al.* 2012; Gui *et al.* 2009; Hampton *et al.* 2004; Yang *et al.* 2008).

4.2.1 Factors which can influence population genetic structure

There exists a plethora of factors that can influence the genetic population structure of a species. Without major disruption or human influence, populations may be expected to show isolation-by-distance patterns of structure, in which genetic differentiation is largely concordant with geographical distance (Hmwe *et al.* 2006b). The exact pattern is related to the dispersal ability of the species, as this determines the extent and range of gene flow, which shapes the structure (Vigilant & Guschanski 2009). Isolation-by-distance has been demonstrated using nuclear genetic markers in many populations, including mainland Scottish red deer (Pérez-Espona *et al.* 2008) and red grouse in north-east Scotland (Piertney *et al.* 1998).

Beyond distance, the quality and topology of a landscape can influence a species' population structure. Perez-Espona *et al.* (2008) demonstrated the significant effect of particular natural landscape features (e.g. the Great Glen, sea lochs, mountain slopes) as barriers to gene flow between red deer populations in the Scottish highlands and others (e.g. inland lochs and rivers) which act to facilitate gene flow. These effects may result in significant differences in genetic variation over small distances as in the case of the

Great Glen in red deer (Pérez-Espona *et al.* 2008). Unexpected landscape features may turn out to affect gene flow (Vigilant & Guschanski 2009). For example, unsuitable ground surrounding a Scottish river system acts as a barrier to red grouse (*Lagopus lagopus scoticus*) (Piertney *et al.* 1998), bais in northern Congo attract and facilitate gene flow in western lowland gorillas (*Gorilla gorilla gorilla*) (Parnell 2002), ice-free corridors in the Canadian Rockies restrict caribou (*Rangifer tarandus*) gene flow (McDevitt *et al.* 2009b) and spatially distributed water reservoirs in Australia influence the population structure of feral pigs (*S. scrofa*) (Hampton *et al.* 2004).

The social structure and behaviour of a species can also influence its genetic population structure. Philopatry risks inbreeding and deleterious effects while dispersal reduces competition with relatives and avoids inbreeding (Piertney *et al.* 1998; Wheelwright & Mauck 1998). This is particularly notable in avian systems, such as Savannah sparrows in the Bay of Fundy, Canada, which exhibit strong regional philopatry ("natal dispersal") and yet appear to be able to recognise and disperse locally away from related individuals during breeding ("breeding dispersal") in order to actively avoid inbreeding depression (Wheelwright & Mauck 1998). Similarly, the breeding system and sex-biased dispersal can shape population structure among polygamous animals, such as brook charr (*Salvelinus fontinalis*) in Canadian lakes (Fraser *et al.* 2004), amongst Soay sheep on St. Kilda (Coltman *et al.* 1998) and primates in Japan (Vigilant & Guschanski 2009).

Beyond these intrinsic factors, anthropogenic disturbances can drastically affect population structure. In addition to habitat exploitation and fragmentation, obstacles such as roads and railways can act as barriers to gene flow, leaving populations isolated and vulnerable. For example, a large highway in Los Angles inhibits genetically effective movement of various species of carnivore (Riley *et al.* 2006); patches of arable and urbanised land in Tibet appear to be shaping population structure of the resident snub-nosed monkeys (*Rhinopithecus bieti*) (Liu *et al.* 2009); and flow regimes from hydropower schemes affect genetic structure of Lake Sturgeon (*Acipenser fulvescens*) in the Great Lakes (Welsh *et al.* 2008).

Anthropogenic influences on population structure may also take the form of translocation and introduction of animals. Perez-Espona *et al.* (2012) considered the impact of red deer introductions on Scottish populations, concluding that introductions of a few animals, to improve trophy quality, to large existing populations in mainland Scotland has generally had minimal impact, whereas larger numbers used to restock (small) populations on Scottish islands have often had relatively greater impact.

4.2.2 Red deer in the British Isles

Since the end of the last glaciation c.11,000 BP, red deer populations in the British Isles have been shaped by the post-glacial expansion of the human population and the consequent deforestation and loss of suitable habitat (Hmwe *et al.* 2006b; Zachos & Hartl 2011). Episodes of population expansion and contraction incurred will have influenced current populations. By the end of the 18th century, it is likely many wild native stocks of red deer were extinct, recovering in the 19th century as a result of the Victorian fashion for highland sporting estates. Populations declined again during the First World War (Hmwe *et al.* 2006b) but have recovered dramatically since. Their distribution and genetics has also been influenced by numerous introductions and translocations made for a variety of purposes (Pérez-Espona *et al.* 2009a). Since the mid-20th century and in particular over the last 30 years, the population of red deer has been expanding at a rate of around 0.3% p.a. in England and Scotland and up to 7% p.a. in Ireland (Carden *et al.* 2010; Ward 2005). Below, the red deer populations residing in Scotland, England and Ireland will be briefly introduced as samples were sourced from each country for this study.

Red deer have existed throughout much of Scotland since at least the 8th century, but on the mainland they were slowly driven northwards into the highlands by deforestation and population of lowland areas by humans (Whitehead 1964). Current estimates suggest there are 450,000 in the country (Clutton-Brock *et al.* 2004). Recent genetic work has shown that on the mainland, the genetic structure of red deer is largely concordant with geographical locality and the modifying influence of particular landscape features on gene flow (Pérez-Espona *et al.* 2009b; Pérez-Espona *et al.* 2008) (see above). On the other hand, a long history of extinctions from, introductions to and isolation of the Hebridean islands has led to more discordant population structure (Pérez-Espona *et al.* 2013; Pérez-Espona *et al.* 2009b). The discovery of a mitochondrial haplotype amongst the animals on Rum closely related to Corsican red deer (*C. elaphus corsicanus*) exemplifies the long lines of translocation that have occurred (Nussey *et al.* 2006) and low mitochondrial variation amongst samples from Islay has been interpreted as a founder effect following introduction to the island (Hmwe *et al.* 2006b). Legislation now in place under the Wildlife and Countryside Act (1981) means it is currently illegal to introduce deer of the genus *Cervus* onto the Outer Hebrides, Rum, Jura, Islay and Arran without genetic testing, in order to protect red deer populations from sika hybridisation.

Estimates for the number of wild red deer resident in England lie between 12,500 and 20,000 and they occupy a far patchier distribution than in mainland Scotland (Díaz *et al.* 2006; Harris 1995; Hmwe *et al.* 2006b; Pérez-Espona *et al.* 2009a; Ward 2005). Many English red deer may be the descendants of introductions, for example, putative native reds in the New Forest were supplemented with park stock around the 1960s and are now of unknown genetic provenance (Díaz *et al.* 2006). The red deer of Grizedale and Martindale in the Lake District, however, are suspected to be native (Lowe & Gardiner 1975; Pérez-Espona *et al.* 2009a; Whitehead 1964).

Since Medieval times, when Royal parks and enclosures were established within which deer where enclosed and hunted, the deer park has been a significant feature of the British Isles, especially England, and many remain to the present (Hingston 1988). Park numbers reached their peak during the 13th century at around 2,000 in England & Wales; however, disruptions during the early 20th century saw them fall to only 112 in England, although there has been some recovery since (Hingston 1988). It is estimated that between 3,000 to over 10,000 red deer are currently in captivity in parks (Hingston 1988; Whitehead 1964). Overall, English wild and captive deer populations have been rather understudied in terms of recent genetic assessment - in previous studies, marker panels and sample sizes have been relatively small and discriminatory power, therefore, low (Díaz *et al.* 2006; Hmwe *et al.* 2006b).

In Ireland, the modern red deer population is descended from ancient and recent postglacial introductions by man (Carden *et al.* 2012). There are currently thought to be around 4,000 red deer in Ireland and they are primarily present in the East (Co. Wicklow; although note that in Chapter 3 we found that all phenotypically red deer sampled were hybrid), the South West (Co. Kerry) and the North West (Co. Galway up to Co. Donegal) and are expanding at a considerable rate (see above) (Carden *et al.* 2010; Pérez-Espona *et al.* 2009a; Whitehead 1964). Recent work has established that the red deer centred on Killarney, Co. Kerry, are descended from a human introduction from Britain during the Neolithic period (Carden *et al.* 2012). The other populations are descended from more recent introductions from Britain and continental Europe, in several cases indirectly through parks, primarily Powerscourt Park, Co. Wicklow, where they may have interacted with other *Cervus* species (Carden *et al.* 2012; McDevitt *et al.* 2009a). This is supported by a genetic study using nine microsatellites and mitochondrial sequence data which suggested the red deer from Co. Kerry formed a distinct cluster, those from the North-west formed a second cluster, whilst the final cluster contained a mix of those from the remaining sites (McDevitt *et al.* 2009a). The haplotype diversity in

these other regions was found to be up to ten times greater than that found in Killarney, Co. Kerry (McDevitt *et al.* 2009a). Further corroboration comes from the results of Chapter 3 which showed that the red deer in Wicklow have become so introgressed that pure red animals were not found in this study, whilst the red deer in Kerry formed a cluster with relatively low genetic diversity but no evidence for sika introgression, based on our panel of 22 microsatellite markers.

4.2.3 Sika in the British Isles

Since the mid-19th century, introductions of sika deer (*C. nippon*) from Japan have taken place at numerous sites, especially deer parks, throughout the British Isles, and many have escaped or been deliberately released to the wild (Ratcliffe 1987). In the last 30 years sika populations are believed to have expanded their range by around 5.3% p.a. in Great Britain and by a similar rate in Ireland (Carden *et al.* 2010; Ward 2005). This has led to inevitable overlap with the range of red deer and hybridisation between these species has since been documented in the wild in the British Isles (Chapter 2 ; Chapter 3 ; Goodman *et al.* 1999; Lowe & Gardiner 1975; Ratcliffe 1987; Senn 2009). Below, we briefly describe the sika populations covered in this study, in Scotland, England and Ireland.

It likely there are now around 15,000-20,000 sika in Scotland (Pérez-Espona *et al.* 2009a) occupying around 40% (c.14,000km^2) of the country, the distribution of which is attributed to twelve separate episodes of introduction, release or escape (Pérez-Espona *et al.* 2009a; Ratcliffe 1987; Ward 2005). Whilst the majority of these likely came to the UK via Powerscourt Estate, Wicklow, populations in Dawyck in Peebles are suspected to have come straight from Japan (Goodman *et al.* 2001; Ratcliffe 1987). Sika in the British Isles are genetically similar to those from Kyushu, Japan and their likely source was Nagasaki, in Kyushu, the only Japanese port open to international trade around the time these animals were exported (Goodman *et al.* 1999). Population diversity and structure within this repeatedly translocated species in Scotland is yet to be assessed in detail.

Within England, there are an estimated 1,500-2,000 sika in the wild, which occupy a very discontinuous distribution (Díaz *et al.* 2006; Pérez-Espona *et al.* 2009a). In captivity, sika have been introduced to numerous parks, several of which still hold herds of several hundred in counties such as Kent, Dorset and Buckinghamshire (Whitehead 1964). Goodman *et al.* (2001) used nine microsatellites to explore and compare the genetic population structure of sika in Japan with those introduced to several British

sites, including sika from Dorset and concluded these introductions clustered with those carrying the southern mitochondrial haplotype in Japan (consistent with origin in Kyushu, see above). In addition, a genetic study using six polymorphic microsatellite loci found no genetic differentiation between sika populations sampled from Dorset (Purbeck) with those from the New Forest, Hampshire, which likely descended from very few founding individuals independently introduced to both counties (Díaz *et al.* 2006; Ratcliffe 1987).

Since their introduction to Ireland over 150 years ago, sika have established and, over the last 30 years, expanded their population range by 353% primarily in three major centres: the east (Co. Wicklow), the south west (Co. Kerry) and the north (Co. Tyrone, Co. Fermanagh) (Carden *et al.* 2010; Whitehead 1964). This expansion is probably driven by the availability of suitable habitat in Ireland (Carden *et al.* 2010). A panel of eight microsatellites suggested two genetic clusters most accurately described the population structure amongst 47 deer sampled from Ireland; the animals from Co. Kerry formed one cluster whilst those from Co. Wicklow and Co. Down were a mix of cluster 1 and 2 (McDevitt *et al.* 2009a). Similarly, based on our panel of 22 microsatellite markers, Structure analysis in Chapter 3 showed sika populations in Co. Kerry and Co. Wicklow formed distinct clusters (Figure 3.A3).

4.2.4 This study

In this study we investigate whether hybridisation between red deer and sika influences estimates of population genetic structure within red deer and within sika in the British Isles. In previous research (Chapter 2 ; Chapter 3 ; Senn 2009) we used 22 nuclear microsatellites and a single mitochondrial DNA marker which are highly diagnostic for red and sika to assign hybrid status to each individual sampled. However, the microsatellite markers are also polymorphic within these two taxa and so can be used to investigate within-taxon population structure (Senn 2009).

Two approaches were used to analyse genotype data and infer population structure and genetic admixture: a Bayesian clustering method implemented in Structure 2.3.3 (Falush *et al.* 2003, 2007; Pritchard *et al.* 2000) and the multivariate method Discriminant Analysis of Principal Components (DAPC; (Jombart 2008)). We also investigated the possibility that any genetic structure detected within red or sika was influenced by the purity criterion in use. Up to this point our research has used a somewhat arbitrary definition of purity (with 'pure' red having Q>0.95 and 'pure' sika having Q<0.05) following Structure analysis (see Senn & Pemberton 2009 for discussion). To a rough

approximation, this means that individuals with no more than one or two alleles (out of 44) that are either not diagnostic or are characteristic of the other taxon are counted as pure. In this study we use both this definition of purity and a more stringent one ($Q \geq 0.99$ for red; $Q \leq 0.01$ for sika), and investigate the consequences for population structure within each taxon, with the expectation that within-species population structure might be reduced when highly introgressed individuals are removed from analysis because spatially-varying introgressed alleles might be removed. The individuals removed by the more stringent definition are of interest in their own right: if they carry alleles characteristic of the opposite taxon (rather than non-diagnostic alleles) and if they are sampled predominantly in areas where there are known hybrid swarms (from Chapters 2 and 3), then this suggests that they are genuinely introgressed individuals.

The specific aims of this study are;

1. *To assess population genetic structure within 'pure' red deer from across the British Isles using two different purity criteria.*
2. *To assess population genetic structure within 'pure' sika from across the British Isles using two different purity criteria.*
3. *To determine whether individuals removed by the more stringent purity criteria are introgressed hybrids or carry non-diagnostic alleles shared by the parental taxa.*
4. *To consider how can this information be used to benefit management of both species.*

4.3 Materials and Methods

4.3.1 Sampling Sites

Samples were obtained from various sites from the south of England, the Lake District, some of the major populations in Ireland, across mainland Scotland and the Outer Hebrides (Table 4.1, Figure 4.1 and 7.0 Appendix 1).

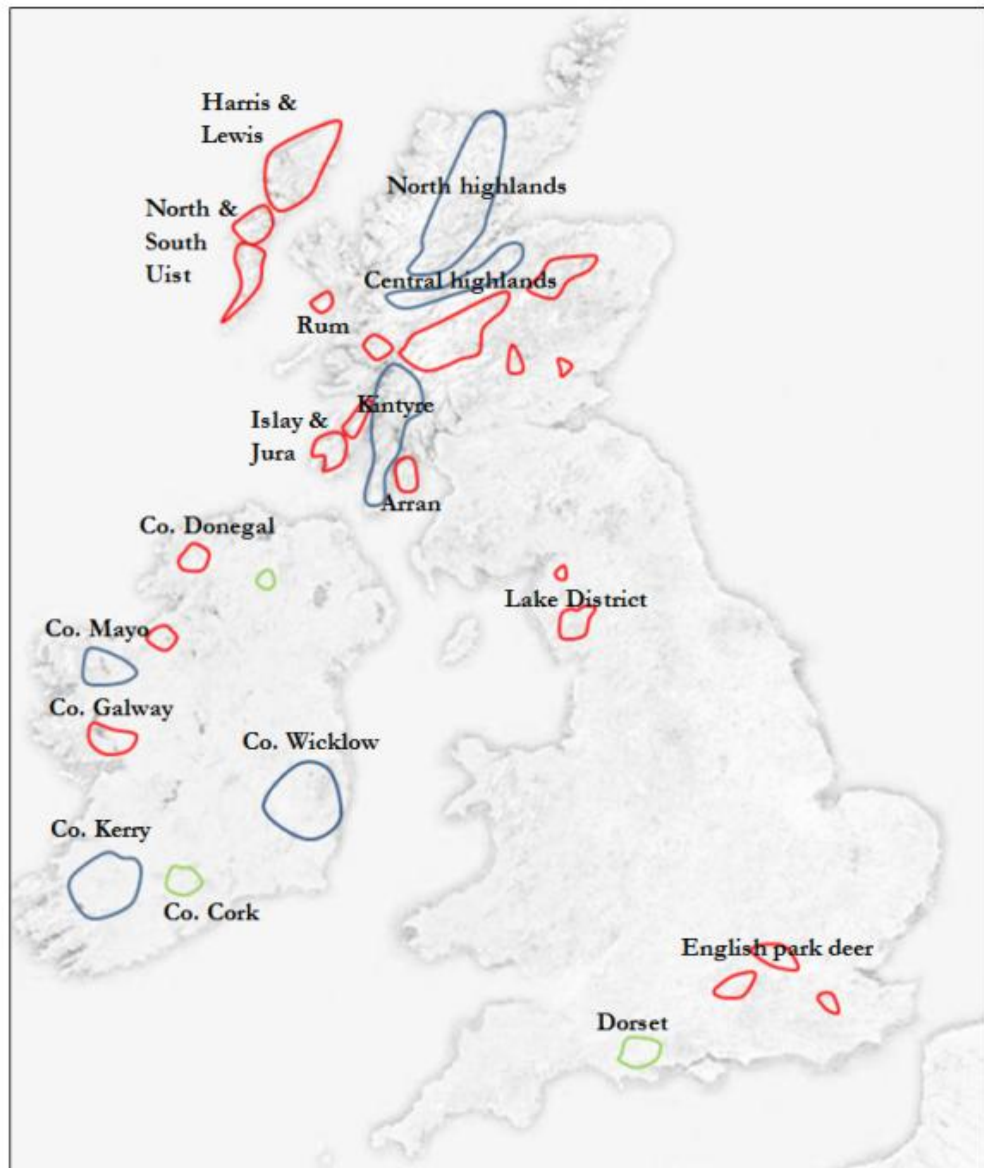


Figure 4.1. Map showing the sites from which samples were obtained; red shows those from which phenotypically red deer only were sampled, green from which phenotypically sika only were sampled and blue from which both species were sampled.

Table 4.1. Sample sizes and genetic data completeness for the 2307 red deer and 752 sika successfully genotyped (at at least 20 out of the 22 markers), shown for five main regions of the British Isles.

	No. genetically pure red (Q > 0.95)	**No. genetically pure sika (Q < 0.05)**	**Nuclear dataset (% complete)**	**MtDNA dataset (% complete)**
Southern England	57	40	98.64	95.88
Lake District, England	132	0	98.73	100
Ireland	80	209	98.33	94.24
Mainland Scotland	1313	503	99.58	99.51
Hebrides, Scotland	725	0	97.50	100

Most of the samples were initially collected for studies of the extent of red-sika hybridisation. Sampling in Ireland is described in Chapter 3, while samples obtained in Scotland and Cumbria are described in Chapter 2. This investigation also included samples collected in southern England during study of the role of vocalisation in hybridisation conducted by Megan Wyman (see author's contributions). The sites in southern England from which red deer were sampled were the deer parks at Wadhurst, Bushy, Badminton, Richmond and Windsor, while the wild sika from England were obtained from Lulworth and Arne, Dorset. Samples were obtained in the course of normal culling operations.

4.3.2 DNA analysis & investigation of hybridisation

See Chapter 2 section 2.3.2 for DNA extraction, microsatellite genotyping and mtDNA haplotyping procedure. Justification for the use of this marker panel comes from the fact it is polymorphic *within* red deer and sika independently, as well as between, whilst the software Structure 2.3 is robust and can account for null alleles (Falush *et al.* 2007; Senn 2009). An analysis of hybridisation in Scottish, Lake District and Irish deer is reported in Chapters 2 and 3, while analysis of hybridisation for the English deer included here is presented in the Appendix to this thesis (section 7.0 Appendix 1).

4.3.3 To assess population genetic structure within 'pure' red deer from across the British Isles using two different purity criteria (objective 1).
To assess population genetic structure within 'pure' sika from across the British Isles using two different purity criteria (objective 2).

From the analyses described in Chapters 2 and 3 and the Appendix (section 7.0) we selected two groups of red deer with different levels of purity and two groups of sika with different levels of purity. The first dataset consisted of all red deer samples from across the British Isles which had a Q value of >0.95 and did not have sika mtDNA (dataset 1, n = 2307) and the second only retained those samples with Q≥0.99 (dataset 2, n = 2201). The same exercise was carried out for the sika: first we analysed all sika samples which returned Q<0.05 and did not have red mtDNA (dataset 3, n = 752) and then we retained those with Q≤0.01 (dataset 4, n = 702).

To investigate population structure within red deer and sika we used Structure 2.3.3 (Falush *et al.* 2003, 2007; Pritchard *et al.* 2000) using the microsatellite genotype data within each dataset. The number of inferred, genetically distinct populations (K) that maximises the likelihood (Ln Pr (X|K)) of the dataset, assuming they are in Hardy-Weinberg equilibrium and linkage equilibrium, was estimated by running five independent replicates at different values of K (1-8). See Chapter 2, section 2.3.3, for Structure methodology and evaluation of the best value of K.

Subsequently, all datasets were analysed using an alternative approach, discriminant analysis of principal components (DAPC) (Jombart *et al.* 2010) using the open source statistical programming language, R. 2.15 (http://www.r-project.org/). This approach seeks to identify clusters by maximising between-population variation, minimising within-population variation (*k-means* approach) and it avoids any assumptions of an explicitly evolutionary model (for example it does not make the assumption of Hardy-Weinberg equilibrium or linkage disequilibrium) (Jombart *et al.* 2010). The number of clusters found within each dataset was determined by retaining the optimal number of principal components for that dataset (established using the a-score) and all the eigenvalues and evaluating K values (number of populations) up to 40. Graphical outputs included tabulating the inferred clusters ("inf") against the populations from which the individuals were obtained ("ori") and producing a scatter plot, in which individuals were located according to coordinates determined by the principal component analysis.

4.3.4 *What are the genetic characteristics of individuals removed by the more stringent purity criteria and where were they sampled (objective 3)?*

In order to address this objective we examined the posterior allele frequencies for the parental taxa generated by Structure following analysis of red and sika together (n = 3059) and assigned these as red-specific, sika-specific or inconclusive, according to conservative criteria (Appendices Table 4.A1). The frequency of sika-specific alleles was then compared amongst the individuals removed from the red dataset ($0.95<Q<0.99$, n=106) in order to meet the criteria of $Q\geq0.99$, with those that satisfied the stricter criterion ($Q\geq0.99$, n=2201). The same exercise was carried out to assess the frequency of red-specific alleles amongst the individuals removed from the sika dataset (n=50, $0.01<Q<0.05$) in order to meet the criteria of $Q\leq0.01$, with those that satisfied the stricter criterion ($Q\leq0.01$, n=702). The location of the removed individuals was also assessed to see if they came preferentially from sites with known red-sika hybridisation.

In order to address whether differences in population structure between this and previous studies are due to the differences in marker informativeness we calculated the allelic diversity for each locus and each population in the more stringent red and sika datasets (2 and 4), respectively, using CERVUS 3.0 (Kalinowski *et al.* 2007). The average number of alleles for all red and all sika populations, alongside other genetic diversity indices for both species were also calculated using CERVUS 3.0.

4.4 Results

Genetic diversity indices are given for each locus within pure red deer (Table 4.2) and sika deer datasets (Table 4.3).

Table 4.2. Genetic diversity indices for each of the 22 loci in our microsatellite marker panel in red deer under the more stringent purity criteria ($0.99 \leq Q \leq 1$, n = 2201) calculated in Cervus 3.0. Parameters are k, the number of alleles at each locus in each species, N, number of samples typed at each locus, H_o, observed heterozygosity, H_e, expected heterozygosity and Null, the frequency of null alleles at each locus.

Locus	k	N	H_O	H_E	Null
AGLA293	3	2156	0.19	0.264	0.1663
BM4006	3	2190	0.334	0.398	0.1017
BM6438	5	2164	0.475	0.585	0.0967
BM757	15	2198	0.598	0.658	0.0543
BOVIRBP	9	2189	0.656	0.751	0.0689
FCB193	20	2146	0.761	0.866	0.0641
FSHB	29	2184	0.823	0.901	0.0454
IDVGA29	3	2156	0.424	0.443	0.0217
IDVGA55	10	2146	0.724	0.793	0.0447
INRA5	3	2196	0	0.001	0.0672
INRA6	5	2199	0.397	0.435	0.0443
INRA131	7	2201	0.525	0.559	0.0334
MM012	5	2199	0.314	0.346	0.0491
RM12	12	2186	0.75	0.862	0.0685
RM188	14	2183	0.634	0.75	0.0877
RM95	13	2191	0.763	0.835	0.0455
RME025	8	2192	0.32	0.361	0.0658
TGLA40	9	2191	0.515	0.637	0.107
TGLA126	6	2201	0.01	0.031	0.3488
TGLA127	13	2195	0.708	0.808	0.0669
TGLA337	11	1949	0.641	0.791	0.104
UWCA47	3	2184	0.134	0.16	0.0838

Table 4.3. Genetic diversity indices for each of the 22 loci in our microsatellite marker panel in sika deer under the more stringent purity criteria ($0 \leq Q \leq 0.01$, n = 702) calculated in Cervus 3.0. Parameters are k, the number of alleles at each locus in each species, N, number of samples typed at each locus, H_o, observed heterozygosity, H_e, expected heterozygosity and Null, the frequency of null alleles at each locus.

Locus	k	N	H_O	H_E	Null
AGLA293	3	702	0.027	0.03	0.0418
BM4006	3	702	0.034	0.034	-0.0025
BM6438	7	686	0.43	0.596	0.1633
BM757	7	700	0.16	0.171	0.0275
BOVIRBP	6	702	0.041	0.046	0.0527
FCB193	6	702	0.192	0.232	0.1057
FSHB	15	700	0.29	0.394	0.1794
IDVGA29	5	690	0.151	0.191	0.1081
IDVGA55	7	702	0.272	0.3	0.0331
INRA5	6	701	0.187	0.191	0.0051
INRA6	3	699	0.05	0.054	0.0399
INRA131	6	701	0.244	0.298	0.1023
MM012	3	701	0.148	0.199	0.1403
RM12	2	702	0.009	0.009	-0.0002
RM188	12	696	0.578	0.637	0.0462
RM95	4	701	0.251	0.325	0.1339
RME025	7	694	0.076	0.087	0.0645
TGLA40	4	699	0.233	0.385	0.2442
TGLA126	5	700	0.504	0.542	0.0356
TGLA127	5	700	0.356	0.511	0.1811
TGLA337	7	685	0.396	0.674	0.2793
UWCA47	2	696	0.135	0.167	0.1057

4.4.1 *To assess population genetic structure within 'pure' red deer from across the British Isles using two different purity criteria (objective 1).*

Analysis 1: 'Pure' red deer (dataset 1, n = 2307) under the Q>0.95 criterion

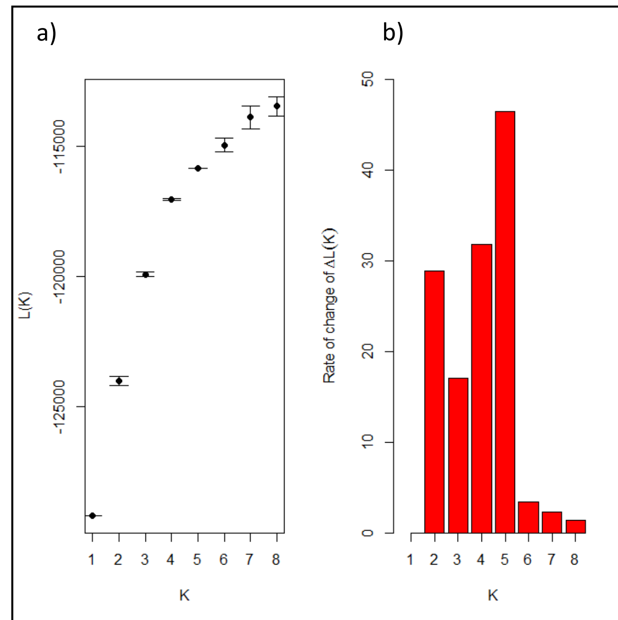


Figure 4.2. Assessment of the most likely number of populations in analysis 1 by Structure 2.3.3. a) shows the log-likelihood (with standard error) of the value of K (no. of populations) given the dataset and b) shows the rate of change in log likelihood between values of K. Results appear to support K = 5 as the most likely number of populations.

Whilst identifying the point at which the log likelihood starts to plateau is somewhat subjective, a case could be made for K = 5, with an average Ln Pr (X|K) (natural logarithm of the probability of data X, conditional on K) of -115849.66 (s.d. 9.83) (Figure 4.2a). K = 5 is also supported by the use of Evanno's rate of change approach (Evanno *et al.* 2005)(Figures 4.2b). At this value of K, the five clusters roughly comprise most of the English parks, the Lake District, Co. Kerry, Co. Galway, Arran and Rum (cluster 1), Kintyre (cluster 2), the North Highlands and Windsor Park (cluster 3), Islay and Jura (cluster 4) and Harris and Lewis (cluster 5), as shown in Figure 4.3. The remaining sites are admixed (Figure 4.3a). The barrier formed by the Great Glen reported by Perez-Espona *et al.* (2008) was recovered by this independent set of markers (Figure 4.3b). Posterior allele frequencies for population clusters for analysis 1 at K = 5 are given in Appendix Table 4.A2. The variation in the log likelihood generated during

replicated simulations at the same value of K may be attributed to slight variation in the sampling (or "mixing") of the Markov chain, as part of the Bayesian analysis, when converging on the posterior distribution of each of the required parameters (Pritchard *et al.* 2000).

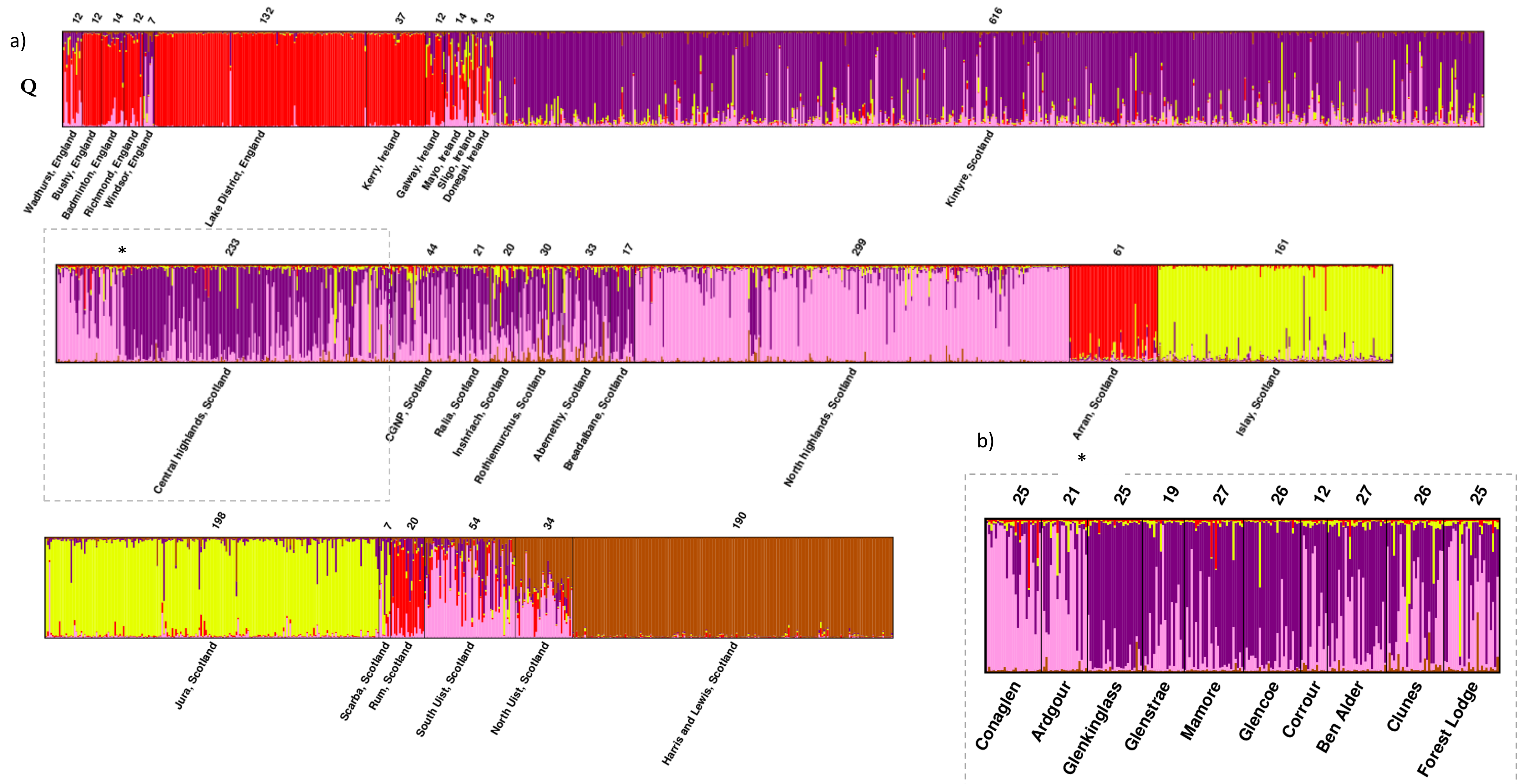


Figure 4.3. a) Bar chart showing the results of *analysis 1* in STRUCTURE at K = 5 for each individual red deer across Britain and Ireland (n = 2307) which met the 0.95 ≤Q ≤ 1 criteria when analysed in a larger dataset with sika and wapiti. The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to each of the five clusters. Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis and b) a zoomed-in image of the samples obtained from the central highlands. The estates from which samples were obtained are on the x-axis; the Great Glen (indicated by *) is positioned between the first two estates (Conaglen and Ardgour) and the remaining estates and is associated with a shift in population cluster membership.

Analysis 2: 'Pure' red deer (dataset 2, n = 2201) under the Q≥0.99 criterion

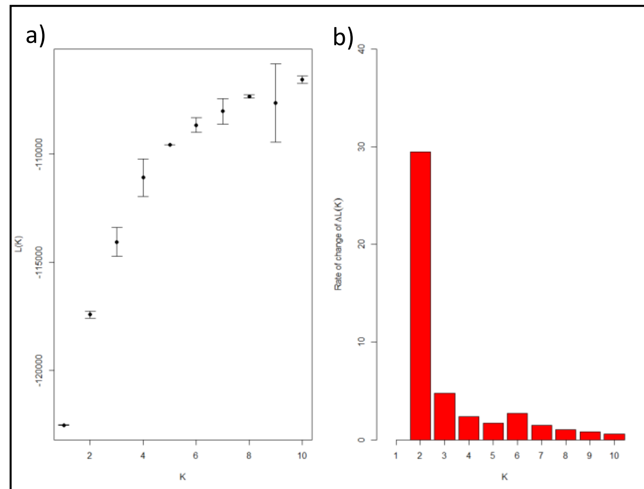


Figure 4.4. Assessment of the most likely number of populations in analysis 2 by Structure 2.3.3. a) Shows the log-likelihood (with standard error) of the value of K (number of populations) given the dataset and b) shows the rate of change in log likelihood between values of K.

Whilst the plot of log likelihoods does not pinpoint a particular value of K, Evanno's rate of change approach (29.50) strongly suggests K = 2 (Ln Pr (X|K) = -117420.86, s.d. 166.43, Figure 4.4). Under this structure, the Harris and Lewis population appears distinct from all other populations in Scotland, England and Ireland (Figure 4.5). Posterior allele frequencies for population clusters at K = 2 are given in Appendix Table 4.A3.

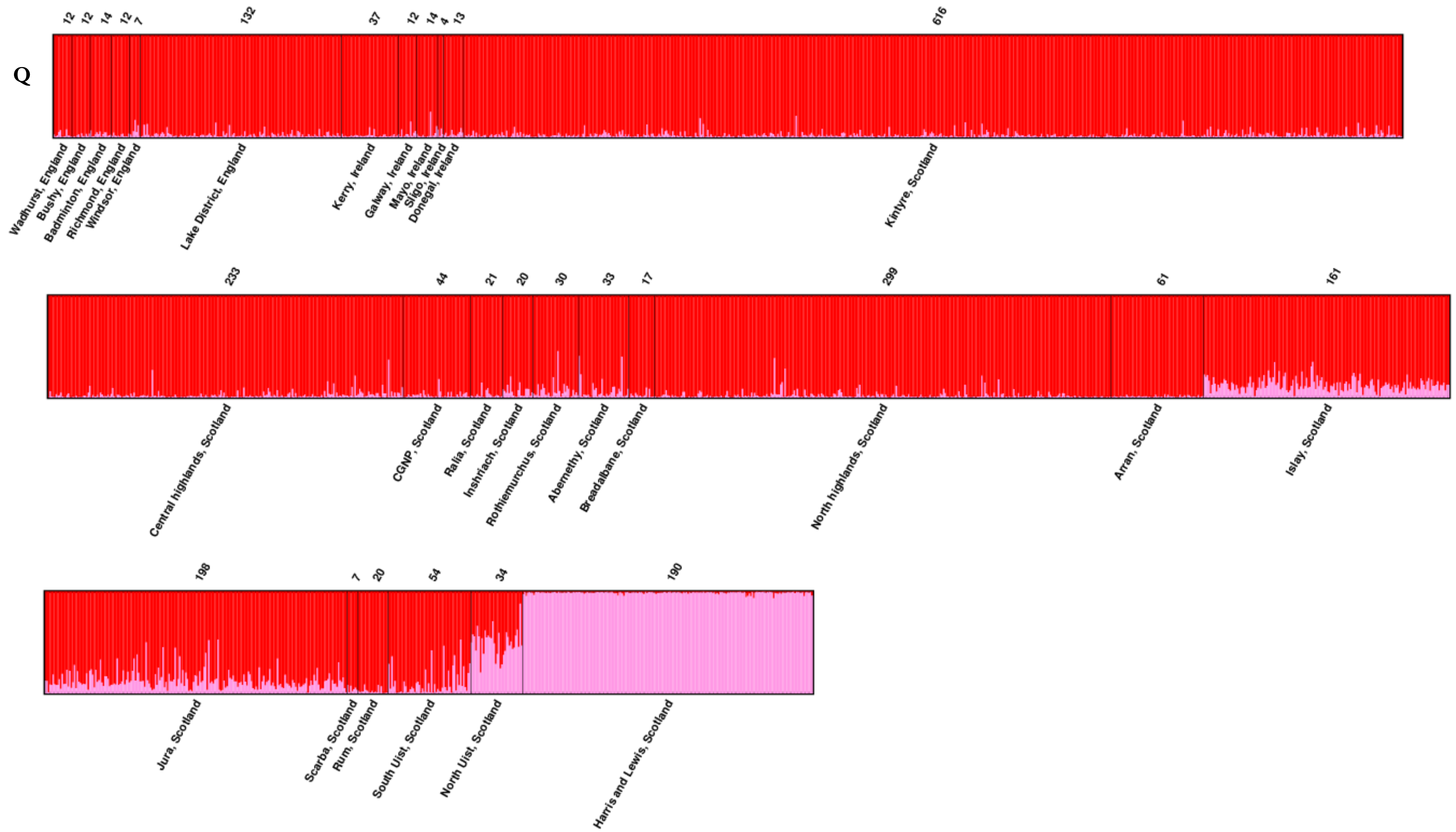


Figure 4.5. Bar chart showing a) the results of *analysis 2* in STRUCTURE at K = 2 for each individual red deer across Britain and Ireland (n = 2201) which met the 0.99 ≤ Q ≤ 1 criteria when analysed in a larger dataset with sika and wapiti. The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to each of the five clusters. Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis.

Analysis 3: Analysis of deer under the less (dataset 1, n = 2307) and more (dataset 2, n = 2201) stringent definitions of 'pure' red deer using DAPC

During analysis, DAPC assigns each value of K a Bayesian Information Criterion (BIC), the lowest value of which should correspond to the optimal number of clusters. BIC values produced for dataset 1 and dataset 2 are shown in Figure 4.6a and and b respectively. Despite slightly different scales on the y-axis, the likelihood plots appear similar to each other and the optimal value of K is not obvious. Using the results from the previous analysis in Structure, spatial clustering at K=2 to K=5 was explored. The optimal number of principal components retained was 50 for dataset 1 and 44 for dataset 2.

The populations from which each individual came ("ori") was then tabulated against the inferred clusters ("inf") at K = 2 (Figure 4.6c) and K = 5 (Figure 4.6d) for dataset 1 only. This was to explore the inference of population assignment by the software in the less stringent dataset at the two values of K that Structure 2.3 had shown support for.

Lastly, clusters were visualised at K = 5 for both dataset 1 (Figure 4.6e) and K = 2 for dataset 2 (Figure 4.6f). Cluster numbers are not consistent between colour plots (e.g. Cluster 1, 2 etc.) but the colours of the plots themselves are (e.g. red, purple etc.). Individuals in each analysis were located on the factorial planes by coordinates determined by the principal component analyses. Inferred clusters in these plots largely encompass individuals from the Lake District and most of the English parks (green), Kintyre, Co. Kerry and N.W. Ireland (orange), North highlands (red), Islay and Jura (dark blue) and Harris and Lewis (purple) (Figure 4.6e, f). Whilst the Lake District (green) and Harris and Lewis (purple) clusters are relatively distinct the others including central Scotland, the Irish sites, the English sites, Kintyre and the North highlands are only subtly differentiated and relatively admixed.

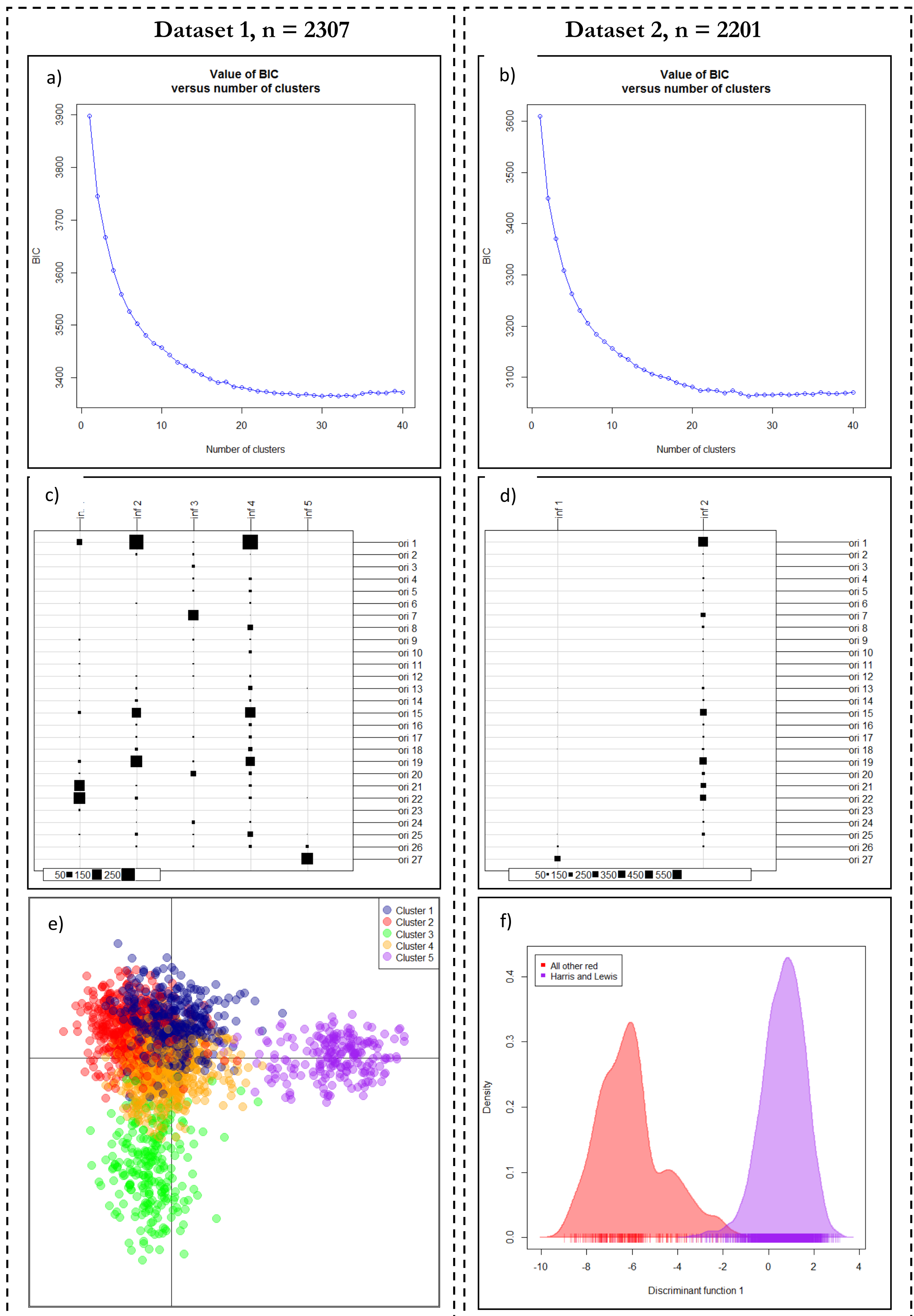


Figure 4.6. a) BIC values against K for a) dataset 1 (n = 2307) and b) dataset 2 (n= 2201). Results were tabulated between the inferred clusters ("inf") and the originally assigned ("ori") clusters for c) dataset 1 at K = 5 and d) dataset 2 at K = 2. Original populations were: ori1, Kintyre; ori2, Wadhurst; ori3, Bushy; ori4, Badminton; ori5, Richmond; ori6, Windsor; ori7, Cumbria; ori8, Co. Kerry; ori9, Co. Galway; ori10, Co. Mayo; ori11, Co. Sligo; ori12, Co. Donegal; ori13, Abernethy; ori14; Breadalbane; ori15, Central Scotland; ori16, Ralia; ori17, Inshriach; ori18, Rothiemurchus; ori19, North highlands; ori20, Arran; ori21, Islay; ori22, Jura; ori23, Scarba; ori24, Rum; ori25, South Uist; ori26, North Uist, ori27; Harris and Lewis. Lastly, the scatter plot of inferred clusters is given for e) dataset 1 at K = 5 and f) dataset 2 at K = 2. Inferred clusters in e) largely encompass individuals from the Lake District and most of the English parks (green), Kintyre, Co. Kerry and N.W. Ireland (orange), North highlands (red), Islay and Jura (dark blue) and Harris and Lewis (purple), whilst the distributions in f) represent all other red populations (red) and the Harris and Lewis population (purple).

4.4.2 *To assess population genetic structure within 'pure' sika from across the British Isles using two different purity criteria (objective 2).*

Analysis 4: 'Pure' sika (dataset 3, n = 752) under the Q<0.05 criterion

The logarithmic likelihoods calculated in Structure suggested K = 2 is the smallest number of genetic clusters that is optimal to describe the majority of the population structure amongst the 'pure' sika animals, with an average log likelihood of -7796.82 (s.d. 4.99) and a rate of change of 270.31 (Figure 4.7). At this value of K, it appears that the sika in Co. Kerry and Kintyre cluster together (cluster 1) while the sika in Co. Wicklow, parts of the North highlands and the English sika form a second cluster (cluster 2) (Figure 4.8). Although sika from the south of England predominantly group with cluster 2 they show substantial admixture with cluster 1. Sika from the North Highlands, although also quite admixed, show blocks of animals assigned to cluster 1 (predominantly around South Loch Ness and Moriston) and those with cluster 2 (remaining sites mostly west of Loch Ness). The relative integrity of sika in Kintyre is apparent against the more admixed situation in the rest of Scotland and Ireland. Posterior allele frequencies for population clusters at K = 2 are given in Appendix Table 4.A4.

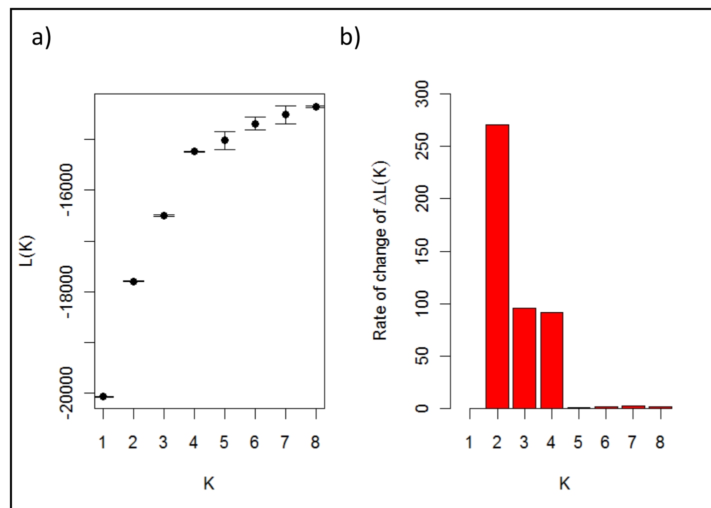


Figure 4.7. Assessment of the most likely number of populations in analysis 4 by Structure 2.3.3. a) shows the log-likelihood (with standard error) of the value of K (number of populations) given the dataset and b) shows the rate of change in log likelihood between values of K. Results support K = 2 as the most likely number of populations.

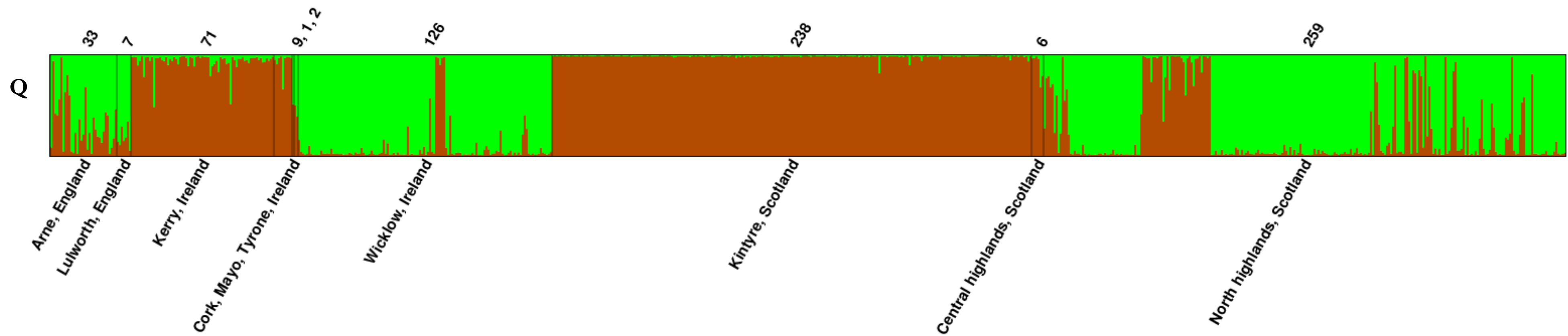


Figure 4.8. Bar chart showing a) the results of *analysis 4* in STRUCTURE at K = 2 for each individual sika deer across Britain and Ireland (n = 752) which met the 0.95 ≤ Q ≤ 1 criteria when analysed in a larger dataset with red and wapiti. The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to each of the two clusters (green and brown). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis.

Analysis 5: 'Pure' sika (dataset 4, n = 702) under the Q≤0.01 criterion

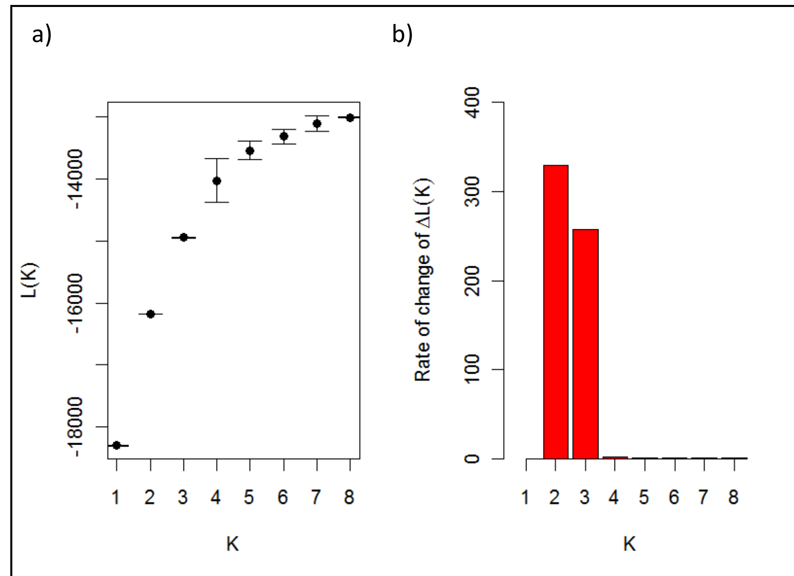


Figure 4.9. Assessment of the most likely number of populations in analysis 5 using Structure 2.3.3. a) shows the log-likelihood (with standard error) of the value of K (number of populations) given the dataset and b) shows the rate of change in log likelihood between values of K.

The logarithmic likelihoods calculated in Structure revealed K = 2 was the smallest number of genetic clusters that is optimal to describe the majority of the population structure amongst the 'pure' sika animals in analysis 5, with an average log likelihood of -16182.24 (s.d. 4.52) and a rate of change of 329.5 (Figure 4.9). This supports the outcome of analysis 4 i.e. that the population structure is best explained by K = 2; however, this is closely followed by support for three population clusters in this dataset (which was not so apparent in analysis 4). When K = 3, southern English and Irish sika become differentiated as a cluster (Figure 4.10).

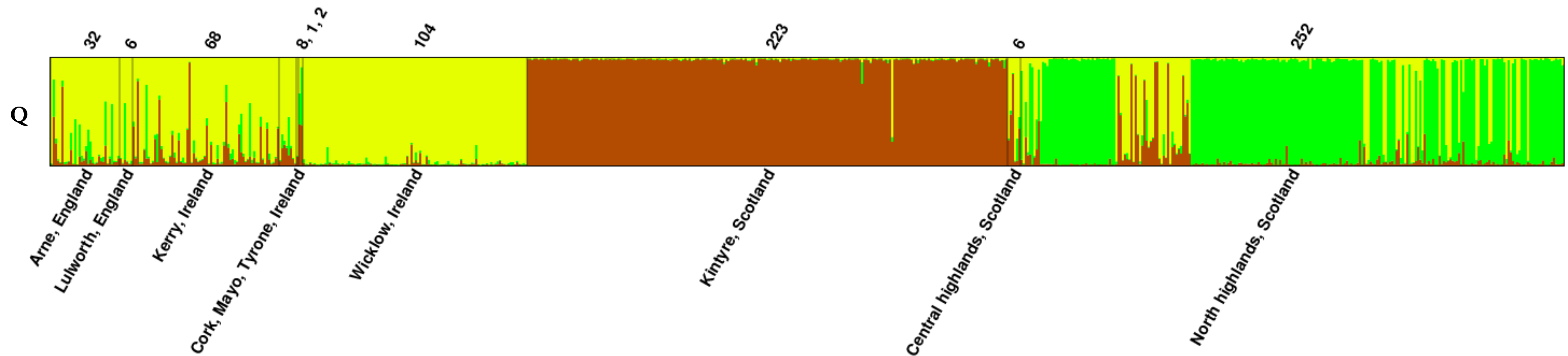


Figure 4.10. Bar chart showing a) the results of *analysis 5* in STRUCTURE at K = 4 for each individual sika deer across Britain and Ireland (n = 702) which met the 0.99 ≤ Q ≤ 1 criteria when analysed with red and wapiti. The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to each of the three clusters (yellow, green and brown). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis.

Analysis 6: Analysis of deer under the less (dataset 3, n=752) and more (dataset 4, n=702) stringent definitions of 'pure' sika deer using DAPC

The BIC by K plot produced for dataset 3 (Figure 4.11a) and that for dataset 4 (Figure 4.11b) appear exceptionally similar and as for red deer, the optimal value of K is not obvious. Using the results from the previous analysis in Structure, K values of two and three were explored. The optimal number of principal components retained was 22 for dataset 3 and 11 for dataset 4.

The populations from which each individual came ("ori") was then tabulated against the inferred clusters ("inf") at K = 2 (Figure 4.11c) and K = 3 (Figure 4.11d) for dataset 3 only. This was to explore the inference of populations by the software in the less stringent dataset at the two values of K that Structure 2.3 had shown support for.

Lastly, clusters were visualised at K = 3 for both dataset 1 (Figure 4.11e) and dataset 2 (Figure 4.11f); cluster numbers and colours are consistent between plots. Individuals in each analysis were located on the factorial planes by coordinates determined by the principal component analyses. Inferred clusters in these plots largely encompass individuals from Co. Wicklow, Co. Kerry, Co. Cork and the southern English sites (dark blue, cluster 1), the North highlands (green, cluster 2) and Kintyre (purple, cluster 3) (Figure 4.11 e, f). Individuals from the remaining sites including NW Ireland and central Scotland were relatively admixed and few in number.

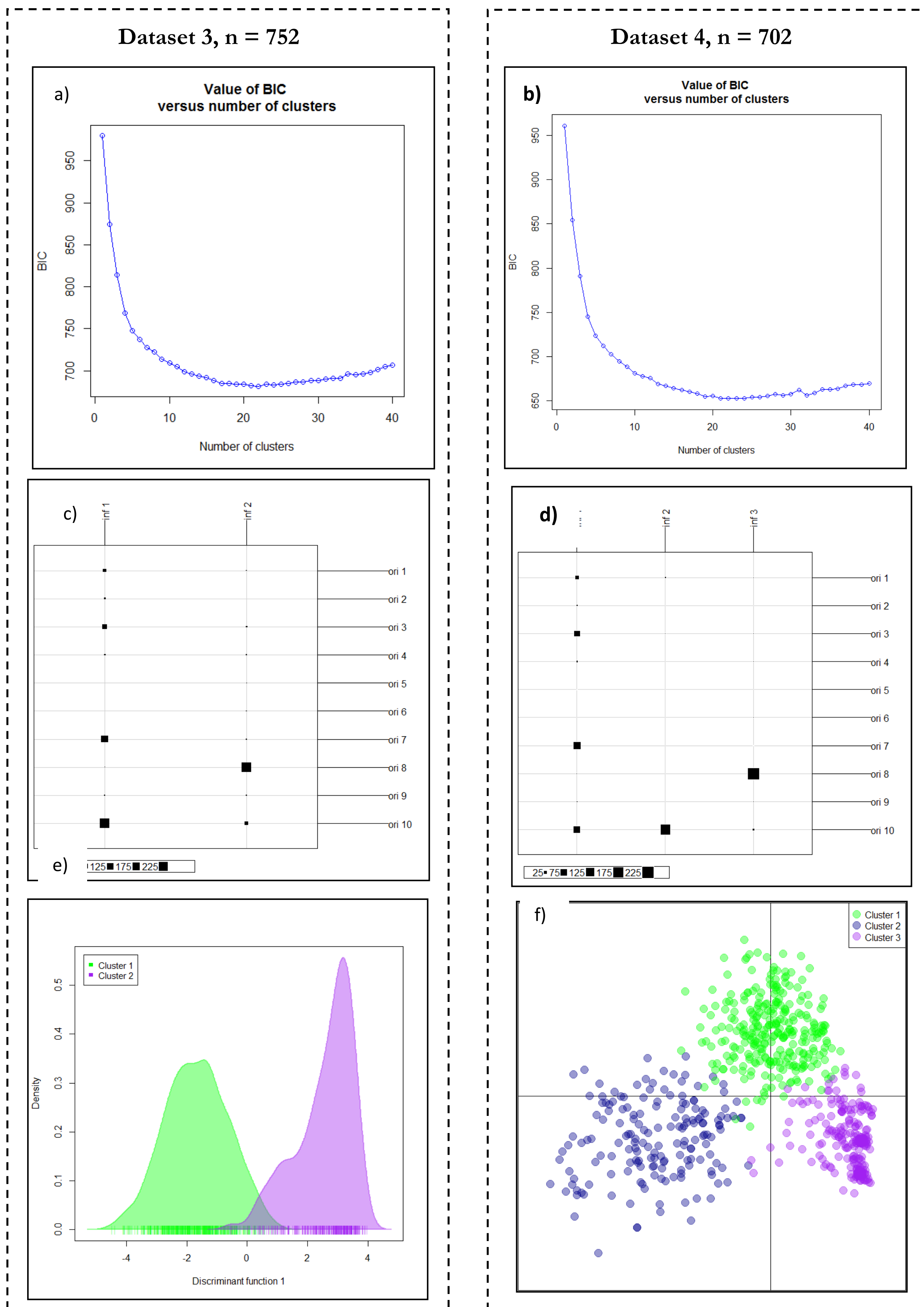


Figure 4.11. BIC values against K for a) dataset 3 (n=752) and b) dataset 4 (n=702). Results were tabulated between the inferred clusters ("inf") and the originally assigned ("ori") clusters for c) dataset 3 at K = 2 and d) dataset 4 at K = 3. Original populations were: ori1, Arne; ori2, Lulworth; ori3, Co. Kerry; ori4, Co. Cork; ori5, Co. Mayo; ori6, Co. Tyrone; ori7, Co. Wicklow; ori8, Kintyre; ori9, Central Scotland; ori10, North highlands. Lastly, the scatter plot of inferred clusters is given for e) dataset 3 at K = 2 and f) dataset 4 at K = 3. In 12 e) inferred clusters largely encompass individuals from Co. Wicklow, the southern English sites and the North highlands (green) and Kintyre, Co. Kerry and Co. Cork (purple)., whilst in f) the Kintyre group remains distinct (purple) and Co. Wicklow, Co. Kerry, Co. Cork and the southern English sites form a cluster (dark blue) whilst the North highlands forms another (green).

4.4.3 *To determine whether individuals removed by the more stringent purity criteria are introgressed hybrids or carry non-diagnostic alleles shared by the parental taxa (objective 3)?*

Of the 106 red deer removed from analysis by the purity criterion $0.95<Q<0.99$, 81 (76.4%) were from Kintyre (Table 4.4). Kintyre is the one area in all our surveys where introgression into red deer is already known to be extensive in places (Senn & Pemberton, 2009; Chapter 2). Furthermore, the frequency of sika-specific alleles was significantly greater in the 106 red deer with $0.95<Q<0.99$ than in the 2201 red animals with $Q \geq 0.99$ (Wilcoxon test, $p = 4.631 \times 10^{-5}$, Figure 4.12).

Of the 50 sika removed from analysis by the purity criterion $0.01<Q<0.05$, 22 (44%) were from Co. Wicklow, Ireland and 15 (30%) were from Kintyre (Table 4.4). Wicklow and Kintyre are the two areas in all our surveys where introgression into sika is already known to be extensive (Chapters 2 and 3). The frequency of red-specific alleles was significantly greater in the 50 sika with $0.01<Q<0.05$ than the 702 sika animals with $Q \leq 0.01$ (Wilcoxon test, p = 8.155e-06, Figure 4.13).

Table 4.4. Animals removed when moving from the less to more stringent dataset in red-like animals (those with $0.95<Q<0.99$, column 2) and sika-like animals (those with $0.01<Q<0.05$, column 3). The percentage of the total animals removed in each species dataset (red = 106; sika = 50) is given in parenthesis.

	No. red-like animals ($0.95<Q<0.99$) removed from sample *(% total removed)*	**No. sika-like animals ($0.01<Q<0.05$) removed from sample** *(% total removed)*
Southern England	0 (0.0)	2 (4)
Lake District, England	13 (12.3)	0 (0.0)
Wicklow, Ireland	0 (0.0)	22 (44)
Kerry, Ireland	0 (0.0)	3 (6)
Cork, Ireland	0 (0.0)	1 (2)
Mayo, Ireland	2 (1.9)	0 (0.0)
Kintyre, Scotland	81 (76.4)	15 (30)
North highlands, Scotland	5 (4.7)	7 (14)
Hebrides, Scotland	5 (4.7)	0 (0.0)

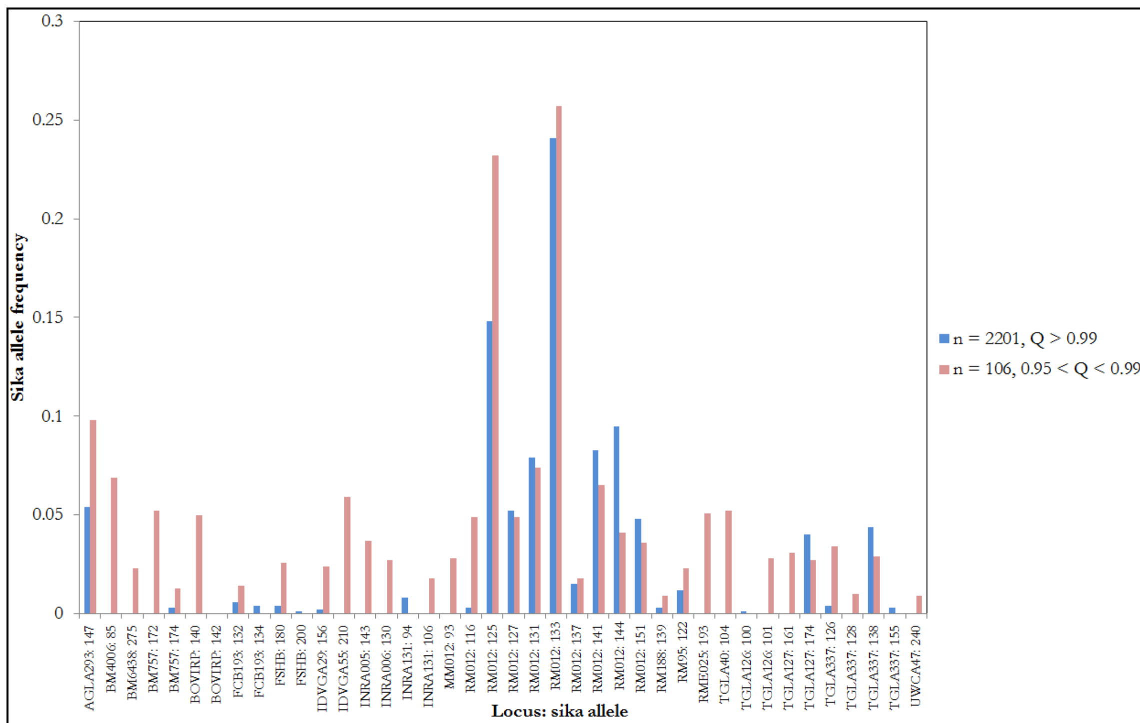


Figure 4.13. Frequency of red-specific alleles in sika-like animals with 0.01<Q<0.05 (n = 50, blue) compared to those in sika-like animals with Q≤ 0.01 (n = 702, green).

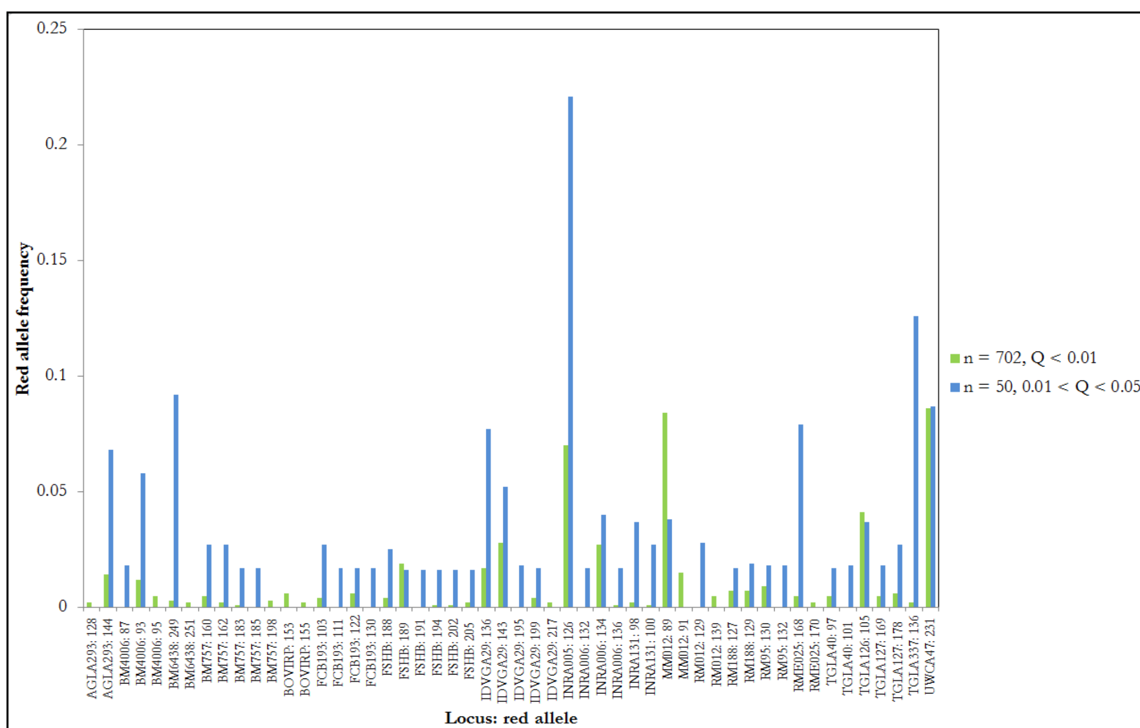


Figure 4.12. Frequency of sika-specific alleles in red-like animals with 0.95<Q<0.99 (n = 106, pink) compared to those in red-like animals with Q≥0.99 (n = 2201, blue).

After the removal of the animals presented in Table 4.4, the mean allelic diversity was assessed within the remaining red deer and sika deer at each sampling site (Tables 4.5 and 4.6 respectively) and averaged across all loci (Table 4.7) for comparison with the markers used by Perez-Espona et al. (2013) (see also supplementary table S1 in Pérez-Espona *et al.* (2009b)). It is important to note the large variation in sample sizes between populations presented in Tables 4.5 and 4.6 complicates making direct comparisons of allelic diversity between sites.

Table 4.5. The number of animals and mean number of alleles per microsatellite locus for each of the populations from which red deer that met the more stringent purity criteria (dataset 2, n = 2201) were sampled as well as the expected and observed heterozygosity from each locus based on analysis in Cervus 3.0.

Population	*Sample Size*	*Mean No. alleles per locus*	H_E	H_O
English park deer	57	5.59	0.50	0.43
Lake District, England	119	4.82	0.46	0.47
Co. Kerry, Ireland	37	3.27	0.39	0.35
Co. Galway, Ireland	12	3.86	0.51	0.50
Co. Mayo, Ireland	12	4.86	0.53	0.50
Co. Sligo, Ireland	4	2.68	0.50	0.49
Co. Donegal, Ireland	13	3.95	0.54	0.46
Argyll, Scotland	535	6.91	0.55	0.52
Central highlands, Scotland	398	8.05	0.56	0.52
North highlands, Scotland	294	7.45	0.58	0.55
Arran, Scotland	60	4.32	0.51	0.43
Islay, Scotland	161	5.77	0.49	0.48
Jura, Scotland	197	6	0.48	0.47
Scarba, Scotland	7	3.32	0.49	0.48
Rum, Scotland	20	4.27	0.45	0.44
North and South Uist, Scotland	85	4.95	0.50	0.47
Harris and Lewis, Scotland	190	3.77	0.33	0.32

Table 4.6. The number of animals and mean number of alleles per microsatellite locus for each of the populations from which sika deer that met the more stringent purity criteria (dataset 4, n = 702) were sampled as well as the expected and observed heterozygosity from each locus based on analysis in Cervus 3.0.

Population	*Sample Size*	*Mean No. alleles per locus*	H_E	H_O
Lulworth, England	32	2.36	0.28	0.27
Arne, England	6	1.82	0.26	0.23
Co. Kerry, Ireland	68	1.68	0.16	0.15
Co. Cork, Ireland	8	1.64	0.19	0.15
Co. Mayo, Ireland	1	1.05	0.09	0.09
Co. Tyrone, Ireland	2	1.18	0.10	0.14
Co. Wicklow, Ireland	104	3.09	0.32	0.31
Argyll, Scotland	223	2.73	0.11	0.10
Central highlands, Scotland	6	1.55	0.17	0.18
North highlands, Scotland	252	3.36	0.33	0.29

Table 4.7. Genetic diversity indices for the 22 microsatellite markers in our study within the red (Q≥0.99) and sika (Q≤0.01) datasets (2 and 4) and within the 15 microsatellite markers used to assess red deer populations by Pérez-Espona *et al.* (2013). Indices include: N alleles = average number of number of alleles at each locus in the marker panel, H_O = Average observed heterozygosity across all loci, H_E = Average expected heterozygosity across all loci, PIC = Average Polymorphic Information Content across all loci, F(Null) = Mean estimated null allele frequency (no information for Pérez-Espona *et al.* (2013)).

	N alleles	H_O	H_E	*PIC*	*F(Null)*
Red (Q≥0.99)	9.36	0.49	0.56	0.53	0.08
Red, Perez-Espona et al., 2012	16.80	0.76	0.73	0.81	NA
Sika (Q≤0.1)	5.68	0.22	0.28	0.25	0.09

4.5 Discussion

4.5.1 To assess population genetic structure within 'pure' red deer from across the British Isles using two different purity criteria (objective 1).

A number of previous studies have investigated the genetic population structure of red deer in Europe (Frantz *et al.* 2008; Gyllensten *et al.* 1983; Kuehn *et al.* 2003) and Scotland (Nussey *et al.* 2005; Pérez-Espona *et al.* 2013). The most comprehensive of these to date was a recent study that used 15 microsatellite loci to assess genetic population structure amongst 1152 red deer from across the Scottish mainland (14 estates) and islands (7 on the west coast) in order to determine the extent to which non-native red deer and wapiti introductions have impacted on it (Pérez-Espona *et al.* 2013). It is important to note that the latter study did not include sika deer or red deer from areas where this project has shown hybridisation to occur. The results from that analysis can be compared and contrasted to our own. The two studies are based on a similar approach as they have overlapping sampling areas and 235 of the same individuals from the central highlands; however, they have only two markers in common (BM757, RM188).

Pérez-Espona *et al.* (2013) concluded that the population structure amongst the red deer examined in their study could be best explained by both K = 7 and K = 10. This structure was largely driven by differentiation amongst the Hebridean islands, the clustering of Scottish mainland samples according to particular landscape features and the inclusion of wapiti (Pérez-Espona *et al.* 2013; Pérez-Espona *et al.* 2008). However, the most likely structure inferred using Evanno's approach (Evanno *et al.* 2005), was K = 5; in which case clusters were (1) west of the Great Glen, (2) east of the Great Glen, (3) Harris and Lewis, (4) Arran, Rum, English deer park and Eastern European and German samples and (5) wapiti (see supplementary material, figure S3c, (Pérez-Espona *et al.* 2013). The application of Evanno's approach to our red deer dataset at the less stringent purity criteria (Q>0.95) also suggested K = 5 was most likely, of which the clusters contained the Lake District, the English parks, Arran and Rum (cluster 1, Figure 4.3 & 4.6d, e, f) and Harris and Lewis (cluster 5, Figure 4.3 and 4.6), which could parallel clusters 3 and 4 from Pérez-Espona *et al.* (2013). This corroborates the affiliation between the English park deer and those on Rum and Arran, which are known to have been restocked from English deer parks (Nussey *et al.* 2006; Pérez-Espona *et al.* 2013) and the distinctiveness of the Harris and Lewis population (Chapter 2). Amongst our red deer dataset at the less stringent purity criteria (Q>0.95) we were also able to locate the genetic signature of the Great Glen; a large landscape feature which Perez-Espona *et*

al. (2008) identified as causing significant population structure in nuclear data (see * in Figure 4.3b).

Due to the inclusion of more deer from different sites, some of our conclusions go beyond those of Pérez-Espona *et al.* (2013). The affinity of the Lake District with the English parks and Rum and Arran are likely due to the use of red deer from Knowsley, Lancashire and from the southern English parks to restock Rum and Arran around the 1850s (Whitehead 1964). Similarly, the inclusion of the Irish animals in this cluster reflects the movement of continental, Scottish, English and Irish stocks to found populations at Screebe Estate and Connemara Estate, Co. Galway (1980s, 1990s respectively) and Glenveagh, Co. Donegal (1891), as well as the introduction of red deer from British parks to the north west of Ireland (19th century) (Carden *et al.* 2010; McDevitt *et al.* 2009a).

The relative distinctness of Islay and Jura (cluster 4) and Harris and Lewis (cluster 5), is consistent with the greater impact of introductions to island systems (and the long periods between such events with no introductions) than to a more continuous mainland (Pérez-Espona *et al.* 2013). The other islands in the Hebridean archipelago are more admixed and have probably been subject to a more continuous succession of translocations. North Uist, for example, is known to have received red deer from Kerry in the early 1900s, whilst South Uist has been subject to relatively recent introductions in 1970s-1980s, following the near-extinction of its population in the later part of the 18th century (Carden *et al.* 2010; Pérez-Espona *et al.* 2009a; Whitehead 1964). The highland region of mainland Scotland is also relatively mixed - it does contain the highest densities of red deer, in a large continuous population with free movement, such that genetic drift is less likely and may contribute to the weaker population structure (Pérez-Espona *et al.* 2009b).

When we tightened the purity criteria of our red deer dataset to $Q \geq 0.99$, the most likely population structure obtained was, in fact, that between Harris and Lewis and all remaining populations of red deer (Figure 4.5, 4.6c). Whilst Pérez-Espona *et al.* (2013) did not present statistical support for $K = 2$ explicitly, their analysis drew attention to the low genetic diversity in the Harris and Lewis population (an average of 4.0 ± 1.46 alleles per locus amongst 34 animals from Harris and Lewis, less than half the average sampled from the Scottish mainland in that study) and, similarly, we recorded 3.77 ± 2.51 alleles per locus amongst 190 animals from Harris and Lewis (less than two thirds of that from our Scottish mainland sample; Table 4.3 and supplementary table S1 from Pérez-Espona *et al.* (2013). The distinctiveness of the Harris and Lewis population has

been attributed to a severe bottleneck effect following the introduction of deer to this remote island (resident since at least the 16th century (Ratcliffe 1987)) despite the fact it was supplemented with Scottish mainland and English park animals in the mid to late 19th century (Pérez-Espona *et al.* 2013; Whitehead 1964). Situated approximately 50 miles off the coast of mainland Scotland and exposed to the Atlantic Ocean and Gulf Stream, the animals which survive here are also likely highly adapted to the cold and windy climatic conditions on this island (Pérez-Espona *et al.* 2013; Whitehead 1964). In addition, Whitehead (1964) describes the tendency of the red deer on Harris and Lewis to chew cast antlers due to the absence of bone-producing elements on the islands: another, geological aspect of the area that may lead to specific adaptation. Since around the start of the 20th century, red deer stocks on Harris and Lewis appear to have been in decline, the cause of which is likely attributed to extensive poaching, unproductive grounds with low carrying capacity for deer and persistent culling (Whitehead 1964). This will inevitably have lowered the effective population size of red deer and may have encouraged inbreeding on the island which would serve to homogenise population. The red deer on Harris and Lewis carry the same *C. elaphus* mitochondrial haplotype as the mainland red deer and carry no private red deer nuclear alleles, instead having the distribution of their nuclear allele frequencies skewed toward one or two red deer specific alleles at each locus. This, and the emergence of the Harris and Lewis red deer population during Structure analyses *before* the wapiti species in Chapter 2 and 4, suggest that this island population simply represents a small, restricted portion of the potential red deer genetic variation that causes it to become differentiated, rather than divergent branch of the *C. elaphus* phylogeny. Similarly, without the Harris and Lewis red deer population in Chapter 3, the red, sika and wapiti were best described by three clear clusters, highlighting the distinctness of this island population.

Overall, our analysis and the analysis by Pérez-Espona *et al.* (2013) have common findings, most notably they both show genetic differentiation of the Hebridean islands and genetic evidence likely to reflect known translocations and introductions. In addition, we show that tightening the purity criteria for red deer in our dataset simplifies the population structure as a consequence of removing introgressed alleles and/or non-diagnostic alleles that differentiate red deer clusters. This is discussed below in section 4.5.3.

4.5.2 To assess population genetic structure within 'pure' sika from across the British Isles using two different purity criteria (objective 2).

Previous studies have also looked into the population structure amongst sika deer, primarily in their native range in Japan but also within British sites (Goodman *et al.* 2001; Liu *et al.* 2003; Nagata *et al.* 1998). Animals translocated outside of Japan are suspected to have retained genetic diversity originally present in their native range, but it has now been lost in contemporary sika populations in Japan (Goodman *et al.* 2001). Here we carried out the first attempt at determining the genetic population structure amongst sika deer sampled from sites in Scotland, Ireland and England. When animals were assigned as 'pure' sika if they had $Q<0.05$, the population structure was best described by two clusters comprising sika from Co. Kerry, Co. Cork and Kintyre (cluster 1) and those from Co. Wicklow and the rest of the English sites and North highlands sites (cluster 2). The distinction between the sika from Co. Kerry and Co. Wicklow is consistent with the analyses performed by McDevitt et al. (2009) using eight microsatellite markers and with historical records. Relatively soon after their introduction to Powerscourt Park, Co. Wicklow in 1860, two sika hinds and a sika stag were translocated to Co. Kerry in 1864 (McDevitt *et al.* 2009a). It is thus unlikely they carried red alleles from hybridisation in the park (three to four hybrids observed by Powerscourt by 1884). Such a small founding group explains the low genetic diversity found in Kerry in this study (average number of alleles recorded 1.68 ± 0.89; Table 4.4). The sika in Cork cluster with those from Kerry, as they are suspected to have dispersed from the latter county (Chapter 3). Lower genetic diversity present in sika compared to red deer in Britain and Ireland has been validated by numerous studies using molecular approaches and is unsurprising given the genetic bottlenecks, founder events and predominance of drift this exotic species has experienced during introduction to unfamiliar territory (Goodman *et al.* 2001; McDevitt *et al.* 2009a; Senn & Pemberton 2009). In addition, habitat fragmentation and exploitation of sika in their native range of Japan may have reduced genetic diversity of this species, prior to its introduction to Britain and Ireland (Goodman *et al.* 2001).

Genetic affinity between populations may also be linked to separate introductions from the same source. The block of animals in the North highlands, for example, which cluster with Kerry and Kintyre are primarily based around South Loch Ness and Moriston; a region within which independent introductions of sika were made to both Aldourie and Glenmazeran in 1900 (Ratcliffe 1987). The source of this later introduction was Fawley Court from which sika have also been introduced to Kintyre in

the late 19th century and may explain the affiliation between these northern sika and those from Kintyre (Whitehead 1964). The remaining sika animals in the North Highlands which cluster with Wicklow may have derived from sika introduced west of Loch Ness (e.g. Achanalt; which themselves were introduced from Powerscourt in 1889 (Ratcliffe 1987)).

Tightening the criteria on the purity of sika animals to those with a Q≤0.01 similarly suggested the most likely population structure is best represented by two or (to a slightly lesser extent), three clusters. Use of DAPC supports and differentiates three clusters in a way consistent with Structure, namely Ireland and the southern English sites (1), the North highlands (2) and Kintyre (3) (Figure 4.10 and 4.11 e, f). Such differentiation amongst the sika populations may be attributed to the different introductions of this species during the late 19th and early 20th centuries (Ratcliffe 1987). Other molecular and morphological studies have exposed population variation in sika populations (Díaz *et al.* 2006; Goodman *et al.* 2001; Swanson 2000). Within Japan, Goodman *et al.* (2001) suggests that sika population structure is largely shaped by the effects of drift and mutation (occuring across an ancient timescale) combined with more recent anthrogenic-induced disruptions to gene flow and drift by activities such as habitat fragmentation, economic development and overexploitation (Goodman *et al.* 2001). Similar sorts of activities (e.g. patterns of forestry, urbanisation) may have influenced the amount of available habitat for sika across the British Isles, as well as the natural landscape features shown to affect red deer (Pérez-Espona *et al.* 2008).

4.5.3 To determine whether individuals removed by the more stringent purity criteria are introgressed hybrids or carry non-diagnostic alleles shared by the parental taxa (objective 3).

When applying stricter criteria concerning the purity of the parental populations, the majority of both red deer with low-level introgression and sika with low-level introgression which were removed were largely sampled from sites within which hybrid swarms have been identified (Table 4.2; Chapter 2 and 3). In addition, the frequency of sika-specific alleles in the subset of red deer removed which did not meet the stricter criteria (0.95<Q<0.99), was significantly greater than their frequency amongst those with Q≥0.99 (Figure 4.12). Similarly, the frequency of red-specific alleles in the subset of sika removed which did not meet the stricter criteria (0.01<Q<0.05), was significantly greater than their frequency amongst those with Q≤0.01 (Figure 4.13). Both these lines of evidence strongly suggest that the apparent population structure revealed when applying the threshold of Q>0.95 and Q<0.05 could be confounded by the influence of

sika and red deer introgression, respectively. The more stringent purity criteria adopted here was similarly adopted by McDevitt *et al.* (2009a) in their analysis of red and sika population structure in Ireland (using a different and non-diagnostic panel of markers) in order to increase the chance of identifying and removing hybrid animals from within-species population analysis. Overall, we have shown that the precise definition of a 'pure' animal of a particular species, prior to investigating within-species population structure, can determine the outcome. In this study Kintyre red deer was differentiated as a cluster when a less stringent purity definition was used for red deer, but then lost when the purity criteria was made more stringent. Therefore, care needs to be taken when investigating genetic population structure in the presence of potentially uneven levels of introgression from another taxon.

Inspection of levels of polymorphism within our most stringently pure datasets may also explain why our study found relatively low red deer population genetic structure compared to Pérez-Espona *et al.* (2013) and why sika showed relatively low population structure. Although the marker panel used by Pérez-Espona *et al.* (2013) is smaller than ours, the microsatellites are more polymorphic (Table 4.5) and may be more sensitive to population structure at smaller scales, explaining why seven to ten population clusters were deemed most likely amongst red deer in that study. Our marker panel was designed to differentiate red and sika and although most of the loci are polymorphic within species, they are together less powerful at resolving within-species population structure (Table 4.5). In conclusion, when investigating the within-species population genetic structure of species which are also involved in introgressive hybridisation, we would recommend using an appropriate number of suitably polymorphic molecular markers, a biologically stringent genetic purity criteria and large, representative sample sizes, especially when the results may have management and legislative consequences.

4.5.4 *To consider how can this information be used to benefit management of both species (objective 4).*

It is evident that, in addition to landscape features (e.g. the Great Glen), the impact of human-mediated introduction and translocation of con- and hetero-specific deer from various locations can impact population structure within-species as well as between species (Chapter 5). This is evident in the distinctiveness of the red deer on Harris and Lewis, resident since at least the mid-16th century with infrequent introductions made (Ratcliffe 1987), the affinity between the English parks with islands such as Arran and Rum, due to restocking activities between them and the differentiation amongst sika

populations likely influenced by the location of their introduction and the number of animals introduced. As well as site and number of founding individuals, the sex of those introduced and the purpose for doing so can be important; the introduction of females, for example, is likely to impact population growth and structure to a greater extent than males, due to philopatry and amount of polygyny and whether the introduction is being made to improve appearance or restock a population can influence its relative impact in the population or site to which they are introduced (Pérez-Espona *et al.* 2013). Within the native range of sika deer, recent anthropogenic disturbances to gene flow have modified their population structure and increased the influence of genetic drift (Goodman *et al.* 2001). This and previous studies highlight that the impact of introductions and translocations on within-species population structure should not be underestimated. Whilst the history of deer introductions into Scotland may be relatively well documented, not all translocations will have been recorded (Pérez-Espona *et al.* 2013); the identification of a highly divergent haplotype on the isle of Rum most closely related to the Corsican red deer (*C. elaphus corsicanus*), for example, was surprising considering there are no current records of introductions of this subspecies into Britain (Nussey *et al.* 2006). Overall, if the record-keeping and policing of deer translocations between sites is improved and regulated via appropriate scientific advice, this should benefit the management of red and sika populations.

Such an approach can also be taken when considering restocking or 'rescuing' a population from potential decline or extinction. The population of red deer on Harris and Lewis, evidently have limited genetic diversity. If it were ever apparent that red populations were in decline, with inbreeding depression as the possible cause (which would require evidence), genetic rescue by introduction from appropriate sources determined by genetics might be appropriate. Genetic tools have been similarly used to the benefit of other species' management; for example, to identify and counteract habitat fragmentation by providing migration corridors (Mech & Hallett 2001), to buffer against the effects of inbreeding in populations, such as the introduction of male adders (*Vipera berus)* to an isolated and inbred population of this species in Sweden (Madsen *et al.* 1999) and by guiding genetic recovery of populations such as the Mesola red deer (Zachos & Hartl 2011).

4.6 Acknowledgements

We thank the Forestry Commission Scotland rangers, including Kevin McKillop, Willie Lamont, Derick Macaskill, David Bain, staff of the Deer Commission Scotland (now Scottish Natural Heritage), Will Boyd Wallis and the Cairngorm National Park Authority

and the private estate owners and stalkers who cooperated within this project and provided samples. Thanks to Eleni Socratous, Eleanor A. M. Graham and Guy N. Rutty from the University of Leicester and Mike Thornley (Forestry Commission Chief Ranger) and the rangers of the Forestry Commission Cumbria for the samples they provided from the Lake District and to Elizabeth Mittell and Sarah MacDonald for their genotyping work on them. Thanks also to genotyping conducted by Elizabeth Heap, Sheena Morrissey and Phil Ellis. Lastly, thanks to the GenePool for their excellent service and NERC for funding.

4.7 Appendices

Table 4.A1. Posterior allele frequencies from analysis of dataset containing all pure red and sika from the British Isles under the less stringent purity criteria (n = 3059) at K = 2 and species specific allele assignment. An allele was not assigned to a species if its frequency was less than 1% (0.01) for both species. Alleles were assigned to a species (red = red, green = sika) if its frequency in the other species was 0 or if its frequency was five-fold larger than the other species.

Locus	% Missing data	Allele Size	Estimated allele frequency in Red	Estimated allele frequency in Sika	Allele species - specific assignment
AGLA293	1.50%	128	0.08	0.001	R
		144	0.775	0.015	R
		147	0.054	0.974	S
BM4006	0.40%	85	0.001	0.977	S
		87	0.1	0.001	R
		93	0.723	0.014	R
		95	0.13	0.005	R
BM6438	1.80%	249	0.549	0.007	R
		251	0.196	0.001	R
		253	0.096	0	R
		257	0.004	0	NA
		259	0	0.066	S
		261	0.071	0	R
		263	0	0.002	NA
		265	0	0.416	S
		273	0	0.021	S
		275	0.001	0.384	S
BM757	0.20%	160	0.07	0.006	R
		162	0.543	0.002	R
		164	0.007	0	NA
		172	0	0.896	S
		174	0.003	0.079	S
		179	0.05	0	R
		183	0.075	0.001	R
		185	0.045	0	R
		187	0.04	0	R
		189	0.002	0	NA
		196	0	0	NA
		197	0	0	NA
		198	0.052	0.003	R
		200	0.066	0	R
		202	0.01	0.005	NA
		210	0.004	0	NA
BOVIRP	0.40%	140	0.001	0.958	S
		142	0	0.012	S
		144	0	0.001	NA
		145	0	0.001	NA
		147	0.059	0	R
		149	0.059	0	R
		151	0.181	0	R
		153	0.376	0.005	R
		155	0.055	0.002	R
		157	0.192	0	R
		159	0.023	0	R
		163	0	0	NA
FCB193	1.90%	101	0.006	0	NA
		103	0.036	0.004	R
		105	0.001	0	NA
		107	0.088	0	R
		109	0.152	0	R
		111	0.021	0	R
		113	0.245	0	R
		115	0.008	0	NA
		118	0.04	0	R
		120	0.103	0	R
		122	0.098	0.006	R
		124	0.041	0	R
		126	0.01	0.026	NA
		128	0.011	0.045	NA
		130	0.058	0	R
		132	0.006	0.831	S
		134	0.004	0.038	S
		140	0.003	0	NA
		141	0	0	NA
		143	0.012	0	R
FSHB	0.70%	179	0	0.007	NA
		180	0.004	0.718	S
		181	0	0.092	S
		182	0	0.034	S
		183	0	0	NA
		184	0.047	0	R
		185	0.181	0	R
		186	0.002	0	NA
		187	0.003	0	NA
		188	0.125	0.004	R
		189	0.127	0.018	R
		190	0.008	0.021	NA
		191	0.087	0	R
		192	0.018	0	R

		193	0	0	NA
		194	0.021	0.001	R
		195	0	0	NA
		196	0.009	0	NA
		197	0.006	0	NA
		198	0.079	0	R
		199	0.021	0.007	NA
		200	0.001	0.011	S
		201	0.007	0	NA
		202	0.025	0.001	R
		203	0.02	0	R
		204	0.012	0	R
		205	0.077	0.002	R
		206	0.031	0	R
		207	0.037	0.01	NA
		208	0.002	0.001	NA
		209	0	0.001	NA
		210	0.01	0	R
		211	0.002	0	NA
IDVGA29	2.10%	136	0.66	0.019	R
		143	0.32	0.028	R
		145	0	0.026	S
		146	0	0.025	S
		156	0.002	0.832	S
IDVGA55	1.80%	191	0.04	0	R
		193	0.08	0	R
		195	0.218	0	R
		197	0.295	0	R
		199	0.208	0.004	R
		202	0.023	0	R
		204	0.039	0.032	NA
		210	0.001	0.797	S
		212	0	0.08	S
		214	0	0.042	S
		215	0	0.001	NA
		217	0.037	0.001	R
		219	0.015	0	R
		221	0	0	NA
INRA005	0.20%	124	0	0.027	S
		126	0.983	0.078	R
		129	0	0.001	NA
		136	0	0.003	NA
		137	0	0.001	NA
		143	0.001	0.875	S
INRA006	0.20%	128	0	0	NA
		130	0.001	0.955	S
		132	0.039	0.001	R
		134	0.693	0.027	R
		136	0.23	0.001	R
		138	0.012	0	R
INRA131	0.00%	87	0	0	NA
		92	0.039	0	R
		94	0.008	0.093	S
		98	0.603	0.003	R
		100	0.227	0.001	R
		102	0.069	0	R
		104	0.037	0	R
		106	0	0.784	S
		113	0	0.053	S
		115	0	0.009	NA
INRA131	0.00%	87	0	0	NA
		92	0.039	0	R
		94	0.008	0.093	S
		98	0.603	0.003	R
		100	0.227	0.001	R
		102	0.069	0	R
		104	0.037	0	R
		106	0	0.784	S
		113	0	0.053	S
		115	0	0.009	NA
MM012	0.10%	89	0.75	0.08	R
		91	0.216	0.013	R
		93	0	0.837	S
		95	0	0	NA
		97	0.001	0	NA
		104	0	0	NA
RM012	0.50%	125	0.152	0	R
		151	0.047	0	R
		127	0.052	0	R
		133	0.242	0	R
		131	0.08	0	R
		120	0.007	0	NA
		141	0.082	0	R
		137	0.015	0	R
		144	0.093	0	R
		129	0.085	0.001	R
		139	0.081	0.004	R
		116	0.003	0.99	S
RM188	0.80%	113	0.006	0	NA
		115	0.019	0	R
		117	0.035	0	R
		121	0.001	0	NA
		123	0.047	0	R
		125	0.076	0	R
		127	0.401	0.006	R
		129	0.202	0.007	R
		131	0.032	0	R
		132	0.033	0	R
		133	0.001	0	NA
		134	0.038	0	R
		137	0.042	0	R
		139	0.003	0.026	S
		141	0	0.007	NA
		143	0.001	0.528	S

		145	0	0.015	S
		153	0	0.022	S
		161	0	0.175	S
		163	0	0.003	NA
		176	0	0.025	S
		178	0	0.003	NA
		182	0	0.139	S
RM95	0.40%	116	0	0.173	S
		118	0.053	0	R
		120	0.002	0	NA
		122	0.012	0.753	S
		124	0.083	0	R
		126	0.036	0	R
		128	0.173	0	R
		130	0.293	0.008	R
		132	0.109	0	R
		134	0.005	0	NA
		136	0.079	0	R
		138	0.088	0	R
		140	0.026	0	R
		142	0.002	0	NA
		147	0	0.001	NA
RME025	0.60%	151	0.019	0	R
		155	0.07	0	R
		157	0.001	0	NA
		159	0.003	0.001	NA
		168	0.758	0.007	R
		170	0.093	0.001	R
		183	0.001	0	NA
		193	0.001	0.914	S
		195	0	0.009	NA
		203	0	0.003	NA
		207	0.012	0.028	NA
TGLA40	0.40%	91	0.191	0	R
		95	0.005	0	NA
		96	0.003	0	NA
		97	0.478	0.004	R
		98	0	0	NA
		99	0.041	0	R
		101	0.197	0	R
		102	0.002	0	NA
		104	0.001	0.661	S
		106	0	0.218	S
		108	0.001	0.001	NA
TGLA126	0.10%	99	0	0.001	NA
		100	0.001	0.444	S
		101	0	0.476	S
		105	0.934	0.039	R
		130	0	0.003	NA
		132	0.002	0	NA
		134	0.006	0	NA
		136	0.003	0	NA
		138	0	0	NA
TGLA127	0.30%	161	0	0.477	S
		167	0.014	0	R
		169	0.301	0.005	R
		171	0	0	NA
		172	0.003	0.003	NA
		174	0.039	0.41	S
		176	0.024	0	R
		178	0.238	0.006	R
		180	0.05	0	R
		184	0.097	0	R
		186	0.073	0	R
		188	0.002	0	NA
		190	0.067	0	R
		192	0.04	0	R
TGLA337	9.20%	126	0.005	0.396	S
		138	0.043	0.22	S
		145	0.237	0.002	R
		128	0	0.093	S
		147	0.072	0.096	NA
		155	0.003	0.037	S
		136	0.241	0.001	R
		134	0.002	0.003	NA
		130	0.205	0	R
		132	0.106	0	R
		153	0.001	0	NA
		142	0.001	0	NA
UWCA47	0.80%	225	0.03	0	R
		229	0.049	0	R
		231	0.87	0.085	R
		240	0	0.868	S

Table 4.A2. Posterior allele frequencies from analysis 1 (n = 2307) at K = 5.

Locus	% Missing data	Allele Size	Estimated allele frequency in Cluster IV	Estimated allele frequency in Cluster II	Estimated allele frequency in Cluster V	Estimated allele frequency in Cluster I	Estimated allele frequency in Cluster III
AGLA293	2.00%	128	0.003	0.074	0.000	0.158	0.048
		144	0.906	0.786	0.993	0.657	0.850
		147	0.046	0.094	0.001	0.078	0.004
		Null	0.045	0.046	0.006	0.107	0.098
BM4006	0.50%	85	0.000	0.000	0.000	0.009	0.000
		87	0.086	0.092	0.360	0.051	0.094
		93	0.774	0.850	0.125	0.812	0.850
		95	0.127	0.053	0.515	0.124	0.054
		Null	0.013	0.004	0.001	0.004	0.002
BM6438	1.60%	249	0.590	0.528	0.423	0.536	0.766
		251	0.175	0.220	0.001	0.322	0.081
		253	0.053	0.160	0.001	0.123	0.062
		257	0.024	0.000	0.000	0.000	0.007
		261	0.001	0.033	0.569	0.003	0.072
		275	0.005	0.000	0.000	0.001	0.000
		Null	0.152	0.059	0.005	0.015	0.011
BM757	0.10%	160	0.117	0.044	0.029	0.048	0.139
		162	0.499	0.545	0.966	0.507	0.442
		164	0.000	0.026	0.000	0.003	0.000
		172	0.000	0.000	0.000	0.006	0.000
		174	0.000	0.001	0.000	0.007	0.000
		179	0.072	0.025	0.000	0.075	0.044
		183	0.042	0.120	0.000	0.094	0.057
		185	0.001	0.098	0.000	0.045	0.039
		187	0.203	0.024	0.002	0.008	0.017
		189	0.000	0.005	0.000	0.002	0.001
		196	0.000	0.002	0.000	0.000	0.000
		197	0.000	0.000	0.000	0.001	0.000
		198	0.001	0.019	0.000	0.084	0.110
		200	0.004	0.060	0.000	0.078	0.146
		202	0.001	0.002	0.000	0.028	0.000
		210	0.000	0.017	0.000	0.000	0.000
		Null	0.058	0.009	0.003	0.012	0.004
BOVIRP	0.50%	140	0.007	0.002	0.000	0.003	0.000
		142	0.000	0.000	0.000	0.001	0.000
		147	0.017	0.061	0.000	0.104	0.035
		149	0.257	0.074	0.013	0.012	0.015
		151	0.081	0.163	0.485	0.150	0.196
		153	0.433	0.394	0.019	0.438	0.428
		155	0.013	0.035	0.068	0.040	0.148
		157	0.088	0.183	0.412	0.196	0.172
		159	0.083	0.024	0.000	0.016	0.001
		163	0.000	0.001	0.000	0.000	0.000
		Null	0.020	0.064	0.002	0.041	0.005
FCB193	2.50%	101	0.034	0.004	0.000	0.000	0.000
		103	0.001	0.012	0.000	0.093	0.000
		105	0.006	0.001	0.000	0.001	0.000
		107	0.021	0.193	0.000	0.059	0.131
		109	0.028	0.270	0.001	0.213	0.081
		111	0.138	0.001	0.003	0.001	0.007
		113	0.274	0.283	0.035	0.291	0.224
		115	0.000	0.000	0.000	0.000	0.044
		118	0.025	0.024	0.004	0.080	0.013
		120	0.228	0.057	0.315	0.028	0.117
		122	0.046	0.092	0.129	0.095	0.143
		124	0.001	0.009	0.000	0.099	0.019
		126	0.001	0.009	0.082	0.000	0.001
		128	0.028	0.002	0.048	0.003	0.006
		130	0.001	0.012	0.348	0.009	0.131
		132	0.006	0.006	0.033	0.003	0.002
		134	0.000	0.001	0.000	0.002	0.019
		140	0.024	0.000	0.000	0.000	0.000
		141	0.000	0.002	0.000	0.000	0.000
		143	0.030	0.008	0.000	0.016	0.001
		Null	0.109	0.015	0.003	0.008	0.060
FSHB	0.80%	180	0.018	0.000	0.000	0.006	0.000
		183	0.000	0.001	0.000	0.000	0.000
		184	0.000	0.002	0.000	0.006	0.281
		185	0.087	0.182	0.000	0.297	0.146
		186	0.001	0.003	0.000	0.002	0.000
		187	0.000	0.000	0.000	0.000	0.019
		188	0.122	0.146	0.297	0.089	0.083
		189	0.174	0.167	0.050	0.117	0.111
		190	0.053	0.002	0.001	0.000	0.000
		191	0.074	0.082	0.202	0.076	0.062
		192	0.000	0.011	0.000	0.041	0.005

		193	0.000	0.001	0.000	0.000	0.000
		194	0.000	0.035	0.000	0.035	0.000
		195	0.000	0.001	0.000	0.000	0.000
		196	0.001	0.037	0.000	0.001	0.000
		197	0.002	0.004	0.005	0.009	0.006
		198	0.104	0.051	0.103	0.084	0.081
		199	0.014	0.030	0.005	0.032	0.002
		200	0.004	0.000	0.000	0.000	0.000
		201	0.016	0.011	0.000	0.007	0.003
		202	0.001	0.022	0.000	0.047	0.019
		203	0.018	0.032	0.000	0.022	0.017
		204	0.067	0.001	0.025	0.000	0.000
		205	0.089	0.032	0.306	0.077	0.004
		206	0.002	0.102	0.000	0.023	0.000
		207	0.066	0.021	0.000	0.018	0.105
		208	0.011	0.000	0.000	0.000	0.000
		210	0.000	0.000	0.000	0.002	0.054
		211	0.000	0.009	0.000	0.000	0.000
		Null	0.076	0.014	0.003	0.007	0.002
IDVGA29	2.00%	136	0.708	0.712	0.397	0.692	0.663
		143	0.245	0.277	0.602	0.297	0.335
		156	0.019	0.001	0.000	0.001	0.000
		Null	0.028	0.011	0.001	0.010	0.002
IDGVA55	2.40%	191	0.069	0.016	0.038	0.064	0.002
		193	0.021	0.113	0.000	0.134	0.024
		195	0.240	0.260	0.005	0.299	0.134
		197	0.282	0.345	0.300	0.326	0.213
		199	0.135	0.152	0.652	0.120	0.329
		202	0.001	0.042	0.000	0.037	0.003
		204	0.000	0.007	0.000	0.002	0.266
		210	0.005	0.000	0.000	0.006	0.000
		217	0.220	0.005	0.003	0.002	0.025
		219	0.000	0.051	0.000	0.008	0.000
		221	0.000	0.000	0.000	0.001	0.000
		Null	0.029	0.008	0.002	0.004	0.003
INRA005	0.20%	126	0.969	0.982	0.998	0.985	0.991
		136	0.002	0.000	0.000	0.000	0.000
		143	0.002	0.002	0.000	0.003	0.000
		Null	0.028	0.016	0.002	0.011	0.008
INRA006	0.10%	128	0.001	0.001	0.000	0.000	0.000
		130	0.000	0.002	0.000	0.002	0.002
		132	0.001	0.089	0.000	0.033	0.043
		134	0.917	0.654	0.897	0.654	0.575
		136	0.048	0.215	0.099	0.291	0.375
		138	0.012	0.030	0.000	0.010	0.002
		Null	0.020	0.010	0.003	0.010	0.004
INRA131	0.00%	87	0.000	0.001	0.000	0.000	0.000
		92	0.006	0.078	0.000	0.052	0.013
		94	0.005	0.015	0.000	0.009	0.004
		98	0.595	0.547	0.897	0.523	0.713
		100	0.380	0.239	0.002	0.262	0.162
		102	0.010	0.073	0.000	0.121	0.049
		104	0.001	0.042	0.098	0.022	0.056
		106	0.000	0.000	0.000	0.002	0.000
		Null	0.004	0.004	0.002	0.010	0.001
MM012	0.10%	89	0.833	0.683	0.995	0.705	0.775
		91	0.145	0.296	0.003	0.262	0.219
		93	0.000	0.000	0.000	0.003	0.000
		95	0.001	0.000	0.000	0.000	0.000
		97	0.000	0.004	0.000	0.001	0.000
		104	0.000	0.001	0.000	0.000	0.000
		Null	0.021	0.015	0.002	0.029	0.006
RM012	0.70%	116	0.000	0.000	0.000	0.014	0.000
		120	0.030	0.003	0.015	0.000	0.001
		125	0.165	0.157	0.000	0.250	0.055
		127	0.007	0.154	0.029	0.037	0.004
		129	0.155	0.127	0.001	0.062	0.082
		131	0.039	0.057	0.063	0.032	0.277
		133	0.172	0.302	0.001	0.324	0.241
		137	0.001	0.011	0.000	0.034	0.002
		139	0.022	0.073	0.190	0.113	0.022
		141	0.090	0.061	0.001	0.066	0.199
		144	0.001	0.039	0.674	0.024	0.097
		151	0.253	0.005	0.010	0.027	0.016
		Null	0.063	0.010	0.015	0.016	0.003
RM188	0.80%	113	0.041	0.000	0.000	0.000	0.000
		115	0.002	0.050	0.001	0.018	0.002
		117	0.004	0.019	0.000	0.075	0.022
		121	0.002	0.002	0.000	0.000	0.000
		123	0.075	0.042	0.000	0.072	0.012
		125	0.095	0.110	0.014	0.019	0.183
		127	0.218	0.325	0.965	0.452	0.325
		129	0.078	0.269	0.001	0.192	0.402
		131	0.001	0.069	0.000	0.043	0.003
		132	0.251	0.001	0.006	0.000	0.005
		133	0.003	0.004	0.000	0.000	0.000
		134	0.001	0.059	0.000	0.067	0.001
		137	0.111	0.037	0.001	0.040	0.024
		139	0.003	0.002	0.000	0.003	0.009
		143	0.001	0.001	0.000	0.003	0.000

		161	0.000	0.000	0.000	0.001	0.000
		182	0.000	0.000	0.000	0.001	0.000
		Null	0.115	0.010	0.012	0.013	0.013
RM95	0.50%	118	0.001	0.011	0.261	0.046	0.061
		120	0.000	0.009	0.000	0.000	0.000
		122	0.024	0.005	0.052	0.008	0.001
		124	0.001	0.187	0.000	0.087	0.064
		126	0.001	0.022	0.000	0.073	0.035
		128	0.135	0.126	0.324	0.223	0.101
		130	0.222	0.274	0.003	0.326	0.538
		132	0.370	0.065	0.008	0.095	0.051
		134	0.000	0.023	0.000	0.001	0.000
		136	0.080	0.038	0.350	0.039	0.075
		138	0.045	0.190	0.001	0.084	0.062
		140	0.085	0.036	0.000	0.014	0.006
		142	0.000	0.009	0.000	0.000	0.000
		Null	0.037	0.005	0.002	0.005	0.006
RME025	0.40%	151	0.002	0.017	0.000	0.044	0.002
		155	0.050	0.150	0.000	0.051	0.067
		157	0.000	0.001	0.005	0.000	0.000
		159	0.000	0.006	0.000	0.004	0.000
		168	0.765	0.609	0.991	0.744	0.876
		170	0.029	0.178	0.000	0.124	0.037
		183	0.000	0.000	0.000	0.003	0.001
		193	0.002	0.003	0.000	0.004	0.000
		207	0.022	0.022	0.000	0.010	0.006
		Null	0.130	0.013	0.003	0.016	0.011
TGLA40	0.40%	91	0.197	0.204	0.001	0.249	0.196
		95	0.040	0.000	0.000	0.000	0.000
		96	0.020	0.000	0.000	0.000	0.000
		97	0.112	0.385	0.734	0.537	0.754
		98	0.001	0.001	0.000	0.000	0.000
		99	0.006	0.117	0.001	0.040	0.002
		101	0.508	0.200	0.247	0.154	0.033
		102	0.000	0.009	0.000	0.000	0.000
		104	0.001	0.001	0.000	0.005	0.000
		108	0.003	0.000	0.000	0.000	0.000
		Null	0.112	0.084	0.017	0.014	0.013
TGLA126	0.00%	100	0.000	0.006	0.000	0.000	0.000
		101	0.000	0.000	0.000	0.003	0.000
		105	0.984	0.847	0.996	0.976	0.985
		132	0.000	0.007	0.000	0.000	0.000
		134	0.000	0.027	0.000	0.000	0.000
		136	0.000	0.013	0.000	0.000	0.000
		138	0.000	0.002	0.000	0.000	0.000
		Null	0.015	0.098	0.004	0.020	0.014
TGLA127	0.30%	161	0.000	0.000	0.000	0.004	0.000
		167	0.001	0.049	0.000	0.009	0.000
		169	0.076	0.283	0.184	0.418	0.398
		171	0.001	0.001	0.000	0.000	0.000
		172	0.000	0.012	0.000	0.000	0.000
		174	0.135	0.017	0.004	0.037	0.022
		176	0.164	0.005	0.000	0.000	0.000
		178	0.134	0.219	0.732	0.280	0.024
		180	0.129	0.052	0.001	0.052	0.015
		184	0.016	0.108	0.074	0.074	0.234
		186	0.076	0.025	0.000	0.041	0.258
		188	0.000	0.007	0.000	0.000	0.000
		190	0.165	0.122	0.004	0.027	0.042
		192	0.054	0.081	0.000	0.040	0.001
		Null	0.048	0.017	0.002	0.018	0.004
TGLA337	11.30%	126	0.021	0.003	0.001	0.006	0.000
		128	0.000	0.001	0.000	0.000	0.000
		130	0.377	0.164	0.470	0.125	0.199
		132	0.003	0.136	0.000	0.181	0.032
		134	0.000	0.002	0.000	0.004	0.000
		136	0.151	0.205	0.376	0.255	0.269
		138	0.012	0.100	0.000	0.045	0.005
		142	0.000	0.003	0.000	0.000	0.000
		145	0.251	0.205	0.117	0.224	0.399
		147	0.135	0.094	0.006	0.077	0.024
		153	0.000	0.001	0.000	0.001	0.000
		155	0.000	0.009	0.000	0.000	0.001
		Null	0.049	0.077	0.029	0.080	0.070
UWCA47	0.70%	225	0.196	0.001	0.000	0.008	0.006
		229	0.020	0.086	0.001	0.066	0.021
		231	0.751	0.875	0.976	0.884	0.964
		240	0.000	0.000	0.000	0.001	0.000
		Null	0.032	0.037	0.022	0.041	0.009

Table 4.A3. Posterior allele frequencies from analysis 2 (n = 2201) at K = 2.

Locus	% Missing data	Allele Size	Estimated allele frequency in Red Cluster II (HL)	Estimated allele frequency in Red Cluster I
AGLA293	2.00%	128	0.001	0.106
		144	0.945	0.740
		147	0.018	0.067
		Null	0.036	0.088
BM4006	0.50%	87	0.194	0.072
		93	0.451	0.836
		95	0.266	0.085
		Null	0.088	0.007
BM6438	1.70%	249	0.495	0.580
		251	0.083	0.231
		253	0.023	0.125
		257	0.016	0.001
		261	0.230	0.025
		Null	0.154	0.038
BM757	0.10%	160	0.079	0.070
		162	0.672	0.499
		164	0.000	0.010
		174	0.000	0.004
		179	0.038	0.054
		183	0.017	0.097
		185	0.000	0.060
		187	0.121	0.010
		189	0.000	0.003
		196	0.000	0.001
		197	0.000	0.000
		198	0.000	0.071
		200	0.000	0.090
		202	0.000	0.013
		210	0.000	0.006
		Null	0.071	0.014
BOVIRP	0.50%	142	0.000	0.000
		147	0.003	0.073
		149	0.150	0.030
		151	0.245	0.165
		153	0.229	0.430
		155	0.039	0.063
		157	0.223	0.183
		159	0.047	0.015
		163	0.000	0.000
		Null	0.064	0.042
FCB193	2.50%	101	0.022	0.001
		103	0.000	0.046
		105	0.003	0.001
		107	0.016	0.117
		109	0.010	0.199
		111	0.080	0.001
		113	0.157	0.278
		115	0.000	0.011
		118	0.017	0.045
		120	0.254	0.056
		122	0.080	0.106
		124	0.000	0.052
		126	0.035	0.003
		128	0.035	0.003
		130	0.149	0.030
		132	0.021	0.002
		134	0.000	0.005
		140	0.013	0.000
		141	0.000	0.001
		143	0.013	0.012
		Null	0.094	0.031
FSHB	0.80%	180	0.010	0.002
		183	0.000	0.000
		184	0.000	0.065
		185	0.039	0.229
		186	0.000	0.002
		187	0.000	0.005
		188	0.193	0.099
		189	0.124	0.132
		190	0.032	0.000
		191	0.126	0.076
		192	0.000	0.022

		193	0.000	0.000
		194	0.000	0.027
		195	0.000	0.000
		196	0.002	0.012
		197	0.004	0.006
		198	0.102	0.073
		199	0.009	0.024
		200	0.002	0.000
		201	0.008	0.008
		202	0.000	0.032
		203	0.012	0.024
		204	0.050	0.000
		205	0.173	0.045
		206	0.006	0.039
		207	0.036	0.039
		208	0.006	0.000
		210	0.000	0.013
		211	0.000	0.003
		Null	0.063	0.023
IDVGA29	2.00%	136	0.566	0.697
		143	0.402	0.297
		156	0.009	0.000
		Null	0.023	0.006
IDVGA55	2.50%	191	0.051	0.038
		193	0.009	0.105
		195	0.116	0.248
		197	0.287	0.306
		199	0.341	0.167
		202	0.000	0.032
		204	0.000	0.056
		217	0.133	0.004
		219	0.000	0.021
		221	0.000	0.000
		Null	0.063	0.023
INRA005	0.20%	126	0.973	0.988
		136	0.001	0.000
		143	0.001	0.000
		Null	0.025	0.012
INRA006	0.10%	128	0.001	0.000
		132	0.002	0.054
		134	0.888	0.637
		136	0.074	0.281
		138	0.006	0.015
		Null	0.029	0.013
INRA131	0.00%	87	0.000	0.000
		92	0.001	0.054
		94	0.003	0.010
		98	0.707	0.569
		100	0.190	0.236
		102	0.002	0.089
		104	0.046	0.035
		Null	0.052	0.007
MM012	0.10%	89	0.893	0.717
		91	0.079	0.264
		95	0.001	0.000
		97	0.000	0.002
		104	0.000	0.000
		Null	0.027	0.018
RM012	0.70%	116	0.000	0.004
		120	0.025	0.001
		125	0.075	0.177
		127	0.018	0.065
		129	0.078	0.089
		131	0.049	0.092
		133	0.091	0.298
		137	0.000	0.020
		139	0.089	0.078
		141	0.049	0.095
		144	0.266	0.040
		151	0.141	0.018
		Null	0.120	0.022
RM188	0.80%	113	0.023	0.000
		115	0.009	0.022
		117	0.001	0.044
		121	0.001	0.001
		123	0.037	0.051
		125	0.056	0.085
		127	0.471	0.381
		129	0.048	0.264
		131	0.003	0.043
		132	0.132	0.000
		133	0.001	0.002
		134	0.000	0.048
		137	0.057	0.036
		139	0.002	0.004
		Null	0.158	0.019

RM95	0.50%	118	0.102	0.039
		120	0.000	0.003
		122	0.033	0.005
		124	0.001	0.114
		126	0.000	0.049
		128	0.209	0.155
		130	0.118	0.362
		132	0.200	0.075
		134	0.000	0.008
		136	0.186	0.045
		138	0.025	0.113
		140	0.046	0.020
		142	0.000	0.003
		Null	0.079	0.009
RME025	0.40%	151	0.001	0.026
		155	0.033	0.085
		157	0.003	0.000
		159	0.000	0.004
		168	0.830	0.730
		170	0.015	0.122
		183	0.000	0.002
		207	0.012	0.013
		Null	0.107	0.018
TGLA40	0.50%	91	0.095	0.227
		95	0.022	0.000
		96	0.012	0.000
		97	0.356	0.532
		98	0.000	0.000
		99	0.004	0.054
		101	0.368	0.137
		102	0.000	0.003
		108	0.002	0.000
		Null	0.142	0.046
TGLA126	0.00%	100	0.000	0.002
		105	0.985	0.920
		132	0.000	0.002
		134	0.000	0.008
		136	0.000	0.004
		138	0.000	0.001
		Null	0.015	0.063
TGLA127	0.30%	167	0.001	0.020
		169	0.118	0.363
		171	0.000	0.001
		172	0.000	0.004
		174	0.075	0.028
		176	0.093	0.000
		178	0.363	0.201
		180	0.071	0.046
		184	0.046	0.115
		186	0.037	0.089
		188	0.000	0.002
		190	0.091	0.057
		192	0.022	0.043
		Null	0.084	0.031
TGLA337	11.40%	126	0.011	0.002
		130	0.415	0.150
		132	0.001	0.135
		134	0.000	0.003
		136	0.253	0.241
		138	0.003	0.056
		142	0.000	0.001
		145	0.190	0.250
		147	0.070	0.072
		153	0.000	0.001
		155	0.000	0.003
		Null	0.057	0.085
UWCA47	0.80%	225	0.102	0.005
		229	0.009	0.066
		231	0.826	0.896
		Null	0.063	0.033

Table 4.A4. Posterior allele frequencies from analysis 4 (n = 752) at K = 2.

Locus	% Missing data	Allele Size	Estimated allele frequency in Red	Estimated allele frequency in Sika
AGLA293	0.00%	128	0.003	0.000
		144	0.014	0.020
		147	0.972	0.970
		Null	0.011	0.010
BM4006	0.00%	85	0.969	0.979
		87	0.001	0.001
		93	0.027	0.000
		95	0.000	0.010
		Null	0.003	0.009
BM6438	2.40%	249	0.004	0.014
		251	0.003	0.000
		259	0.001	0.144
		263	0.000	0.005
		265	0.444	0.391
		273	0.000	0.045
		275	0.457	0.306
		Null	0.092	0.095
BM757	0.30%	160	0.011	0.000
		162	0.005	0.000
		172	0.975	0.809
		174	0.000	0.170
		183	0.000	0.003
		185	0.001	0.001
		198	0.005	0.000
		202	0.000	0.010
		Null	0.002	0.005
BOVIRP	0.00%	140	0.973	0.948
		142	0.000	0.026
		144	0.000	0.002
		145	0.000	0.003
		153	0.011	0.000
		155	0.000	0.005
		Null	0.016	0.016
FCB193	0.00%	103	0.009	0.000
		111	0.000	0.002
		122	0.001	0.013
		126	0.000	0.057
		128	0.000	0.098
		130	0.001	0.000
		132	0.981	0.705
		134	0.000	0.085
		Null	0.008	0.040
FSHB	0.30%	179	0.000	0.016
		180	0.978	0.500
		181	0.000	0.217
		182	0.009	0.064
		185	0.000	0.003
		188	0.000	0.009
		189	0.000	0.039
		190	0.000	0.046
		191	0.000	0.002

		194	0.003	0.000
		199	0.000	0.016
		200	0.000	0.023
		202	0.000	0.003
		205	0.004	0.000
		207	0.000	0.022
		208	0.000	0.002
		209	0.000	0.002
		Null	0.004	0.036
IDVGA29	2.30%	136	0.008	0.036
		143	0.004	0.060
		145	0.000	0.058
		146	0.000	0.057
		156	0.971	0.715
		Null	0.017	0.074
IDVGA55	0.00%	195	0.001	0.000
		199	0.008	0.000
		204	0.000	0.071
		210	0.854	0.761
		212	0.123	0.037
		214	0.001	0.092
		215	0.001	0.001
		217	0.000	0.003
		Null	0.012	0.034
INRA005	0.30%	124	0.003	0.057
		126	0.024	0.149
		129	0.000	0.003
		136	0.000	0.006
		137	0.000	0.003
		143	0.967	0.775
		Null	0.006	0.007
INRA006	0.40%	130	0.955	0.942
		132	0.000	0.002
		134	0.017	0.040
		136	0.003	0.000
		Null	0.025	0.015
INRA131	0.10%	94	0.000	0.211
		98	0.005	0.002
		100	0.000	0.005
		106	0.878	0.699
		113	0.068	0.038
		115	0.000	0.020
		Null	0.049	0.024
MM012	0.10%	89	0.001	0.180
		91	0.010	0.019
		93	0.976	0.739
		Null	0.014	0.062
RM012	0.00%	116	0.984	0.987
		125	0.001	0.000
		129	0.003	0.000
		139	0.008	0.000
		Null	0.004	0.012
RM188	0.80%	127	0.004	0.010
		129	0.000	0.015
		139	0.000	0.056
		141	0.000	0.016
		143	0.505	0.562

		145	0.000	0.033
		153	0.000	0.049
		161	0.263	0.080
		163	0.005	0.000
		176	0.037	0.012
		178	0.000	0.007
		182	0.149	0.130
		Null	0.037	0.030
RM95	0.10%	116	0.073	0.301
		122	0.881	0.638
		130	0.000	0.019
		132	0.000	0.002
		147	0.001	0.000
		Null	0.044	0.039
RME025	1.30%	159	0.000	0.002
		168	0.010	0.008
		170	0.003	0.000
		193	0.981	0.856
		195	0.000	0.020
		203	0.000	0.008
		207	0.000	0.064
		Null	0.006	0.042
TGLA40	0.40%	97	0.009	0.000
		101	0.001	0.000
		104	0.903	0.472
		106	0.035	0.473
		108	0.001	0.000
		Null	0.051	0.054
TGLA126	0.30%	99	0.001	0.000
		100	0.428	0.464
		101	0.512	0.432
		105	0.016	0.070
		130	0.001	0.006
		Null	0.042	0.028
TGLA127	0.30%	161	0.534	0.406
		169	0.001	0.011
		172	0.001	0.006
		174	0.293	0.559
		178	0.012	0.001
		Null	0.159	0.017
TGLA337	2.80%	126	0.653	0.145
		128	0.096	0.092
		134	0.002	0.006
		136	0.003	0.000
		138	0.088	0.417
		145	0.003	0.003
		147	0.000	0.228
		155	0.028	0.049
		Null	0.126	0.060
UWCA47	0.80%	231	0.003	0.191
		240	0.992	0.790
		Null	0.005	0.019

Chapter 5: Phenotypic consequences of introgression.

Author's contributions:

Of the 1461 individuals analysed in this chapter 720 had previously been genotyped by Helen Senn, 639 by SS and 102 by Megan Wyman and SS. All statistical analyses were performed by SS with Matt Bell providing some advice on the stalker identification analysis and Hanna Granroth-Wilding providing some advice on plots in R. SS wrote the MS. JMP guided the study and edited and commented on the MS.

5.1 Abstract

The impact of hybridisation on phenotype is difficult to assess in the absence of molecular genetic data. Using carcass data, some case studies and stalker assessments, we explore the phenotypic consequences of hybridisation between native red deer (*Cervus elaphus*) and introduced Japanese sika deer (*C. nippon)* in the British Isles. Firstly, we test the extent to which genetically-determined hybrid score (Q) and heterozygosity (of red and sika-specific alleles) explain variation in carcass weight in hybrids. Carcass weight increases with Q amongst hybrid animals and, amongst the dataset that satisfied the less stringent purity criteria, there was evidence for a slight positive effect of heterozygosity on carcass weight. This suggests that additive genetic variation explains variation in carcass weight to a greater extent than by heterosis and that hybridisation introduces a burst of additive genetic variation on which selection may subsequently act. Secondly, I assessed samples from five animals ('case studies') sent to the lab as possible hybrids from areas without known hybrids. Two of the five cases were hybrids. Thirdly, I assess the ability to identify introgressed animals in regions containing hybrids and show that in Scotland accuracy tends to decline as an individual becomes more genetically intermediate, whilst in Co. Wicklow identification of animals is not significantly related to its hybrid score. I discuss the ramifications of these three sets of observations for the management of red-sika hybridisation and introgression in the future.

Key words: Cervus, phenotype, sika, red, carcass weight.

5.2 Introduction

5.2.1 *Phenotypic consequences of hybridisation*

Hybridization is the interbreeding of genetically distinct taxa and is widespread amongst eukaryotes (Allendorf *et al.* 2001). Introgression is the resultant gene flow (described as 'horizontal') between populations whose members are hybridising and can dramatically influence the evolutionary trajectory of a species (Allendorf *et al.* 2001). Where introgression and hybridisation occur, the phenotypic consequences of new recombinant genotypes in hybrids can vary. The most common outcomes are summarised below in terms of the initial effect of introgressed genes on phenotype and the predicted trajectory of phenotypic change with selection.

Introgressive hybridisation may disrupt gene complexes that adapt a population to its environment, a situation also known as 'outbreeding depression', such that over time selection will act to remove hybrid animals and reinforce the integrity of the parental species. This was found for hybrids between the clam species *Mercenaria mercenaria* and *M. campechiensis* in Florida, which are more susceptible to gonadal neoplasia (Bert *et al.* 1993).

Alternatively hybrids may be intermediate in phenotype or change linearly with hybrid score. Intermediate plumage colouration has been shown in the Italian sparrow (*Passer italiae*), a hybrid between the house sparrow (*P. domesticus*) and the Spanish sparrow (*P. hispaniolensis*) (Hermansen *et al.* 2011); of 'labyrinthine' (intricate arrangement of stripes) body patterns in salmonid hybrids (Miyazawa *et al.* 2010); in skull metrics and dentition in hybrid *Myotis* bat species in the Carparthian basin (Bachanek & Postawa 2010); indirectly in growth rates in *Chrondrostoma* species hybrids (Stolzenberg *et al.* 2009); with skull and horn shape in hybrids between the black (*Connochaestes gnou*) and the blue wildebeest (*C. taurinus*) in South Africa (Grobler *et al.* 2011) and with size between the South polar skua (*Catharacta maccormicki)* and the Brown skua (*C. antarctica lonnbergi)* in the Antarctic (Ritz *et al.* 2006). However, if over time, hybridisation results in a phenotype that is not intermediate or conspicuously different from either parental species (e.g. between *Streptopelia spp.,* Rhymer & Simberloff 1996) or occurs by homoploidy hybrid speciation (creation of a new species by hybridisation between genetically or chromosomally different parents, e.g. *Helianthus spp*.; Ungerer *et al.* (1998)) hybrids can be far more difficult to detect. This appears the case between the Seychelles turtle dove (*Streptopelia picturata rostrata*) and the introduced *S. p. picturata,* in which hybrids are more

similar to the phenotype of the latter species (Rhymer & Simberloff 1996). If undetected, such introgressive hybridisation has the potential to erode the genetic integrity of one of the two hybridising species and has led to the extinction of the Tecopa pupfish (*Cyprinodon nevadensis calidae*), the Amista gambusia (*Gambusia amistadensis*) and the longjaw cisco (*Coregonus alpenae*) (Rhymer & Simberloff 1996).

Lastly, hybridisation may lead to heterosis if the reshuffled genetic constitution of the hybrid animal is at a selective advantage in its environment, or areas adjacent to it. This is the case for the hybridogenetic species of frog, *Rana esculenta* produced between *R. lessonae* and *R. ridibunda* in central Spain, which show heterosis in fitness-related traits such as a faster growth, better disease resistance and lower metabolic demands (Arano *et al.* 1995).

Phenotypes are, therefore, often poor indicators of genotypes. Great progress in molecular approaches has enabled us to obtain more detail on the genotype with which particular phenotypes are associated. Allozymes were first used in studies such as that regressing quantitative traits on diagnostic allozyme variants in *Bombina spp.* (Nurnberger *et al.* 1995). This approach has been superceded by the use of DNA markers, for example, Charpentier *et al.* (2008) used 14 microsatellite loci to correlate phenotypic parameters with genetic introgression between yellow and anubis baboons (*Papio cynocephalus* and *P. anubis* respectively) in the Amboseli basin. As well as improving understanding of the phenotypic consequences of introgressive hybridisation on fitness-related traits, genetic analysis may also guide identification of hybrid animals in the field and help to evaluate how effective control measures based on observation are likely to be.

5.2.2 Red-sika hybridisation and its potential phenotypic consequences

Red deer (*Cervus elaphus)* are one of the most widespread species globally and their largest population in Europe resides in the British Isles (Hmwe *et al.* 2006a; Ludt *et al.* 2004). They have been present in Europe since the middle of the Pleistocene, impacted by the end of the last glaciation (~10-11,000 years BP), after which they recolonized the UK and spread into forested areas (Sommer *et al.* 2008). Since the mid-19th century, a series of introductions of exotic deer including Japanese sika (*C. nippon*) into the British Isles has created many opportunities for hybridisation with the native red deer, which has now been documented in captivity (Harrington 1973) and the wild ((Goodman *et al.* 1999; Senn & Pemberton 2009) Chapters 2 and 3). It likely there are now more than the 15,000-20,000 sika in Scotland, around 1,500-2,000 in England, whilst the population in

Ireland is unknown but has increased in range by about 353% over the last 30 years (Carden *et al.* 2010; Díaz *et al.* 2006; Pérez-Espona *et al.* 2009a; Ratcliffe 1987; Whitehead 1964). The distribution of these populations is attributed to many separate episodes of introduction, release or escape (Ratcliffe 1987).

Hybridisation events between these two species appear to be rare, but introgression is extensive at two sites in Kintyre, Argyll: at West Loch Awe and the most southern site in South Kintyre, 50.4% and 61.8% of sampled individuals were hybrid respectively and these populations are best described as 'hybrid swarms' (see Figure in Chpater 2). Similarly, in Co. Wicklow, Ireland, 41% of 197 deer sampled in Co. Wicklow were red-sika hybrids according to either their nuclear genome or mitochondrial haplotype. Elsewhere in Scotland, low-level introgression is apparent in a few individuals ((Senn & Pemberton 2009); Chapter 2). Since genetic introgression has occurred, it is possible that many traits could also have introgressed between species, altering their behaviour, ecology and ultimately, management requirements.

Red and sika deer differ in many morphometric traits; with red deer being very substantially larger than sika in all morphological traits (see Table 1.1). Previous research has shown that red-sika genetic hybrids show intermediate phenotypes between those of the parental taxa. Senn *et al.* (2010) regressed phenotypic trait values against genetically-determined hybrid scores for animals from the hybrid swarm in West Loch Awe and concluded that carcass weight was greater in sika-like hybrids than in putative pure sika and lower in red-like hybrid females than in putative pure red females. Similarly, sika-like hybrids had increased jaw length and incisor arcade breadth (IAB) compared with putative pure sika, whilst IAB was lower in red-like hybrids compared to putative pure red (Senn *et al.* 2010b). The latter study adopted a conservative approach in which analyses were performed independently on the two halves of the distribution ($Q<0.5$ and $Q>0.5$), since an analysis including both parental taxa is strongly influenced by the large number of pure individuals and their phenotypes. As well as morphometric traits, there is a strong likelihood that life history traits could introgress between red and sika. Sika are more fecund than red deer, have a longer rutting season and are more cautious, making them harder to find and shoot (Chadwick *et al.* 1996). Senn *et al.* (2010b) confirmed a difference in pregnancy rate between pure red and pure sika, but did not find a relationship with hybrid score. In addition, sika can live on a much poorer and more fibrous diet and exhibit a greater resistance to lungworm, *Elaphostrongylus spp;* (Bohm *et al.* 2006; Chadwick *et al.* 1996), genetic determinants of which could introgress.

Considering morphometric and life history traits together, it is clear that introgression could result in some very undesirable scenarios: a reduction in body and antler size of red deer and a 'mongrel of the glen scenario', reducing the profitability of sporting stalking on red stags; an increase in red deer fecundity, meaning that larger numbers need to be culled; and an increase in the difficulty of finding and culling red deer. Under these circumstances, managers should try to prevent hybridisation and introgression. In populations where hybrid animals are rare or apparently absent, stalkers should be vigilant in selectively culling any early generation (e.g. F1, F2, early backcrosses) hybrids, should they occur. The presence of only a few hybrid animals can breach the species barrier and lead to substantial introgression due to the fact that during each round of backcrossing with a parental species will reduce the amount of introgressed alleles or genetic material in the offspring by a half leading to widespread, low-level introgression (Fitzpatrick *et al.* 2010; Pemberton 2000; Schwenk *et al.* 2008). In populations which have collapsed into a hybrid swarm, however, stalkers would ideally be able to identify and cull hybrids of all kinds to reinforce assortative mating and maintain the parental species. This is confounded by the likelihood that by the time hybrids are conspicuous or numerous enough to be identified, introgression is likely to be extensive. Thus, the ability of those shooting deer to identify hybrids becomes an important issue for future management of the hybridisation process.

5.2.3 *This study*

This study uses larder data, various case studies and stalker assessments to explore the phenotypic consequences of hybridisation between native red deer and sika in Britain and Ireland. Initially we test the extent to which genetically-determined hybrid score (Q) and heterozygosity (of red and sika-specific alleles) explain variation in carcass weight in hybrids. This study builds upon that of Senn *et al.* (2010b) as a greater number of animals have been genotyped across the full panel of 22 nuclear microsatellite markers and their associated larder data retrieved from the Forestry Commission Larder Record database. This provided sufficient numbers of animals to model carcass weight using exclusively hybrid animals, rather than by more conservative approaches (in which trend only considered significant overall if significant within red-like animals and sika-like animals independently) (Senn *et al.* 2010b). When modelling weight we also incorporate a parameter that accounts for the proportion of loci in an animal's genotype which are heterozygous between red and sika, in order to test for heterosis or outbreeding depression. Lastly, we also vary the genetic criteria over what we regard as a "hybrid". Previous work ((Senn *et al.* 2010a; Senn & Pemberton 2009; Senn *et al.* 2010b); Chapters

2 and 3) defined recent red-sika hybrids as individuals returning a Q value (parameter estimating the membership of an individual to red ancestry, with Q = 1 a pure red and Q = 0 a pure sika) of $0.05 \leq Q \leq 0.95$, whilst under more stringent criteria those further in the extremities of the distribution with $0.01 \leq Q < 0.05$ and $0.95 < Q \leq 0.99$ are described as 'distant' hybrid animals (Senn *et al.* 2010b). The results from Chapter 4 showed that applying the more stringent purity criteria probably does exclude animals with low-level introgression through hybridisation. This study therefore goes beyond that of Senn *et al.* (2010b) by also analysing weight at both definitions of purity.

This study also explores the accuracy of ranger-assigned phenotypes in the form of five individual case studies and four population datasets. The five case studies consist of samples from ambiguous-looking animals, suspected to be hybrid, from regions where red-sika hybrids had not previously been documented. The population data came from four regions in which this project has confirmed the existence of hybrids (in Chapters 2 and 3) and where the stalkers had routinely stated their opinion on the genetic status of the animals shot. We analyse stalker accuracy in relation to Q value and stalker phenotype, and consider whether there is variation in success rate between areas and why.

In more detail the objectives of this chapter are to answer the following questions:

1. *Is carcass weight associated with hybrid status and red-sika heterozygosity?*
2. *In areas where hybrids are rare or absent, do stalkers correctly identify hybrids?*
3. *In areas containing hybrids, how accurately do stalkers identify the taxonomic status of the deer they have shot?*
4. *What implications do the results have for future management of hybridisation?*

5.3 Materials, Methods and Results

5.3.1 Is carcass weight associated with hybrid status and red-sika heterozygosity (objective 1)?

5.3.1.1 Data selection

Deer samples used in this study were initially collected for a study into the extent of red-sika hybridisation across Scotland and Ireland (Chapters 2 and 3), from which a subset was selected from regions known to contain hybrids and which had the all the appropriate phenotypic information. Forestry Commission Scotland rangers were the only stalkers to provide carcass data and weight data came from larder records on culled

animals from rangers in the appropriate regions. Individuals were obtained from Kintyre, Argyll because previous studies have identified extensive hybrid activity in both the south and central region of Kintyre ((Senn & Pemberton 2009); Chapter 2) and from the North highlands because several animals have been identified with low-level nuclear introgression or discordant mitochondrial DNA (Figure 5.1a; Chapter 2). Weight data on all samples genotyped were first plotted, by sex and age, to visualise their distributions (n = 917). The genetically determined hybrid animals were then extracted to form two datasets, adhering to different purity criteria. The first dataset included only animals that met the more stringent hybrid definition (dataset 1, $0.05 \leq Q \leq 0.95$, n = 99) and the second those which met the less stringent hybrid definition (dataset 2, $0.01 \leq Q \leq 0.99$, n = 171). See section *5.2.3* for more details.

Figure 5.1. .Map showing areas studied in order to address each objective of which a) shows the sites from which samples were obtained for objective 1; red shows those from which phenotypic red deer only were sampled, green from which phenotypic sika only were sampled and blue from which both species were sampled. Each of the four sites is outlined by black dashed lines with 1 = South Kintyre (south of Carradale), 2 = West Loch Awe and adjacent, 3 = North Kintyre, 4 = North highlands. b) shows the sampling location for four of the five case study animals obtained from the British Isles (objective 2). Lastly, c) gives the four areas in which the accuracy of stalker-assigned phenotypes were analysed in objective 3.

5.3.1.2 Statistical analysis

The open source statistical programming language, R. 2.15, was used to construct linear models (lm) of carcass weight using the two hybrid datasets, fitting as explanatory variables: age, sex, site shot, hybrid score (Q), and the proportion of loci that were red-sika heterozygous (HET score). All two-way interactions were investigated. Non-significant explanatory variables were then removed from the model in a sequential process until only those that were significant ($p<0.05$) remained (Crawley 2007). A more detailed description of variables follows:

Response variable:

Weight

The response variable in this analysis was the carcass weight, estimated in kilograms after gralloching (removal of gut) removal of the head, remaining internal organs and lower legs. Weight was log transformed (natural logarithm) prior to inclusion in the model because this normalised its distribution (use of the Shapiro-Wilk test in R 2.15). A single male hybrid from Argyll weighing 112kg was removed as likely erroneous.

Explanatory variables:

Age

Age was fitted as the estimated age in days of the animal at death, calculated using the age of the animal in years (estimated by the stalker) and the date on which it was shot. A birth date of 1st June was assumed for all animals and in situations where the date shot was not recorded, the median day in the shooting season was used, specific to both the sex and species of the individual, according to Forestry Commission Scotland seasons: red hinds from 1st Oct - 20th Oct and 16th Feb - 31st March, red stags from 21st Oct - 30th June, sika hinds from 15th Sep - 20th Oct and 16th Feb - 31st March and sika stags from 21st Oct - 30th June (Forestry Commission 2005).

Sex

Sex was fitted as a categorical variable. The sex of each sample was either provided with the tissue or determined by the use of a set of markers designed to amplify a region of the *Zfy* intron, present on both the X and the Y chromosome (Cathey *et al.* 1998; Shaw *et al.* 2003).

Site

Site was fitted as a categorical variable with four levels. Site 1 = South Kintyre (all sites south of and including Carradale), 2 = Central Kintyre including West Loch Awe (WLA) and adjacent areas, 3 = north and north-east Kintyre and 4 = the North highlands (see delimited regions in Figure 5.1a).

Q

Hybrid score was fitted as Q, the membership to red ancestry score from Structure and was restricted to animals with $0.05 \leq Q \leq 0.95$ (dataset 1) and $0.01 \leq Q \leq 0.99$ (dataset 2).

Heterozygosity

The extent to which loci were heterozygous for a red and a sika allele across an animal's genome was also assessed to determine its significance in predicting carcass weight. A measure of heterozygosity was taken as the number of heterozygote loci in an individual's genotype, divided by the number of loci scored (i.e. 22 minus any loci with missing data; see Table 2.A2).

5.3.1.3 Results

The distribution of weight data from all deer in the selected regions (n = 917) with respect to their genetically determined status from across Kintyre and the North highlands is shown in figure 5.2 and clearly shows the substantial differences between sika and red carcass weight in all sex-age classes.

Subsequent models were based on hybrid individuals only. The first model of carcass weight included hybrid animals defined under the criteria of $0.05 \leq Q \leq 0.95$ (dataset 1). The final (minimal) model for this dataset contained age, hybrid score (Q), sex, heterozygosity and an age:sex interaction (Table 5.1); site was rejected as an explanatory variable. Ln weight increased linearly with Q (Figure 5.3a) and had a marginally positive relationship with the heterozygosity index (Table 5.1). The second model of carcass weight data included hybrid animals defined under the criteria of $0.01 \leq Q \leq 0.99$ (dataset 2). The final model included age, hybrid score (Q) and sex and an age:sex interaction (Table 5.2; Figure 5.3b); site and heterozygosity were rejected terms. Ln weight increased linearly with Q (Figure 5.3b). The normality plots for both models are given in Appendix Figure 5.A1.

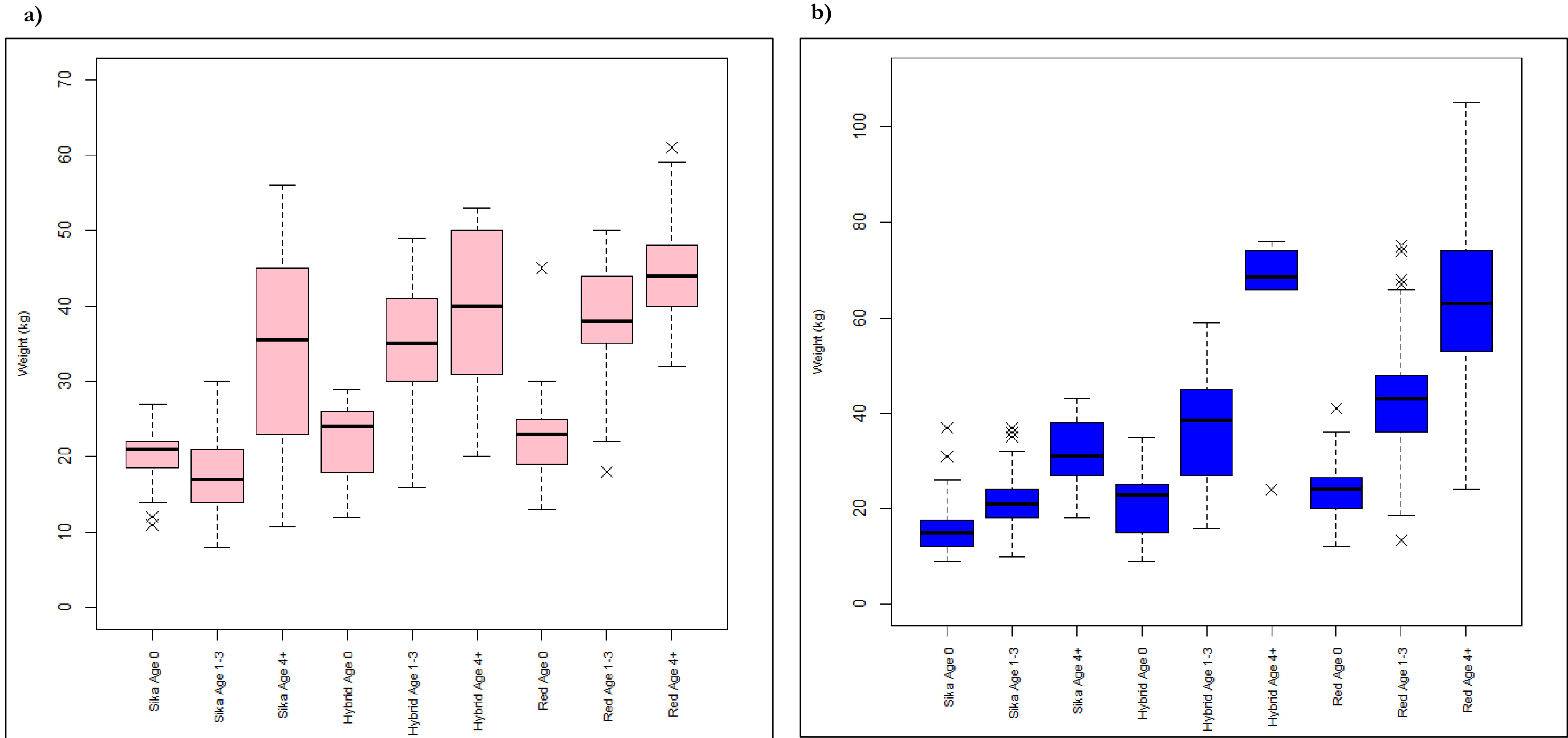


Figure 5.2. Box plots showing the weight distribution of a) females (n = 425) and b) males (n = 492) by age category from Kintyre, Argyll and the North Highlands of Scotland. Note that the weight axis is scaled differently as males are heavier than females.

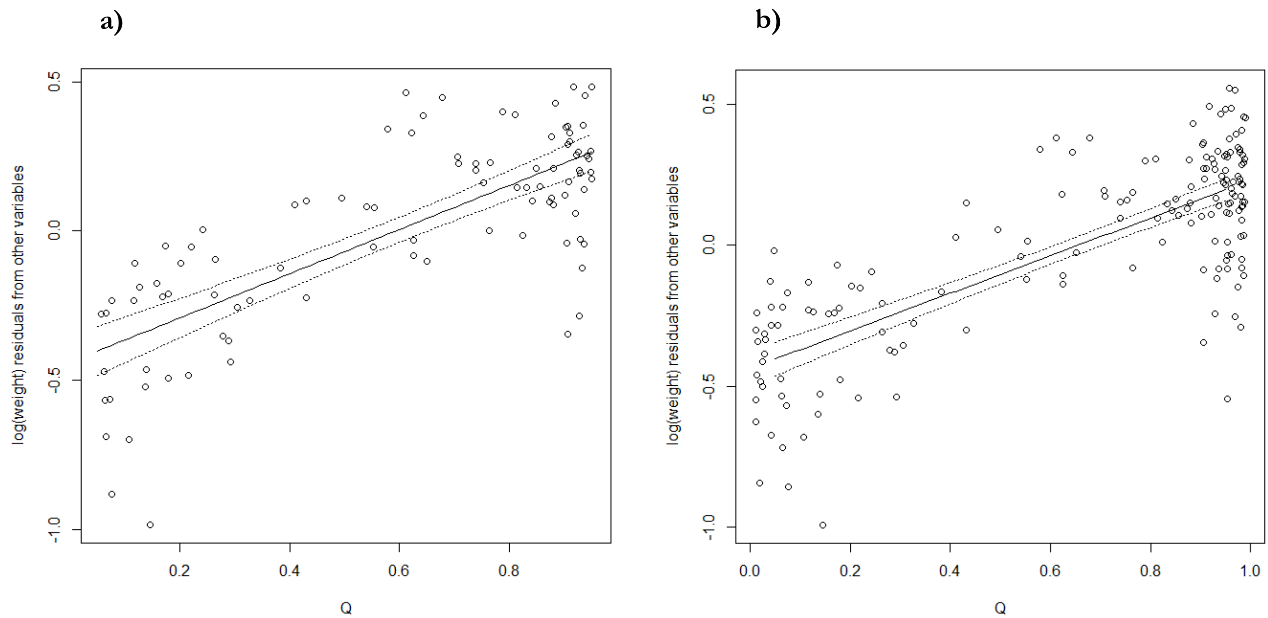


Figure 5.3. The effect of membership to red (Q) on log (weight) in two hybrid datasets. a) Analysis based on hybrids that satisfied the criteria $0.05 \leq Q \leq 0.95$. In order to account for the significance of Age, Sex, Het and Age:Sex interaction this plot shows the residuals of log(weight) after fitting these explanatory variables. b) Analysis based on hybrids which satisfied the criteria $0.01 \leq Q \leq 0.99$. In order to account for the significance of Age, Sex, and Age:Sex interaction this plot shows the residuals of log(weight) after fitting these explanatory variables. The solid line in both plots represent the main trend line through the data and the broken lines the 95% confidence interval around them.

Table 5.1. Final model of log(weight) for hybrid animals with $0.05 \leq Q \leq 0.95$ (n = 99). Adjusted $R^2 = 0.7821$

	Estimate	**SE**	**t-value**	**P-value**	**Significance**
Intercept	2.48	0.086	28.80	$< 2 \times 10^{-16}$	***
Age	3.6×10^{-4}	5×10^{-5}	7.68	1.9×10^{-11}	***
Q	0.84	0.065	12.93	$< 2 \times 10^{-16}$	***
SexM	0.07	0.078	0.85	0.40	
HET	0.27	0.123	2.20	0.03	*
Age:SexM	1.5×10^{-4}	7×10^{-5}	2.17	0.03	*

Table 5.2. Final model of log(weight) for hybrid animals with $0.01 \leq Q \leq 0.99$ ($n = 171$). Adjusted $R^2 = 0.7698$.

	Estimate	**SE**	**t-value**	**P-value**	**Significance**
Intercept	2.66	0.05	53.03	$< 2 \times 10^{-16}$	***
Age	3.2×10^{-4}	3.3×10^{-5}	9.86	$< 2 \times 10^{-16}$	***
Q	0.69	0.04	16.12	$< 2 \times 10^{-16}$	***
SexM	0.09	0.06	1.51	0.13	
Age:SexM	1.3×10^{-4}	5.2×10^{-5}	2.53	0.01	*

5.3.2 In areas where hybrids are rare or absent, do stalkers correctly identify hybrids (objective 2)?

5.3.2.1 *Description of case studies*

During this project, tissue samples and photographs of five different individuals were received which had been shot in areas where hybrids were not expected, rare or absent. These cases have not been part of any of the analyses presented so far. The sites from which four of the five individual case studies are shown in Figure 5.1b and the fifth was obtained from Texas.

Case Study 1: Glenlivet, Moray

This case concerns an animal from Glenlivet, Moray (Figure 5.1b) in 2009. Very few hybrids were recorded in the North of Scotland survey, and none in this area (Chapter 2). A sika stag jumped into a red deer farm field and took up residence alongside a red stag. A late and runty calf was born that was assumed by the owner to be an F1 hybrid. It died of natural causes in its second year while in velvet, and the frozen head was brought to Edinburgh allowing both tissue sampling and photography (Figure 5.4). In terms of phenotype, the head colouration was red-like, and the ears were long and red-like.

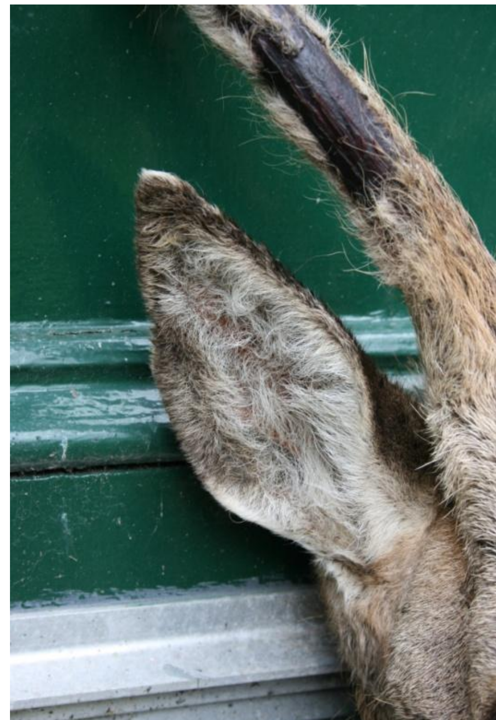


Figure 5.4. Images of case study 1 from Glenlivit. See text for details.

Case Study 2: Devilla, Fife

This case concerns an animal collected from Devilla forest, Fife (Figure 5.1b) in November 2011. The sika population here is thought to be descended from the introduction at Tulliallen estate, Fife circa 1870 (Ratcliffe 1987). Prior to this, it is thought they came from Ireland (Ben Harrower, pers. comm.). The sika population has established locally, however, no hybrids have been recorded in the area (Ratcliffe 1987). An ear tip preserved in ethanol was provided as a sample and pictures of the animal's external phenotype were sent by Ben Harrower (Figure 5.5). The individual is clearly mainly of sika phenotype with dark winter pelage, short, rounded ears with distinctive black half-moon on inside lower rims. However, for a sika stag this animal had very heavy and highly branched antlers including large trez tines and unusual, backwards pointing top tines. Note that while the red pelage colouration on the head could be indicative of red deer, it could also be unmoulted fur from the normal sika summer coat.

Figure 5.5. Images of case study 2 from Devilla. See text for details.

Case Study 3: Wester Ross

This case concerns an animal shot in Wester Ross in 2009 (Figure 5.1b). Very few hybrids were recorded in the North of Scotland survey and not in this area (Chapter 2). A section of tongue tissue and pictures of the animal's external phenotype (Figure 5.6) were provided by the stalker K. Urquhart.

Although generally of sika phenotype its long antlers and red fur on the top of the head led to doubt.

Figure 5.6. Images of case study 3 from Wester Ross. See text for details.

Case Study 4: West Cork

This animal was collected from a site in West Cork (Figure 5.1b) in September 2010. A sample came in the form of an ear tip preserved in ethanol and pictures of the animal's external phenotype were also provided (Figure 5.7). At the time it was sent, there was no public knowledge of hybrids in this area.

Figure 5.7. Image of case study 4 from West Cork. See text for details.

Phenotypically, the animal appeared to be a hybrid. It was physically large for a sika, with a long head and pointed ears (suggesting red ancestry) and yet its spotted flank suggests sika ancestry.

Case Study 5: Texan hunting ranch

This case concerns an animal shot on a Texas hunting ranch in or before 2011. Extracted dehydrated DNA was provided by Prof. James Derr (Texas A&M University). Pictures of the animal were also provided (Figure 5.8).

The animal was claimed to be a record Japanese sika trophy. However the trophy has an exceptionally red-like appearance, including large size, long, thin and pointed ears and large complex antlers with strong trez tines, all more characteristic of red deer. The mtDNA was known from sequencing to be that of Japanese sika (J. Derr, pers. comm.)

Figure 5.8. Images of case study 5 from Texas. See text for details.

5.3.2.2 *Genetic analysis*

See section 2.3.2 in Chapter 2 for DNA extraction, microsatellite genotyping and mtDNA haplotyping procedures. All individuals explored in the case studies were analysed using Structure in an analysis including over 2,500 red and sika animals from across Scotland and Cumbria (see section 2.3.3 in Chapter 2 for parameters).

5.3.2.3 *Results*

Case Study 1: Glenlivit

The suspected F1 stag from Glenlivet returned a Q value of 0.996 (credible region, 0.975 – 1) and carried a red mitochondrial haplotype, confirming that this animal was a pure red deer according to present methods.

Case Study 2: Devilla

The animal from Devilla returned a Q value of 0.015 (c.r. 0.000 – 0.068) and carried a sika mitochondrial haplotype. This suggests that it was a pure sika according to the less stringent purity criteria ($0.05 \leq Q \leq 0.95$). If interpreted under the more stringent purity criteria ($0.01 \leq Q \leq 0.99$), however, it would be defined as a distant hybrid with low-level red introgression.

Case Study 3: Wester Ross

The animal from Wester Ross returned a Q value of 0.002 (c.r. 0.000 – 0.008) suggesting this animal was a pure sika, according to our more stringent purity criteria. The mitochondrial haplotype of this animal failed to amplify in several trials so whether its haplotype was consistent with its nuclear background is, at this stage, unknown.

Case Study 4: West Cork

The suspected 'hybrid' animal from West Cork returned a Q value of 0.249 (c.r. 0.137 – 0.375) and carried a red mitochondrial haplotype. It was clearly a hybrid, matching its extremely intermediate features. Given that it was not an F1, it came from a population in which introgression had proceeded beyond this generation. It was in response to this finding that further samples were obtained from Co. Cork (see Chapter 3).

Case Study 5: Texas

The unusually large 'Japanese sika' animal from Texas returned a Q value of 0.790 (c.r. 0.667 – 0.895) and carried the sika mitochondrial haplotype (in agreement with the sequencing results of J. Derr). Therefore, in common with the last example, it was a hybrid. Again this animal has come from a population in which introgression has extended beyond the F1 hybrid generation. This result should be viewed with some caution, since no other animals were genotyped from this population and it is possible that the current test is not robust when other subspecies of red or sika than are present in the British Isles are involved.

5.3.3 *In areas containing hybrids, how accurately do stalkers identify the taxonomic status of the deer they have shot (objective 3)?*

5.3.3.1 *Data selection*

During sample collection for Chapters 2 and 3, stalkers were asked to identify each sample as either 'red', 'sika' or 'hybrid' based on their assessment of the shot animal. Combining the stalker-assigned phenotype of each individual with its genetically-determined hybrid score and the typing of its mitochondrial haplotype, we were able to assess the accuracy with which this phenotype was assigned. Animals were placed in the 'hybrid' category if they were described as a 'red-like' or a 'sika-like' hybrid by the ranger and were excluded if they weren't assigned a phenotype at all. Whilst previous chapters have made reference to the accuracy of ranger-assigned phenotype across Scotland and Ireland (Chapters 2 and 3), this analysis examined the assignment of phenotype in much greater detail and at the scale of three smaller regions of Scotland: North highlands, South Kintyre and West Loch Awe and one in Ireland, Co. Wicklow (Figure 5.1c). As described in Chapters 2 and 3, South Kintyre, West Loch Awe and Co. Wicklow all contain hybrids swarms, with the Co. Wicklow swarm being very well established and reported since Harrington (1973), the West Loch Awe swarm being discovered and reported relatively recently (Goodman *et al.* 1999; Senn & Pemberton 2009) and the South Kintyre swarm documented for the first time in this thesis (Chapter 2). The North highlands area is characterised by mainly pure red and sika with occasional individuals showing advanced introgression, i.e. a few red alleles in a mainly sika background, or a pure sika Q value with red mtDNA (Chapter 2). Individuals used in this analysis were drawn from the data collected in Chapters 2 and 3 and sample sizes were South Kintyre (south of Carradale) n = 246, West Loch Awe and adjacent n = 369, North highlands n = 198 and Co. Wicklow n = 173.

5.3.3.2 *Statistical analysis*

Analysis of this data took three approaches. The first approach asked whether the number of stalker-assigned pure red, hybrid and pure sika shot in an area came from the same distribution as the genetically-assigned pure red, hybrid and pure sika. For this we constructed a contingency table for each area recording the number of stalker-assigned phenotypes which fell into each genetically-determined category (using the less stringent definition of purity but including hybrids identified through discordant mtDNA). A chi-squared test was conducted in R 2.15 to test whether the stalker assignments (treated as observed numbers) were different from the genetic assignments (treated as the expected numbers with 2 d.f.). Note that in this analysis the genotypic and phenotypic category of 'red' animals was removed from the Co. Wicklow dataset due to low expected counts and the test carried out on sika and hybrid categories only with 1.d.f.

A binary response variable was then created that recorded whether the stalker assigned each animal correctly (1) or incorrectly (0). In the second approach we asked whether the probability of correct stalker assignment was related to the phenotype the stalker assigned – i.e. whether, for example, stalkers were systematically more likely to be wrong when they assigned a deer as a sika. For each site correct/incorrect was modelled by logistic regression using stalker-assigned phenotypes as a categorical explanatory variable with three levels (red, hybrid, sika). Again, the red category in Co. Wicklow was removed prior to this analysis, due to low numbers (6 called by stalkers but none found).

Using this binary response variable, the third approach asked whether the probability of correct stalker assignment was related to the genetic status of an individual. For each site correct/incorrect was modelled by logistic regression using as the explanatory variable the absolute deviation from purity ($|Q2|$). The absolute deviation from purity is a collapsed version of the Q score and is calculated by: if $Q < 0.5$, $Q2 = Q$ and if $Q > 0.5$, $Q2 = 1 – Q$ and was fitted as a continuous variable.

5.3.3.3 *Results*

Comparison of categorical assignments by stalkers and genetic methods.

The number of stalker-assigned phenotypes which fell into each genetically-determined category is given in Table 5.3. The overall accuracy of ranger-assigned phenotype across each of the sites was as follows: 45 of 246 (17.89%) were incorrect with regards to their nuclear genotype or mitochondrial haplotype in South Kintyre, as were 110 of 369 (29.81%) animals obtained from WLA and adjacent sites, 33 of 198 (16.67%) from the North highlands and 28 of 173 (16.8%) from Co. Wicklow (Table 5.3). Application of a chi-squared test showed that within each site, the frequencies of stalker-assignments were significantly different from the frequencies of genetic assignments (χ^2 = 217.36, 151.7, 180.99, all $p < 2.20 \times 10^{-16}$ with 2 d.f. for South Kintyre, WLA and North highlands respectively and $\chi^2 = 84.29$, $p < 2.20 \times 10^{-16}$ for Co. Wicklow with 1.d.f).

Closer inspection of the nature of errors made by stalkers (Table 5.4) suggests differences between sites. At South Kintyre all categories were reasonably accurately identified. At WLA most animals identified as sika were wrongly identified as they were in fact hybrids (Table 5.3). Conversely, in the North highlands, hybrids were over-reported and generally turned out to be sika (Table 5.3). In Co. Wicklow all reds were hybrids but numbers were too low (n= 6) to be included in analyses.

Probability a stalker correctly identified deer in relation to the phenotype he called.

Logistic regression analysis confirmed the observations made above. In south Kintyre and Co. Wicklow there was no difference in the probability a stalker was correct and the phenotype a stalker assigned an animal (Table 5.5; Figure 5.10a, d). At West Loch Awe the stalker was more likely to assign an animal incorrectly if he had called it a sika (Table 5.5; Figure 5.10b). In the North highlands a stalker was significantly more likely to assign an animal incorrectly if they called it a hybrid compared to if they called it a red or sika (Table 5.5; Figure 5.10c).

Probability a stalker correctly identified deer in relation to its genetic status.

At three sites, the probability that a stalker correctly assigned a deer to the categories red, hybrid or sika was associated with Q2 (Table 5.6; Figure 5.11). Stalkers in South Kintyre, the North highlands and WLA were more likely to be wrong the more intermediate the deer was (Figure 5.11a, b, c). In Co. Wicklow, the probability of correctly assigning an individual's status was not associated with Q2 (Figure 5.11d).

		Stalker-Assigned Phenotype											
		South Kintrye			West Loch Awe			North highlands			Co. Wicklow		
		Sika (*n* = 202)	Hybrid (*n* = 17)	Red (*n* = 27)	Sika (*n* = 54)	Hybrid (*n* = 13)	Red (*n* = 302)	Sika (*n* = 89)	Hybrid (*n* = 25)	Red (*n* = 84)	Sika (*n* = 119)	Hybrid (*n* = 48)	Red (*n* = 6)
Genetically determined status	Red (Q> 0.95)	1	0	20	1	4	242	0	4	83	0	0	0
	Hybrid (0.05≤Q≤0.95) or discordant mtDNA	35	15	5	45	9	60	8	1	0	19	45	6
	Sika (Q<0.05)	166	2	2	8	0	0	81	20	1	100	3	0

Table 5.3. Stalker assignments in relation to genetically determined status in four study areas.

		Proportion of phenotypes assigned which were correct		
		Sika	Hybrid	Red
Site	South Kintyre	166/ 202	15/ 17	20/ 27
	WLA	8/ 54	9/ 13	242/ 302
	North highlands	81/ 89	1/ 25	83/ 84
	Co. Wicklow	100/ 119	45/ 48	0 / 6

Table 5.4. The proportion of sika, hybrid and red deer assignments which were made correctly by the stalkers, from each of the four areas.

Table 5.5. Results of logistic regression analysis of the probability a stalker correctly identified the phenotype of a deer as a function of the phenotype called in each of four sites. Only the explanatory variables which were retained in the minimal model for each site are shown. At WLA the stalker was more likely to assign an animal incorrectly if he called it a "sika", whilst in the North highlands a stalker was more likely to assign an animal incorrectly if he called it a 'hybrid' compared to if he called it a 'red' or 'sika'. Phenotype was not significant for South Kintyre or Co. Wicklow.

Model 2	*Explanatory Variable*	**Estimate**	**SE**	**t-value**	**P-value**	**Significance**
West Loch Awe	PhenotypeR	0.9455	0.5883	1.607	0.10799	
	PhenotypeS	-2.2192	0.6868	-3.231	0.00123	**
North highlands	PhenotypeR	6.474	1.178	5.496	3.9×10^{-8}	***
	PhenotypeS	5.493	1.086	5.059	4.2×10^{-7}	***

Table 5.6. Results of logistic regression of the probability a stalker correctly identified the phenotype of a deer as a function of Q2 in each of four sites. The significance of this explanatory variable is shown in each case. In all Scottish sites Q2 is significant in explaining variation in the phenotype assigned by the stalker.

Model 2	**Explanatory Variable**	**Estimate**	**SE**	**t-value**	**P-value**	**Significance**
South Kintyre	Q2	-5.656	1.421	-3.981	6.9×10^{-5}	***
West Loch Awe	Q2	-18.5685	2.3091	-8.041	8.9×10^{-16}	***
North highlands	Q2	-29.0603	12.7414	-2.281	2.3×10^{-2}	*
Co. Wicklow	Q2	-1.185	1.2268	-0.966	3.3×10^{-1}	

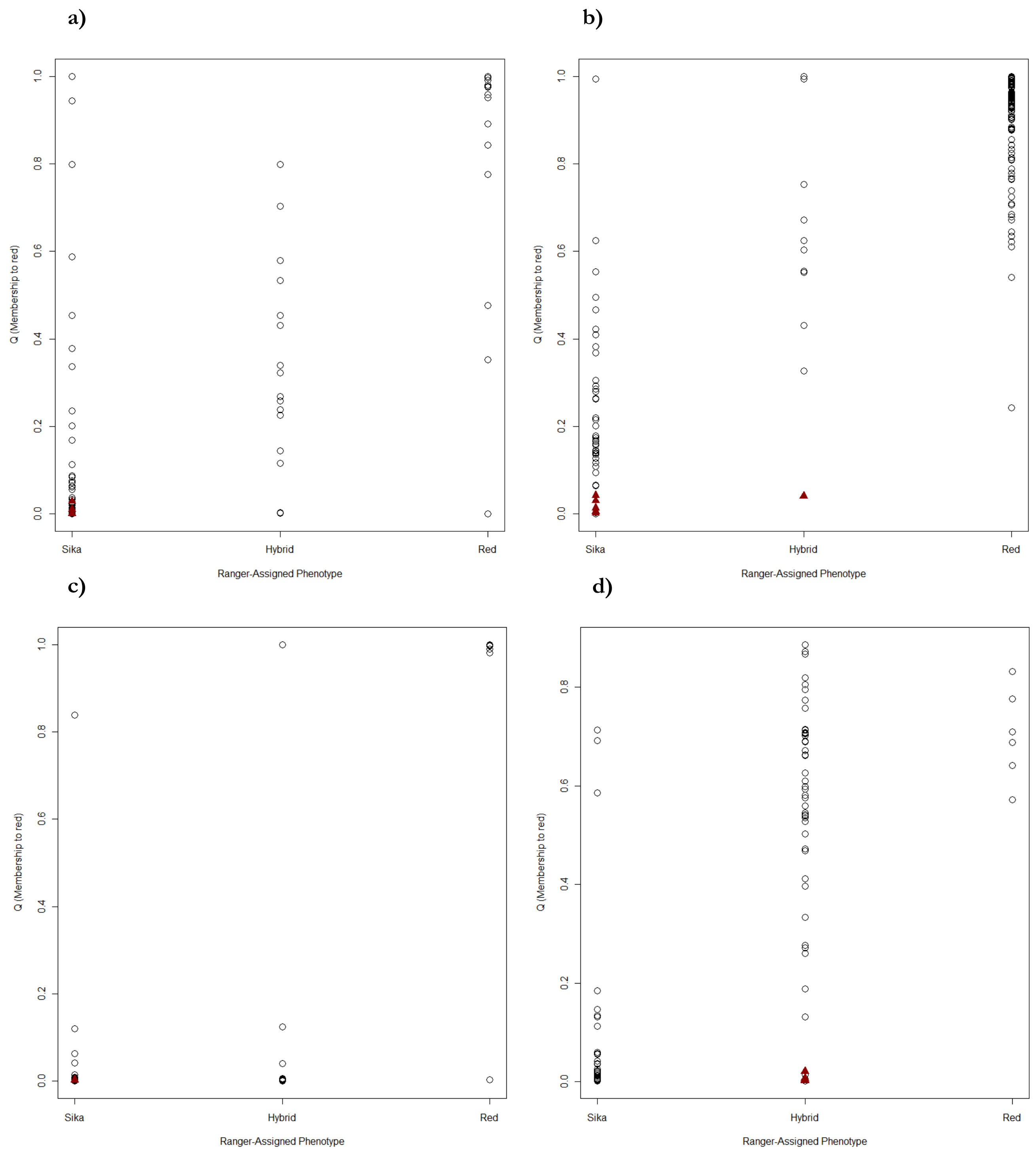


Figure 5.10. The genetically-determined proportion membership to red deer (Q) plotted against stalker-assigned phenotype for animals from a) South Kintyre (n = 246), b) West Loch Awe and adjacent (n = 369), c) North highlands (n = 198) and d) Co. Wicklow (n = 173). Unfilled circles represent animals with the mtDNA haplotype according to their nuclear genetic background and filled triangles to those with discordant mtDNA (i.e. mitochondrial hybrids). Note that the six animals identified as red deer in Co. Wicklow (d), were excluded from the statistical analysis shown in Table 5.5.

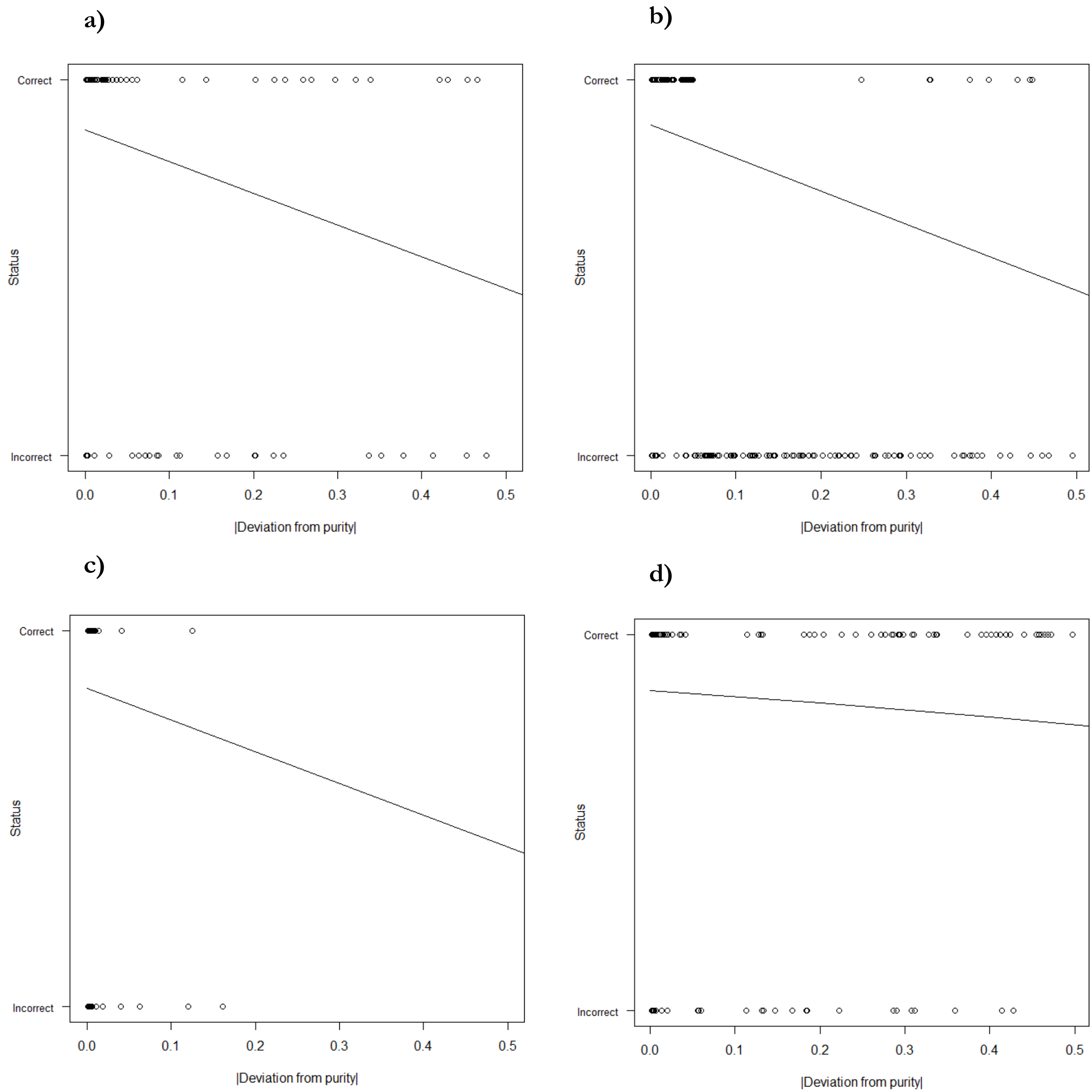


Figure 5.11. The probability a stalker correctly assigned an animal's status plotted against absolute deviation from purity (Q2) for those from a) South Kintyre (n = 246), b) West Loch Awe (n = 369), c) North highlands (n = 198) and d) Co. Wicklow (n = 173). The relationship is significant in all but Co. Wicklow (see Table 5.6).

5.4 Discussion

5.4.1 Is carcass weight associated with hybrid status and red-sika heterozygosity (objective 1)?

Testing whether the level of introgression (Q) or HET score explained variation in carcass weight in hybrid animals only is an improvement on previous studies (Senn *et al.* 2010b) because it reduces the chance of the relationship being driven by the extreme weight differences in pure red and sika animals. By collecting additional samples and making the purity definition more stringent we were able to increase the number of hybrid animals by over 200% compared with Senn et al. (2010b).

Weight models based on the two different definitions of a hybrid did confirm that carcass weight has a significant positive linear relationship with Q. The fact that the inclusion of more animals in the second dataset slightly lowered the slope of Q supports the likelihood these additional animals harbour low-level introgression from the other species and represent 'distant' hybrid animals. This is because the slope would be expected to increase if we had actually added pure parental animals with Q values that reflected small-scale sharing of ancestral polymorphisms.

The linear relationship between carcass weight and Q can be interpreted biologically. Weight is a quantitative trait that is determined by the combined effect of multiple genes (polygenic), rather than one or a few genes. The sequential addition of QTL with small additive effects introduced by introgressive hybridisation could, therefore, increase the weight of the animal.

The proportion of loci in the genotype generated from our marker panel that were heterozygous for red and sika alleles was marginally significant in explaining variation in carcass weight in hybrids, which satisfied the less stringent purity criteria only. This may hint at evidence for heterosis in carcass weight. Visually inspecting all the data (Figure 5.2), however, suggests that the weight for hybrids seems to be slightly closer to the weight distribution for the pure red deer, such that this trait may not be determine purely by additive genetic variation and heterosis may exist. However this trend was lost in the larger dataset under the more stringent purity criterion, such that there was no evidence in this dataset for heterosis (or hybrid vigour) or for outbreeding depression in this trait, a similar finding to that on maturation scheduling in wild baboons (*Papio spp.*; (Charpentier *et al.* 2008)).

Age and sex were incorporated as parameters in the model to improve its inference but are not discussed in this study.

5.4.2 *In areas where hybrids are rare or absent, do stalkers correctly identify hybrid (objective 2)?*

Results from the case studies show that in three cases in which the situation or phenotype suggested animals might be hybrid (Glenlivet, Devilla, Wester Ross) the animals were, in fact, pure members of one or other parental species (red, sika and sika respectively). The unusual looking animal from Cork was confirmed to be a very intermediate hybrid and the suspected Japanese sika trophy from Texas was a red-like hybrid. These cases highlight the phenotypic variation amongst sika (highly branched and backwards pointing tines on the sika from Devilla and unusual long antlers on the Wester Ross animal). Phenotypic variation amongst British sika populations has been noted previously, despite the fact they are all likely to be descended from the same source area in Japan (Goodman *et al.* 2001; Swanson & Putman 2009).

The case studies also highlight that the presence of a sika stag amongst enclosed red hinds doesn't necessarily lead to successful fertilisation or that small runty calves are indicative of hybrid origin (Glenlivit sample). However, when hybrid animals do occur they can have a very intermediate appearance (Cork animal). Neither the Cork nor the Texas hunting ranch example were F1s, and at both sites a substantially introgressed population is likely to exist. The discovery of the Cork hybrid in this work was responsible for the additional sampling of the area reported in Chapter 3. Since these cases are anecdotal, it is hard to draw strong conclusions from them, except that it is clearly possible for stalkers to over-report as well as under-report hybrids (see more below).

5.4.3 *In areas containing hybrids, how accurately do stalkers identify the taxonomic status of the deer they have shot (objective 3)?*

Overall, the accuracy of ranger-assigned phenotype varied from 70.19% - 83.3%. In all four areas containing hybrids that were examined, the ranger-assigned observations were significantly different from their genetically-determined classification. Within the three categories of phenotype assigned there were significant differences in accuracy in WLA and the North highlands only. In all three of the Scottish sites the relationship was that the stalker was more likely to incorrectly identify an animal the more genetically intermediate it was. It must be noted, however, that within these sites the steeper decline in accuracy in the North highlands means that animals here were more likely to be wrongly identified as they became intermediate than at WLA or South Kintyre. In Co. Wicklow, however, Q2 was not significant and a flatter trend suggests the stalkers

assigned individuals (either correctly or incorrectly) more consistently regardless of the genetic status of the animal.

These results can be understood in terms of the pattern and previous knowledge of introgression at each site. In South Kintyre, the accuracy of animals in each category were relatively consistent and reasonably high, despite analyses showing nuclear hybrids were more likely to go undetected by the stalker than putatively pure animals of the parental species (Table 5.6; Figure 5.11). Previous work in this area found that one of two sampled deer was hybrid (Senn & Pemberton 2009) and the sampling reported in Chapter 2 was conducted to investigate the situation more closely. The stalker who provided samples from this site was well informed in terms of this project and may have been more primed to suspect hybrid animals than previously, making the overall accuracy relatively higher.

At West Loch Awe the probability of correctly identifying a hybrid individual was again significant and also slightly lower than South Kintyre (Table 5.6; Figure 5.11). Regardless of the fact that hybrids are well known in this area of Scotland (Abernethy 1994; Goodman *et al.* 1999; Senn & Pemberton 2009) and that the positive linear trend confirmed between weight and hybrid score was based primarily on these animals (section 5.3.1), it appears identification of hybrids in the field remains extremely difficult. Instead red deer were most accurately identified at this site. They make up over 70% of those sampled from WLA, suggesting the abundance and presence of individuals in each of the categories may affect the likelihood of their identification at a particular site.

In the North highlands the probability of correctly identifying a hybrid animal was the lowest amongst all sites (Table 5.6). Of the three nuclear hybrids in the North highlands, only one was correctly identified despite the fact 25 'hybrid' animals were assigned. At this site, stalkers were over-zealous in assigning hybrids but actually missed the genetically-confirmed hybrids.

Lastly, in Co. Wicklow, the relative accuracy of hybrid identification (45/48 correct) accounts for the fact that Q2 is not significant with respect to phenotype call accuracy. The hybrid swarm in Co. Wicklow has been long established (>80 years) and well-reported (Harrington 1973; McDevitt *et al.* 2009a), such that stalkers may be more adept at this site at spotting hybrids.

5.4.4 *What implications do the results have for future management of hybridisation (objective 4)?*

It is apparent that hybridisation is likely to have differential effects on different phenotypic attributes depending on the mechanism behind their genetic regulation and the selection pressures acting on them. Ackermann *et al.* (2006) for example, shows that for some traits Anubis-yellow hybrid baboons showed heterosis and for others, outbreeding depression. In this study, it is evident that the introduction of sika to Britain has altered the genetic, phenotypic and ecological integrity of the native red deer. We have shown that deer weight is significantly related to hybrid score. This, however, is not readily identified in the field, as gauging relative size when only looking at one deer is hard. The phenotype assigned by the ranger may be biased toward the more abundant species at that site or they may be over-sensitive to suspected hybrid reports, both outcomes of which are compounded by the still largely unknown holistic consequences of hybridisation on phenotype. Control of hybridisation and introgression by selective culling based on observation may, therefore, have limited success.

One approach that may help ameliorate this situation is to collect and analyse more detailed phenotypic information on culled animals and provide better of the results to the stalkers in the field. Collecting more detailed morphometric parameters on culled individuals as well as fitness-related traits (such as pregnancy rates) would allow more powerful analyses of the phenotypic consequences of hybridisation, over and above previous work (Senn *et al.* 2010b). Improving liaison with stalkers and ranger managers would benefit both parties; understanding the genetic structure of populations may aid deer management and observations and expertise from those in the field can be invaluable to researchers. Bringing the stalking community up to date with the phenotypic impacts of hybridisation that we have evidence for so far ((Senn *et al.* 2010b); this chapter) through the use of images of hybrids and clearly-explained reports should put us in a much better position to deal with the problem of hybridisation between red and sika in the British Isles. Further, regular meetings with regional sites or larger gatherings of stakeholders would help ensure consistency across management units and encourage progress.

5.5 Acknowledgements

We thank the Forestry Commission Scotland rangers, including Kevin McKillop, Willie Lamont, Derick Macaskill, David Bain, staff of the Deer Commission Scotland (now Scottish Natural Heritage) and stalkers who cooperated within this project and provided samples. Thanks to E. Smith, Ben Harrower, K. Urquhart and J. Derr who provided case study samples. Lastly, thanks to NERC for funding.

5.6 Appendices

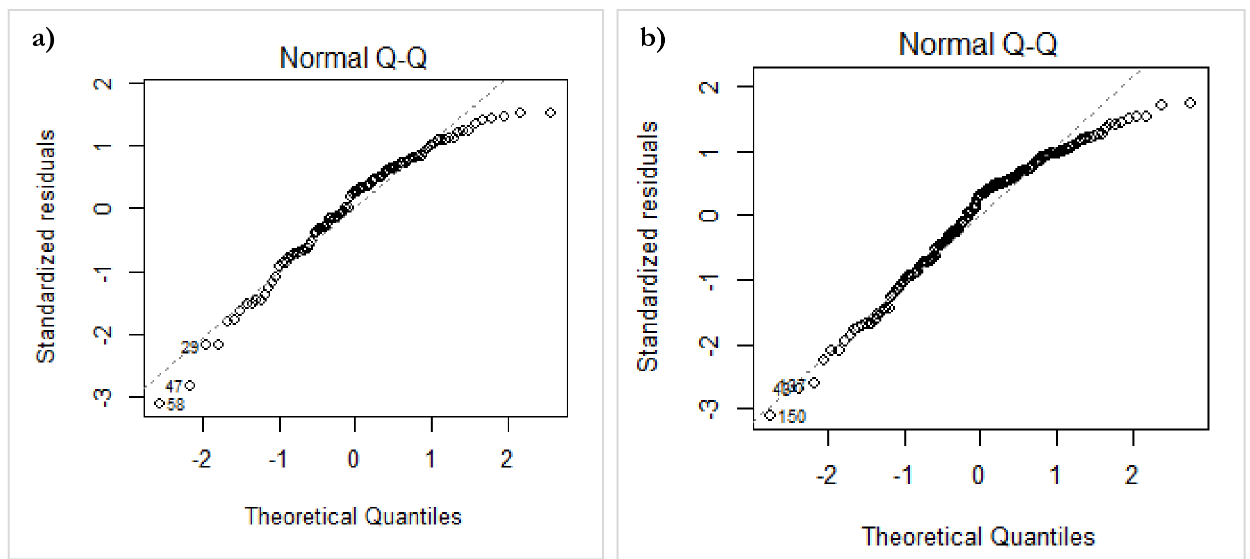


Figure 5.A1. Normality plots generated for the linear model for weight fitted to a) hybrid animals with $0.05 \leq Q \leq 0.95$ and b) hybrid animals $0.01 \leq Q \leq 0.99$. Outliers are identified by a number indicating their position in the dataset.

Chapter 6: General Conclusions.

6.1 Summary of main findings

This thesis builds on previous work spanning almost 20 years, led by numerous contributors (Abernethy 1994; Goodman *et al.* 1999; Senn & Pemberton 2009; Swanson 2000). This previous work showed that hybridisation events between red deer and sika deer are infrequent, yet the consequences in terms of introgression can be extensive. Alongside all the sample sites studied previously by these authors, this project includes new sites in Scotland, England and Ireland and brings the total sample size of deer genotypes at a diagnostic set of genetic markers to over 3,000 individuals.

This thesis corroborates the finding that red-sika hybridisation is infrequent, yet not as infrequent as previously documented. The likelihood of a sika stag managing to secure and mate with a female hind successfully and produce fertile offspring, which survives to reproductive age is low, then such individual F1 animal not only has to be sampled but is likely to be surrounded by individuals of the pure parental species such that subsequent backcrossing occurs (Senn *et al.* 2010a). If hybrid animals reach a threshold at which assortative mating collapses, the occurrence of F1 animals may be relatively more frequent. Beyond the well-established hybrid swarm around West Loch Awe (WLA) in Kintyre this study found a similar swarm in the south of Kintyre (Chapter 2), confirmed the swarm in Co. Wicklow, Ireland (Chapter 3) and identified hybrids (based on both their nuclear genotype and mitochondrial haplotype) in regions within which hybridisation has not been previously shown. These areas include Co. Cork, where a hybrid swarm is present and the North highlands and the Lake District, Cumbria, where occasional advanced back-cross individuals are present. However, overall, only one putative F1 animal was sampled (south of Kintyre) emphasising that the likelihood of an F1 individual being generated, surviving and being sampled appears to remain low.

Among the four hybrid swarms now identified, the proportion of hybrids based on either their nuclear genotype or mitochondrial haplotype are as follows: 114 of 224 (50.9%) animals from WLA, 21 of 34 (61.8%) animals from the most southern site in South Kintyre (or 76/246, 30.9%, of sites south of Carradale), 80 of 197 (41%) from Co. Wicklow, and seven out of 15 animals (47%) sampled from Co. Cork. No pure red deer were detected in Co. Wicklow, suggesting that in this region the red deer has disappeared following hybridisation. The swarm reported in Co. Cork is alarming due to

its proximity to the large population of ancient-origin red deer and introduced sika in Co. Kerry, where no hybridisation has been detected to date.

There has been an over five-fold increase in the number of hybrids genotyped in this study compared to Senn & Pemberton (2009). Despite this much larger sample, only a single putative F1 individual was identified. This animal, shot at West Loch Awe, had a Q value of 0.465 and was heterozygous for a red and a sika allele at 21/22 markers with a single locus (RM95) apparently homozygous for two sika alleles (however the red null allele frequency at this site was 0.035). This animal carried a red mitochondrial haplotype suggesting that, if this was an F1, the sire was a sika and the dam was a red deer. This supports previous suggestions concerning the direction of mating (red hind with sika stag) at a hybridisation event (McDevitt *et al.* 2009a; Powerscourt 1884; Senn & Pemberton 2009) and the idea that hybridisation is initiated by pioneering sika stags that migrate ahead of sika hinds into herds of red hinds. Mating in the opposite direction (red stag-sika hind) has been documented (Harrington 1973) although if a large red stag mates with a young sika hind she may be injured in the process (Harrington 1979; Ratcliffe 1987). However, the preponderance of sika mtDNA in hybrids in South Kintyre (17 of 27 nuclear hybrids carried sika mtDNA) suggests this swarm may have involved mating between red-like stags and sika-like hinds. This swarm is suspected to have been triggered by the escape of red deer from a local deer farm which then hybridised with the abundant sika in the south (K. McKillop pers. comm.). Hybridisation proceeding in a primarily (but not exclusively) unidirectional manner has also been observed in other hybrid systems; for example, between native black wildebeest cows and blue wildebeest bulls (Grobler *et al.* 2011) and between white-tailed deer (*Odocoileus virginianus texanus*) hinds and mule deer stags (*O. hemionus crooki*) (Carr & Hughes 1993).

A hybrid swarm, or complete admixture, represents one end-point that the trajectory of events after a hybridisation event can arrive at. At several sites (excluding hybrid swarms noted above) we observed occasional animals with low-level introgression at nuclear loci and mitochondrial hybrids ($Q<0.05$ with red mtDNA haplotype) within areas where most other sampled animals were primarily pure parental species and assortative mating remained strong. These nuclear hybrids (e.g. the North highlands = 3/568, the Lake District = 3/137) and the mitochondrial hybrids (North highlands = 6/568) represent a different end-point in which F1 hybrids and their descendants backcrossed exclusively into one parental species and many alleles of the other species were lost from the

genome. These individuals serve as a caution that hybridisation events occur more frequently than is indicated by conspicuous hybrid swarms.

Despite the presence of hybrid swarms and various 'smoking guns' of hybrid activity, many other sites that were screened in this study were effectively free of introgressive hybridisation. This may be interpreted as the third end-point of hybridisation; one in which despite the presence of both species, so few F1 hybrids have arisen or that these have been inviable, sterile or left no descendants for some other reason. From the central highlands, including in and around the Cairngorm National Park, large regions throughout the North highlands as well as from the counties in the north west of Ireland and from Co. Kerry we sampled putatively pure red and sika only. The islands sampled in the Hebrides were also free of sika introgression and the 735 animals sampled from these were all pure reds. These islands are protected against *Cervus* introductions by the Wildlife and Countryside act (1981) and this study indicates that this coordinated management remains effective.

This study also demonstrates the use of our marker panel (designed by Senn & Pemberton 2009) for assessing the extent of hybridisation and introgression of red and sika deer with the North American wapiti (*C. canadensis*) across large regions of Scotland, northern and southern parts of England and Ireland. Most wapiti introductions went extinct in the British Isles due to being ill-adapted to its conditions, and previous studies by Perez-Espona *et al.* found no evidence for wapiti introgression using both a mitochondrial and a Y chromosome marker for red deer across central Scotland (Pérez-Espona *et al.* 2010b). Most recently, using a panel of 15 microsatellite markers, only one individual out of 1152 deer samples from the Scottish highlands, islands and English parks was found to have low-level wapiti introgression (Pérez-Espona *et al.* 2013). This study reports a similarly low level of introgression amongst over 3000 individuals analysed. Only 0.24% of the animals sampled from Scotland and just 0.53% of animals from Ireland were marginally introgressed ($Q \geq 0.05$ membership to wapiti) and there was no evidence of introgression amongst the samples from Cumbria. Of a total of nine individuals found to harbour low-level wapiti introgression (average membership to wapiti of $Q = 0.08$) from Scotland and Ireland, only one was from Mamore, a forest into which around 30 wapiti were introduced in 1900 and were known to have crossed with red deer (Whitehead 1964). It may, therefore, be concluded that the impact of North American wapiti on red and sika deer in the British Isles has been negligible.

With a substantial number of putatively pure parental species (>2300 red, >750 sika) this study also explored population structure within red deer, complementing the work of others (Hmwe *et al.* 2006b; McDevitt *et al.* 2009a; Pérez-Espona *et al.* 2009b; Pérez-Espona *et al.* 2008) and, for the first time, that within sika. This was achieved using the clustering program Structure (Pritchard *et al.* 2000) and the multivariate approach of Discriminant Analysis of Principal Components, DAPC (Jombart *et al.* 2010). When we made the definition of a pure parental animal more stringent by increasing the threshold to Q>0.99 for a 'pure' red deer and Q<0.01 for a 'pure' sika, the animals removed were primarily from areas known to contain advanced backcrosses, such that they were suspected to contain low-level introgression. Application of these more stringent purity criteria, therefore, may clarify population structure underlying both parental species.

The within-red population analyses confirmed the strong differentiation of the red deer on Harris and Lewis from the remaining populations in the Hebrides and on the mainland. There was also evidence for sub-structuring within the remaining red deer which fell into four groups: the Irish, the English park, Arran and Rum; Kintyre; the North highlands; Islay and Jura. The differentiation observed is likely to be due to a combination of isolation by distance patterns, significant landscape features, translocations, introductions and past management practices (Hmwe *et al.* 2006b; McDevitt *et al.* 2009a; Pérez-Espona *et al.* 2008; Whitehead 1964). Amongst sika populations sampled, both analytical approaches supported the three clusters which are presumably the result of bottleneck events on introduction and the translocations which followed. The sika clusters primarily constituted the Irish, the North highlands and the Kintyre sika.

Hybridisation between red and sika has important effects on phenotype. Using linear models this study concludes that carcass weight is linearly related to hybrid score (Q) when based exclusively on genetically-determined hybrids from Scotland, rather than approaches which modelled weight within parental taxa datasets separately (red-like, Q>0.5 and sika-like, Q<0.5) and were, therefore, more conservative in their inferences (Senn *et al.* 2010b). We found no strong evidence that the level of red-sika heterozygosity explained variation in carcass weight and, therefore, evidence for heterosis in this trait.

Lastly, this study performed analyses of the ability of stalkers to correctly assign the phenotype of an animal by regions where both hybrids were unreported and well-known. Identification of hybrid animals based on observation has been shown to be

difficult by others (Grobler *et al.* 2011; Senn & Pemberton 2009; Tung *et al.* 2008). Use of logistic regression models on data from animals across regions known to contain hybrids showed the accuracy of ranger-assigned phenotype averaged 78% and appeared site-specific. For example, in regions with backcrossed hybrids (e.g. North highlands) these were generally not identified, in regions with many hybrid animals (e.g. Kintyre), detecting intermediacy was very difficult and in regions with no genetically pure red deer (e.g. Co. Wicklow), animals were still erroneously designated as red deer.

6.2 Implications for management

A major reason to study red-sika hybridisation is to guide future management of the situation. Management of British woodland deer in the past has primarily been reactive, but now there is movement toward more proactive, predictive management (Mayle 1996). Better management is likely to have indirect benefits on reducing forestry and agricultural damage and deer-vehicle collisions (Mayle 1996). Here we consider whether it may also be possible to reduce the rate of hybridisation and introgression through management. Three main scenarios are presented and management suggestions made in light of this study.

6.2.1 Hybrid swarm

Hybrid swarms are a concern as they contain a concentration of introgressed animals which could migrate out and (as they are already primed for hybridisation) initiate hybridisation events elsewhere. When faced with the scenario of a hybrid swarm, as in West Loch Awe, South Kintyre, Wicklow or Cork, recovering threatened taxa or pure parental species can be extremely difficult (Allendorf *et al.* 2001). Management approaches may include enclosing the population with fencing and eliminating the population, as was carried out on 160 blue and black wildebeest in Spioenkop Nature Reserve, South Africa, which were suspected to have almost entirely hybridised (Grobler *et al.* 2011). This may be the optimal approach to eliminating the hybrid swarm at West Loch Awe. It could also be applied to South Kintyre with the use of a 'top down' approach, driving animals towards the Mull of Kintyre. The width of the southern part of the Kintyre peninsula varies from around 12km across at Carradale to less than 8km across at Campbeltown; distances across which deer fencing would be feasible. The dimensions of such fences are important as they should account for the fact that sika and red deer can jump up to 2.5m and can be designed toward different species' dimensions (Honda *et al.* 2011). Without any remaining red deer, the situation in Co. Wicklow would be one of preserving the numerous pure sika. For example, deer

populations could be eliminated from sites known to contain large number of hybrids (e.g. Kippure, Ballyknockan, Derrybawn) and those known to contain mostly pure sika (e.g. Luggala) could be monitored for unusual animals or conspicuous hybrids. Overall, site-specific integrated approaches may best deal with different swarms.

In cases where intensive culling is adopted, it should be consistent and well-planned; partial or ill-informed culling may exacerbate the situation by displacing hybrids and initiating new swarms (Senn & Pemberton 2009). This is exemplified by the problem we have identified in southern Ireland. Management by isolation and elimination of the deer population in Co. Cork will have to ensure it does not displace animals the mere 20km into the range of the protected red deer of ancient origin in Killarney, Co. Kerry. This study confirmed the absence of hybrids in Co. Kerry to date, however, both insufficient culling *and* inaction is eventually likely to see the migration of hybrid animals into the area.

Further suggestions to managing a hybrid swarm, particularly in regions of dense forestry, include the use of wildlife contraceptives in the form of synthetic hormones or immunocontraception (Patton *et al.* 2007). For example, the contraceptive Melengestrol Acetate (MGA) has been used to manage captive populations of antelope and deer in New York (Raphael *et al.* 2003) and they have also been applied to wild populations (DeNicola *et al.* 1997; McShea *et al.* 1997). Despite being highly effective there are concerns about their impact on animal development and the development of pathologies and their biological and toxicological impact on the environment (Patton *et al.* 2007; Raphael *et al.* 2003). Also, ensuring the contraceptive is consumed only by the target species is difficult and they are generally ineffective at preventing male dispersal and reproduction.

6.2.2 Low-level introgression

If deer managers are presented with a situation in which there have been infrequent hybrid individuals reported or evidence for low-level introgression (e.g. North highlands, Lake District), it is very unlikely that stalkers will be able to pick out such individuals (Chapter 5). It is also much debated what proportion of admixture between species is actually "acceptable" (Allendorf *et al.* 2001). In such a scenario it would seem necessary for deer managers to be aware and vigilant against the likely triggers of hybridisation; namely pioneering sika stags (see below) and the migration and translocation of introgressed animals (Pérez-Espona *et al.* 2009a; Senn & Pemberton 2009).

6.2.3 No hybridisation

In regions where no hybrid animals or introgressed material has been found to date it may be necessary for deer stalkers to protect species integrity by remaining vigilant against the appearance of pioneering sika stags within a red deer area (Pérez-Espona *et al.* 2009a; Senn & Pemberton 2009). Studies have suggested stags may arrive in an area up to 10 – 15 years before females or at an advance of 1.3 – 11km/ year and that they initiate hybridisation events beyond the range of sika females (Clarke 1972; Goodman *et al.* 1999; Swanson & Putman 2009). Culling these individuals before they reproduce could lower the risk of hybridisation.

6.2.4 Further aspects to consider

These scenarios raise further issues, the first of which is the detrimental effect of translocation and introduction of individuals and the need for this to be tightly controlled and policed. The widespread impact of the movements to and from Powerscourt estate in the 19th century is an example with numerous detrimental impacts. More stringent control measures and improved record-keeping has been implemented for the transfer of the blue wildebeest (*Connochaetes taurinus*) between reserves in South Africa in order to minimise introgressive hybridisation with the endemic black wildebeest (*C. gnou*) (Grobler *et al.* 2011) and for Roosevelt elk (*Cervus elaphus roosevelti)* and Rocky Mountain elk (*C. e. nelson)* in California, due to their propensity to hybridise (Meredith *et al.* 2007). This has also been the function of the Island Refugia Policy for the Hebrides; legislation established in 1999 as part of the Wildlife and Countryside Act (1981), which makes it illegal to introduce any *Cervus* species into the wild on the Outer Hebrides, Arran, Jura, Islay and Rum.

The second issue raised from these scenarios is that beyond culling and fencing another important facet of deer management is education. In some cases efforts may be better focused on educating landowners and stalkers about the appearance of sika animals, the importance of selectively culling unusual looking animals and raising public awareness, rather than by investing all resources on searching for introgression (Senn 2009). At the same time, the cost of genotyping continues to decline as throughput capacity and resolution accelerates; therefore, use of both approaches could be optimal in the future. As with any situation, an integrated and flexible approach to management is likely to be best.

6.3 Future research

Following the generation of an F1 individual, subsequent backcrossing into one of the two parental species will reduce the introgressed material by 50% each generation, leading to progressively less evidence for hybridisation. With our panel of 22 microsatellites we would no longer expect to detect introgressed material after six generations of backcrossing (Senn & Pemberton 2009). The detection of mitochondrial discordance, however, in an otherwise pure parental animal can provide evidence for a hybridisation event beyond six generations ago. Despite our marker panel being robust and effective, such microsatellite panels can have relatively low coverage and resolution, show variable rates of mutation, null alleles or can exhibit homoplasy (identical character states due to multiple mutation to the same allele size) at particular loci, which can lead to population structure being underestimated (Coates *et al.* 2009; Morin *et al.* 2004). Further, standardising allele sizes for comparison between laboratories (e.g. electrophoresis methods and specific standards used) can be difficult (Coates *et al.* 2009; Morin *et al.* 2004). In addition, our use of a single mitochondrial marker has helped identify past hybridisation and directionality; however, it represents a single, maternally-inherited marker with limited molecular resolution and may be experiencing different selection pressures to the nuclear genome (Hurst & Jiggins 2005; Twyford & Ennos 2012). In conclusion, microsatellites are informative for population-level questions; however, analysis would be improved by more markers (Morin *et al.* 2004). The first improvement in marker panel could, therefore, be the detection and application of further diagnostic loci between red and sika deer. A further, would be the acquisition of a Y-chromosome marker, as used in another study looking for wapiti introgression in Scottish red deer (Pérez-Espona *et al.* 2010b). This would allow the paternal line to be traced and would complement the use of nuclear and mitochondrial markers (Isoda *et al.* 2000). However, Y-chromosome markers provide relatively low molecular resolution as they are a single non-recombining locus, and their inference power is compromised by high variability in male reproductive success lowering the probability that they would persist as introgressed material in the opposite species (Twyford & Ennos 2012).

A more powerful approach would be to genotype numerous single nucleotide polymorphisms (SNPs) in order to detect admixture between red and sika. Despite high initial isolation costs and the fact that more SNPs are generally required than microsatellite markers due to their diallelic nature, it is the high density and uniform distribution that can be achieved with SNPs that makes them superior (Coates *et al.*

2009; Schaid *et al.* 2004). SNPs occur every 200-500bp in most species, highlighting the high genome coverage they permit (Brumfield *et al.* 2003). If SNP genotyping is performed on whole-genome libraries or pooled, enriched regions containing exome and regulatory sequences, as well as noncoding, this could help identify SNPs in functional regions with significance in adaptation. In a study searching for susceptible loci for prostate cancer, SNPs (10k array, average spacing 0.34cM) were shown to provide an average information content of 61% and microsatellites (402 markers used, average spacing 10cM), 41%, primarily due to their higher density and, therefore, association ("linkage") with remaining untyped genomic regions, giving a greater overall representation of the genome (Schaid *et al.* 2004). There are also more flexible approaches to SNP detection, they can be less expensive, have lower error rates in genotyping and the data they generate is far more comparable between studies as they are represented according to the DNA code (G, C, A, T) (Coates *et al.* 2009; Schlotterer 2004). Despite the fact they are usually diallelic, the study of local haplotypes of linked SNPs can act as "super" alleles (Schaid *et al.* 2004). They can also be applied to non-model organisms or those for whom parental genotypes are not available. SNP genotyping has now been applied to numerous hybrid studies, including that between Chinese rhesus macaque and mainland longtails in Indochina (*Macaca* spp.) (Kanthaswamy *et al.* 2010), between the blue catfish (*Ictalurus furcatus*) and channel catfish (*I. punctatus*) (He *et al.* 2003) and between *Acacia auriculiformis* and *A. mangium* (Wong *et al.* 2012).

Massive improvements in sequencing technology have provided the infrastructure to automate SNP genotyping and provide more and better quality information than by previous approaches such as PCR-RFLP, single-base extension, gel electrophoresis or microarrays (Morin *et al.* 2004). One way of obtaining a large amount of SNP genotype data in the absence of any other genomic information for a species is the use of restriction-site associated DNA tag (RAD) sequencing (Baird *et al.* 2008). This recently developed method allows the simultaneous identification and genotyping of thousands of SNPs, removes the issue of ascertainment bias, enabling detailed mapping and providing an integrated platform that uses existing infrastructure (Baird *et al.* 2008). High-throughput is attained by an automated multiplex sequencing and novel barcoding approach for individual identification and can be tailored toward experimental objectives by the choice of restriction enzyme (Baird *et al.* 2008). The application of RAD sequencing to hybrid studies include that between rainbow trout (*Oncorhynchus mykiss*) and the cutthroat trout (*O. clarkii lewisi),* providing the detail to distinguish

homeologs (duplicate homologous genes from same parent) from homologs (informative SNPs between parental species) (Hohenlohe *et al.* 2011) and between *Populus alba* and *P. tremula* highlighting the recurrent gene flow between these tree species and the 'porous' nature of their genomes (Stolting *et al.* 2013).

Such methods, however, target neutral polymorphisms, which are limited in their association with fitness. One way of enabling sites associated with fitness to be screened could be by sequencing the entire genome. This moves from a marker-based system to one of using all the information from the DNA sequence (Schlotterer 2004). Being the ultimate resolution, it has a suite of benefits: it is optimal for detecting candidate genes and quantitative trait loci associated with fitness and can trace the exact sites of recombination, mutation and inheritance (Allendorf *et al.* 2010; Roach *et al.* 2010; Schlotterer 2004). However, it is a far more expensive and laborious approach as it involves sequencing invariant and uninformative sites and would therefore be impractical in application to hybridisation studies and better suited to targeted candidate gene surveys using smaller sample sizes (Allendorf *et al.* 2010; Schlotterer 2004).

Overall, the number and type of markers selected is important, as is the range of animals and populations sampled and the genetic polymorphism within these (Witherspoon *et al.* 2007). The latest developments in next-generation sequencing allow integration of the task of marker design and application in automated processes, have extremely high throughput and, therefore, may provide the resolution to document the descendants that occur after the generation of an F1 hybrid and to distinguish between incomplete lineage sorting and recent introgression (Twyford & Ennos 2012). One hindrance to the application of these technologies is the lagging progress in the analytical bioinformatics, primarily software, which can deal with large number of markers and account for linkage disequilibrium (necessary when the number of markers is much increased).

Aside from molecular approaches, future analyses would greatly benefit from better phenotypic data collection to explore its association with degree of hybridism. The phenotypic outcomes of this project may have major economic incentives for the Scottish economy in terms of deer management and stalking (estimated at £105 million in 2005; (PACEC 2006)) as revenue from sport stalking and recreational activities rely partially on the aesthetic qualities of red deer. Work by Senn *et al.* (2010b) searched for an association between a genetically-determined hybrid score and numerous phenotypic parameters including dentition, body and kidney weights and fertility and highlighted the

extensive analyses that can be done with such phenotypic information. Overall, if we are able to combine higher-resolution genetic information with more consistent and detailed phenotypic information on both species and their hybrids, we could better address the impacts of hybridisation.

7.0 Appendix 1: Supplementary samples

The taxonomic and hybridisation status of *Cervus* deer species in various populations in the south of England and France.

Author's contributions:

An additional 140 samples from southern England and 13 samples from France were provided by Megan Wyman and genotyped by SS. The statistical analysis was performed by SS. SS wrote the MS. JMP guided the study and edited and commented on the MS.

7.1 Abstract

Since the mid-19th century, multiple introductions of Japanese sika deer (*Cervus nippon nippon*), Manchurian sika (*C. nippon mantchuricus)* and wapiti (*C. canadensis*) have taken place across Britain and Europe, primarily into deer parks and enclosures. While wapiti introductions have generally gone extinct, sika have thrived. Hybridisation between this species and the native red deer (*C. elaphus*) has been demonstrated in captivity and in the wild in various parts of the world. Using a panel of 22 microsatellite loci that are highly diagnostic for Japanese sika-red and strongly diagnostic for wapiti-red, a mitochondrial marker which distinguishes red, Japanese sika and Manchurian sika but not wapiti and the Bayesian clustering software Structure 2.3.3*,* we analysed 110 deer comprising phenotypic red deer from five deer parks in the south of England (Badminton, Windsor, Richmond, Bushy and Wadhurst), Japanese sika from two regions of Dorset (Lulworth and Arne) and sika from a park in central France, to investigate the taxonomic status and extent of introgression between species. The integrity of the red deer from the five parks was complete except for the presence of one red-sika hybrid animal in Wadhurst. There was no evidence for red introgression amongst the Dorset sika. All the French samples carried the Manchurian sika mitochondrial haplotype. Two of eleven French samples showed signs of nuclear introgression, one from red and one from wapiti, but since the nuclear marker panel was not designed to discriminate Manchurian sika from other taxa, this is a very tentative finding.

7.2 Introduction

Keeping deer in parks is a form of husbandry that has existed in Great Britain for centuries, principally using red deer and fallow deer (*Dama dama*) which do not hybridise. Numerous deer species from across the globe have been introduced to deer parks for aesthetic purposes and to improve trophy quality. Since the mid-19th century, for example a series of introductions of both North American wapiti (*C. canadensis)* and sika (*C. nippon*) have occurred into the British Isles and this has created many opportunities for hybridisation with resident red deer (*C. elaphus)*. In England, wapiti were introduced to Derby around the 1790s and herds kept in Woburn (Bedfordshire), Buckinghamshire, Kent, Sussex, Cumbria and Northamptonshire, all around the turn of the 20th century (Whitehead 1964). Hybridisation between red and wapiti has occurred both in the wild in Britain (Whitehead 1964) and in captivity (Moore & Littlejohn 1989; Shackell *et al.* 2003). Overall, however, their impact has been limited; wapiti are highly susceptible to lung disease and foot malformation, delayed female maturity and lower

levels of stag aggression than red deer in the rut (Asher *et al.* 2005; Pérez-Espona *et al.* 2010a).

On the other hand, Japanese sika deer have habituated better to the British Isles and have hybridised with red deer in captivity (Harrington 1973) and the wild in Britain (Goodman *et al.* 1999; Lowe & Gardiner 1975; Senn & Pemberton 2009). Whilst much research has been conducted on red-sika populations in Scotland, studies in England have been less extensive and lacked power. Diaz *et al.* (2006) reported low level introgression in New Forest and Purbeck, however, marker numbers and sample sizes were small and little account was made for ancestral polymorphism.

Another sika subspecies, Manchurian sika has also been introduced to England, Ireland and other parts of Europe. Currently it is largely confined to enclosed populations. It is physically larger than the Japanese sika and hybridisation with red has been documented (Powerscourt 1884; Ratcliffe 1987; Whitehead 1964). The hybrids of Manchurian sika with Japanese sika deer are thought to be fertile and this cross was repeatedly made at Frankfurt zoo (Gray 1971).

In this study we sampled red deer from five deer parks in the south of England (Badminton, Windsor, Richmond, Bushy and Wadhurst), Japanese sika from two unenclosed regions of Dorset (Lulworth and Arne) and sika from a park in central France. The objective of this study was to clarify the taxonomic status of the deer sampled and whether any populations contained hybrids.

7.3 Materials and Methods

7.3.1 *Study area and sampling*

We collected samples from red deer in five deer parks in the south of England, wild sika from two regions of Dorset and sika from within a park in Central France (Figure 7.1). Those sampled from Badminton, Bushy, Richmond and Windsor Great Park were all assigned 'pure' red phenotypically (Table 8.1). The old deer park in Windsor was disbanded around the 19th century and the red deer stock extirpated. It was restocked recently with further reds from highland Scotland. It was subject to sika deer introduction around the 1900s (Ratcliffe 1987; Whitehead 1964); however, all the animals sampled from this site in this study were pure red, according to our definitions. The sika at Wadhurst are believed to be Manchurian sika (Neil Brookes, pers. comm).

Table 7.1. Summary of the English parks from which we obtained samples.				
Badminton	1656	330	Red, fallow	Hingston 1988
Bushy	>1514	1,099	Red, fallow	Hingston 1988
Richmond	1637	2,500	Red, fallow	Hingston 1988
Windsor	1247	167	Red	Hingston 1988; Whitehead 1964
Wadhurst	1976	600	Manchurian sika, Pere David, Barasingha, Axis, Menil Fallow	Ratdiffe 1987; Hingston 1988

The free-living sika sampled from Dorset were collected from Lulworth and Arne, regions from which sika have also been studied in the context of their demographic history (Goodman *et al.* 2001) and when looking for red introgression (Díaz *et al.* 2006).

The French sika samples were obtained from a wildlife park that holds various species of Cervid (pers. comm M. Wyman). The enclosure of such numerous closely related species is likely to increase interaction between heterospecific animals and could facilitate hybridisation, as such has been proposed between captive populations of Vietnamese sika, Manchurian sika and other sika subspecies in European zoological parks (Thevenon *et al.* 2003; Thevenon *et al.* 2004).

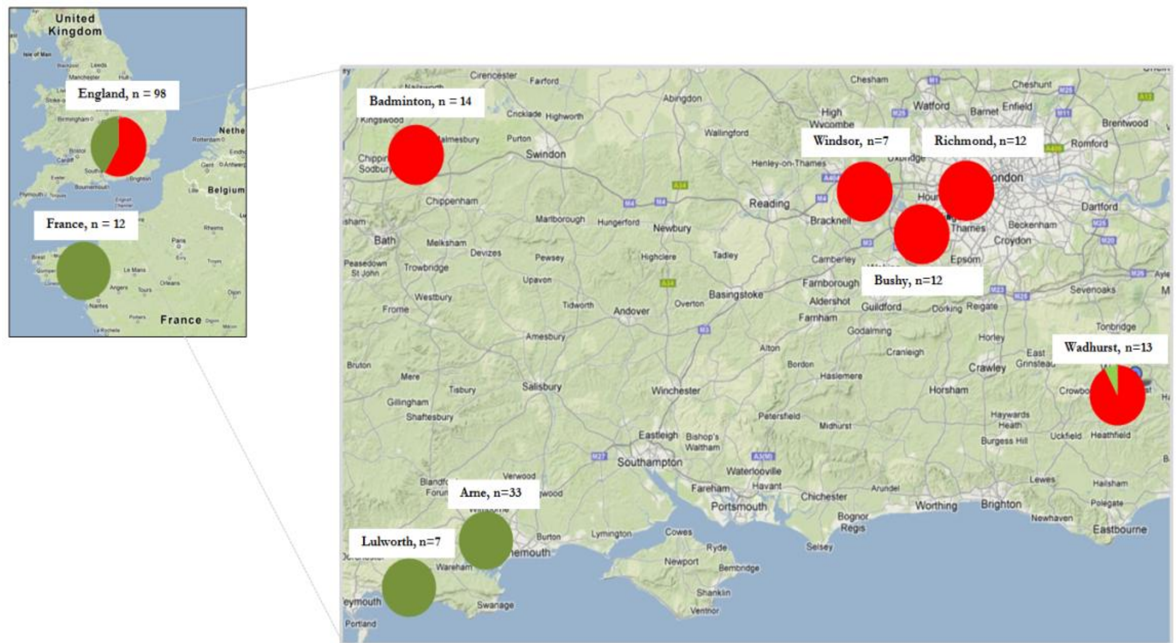


Figure 7.1. Google map image showing the proportion of red (red) and sika (green) based on phenotype, obtained from both south England and France, represented by a pie chart. The larger, right-hand image gives a zoomed in view of the various sample sites from which samples were collected in England. The name of each site and sample size is given above each pie chart.

7.3.2 *DNA analysis*

See Chapter 2 section 2.3.2 for DNA extraction and genotyping procedure. Individuals were also screened for their haplotype in the mitochondrial control region that in deer includes a diagnostic number of 39bp tandem repeats: red deer have a single repeat, Japanese sika have three and Manchurian sika have seven (Cook *et al.* 1999).

7.3.3 *To clarify the taxonomic status of the deer sampled and whether any populations contained hybrids.*

See Chapter 2, section 2.3.3, for Structure methodology.

Two datasets were analysed sequentially using Structure 2.3.3, in order to address the first study aim. Initially, all red and sika sampled from England and France were analysed together with 50 control red animals from central Scotland, 50 control Japanese sika animals from Kintyre, Argyll and 49 wapiti animals from Canada to resolve the most likely population structure which recognises these three species (analysis 1, n = 259). Secondly, the wapiti controls and any animals showing evidence

for wapiti introgression were excluded, in order to assess the extent of red-sika hybridisation only across all sample sites (analysis 2, n = 209).

7.4 Results

7.4.1 *Genotypes*

In total, nuclear genotypes were obtained from 110 animals (excluding controls) which amplified successfully for at least 20 out of 22 of the nuclear loci and had their mitochondrial haplotype scored.

To clarify the taxonomic status of the deer sampled and whether any populations contained hybrids.

7.4.2 To clarify the taxonomic status of the deer sampled and whether any populations contained hybrids.

Analysis 1: Red, Japanese sika and wapiti genotypes (n = 259)

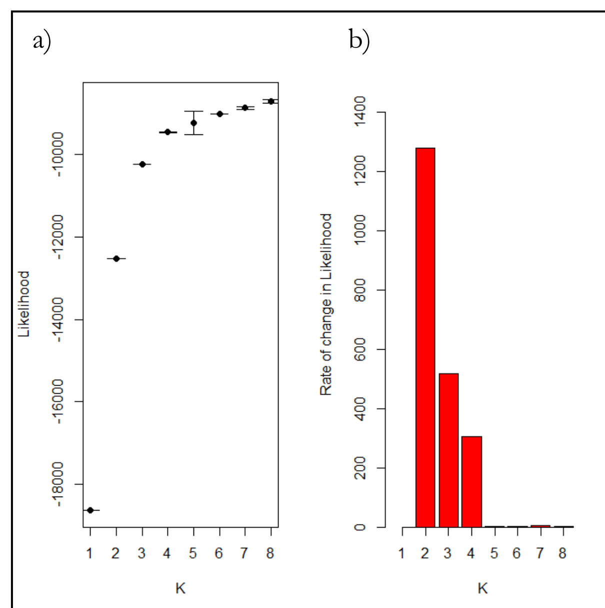


Figure 7.2. Assessment of the most likely number of populations by Structure 2.3.3 in analysis 1. a) Shows the log-likelihood (with standard error) of the value of K (number of populations) given the dataset and b) shows the rate of change in log likelihood between values of K. Although K = 2 appears most likely, wapiti do not differentiate from red until K = 3.

The log likelihood and rate of change in likelihood calculated from Structure would suggest K=2 was the smallest number of genetic clusters that was optimal to describe the population structure (average likelihood (Ln Pr (X|K)) = -12518.56, s.d. 2.70, rate of change in K (ΔL(K)) = 1279.24). At this value of K, red and Japanese sika are differentiated, but not wapiti, which cluster with red (see Appendix Figure 7.A1). The population structure postulated at K=3 defines the point at which wapiti become differentiated from red and sika and, whilst not most likely, for our purposes it is the most appropriate since we are interested in the three potentially hybridising taxa (Figure 7.3). There also remains some support for 4 population clusters and this is illustrated in Appendix Figure 7.A2.

From analysis 1 at K=3, there are three hybrid animals, according to our definitions (Chapter 2, Table 2.2): a red-sika hybrid from Wadhurst and a red-sika hybrid and a sika-wapiti hybrid from the French site. Except for the single hybrid from Wadhurst the remaining red deer from the English parks and the sika deer from Dorset appeared pure and gave no evidence for recent hybridisation or introgression between the three species according to our markers. Since it is possible that the inclusion of wapiti genotypes could confound the analysis of red-sika hybridisation, in analysis 2 we repeated the Structure analysis after removing the 49 wapiti control samples and the single French sample showing wapiti introgression.

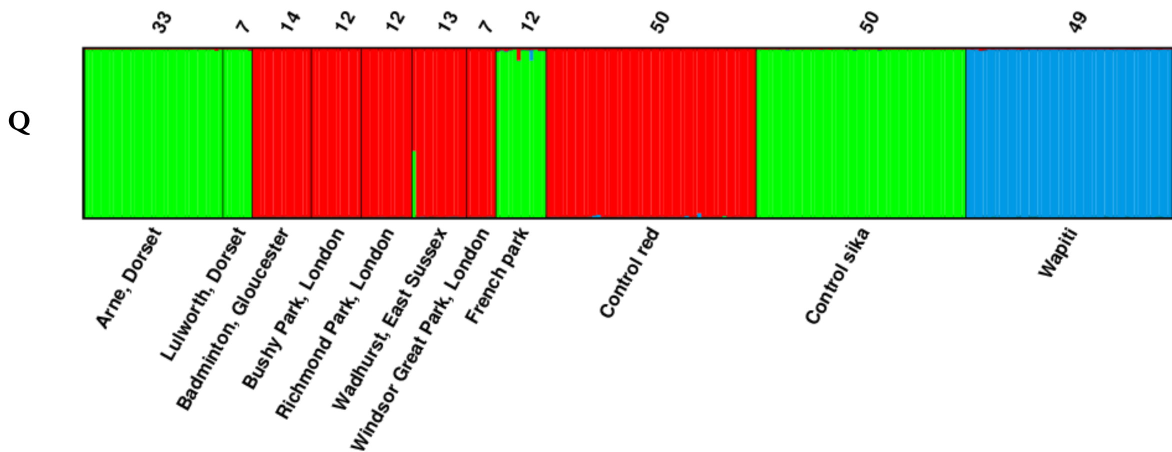


Figure 7.3. Bar chart showing the results of *analysis 1* in STRUCTURE at K = 3 for each individual in the dataset consisting of red deer, sika and wapiti animals (n = 259). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (shown in red), the proportion attributable to Japanese sika ancestry (green) and that attributed to wapiti ancestry (blue). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis.

Analysis 2: Red and Japanese sika genotypes (n = 209)

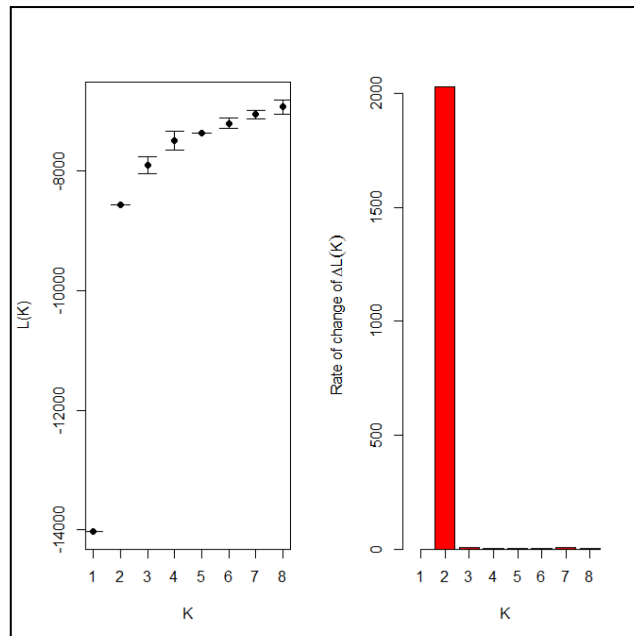


Figure 7.4. Assessment of the most likely number of populations using Structure 2.3.3 in analysis 2. a) Shows the log-likelihood (with standard error) of the value of K, given the dataset and b) shows the rate of change in log likelihood between values of K. Both provide evidence that K = 2 is the most likely.

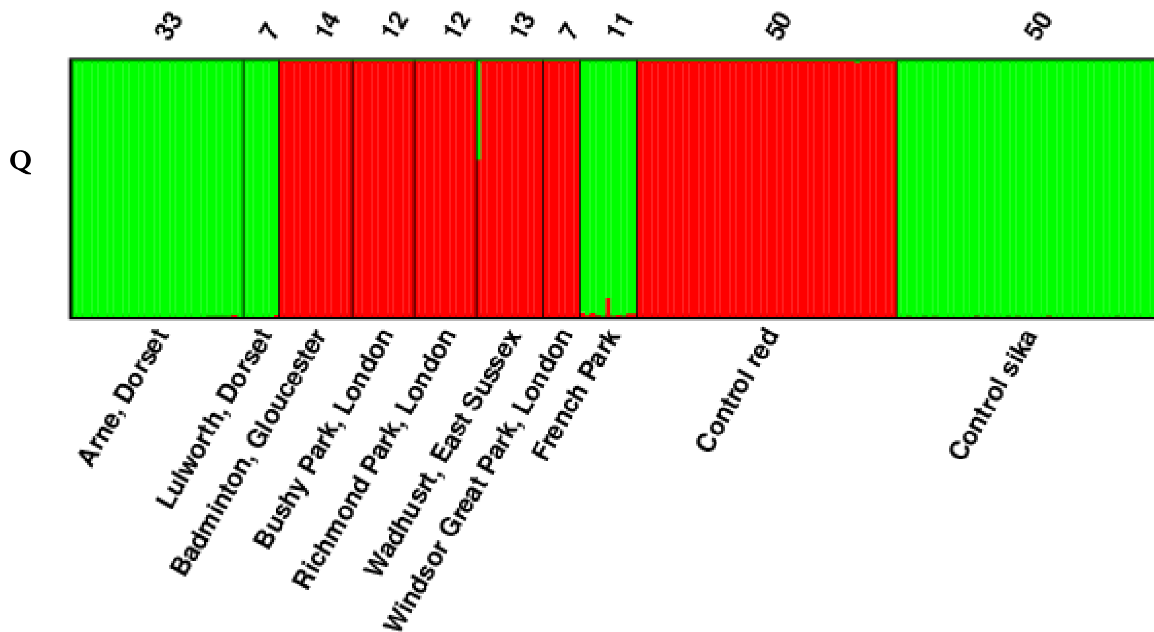


Figure 7.5. Bar chart showing the results of *analysis 2* in STRUCTURE at K = 2 for each individual in the dataset consisting of red deer and sika animals only (n = 209). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (shown in red) and the proportion attributable to sika ancestry (green). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis.

In analysis 2, the logarithmic likelihoods calculated in Structure support K=2 as the smallest number of genetic clusters that describes the majority of the population structure, with an average likelihood of -8557.3 (s.d. = 2.61) and is consistent with the largest rate of change of 2030.18 (Figure 7.4). At this value of K, red deer and sika differentiate and both the red-sika hybrids identified in analysis 1 were again apparent; one from Wadhurst and one from the French park (Figure 7.5).

Incorporating the information from the mitochondrial marker adds further support and resolution to the results from the nuclear markers. All the individuals sampled from England had mitochondrial haplotypes consistent with their nuclear markers, i.e. all park red deer had red mtDNA and all Dorset sika has Japanese sika mtDNA. The Wadhurst hybrid (Q=0.614) had the Japanese sika mtDNA rather than the expected Manchurian haplotype. All 12 deer from the French park carried the Manchurian sika haplotype.

7.5 Discussion

7.5.1 *To clarify the taxonomic status of the deer sampled and whether any populations contained hybrids.*

Bearing in mind that sample sizes are modest, it appears from analysis 1 at K=3 that there is no recent hybridisation and introgression from wapiti in the red deer sampled from the five English parks sampled or the feral Japanese sika sampled from Dorset. However, one of the sika-like animals sampled from France, had a level of apparent wapiti introgression great enough to define it as a sika-wapiti hybrid (Q = 0.005/ 0.063/ 0.932 membership to red/wapiti/sika respectively). Since the French animals all had the Manchurian sika mtDNA haplotype, this interpretation is very provisional. The nuclear markers used here were developed for discriminating red and Japanese sika (Senn & Pemberton 2009) and this is the first time, to our knowledge, they have been applied to possible Manchurian sika. It is possible that Manchurian sika and wapiti share alleles at these markers.

Regarding red-sika nuclear and mitochondrial introgression, the deer from Dorset were all 'pure' sika and the deer from the English parks were all 'pure' red, except for a single red-like hybrid individual from Wadhurst (Q=0.614) that carried a Japanese sika mitochondrial haplotype. Of the 11 phenotypic sika sampled from France (single animal with wapiti introgression removed), one had a level of red introgression (Q = 0.075) great enough to define it as a red-sika hybrid and all deer from this location carried a Manchurian sika mitochondrial haplotype. The results for each sample site will now be discussed in turn.

The genetic integrity of four of the five park red deer populations appears largely intact. Those sampled from Badminton, Bushy, Richmond and Windsor Great Park were all pure red. However, of the 13 animals sampled from Wadhurst, all were 'pure' red except a single intermediate hybrid individual (Q=0.614) that carried the Japanese sika mtDNA haplotype. Care should be taken when interpreting this 'hybrid' animal from Wadhurst, as the sika here are supposed to be Manchurian sika. Our nuclear marker panel is designed to discriminate Japanese sika and red and until now it has not been applied to Manchurian sika. However, if the French samples screened here are genuine Manchurian sika (we have no knowledge of which subspecies they are supposed to be), then the markers appear to strongly cluster Manchurian and Japanese sika.

Remembering that the Wadhurst hybrid was shot as a phenotypic sika, the most likely interpretation is that this animal is hybrid between the red and Japanese sika in the park, and that the Wadhurst sika, although putatively Manchurian, are themselves introgressed by Japanese sika. Wadhurst obtained its Manchurian sika from Woburn Park, Bedfordshire in 1976 ((Hingston 1988); Neil Brookes, pers. comm), and Japanese sika deer were present in Woburn park around the early 20th century until they contracted Johne's disease and were destroyed (Whitehead 1964). Reports suggest Manchurian sika would have also been present in Woburn at this time (Glover 1956; Gustavss I. & Sundt 1969). It is perhaps possible that hybridisation and introgression between the two sika subspecies took place at Woburn before the Japanese sika were destroyed.

The genetic integrity of the Japanese sika in Dorset appears strong since all individuals typed as pure sika at both nuclear and mtDNA markers. This corroborates the work of Diaz *et al.* (2006) who found neglible introgression amongst sika in the New Forest (Hampshire) and Purbeck (Dorset) regions, despite the fact this analysis was based on only eight mircosatellite loci. We suspect that the sika sampled from the park in France were Manchurian sika, since they all had the distinctive Manchurian mtDNA haplotype and all 11 samples were very sika-like at their nuclear markers. If this is correct, then it appears that the markers originally developed to discriminate red and Japanese sika (Senn & Pemberton 2009) also largely discriminate red and Manchurian sika. On the other hand, the detection of two possible hybrids, one with a small amount of wapiti introgression and one with a small amount of red introgression, suggests we should be cautious in our in our interpretation; it is possible that these animals are all pure Manchurian sika but that this subspecies shares low frequency alleles with both red and wapiti.

7.6 Acknowledgments

We would like to thank all the stalkers who provided samples from deer. This work was supported by a Natural Environment Research Council PhD studentship.

7.7 Appendices

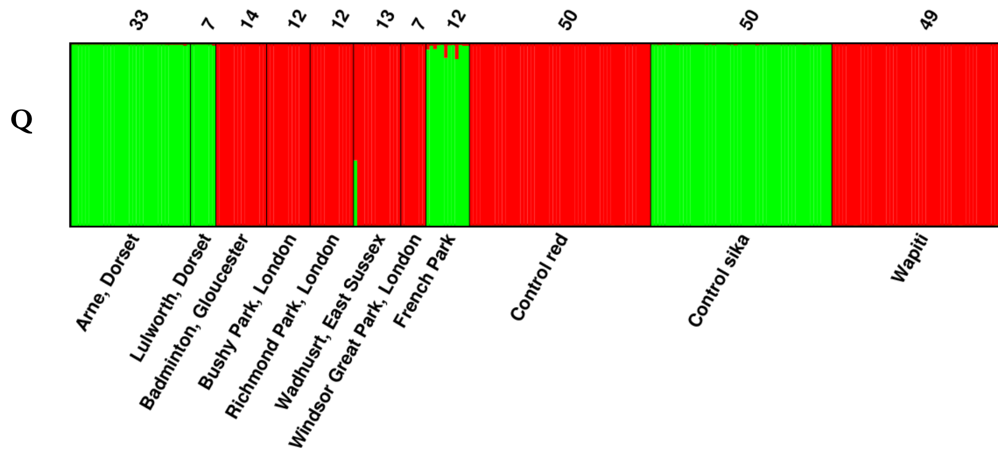


Figure 7.A1. Bar chart showing the results of *analysis 1* in STRUCTURE at K = 2 for each individual in the dataset consisting of red deer, sika and wapiti animals (n = 259). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (shown in red) and the proportion attributable to sika ancestry (green). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis.

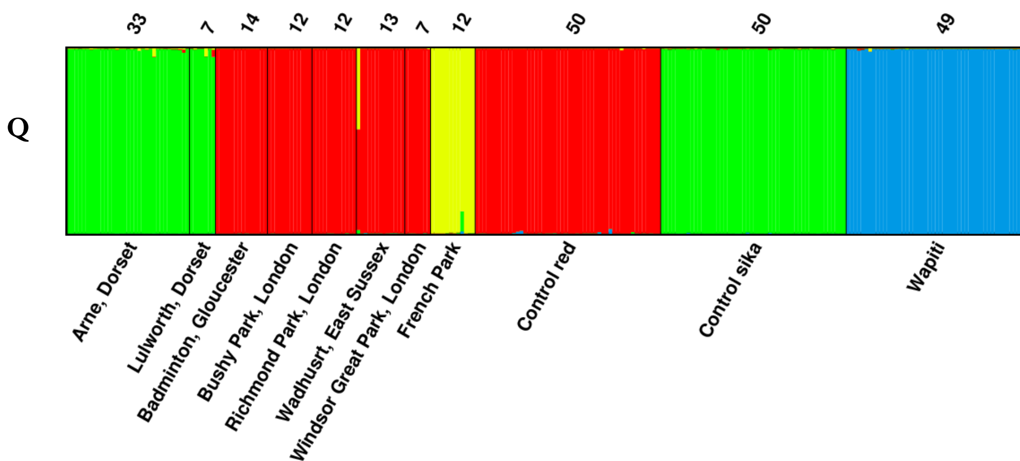


Figure 7.A2. Bar chart showing the results of *analysis 1* in STRUCTURE at K = 4 for each individual in the dataset consisting of red deer, sika and wapiti animals (n = 259). The Q value on the y-axis indicates the proportion of an individual's nuclear genome attributable to red ancestry (red) and the proportion attributable to sika ancestry (green) and that to wapiti ancestry (blue). Populations from where samples were collected are indicated in the x-axis (lower) and the number of animals sampled from each site on the upper x-axis.

8.0 References

Abernethy K (1994) The Establishment of a hybrid zone between red and sika deer (Genus Cervus). *Molecular Ecology* 3, 551-562.

Ackermann RR, Rogers J, Cheverud JM (2006) Identifying the morphological signatures of hybridization in primate and human evolution. *Journal of Human Evolution* 51, 632-645.

Agapow PM, Bininda-Emonds ORP, Crandall KA*, et al.* (2004) The impact of species concept on biodiversity studies. *Quarterly Review of Biology* 79, 161-179.

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics* 11, 697-709.

Allendorf FW, Leary RF, Spruell P, Wenburg JK (2001) The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution* 16, 613-622.

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 160, 1217-1229.

Arano B, Llorente G, Garciaparis M, Herrero P (1995) Species translocation menaces Iberian waterfrogs. *Conservation Biology* 9, 196-198.

Arnold J (1993) Cytonuclear disequilibria in Hybrid Zones. *Annual Review of Ecology and Systematics* 24, 521-554.

Arnold ML, Sapir Y, Martin NH (2008) Genetic exchange and the origin of adaptations: prokaryotes to primates. *Philosophical Transactions of the Royal Society B-Biological Sciences* 363, 2813-2820.

Asher GW, Archer JA, Scott IC*, et al.* (2005) Reproductive performance of pubertal red deer (Cervus elaphus) hinds: Effects of genetic introgression of wapiti subspecies on pregnancy rates at 18 months of age. *Animal Reproduction Science* 90, 287-306.

Bachanek J, Postawa T (2010) Morphological evidence for hybridization in the sister species Myotis myotis and Myotis oxygnathus (Chiroptera: Vespertilionidae) in the Carpathian Basin. *Acta Chiropterologica* 12, 439-448.

Baird NA, Etter PD, Atwood TS*, et al.* (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *Plos One* 3, 7.

Barton NH, Hewitt GM (1989) Adaptation, Speciation and Hybrid Zones. *Nature* 341, 497-503.

Bartos L, Hyanek J, Zirovnicky J (1981) Hybridisation between red and sika deer, 1. Craniological analysis. *Zoologischer Anzeiger* 207, 260-270.

Bartos L, Zirovnicky J (1981) Hybridisation between red and sika deer. 2. Phenotype analysis. *Zoologischer Anzeiger* 207, 271-287.

Bartos L, Zirovnicky J (1982) Hybridisation between red and sika deer. 3. Interspecific behavior. *Zoologischer Anzeiger* 208, 30-36.

Baum DA, Donoghue MJ (1995) Choosing among alternative phylogenetic species concepts. *Systematic Botany* 20, 560-573.

Beebee TJC, Rowe G (2004) *An introduction to molecular ecology* Oxford University Press, 198 Madison Avenue, New York, NY, 10016, USA.

Bert TM, Hesselman DM, Arnold WS*, et al.* (1993) High frequenct of gonadal neoplasia in a hard clam (Mercenaria spp.) hybrid zone. *Marine Biology* 117, 97-104.

Biedrzycka A, Solarz W, Okarma H (2012) Hybridization between native and introduced species of deer in Eastern Europe. *Journal of Mammalogy* 93, 1331-1341.

Bohm M, White PCL, Daniels MJ, *et al.* (2006) The health of wild red and sika deer in Scotland: An analysis of key endoparasites and recommendations for monitoring disease. *Veterinary Journal* 171, 287-294.
Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology & Evolution* 18, 249-256.
Carden RF, Carlin CM, Marnell F, *et al.* (2010) Distribution and range expansion of deer in Ireland. *Mammal Review* 41, 313-325.
Carden RF, McDevitt AD, Zachos FE, *et al.* (2012) Phylogeographic, ancient DNA, fossil and morphometric analyses reveal ancient and modern introductions of a large mammal: the complex case of red deer (Cervus elaphus) in Ireland. *Quaternary Science Reviews* 42, 74-84.
Carr SM, Hughes GA (1993) Direction of introgressive hybridisation between species of North-American deer (*Odocoileus*) as inferred from mitochondrial-cytochrome-B sequences. *Journal of Mammalogy* 74, 331-342.
Cathey JC, Bickham JW, Patton JC (1998) Introgressive hybridization and nonconcordant evolutionary history of maternal and paternal lineages in North American deer. *Evolution* 52, 1224-1229.
Chadwick AH, Ratcliffe PR, Abernethy K (1996) Sika deer in Scotland: Density, population size, habitat use and fertility: Some comparisons with red deer. *Scottish Forestry* 50, 8-16.
Chapter 2.
Chapter 3.
Charpentier MJE, Tung J, Altmann J, Alberts SC (2008) Age at maturity in wild baboons: genetic, environmental and demographic influences. *Molecular Ecology* 17, 2026-2040.
Childs MR, Echelle AA, Dowling TE (1996) Development of the hybrid swarm between pecos pupfish (Cyprinodontidae: Cyprinodon pecosensis) and sheepshead minnow (Cyprinodon variegatus): A perspective from allozymes and mtDNA. *Evolution* 50, 2014-2022.
Clarke B, Johnson MS, Murray J (1998) *How 'molecular leakage' can mislead us about island speciation.* Oxford University Press, 200 Madison Avenue, New York, New York 10016, USA; Oxford University Press, Walton Street, Oxford OX2 6DP, England.
Clarke CMH (1972) Red deer in the Northern South Island region - Their early impact. *New Zealand Journal of Forestry* 17, 37-42.
Clutton-Brock TH, Albon SD, Guinness FE (1989) Fitness costs of gestation and lactation in wild mammals. *Nature* 337, 260-262.
Clutton-Brock TH, Coulson T, Milner JM (2004) Red deer stocks in the Highlands of Scotland. *Nature* 429, 261-262.
Coates BS, Sumerford DV, Miller NJ, *et al.* (2009) Comparative Performance of Single Nucleotide Polymorphism and Microsatellite Markers for Population Genetic Analysis. *Journal of Heredity* 100, 556-564.
Coltman DW, Bowen WD, Wright JM (1998) Birth weight and neonatal survival of harbour seal pups ape positively correlated with genetic variation measured by microsatellites. *Proceedings of the Royal Society B-Biological Sciences* 265, 803-809.
Cook CE, Wang Y, Sensabaugh G (1999) A mitochondrial control region and cytochrome b phylogeny of sika deer (Cervus nippon) and report of tandem repeats in the control region. *Molecular Phylogenetics and Evolution* 12, 47-56.

Couceiro L, Lopez L, Ruiz JM, Barreiro R (2012) Population structure and range expansion: the case of the invasive gastropod Cyclope neritea in northwest Iberian Peninsula. *Integrative Zoology* 7, 286-298.

Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends in Ecology & Evolution* 15, 290-295.

Crawley MJ (2007) The R Book *John Wiley & Sons, Ltd*.

Davidson MM (1973) Characteristics, liberation and diseprsal of sika deer, *Cervus nippon,* in New Zealand. *New Zealand Journal of Forestry Science* 3, 153-180.

DCS (2008) Scotland's Wild Deer: A National Approach *Deer Commission for Scotland*, 9.

Delap P (1936) Deer in Wicklow. *The Irish Naturalists' Journal* VI, 82 - 89.

DeNicola AJ, Kesler DJ, Swihart RK (1997) Dose determination and efficacy of remotely delivered norgestomet implants on contraception of white-tailed deer. *Zoo Biology* 16, 31-37.

Díaz A, Hughes S, Putman R, Mogg R, Bond JM (2006) A genetic study of sika (Cervus nippon) in the New Forest and in the Purbeck region, southern England: is there evidence of recent or past hybridization with red deer (Cervus elaphus)? *Journal of Zoology* 270, 227-235.

Donoghue MJ (1985) A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist* 88, 172-181.

Ellstrand NC (2003) Current knowledge of gene flow in plants: implications for transgene flow. *Philosophical Transactions of the Royal Society B-Biological Sciences* 358, 1163-1170.

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.

Facon B, Genton BJ, Shykoff J*, et al.* (2006) A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology & Evolution* 21, 130-135.

Fain SR, Straughan DJ, Taylor BF (2010) Genetic outcomes of wolf recovery in the western Great Lakes states. *Conservation Genetics* 11, 1747-1765.

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

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

Fitzpatrick BM, Johnson JR, Kump DK*, et al.* (2010) Rapid spread of invasive genes into a threatened native species. *Proceedings of the National Academy of Sciences of the United States of America* 107, 3606-3610.

Forestry Commission (2005) Deer Management: Operational Guidance Booklet. Version 2.1.

Frantz AC, Hamann JL, Klein F (2008) Fine-scale genetic structure of red deer (Cervus elaphus) in a French temperate forest. *European Journal of Wildlife Research* 54, 44-52.

Fraser DJ, Lippe C, Bernatchez L (2004) Consequences of unequal population size, asymmetric gene flow and sex-biased dispersal on population structure in brook charr (Salvelinus fontinalis). *Molecular Ecology* 13, 67-80.

Gligor M, Ganzhorn JU, Rakotondravony D*, et al.* (2009) Hybridization between mouse lemurs in an ecological transition zone in southern Madagascar. *Molecular Ecology* 18, 520-533.

Glover R (1956) Notes on the Sika deer. *Jour Mammal* 37, 99-105.
Goodman SJ, Barton NH, Swanson G, Abernethy K, Pemberton JM (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.
Goodman SJ, Tamate HB, Wilson R*, et al.* (2001) Bottlenecks, drift and differentiation: the population structure and demographic history of sika deer (Cervus nippon) in the Japanese archipelago. *Molecular Ecology* 10, 1357-1370.
Grant PR, Grant BR (1992) Hybridisation of bird species. *Science* 256, 193-197.
Gray AP (1971) Mammalian Hybrids. A Check-list with Bibliography. *Commonwealth Agricultural Bureaux*, 155.
Grobler JP, Rushworth I, Brink JS*, et al.* (2011) Management of hybridization in an endemic species: decision making in the face of imperfect information in the case of the black wildebeest-Connochaetes gnou. *European Journal of Wildlife Research* 57, 997-1006.
Gui FR, Wan FH, Guo JY (2009) Determination of the population genetic structure of the invasive weed Ageratina adenophora using ISSR-PCR markers. *Russian Journal of Plant Physiology* 56, 410-416.
Guillot G (2008) Inference of structure in subdivided populations at low levels of genetic differentiation-the correlated allele frequencies model revisited. *Bioinformatics* 24, 2222-2228.
Guillot G, Estoup A, Mortier F, Cosson JF (2005a) A spatial statistical model for landscape genetics. *Genetics* 170, 1261-1280.
Guillot G, Mortier F, Estoup A (2005b) GENELAND: a computer package for landscape genetics. *Molecular Ecology Notes* 5, 712-715.
Guillot G, Santos F (2010) Using AFLP markers and the Geneland program for the inference of population genetic structure. *Molecular Ecology Resources* 10, 1082-1084.
Guillot G, Santos F, Estoup A (2008) Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics* 24, 1406-1407.
Gustavss I., Sundt CO (1969) 3 Polymorphic chromosome systems of centric fusion type in a population of Manchurian sika deer *(Cervus nippon hortulorum* Swinhoe). *Chromosoma* 28, 245-&.
Gyllensten U, Ryman N, Reuterwall C, Dratch P (1983) Genetic differentiation in 4 European subspecies of red deer (Cervus elaphus L). *Heredity* 51, 561-580.
Hails RS, Morley K (2005) Genes invading new populations: a risk assessment perspective. *Trends in Ecology & Evolution* 20, 245-252.
Hajji GM, Zachos FE, Charfi-Cheikrouha F, Hartl GB (2007) Conservation genetics of the imperilled Barbary red deer in Tunisia. *Animal Conservation* 10, 229-235.
Hall L, Topinka K, Huffman J, Davis L, Good A (2000) Pollen flow between herbicide-resistant Brassica napus is the cause of multiple-resistant B-napus volunteers. *Weed Science* 48, 688-694.
Hampton JO, Spencer PBS, Alpers DL*, 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.
Harrington (1979) Some Aspects of the Biology and Taxonomy of the Deer of the County Wicklow Region, Ireland. *National University of Ireland*.
Harrington R (1973) Hybridisation among deer and its implications for conservation. *Irish Forestry* 30, 64-78.

Harris SM, P., Wray, S. & Yalden, D. (1995) A review of British mammals: population estimates and conservation status of British mammals other than cetaceans. *Joint Nature Conservation Committee*.

Hauffe HC, Searle JB (1993) Extreme karyotypic variation in a *Mus musculus domesticus* hybrid zone - The Tobacco Mouse story revisited. *Evolution* 47, 1374-1395.

He C, Chen L, Simmons M*, et al.* (2003) Putative SNP discovery in interspecific hybrids of catfish by comparative EST analysis. *Animal Genetics* 34, 445-448.

Hermansen JS, Saether SA, Elgvin TO*, et al.* (2011) Hybrid speciation in sparrows I: phenotypic intermediacy, genetic admixture and barriers to gene flow. *Molecular Ecology* 20, 3812-3822.

Hewitt GM (1988) Hybrid Zones - Natural Laboratories for Evolutionary Studies. *Trends in Ecology & Evolution* 3, 158-167.

Hey J (2001) The mind of the species problem. *Trends in Ecology & Evolution* 16, 326-329.

Hey J, Waples RS, Arnold ML, Butlin RK, Harrison RG (2003) Understanding and confronting species uncertainty in biology and conservation. *Trends in Ecology & Evolution* 18, 597-603.

Hingston F (1988) Deer Parks and Deer of Gt Britain *Sporting Leisure and Press*, 87.

Hmwe SS, Zachos FE, Eckert I*, et al.* (2006a) Conservation genetics of the endangered red deer from Sardinia and Mesola with further remarks on the phylogeography of Cervus elaphus corsicanus. *Biological Journal of the Linnean Society* 88, 691-701.

Hmwe SS, Zachos FE, Sale JB, Rose HR, Hartl GB (2006b) Genetic variability and differentiation in red deer (Cervus elaphus) from Scotland and England. *Journal of Zoology* 270, 479-487.

Hobbs GI, Chadwick EA, Bruford MW, Slater FM (2011) Bayesian clustering techniques and progressive partitioning to identify population structuring within a recovering otter population in the UK. *Journal of Applied Ecology* 48, 1206-1217.

Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. *Molecular Ecology Resources* 11, 117-122.

Honda T, Kuwata H, Yamasaki S, Miyagawa Y (2011) A low-cost, low-labor-intensity electric fence effective against wild boar, sika deer, Japanese macaque and medium-sized mammals. *Mammal Study* 36, 113-117.

Huang L, Chi JX, Nie WH, Wang JH, Yang FT (2006) Phylogenomics of several deer species revealed by comparative chromosome painting with Chinese muntjac paints. *Genetica* 127, 25-33.

Hurst GDD, Jiggins FM (2005) Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings of the Royal Society B-Biological Sciences* 272, 1525-1534.

Isoda K, Shiraishi S, Watanabe S, Kitamura K (2000) Molecular evidence of natural hybridization between Abies veitchii and A-homplepis (Pinaceae) revealed by chloroplast, mitochondrial and nuclear DNA markers. *Molecular Ecology* 9, 1965-1974.

Jepsen BI, Siegismund HR, Fredholm M (2002) Population genetics of the native caribou (Rangifer tarandus groenlandicus) and the semi-domestic reindeer (Rangifer tarandus tarandus) in Southwestern Greenland: Evidence of introgression. *Conservation Genetics* 3, 401-409.

Jiggins CD, Salazar C, Linares M, Mavarez J (2008) Hybrid trait speciation and Heliconius butterflies. *Philosophical Transactions of the Royal Society B-Biological Sciences* 363, 3047-3054.

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403-1405.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *Bmc Genetics* 11.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16, 1099-1106.

Kameyama Y, Kudo G (2011) Clarification of the genetic component of hybrids between Phyllodoce caerulea and Phyllodoce aleutica (Ericaceae) in Hokkaido, northern Japan. *Plant Species Biology* 26, 93-98.

Kaneshiro KY, Val FC (1977) Natural hybridization between a sympatric pair of Hawaiian Drosophila. *American Naturalist* 111, 897-902.

Kanthaswamy S, Sakoski JA, Smith DG (2010) Detecting signatures of inter-regional and inter-specific hybridisation among the Chinese Rhesus macaque SPF population using single nucleotide polymorphic (SNP) markers. *Journal of Medical Primatology* 39, 280-281.

Karl SA, Bowen BW, Avise JC (1995) Hybridisation among the ancient mariners - Characterisation of marine turtle hybrids with molecular genetic assays. *Journal of Heredity* 86, 262-268.

Kidd AG, Bowman J, Lesbarreres D, Schulte-Hostedde AI (2009) Hybridization between escaped domestic and wild American mink (Neovison vison). *Molecular Ecology* 18, 1175-1186.

Konishi M, Takata K (2004) Impact of asymmetrical hybridization followed by sterile F-1 hybrids on species replacement in Pseudorasbora. *Conservation Genetics* 5, 463-474.

Kraaijeveld K (2000) The phylogenetic species concept and its place in modern evolutionary thinking. *Ardea* 88, 265-267.

Kuehn R, Schroeder W, Pirchner F, Rottmann O (2003) Genetic diversity, gene flow and drift in Bavarian red deer populations (Cervus elaphus). *Conservation Genetics* 4, 157-166.

Kuwayama R, Ozawa T (2000) Phylogenetic relationships among European red deer, wapiti, and sika deer inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution* 15, 115-123.

Lamb T, Avise JC (1986) Directional introgression of mitochondrial DNA in a hybrid population of tree frogs - The influence of mating behaviour. *Proceedings of the National Academy of Sciences of the United States of America* 83, 2526-2530.

Lancaster ML, Gemmell NJ, Negro S, Goldsworthy S, Sunnucks P (2006) Menage a trois on Macquarie Island: hybridization among three species of fur seal (Arctocephalus spp.) following historical population extinction. *Molecular Ecology* 15, 3681-3692.

Leary RF, Allendorf FW, Forbes SH (1993) Conservation Genetics of Bull Trout in the Columbia and Klamath River Drainages. *Conservation Biology* 7, 856-865.

Liu H, Yang G, Wei F-W, Li M, Hu J-c (2003) Sequence variability of the mitochondrial DNA control region and population genetic structure of sika deers (*Cervus nippon*) in China. *Acta Zoologica Sinica* 49, 53-60.

Liu ZJ, Ren BP, Wu RD*, et al.* (2009) The effect of landscape features on population genetic structure in Yunnan snub-nosed monkeys (*Rhinopithecus bieti*) implies an anthropogenic genetic discontinuity. *Molecular Ecology* 18, 3831-3846.

Lowe VPW, Gardiner AS (1975) Hybridisation between red deer (*Cervus elaphus*) and sika deer (*Cervus nippon*) with particular reference to stocks in NW England. *Journal of Zoology* 177, 553-566.

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

Madsen T, Shine R, Olsson M, Wittzell H (1999) Conservation biology - Restoration of an inbred adder population. *Nature* 402, 34-35.

Mahmut H, Masuda R, Onuma M*, et al.* (2002) Molecular phylogeography of the red deer (*Cervus elaphus*) populations in Xinjiang of China: Comparison with other Asian, European, and North American populations. *Zoological Science* 19, 485-495.

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

Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B-Biological Sciences* 363, 2971-2986.

Mank JE, Carlson JE, Brittingham MC (2004) A century of hybridization: Decreasing genetic distance between American black ducks and mallards. *Conservation Genetics* 5, 395-403.

Marie AD, Bernatchez L, Garant D (2011) Empirical assessment of software efficiency and accuracy to detect introgression under variable stocking scenarios in brook charr (Salvelinus fontinalis). *Conservation Genetics* 12, 1215-1227.

Mayle BA (1996) Progress in predictive management of deer populations in British woodlands. *Forest Ecology and Management* 88, 187-198.

Mayr E (1942) *Systematics and the Origin of Species from the Viewpoint of a Zoologist.* Columbia University Press, New York.

McDevitt AD, Edwards CJ, O'Toole P*, et al.* (2009a) Genetic structure of, and hybridisation between, red (*Cervus elaphus*) and sika (*Cervus nippon*) deer in Ireland. *Mammalian Biology* 74, 263-273.

McDevitt AD, Mariani S, Hebblewhite M*, et al.* (2009b) Survival in the Rockies of an endangered hybrid swarm from diverged caribou (*Rangifer tarandus*) lineages. *Molecular Ecology* 18, 665-679.

McDougal EI, Lowe VPW (1968) Transferrin Polymorphism and Serum Proteins of some British Deer. *Journal of Zoology* 155, 131-&.

McShea WJ, Monfort SL, Hakim S*, et al.* (1997) The effect of immunocontraception on the behavior and reproduction of white-tailed deer. *Journal of Wildlife Management* 61, 560-569.

Mech SG, Hallett JG (2001) Evaluating the effectiveness of corridors: a genetic approach. *Conservation Biology* 15, 467-474.

Meredith EP, Rodzen JA, Banks JD*, et al.* (2007) Microsatellite analysis of three subspecies of elk (*Cervus elaphus*) in California. *Journal of Mammalogy* 88, 801-808.

Miyazawa S, Okamoto M, Kondo S (2010) Blending of animal colour patterns by hybridization. *Nature Communications* 1.

Moore GH, Littlejohn RP (1989) Hybridisation of farmed wapiti (*Cervus elaphus manitobensis*) and red deer (*Cervus elaphus*). *New Zealand Journal of Zoology* 16, 191-198.

Moore WS, Price JT (1993) *Nature of selection in the Northern flicker hybrid zone and its implications for speciation theory.* Oxford Univ Press, New York.

Morin PA, Luikart G, Wayne RK, Grp SNPW (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution* 19, 208-216.

Nagata J, Masuda R, Kaji K, Kaneko M, Yoshida MC (1998) Genetic variation and population structure of the Japanese sika deer (*Cervus nippon*) in Hokkaido Island, based on mitochondrial D-loop sequences. *Molecular Ecology* 7, 871-877.

Nixon KC, Wheeler QD (1990) An amplification of the phylogenetic species concept. *Cladistics-the International Journal of the Willi Hennig Society* 6, 211-223.

Noor MAF, Grams KL, Bertucci LA, Reiland J (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences of the United States of America* 98, 12084-12088.

Nurnberger B, Barton N, MacCallum C, Gilchrist J, Appleby M (1995) Natural selection on quantitative traits in the Bombina hybrid zone. *Evolution* 49, 1224-1238.

Nussey DH, Coltman DW, Coulson T*, et al.* (2005) Rapidly declining fine-scale spatial genetic structure in female red deer. *Molecular Ecology* 14, 3395-3405.

Nussey DH, Pemberton J, Donald A, Kruuk LEB (2006) Genetic consequences of human management in an introduced island population of red deer (*Cervus elaphus*). *Heredity* 97, 56-65.

Okello JBA, Masembe C, Rasmussen HB*, et al.* (2008) Population genetic structure of savannah elephants in Kenya: Conservation and management implications. *Journal of Heredity* 99, 443-452.

PACEC (2006) Shooting Sports: Findings of an economic and environmental survey 1 - 23.

Parnell RJ (2002) Group size and structure in western lowland gorillas (*Gorilla gorilla gorilla*) at Mbeli Bai, Republic of Congo. *American Journal of Primatology* 56, 193-206.

Parsons TJ, Olson SL, Braun MJ (1993) Unidirectional spread of secondary sexual plumage traits across an avian hybrid zone. *Science* 260, 1643-1646.

Patton ML, Jochle W, Penfold LM (2007) Review of contraception in ungulate species. *Zoo Biology* 26, 311-326.

Pemberton JM, Jones, F. & Goodman, S. J. (2000) Unpublished report to Deer Commission for Scotland.

Pérez-Espona S, Hall RJ, Pérez-Barbería FJ*, et al.* (2013) The impact of past introductions on an iconic and economically important species, the red deer of Scotland. *The Journal of heredity* 104, 14-22.

Pérez-Espona S, Pemberton JM, Putman R (2009a) Red and sika deer in the British Isles, current management issues and management policy. *Mammalian Biology* 74, 247-262.

Pérez-Espona S, Pérez-Barbería FJ, Goodall-Copestake WP*, et al.* (2009b) Genetic diversity and population structure of Scottish Highland red deer (*Cervus elaphus*) populations: a mitochondrial survey. *Heredity* 102, 199-210.

Pérez-Espona S, Pérez-Barbería FJ, Jiggins CD, Gordon IJ, Pemberton JM (2010a) Variable extent of sex-biased dispersal in a strongly polygynous mammal. *Molecular Ecology* 19, 3101-3113.

Pérez-Espona S, Pérez-Barbería FJ, McLeod JE*, et al.* (2008) Landscape features affect gene flow of Scottish Highland red deer (*Cervus elaphus*). *Molecular Ecology* 17, 981-996.

Pérez-Espona S, Pérez-Barbería FJ, Pemberton JM (2010b) Assessing the impact of past wapiti introductions into Scottish Highland red deer populations using a Y chromosome marker. *Mammalian Biology* 76, 640-643.

Piertney SB, MacColl ADC, Bacon PJ, Dallas JF (1998) Local genetic structure in red grouse (*Lagopus lagopus scoticus*): evidence from microsatellite DNA markers. *Molecular Ecology* 7, 1645-1654.

Polziehn RO, Strobeck C (1998) Phylogeny of wapiti, red deer, sika deer, and other North American cervids as determined from mitochondrial DNA. *Molecular Phylogenetics and Evolution* 10, 249-258.

Powerscourt V (1884) On the acclimatization of the Japanese deer at Powerscourt. *Proceedings of the Zoological Society of London*, 207 - 209.

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

Ramsey J, Schemske DW (1998) Pathways, mechanisms, and rates of polyploid formation in flowering plants. In: *Annual Review of Ecology and Systematics* (ed. Fautin DG), pp. 467-501. Annual Reviews Inc. {a}, P.O. Box 10139, 4139 El Camino Way, Palo Alto, California 94306, USA.

Rand DM, Harrison RG (1989) Ecological genetics of a mosaic hybrid zone - Mitochondrial, nuclear and reproductive differentiation of crickets by soil type. *Evolution* 43, 432-449.

Randi E, Mucci N, Claro-Hergueta F, Bonnet A, Douzery EJP (2001) A mitochondrial DNA control region phylogeny of the Cervinae: speciation in *Cervus* and implications for conservation. *Animal Conservation* 4, 1-11.

Raphael BL, Kalk P, Thomas P*, et al.* (2003) Use of melengestrol acetate in feed for contraception in herds of captive ungulates. *Zoo Biology* 22, 455-463.

Ratcliffe PR (1987) Distribution and current status of sika deer, *Cervus nippon,* in Great Britain. *Mammal Review* 17, 39-58.

Rhymer JM, Simberloff D (1996) Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics* 27, 83-109.

Rhymer jMMJW, Braun MJ (1994) Mitochondrial analysis of gene flow between New Zealand mallards (*Anas platyrhynchos*) and grey ducks (*A. superciliosa*). *Auk* 111, 970-978.

Riley SPD, Pollinger JP, Sauvajot RM*, et al.* (2006) A southern California freeway is a physical and social barrier to gene flow in carnivores. *Molecular Ecology* 15, 1733-1741.

Ritz MS, Hahn S, Janicke T, Peter HU (2006) Hybridisation between South polar skua (*Catharacta maccormicki*) and Brown skua (*C. antarctica lonnbergi*) in the Antarctic Peninsula region. *Polar Biology* 29, 153-159.

Roach JC, Glusman G, Smit AFA*, et al.* (2010) Analysis of Genetic Inheritance in a Family Quartet by Whole-Genome Sequencing. *Science* 328, 636-639.

Rosenberg NA, Pritchard JK, Weber JL*, et al.* (2002) Genetic structure of human populations. *Science* 298, 2381-2385.

Schaid DJ, Guenther JC, Christensen GB*, et al.* (2004) Comparison of microsatellites versus single-nucleotide polymorphisms in a genome linkage screen for prostate cancer-susceptibility loci. *American Journal of Human Genetics* 75, 948-965.

Schlotterer C (2004) The evolution of molecular markers - just a matter of fashion? *Nature Reviews Genetics* 5, 63-69.

Schwenk K, Brede N, Streit B (2008) Introduction. Extent, processes and evolutionary impact of interspecific hybridization in animals. *Philosophical Transactions of the Royal Society B-Biological Sciences* 363, 2805-2811.

Senn HV (2009) Hybridisation between red deer (*Cervus elaphus*) and Japanese sika (*C. nippon*) on the Kintyre Peninsula, Scotland. *Unpublished PhD thesis, University of Edinburgh*.

Senn HV, Barton NH, Goodman SJ*, et al.* (2010a) Investigating temporal changes in hybridization and introgression in a predominantly bimodal hybridizing population of invasive sika (*Cervus nippon*) and native red deer (*C. elaphus*) on the Kintyre Peninsula, Scotland. *Molecular Ecology* 19, 910-924.

Senn HV, Pemberton JM (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.

Senn HV, Swanson GM, Goodman SJ, Barton NH, Pemberton JM (2010b) Phenotypic correlates of hybridisation between red and sika deer (genus *Cervus*). *Journal of Animal Ecology* 79, 414-425.

Shackell GH, Drew KR, Pearse AJT, Amer PR (2003) A cost-benefit analysis of the value of investment in a wapiti hybridisation research programme and the returns to New Zealand venison producers. *New Zealand Journal of Agricultural Research* 46, 133-140.

Shaw CN, Wilson PJ, White BN (2003) A reliable molecular method of gender determination for mammals. *Journal of Mammalogy* 84, 123-128.

Slate J, Coltman DW, Goodman SJ*, et al.* (1998) Bovine microsatellite loci are highly conserved in red deer (*Cervus elaphus*), sika deer (*Cervus nippon*) and Soay sheep (*Ovis aries*). *Animal Genetics* 29, 307-315.

Sommer RS, Zachos FE, Street M*, et al.* (2008) Late Quaternary distribution dynamics and phylogeography of the red deer (*Cervus elaphus*) in Europe. *Quaternary Science Reviews* 27, 714-733.

Stolting KN, Nipper R, Lindtke D*, et al.* (2013) Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. *Molecular Ecology* 22, 842-855.

Stolzenberg N, The BN, Salducci MD, Cavalli L (2009) Influence of Environment and Mitochondrial Heritage on the Ecological Characteristics of Fish in a Hybrid Zone. *Plos One* 4.

Swanson GM (2000) The Genetic and Phenotypic Consequences of Translocations of Deer (Genus *Cervus*) in Scotland. *Unpublished PhD thesis, University of Edinburgh*.

Swanson GM, Putman R (2009) *Sika Deer in the British Isles.* Springer, 233 Spring Street, New York, Ny 10013, United States.

Tammeleht E, Remm J, Korsten M*, et al.* (2010) Genetic structure in large, continuous mammal populations: the example of brown bears in northwestern Eurasia. *Molecular Ecology* 19, 5359-5370.

Thevenon S, Bonnet A, Claro F, Maillard JC (2003) Genetic diversity analysis of captive populations: The Vietnamese sika deer (*Cervus nippon pseudaxis*) in zoological parks. *Zoo Biology* 22, 465-475.

Thevenon S, Thuy LT, Ly LV*, et al.* (2004) Microsatellite analysis of genetic diversity of the Vietnamese sika deer (*Cervus nippon pseudaxis*). *Journal of Heredity* 95, 11-18.

Trotter PC, Behnke RJ (2008) The case for Humboldtensis: A subspecies name for the indigenous cutthroat trout (*Oncorhynchus clarkii*) of the Humboldt River, Upper Quinn River, and Coyote Basin drainages, Nevada and Oregon. *Western North American Naturalist* 68, 58-65.

Tung J, Charpentier MJE, Garfield DA, Altmann J, Alberts SC (2008) Genetic evidence reveals temporal change in hybridization patterns in a wild baboon population. *Molecular Ecology* 17, 1998-2011.

Twyford AD, Ennos RA (2012) Next-generation hybridization and introgression. *Heredity* 108, 179-189.

Ungerer MC, Baird SJE, Pan J, Rieseberg LH (1998) Rapid hybrid speciation in wild sunflowers. *Proceedings of the National Academy of Sciences of the United States of America* 95, 11757-11762.

Velickovic N, Djan M, Obreht D, Vapa L (2012) Population genetic structure of wild boars in the West Balkan region. *Russian Journal of Genetics* 48, 859-863.

Vigfusdottir F, Palsson S, Ingolfsson A (2008) Hybridization of glaucous gull (*Larus hyperboreus*) and herring gull (*Larus argentatus*) in Iceland: mitochondrial and microsatellite data. *Philosophical Transactions of the Royal Society B-Biological Sciences* 363, 2851-2860.

Vigilant L, Guschanski K (2009) Using genetics to understand the dynamics of wild primate populations. *Primates* 50, 105-120.

Vilà C, Walker C, Sundqvist AK*, et al.* (2003) Combined use of maternal, paternal and bi-parental genetic markers for the identification of wolf-dog hybrids. *Heredity* 90, 17-24.

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

Ward AI (2007) Trends in deer distribution and abundance within the UK. *Presentation to the Deer Initiative Conference "Deer, Habitats and Impacts", Buxton UK, 23 March 2007.*

Welsh A, Hill T, Quinlan H, Robinson C, May B (2008) Genetic assessment of lake sturgeon population structure in the Laurentian Great Lakes. *North American Journal of Fisheries Management* 28, 572-591.

Wheelwright NT, Mauck RA (1998) Philopatry, natal dispersal, and inbreeding avoidance in an island population of Savannah Sparrows. *Ecology* 79, 755-767.

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

Witherspoon DJ, Wooding S, Rogers AR*, et al.* (2007) Genetic similarities within and between human populations. *Genetics* 176, 351-359.

Wong MM, Cannon CH, Wickneswari R (2012) Development of high-throughput SNP-based genotyping in *Acacia auriculiformis x A. mangium* hybrids using short-read transcriptome data. *BMC genomics* 13, 726.

Wright S (1932) The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proc Sixth Internat Congr Genetics Ithaca New York* 1, 356-366.

Yang CC, Shoemaker DD, Wu WJ, Shih CJ (2008) Population genetic structure of the red imported fire ant, *Solenopsis invicta*, in Taiwan. *Insectes Sociaux* 55, 54-65.

Zachos FE, Hartl GB (2011) Phylogeography, population genetics and conservation of the European red deer *Cervus elaphus*. *Mammal Review* 41, 138-150.