Hybridisation between red deer (*Cervus elaphus*) and Japanese sika (*C. nippon*) on the Kintyre Peninsula, Scotland

Helen V. Senn

(BSc. Hons. University of St Andrews)

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 2009

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.

Helen Victoria Senn

Edinburgh

May 2009

Front cover photo credit: DCS

Abstract

Hybridisation between introduced and endemic species causes conservation concerns, but also provides us with an opportunity to study the dynamics of gene flow between two species as they first meet. Japanese sika deer (*Cervus nippon*) were introduced to the British Isles at a number of locations at the beginning of the 20th century. In the intervening time, sika have spread and their range now extends across approximately 40% of Scotland, where they overlap with that of native red deer (*C. elaphus*), with which they hybridise. In this study we focus on the consequences of one particular introduction that took place at Carradale, on the Kintyre Peninsula in 1893.

First, I assessed the current state of hybridisation using a sample of 735 red and sika deer samples collected in 2006/7 from forestry blocks throughout the Kintyre Peninsula. Genetic analysis was conducted with a panel of 22 highly differentiated microsatellite loci and one mtDNA marker. Population admixture analysis of the microsatellite data was conducted with the Bayesian clustering programme STRUCTURE. Over most of the study area, levels of introgression into red and sika deer were low and were consistent with a scenario of very occasional F1 hybridisation followed by backcrossing. There was, however, one forestry block where 43% of individuals could be defined as hybrids.

Second, I developed a branching process model of introgression via backcrossing, to assess whether variation in introgression across microsatellite loci could be interpreted as a signature of selection, or could in fact be attributed to stochastic processes. If only a few hybridisation events have contributed to the hybridising population, the pattern of introgression, even with a large number of genetic markers, will be highly stochastic. This pattern of neutral variation in introgression can have high enough variance that it could be mistaken for selection. Therefore, even if strong selection is acting, it may not be possible to distinguish its effects from neutral variation.

Third, I analysed trends in hybridisation and introgression over 15 years on the peninsula, through analysis of a dataset of 1513 red and sika deer samples at 20 microsatellite and a mtDNA marker. There was little evidence of change in the extent of hybridisation and introgression over time. MtDNA introgression was predominantly from red deer into sika. Recent introgression into sika on the peninsula can be explained by a very small number of F1 hybridisation events (~10) via analysis of the number of alleles that have introgressed from polymorphic red deer into the genetically homogenous sika population (a similar analysis cannot be conducted for introgression into red deer).

Finally, I conducted a regression analysis of genetic hybrid scores against phenotypic traits to assess the effect of hybridisation on phenotype. Hybridisation has caused changes in the weight of sika-like deer and red-like females. Hybridisation has caused changes in incisor arcade breadth of both populations and jaw length (a proxy for skeletal size) in sika-like females. However, there is no evidence that hybridisation has caused changes in kidney fat (a measure of condition) or pregnancy rates in either population.

In conclusion, even a small number of F1 hybridisation events can lead to extensive introgression and the timing and spatial distribution of these events is likely to have a large impact on the structure of a recently hybridising population - stochastic factors dominate both the distribution of hybrid individuals and the distribution of the genes that introgress following a hybridisation event. In red deer and sika deer, increasing phenotypic similarities of the two populations caused by hybridisation are likely to facilitate further breakdown between the two species. It is possible that breakdown in assortative mating between the two species could occur across their range.

Acknowledgements

First and foremost I would like to say a big thank you to my supervisors Josephine Pemberton and Nick Barton. They have not only given constant, enthusiasm, support, advice and encouragement, but also enormous amounts of incredibly swift feedback on many draft chapters. I really am incredibly grateful for all of this. I would also like to thank Josephine for all the additional support that she has provided and for taking me on field trips to Rum, Kintyre and St Kilda which count as some of the highlights during the time of my PhD.

Secondly this project would not have been possible without the help of the Forestry Commission Scotland. I am incredibly grateful to them for their careful and industrious help with collecting samples for me. In particular I would like to thank Colin Lavin, Kevin McKillop, Alasdair Cameron, John Jackson and Donnie Cameron for many useful discussions and much enthusiasm. I would also like to thank Helen Armstrong at Forest Research for much help with the Forestry Commission Larder Record database.

Thirdly I am incredibly grateful to Kate Abernethy, Simon Goodman and Graeme Swanson for allowing me to use their data for various parts of my thesis and for comments and advice.

I would like to thank Will Goodall-Copestake for spending a lot of time advising me on the development of microsatellite marker multiplexing at the start of my PhD. I also want to thank Silvia Pérez Espona for advice on the sampling of deer and for many other useful discussions. Without Silvia and Will's input I feel many lab-related things might have taken me much longer to get to grips with! Additionally I would like to thank the GenePool at IEB for expertly processing many Genemapper plates for me.

I thank my office mates Dario Beraldi and Lel Eory. Dario has been a constant source of support and help for many diverse issues. Lel Eory helped me on numerous occasions with problems relating to computer freeware and computers in general. Kati Csillery, Jonathan Coe and Sayanthan Logeswaran are thanked for many insightful discussions and comments on Chapter 2 and support with Mathematica software.

I thank Dan Nussey, Alastair Wilson, Jarrod Hadfield and Matt Robinson for statistical advice at various points. Most of all, I thank my father Stephen, who has given me a huge amount of statistical support. He particularly helped me with understanding the statistics necessary for Chapter 5 and also gave me advice on using Genstat. Any statistical blunders are my own!!

I thank Ellie, Tabi and Kath and many other IEB and IIR friends for lots of good times and supportive PhD chats!

I thank Victoria Senn and Edward Watson for a large amount of proof reading.

Finally, and most importantly, I thank my parents, Victoria and Stephen and my brother Mark for constant love, support and encouragement. I thank my grandparent Elizabeth and Teddy for always being interested. I also thank my husband Ed who is always supportive and loving and keeps me sane! Thank you.

Table of contents

Chapter 1

General Introduction

"Those forms which possess in some considerable degree the character of species, but which are so closely similar to some other forms, or are so closely linked to them by intermediate graduations, that naturalists do not like to rank them as distinct species, are in several respects the most important for us."

Charles Darwin 1859, Origin of Species (Chapter 2)

1.1 Hybridisation

1.1.1 The long view of hybridisation

Living things can be divided into assemblages of individuals that are similar in appearance, physiology, ecology and behaviour; these we call species. If a species could be followed back through its evolutionary history, we would see its form alter, becoming more similar and then indistinguishable from other related species. In practice of course we observe this, with some serendipity, in the fossil record (Benton & Pearson 2001). But if we scratch away the calciferous exterior from the branches of the tree of life, and view the genes beneath, the situation becomes a lot more complex. Whilst each gene can, in theory, be followed back in time, through its carriers (individual organisms that would be grouped at varying time points by a palaeontologist as belonging to different species) the backwards paths taken by all the genes found in a single individual in the present are not congruent. There are cases where the paths of genes disagree fundamentally with each other, where one gene's path breaks away from a group whose rough trajectory it had accompanied, and joins another group on a different trajectory. This 'horizontal gene transfer' across the branches of life is ubiquitous and exists both within and between the three domains of Bacteria, Archaea and Eukaryota (Lane & Archibald 2008; McInerney *et al.* 2008). When it occurs in the context of

eukaryotic sexually reproducing diploids, the process that causes it is known as hybridisation (Doyle 1992; Figure 1).

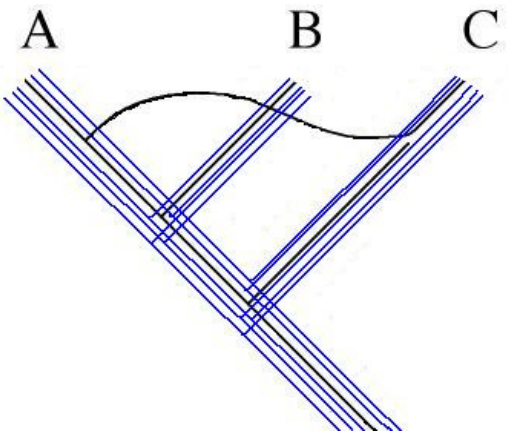


Figure 1: A gene tree view of hybridisation. Gene trees in blue are concordant and represent the general view of speciation between the three species A, B and C. A hybridisation event between A and C creates a gene tree that is non-concordant with the general view (black). Modified from a figure found at http://spot.colorado.edu/~am/Hybrid.jpg.

1.1.2 Defining species defines hybrids

Hybridisation appears to be relatively common in eukaryotic organisms. Obtaining objective estimates for the incidence of hybridisation amongst the taxa within this group is far from easy: firstly, because of taxonomic bias, the tendency of biologists to focus on certain taxa (Schwenk *et al.* 2008) and secondly because uneven application of the Biological Species Concept (Mayr 1963) is likely to confound the studies. If species are defined as "groups of actually or potentially interbreeding natural populations which are reproductively isolated from other such groups", then hybridisation is precluded by definition. The problem arises because in practice groups are often defined as species according to other criteria, such as phenotype, biogeography, or, according to other species concepts (e.g. Cracraft (1989) and

Mallet (1995)), the biological species concept is ignored, and so hybridisation is suddenly 'possible'. A recent survey of hybridisation in animal taxa, which controlled for research effort (but not different taxonomic practices) found the incidence of hybridisation to be in the range of 0.1-3% in all groups investigated (Schwenk *et al.* 2008). This is lower than previous estimates of around 1-25% in various groups of animals and plants that did not control for study effort (Grant & Grant 1992; Mallet 2005). In this thesis, hybridisation is defined as interbreeding between genetically distinct populations and therefore is considered to be synonymous with the term 'admixture'. Introgression is gene flow from one population into another following hybridisation. A hybrid is an individual carrying a (detectable) mixture of genes from the distinct populations.

Hybridisation has been the subject of numerous studies and a large number of reviews (e.g. Arnold (1992), Arnold & Hodges (1995), Barton & Hewitt (1985), Barton & Hewitt (1989), Buggs (2007), Dowling & Secor (1997), Grant & Grant (1992), Jiggins & Mallet (2000), Mallet (2005), Mallet (2008), Rhymer & Simberloff (1996), Rieseberg (1997) and Seehausen (2004)) and in 2006 around 425 publications dealt with the subject (Schwenk *et al.* 2008). Studies carried out on hybridisation can be classified into three broad and not mutually exclusive approaches which are outlined in the next subsections.

1.1.3 Hybrid zones reveal the nature of species boundaries

'Hybrid zones are narrow geographic regions in which populations meet, mate and produce hybrids' (Barton & Hewitt 1985). Most examples come from hybrid populations that appear to have formed under secondary contact, often seemingly due to re-expansion of species from glacial refugia (some examples with a seminal reference: chromosomal races of *Podisma pedestris* (Barton & Hewitt 1981) *Bombina bombina x B. variegata* (Szymura & Barton1986)*, Chorthippus parallelus parallelus* x *C. p. erythropus* (Butlin & Hewitt 1985) *, Mus musculus x M. domesticus* (Hunt & Selander 1973)), although for some, primary contact

(*in situ* origin) appears to be a more likely explanation (e.g. *Littorina saxatilis* morphs (Wilding *et al.* 2001)). In many hybrid zones (see examples above) species meet forming a boundary along which they hybridise creating parallel gradients in introgressing allele frequency (clines) that change across the hybrid zone. Hybrid zones may only be a few hundred metres wide but may stretch for hundreds of kilometres along the edge of the species range. At fine spatial scales (i.e. if the centre of the zone is examined) some zones exhibit a mosaic structure, in which species come in contact at the edges of interspersed patches of habitat (e.g. *Gryllus firmus x G. pensylvanicus* (Ross & Harrison 2002), some regions of the *Bombina* hybrid zone (Vines *et al.* 2003), *Allonemobuis socius x A. fasciatus* (Ross *et al.* 2008)). Regardless of its structure, the maintenance of a hybrid zone depends on the same forces: reproduction and dispersal of hybrids act to homogenise the two populations, widening the cline of introgression, whereas selection against hybrids acts to keep the two populations apart, narrowing it. At a single locus, the minimum distance over which an allele frequency can respond to selection (the characteristic scale, Slatkin (1973)) is $\sigma/s^{1/2}$, where σ^2 is the variance in distance between parent and offspring (dispersal rate) and s is the strength of selection. The width of the cline (1/maximum gradient) is roughly proportional to the characteristic scale (Barton 1979a). If selection against an introgressing allele at a particular locus is strong, this will have the effect of narrowing the clines of neutral alleles introgressing at other loci to an extent proportional to the linkage disequilibrium between them and the selected locus (Barton 1979b). Thus, even weak negative selection acting on multiple loci can produce a sharp gradient in the cline acting as a barrier to gene flow. However, providing F1 hybrids are fertile and introgression happens at all, then recombination in subsequent generations breaks up the associations between introgressing alleles and they will introgress into their new genetic background, subject only to their individual effects on fitness (Barton 1983). Even if negative selection is strong, advantageous and neutral genes can penetrate across the hybrid zone over time.

Because hybrid zones present scenarios of incomplete speciation, they offer an insight into the nature of reproductive isolation. One question centres on whether the nature of selection maintaining the zone is endogenous, arising from intrinsic hybrid inability, or whether selection is exogenous, arising from a mismatch between genotype and habitat. Zones maintained by endogenous selection are called tension zones and, in them, the centre of the cline, which is not fixed by any environmental gradient, is pulled towards areas of low population density (Barton 1979a; Barton & Hewitt 1981). Many hybrid zones appear, in fact, to be maintained by a mixture of endogenous and exogenous selection, and in animals, habitat preference (e.g. Cruzan & Arnold (1993), Grant & Grant (1992), Ross & Harrison (2002) and Vines *et al.* (2003)). In addition to this, reinforcement (Dobzhansky 1940), which is selection for assortative mating forced either by intrinsic hybrid inviability, or by hybrid-environment mismatch, appears to have a role in the maintenance of a number of hybrid zones, for example in *Drosophila pseudoobscura* x *D. persimilis* (Noor 1995) and *Mus musculus* x *Mus domesticus* (Smadja & Ganem 2005) and it seems likely to be a pervasive phenomenon (see Noor (1999) and Servedio & Noor (2003)) for reviews).

1.1.4 Hybrid speciation

Hybridisation can also result in the formation of new species by polyploidy (Otto & Whitton 2000). However, reproductive isolation can occur even without a change in chromosome number (homoploid hybrid speciation). This type of speciation was, at one point, considered to be rare, because of objections to how a new species could become established in the presence of its parent species, either because of ecological competition or because of genetic swamping from the parental taxa. However, homoploid hybridisation can give rise to reproductive isolation: hybrid offspring between *Helianthus annus* and *H. petiolaris* not only resemble other existing *Helianthus* species genetically, but are also better adapted than *H. annus* and *H petiolaris* to surviving in the extremely arid environments in which the other species exist (Rieseberg *et al.* 2003; Ungerer *et al.* 1998). Evidence for the role of homoploid

hybrid speciation in evolution is growing (see reviews by Jiggins *et al.* (2008), Mallet (2007), Mavarez & Linares (2008), Rieseberg (1997) and Seehausen (2004)). Movement of tension zones could also be a path to homoploid hybrid speciation in the absence of spatial segregation of hybrids from parents, via phase III of Wright's shifting balance theory (Wright 1931). Selectively advantageous gene combinations, newly arisen within the tension zone, could spread through a population through tension zone movement without being broken up by recombination with genes from the outside (Barton & Hewitt 1989). This theory has been hypothesised as the origin of diversity of *Heliconius* warning-colour morphs (Blum 2002; Mallet & Joron 1999; Turner & Mallet 1996), although the role of the shifting balance theory in species formation remains controversial (Coyne *et al.* 1997; Coyne *et al.* 2000).

Within the last 15 years, advances in molecular genetic techniques have allowed quantitative trait locus (QTL) approaches to investigate the genetic architecture of hybrid zones (i.e. identification of the number, effect size and location of genomic regions contributing to differentiation within and among populations (Rieseberg & Buerkle 2002)). For example such QTL studies have been carried out in *Helianthus* (Lexer *et al.* 2005; Rieseberg *et al.* 1999) *Iris* (Bouck *et al.* 2007; Martin *et al.* 2007; Martin *et al.* 2008) and *Coregonus* (Rogers *et al.* 2007) hybrid zones. Comparative microarray studies have also enabled identification of the likely genomic regions involved in reproductive isolation (Turner *et al.* 2005). See Chapter 3 for more detail on the methods used to analyse hybrid zones.

1.1.5 Hybridisation between invasive and endemic species

Human activity leads to hybridisation: hybridisation occurs because of deliberate or accidental movement of one species into the range of another, because of the presence of domesticated species in the ranges of wild counterparts, and because habitat destruction and climate change can bring previously isolated species together (Rhymer & Simberloff 1996).

Hybridisation between invasive and native species is, in essence, no different from natural hybridisation: dispersal, selection and drift all act to shape introgression as in any other hybridising taxa (Section 1.1.3). There are, however some general, but by no means universal, differences to consider: many natural hybrid zones appear to be stable, seemingly having existed for at least thousands of years (Barton & Hewitt 1985); in contrast, hybridisation resulting from alien introduction is recent. This means that gene flow between the two populations is also more recent and selection has had less time to shape it and promote processes such as reinforcement (Section 1.3). Humans may have changed the alien population through selective breeding (Beaumont *et al.* 2001; Hutchings & Fraser 2008; Kidd *et al.* 2009; Rand & Lucchini 2002), through artificial hybridisation prior to release or escape into the wild (Barilani *et al.* 2007) or through genetic engineering (Hails & Morley 2005). The invasive species may itself be a hybrid, arising from multiple introductions of distinct populations with novel variation that may be advantageous to the invader (Kolbe *et al.* 2007; Milne & Abbott 2000; Shields *et al.* 2008; Zardus & Hadfield 2005). Alternatively, introductions may consist of very few individuals and lead to founder effects (Tsutsui *et al.* 2000). The distribution of the invasive organism may reflect human intervention and movement as well as natural dispersal of the invasive species in its new environment (e.g. spread of invasive marine organisms along shipping routes (Grosholz 2002)). Arguably, human intervention can unite species from more disparate geographic locations than would occur naturally, and so these populations will potentially have spent longer evolving in allopatry and come from more diverse ecosystems than their counterparts in natural hybrid zones. This means that the introduced species will be involved in a barrage of novel ecological interactions, to which it, and the community into which it arrives, has had little time to adapt. Thus a complex interplay of gene exchange with the native taxa and ecological interactions with the native community will determine whether the invasive species establishes itself and how it spreads (Facon *et al.* 2006; Hastings *et al.* 2005) and the invasive species may only become successful following the acquisition of locally adapted

genes or new recombinant genotypes (Anderson & Stebbins 1954; Ellstrand & Schierenbeck 2000).

Hybridisation between invasive and endemic species provides us with an important evolutionary insight into species boundaries: for example studies of invasive species have shown that hybridisation can result in the extremely rapid formation of hybrid swarms. Hybridisation of the invasive sheepshead minnow (*Cyprinodon variegates*) with the endemic Pecos pupfish (*C. pecosensis*), led to formation of a hybrid swarm over a distance of more than 500km within 5 years of introduction (Childs *et al.* 1996). To observe such rapid evolution in a natural system would require extreme serendipity. In contrast to ancient hybrid zones, hybridisation between introduced and endemic species provides us with the opportunity to observe gene flow between two populations at an early stage before selection acts to stabilise hybridisation, if indeed it does so.

Many genetic studies of invasive hybridisation have, however, concentrated on practical questions relating to conservation: for example, on identifying hybrid individuals and levels of introgression within populations (e.g. Beaumont *et al.* (2001), Halbert *et al.* (2005), Miller *et al.* (2003), Munoz-Fuentes *et al.* (2007) and Randi & Lucchini (2002)), identifying the sources of introduction of the alien species (e.g. (Driscoll *et al.* 2007; Munoz-Fuentes *et al.* 2006)) and investigating the fitness of hybrids (McGinnity *et al.* 2003), in order to assess and monitor the extent of introgression within the native population, to better understand the history and ecology of the invasion and forecast the impact of hybridisation. Many of these studies have used Bayesian admixture analysis of variation at neutral molecular markers, such as microsatellites (Anderson & Thompson 2002; Corander & Marttinen 2006; Pritchard *et al.* 2000) to answer some of these questions, because they require the development of differentiated (Vaha & Primmer 2006), but not diagnostic molecular markers (see much more on this approach in Chapters 2 & 4).

1.2 Context of the thesis work

1.2.1 Phylogenetics of Cervus

The phylogenetic history of the genus *Cervus* is complex, heavily researched and a detailed description is beyond the scope of this introduction (Goodman *et al.* 2001; Kuwayama & Ozawa 2000; Ludt *et al.* 2004; Nagata *et al.* 1999; Pitra *et al.* 2004; Polziehn & Strobeck 1998; Polziehn & Strobeck 2002; Randi *et al.* 2001; Randi *et al.* 1998; Wu *et al.* 2004). In brief, *Cervus* has a ring distribution in the mid to high latitude of the northern hemisphere. The most recent common ancestor of red and sika deer probably originated in an area between northern India and Kyrgyzstan (Ludt *et al.* 2004) around 4.5 million years ago (fossil calibrated mtDNA chytochrome b (Pitra *et al.* 2004)). Divergence of westward migrating deer, aided by geographical barriers and repeated glaciations, resulted in the formation of the medium-sized *Cervus elaphus* (found in Europe, the Middle East and North Africa). The eastern clade became the progenitor of the physically large wapiti or elk *Cervus canadensis* (North Asia) and the physically small sika, including the diminutive Japanese sika, *Cervus nippon* (South East Asia). Thus the hybridisation of western European red deer with Japanese sika represents the uniting of two of the extreme ends of the *Cervus* distribution. See Figure 2 for photographs of red and sika deer and Chapters 2 and 5 for descriptions of other phenotypic differences.

Figure 2: Clockwise from top left, a red hind (Alison Morris), a sika hind (DCS), a sika stag with a red hind (DCS), a red stag (Alison Morris). Further details on phenotypic differences between the two species can be found in Chapter 2 & 5.

1.2.2 Population history of Cervus in Britain

Red deer recolonised the British Isles ~10 000 years ago after the end of the last glaciation (Sommer *et al.* 2008). Initially, extensive forestation meant red deer were widely distributed, but with the spread of neolithic farming culture across Britain (<5000 years ago) deforestation began and red deer populations became concentrated into remoter areas of upland and shrinking forests (Lister 1984; Whitehead 1964). By the late 18th century, despite the fact that their primary natural predators, *Ursa arctos* and *Canis lupus* were already extinct due to anthropogenic causes, red deer populations fell to a minimum: continued habitat loss and over-hunting by humans had led to their extinction across most of England, Wales and lowland Scotland (Clutton-Brock & Albon 1989). In the 19th century red deer populations grew again because an increased fashion for hunting, combined with the fact that sheep farming was becoming unprofitable, meant that many deer forests (reserves or estates, not necessarily with trees!) were established, particularly across Scotland (Clutton-Brock & Albon 1989). The tradition of stalking has been maintained up until the current day and is an important component of the rural economy, bringing in money both through stalking lets and through venison sales. The combination of the absence of substantial natural predation, management of deer populations at densities optimal for stalking and possibly reduced competition from sheep and milder climates has resulted in high population numbers (Clutton-Brock *et al.* 2004). Red deer populations can cause damage to forestry, preventing tree regeneration and growth (Scott *et al.* 2000). Additionally, deer grazing, especially at high density, can have adverse effects on specific plant communities (blanket bog, heath and coarse grassland), but deer presence, provided density is not excessively high, appears to have negligible to positive effects on the diversity of other upland plant communities (Albon *et al.* (2007), Virtanen *et al.* (2002) also reviewed in Perez-Espona *et al.* (in press)). Thus, many land owners are involved in deer management for conservation purposes or to protect forestry, whilst some combine these objectives with sport stalking (PACEC 2006). Currently, there are around 400 000 red deer across Scotland (Clutton-Brock *et al.* 2004).

Probably because of its abundance, the fact it is the largest extant British land mammal, and its association with traditional Scottish stalking culture, red deer (combined with the physically much smaller roe deer) were voted the most iconic animal species by the Scottish public in 2005 (Stewart 2006).

On many Scottish estates, over the years, deer management has involved the introduction of exotic red deer from other parts of Europe and English deer parks, in order to improve trophy quality (Whitehead 1964), which has undoubtedly led to admixture in some areas. For example, this is the likely origin of a mtDNA haplotype closely related to *C. elaphus corsicanus* (Corsican red deer), on the Isle of Rum, Scotland (Nussey *et al.* 2006). Another small study of seven British populations discovered that differentiation of mtDNA and microsatellite data between populations was generally high, pointing towards possible human introductions (Hmwe *et al.* 2006). Additionally, introductions of *Cervus canadensis* (formerly *C. elaphus canadensis*, North American wapiti) have occurred (Whitehead 1964). Wapiti have either died out or interbred with red deer as there is no evidence for phenotypically distinct populations in Scotland. A large scale survey of mtDNA variation of red deer in the central Scottish highlands found no evidence of introgression from wapiti (or indeed sika or central European and N Africa red deer). However, introductions may have been of males, so studies using Y-chromosome markers or admixture studies comparing red deer at multiple microsatellite markers to individuals from putative source populations are needed, in addition to mtDNA studies, to address these questions (Perez-Espona *et al.* 2009). Microsatellite studies of the same central Scottish samples revealed that geographic factors (isolation by distance and barriers) are responsible for population structure, but without comparison with samples from putative source populations, it is not possible to say whether geography is the only contributing factor to population structure (Perez-Espona 2008).

Sika are physically smaller than red deer (Table 1, Chapter 2) and so were not introduced as a means of improving red deer, but rather as an ornamental park deer. However, as a result of at least 34 introductions to the wild they now range across a large expanse of Great Britain (Ratcliffe 1987). They are particularly abundant in Scotland as the result of 12 separate introductions, nine of which appear to have resulted in established populations (Ratcliffe 1987). In Scotland, introductions were the results of escapes or deliberate releases from deer parks dating between c1870-1930 (see Ratcliffe (1987)) for a full account of introduction history). Viscount Powerscourt was the first person to introduce sika to the British Isles in about 1860. In a correspondence to the Zoological Society in London (1884) about a whole range of deer species he introduced into his park in County Wicklow, Ireland, he expounds the virtues of sika as a park deer and writes ominously: " [the] other animals all died off; and these pretty little deer are the only ones that have multiplied". These sika later directly stocked five parks in England and one in Scotland (Achanalt, Ross-shire) and probably indirectly stocked most other parks in the British Isles for which no records exist (Ratcliffe 1987). Only introductions to North Devon (now possibly extinct) and Peebles-shire are definitely known to have come directly form Japan and thus independently of Powerscourt (Ratcliffe 1987). Sika that come from Powerscourt may have a questionable origin as it is known that red deer, sambur (*Cervus unicolor*), axis deer (*Axis axis*), wapiti and Japanese sika were all kept together in an enclosure. Powerscourt reported hybridisation between red and sambur and red and sika (Powerscourt 1884), but hybridisation of the other species is also conceivable. Most wild sika in Britain are considered to be phenotypically of Japanese origin (Ratcliffe 1987). This was confirmed genetically for some populations when sika from Argyll, Fife, Peebles and Dorset were all found to be carrying mitochondrial haplotypes closely related to those found in southern Japan and microsatellite variation also placed them within Japanese sika (Goodman *et al.* 2001). There are also two populations that are possibly not of Japanese origin: sika from the Charing area, Kent (if this population still exists) are reputedly of Japanese/Manchurian hybrid origin (Whitehead 1964) and there appears to be a

population of Manchurian sika at the Teifi marshes in Wales as the result of a recent introduction or escape (personal communication L. Wilberforce).

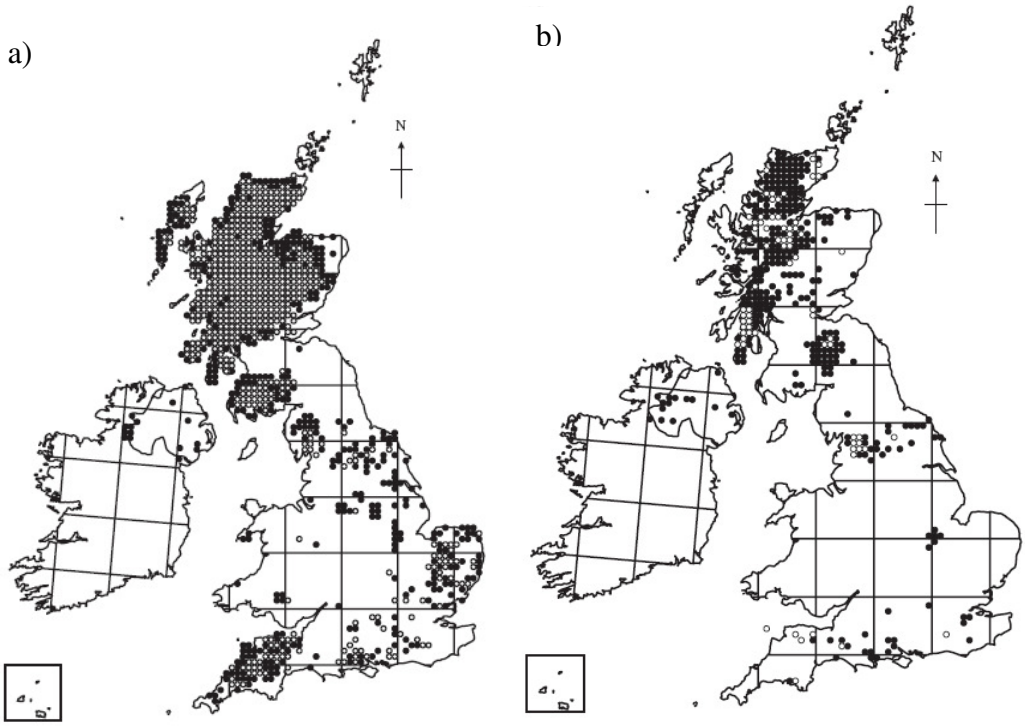


Figure 3: Distribution and range expansion of a) red deer and b) sika deer in Britain. The open circles represent observation made up to 1972. The closed circles represent additional observations made between 1973 and 2002. Modified from Ward (2005).

Since their introduction sika populations have expanded considerably and their range now occupies around 40% of the Scottish mainland (Figure 3) and isolated populations exist across the rest of Britain (Livingstone 2001; Ward 2005). Most of the sika populations in Britain seem to be confined to forested areas, although there are reports of sika on the open hill in Scotland (personal communication Colin McLean). Sika cause considerable damage to forestry through browsing and bark stripping in Britain (Lowe 1994; McLean 1993;

Swanson & Abernethy 2002; Swanson & Putman 2008), Ireland (Perrin *et al.* 2006) and New Zealand (Husheer *et al.* 2006). Currently around 5000 sika are culled annually in Scotland (www.dcs.gov.uk) of which the Forestry Commission Scotland culls 3000 (FCS larder record database).

1.2.3 Hybridisation between red deer and sika deer

Captive bred hybrids between red deer and Japanese sika do not appear to show abnormalities and although red and sika deer karyotypes differ by two Robertsonian rearrangements (2n sika = 64, 2n red = 68), there is no evidence for nondisjunction and the hybrids are fertile (Harrington 1979; Herzog & Harrington 1991).Harrington (1979) successfully crossed red deer males with sika females. He did not manage to cross sika males with red deer females as the males appeared to be unable to reach their vulvas. Whilst some reciprocal crosses occurred successfully, others caused mortality: crosses of particularly large red deer stags with sika hinds resulted in some female deaths from perforation of the vagina or anus and so only young red stags (aged 1-3) successfully sired F1 hybrids (Harrington 1979). However, it is worth mentioning that Harrington was using large park red deer and the very small-bodied sika from Killarney in his experiments. Below, I review previous studies of hybridisation by chronology and location, since the methods applied have been somewhat specific to each location.

Ireland

In County Wicklow, Ireland, a hybrid swarm between red deer and sika deer exists as a result of the stock which escaped (sika and most probably red-sika hybrids) from Powerscourt park around 1922 (see above) (Harrington 1973). Based on immunoelectrophoretic techniques and phenotypic observation Harrington (1979 & 1982) concluded that introgression between red deer and sika deer was extensive. A more recent,

fairly limited, genetic study using 8 randomly-chosen microsatellites and a mtDNA marker, of red and sika deer in Ireland, found evidence of hybridisation in Wicklow (McDevitt *et al.* in review). This study found both sika-like and red-like hybrids in Wicklow, and suggested that the red deer population was heavily introgressed, although sample sizes were fairly small for this locality (red deer=10, sika= 33). MtDNA introgression in both directions was discovered. The study surveyed other parts of Ireland and found no evidence of introgression.

The Lake District, England

Lowe and Gardiner (1975) carried out a craniological analysis of red and sika deer culled on Cartmel Fell in the Lake District, an area of putative hybrids derived from sika originally introduced at Rigmaden (late 19thC) and Park Nook (1907). When they examined the putative hybrids they found them to be intermediate in craniological characters to samples of supposedly pure red and sika deer collected from around Britain. They concluded from their analysis that there were no pure red deer remaining in the area. No subsequent studies of these deer have been conducted.

The Czech Republic

Studies of hybridisation in the former Czechoslovakia investigated hybridisation through phenotypic surveys and craniological analysis. These studies claim that hybrids were found across Western and Central Bohemia and Moravia (Czech Republic) and that a hybrid swarm exists at a game reserve in Janovice (Bartos *et al.* 1981; Bartos & Zirovnicky 1981). In a behavioural study, Bartos & Zirovnicky (1982) reported that during the rut, fights between red deer and sika deer were fairly frequent and that sika stags showed high levels of aggression towards red deer stags.

Scotland

The first genetic study of hybridisation was carried out on the Kintyre and Cowal Peninsulas, Argyll, Scotland, using two microsatellite markers, two allozyme markers and a mtDNA marker on a sample of 246 individuals (Abernethy 1994). The samples were later reanalysed at 11 diagnostic loci and the two allozyme markers were dropped because it was discovered they were not diagnostic (Goodman *et al.* 1999). Goodman *et al.* (1999) worked from the premise that although markers were apparently diagnostic on a test panel of 44 red and 44 sika of diverse geographic origins, it could not be assumed that the species shared no ancestral polymorphisms. So instead of calculating individuals hybrid indices for each animal, they assumed that individuals carrying multiple introgressed alleles were likely to be hybrids and used these individuals to estimate the extent of linkage disequilibrium in each population. From this they estimated the current rate of hybridisation to be 0.001-0.002 per generation or 1:500-1:1000 matings (indeed no F1 hybrids were found in the sample). Using this rate of hybridisation and models of population growth history of red and sika deer on the peninsula, the authors asked whether the levels of past introgression (individuals carrying single introgressed alleles) were consistent with those expected given current rates of hybridisation. Despite the fact that in one population up to 40% of individuals carried apparently introgressed alleles, levels of past introgression were 30-40% lower than those expected by recent hybridisation. This is indicative of possible negative selection against introgressing alleles, although a test for selection was not statistically significant. Further aspects of this study will be discussed throughout this thesis.

In the 1996-1997 stalking season, a large sample of 664 Kintyre deer was collected by Graeme Swanson and Simon Goodman through the Forestry Commission Scotland ranger network. Simon Goodman analysed the samples at 25 microsatellites chosen for red-sika differentiation and a mitochondrial marker. Although the genetic data were analysed using

the techniques of the Goodman *et al.* (1999) paper, the results remain unpublished. The microsatellite data from this study has been reanalysed and the results are incorporated into this thesis in Chapter 4. Graeme Swanson collected extensive post mortem information on the sampled animals and although this data was analysed in relation to hybrid index and appears in his thesis (Swanson 2000), this analysis again remains unpublished. The phenotypic data has been reanalysed and is incorporated into this thesis in Chapter 5.

Additional (unpublished) genetic studies of hybridisation between red deer and sika deer using up to ten differentiated microsatellite markers and one mtDNA marker reported evidence of introgression into red, sika or both at the Great Glen, Easter Ross, Sutherland, Loch Morar, Angus, Fife and the Borders (Swanson 2000).

New Forest and Purbeck, England

A study of red and sika deer from the New Forest and sika deer from Purbeck, Dorset, using eight differentiated microsatellite markers found limited evidence of introgression (Diaz *et al.* 2006). Multidimensional scaling of the genetic data (a type of clustering analysis) revealed that red and sika deer were distinct and that sika showed limited diversity in comparison to red deer. Analysis with the Bayesian clustering software STRUCTURE (Pritchard *et al.* 2000) and examination of alleles present in each population revealed possible low-level introgression within all populations investigated, although the results of this paper are not easy to interpret.

1.3 Objectives of this thesis

This thesis details a study of hybridisation between red deer (*Cervus elaphus*), native to Britain, and the introduced Japanese sika deer (*Cervus nippon*). The study focuses on the genetic and phenotypic effects of one particular introduction of sika deer at the Carradale Estate on the Kintyre Peninsula in 1893. The general aims of this study are both to add to the understanding of the evolutionary biology of hybridisation between two species as they first come into contact and to investigate the specific problems that hybridisation with sika pose for red deer in Britain. More specifically I aim to:

1. Describe the current genetic structure of red and sika deer populations on the Kintyre peninsula.
2. Investigate the potential for detecting selection on introgressing alleles in hybridising populations where F1 hybridisation is rare, for example as in the red-sika hybridisation study population on Kintyre.
3. Investigate how gene flow between red and sika populations in Kintyre has changed over time through comparison with previous studies (Goodman *et al.* 1999; Goodman et al unpublished data).
4. Investigate the extent to which hybridisation is affecting phenotypic traits in red and sika deer in Kintyre.

1.4 References

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

Albon SD, Brewer MJ, O'Brien S, Nolan AJ, Cope D (2007) Quantifying the grazing impacts associated with different herbivores on rangelands. *Journal of Applied Ecology* **44**, 1176-1187.

Anderson E, Stebbins GL (1954) Hybridization as an evolutionary stimulus. *Evolution* **8**, 378-388.
Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* **160**, 1217-1229.
Arnold ML (1992) Natural hybridization as an evolutionary process. *Annual Review of Ecology and Systematics* **23**, 237-261.
Arnold ML, Hodges SA (1995) Are natural hybrids fit or unfit relative to their parents?. *Trends in Ecology & Evolution* **10**, 67-71.
Barilani M, Bernard-Laurent A, Mucci N*, et al.* (2007) Hybridisation with introduced chukars (*Alectoris chukar*) threatens the gene pool integrity of native rock (*A. graeca*) and red-legged (*A. rufa*) partridge populations. *Biological Conservation* **137**, 57-69.
Barton NH (1979a) The dynamics of hybrid zones. *Heredity* **43**, 341-359.
Barton NH (1979b) Gene flow past a cline. *Heredity* **43**, 333-339.
Barton NH (1983) Multilocus clines. *Evolution* **37**, 454-471.
Barton NH, Hewitt GM (1981) A chromosomal cline in the grasshopper *Podisma pedestris. Evolution* **35**, 1008-1018.
Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**, 113-148.
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, Jena* **207**, 260-270.
Bartos L, Zirovnicky J (1981) Hybridization between red and sika deer II phenotype analysis. *Zoologischer Anzeiger* **207**, 271-287.
Bartos L, Zirovnicky J (1982) Hybridisation between red and sika deer 3. Interspecific behaviour. *Zoologischer Anzeiger, Jena* **208**, 30-36.
Beaumont M, Barratt EM, Gottelli D*, et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology* **10**, 319-336.
Benton MJ, Pearson PN (2001) Speciation in the fossil record. *Trends in Ecology & Evolution* **16**, 405-411.
Blum MJ (2002) Rapid movement of a *Heliconius* hybrid zone: Evidence for phase III of Wright's shifting balance theory? *Evolution* **56**, 1992-1998.
Bouck A, Wessler SR, Arnold ML (2007) QTL analysis of floral traits in Louisiana Iris hybrids. *Evolution* **61**, 2308-2319.
Buggs RJA (2007) Empirical study of hybrid zone movement. *Heredity* **99**, 301-312.
Butlin RK, Hewitt GM (1985) A hybrid zone between *Chorthippus parallelus parallelus* and *Chorthippus parallelus erythropus* (Orthoptera: Acrididae): behavioural characters. *Biological Journal of the Linnean Society* **26**, 287-299.
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.
Clutton-Brock TH, Albon SD (1989) *Red deer in the Highlands*, BSP Professional Books.
Clutton-Brock TH, Coulson T, Milner JM (2004) Red deer stocks in the Highlands of Scotland. *Nature* **429**, 261-262.
Corander J, Marttinen P (2006) Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology* **15**, 2833-2843.
Coyne JA, Barton NH, Turelli M (1997) Perspective: A critique of Sewall Wright's shifting balance theory of evolution. *Evolution* **51**, 643-671.
Coyne JA, Barton NH, Turelli M (2000) Is Wright's shifting balance process important in evolution? *Evolution* **54**, 306-317.

Cracraft J (1989) Speciation and its ontology the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. Otte D and JA Endler (Eds). *Speciation and Its Consequences. Sinauer Associates, Inc. Sunderland, Massachusetts, USA*.

Cruzan MB, Arnold ML (1993) Ecological and genetic associations in an *Iris* hybrid zone. *Evolution* **47**, 1432-1445.

Diaz 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.

Dobzhansky T (1940) Speciation as a stage in evolutionary divergence. *American Naturalist* **74**, 312-321.

Dowling TE, Secor CL (1997) The role of hybridization and introgression in the diversification of animals. *Annual Review of Ecology and Systematics* **28**, 593-619.

Doyle JJ (1992) Gene trees and species trees – molecular systematic as one-character taxonomy. *Systematic Botany* **17**, 144-163.

Driscoll CA, Menotti-Raymond M, Roca AL*, et al.* (2007) The Near Eastern origin of cat domestication. *Science* **317**, 519-523.

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7043-7050.

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

Goodman SJ, Barton NH, Swanson G, Abernethy K, Pemberton JM (1999) Introgression through rare hybridisation: 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) Hybridization of bird species. *Science* **256**, 193-197.

Grosholz E (2002) Ecological and evolutionary consequences of coastal invasions. *Trends in Ecology & Evolution* **17**, 22-27.

Hails RS, Morley K (2005) Genes invading new populations: a risk assessment perspective. *Trends in Ecology & Evolution* **20**, 245-252.

Halbert ND, Ward TJ, Schnabel RD, Taylor JF, Derr JN (2005) Conservation genomics: disequilibrium mapping of domestic cattle chromosomal segments in North American bison populations. *Molecular Ecology* **14**, 2343-2362.

Harrington R (1973) Hybridisation among deer and its implication for conservation. *Irish Forestry* **30**, 64-78.

Harrington R (1979) *Some Aspects of the Biology and Taxonomy of the Deer of the County Wicklow Region, Ireland.*, National University of Ireland.

Harrington R (1982) Hybridization of red deer (*Cervus elaphus L.* 1758) and Japanese sika deer (*Cervus nippon* Temminck, 1838). *Transactions of the International Congress of Game Biology* **14**, 559-571.

Hastings A, Cuddington K, Davies KF*, et al.* (2005) The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters* **8**, 91-101.

Herzog S, Harrington R (1991) The role of hybridization in the karyotype evolution of deer (Cervidae; Artiodactyla; Mammalia). *Theoretical and Applied Genetics* **82**, 425-429.

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

Hunt GW, Selander RK (1973) Biochemical genetics of hybridisation in European house mice. *Heredity* **31**, 11-33.

Husheer SW, Allen RB, Robertson AW (2006) Suppression of regeneration in New Zealand mountain beech forests is dependent on species of introduced deer. *Biological Invasions* **8**, 823-834.
Hutchings JA, Fraser DJ (2008) The nature of fisheries- and farming-induced evolution. *Molecular Ecology* **17**, 294-313.
Jiggins CD, Mallet J (2000) Bimodal hybrid zones and the scale of a snail - Reply. *Trends in Ecology & Evolution* **15**, 469-469.
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.
Kidd AG, Bowman J, Lesbarreres D, Schulte-Hostedde AI (2009) Hybridisation between escaped dometic and wild American mink (*Neovison vison*). *Molecular Ecology*.
Kolbe JJ, Larson A, Losos JB (2007) Differential admixture shapes morphological variation among invasive populations of the lizard *Anolis sagrei*. *Molecular Ecology* **16**, 1579-1591.
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.
Lane CE, Archibald JM (2008) The eukaryotic tree of life: endosymbiosis takes its TOL. *Trends in Ecology & Evolution* **23**, 268-275.
Lexer C, Rosenthal DM, Raymond O, Donovan LA, Rieseberg LH (2005) Genetics of species differences in the wild annual sunflowers, *Helianthus annuus* and *H. petiolaris*. *Genetics* **169**, 2225-2239.
Lister AM (1984) Evolutionary and ecological origins of British red deer. *Proceedings of the Royal Society of Edinburgh Section B-Biological Sciences* **82**, 205-229.
Livingstone SR (2001) *The Application of GIS to the Spread of Introduced Japanese Sika Deer (Cervus nippon) in Scotland.*, University of Edinburgh.
Lowe (1994) Deer Management: developing the requirements for the establishment of diverse coniferous and broadleaf forests. Coilte, Co.Wicklow, Ireland.
Ludt JC, Schroeder W, Rottmann O, Kuehn R (2004) Mitochondrial DNA phylogeography of red deer *(Cervus elaphus)*. *Molecular Phylogenetics and Evolution* **31**, 1064-1083.
Mallet J (1995) A species definition for the modern synthesis. *Trends in Ecology & Evolution* **10**, 294-299.
Mallet J (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* **20**, 229.
Mallet J (2007) Hybrid speciation. *Nature* **446**, 279-283.
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.
Mallet J, Joron M (1999) Evolution of diversity in warning color and mimicry: polymorphisms, shifting balance, and speciation. *Annual Review of Ecology and Systematics* **30**, 201-233.
Martin NH, Bouck AC, Arnold ML (2007) The genetic architecture of reproductive isolation in Louisiana irises: Flowering phenology. *Genetics* **175**, 1803-1812.
Martin NH, Sapir Y, Arnold ML (2008) The genetic architecture of reproductive isolation in louisiana irises: Pollination syndromes and pollinator preferences. *Evolution* **62**, 740-752.
Mavarez J, Linares M (2008) Homoploid hybrid speciation in animals. *Molecular Ecology* **17**, 4181-4185.
Mayr E (1963) *Animal Species and Evolution,* Belknap Press of Harvard University, Cambridge, MA.

McDevitt AD, Edward JE, O'Toole P*, et al.* (in review) Genetic structure of, and hybridisation between red (*Cervus elaphus*) and sika (*Cervus nippon*) deer in Ireland. *Mammalian Biology*.

McGinnity P, Prodohl P, Ferguson K*, et al.* (2003) Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 2443-2450.

McInerney JO, Cotton JA, Pisani D (2008) The prokaryotic tree of life: past, present ... and future? *Trends in Ecology & Evolution* **23**, 276-281.

McLean C (1993) Sika Deer Control. Red Deer Commision.

Miller CR, Adams JR, Waits LP (2003) Pedigree-based assignment tests for reversing coyote (*Canis latrans*) introgression into the wild red wolf (*Canis rufus*) population. *Molecular Ecology* **12**, 3287-3301.

Milne RI, Abbott RJ (2000) Origin and evolution of invasive naturalized material of *Rhododendron ponticum* L. in the British Isles. *Molecular Ecology* **9**, 541-556.

Munoz-Fuentes V, Green AJ, Sorenson MD, Negro JJ, Vila C (2006) The ruddy duck *Oxyura jamaicensis* in Europe: natural colonization or human introduction? *Molecular Ecology* **15**, 1441-1453.

Munoz-Fuentes V, Vila C, Green AJ, Negro JJ, Sorenson MD (2007) Hybridization between white-headed ducks and introduced ruddy ducks in Spain. *Molecular Ecology* **16**, 629-638.

Nagata J, Masuda R, Tamate HB*, et al.* (1999) Two genetically distinct lineages of the sika deer, *Cervus nippon*, in Japanese Islands: Comparison of mitochondrial D-Loop region sequences. *Molecular Phylogenetics and Evolution* **13**, 551-519.

Noor MA (1995) Speciation driven by natural selection in drosophila. *Nature* **375**, 674-675.

Noor MAF (1999) Reinforcement and other consequences of sympatry. *Heredity* **83**, 503-508.

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.

Otto SP, Whitton J (2000) Polyploid incidence and evolution. *Annual Review of Genetics* **34**, 401-437.

PACEC (2006) The contribution of deer management to the Scottish economy. Public and Corporate Economic Consultants, London and Cambridge.

Perez-Espona S, Pemberton JM, Putman R (in press) Red and sika deer in the British Isles, current management issues and management policy. *Mammalian Biology*.

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

Perez-Espona S, Perez-Barberia FJ, McLeoad JE, Jiggins CD, Gordon IJ, Pemberton JM (2008) Landscape features affect gene flow in Scottish Highland red deer. *Molecular Ecology* **17,** 981-996.

Perrin PM, Kelly DL, Mitchell FJG (2006) Long-term deer exclusion in yew-wood and oakwood habitats in southwest Ireland: Natural regeneration and stand dynamics. *Forest Ecology and Management* **236**, 356-367.

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

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.

Polziehn RO, Strobeck C (2002) A phylogenetic comparison of red deer and wapiti using mitochondrial DNA. *Molecular Phylogenetics and Evolution* **22**, 342-356.

Powerscourt V (1884) On the acclimatization of the Japanese deer at Powerscourt. *Proceedings of the Zoological Society of London* **1884**, 207-209.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
Randi E, Lucchini V (2002) Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. *Conservation Genetics* **3**, 31-45.
Randi E, Mucci N, Carlo-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.
Randi E, Mucci N, Pierpaoli M, Douzery E (1998) New phylogenetic perspectives on the Cervidae (Artiodactyla) are provided by the mitochondrial cytochrome b gene. *Proceedings of the Royal Society (London) B* **265**, 793-801.
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 hybridisation and introgression. *Annual Review of Ecology and Systematics* **27**, 83-109.
Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review of Ecology and Systematics* **28**, 359-389.
Rieseberg LH, Buerkle CA (2002) Genetic mapping in hybrid zones. *American Naturalist* **159**, S36-S50.
Rieseberg LH, Raymond O, Rosenthal DM*, et al.* (2003) Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* **301**, 1211-1216.
Rieseberg LH, Whitton J, Gardner K (1999) Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. *Genetics* **152**.
Rogers SM, Isabel N, Bernatchez L (2007) Linkage maps of the dwarf and normal lake whitefish (*Coregonus clupeaformis*) species complex and their hybrids reveal the genetic architecture of population divergence. *Genetics* **175**, 375-398.
Ross CL, Benedix JH, Garcia C*, et al.* (2008) Scale-independent criteria and scale-dependent agents determining the structure of a ground cricket mosaic hybrid zone (*Allonemobius socius - Allonemobius fasciatus*). *Biological Journal of the Linnean Society* **94**, 777-796.
Ross CL, Harrison RG (2002) A fine-scale spatial analysis of the mosaic hybrid zone between *Gryllus firmus* and *Gryllus pennsylvanicus*. *Evolution* **56**, 2296-2312.
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.
Scott D, Welch D, Thurlow M, Elston DA (2000) Regeneration of *Pinus sylvestris* in a natural pinewood in NE Scotland following reduction in grazing by *Cervus elaphus*. *Forest Ecology and Management* **130**, 199-211.
Seehausen O (2004) Hybridization and adaptive radiation. *Trends In Ecology & Evolution* **19**, 198-207.
Servedio MR, Noor MAF (2003) The role of reinforcement in speciation: Theory and data. *Annual Review of Ecology Evolution and Systematics* **34**, 339-364.
Shields JL, Barnes P, Heath DD (2008) Growth and survival differences among native, introduced and hybrid blue mussels (*Mytilus* spp.): genotype, environment and interaction effects. *Marine Biology* **154**, 919-928.
Slatkin M (1973) Gene flow and selection in a cline. *Genetics* **75**, 733-756.
Smadja C, Ganem G (2005) Asymmetrical reproductive character displacement in the house mouse. *Journal of Evolutionary Biology* **18**, 1485-1493.
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.

Stewart D (2006) Scottish biodiversity list social criterion: results of a survey of the Scottish population. In: *Scottish Executive Social Research Findings No 26*.
Swanson GM (2000) *The Genetic and Phenotypic Consequences of Translocations of Deer (Genus Cervus) in Scotland*, Unpublished PhD thesis, University of Edinburgh.
Swanson GM, Abernethy K (2002) Sika deer in Scotland. *Report to the Deer Commission for Scotland*.
Swanson GM, Putman RJ (2008) Sika Deer in the British Isles. In: *Sika Deer: Biology, Conservation and Management* (eds. McCullough DR, Takatsuki S, Kaji K). Springer, Japan.
Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*, near Cracow in Southern Poland. *Evolution* **40**, 1141-1159.
Tsutsui ND, Suarez AV, Holway DA, Case TJ (2000) Reduced genetic variation and the success of an invasive species. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 5948-5953.
Turner JRG, Mallet JLB (1996) Did forest islands drive the diversity of warningly coloured butterflies? Biotic drift and the shifting balance. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **351**, 835-845.
Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *Plos Biology* **3**, 1572-1578.
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.
Vaha J-P, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15**, 63-72.
Vines TH, Ghira I, Koehler SC*, et al.* (2003) On the maintenance of reproductive isolation in a mosaic hybrid zone between the toads *Bombina bombina* and *B. variegata*. *Evolution*.
Virtanen R, Edwards GR, Crawley MJ (2002) Red deer management and vegetation on the Isle of Rum. *Journal of Applied Ecology* **39**, 572-583.
Ward AI (2005) Expanding ranges of wild and feral deer in Great Britain. *Mammal Review* **35**, 165-173.
Whitehead GK (1964) *The Deer of Great Britain and Ireland* Routledge and Kegan Paul, London.
Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *Journal of Evolutionary Biology* **14**, 611-619.
Wright S (1931) Evolution in Mendelian populations. *Genetics* **16**, 97-159.
Wu H, Wan Q-H, Fang S-G (2004) Two genetically distinct units of the Chinese sika deer (*Cervus nippon*): analyses of mitochondrial DNA variation. *Biological Conservation*, 183-190.
Zardus JD, Hadfield MG (2005) Multiple origins and incursions of the Atlantic barnacle *Chthamalus proteus* in the Pacific. *Molecular Ecology* **14**, 3719-3733.

Chapter 2

Variable extent of hybridisation between invasive sika (*Cervus nippon)* and native red deer (*C. elaphus*) in a small geographic area.

Helen V. Senn [1 2], Josephine M. Pemberton[1]

[1] Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JT, UK

[2] Macaulay Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK

Corresponding author:

Helen V. Senn
Institute of Evolutionary Biology,
West Mains Road
Edinburgh, EH9 3JT, UK

Tel: 0131 6513612
Fax: 0131 5066564
Email: h.v.senn@sms.ed.ac.uk

Authors' contributions:

HVS organised the sampling, carried out the laboratory work, performed the statistical analysis and wrote the manuscript. JMP supervised the project and commented on all draft manuscripts.

This chapter has been published as a paper with the same title in the journal Molecular Ecology, 2009, volume 18, pages 862-876.

2.1 Abstract

Instances of hybridisation between endemic and alien species pose a threat to species integrity, but also provide us with an opportunity to study the dynamics of gene flow between two species as they first meet. Here we used variation at 22 highly differentiated microsatellite loci and one mtDNA marker in a sample of 735 individuals, to investigate the genetic consequences of a Japanese sika deer (*Cervus nippon*) introduction for native red deer (*C. elaphus*) on the Kintyre Peninsula, Scotland. We investigated population structure, estimated null allele frequency and assigned individual hybrid scores using a Bayesian clustering algorithm implemented in STRUCTURE 2.2. The dataset clearly divided into two clusters and generally introgression into red and sika was low. However at one site, West Loch Awe, 43% of individuals were hybrids. MtDNA introgression indicated that hybridisation was occurring between red deer hinds and sika deer stags. We argue that the pattern of differential introgression across the study area is primarily due to the rarity of hybridisation events between the two species and the limited time the two species have been in contact (<120 years). This contrasts with the causes of classic mosaic hybrid zones (selection induced by habitat variability). Currently it seems possible that, in time, the level of hybridisation found at West Loch Awe could also be found across the whole of the peninsula.

2.2 Introduction

Hybridisation of exotic species with natives provides us with unequivocal examples of the early dynamics of hybridisation and gene flow following secondary contact. Upon introduction of an exotic species, it is not immediately apparent what its fate (and ecological impact) will be and many factors such as life history traits, pre-adaptation to new habitat, genetic compatibility and evolution subsequent to introduction, as well as chance appear to play a role in determining if, and how, invasion and hybridisation proceed (Ellstrand & Schierenbeck 2000; Mack *et al.* 2000). In many ancient hybrid zones it has been possible to establish, at least some, of the factors governing gene flow (e.g. Yanchukov *et al.* 2006; Ross & Harrison 2002; Johnston *et al.* 2001). Ancient hybrid zones however, represent one 'end point' in the process of invasion (others being total admixture, or rapid reinforcement) and so, to some extent, may be limited in their potential to explain the early dynamics of gene flow *en route* to any outcome of hybridisation including the early stages in the formation of hybrid zones themselves. Here we describe an example of recent hybridisation and introgression between native red deer (*Cervus elaphus*) and introduced sika deer (*C. nippon*) in contact on the Kintyre Peninsula, Scotland, for ~115 years. We show how, currently, the structure of gene flow between the two populations varies considerably within a relatively small geographic area. We hypothesise that this is mainly due to the recent contact of these two species and discuss this finding in the light of hybrid zone theory and also discuss its specific impacts on red deer conservation in Scotland.

2.2.1 Natural history

Red deer are native to the British Isles and are the largest extant land mammal. As such, red deer are generally perceived as an iconic native species, especially in Scotland which is home to Europe's largest population (300 000 - 400 000; Clutton-Brock *et al.* 2004).

Landowners in Scotland have a complex economic relationship with red deer: due to the lack of natural predators local densities of deer can be high and animals are culled to protect crops and the natural heritage. On the other hand, deer shooting for sport (stalking) is a source of revenue and many landowners combine management for control and sport (PACEC 2006).

Most wild sika in Britain are considered to be of Japanese origin and their presence in the wild is as a result of accidental or deliberate releases from deer parks (Goodman *et al.* 2001; Ratcliffe 1987). Fossil-calibrated molecular clock estimates for the divergence dates between sika and Western European red deer range from 5.2 to 7 million years ago (Ludt *et al.* 2004; Randi *et al.* 2001). The two species exhibit many physical and behavioural differences (Table 1). Notably, Scottish red deer are much larger than Japanese sika, weighing approximately twice as much (Harris & Yalden 2008). Despite these differences, morphological and molecular studies have shown that these species hybridise in captivity (Harrington 1973) and in the wild in the UK, the Czech Republic and New Zealand (Goodman *et al.* 1999; Harrington 1979; Lowe & Gardiner 1975; Davidson 1973; Bartos *et al.* 1981). However, a study based on 8 microsatellite markers, of deer in the New Forest, Hampshire and Purbeck, Dorset, UK found little compelling evidence of hybridisation between red and sika deer (Diaz *et al.* 2006).

In 1893, nine female and two male Japanese sika were released onto Carradale Estate on the Kintyre Peninsula (Argyll, West coast of Scotland, Figure 1) and they escaped from their poorly fenced enclosure soon after. The putative Japanese origin of these animals has been verified by a study of microsatellite (9 loci) and mitochondrial control region variation which places Kintyre sika with those from the island of Kyushu, Japan (Goodman *et al.* 2001). By 1925 the number of feral sika had increased enough for them to be considered an annoyance and require culling. By 1937 the population was estimated at 300-400 individuals

(Whitehead 1964). Subsequent to the escape, the sika population is estimated to have, increased by an average of 9.2% per year and expanded its range northwards at 3.7 km per year (Goodman *et al.* 1999). At the time of escape, red deer were apparently at low density in the south of the peninsula, but their numbers have increased since, probably through an influx from the north (Ratcliffe 1987; Whitehead 1964). Red and sika ranges probably came into permanent contact in the region of Knapdale (site 5, Figure 1) in the 1960's and by 1975 sika stags were being shot at Eredine (site 10, Figure 1) (Ratcliffe 1987).

Table 1: Some phenotypic differences between red and sika deer. Local variation in size and markings of animals occur; these differences are an approximate guide to typical animals found in Scotland. (After Harrington 1973 and Whitehead 1964)

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: summer:	Grey-brown to dark-brown Chestnut red	Grey-brown, spots invisible 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 twelve points or more	Up to eight points
Rutting call (males only)	Deep roar	Whistle, sometimes high pitched scream

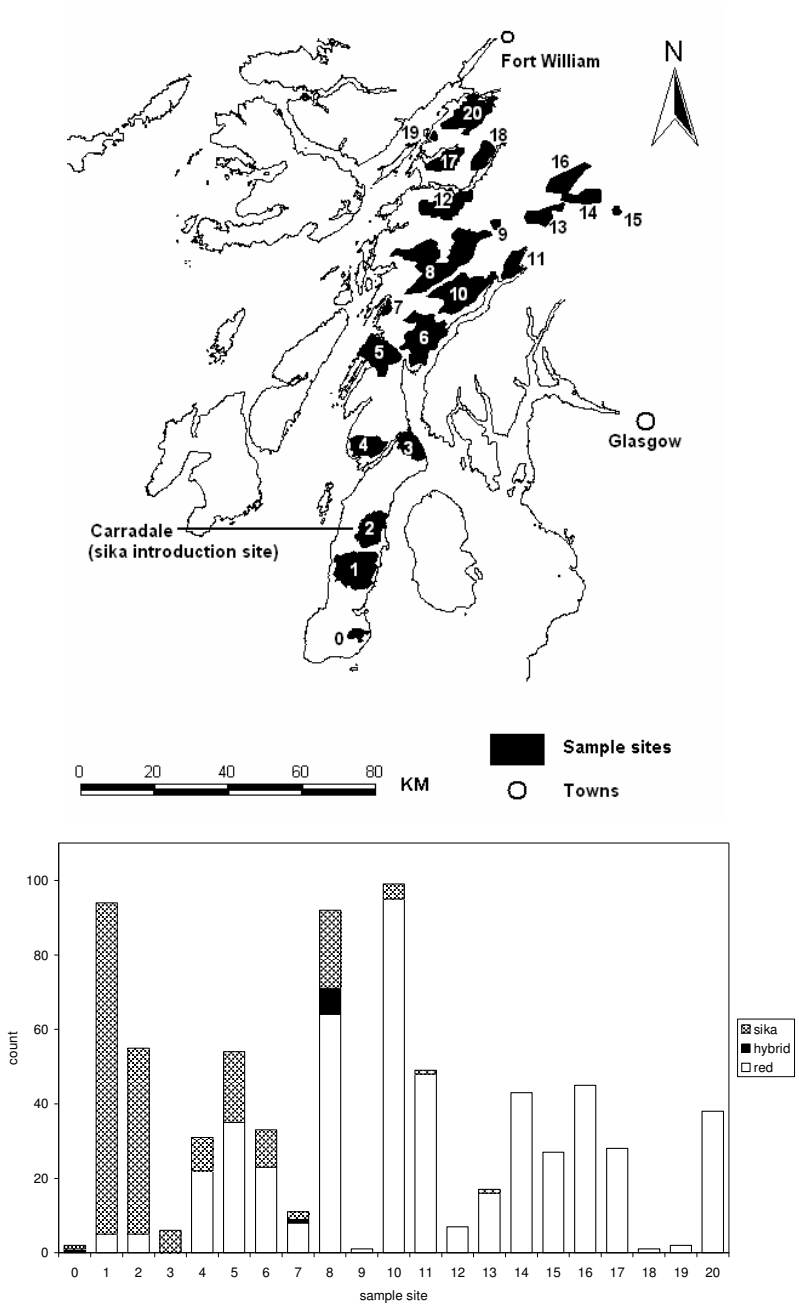


Figure 1: a) Map of the study area with sampling sites on the Kintyre Peninsula shown and numbered (0-20) approximately from south to north. b) The number of samples collected at each site shaded according to the phenotype identified by the ranger that collected the sample. Sika deer are more numerous in the south of the study area (near the introduction site, Carradale) and red deer are more numerous in the north. The names of the sample sites (with sample sizes) are: 0 (South Kintyre, n=2), 1 (Lussa, n=94), 2 (Carradale, n=55), 3 (South Tarbert, n=6), 4 (Achaglachach, n=31), 5 (Knapdale, n=54), 6 (Kilmichael, n=33), 7(Ormaig, n=11), 8 (West Loch Awe, n=92), 9 (Collaig, n=1), 10 (Eredine/Birdfield, n=99), 11 (Shira, n=49), 12 (Oban, n=7), 13 (Succoth, n=17), 14 (Glen Lochy, n=43), 15 (Ben More, n=27), 16 (Glen Orchy, n=45), 17 (Barcaldine, n=28), 18 (Barrs, n=1), 19 (Appin, n=2), 20 (South Ballachulish, n=38). Sites 0, 2, 4, 5, 6 & 10 were sampled by Goodman *et al.* (1999), they have the same names but are numbered differently in that paper.

2.2.2 Study system

The Kintyre peninsula has been the focus of previous population genetic studies of red-sika hybridisation. 246 (red and sika) deer from the Kintyre and Cowal peninsula (to the east of Kintyre) were initially screened at 4 nuclear loci (2 allozymes and 2 microsatellites) and one mtDNA marker (Abernethy 1994a; Abernethy 1994b). Later, the same sample set was screened for a total of 11 microsatellite loci (Goodman *et al.* 1999). The two allozyme markers (superoxide dismutase and 6-phosphogluconate dehydrogenase) were dropped because it emerged that previously assumed fixed differences between the two species were incorrect and gave inflated estimates of hybridisation.

Goodman *et al.* (1999) described the situation as a bimodal hybrid zone (Harrison & Bogdanowicz 1997) with deer falling into two distinct classes (red-like and sika-like). Their analysis methods treated the taxa as separate populations, in which occasional hybridisation is followed by introgression through backcrossing into either parental species. The 11 loci used by Goodman *et al.* (1999) had very different allele frequencies in the two species and their analysis was based on assigning alleles *a priori* to red or sika based on their frequency in each of these sample subsets. The analysis used the pattern of linkage disequilibrium at the screened loci to distinguish recent hybridisation and introgression from the possibility that the parental populations had alleles in common prior to hybridisation (ancestral polymorphism). From this, the rate of hybridisation was estimated as 1 in 500 and 1 in 1000 matings per generation into sika and red respectively. Both red and sika populations contained a large number of individuals with a small number of apparently introgressed alleles (the most hybridised animals had five apparently introgressed alleles out of a possible 22). Ancestral polymorphisms probably also accounted for some of the apparent introgression into the red population. The highest levels of introgression (40% of animals) occurred at Knapdale (site 5, Figure 1), 50 km north of the introduction site at Carradale.

2.2.3 Aims

In this study we revisited the Kintyre peninsula again 15 years later, this time analysing a sample of 735 red and sika deer collected in a year spanning 2006 and 2007 at a new panel of 22 microsatellite loci and one mtDNA marker. We adopted a different approach to analysis from Goodman *et al.* (1999), using a Bayesian clustering method implemented in the software STRUCTURE (Pritchard *et al.* 2000, Falush *et al.* 2007) to infer the extent of hybridisation. The aims of this investigation were:

1. To assess the current extent and distribution of gene flow between red deer and sika deer on the Kintyre peninsula.
2. To assess whether the direction of hybridisation (i.e. sika stag with red deer hind or *vice versa*) was having an effect on gene flow.
3. To make an informal comparison with the previous study by Goodman *et al.* (1999) with a view to extending this to a thorough comparison of the dynamics of hybridisation between red and sika deer on this peninsula in the future.

2.3 Materials and methods

2.3.1 Study area and sampling

The study area stretched from the southern tip of the Kintyre Peninsula to Glencoe in the north and Crianlarich in the east. Samples were collected from 21 commercial forestry sites across this region by Forestry Commission Scotland rangers as part of their normal culling operations between 06/2006 and 06/2007 (Figure 1). Rangers were instructed to sample both species and no limit was placed on the number of a given sex or species gathered from each site. Samples consisted of an ear tip that was preserved directly in 100% ethanol. Rangers

were asked to classify each sample as either 'red', 'sika' or 'hybrid' based on their assessment of the phenotype of the shot animal.

2.3.2 Laboratory procedures

Genomic DNA was extracted with the DNeasy Blood and Tissue Kit (QIAGEN) according to the manufacturer's instructions.

A 22- microsatellite locus protocol was developed which enabled efficient polymerase chain reaction (PCR) and sequencer load multiplexing (Table 2). The primers for these loci were originally derived from cattle (*Bos taurus*) apart from oarFCB193 which was developed in sheep (*Ovis aries*) and were screened for cross-species polymorphisms as detailed by Slate *et al.* (1998). The 22 loci that were selected were also chosen for having no shared alleles between test panels of 44 red and 44 sika of diverse geographic origins (Goodman *et al.* 1999*).* Of the 22 loci, 18 have been mapped to red deer linkage groups (Slate *et al.* 2002). Two pairs map to the same linkage group (INRA6 & TGLA127, BM6438 & RM95), but in both cases they are located at the opposite ends of the group and are expected to recombine freely. This yields markers on 16 of the 34 red deer linkage groups, with 4 of unknown location. Six of the 22 loci (BOVIRBP, FCB193, RM95, RME25, BM6438 and MM12) were used in the previous study (Goodman *et al.* 1999).

During the development of multiplexed PCR protocols, visualisation of PCR products was carried out on 4% agarose gels stained with ethidium bromide. Later, multiplexed PCR products were run on an ABI3730 capillary sequencer (Applied Biosystems) together with an internal size standard GENESCAN 500 LIZ (Applied Biosystems). Analysis of the fragments was carried out using the software GENEMAPPER version 4.0 (Applied Biosystems). Full details of all loci and protocols can be found in Table 2.

A mitochondrial DNA haplotype was assigned using a 39-bp tandem repeat in the mitochondrial control region (Cook 1993). Red deer have a single repeat, whereas sika have multiple repeats. Length variation in repeats was assayed on 4% agarose gel stained with ethidum bromide (Goodman *et al.* 1999); further details in Table 2.

Table 2: Marker details and multiplex PCR panel conditions

Loading panel	PCR plex	Locus name	Annealing temperature (°C)	Primer concentration (µM)	Label	Size range (bp)	Deer group	linkage	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***	52	0.5	-	430(sika)	-		TTAAACTATTCCCTGACGCTT
		H16498§§§				350(red)			CCTGAAGTAGGAACCAGATG

Conditions for 10µl PCR reactions: ~2ng DNA, 2.5mM $MgCl_2$, 0.2mM dNTPs, 1.5x NH_4 buffer; 0.06-0.5µM primer, 0.25 units of BioTaq polymerase (Bioline). Thermocycling: 95°C for 4 min, followed by 30 cycles of denaturation at 94 °C for 30 sec, annealing (variable temperatures, see table) for 60 sec and extension at 72 °C for 90 sec (ramping to extension temperature at 0.3°C/sec). Final extension at 60 °C for 30 min. References: *Georges & Massey (1992); §Barendse *et al.* (1994); ¶Viaman *et al.* (1992);¥ Mezzalani *et al.* (1995);† Bishop *et al.* (1994);‡ Moore *et al.* (1992);** Moore *et al.* 1994; §§Viaman *et al.* (1994); ¶¶Mommens *et al.* (1994); ¥¥ Buchanan and Crawford (1993);†† Barendse *et al.* (1997);‡‡ Grosse *et al.* (1995); ***Cook (1993); §§§ Kocher *et al.* (1989)

2.3.3 Genetic diversity analysis and statistics

Initial locus-specific diversity measures such as allele frequencies, observed and expected heterozygosities (H_O and H_E) were calculated using CERVUS version 3.0 (Kalinowski *et al.* 2007). F-statistics were calculated with FSTAT 2.9.3.2 (Goudet 1995). Graphics and statistics were produced in R version 2.6.2.

2.3.4 Population admixture analysis

Analysis of population and individual admixture, using the microsatellite data only, was carried out with a Bayesian clustering algorithm implemented in STRUCTURE 2.2 (Falush *et al.* 2007). This model assumes that there are K populations which are each characterised by a set of allele frequencies at each locus. Within populations, loci are assumed to be in Hardy-Weinberg equilibrium and linkage equilibrium. The model introduces population structure by assigning individuals probabilistically to a population, or more than one population if individuals are admixed, in such a way as to produce K clusters that are, as far as possible in Hardy-Weinberg and linkage equilibrium (Pritchard *et al.* 2000). This method generates estimates of the admixture proportion (known as Q) for each individual in the sample set. A STRUCTURE analysis was chosen for this study because assumptions about the population origins of alleles at each locus are not needed. Despite the known allelic differentiation between the two populations (see above), it was deemed to be more robust not to make prior assumptions about the origin of specific alleles, as the data set contained a number of (mostly rare) alleles whose population origin it had not been possible to assign convincingly.

The most likely number of populations in the data set (K) was estimated by conducting 10 independent replicates of K=1-5. The model was run using a burn-in of $5x10^4$ and a run of 10^6 Markov chain Monte Carlo (MCMC) steps, under 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 (λ=1) (see program manual for further details). The presence of

null alleles was estimated simultaneously, using the "RECESSIVEALLELES=1" option, new to STRUCTURE 2.2 (see Supplementary Material for further information).

2.4 Results

2.4.1 Samples

In total 735 samples were collected (312 females, 423 males). The number of animals collected at each site (Figure 1) was highly variable (n=1-99). This variability reflects the varying area of the sites and variations in culling and sample return effort across sites. Since rangers shoot deer that they see regardless of species, we make the assumption that red and sika are sampled in approximate proportion to their abundance at a site. Sika dominate in the south of the study area and red deer in the north. Both species were collected at 10 of the 21 sites and individuals nominated as phenotypic hybrids were present at three sites (0, 1 & 8). In total there were 513 animals designated by rangers as red, 213 as sika and 9 as hybrids.

2.4.2 Genotypes

The nuclear data set consisted of a 99.82% complete matrix of data and the MtDNA data set was 99.72% complete. 6.5% of the nuclear data set was retyped, and genotyping error rate (Bonin *et al.* 2004) was estimated as 0.3-0.8% (see Supplementary Material for further information). MtDNA error rate was estimated at 1.5% from a retyped set of 131 samples (2 errors).

2.4.3 Genetic diversity

Genetic diversity indices are summarised in Table 3. Across the entire dataset, the number of alleles per locus ranged from 2 (INRA5) to 19 (FSHB) with a median of 7.0 per locus. Median allelic diversity in sika (3.0) was significantly lower than in red (6.5) (Wilcoxon's signed rank test=168, $p<0.001$). Estimated allele frequencies in the parental populations calculated by STRUCTURE can be found in the Supplementary Material (Table S1). Typically the distributions of allele frequencies at loci in sika were highly skewed with one allele having a frequency of >0.9 and the remaining alleles being at low frequency. Allele frequencies in red were more evenly distributed across alleles (see Supplementary material: Table S1). Fifteen alleles at twelve loci have estimated allele frequencies of >1% in both red and sika populations. These loci (alleles) are: AGLA293 (147 & 144), BM4006 (93), BM757 (160), BOVIRBP (153), FCB193 (103), INRA5 (126), RM12 (139), RME25 (168 & 170), TGLA40 (97), TGLA126 (105), TGLA127 (174 & 178), TGLA337 (138).

Mean observed heterozygosity (H_O) and expected heterozygosity (H_E) across the whole sample population were 0.41 and 0.69 respectively. Mean observed and expected heterozygosity in phenotypic red deer as assigned by the rangers ($H_{O\,r}$ and $H_{E\,r}$) were 0.52 and 0.56 respectively and mean observed and expected heterozygosity in phenotypic sika ($H_{O\,s}$ and $H_{E\,s}$) were 0.14 and 0.15 respectively (note that both red and sika phenotypic classes contain genetic hybrids, see Figure 3). Weir and Cockerham's (1984) estimate of F_{ST} between phenotypic sika and phenotypic red deer for these *non-randomly selected markers* was 0.58 (see Discussion).

2.4.4 Population admixture analysis

The results of the replicate STRUCTURE 2.2 simulations at each value of K were highly consistent. The most parsimonious division of the dataset was into two clusters (K=2), with an average Ln Pr (X | K) (natural logarithm of the probability of the data X, given K) of -

30852. (s.d.=12.78). Average Ln Pr (X | K) is slightly higher at higher values of K (mean (s.d) for K=3, 4 & 5 respectively: -30834. (23.39), -30755. (86.82), -30831. (162.19)).

In the red population, five loci (AGLA293, MM12, BOVIRPB, TGLA337& UWCA47) had estimated null allele frequencies greater than 0.04 (see Table 3). In sika, the estimated null allele frequency was <0.01 at all loci (see Supplementary Material for further information on null alleles).

Table 3: Diversity indices for each of the 22 loci studied (n=735 samples)

Locus	N	k_r	$H_{O\,r}$	$H_{E\,r}$	$Null_r$	k_s	$H_{O\,s}$	$H_{E\,s}$	$Null_s$
AGLA293	725	3	0.298	0.45	**0.163**	3	0.07	0.082	0.008
BM4006	733	4	0.295	0.308	0.007	2	0.141	0.144	0.001
BM6438	733	4	0.573	0.583	0.005	3	0.413	0.406	0.002
BM757	733	13	0.715	0.736	0.01	4	0.062	0.07	0.001
BOVIRBP	734	8	0.631	0.723	**0.06**	2	0.061	0.078	0.004
FCB193	734	14	0.827	0.832	0.004	2	0.085	0.087	0.001
FSHB	734	19	0.852	0.865	0.008	3	0.075	0.078	0.001
IDVGA29	735	3	0.425	0.442	0.011	3	0.033	0.055	0.002
IDVGA55	734	11	0.736	0.769	0.004	4	0.174	0.187	0.001
INRA5	735	2	0.018	0.021	0.014	2	0.08	0.094	0.003
INRA6	735	5	0.47	0.484	0.005	4	0.066	0.095	0.001
INRA131	735	6	0.575	0.617	0.005	2	0.042	0.051	0.001
MM12	734	4	0.353	0.42	**0.043**	3	0.061	0.069	0.001
RM12	734	11	0.777	0.815	0.013	3	0.08	0.078	0.001
RM188	734	12	0.709	0.726	0.01	6	0.577	0.601	0.001
RM95	735	11	0.774	0.806	0.009	2	0.033	0.042	0.001
RME25	735	7	0.361	0.377	0.017	3	0.085	0.095	0.004
TGLA40	734	5	0.582	0.618	0.015	5	0.085	0.103	0.003
TGLA126	735	2	0.018	0.017	0.014	3	0.474	0.524	0.001
TGLA127	735	10	0.698	0.711	**0.078**	3	0.221	0.249	0.001
TGLA337	730	8	0.678	0.8	0.015	3	0.142	0.139	0.002
UWCA47	735	3	0.117	0.162	**0.054**	2	0.019	0.046	0.004

N, number of samples typed at each locus; k_r and k_s, number of alleles considered to be in the red and sika populations according to the STRUCTURE analysis, these values were obtained from the allele frequency data (see Supplementary Material: Table S1); $H_{O\,r}$ and $H_{O\,s}$, observed heterozygosity in phenotypic red (n= 513) and sika (n=213) ; $H_{E\,r}$ and $H_{E\,s}$, expected heterozygosity in phenotypic red and sika; $Null_r$, estimated null allele frequency in red; $Null_s$, estimated null allele frequency in sika. Both these estimates were obtained with STRUCTURE 2.2. Value in bold indicate null frequency >0.04.

STRUCTURE was used to obtain an estimate of an individual's proportion of ancestry from each of the clusters. From here on this estimate will always be referred to in terms of membership to red (1=red, 0=sika) and will be called Q. In total 13 individuals had Q values of $0.25 \leq Q \leq 0.75$, 14 individuals had $0.05 \leq Q < 0.25$ (sika-like hybrids) and 24 individuals had $0.75 < Q \leq 0.95$ (red like hybrids). The reason for choosing a 0.05 bound for distinguishing a hybridised animal from pure is discussed below. The distribution of Q and 90% posterior probability intervals can be found in Figure 2.

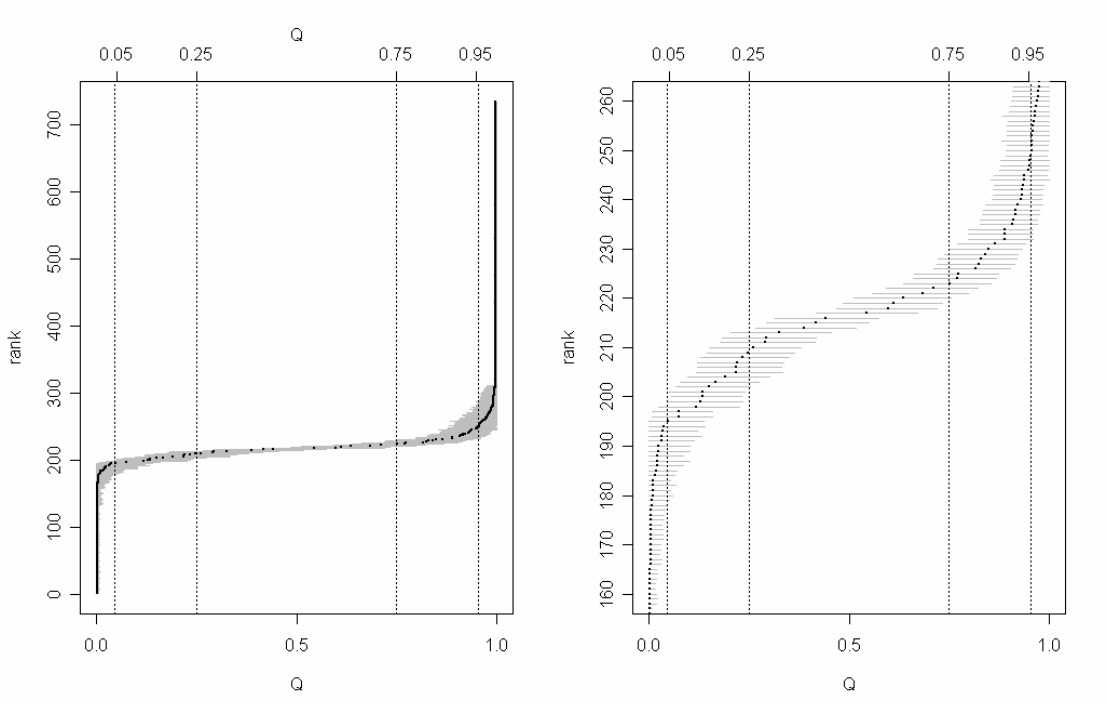


Figure 2: The distribution of estimated proportion of ancestry (Q) against rank order of Q, for the 735 individuals in the data set. Horizontal lines are 90%, equal-tail posterior probability intervals for each individual. Dashed vertical lines bound arbitrary cut-offs for red-like hybrids, hybrids and sika-like hybrids. For ease of viewing, the right hand figure depicts a magnified section of the left hand one for the range rank=160-260. Figure after Beaumont *et al.* (2001).

The phenotypic designations provided by the rangers corresponded well to Q in the case of red and sika (Figure 3). The mean (and variance) of estimated Q were 0.99 (0.028) and 0.02 (0.025) for phenotypic red and sika respectively. The individuals designated as hybrids had a much poorer relationship to estimated Q (mean=0.65, variance = 0.12) and there was considerable overlap in the range of estimated Q between the phenotypic hybrids and the two other categories.

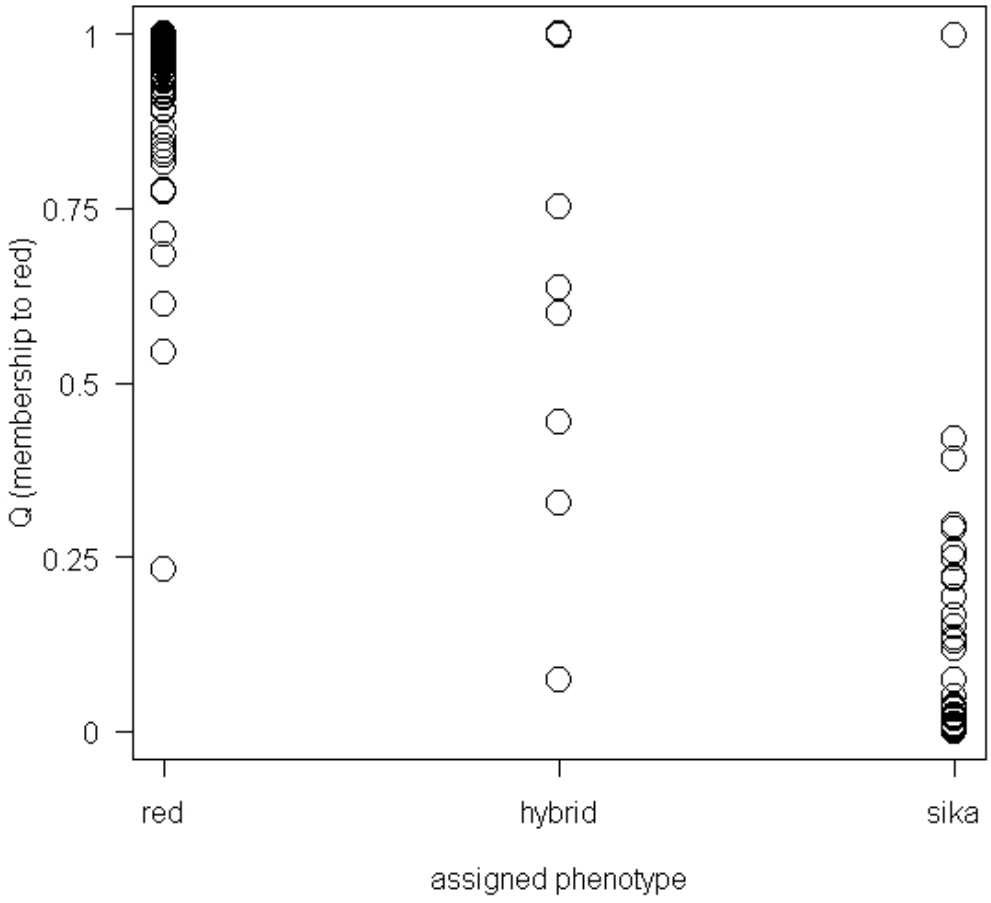


Figure 3: The estimated proportion of ancestry (Q) of each of the 735 individuals plotted against ranger-assigned phenotype. A Q-value of zero indicates that the individual is pure sika; a Q-value of one indicates the individual is pure red; intermediate values indicate a degree of hybridism. In total 513 animals were designated as red, 213 as sika and 9 as hybrids by the rangers that shot them.

The spatial distribution of hybrids was highly clumped (Kruskal-Wallis χ^2=177.77, df=10, p<0.001) (see also Figure 4). To carry out this test, we transformed the original measure Q into a collapsed hybrid index. The intention of the new measure, $Q_{collapsed}$, was to capture the extent of hybridisation without a direction (i.e. how red or sika-like the animal is). $Q_{collapsed}$ was calculated as follows: if Q>0.5 then $Q_{collapsed}$=1 - Q; if Q<0.5 then $Q_{collapsed}$=Q. Sites where fewer than 30 individuals were sampled were excluded from the test (i.e. sites 0, 3 ,7 ,9 ,12 ,15 ,17 ,18 & 19). Site 8 (West Loch Awe) had 78 % of all hybrids ($0.05 \leq Q < 0.95$) in the study area and 12 out of the 13 intermediate hybrids ($0.25 \leq Q \leq 0.75$). The only other site with an intermediate hybrid was South Kintyre (site 0), although neighbouring Lussa (site 1) had one sika-like animal that was borderline (Q=0.248). Red-like hybrids ($0.75 < Q \leq 0.95$) were found at sites 5, 7, 12, 16 & 17 (Figure 4). It is particularly striking that sites 8 and 10 have an extremely different pattern of hybridisation despite being approximately equally far from the introduction point of sika (Figure 5).

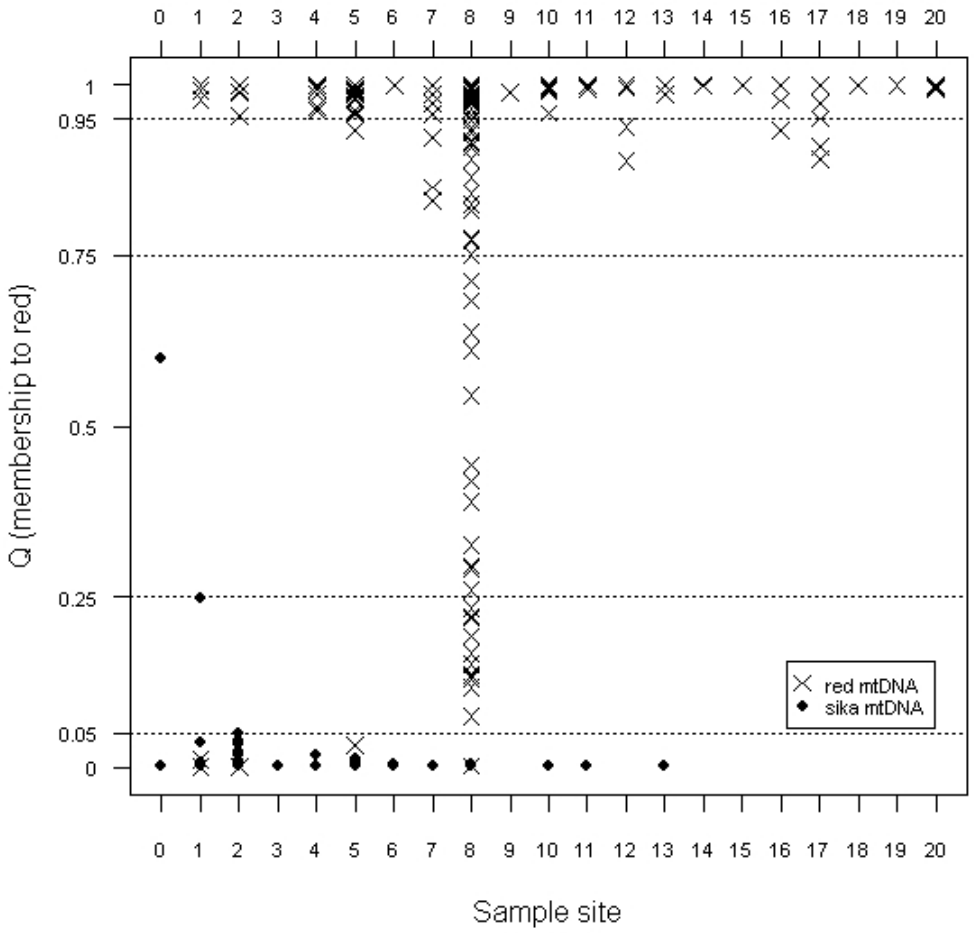


Figure 4: Distribution of Q by sample site. Individuals are also coded according to mtDNA haplotype. Note that in upper and lower regions of the graph many points may overlap; for a guide to sample sizes at each site see Figure 1b. Horizontal dashed lines represent cut off points for (from bottom to top) 'pure' sika, sika-like hybrids, intermediate hybrids, red-like hybrids and 'pure' red deer.

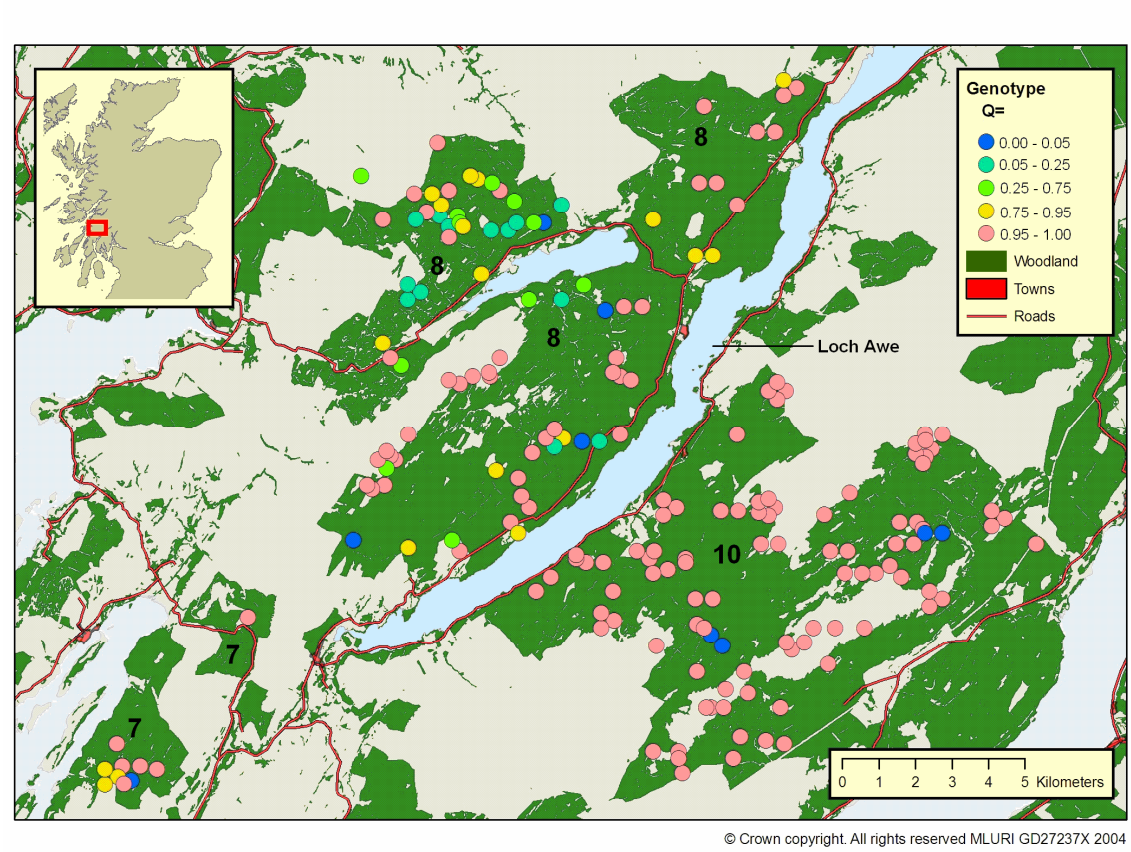


Figure 5: Animals from Ormaig (site 7), West Loch Awe (site 8) and Eredine/Birdfield (site 10) plotted by the grid reference where they were shot. Q values have been binned into various categories by which the animals are colour coded (see legend). Nine animals in the data set are not plotted because they did not have grid references. Hybridisation is extensive at West Loch Awe, but Eredine/Birdfield appears to be free from hybridisation despite the presence of sika at this site.

Intermediate hybrid genotypes, mostly found at site 8, West Loch Awe, generally had a mixture of loci that were homozygous for red-typical alleles, loci which were homozygous for sika-typical alleles, and loci which were heterozygous for red- and sika-typical alleles (being homozygous for say, red-typical alleles does not imply that the individual was homozygous at that locus for allele size as there were a number of red typical alleles). Sika-like and red-like hybrids found at other sites typically had genotypes in which loci are mostly homozygous for sika- *or* red- type alleles, with a few heterozygous loci, whereas many sika-

like and red-like hybrids at site 8 had genotypes with loci that were homozygous for both species-type alleles.

An animal was defined as carrying introgressed mtDNA if it carried a red mtDNA haplotype and Q was less than 0.5, or if it carried a sika haplotype and Q was greater than 0.5. Only one individual (from site 0) of 519 deer with Q>0.5 carried the sika mtDNA haplotype, whereas 36 individuals of 216 with Q<0.5 carried the red haplotype (Fisher's exact test=0.0097, p<0.001). The majority (22) of these individuals were found at West Loch Awe, but the pattern still remains significant if West Loch Awe is excluded (Fisher's exact test=0.0028, p<0.001). Excluding data from sites where no hybridisation was detected does not alter the significance of the results. There were 16 animals that were genotypically 'pure' sika at nuclear microsatellite loci (Q<0.05) but carried the red deer mtDNA haplotype. Twelve were found at Lussa (site 1) and the rest at sites 2(1), 5(1) and 8(2) (see Figure 4). There were no animals that were genotypically 'pure' red (Q>0.95) and carried the sika mtDNA haplotype.

2.5 Discussion

2.5.1 Summary

Our analysis of 22 unlinked loci and one mtDNA marker in a sample of 735 deer reveals that, whilst most of the sample set can be clearly be divided into red or sika, there is an appreciable proportion of hybrids (6.9%). These hybrids are not generally identified by the rangers shooting them. The distribution of hybrids is highly clumped with some sites apparently having no introgressed animals (defined as $0.05 \leq Q \leq 0.95$) despite the two species being present, whilst West Loch Awe had 43% introgressed animals. The pattern of mtDNA

introgression is also spatially clumped with the majority of introgression occurring at Lussa and West Loch Awe. The pattern of mtDNA introgression is one-sided, with most introgression of mtDNA occurring from red into sika.

2.5.2 Dealing with ancestral polymorphism

The 22 microsatellite loci used in this study show high levels of differentiation between red and sika (Table 3; Supplementary Material: Table S1), which is hardly surprising given the methods used to select them (see Materials and Methods). Given this differentiation, they should provide reliable estimates of hybridism using the software STRUCTURE (Pritchard *et al.* 2000; Vähä & Primmer 2006). Before we discuss the structure of the hybridising populations, we deal first with the issue of ancestral polymorphism (the possible presence of alleles shared between the populations prior to contact).

STRUCTURE's Q (the proportion of ancestry of each individual) is generated by taking allele frequencies and linkage disequilibrium between alleles into account. Here we define any animal where Q is greater than 0.05 and less than 0.95 as a hybrid (a 0.05 cut-off level). This is a less stringent definition of hybridism than in a number of other studies of hybridisation using STRUCTURE (e.g. Beaumont *et al.* 2001; Lecis *et al.* 2006; Tung *et al.* 2008). However, an assessment of the performance of STRUCTURE on simulated data, (Vähä & Primmer 2006) found that when using a cut-off of Q=0.05, STRUCTURE was 90% accurate at identifying hybrids in studies with 12 or more loci where F_{st} between populations was 0.21 (see Figure 2, Vähä and Primmer 2006). In our study, population differentiation (at the *non-randomly* chosen markers) is considerably higher (F_{ST}=0.58) and we use 22 loci. Ideally, simulated hybrid frequencies would be used to assess the statistical power for correctly identifying the origin of individuals as parental or hybrid (Nielsen *et al.* 2001; Randi 2008). Such an analysis relies on having parental allele frequencies available for each population. Although this might be possible for the red population by taking a subset of deer sampled a long way

from the introduction site, obtaining parental allele frequencies for sika is much harder as there is no region on the Kintyre Peninsula where we can be sure sika are not hybridised.

Despite considerable differentiation in allele frequency between the two parental populations, a number of alleles appear to be shared (Supplementary Material: Table S1). Shared alleles could either be due to hybridisation of sika with red deer in captivity prior to their release (Powerscourt 1884), due to alleles that have persisted in the two species since their split, or due to convergent mutation of microsatellite allele lengths (i.e. alleles are identical by state but not by descent). Examination of individual genotypes reveals that many of the individuals that indicate some introgression ($0.95<Q\leq0.999$ and $0.001\leq Q<0.05$) carry one or more of the 15 alleles that seem to be shared between the two populations (see Results). The distribution of allele sizes at loci at which shared alleles are present is not entirely consistent with the shared alleles having persisted since the two populations split or convergent evolution of microsatellite lengths: At BM4006, BOVIRP, INRA5, RM12 & RME25 STRUCTURE estimates one allele in sika to be at very high frequency <0.9; other alleles, at low frequency, are shared with the red population and are not close in size to the dominant sika allele (Supplementary Material: Table S1). This pattern is possibly more consistent with the idea that sika were hybridised with red prior to their release onto the peninsula, or that hybridisation occurred soon after their release so that alleles introgressed into sika from red have spread with them as they expanded. Both these scenarios are expected to result in the alleles having reduced linkage disequilibrium across the two populations. This would lead STRUCTURE to estimate their presence in both parental populations. The scenario of hybridisation early on in the spread of sika was suggested by Goodman *et al*. (1999) as they noted that the level of apparent introgression into sika was fairly uniform across the Kintyre Peninsula. It is highly probable then, that some individuals currently not classified as hybrids have distant hybrid ancestry. We could address this issue

by lowering the cut-off level for Q, but then we would increase the risk of identifying pure individuals as hybrids.

2.5.3 Patterns of introgression

The previous finding that hybridisation events between red and sika are rare (Goodman *et al.* 1999) is supported by our failure to find any F1 hybrids in a sample of 735 animals (animals with Q values close to 0.5 do not have genotypes with all loci heterozygous for red- and sika-type alleles, see Results). Despite this, introgression has a noticeable effect on the genetic structure of the population. The pattern of introgression varies considerably across sites (Figure 4). When hybridisation is rare, F1s that arise must backcross into the parental populations. Each generation of backcrossing reduces the contribution of genes from the other parent species by 50%. Over time, and in the absence of selection against hybrids, this scenario has the potential to generate a large number of individuals with a small proportion of introgressed alleles. This is an explanation for sites where individuals which only have a small proportion of their genome introgressed are found. For example at sites 2, 5, 12, 16 and 17 (Figure 4) a hybridisation event that occurred a number of generations in the past may have generated the current pattern of introgression. Discordant mtDNA haplotypes in individuals with genomes that appear otherwise free from introgression can indicate older hybridisation events (with a panel of 22 markers, after 6 generations of backcrossing one would no longer expect to find an introgressed allele). This is also an explanation for the pattern of mtDNA introgression found at sites 1, 2 and 5 (Figure 4). Discordant mtDNA in otherwise non-introgressed individuals could also indicate that this marker is not truly diagnostic and that ancestral polymorphism is causing the pattern; however the spatial distribution of the pattern of mtDNA discordance makes this unlikely (see Figure 4). Introgression despite apparent absence of F1s is also documented in the Louisiana iris (*Iris fulva*, *I. brevicaulis* and *I. hexagona*) hybrid zones (Cruzan & Arnold, 1993).

The pattern of introgression found at West Loch Awe (site 8, Figure 4) is quite different from other sites. Hybridisation is extensive (43% of animals) and a broad range of Q values is found, despite an absence of F1 hybrids. This pattern indicates a departure from the scenario of individual hybridisation events followed by backcrossing. The range of Q values suggests that crosses have occurred between hybrids and examination of the hybrids' genotypes reveals that many are homozygous for red and sika alleles at different loci, indicating that they could not have been generated by backcrossing alone.

The pattern of mitochondrial introgression at West Loch Awe is striking, with 97.8% of animals at the site carrying the red mtDNA haplotype, despite many of them being strongly genotypically sika at nuclear loci. The explanation for this pattern is not clear: During range expansion, sika stags are known to move in advance of hinds (Livingstone 2001; Ratcliffe 1987) and hybridisation events between sika stags and red deer hinds would generate hybrids with red mtDNA. To generate sika-like individuals with red deer mtDNA requires successive backcrossing with sika males and/or generation of F2 (and other F*n*) hybrids. This scenario would either require the absence of sika hinds, sika female choice for only 'pure' sika males, selection against sika mtDNA in hybrids or a combination of these. The sample of deer at West Loch Awe does not contain any sika hinds that carry sika mtDNA (two stags carry sika mtDNA, Figure 4), however two such females are present at neighbouring site 10 and sika hinds were found at site 10 as early as 1976 (Ratcliffe 1987). Negative cytonuclear interactions between sika mtDNA and red deer nuclear genes may explain the pattern of introgression, but in at least some advanced backcrossed (or little hybridised) individuals these negative interactions should have broken down as the nuclear genes responsible should have been eliminated by a number of generations of recombination (Arnold 1993). This would, in theory, enable incoming sika hinds to produce viable offspring with some sika-like hybrids. This study does include one red-like hybridised individual (Q=0.59) with sika mtDNA at site 1 and Goodman *et al.* (1999) found two red individuals with sika mtDNA (at

sites 5 & 10, numbered according to this study), so it is apparent that sika mtDNA introgression cannot be lethal in all cases (although these animals might still be infertile).

Female choice by sika hinds (especially against sika like hybrid males) appears to be an unlikely explanation for this pattern of introgression as deer are generally not thought to exert female choice (Clutton-Brock *et al.* 1982). Anyhow even if 'choosy' (sika mtDNA carrying) females were present, we would expect to have sampled them. It seems likely that initial hybridisation events occurred between resident red hinds and sika stags migrating into the area from the south; further migratory sika males have successfully crossed with hybrids. Whether cytonuclear interactions play a role in keeping sika mtDNA out cannot be answered in this study. It may be that, as suggested by our sample, assuming it is indeed representative of the population, no established population of pure sika hinds exists at West Loch Awe, so that differential rates of migration between the sexes are responsible for this pattern. The spatial distribution of males and females has been suggested as the reason for unidirectional mtDNA introgression in a tree frog (*Hyla*) hybrid zone (Lamb & Avise 1986). The previously published study of *Cervus* hybridisation on the Kintyre peninsula (Goodman *et al.* 1999) did not sample West Loch Awe, but we can assume this hybrid swarm is young (<40 years, judging by dates at which sika were first sited, see Introduction). Perhaps sika mtDNA will eventually penetrate this population as sika hinds colonise the area.

2.5.4 Spatial structure of hybridisation and introgression

Previously the scenario on Kintyre has been modelled as a cline of increasing introgression into sika with distance from the introduction site and increasing introgression into red with distance towards the introduction site; although the fit was not convincing (see Figure 4, Goodman *et al.* 1999). Introgressed individuals were found to be most common (37.9%) where the two taxa have overlapped for longest (Knapdale, site 5 in this study). A thorough direct comparison of these two sampling periods is beyond the scope of this paper and is

made difficult by lack of overlap in number and types of markers used, differing sample sizes and sites, and different analytical approaches. However, it is possible to draw the conclusion that the current structure of the hybrid zone is not clinal, in fact it should not be considered a hybrid zone in the classical sense (Barton & Hewitt 1981; Szymura & Barton 1986). Distance from introduction site is not a very good explanatory factor for levels of hybridisation. This is seen very strikingly in the comparisons of West Loch Awe with its neighbouring site Eredine (Figure 5), which despite being approximately equally distant from the introduction site shows a very different pattern of introgression.

Instead the situation (refer to Figure 4) is probably characterised by sporadic and occasional hybridisation events, mostly between red deer hinds and sika stags. In some areas this has resulted in low level introgression and mtDNA discordance generated by backcrossing, perhaps from as little as one hybridisation event (e.g. site 2, Figure 4). At some sites hybridisation has possibly occurred a long time ago and left little trace, or a faint signature of introgression is present from introgressed individuals migrating into the area (e.g. site 10, Figure 4). At West Loch Awe some critical threshold seems to have been reached at which hybrids have reached a sufficient density to generate complex crosses between themselves. Gene flow has possibly spread introgressing alleles to surrounding sites (sites 7 & 12, Figure 4). The south of Kintyre (sites 0 &1, Figure 4) has a signature of more recent hybridisation on top of older introgression (the sample sizes at site 0 is very small). Further north (sites 13-20, Figure 4), a low level of introgression present at some sites (e.g. sites 16 & 17, Figure 4) may indicate past hybridisation at this site or migration of introgressed animals from sites further south.

Patchy hybrids zones in which stochastic forces appear to play the main role in shaping the structure were termed “mottled hybrid zones” by Hauffe & Searle (1993) as opposed to “mosaic hybrid zones”, in which patchy structure is determined by selection induced by

habitat variability, e.g. in *Gryllus* (Harrison & Rand 1989). Haufe and Searle (1993) argue that patterns of extinction and recolonisation following flooding in 1807 in the Upper Valtellina, Italy have led to a patchy distribution of karyotypic variants of house mice (*Mus musculus domesticus*) in a 20km stretch of the valley. It is clear from this study that initiation of hybridisation between red and sika deer in an area and subsequent spread of introgression is highly stochastic, and this scenario could be termed a mottled hybrid zone, although it is arguable whether this type of classification is helpful since classification requires the assumption that the underlying cause of the hybrid zone's structure is established.

In fact, it may not be the stochastic generation of F1s that really characterises the structure of this hybridising population so much as the combination of this with the relatively short time these two species have been in contact (115 years, probably around 38 generations). Many hybrid zones, whether clinal or mosaic, could be seen as evolved 'structures'- gene flow, selection and competition have had thousands of years to structure the populations (e.g. in *Bombina bombina x variagata* (Yanchukov *et al.* 2006), *Gryllus firmus x pennsylvanicus* (Ross & Harrison 2002), *Littorina saxatilis* morphotypes (Wilding *et al.* 2001), *Heliconius himera x erato* (Jiggins *et al.* 1997)). For example in a hybrid zone between Western European house mice *Mus musculus musculus* and Eastern European *M. m. domesticus*, where the two species have been in contact for 2800-6000 years due to range expansion since the last ice age (Boursot *et al.* 1993), mice from within the hybrid zone show a greater tendency to mate assortatively than those from allopatric populations (Smadja & Ganem 2005). The point here is, that in these hybrid zones, interactions present now were not present at the beginning of contact. Red-sika hybridisation on Kintyre is very young and because there have been so few hybridisation events we can expect the effects of stochasticity to dominate over the effects of exogenous selection in terms of its current measurable structure.

The structure of the hybridising population can be described as bimodal, although the extent of bimodality depends on the scale at which it is measured (Schilthuizen 2000), with West Loch Awe representing a step in the transition toward unimodality (Jiggins & Mallet 2000). In general, if possible, it might be more useful to categorise hybrid zones more clearly in terms of the time of divergence of species prior to contact, time since first contact and current frequency of hybridisation events. For example, although the scenario presented here is superficially similar to that of the Louisiana irises (bimodal structure, low rate of hybridisation, patchy distribution of hybrids), there are some crucial differences. Gene flow between the iris species has presumably been occurring for a much longer time (Arnold *et al.* 1990a; Arnold *et al.* 1990b). This is relevant because it means the potential for total genetic exchange between the two species is considerable - even under rare hybridisation, populations can gradually accumulate shared alleles (see above, 'dealing with ancestral polymorphism'). In fact, the Louisiana iris mosaic hybrid zones (*Iris fulva* x *I. brevicaulis*) appear to be maintained by an interplay of habitat structure and differences in flowering phenology and morphology between the two species (Johnston *et al.* 2001; Martin *et al.* 2007; Martin *et al.* 2008). We can expect that the initial genetic distance between populations and the time for the evolution of each population in response to genetic introgression and the ecological impact of the other population to be important factors in determining the structure of any investigated hybrid zone.

2.5.5 Comparison to previous study

We found differences in the levels of introgression across the study area compared with Goodman *et al.* (1999) (see Table 4). Introgression at Knapdale appears to be much lower than the 37.9% suggested by Goodman *et al.* (1999); we find only 1.85 % of individuals have $0.05 \leq Q \leq 0.95$ (Table 4). Relaxing our criteria for hybrids individuals to $0.01 \leq Q \leq 0.99$ only increases the percentage of hybridised individuals at Knapdale to 20.4%. A comparison of other sites (Table 4) reveals that introgression is estimated as being lower at all sites in

this study in comparison to the previous study according to our $0.05 \leq Q \leq 0.95$ definition of a hybrid. Relaxing the definition of a hybrid to $0.01 \leq Q \leq 0.99$, increases the number of hybrids at all the sites, but only at Carradale is the extent of introgression in the two studies similar (~ 20%) (Table 4).

Table 4: A comparison of the extent of introgression at sites in common between this study and Goodman *et al.* (1999)

Sample period	**1991/92 (Goodman *et al.*1999)**		**2006/07 (this study)**			
Site name (site number, this study)	N	(% individuals carrying 1 or more introgressed allele)	N	% individuals $0.01 \leq Q \leq 0.99$	%individuals $0.02 \leq Q \leq 0.98$	% individuals $0.05 \leq Q \leq 0.95$
Carradale (2)	40	20.0	55	21.8	12.7	0.0
Achaglachach (4)	33	36.4	31	12.9	6.5	0.0
Knapdale (5)	29	37.9	54	20.4	14.8	1.9
Kilmichael (6)	9	33.3	33	0.0	0.0	0.0
Eredine/Birdfield (10)	40	10.0	99	1.0	1.0	0.0

Both studies sampled other sites but only those in common are listed in this table, apart from South Kintyre which is excluded because of small sample size (n=2) for both studies. A true comparison of the two studies is not possible as different loci, scoring systems and analysis methods were used in each study. The % data for the Goodman *et al.* (1999) study is calculated from Tables 3 & 5 of that paper.

As mentioned above, it is difficult to assess whether the criteria we are using here for hybridism is too strict, and to what extent ancestral polymorphism is contributing to apparent introgression. A true comparison of this study with that of Goodman *et al.* (1999) is not possible here, as any increase or decrease in introgression that may have occurred in the fifteen year interval between sampling periods is likely be obscured by the different methods of analysis used. Reanalysis of the Goodman *et al.* (1999) data set with the panel of markers

used in this study is required to further investigate this issue and assess the trend in introgression over time.

2.5.6 Phenotype

The rangers involved in collecting the data had no problem in identifying red deer and sika deer (Figure 3), which is unsurprising given the obvious phenotypic differences between the two species (Table 1). More surprising is the fact that hybrids, even ones that are genetically quite intermediate, are so poorly identified (Figure 3). It is possible that one species' phenotype dominates in hybrids, but in this case we would expect to see an asymmetric pattern in the distribution of intermediate animals between the red and sika phenotype classes. In fact, hybridised animals are present in both classes (Figure 3). Captive bred hybrids have had intermediate phenotypes, but these have generally been F1s, F2s and Bx1s (Harrington 1979). It is possible, that rangers do not believe they are going to see hybrids and so categorise them into the closest 'pure' phenotype. A study of the correlation of phenotypic characteristics with genotype is needed to answer these questions about trait introgression and hybrid phenotype. These issues are important from an ecological standpoint: red deer and sika can inflict extensive damage to forestry and natural heritage and require heavy culling in Scotland. Hybridisation might allow selectively advantageous traits such as the reduced birth interval in sika (Chadwick *et al.* 1996) to introgress into red deer, making control even harder.

2.5.7 Management implications

Despite the fact that the area where extensive hybridisation has occurred between red and sika deer is highly localized (<10km^2), control of hybridisation through culling or fencing of deer may be difficult. Deer are hard to manage in the dense forestry plantations that make up most of the sampling area. The two species are in contact across the whole peninsula and further areas of hybridisation may exist already or may arise at any point. Displacement of

hybrids caused by a heavy but incomplete cull might actually trigger hybridisation event in new areas. Currently, red deer are protected from hybridisation on the Scottish islands of Islay, Jura, Rum, Arran and the Outer Hebrides by legislation which prevents the release of deer from the genus *Cervus* onto the islands (Wildlife and Countryside Act 1981 (variation of schedule 9) Order 1999). The range of sika deer covers 40% of Scotland (Livingstone 2001) as a result of around 12 separate introductions (Ratcliffe 1987). Sika are also present at a number of sites across the rest of Britain (Ward 2005). It seems possible then, that many other populations of mainland British red deer could become introgressed in places where the two taxa overlap.

2.5.8 Conclusions

The extent of gene flow between invasive sika and native red deer on the Kintyre Peninsula, Scotland is extremely variable across different locations sampled in this study. Mitochondrial DNA data indicates that hybridisation is taking place between sika stags and red deer hinds. This study supports the work of a previous study (Goodman *et al.* 1999) showing that hybridisation is rare between red and sika deer. The rarity of hybridisation combined with the fact that the two species have only recently come into contact (<120 years ago), means that the total number of hybridisation events that have occurred on the Peninsula is likely to be low. For this reason the pattern of hybridisation and introgression is dominated by the stochastic nature of the initial hybridisation events. Over time we expect additional hybridisation events, migration of deer and gene flow may smooth the pattern of introgression across sites. Currently, it seems possible that the merger of the two species across the whole peninsula (as at West Loch Awe) could occur eventually.

2.6 Acknowledgements

The authors would like to thank the Forestry Commission Scotland rangers of West Argyll and Lorne Forest Districts for collection of samples and also Kevin McKillop, Alasdair Cameron, Donnie Cameron and Colin Lavin for useful discussion. Jim McLeod from the Macaulay Institute is thanked for the production of Figure 6. Will Goodall-Copestake and Silvia Pérez-Espona are thanked for advice on microsatellite multiplexing and sampling of deer. The authors would like to thank Steve Albon, Nick Barton and two anonymous referees for helpful discussion and/or comments which improved the quality of the paper. H.V.S was funded through a NERC studentship with the Macaulay Institute as a CASE partner.

2.7 Supplementary material

2.7.1 Allele frequencies

Table S1 (See Final Appendix, Table A1)

2.7.2 Genotyping errors

Individuals for which genotype information was not complete after the first round of genotyping were retyped for the relevant loci up to two further times if necessary. Retyping, as with first round genotyping, was carried out with the multiplex panels and this generated a large number of duplicate genotypes. The genotyping error rate was calculated from these duplicate genotypes as the ratio between observed numbers of allelic differences and total number of allelic comparisons (Bonin *et al.* 2004).

There were 29 unscored nuclear loci out of a total of 16170 (735 individuals x 22 loci), and 26 individuals had incomplete genotypes. In the process of re-genotyping individuals for loci that had not amplified, 1049 duplicate genotypes were created (due to the fact that all reruns were carried out using the multiplex panels). This represents a retyping of 6.5% of the data set, with an allelic mismatch rate in this data subset of 0.8%. Of the 17 discovered allelic mismatches, 12 were attributable to incorrect scoring of GENEMAPPER peaks (one aspect of human error). Each data point in the whole data set was then rescored at least once more and a number of corrections made to the final matrix. This leaves an estimate of final error rate for the complete data set (taking into account that some, if not all scoring errors have been eliminated) that lies somewhere between 0.3-0.8 % which compares well with other estimates for microsatellite data (Bonin *et al.* 2004; Hoffman & Amos 2005). Other mismatches were attributable to allelic dropout or to other human error. The heterogeneity in error rate across loci (Table S2) is largely due to the fact that some loci were far more prone to scoring error than others. Inconsistent genotypes were either retyped for a third time and the consensus genotype chosen, or were coded as missing data. The retyping effort across loci was very uneven (Table S2). Ideally, a subset of the data should have been reprocessed blind to get a better estimate of error rate (Bonin *et al.* 2004), but the present method of estimating error from duplicates created through the multiplex process was deemed a good compromise given the extra time and cost that reanalysing a suitable proportion (5-10%) of this dataset would incur.

Table S2: Genotyping error rates

Locus	N	%R_{typ}	InitAER
AGLA293	725	5.2	0.0
BM4006	733	3.7	0.0
BM6438	733	6.0	0.0
BM757	733	8.4	0.0
BOVIRBP	734	3.3	0.0
FCB193	734	3.8	0.0
FSHB	734	2.3	0.0
IDVGA29	735	10.1	0.0
IDVGA55	734	6.9	0.0
INRA5	735	8.3	0.0
INRA6	735	5.9	0.0
INRA131	735	4.1	0.0
MM12	734	9.3	0.0
RM12	734	6.1	1.1
RM188	734	3.1	17.4
RM95	735	8.7	0.0
RME25	735	11.3	0.6
TGLA40	734	9.3	2.9
TGLA126	735	5.9	0.0
TGLA127	735	8.6	0.0
TGLA337	730	4.5	0.0
UWCA47	735	8.2	2.5

N, number of samples typed at each locus; *%R_{typ}*, percentage of retyping at each locus; *InitAER*; error rate at the allelic level before rescoring of all peaks, note that the true allelic error rate in the final data set is expected to be lower, especially at RM188 in which all error could be attributed to incorrect scoring.

2.7.3 Null alleles

The presence of null alleles in the data set was estimated simultaneously, using the "RECESSIVEALLELES=1" option, new to STRUCTURE 2.2. This function enables suspected null alleles in the data set to be nominated and estimates null allele frequency at any desired loci. Under this function, nine missing data points at three loci (AGLA293, BOVIRP and TGLA337) were nominated as null homozygotes. They were chosen because PCR had failed twice or more during attempts to retrieve the data point, but were in individuals where other loci had amplified fine. Null allele frequency was estimated at all loci.

The presence of null alleles has the potential to inflate estimates of hybridisation if alleles are introgressing at a locus where null alleles are present. They have the potential to reduce estimates if the null alleles themselves are introgressing. Null allele frequency is known to increase with increasing phylogenetic distance from the species in which the relevant microsatellite was developed (Chapuis & Estoup 2007), and as such they are likely to be an issue for the microsatellite primers used here that were initially developed in cattle and sheep. Null allele frequency was only appreciable (>0.02) at 5 loci in the red deer population (AGLA293, MM12, BOVIRPB, TGLA337 & UWCA47). Null alleles have been previously noted at BOVIRBP (Goodman *et al.* 1999; Pemberton *et al.* 1995), but interestingly Goodman *et al* (1999) did not find null alleles at MM12 in the previous study of this population.

Allelic diversity is very low in sika in comparison to red. The lack of diversity in sika is due to the bottleneck the population has passed through. It arose *at most* from 14 unrelated individuals introduced at Carradale, but in all likelihood these individuals originated from a much smaller related number of individuals (e.g. see Powerscourt 1884; Goodman *et al.* 1999; Goodman *et al.* 2001). Given the screening techniques used in choosing these markers in the first place (see Methods), it is unsurprising that we find low null allele frequency in sika. Since allelic diversity is low in sika, we would expect null alleles, if present at a locus at all, to be at high frequency (provided they originated before the bottleneck). Such loci would have presented cases of non-amplification on the screening test panel and would have been excluded from the outset.

The average Q values across 10 runs of the model actually changed little when the STRUCTURE analysis was rerun ignoring null alleles, although the variance in Q values was slightly lower.

2.8 References

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

Abernethy K (1994b) *The Introduction of Sika Deer (Cervus nippon nippon) to Scotland*, University of Edinburgh. PhD thesis

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

Arnold ML, Bennett BD, Zimmer EA (1990a) Natural hybridization between *Iris fulva* and *Iris hexagona* - Pattern of ribosomal DNA variation. *Evolution* **44**, 1512-1521.

Arnold ML, Hamrick JL, Bennett BD (1990b) Allozyme variation in Louisiana irises - a test for introgression and hybrid speciation. *Heredity* **65**, 297-306.

Barendse W, Armitage SM, Kossarek LM*, et al.* (1994) A genetic-linkage map of the bovine genome. *Nature Genetics* **6**, 227-235.

Barendse W, Vaiman D, Kemp SJ*, et al.* (1997) A medium-density genetic linkage map of the bovine genome. *Mammalian Genome* **8**, 21-28.

Barton NH, Hewitt GM (1981) A chromosomal cline in the grasshopper *Podisma pedestris*. *Evolution* **35**, 1008-1018.

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

Beaumont M, Barratt EM, Gottelli D*, et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology* **10**, 319-336.

Bishop MD, Kappes SM, Keele JW*, et al.* (1994) A genetic-linkage map for cattle. *Genetics* **136**, 619-639.

Bonin A, Bellemain E, Eidesen PB*, et al.* (2004) How to track and assess genotyping errors in population genetics studies. *Molecular Ecology* **13**, 3261-3273.

Boursot P, Auffray JC, Brittondavidian J, Bonhomme F (1993) The evolution of house mice. *Annual Review of Ecology and Systematics* **24**, 119-152.

Buchanan FC, Crawford AM (1993) Ovine microsatellites at the OarFCB11, OarFCB128, OarFCB193, OarFCB266 and OarFCB304 loci. *Animal Genetics* **24**, 145.

Butlin RK, Ritchie MG (1991) Variation in female mate preference across a grasshopper hybrid zone. *Journal of Evolutionary Biology* **4**, 227-240.

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.

Chapuis MP, Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* **24**, 621-631.

Clutton-Brock TH, Coulson T, Milner JM (2004) Red deer stocks in the Highlands of Scotland. *Nature* **429**, 261-262.

Clutton-Brock TH, Guinness FE, Albon SD (1982) *Red Deer: Behaviour and Ecology of Two Sexes*, 1st edn. The University of Chicago Press.

Cruzan MB, Arnold ML (1993) Ecological and genetic associations in an iris hybrid zone. *Evolution* **47**, 1432-1445.

Cook CE (1993) *Historical Biogeography and Phylogeny of Sika Deer (Cervus nippon) in East Asia*, University of California

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

Ellstrand NC, Schierenbeck KA (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences of the United States of America* **97**, 7043-7050.

Diaz 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.

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.

Georges M, Massey J (1992) Polymorphic DNA markers in bovidae (ed. Org WIP), Geneva, Switzerland.

Goodman SJ, Barton NH, Swanson G, Abernethy K, Pemberton JM (1999) Introgression through rare hybridisation: 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.

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **88**, 485-486.

Grosse WM, Finlay O, Kossarek LM, Clark TG, McGraw RA (1995) 5 Bovine microsatellite markers derived from skeletal-muscle cDNA - Rme01, Rme11, Rme23, Rme25 and Rme33. *Animal Genetics* **26**, 126-127.

Harrington R (1973) Hybridisation among deer and its implication for conservation. *Irish Forestry* **30**, 64-78.

Harrington R (1979) *Some Aspects of the Biology and Taxonomy of the Deer of the County Wicklow Region, Ireland.*, National University of Ireland. PhD thesis

Harris S, Yalden DW (2008) *Mammals of the British Isles*. The Mammal Society. 4th edition

Harrison RG, Rand DM (1989) Mosaic hybrid zones and the nature of species boundaries. In: *Speciation and its consequences* (eds. Otte D, Endler JA), pp. 111-133. Sinauer, Sunderland.

Harrison RG, Bogdanowicz SM (1997) Patterns of variation and linkage disequilibrium in a field cricket hybrid zone. *Evolution* **51**, 493-505.

Hoffman JI, Amos W (2005) Microsatellite genotyping errors: detection approaches common sources and consequences for paternal exclusion. *Molecular Ecology* **14**, 599-612.

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

Jiggins CD, Mallet J (2000) Bimodal hybrid zones and speciation. *Trends in Ecology & Evolution* **15**, 250-255.

Jiggins CD, McMillan WO, King P, Mallet J (1997) The maintenance of species differences across a *Heliconius* hybrid zone. *Heredity* **79**, 495-505.

Johnston JA, Wesselingh RA, Bouck AC, Donovan LA, Arnold ML (2001) Intimately linked or hardly speaking? The relationship between genotype and environmental gradients in a Louisiana iris hybrid population. *Molecular Ecology* **10**, 673-681.

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.

Kocher TD, Thomas WK, Meyer A*, et al.* (1989) Dynamics of mitochondrial-DNA evolution in animals - amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 6196-6200.

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.

Lecis R, Pierpaoli M, Biro ZS*, et al.* (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology* **15**, 119-131.
Livingstone SR (2001) *The Application of GIS to the Spread of Japanese Sika Deer (Cervus nippon) in Scotland*, unpublished MSc thesis, University of Edinburgh
Lowe VPW, Gardiner AS (1975) Hybridization between red deer and sika deer with particular reference to stocks in N.W. England. *Journal of Zoology* **177**, 553-566.
Ludt JC, Schroeder W, Rottmann O, Kuehn R (2004) Mitochondrial DNA phylogeography of red deer *(Cervus elaphus)*. *Molecular Phylogenetics and Evolution* **31**, 1064-1083.
Mack RN, Simberloff D, Lonsdale WM*, et al.* (2000) Biotic invasions: causes, epidemiology, global consequences, and control. *Ecological Applications* **10**, 689-710.
Martin NH, Bouck AC, Arnold ML (2007) The genetic architecture of reproductive isolation in Louisiana irises: Flowering phenology. *Genetics* **175**, 1803-1812.
Martin NH, Sapir Y, Arnold ML (2008) The genetic architecture of reproductive isolation in Louisiana irises: Pollination syndromes and pollinator preferences. *Evolution* **62**, 740-752.
Mezzelani A, Zhang Y, Redaelli L*, et al.* (1995) Chromosomal localization and molecular characterization of 53 cosmid-derived bovine microsatellites. *Mammalian Genome* **6**, 629-635.
Mommens G, Coppieters W, Vandeweghe A, Vanzeveren A, Bouquet Y (1994) Dinucleotide repeat polymorphism at the bovine Mm12e6 and Mm8d3 loci. *Animal Genetics* **25**, 368-368.
Moore SS, Barendse W, Berger KT, Armitage SM, Hetzel DJS (1992) Bovine and ovine DNA microsatellites from the Embl and Genbank databases. *Animal Genetics* **23**, 463-467.
Moore SS, Byrne K, Berger KT*, et al.* (1994) Characterization of 65 bovine microsatellites. *Mammalian Genome* **5**, 84-90.
Nielsen EE, Hansen MM, Bach LA (2001) Looking for a needle in a haystack: discovery of indigenous Atlantic salmon (*Salmo salar* L.) in stocked populations. *Conservation Genetics* **2**, 219-232.
PACEC (2006) *The Contribution of Deer Management to the Scottish Economy*. Public and Corporate Economic Consultants, London and Cambridge.
Pemberton JM, Slate J, Bancroft DR, Barrett JA (1995) Nonamplifying alleles at microsatellite loci - a caution for parentage and population studies. *Molecular Ecology* **4**, 249-252.
Powerscourt V (1884) On the acclimatization of the Japanese deer at Powerscourt. *Proceedings of the Zoological Society of London* **1884**, 207-209.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
Randi E, Mucci N, Carlo-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.
Randi E (2008) Detecting hybridization between wild species and their domesticated relatives. *Molecular Ecology* **17**, 285-293.
Ratcliffe PR (1987) Distribution and current status of sika deer, *Cervus nippon*, in Great Britain. *Mammal Review* **17**, 39-58.
Ross CL, Harrison RG (2002) A fine-scale spatial analysis of the mosaic hybrid zone between *Gryllus firmus* and *Gryllus pennsylvanicus*. *Evolution* **56**, 2296-2312.
Schilthuizen M (2000) Bimodal hybrid zones and the scale of a snail. *Trends in Ecology & Evolution* **15**, 469-469.

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.

Slate J, Van Stijn TC, Anderson RM*, et al.* (2002) A deer (subfamily cervinae) genetic linkage map and the evolution of ruminant genomes. *Genetics* **160**, 1587-1597.

Smadja C, Ganem G (2005) Asymmetrical reproductive character displacement in the house mouse. *Journal of Evolutionary Biology* **18**, 1485-1493.

Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*, near Cracow in southern Poland. *Evolution* **40**, 1141-1159.

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.

Vähä J-P, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15**, 63-72.

Vaiman D, Mercier D, Moazamigoudarzi K*, et al.* (1994) A Set of 99 cattle microsatellites - characterization, synteny mapping, and polymorphism. *Mammalian Genome* **5**, 288-297.

Vaiman D, Osta R, Mercier D, Grohs C, Leveziel H (1992) Characterization of 5 new bovine dinucleotide repeats. *Animal Genetics* **23**, 537-541.

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

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population-structure. *Evolution* **38**, 1358-1370.

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

Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *Journal of Evolutionary Biology* **14**, 611-619.

Yanchukov A, Hofman S, Szymura JM*, et al.* (2006) Hybridization of *Bombina bombina* and *B. variegata* (Anura, Discoglossidae) at a sharp ecotone in western Ukraine: comparisons across transects and over time. *Evolution* **60**, 583-600.

Chapter 3

Variation in introgression under rare hybridisation: the influence of stochasticity

H. V. Senn[1,2], N. H. Barton[3]

[1] Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JT, UK

[2] Macaulay Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK

[3] Institute of Science and Technology, Am Campus 1, Klosterneuburg, Austria A3400

Authors' Contributions:

HVS wrote simulations code and the manuscript. NHB aided HVS with coding and commented on multiple drafts of the manuscript. NHB also contributed the equations in the Appendices.

3.1 Abstract

In a study of hybridisation between red and sika deer on the Kintyre Peninsula, Goodman *et al.* (1999) estimated the rate of hybridisation to be 1:500-1000 matings, using the pattern of Linkage Disequilibrium at 11 diagnostic loci in a sample of 246 deer. Despite the fact that F1 hybridisation is rare, backcrossing of the hybrids into each parental population leads to substantial introgression. Many studies have used the differential pattern of introgression at marker loci to infer selection in hybridising populations. Here we use a branching process model, to simulate introgression via backcrossing when hybridisation is rare. The results of this simulation show that if the number of F1 hybridisation events contributing to introgression is small, this can generate highly differential patterns of introgression at marker loci by chance. This pattern of neutral variation in introgression can have high enough variance that it could be mistaken for selection and so even if strong selection is acting, it may not be possible to distinguish its effect from neutral variation. This indicated that it was not appropriate to look for evidence of selection, using the differential in the pattern of introgression across marker loci, in the red-sika hybridisation system on Kintyre. Our findings also have wider implications which are discussed.

3.2 Introduction

The extent of introgression between distant populations is important both for understanding the evolution of reproductive isolation and also for conserving endangered species. Variation in the degree of introgression across loci may be caused by selection on the introgressing alleles or on linked loci. Several studies have used the pattern of introgressing alleles at a number of loci to identify genes responsible for reproductive isolation (e.g. Rieseberg et al 1999; Wilding *et al.* 2001, Akey *et al.* 2002, Martin *et al.* 2006). However, unless many hybridisation events have taken place between the populations, the degree of introgression at different loci will be highly variable even if they are all neutral. The aim of this paper is to examine how the frequency of hybridisation affects the extent of stochastic fluctuation in gene flow and thus the ability to detect selection.

Various approaches involving the analysis of the pattern of introgression of molecular markers in hybridising taxa exist. These can be categorised into five areas:

1. Cline analysis (Haldane 1948; Endler 1977); If populations meet in a set of continuous clines maintained by a balance between selection and dispersal, then the width of those clines at introgressing molecular markers can be used to infer the strength of selection that maintains them. If the width (w) of the cline is narrower than expected, given its age and the expected diffusion of genes across the hybrid zone under neutrality (σ), then the strength of selection (s) maintaining it can be estimated from $w \sim \frac{\sigma}{\sqrt{s}}$.

2. Frequency of introgressed markers; Here, the frequency, as opposed to the spatial arrangement of introgressing molecular markers, is assessed. The premise is simple: neutral regions of the genome should all introgress to the same frequency, whereas selected genes

should introgress to higher or lower frequencies than average. Markers of known allele frequencies in the parental populations are used to identify blocks of the genome which are introgressed. The frequencies of these is taken to indicate their effect on fitness (Martin *et al.* 2006; Rieseberg *et al.* 1999).

3. The pattern of F_{ST} (Lewontin-Krakauer test) (Lewontin & Krakauer 1973), reviewed in Beaumont (2005); The premise of the Lewontin-Krakauer test is that drift and demographic history should have the same influence on F_{ST} across all loci whereas selection acts on individual loci: High values of F_{ST} (i.e. large allele frequency differences between the populations) indicate loci with alleles selected differently in different populations; low values of F_{st} (similar allele frequencies in populations) indicate balancing selection. This approach is fundamentally the same as comparing the frequency of introgressed markers (previous), except in this case no prior assumptions about the origins of the alleles are made. Large numbers of polymorphic markers are needed to give sufficient power to detect selection (Akey *et al.* 2002; Wilding *et al.* 2001).

4. Assignment methods; These encompass a variety of methods of assigning suspected hybrids to one or more populations (reviewed in Manel *et al.* (2005)). Two approaches exist: either alleles are designated to populations from the outset (Boecklen & Howard 1997; Goodman *et al.* 1999); or no prior assumption about their origin is made. Then the pattern of linkage disequilibrium is used to identify separate gene pools and infer population origin (Anderson & Thompson 2002; Corander & Marttinen 2006; Falush *et al.* 2003; Pritchard *et al.* 2000).

5. Discordant genealogies; This method is used to investigate historical hybridisation between diverged taxa. It involves the comparison of inferred genealogies at several loci

(Kliman *et al.* 2000; Machado *et al.* 2002; Patterson *et al.* 2006). If hybridisation events have occurred since the taxa diverged, then introgressed regions of the genome will exhibit different patterns of divergence to the rest of the genome.

In this paper we will focus on how frequency of hybridisation might affect the pattern of introgression at molecular markers and so also the ability to detect selection in approaches that rely on comparing the pattern of introgression to that expected under neutrality (i.e. 2&3). These two methods of analysis depend on the reasonable assumption that large deviations from the expected pattern of introgression under neutrality reflect the action of selection. Under neutral conditions we would expect genetic markers to introgress to roughly the same extent across all loci within the hybrid population. Markers associated with selected regions of the genome will introgress to a greater or lesser extent than average, and this deviation in introgression can, in principle, be detected by the various methods discussed above. But if stochastic effects are strong, can we expect a pattern of introgression to be a reliable signature for selection?

Consider a single hybridisation event between two diploid individuals: the offspring is fertile and backcrosses into one of the two parent populations. As backcrossing proceeds and produces each subsequent descendant, the fraction of the descendants' genomes that is introgressed falls by, on average, a half every generation. In the initial backcross generations, linkage disequilibrium is high and the fates of many introgressing genes are intertwined. At this point, if some genes confer a severe fitness disadvantage they will drag the remaining genes in the genome to extinction with them. However, as backcross generations proceed, random assortment and recombination free the incoming genes from their original genetic background and linkage equilibrium is approached. *The extent to which any particular gene introgresses depends on the interplay of three factors: 1. the selective advantage of the gene itself, 2. the overall fitness of the genome it is linked to, 3. chance, both because of random*

factors in individual fitness and because of recombination and random assortment of genes during reproduction.

In fact, chance plays a large role in the descent of genes through any one individual**.** Baird *et al* (2003) investigated the survival of a single genome through time. In a sexual population at equilibrium (reproductive rate =2) the expected number of descendants doubles every generation but the expected genetic contribution of any one of their ancestors is halved. In a stable population, the average proportion of ancestral genome remains constant at one through time. However, random reproduction and recombination generate ever increasing variance surrounding this expectation and the probability of survival becomes highly skewed: for example, after 30 generations of reproduction most individuals leave only a few genetic descendants, (although there may be pedigree descendants), but there is an appreciable chance (2.3% for a stable population) that an individual will leave more than 300 genetic descendants. These descendants will each carry short blocks of the original genome. Thus, gene flow following one hybridisation event has the potential to result in very uneven patterns of introgression under neutrality.

One would therefore expect the effect of stochasticity on introgression between hybridising taxa to be high. Firstly, if the two populations are arranged spatially in a narrow hybrid zone, the region in which they come into contact may represent a small fraction of the total population; the genes from one population effectively enter a bottleneck as they flow into the second. Secondly, depending on the strength of reproductive isolation, cross-mating may be rare. Since stochasticity in gene flow from one individual is high, it follows that the lower the number of hybridisation events, the greater the effect of chance on the pattern of introgression, making the signal from selection harder to detect.

Although there are many reviews of the incidence of hybridisation in nature (e.g. Mallet 2005; Randler 2002; Rieseberg *et al.* 2006; Seehausen 2004; Schwenk *et al.* 2008), little work has been carried out on the frequency of hybridisation within hybridising populations and its contribution to their genetic structure (see Discussion). However, different types of hybridising populations clearly exist; in a study of hybridisation between native red deer (*Cervus elaphus*) and introduced sika deer (*C. nippon*) in Argyll, Scotland, hybridisation rates were discovered to be extremely low; around 1 in 500 to 1 in 1000 matings (Goodman *et al.* 1999). This low rate of hybridisation was confirmed when a subsequent study (Chapter 2) failed to find a single F1 in a sample of 735 individuals. Hybridisation has been occurring at the most for around 115 years, meaning that the total number of hybridisation events that have occurred in the history of the hybrid zone may be quite low. Despite this up to 20- 40% of deer sampled in areas of overlap of the two species carry introgressed alleles (Chapter 2, Table 4). This hybrid zone contrasts strongly with hybrid zones that are many thousands of years old in which mating may be approximately random within the hybrid zone (e.g. *Bombina*, Szymura & Barton (1986)).

Here we consider a scenario of rare hybridisation similar to red-sika hybridisation on the Kintyre Peninsula. We ask whether it is possible to detect selection against a background of stochasticity using a panel of unlinked molecular markers. How does the frequency of hybridisation events influence this pattern?

3.3 Methods

We model a scenario where a panel of unlinked, markers are used to survey hybridisation in a natural population where the total number of hybridisation events in the population is low.

When this number is low relative to the population size, the chance of hybrids meeting and mating is small, because F1 hybrids are only likely to encounter ‘pure’ individuals. This means that subsequent generations can be treated as a series of successive backcrosses into one or both ‘pure’ populations. This involves the simplifying assumption that backcross individuals only mate with the parental type that they are genetically closest to (Figure 1).

Provided that the population is large, the generation of individuals by successive backcrosses in this way, can be modelled as a branching process. (For a mathematical explanation, see Grimmett and Stirzaker (2001) section 5.4) Branching processes are used to model the individuals that reproduce independently. Each individual in the t’th generation produces a random number of offspring in generation t+1, according to a chosen probability distribution. Branching processes are memory-less processes, meaning that the number of offspring an individual produces is independent of the family size that it came from.

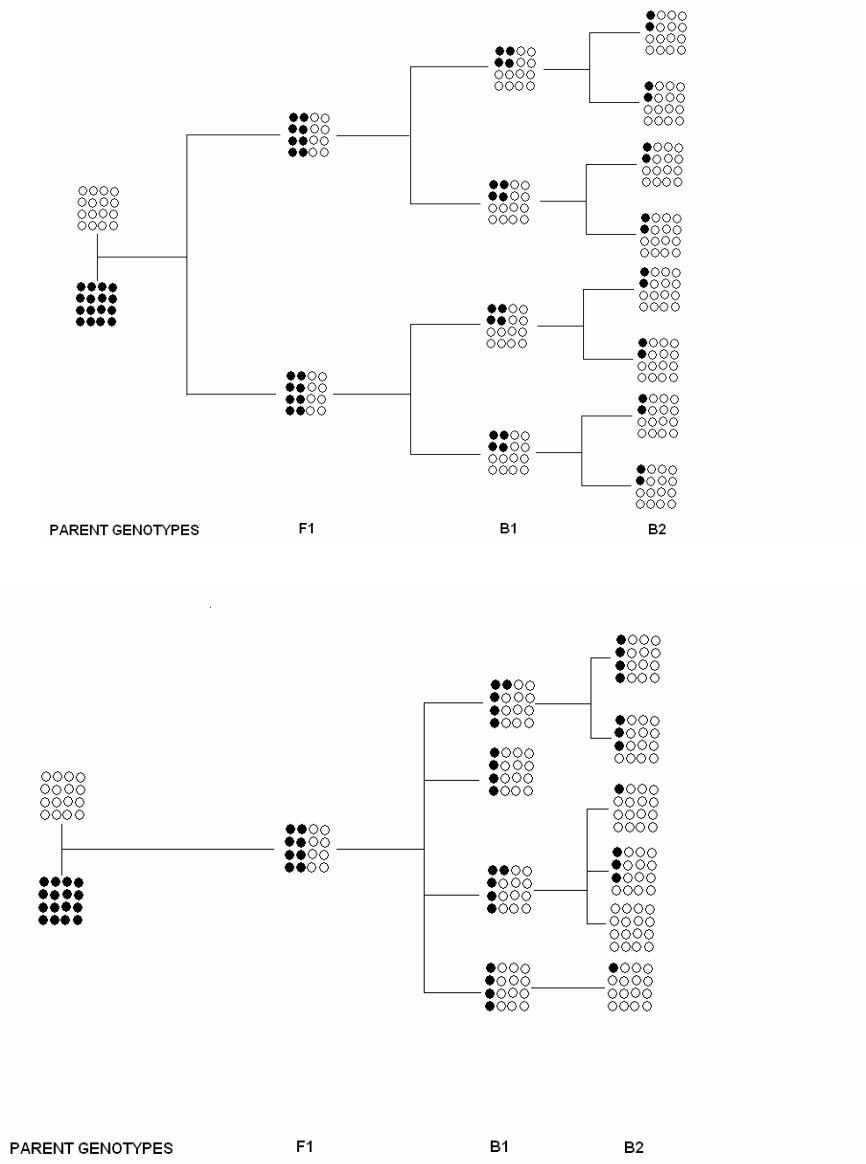


Figure 1: A scenario of backcrossing surveyed at 8 diagnostic loci (16 alleles, represented by circles). Two individuals mate to produce a number of F1 hybrid offspring. The subsequent generations always mate with one parental type (in white, for simplification these joins not shown on the pedigree). Thus the contribution of genes from the other parent (in black) is reduced each generation. The upper cartoon illustrates the deterministic view of this process, where each individual produces a fixed number of offspring (two in this case) and the genetic contribution of the black parent is reduced by exactly 50% each generation. One possible representation of the stochastic view is depicted at the bottom (generated using the branching process model used in this paper). Stochasticity is introduced both through the reproductive success of individuals and the random assortment of genes during meiosis.

This branching process model of introgression via backcrossing follows the individuals generated by a single hybrid cross over any number of generations. The model makes two crucial approximations: firstly, that hybridisation is so rare that introgressed offspring never mate; and secondly, that backcrossing is always into only one of the two populations. Both of these approximations depend on the underlying assumption that the 'pure' population into which the hybrids are backcrossing is large. The initial hybridising individuals have a number of independent marker loci, all of which are bi-allelic and which distinguish them as coming from one or other population (i.e. they are diagnostic). At each generation, each introgressing allele has a 50% chance of being passed on (representing random assortment). In a selectively neutral scenario, the number of individuals produced per individual per generation is drawn from the Poisson distribution with a mean of 2. Selection can be modelled by linking a selectively advantageous gene to a marker locus with a desired probability of them recombining each generation. Selective advantage is modelled by increasing the rate of reproduction in individuals carrying the advantageous gene (i.e. increasing the mean of the Poisson distribution of offspring number). This is quoted as a percentage of the base reproductive rate (2), so that a gene conveying a 100% selective advantage would give the bearer a mean reproductive rate of 4.

In each case the model was run for a number of generations with a desired number of loci for a desired number of replicates. Once an individual had become fixed for the background alleles, its offspring were no longer followed in the simulation to improve computational efficiency. In any case, these pedigree descendants can no longer be distinguished from the background population.

We first investigate backcrossing following a single hybridisation event at a single locus and then extend the investigation to multiple loci and multiple hybridisation events. Next, we examine the possibility of detecting selection in an idealised scenario where large numbers

of diagnostic markers are available and all offspring of a known number of generations of backcrossing have been followed. Finally, we consider the more realistic case where time since the first hybridisation event is not known and the markers are not completely diagnostic.

All simulations and calculations were carried out using Mathematica 5.2.

3.4 Results

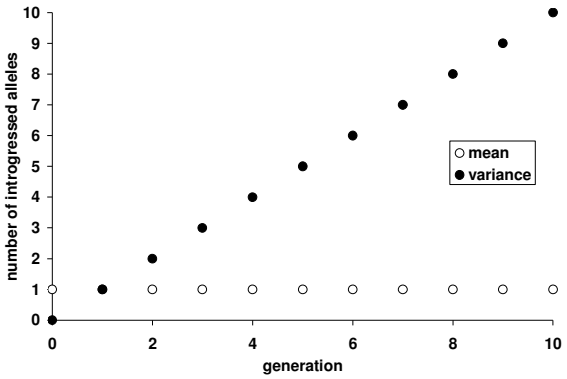
3.4.1 Introgression at one locus

First, we examine the prediction for introgression at a single locus. Each individual produces *n* offspring according to a Poisson distribution with parameter 2λ, with the probability $\frac{e^{-2\lambda}(2\lambda)^n}{n!}$. Since the introgressing allele is only passed 50% of the time, we can expect it to be passed on in j copies with probability $\frac{e^{-\lambda}(\lambda)^j}{j!}$. The probability distribution of the number of copies of a gene after *t* generations can be found using generating functions (see Appendix 1).

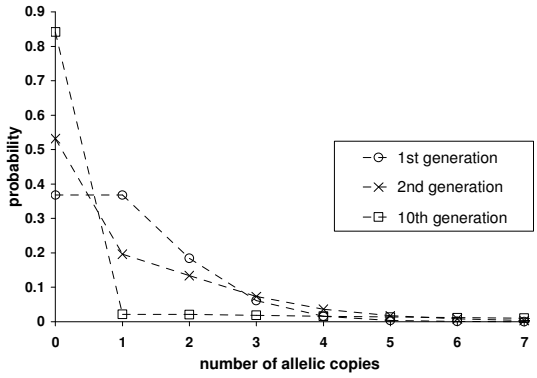
The mean number of individuals carrying introgressed alleles at this one locus remains constant, because although the number of offspring is expected to double per generation, the genetic contribution of the ancestor in question is expected to be halved (i.e. $\mu = 2^t \times \frac{1}{2^t} = 1$). However, the variance increases linearly with time ($\sigma^2 = t$, Figure 2a). In other words as the generation proceeds the variability in introgression increases. For

example, at ten generations most replicates of the model no longer carry any descendants with introgressed alleles, but some have in excess of 40 (Figure 3).

a)



b)



c)

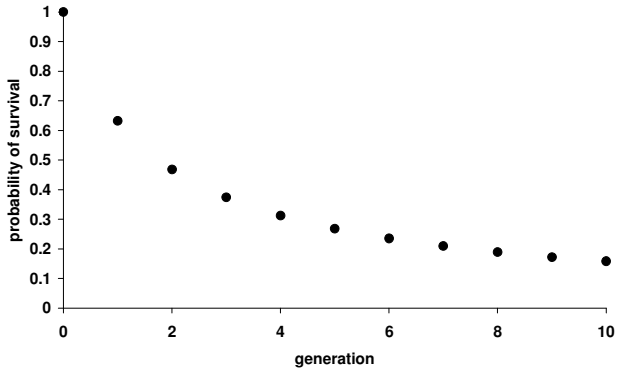
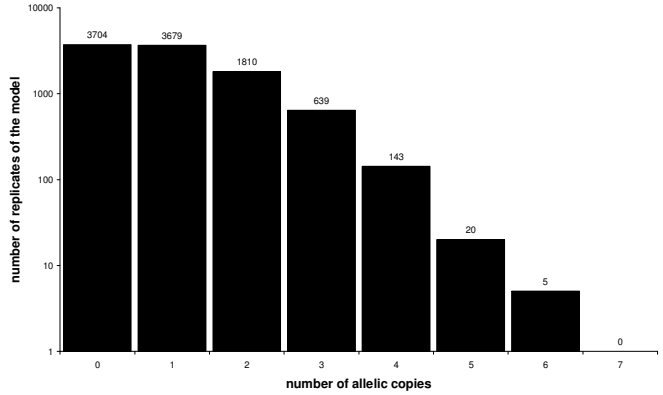


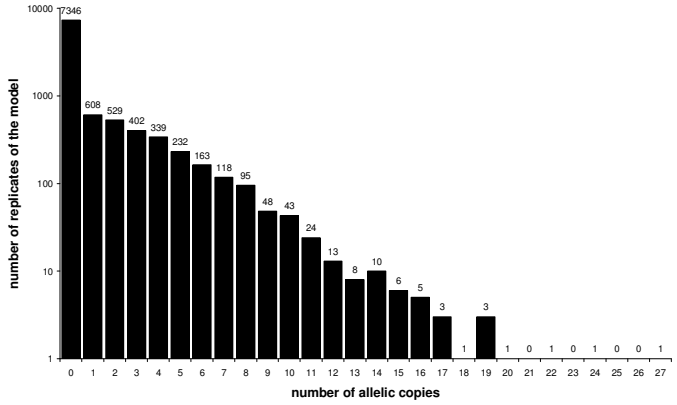
Figure 2: Predictions from the generating function at one locus: a) the expected mean and variance in introgressed allele number over ten generations of backcrossing b) the probability distribution for the numbers of copies of the introgressed alleles at two, three and ten generations. c) the decay in probability of at least one copy of the introgressed allele surviving over ten generations. The lines have been added for clarity.

The probability distribution for the number of copies of alleles passed on to each generation changes over the generations (Figure 2b). The probability of having any genetic descendants (at this locus) decays over time, with the decay being particularly steep initially and then beginning to plateau (Figure 2c). At this point the expected number of pedigree descendants is very high (e.g. 2^{10} at generation 10). Because of such a large number of expected descendants there is still after this time a reasonable chance that at least one of them carries an introgressed allele (at generation ten this is ~0.16). The results of the model match the predictions of the generating function, but because of the high stochasticity around 10,000 replicates are needed for the pattern to emerge convincingly.

a)



b)



c)

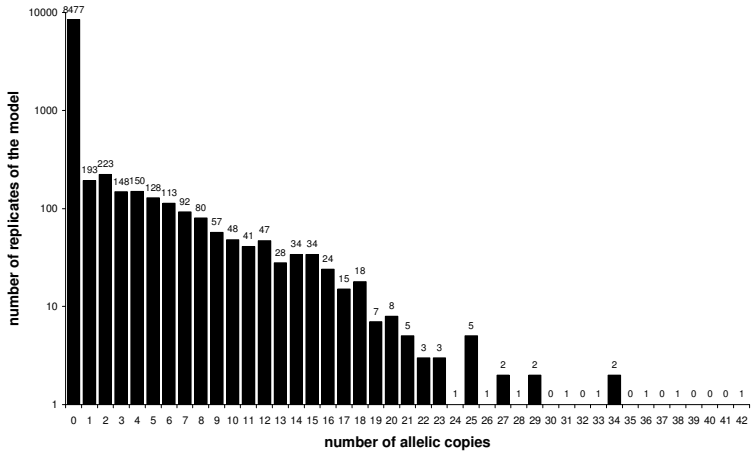


Figure 3: The distribution of the numbers of copies of introgressed alleles in ten thousand replicates of the model of backcrossing at one locus after a) one, b) five, and c) ten generations of backcrossing. As the generations progress, most replicates go extinct for the introgressing allele (in this case 3704 after generation one, 7346 after generation five, 8477 after generation ten). However, a few replicates of the model carry very high numbers of allele copies (e.g. one replicates of the model carries 42 introgressed alleles after 10 generations). Note that the y-axis is on a logarithmic scale. The exact numbers of surviving replicates are given at the top of each bar.

3.4.2 Multiple markers

Before examining the case of multiple markers, it is necessary to consider whether a branching process model is, in fact, an adequate approximation to this scenario. After all, in a finite population related individuals will eventually mate and ancestral blocks will be recombined in the same genome, thus breaking the assumptions of the branching process model (which requires that individuals backcross with unrelated individuals). Baird *et al*. (2003) used a similar branching process model to this one, to follow the erosion of an entire block of genome instead of the fate of alleles at independent loci. They investigated the robustness of the branching process for this scenario.

The branching process in fact turns out to be a good approximation, providing the population is large. This is because in a large population, by the time the descendants are numerous enough to be likely to mate, the blocks of ancestral material they carry are so small that they are highly unlikely to be passed on together for any length of time (a few will become tightly linked and take a long time to break up). More formally, for population size N over t generations, the timescale for the breakdown of the branching process is t ~N (the point at which individual genes are likely to become fixed). In fact, long before that (~log2N), all individuals in the population are pedigree ancestors meaning that the formal assumptions of the branching process model have been broken. Provided N is large the blocks are small enough by the time the assumption is broken for the approximation to remain accurate (Baird *et al.* 2003).

The branching process assumption should then hold for the case of unlinked markers, since if highly backcrossed individuals meet, their blocks of markers are very likely to be broken up

in the subsequent generations. Further to this, we are following hybridisation over a small number of generations, so the breakdown in the assumptions of the branching process is not likely to be reached for any population of appreciable size.

3.4.3 Multiple unlinked markers, multiple genotypes

If one considers only *two* unlinked loci, in the first generation, the probability of them being passed on together is simply half the probability of them being passed on alone, i.e. ¼. More generally, in a bi-allelic system of m unlinked loci each introgressed haplotype has 2^m possible combinations of genotypes. Since only one of these possibilities is introgressed at all loci, it is passed on with a probability of $1/2^m$. The expected number of completely introgressed genotypes under a birth rate, Φ, is $\Phi/2^m$.

However, in order to track an array of different introgressed genotypes the situation becomes more difficult: Of the 2^m possible genotypes, those containing k introgressed alleles can be arranged in $\frac{m!}{k!(m-k)!}$ possible ways. So as m increases, the branching process becomes highly complex. Any genotype involving k introgressed alleles has a prior probability of $2^{-m}\frac{m!}{k!(m-k)!}$, but the expected number of genotypes is *not* simply $\phi 2^{-m}\frac{m!}{k!(m-k)!}$ because the probability of an individual's genotype changes according to its parent's genotype.

The expected number of genotypes in a generation must be $\phi\lambda_{X,Y}2^{-m}$, where $\lambda_{X,Y}$ is the number of possible offspring of type Y from type X. From this a generating function for the probability distribution of numbers of copies of different genotypes can be derived (Appendix 2). The generating function was found to be only computationally tractable for a small number of loci (m<7), because the 2^m possible genotypes rise rapidly with m. Each

possible genotype with *m* loci and *k* introgressed alleles has a different probability of survival over time. Under this model of continuous backcrossing the expectation is that first generation individuals carry introgressed alleles at 50% of surveyed loci, 2nd generation individuals at 25%, 3rd generation individuals and 12.5% and so forth. Probability distributions of the survival of different genotypes over time peak at these expected generations, but the probability of finding a non expected genotype in any given generation can still be high. For example, for the case *m*=4 and *k*=2, the peak of the probability distribution (p=0.53) corresponds to generation 1. However, the probability of finding this genotype at generation 3 is still 0.35 (Figure 4). This clearly illustrates the point, made previously by Boecklen & Howard (1997), that the introgressed proportion of marker loci (hybrid index) cannot be used reliably to classify hybrid individuals into backcross generations unless the number of markers used is very high.

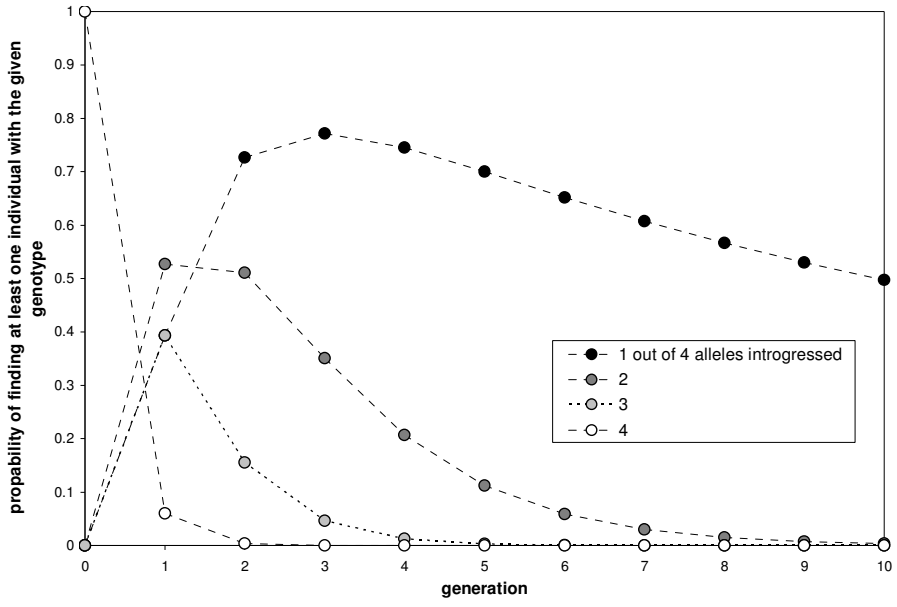
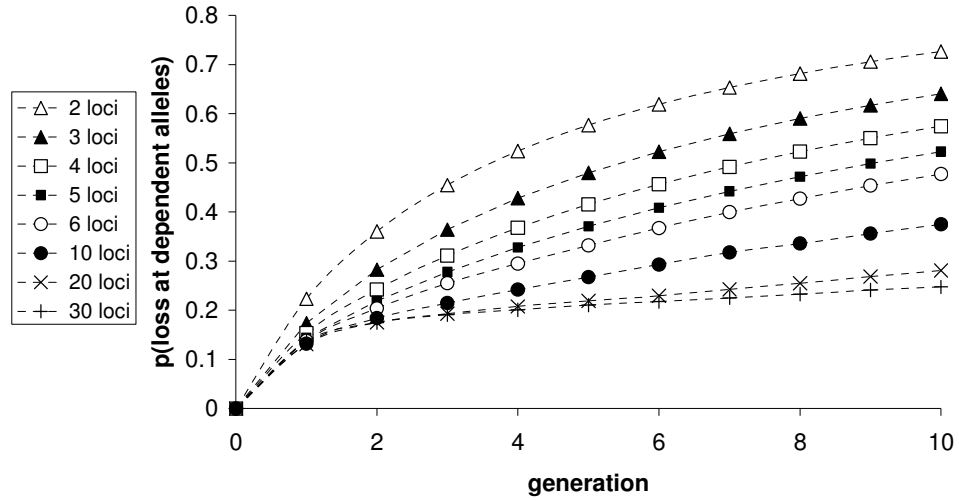


Figure 4: The probability that at least one copy of a genotype containing a given number of alleles survives over 10 generations of backcrossing where the mean reproductive rate is 2 (drawn from the Poisson distribution). Illustrated here is the case were 4 loci are followed. Dashed lines are added for clarity.

Despite the fact that the loci under investigation are chromosomally unlinked, they cannot be treated as independent due to their joint genealogical history. In the initial few generations, while the introgressed alleles are more likely to be present in the same genome, the probability of loss of all introgressing alleles at all *m* surveyed loci, (henceforth, the *probability of total loss* at *dependent loci*), is much higher than the probability of total loss from the hypothetical case of *m* truly independent loci that do not share genealogical history (*not* just independent in the sense of unlinked). This is because for dependent loci, the fate of all the introgressing alleles is greatly influenced by the outcome of the small number of reproductive events at the start of the branching process. This number of events is greater by a factor of *m* for truly independent loci. So, although at increasing numbers of surveyed loci the probability of total loss decreases, the probability of total loss relative to the case of truly independent loci, increases. This is because with a high number of truly independent loci the effect of stochasticity in the early generations is smoothed. Eventually, as the generations proceed assortment breaks up the associations between the dependent loci (Figure 5).

a)



b)

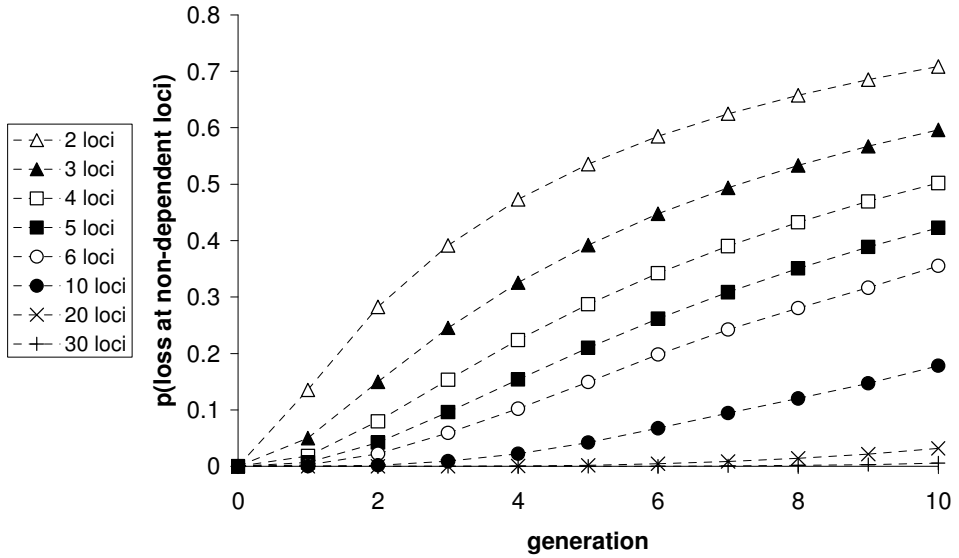


Figure 5: a) The probability of total loss at *m* loci (dependent loci); b)The probability of total loss *m* individual loci (non – dependent loci). In the initial generations the probability of loss is relatively high for the dependent loci, this illustrates the high effect of reproductive stochasticity at this stage. Despite the fact that the probability of total loss is lower at large numbers of loci, the probability relative to the non-dependent case is higher. Failure of one or a few reproductive events can lead to the elimination of introgressed alleles at all loci in early generations, but this cannot happen in the non-dependent case, where increasing number of loci make the probability of total loss very unlikely. As the generations pass, recombination breaks up the alleles and the two distributions begin to look more similar. The probability distributions for 2-5 loci were generated using the generating function, those for larger numbers using the average results from 10000 replicates of the model. Dashed lines are added for clarity.

3.4.4 Variation in introgression

So far we have demonstrated that backcrossing has the potential to result in highly variable levels of introgression at marker loci under neutral conditions. We also expect selection to result in variation in the levels of introgression across loci, with alleles under selection, or linked to genes under selection introgressing to a greater or lesser extent than neutral alleles. In order to understand the pattern of neutral variation in introgression and understand whether we can distinguish it from patterns generated by selection, it is first necessary to develop a convenient measure with which to summarise population-wide variation in introgression across multiple marker loci. An appropriate measure is a variance measure which captures the *difference* in population-wide levels of introgression between multiple loci. Variance in the level of introgression across loci should be high if selection is acting on some alleles and low if they are all neutral.

The measure used will be referred to as the standardised variance of introgression (or SI for ease).

$$SI = \frac{Var[\{x_1, x_2, ...x_m\} / IM]}{m}$$

$$IM = \sum_{i=1}^{m} \frac{x_i}{n}$$

x_i= generation-wide introgression at locus i at time t

m= total number of loci surveyed

n= total number of individuals carrying at least one introgressed allele

IM= mean level of introgression per individual hybrid

Variance is calculated as the true variance (not sample variance) because it is based on all hybrid individuals in the population. It is important to note that this measure of variance only

includes individuals carrying at least one introgressed allele, as it would not be possible to distinguish the other pedigree descendants from the background population. Only individuals in generation *t* are included in the measure. The measure of SI is standardised for the number of hybrid individuals in the population (at generation t) and the number of loci sampled. Only measure of SI of populations in the same generation can be compared, since it will change over time.

SI is highly variable across replicates. An increase from 5 to 30 loci does not noticeably reduce the range in the SI at five hundred replicates of the model (Figure 6a-c). The high peak at SI=1, represents the large number of replicates in which introgressed individuals are all introgressed at the same locus. This pattern of introgression leads to high variance across loci. This peak is greatly reduced when larger numbers of loci are surveyed. This is because individuals are more likely to carry more than one introgressed allele. The proportion of replicates that go to total loss is also lower as the number of loci is increased.

Although not investigated here (because the simulations become very inefficient at higher numbers of loci), these results imply that a limit for *m* will be reached, beyond which point the variance in introgression will not decrease substantially. The pattern of SI for 20 loci, is similar to that for 30. This is hinted at in Figure 5a, where the probability of total loss at generation *t* appears to be converging as loci numbers increase, but this would need to be verified mathematically.

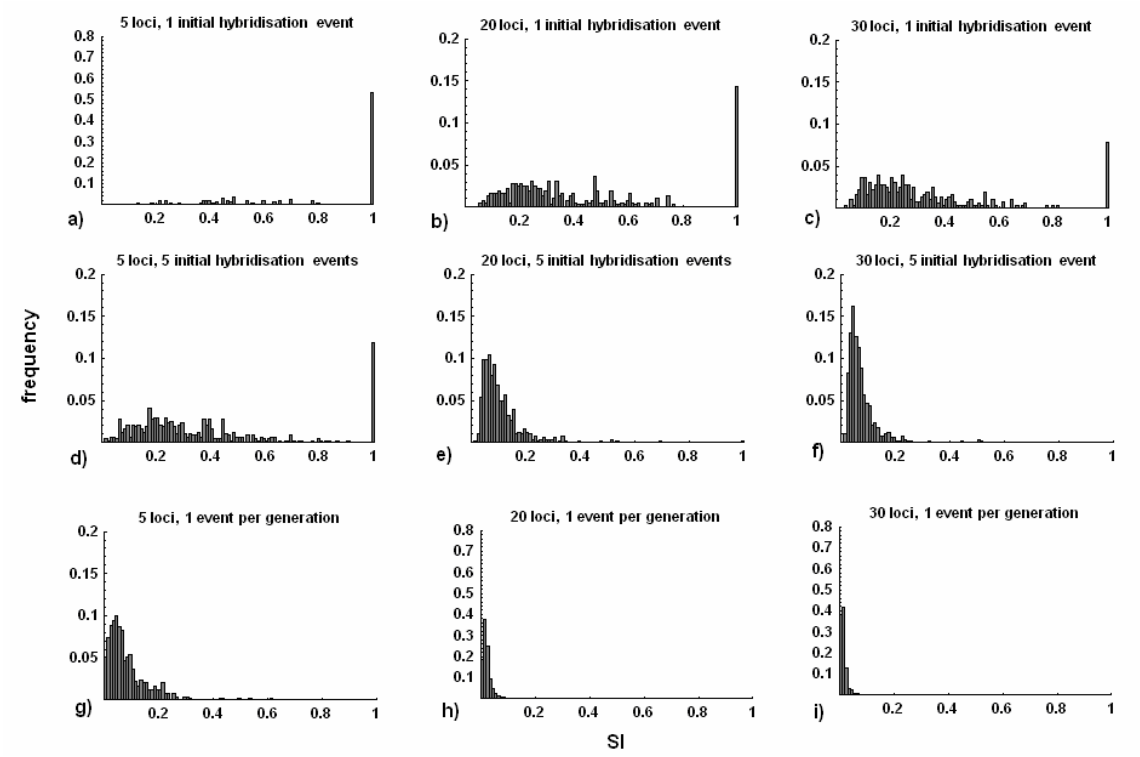


Figure 6: Frequency histogram of the standardised introgression (SI) across loci in five hundred replicates of the model run for ten generations, binned into categories of width 0.01. Y-axis: Frequency, x-axis: SI. The high peak at SI= 1 in some of the plots is caused by replicates in which only introgressed alleles at one of the loci survive. This phenomenon is very common at low number of loci and when hybridisation occurs only once. Increased numbers of loci and, in particular, more hybridisation events make the chances of introgression only occurring at one locus very low and also reduce the range of SI. Note: The axes in plots a, h and i are not to the same scale as the others.

3.4.5 Multiple hybridisation events

How does the pattern of introgression change with more hybridisation events? The scenario of multiple hybridisation events modelled here is the situation where a given number of hybridisations take place but the conditions of infinite backcrossing are still satisfied (i.e. hybridisation is not common relative to the total population size). This can be modelled easily by combining the results for single hybridisation events.

With a single locus, additional initial hybridisation events lead to an increase in the proportion of surviving replicates over time. The mean amount of ancestral material present after each generation is equal to the number of hybridisation events. At multiple loci, even five initial hybridisation events still result in a large range in SI across loci (Figure 6d-f). The addition of new hybridisation events every generation (modelled stochastically by a Poisson distribution with mean of one), results in a considerable reduction in the range of SI, especially at surveys of 20-30 loci (Figure 6g-i). As the number of hybridisation events increases, the range in SI decreases, because fewer replicates exhibit extreme allele frequency differences across loci.

3.4.6 Selection

If neutral markers are dispersed across a genome, necessarily some will be linked to selectively advantageous or disadvantageous genes. If a marker is fairly tightly linked to a selectively advantageous gene then its pattern of introgression may reflect this. It has been shown above, that in scenarios of rare hybridisation, high variance in the level of introgression can be generated even under neutral conditions. The question remains whether, given this high background level of stochasticity, selection can still be detected.

It can be seen from Figure 6, with a low number of markers (<20) and low number of hybridisation events (< 1 per generation), that the range in SI is so high that the neutral scenario will never be distinguished with any certainty from a scenario involving selection. This is because high variability in introgression across loci can be generated even under neutral conditions. However, at high numbers of marker loci and hybridisation at a rate of one event per generation**,** it may be possible to detect selection. The simplest scenario is where one single gene with a selective advantage is linked to one single marker locus with a given recombination rate. This scenario is also the case in which selection is going to be easiest to detect. It is expected that the greater the difference in selective pressures on the

marker loci, the greater the difference in extent of their introgression. Seen from a different view point - if all markers were associated with the same degree of selective advantage, the variance in introgression would be low, making selection harder to detect.

A hypothetical 'best case scenario' to detect selection under rare hybridisation might be a study using 30 loci, in which hybridisation is occurring at a rate of one event per generation. Under this scenario, at neutrality, the range in SI is fairly low (Figure 6i). Selection leads to a large increase in the range in variance, especially if recombination is low (Figure 7). At high levels of recombination (>0.2), even high selective advantage (75%, giving an average reproductive rate of 3.5 to the carrier) gives a probability of less than 0.2 of detecting significantly (5% level) higher variance than expected under neutrality. If recombination is decreased, the probability of detecting selection rises, but even then positive selection on this single gene must be around 75% before it can be detected with any certainty (Figure 8).

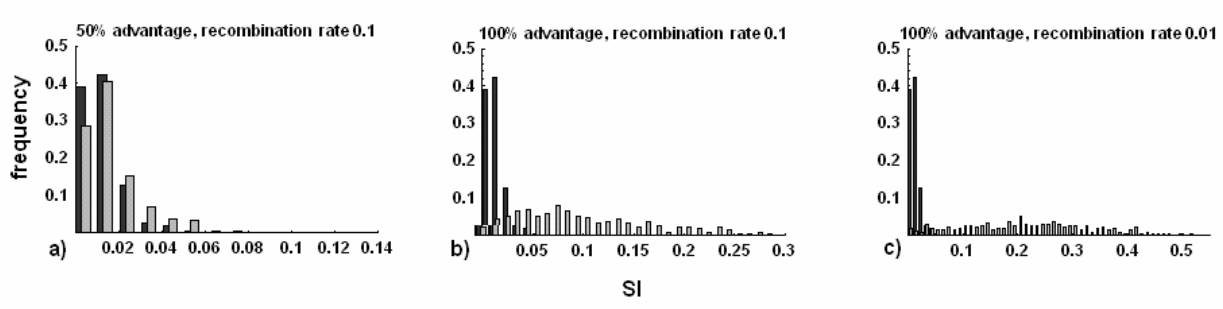


Figure 7: The pattern of the SI for 500 replicates of stochastic backcrossing for ten generations, in a survey of 30 alleles and the case where one hybridisation event is occurring per generation. Backcrossing under neutral conditions (black bars) and with a given positive selective advantage on a gene linked to one of the markers with a given recombination rate (grey bars).

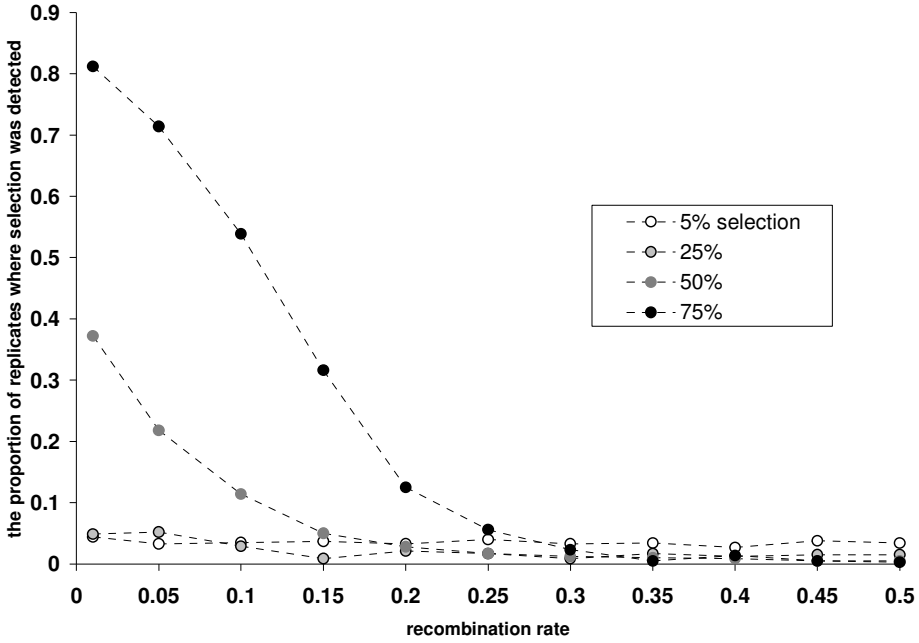


Figure 8: The proportion of replicates where selection was detected (at 95% level of confidence) in a survey of 30 loci where hybridisation occurs at a frequency of one event per generation for ten generations. In this case selection is modelled using one selectively advantageous gene linked to a single marker locus. Values based on finding significantly higher variance in introgression across marker loci than expected under neutrality. The 95% confidence interval was calculated from 10,000 replicates of stochastic backcrossing under neutrality. The probabilities were calculated from 1000 replicates. Plotted are the cases where the selectively advantageous gene conveys a 5, 25, 50 and 75% increase in mean fecundity on the individual carrying the gene.

3.4.7 Shared polymorphisms

So far we have investigated hybridisation using diagnostic markers. For various reasons the scenario in the previous section is highly idealised. Firstly, we can never be 100% certain of the diagnostic capabilities of the markers. Some of the alleles may be shared between the two populations, either because they have been persistent since population division, or because of subsequent mutations resulting in alleles of the same type. Secondly, we are unlikely to know the exact number of generations the population has been hybridising for. Under constant hybridisation the number of introgressed individuals will grow over time as

more hybridisation events add introgressed individuals to the population. This increasing number of independent introductions of foreign alleles into the population will decrease the variation in introgression across loci (lower SI). This means that the measure of SI depends on time since first hybridisation. In fact, providing hybridisation is constant, the number of individuals carrying multiple introgressed alleles will be constant over time. Recombination breaks up associations between alleles at a constant rate, so a balance between incoming and outgoing multiple introgressed genotypes will be struck and the proportions of individuals carrying different numbers of multiple introgressed alleles will be approximately constant. The ever-increasing number of introgressed individuals, are due to the growing number of individuals that carry a single introgressed allele.

The structure of a hybridising population can be thought of as having two components: a proportion of individuals with older hybrid ancestry (carrying a single allele) are in linkage equilibrium but, more recent hybrids (carrying multiple alleles) are in linkage disequilibrium. If hybridisation is occurring regularly, the proportion in linkage equilibrium increases over time. The proportion in linkage disequilibrium depends on the rate of hybridisation. So the rate of recent hybridisation could be estimated from the proportion of the population in linkage disequilibrium and the rate of past hybridisation from the proportion of the population in linkage equilibrium (provided one assumes in the latter case that the two populations were initially free of shared alleles and one knows time of first contact) (Goodman *et al.* 1999).

A time independent measure of SI can be obtained provided that only individuals carrying multiple introgressed alleles ($k \geq 2$) are considered, providing sufficient generations have elapsed to obtain equilibrium (around 6 in a survey of 30 unlinked markers). This approach has the added advantage that it eliminates a certain amount of uncertainty about ancestral polymorphisms and other shared alleles. Provided shared alleles are rare, the chances of an

individual carrying two or more of these is low. This approach has previously been adopted by Goodman *et al.* (1999).

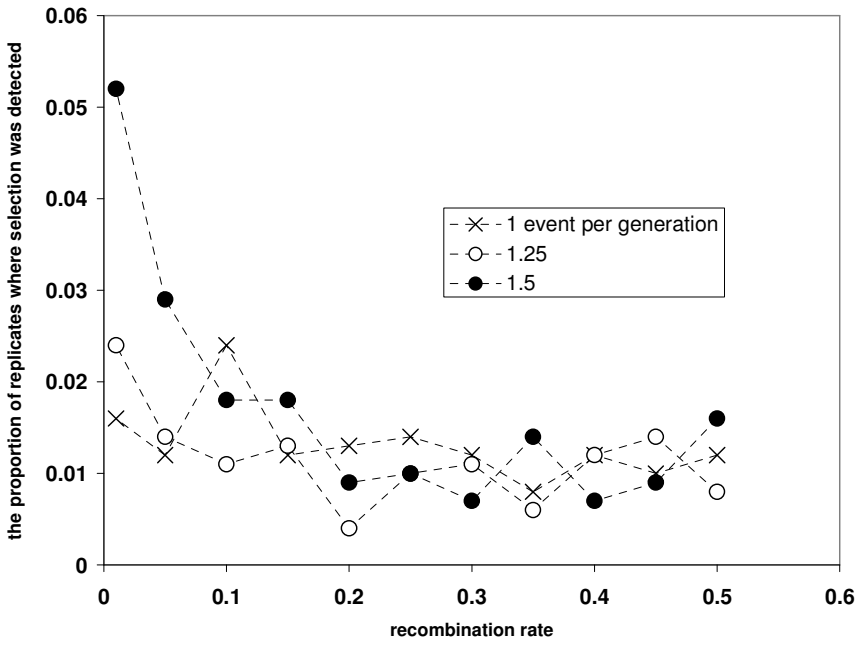


Figure 9: The proportion of replicates where selection was detected (at 95% level of confidence) in a survey of 30 loci in which only multiply introgressed individuals were included in the measure variance. Selection is modelled at one gene that is linked to one of the marker loci with a given recombination rate. In this case the strength of selection is 75%. Values are based on finding significantly higher variances across marker loci than expected under neutrality. The 95% confidence interval was calculated from 10,000 replicates of stochastic backcrossing under neutrality. The probabilities were calculated from 1000 replicates. Even under strong selection (75%) the chance of detecting selection is low. This chance increases slightly as hybridisation increases, but remains far too low to be of any use.

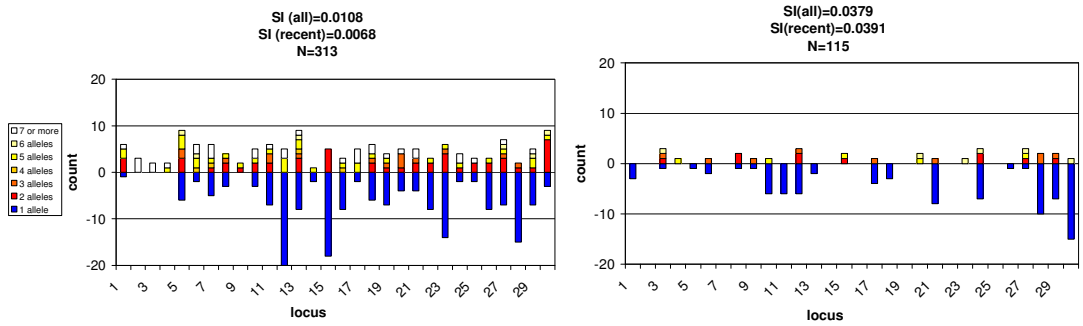
However, this measure of SI makes selection harder to detect because selection only has a few generations to take effect (Figure 9). Even if it is tightly linked to one of the molecular markers, a gene conveying a 75% selective advantage will not result in detectable selection at one hybridisation event per generation. An increase in hybridisation rate does improve the

chance of detecting selection but the probability of detection is still so low as to be irrelevant. There are two reasons why it is harder to detect selection than using the previous method: i) the number of individuals which are informative are fewer and ii) by definition, individuals are carrying multiple markers making selection on a gene associated with any one marker harder to detect - it is only possible to detect broad genomic effects.

This is not surprising, if only multiply introgressed individuals are included, then in effect, the signature of selection is being searched for in the first few generations of hybridisation ($k \approx \frac{1}{t}$, Figure 1), where linkage disequilibrium is high. In this case, one would expect selection to have to be very strong to leave a signature. If later generations (i.e. singly introgressed individuals) are included, then weaker selection can be detected, but the selectively advantageous genes must be tightly linked to the marker locus to survive many generations of recombination (Figure 10).

a)

b)



c)

d)

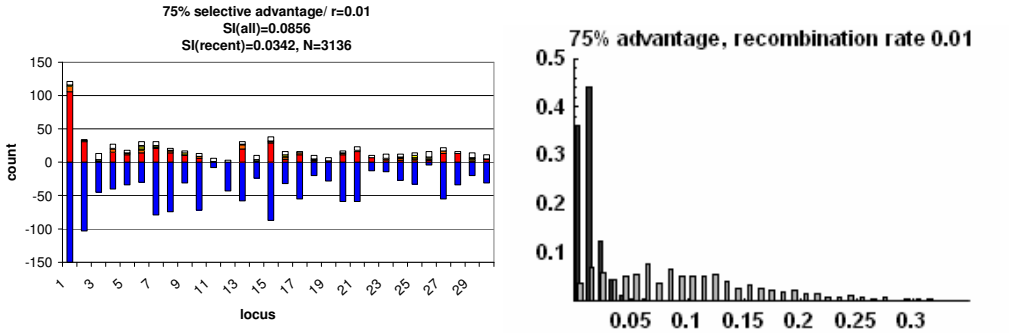


Figure 10: a & b) An illustration of the alleles in linkage equilibrium and linkage disequilibrium for two replicates of the neutral model (1 hybridisation event per generation). The total bar length (above and below the midline) represents the total number of introgressed alleles at each of 30 loci after ten generations. In blue are the number of cases where only one introgressed marker allele is present in the genome; in shades of orange, the cases where that allele is present with a number of others at different loci in the genome. These represent the proportion of the cohort in linkage disequilibrium. The measure of SI that only included recent individuals ignores the blue contribution to introgression. c) Introgression in a replicate in which a gene conveying a 75% selective advantage is linked to marker locus '1' with a recombination rate of 0.01. In this case the measure SI(all) is high and would be detectable as selection (see d). However, if SI is only measured across the LD component of the population, it is lowered greatly. In figure c) there are 1306 individuals with a single copy of marker locus '1', the locus linked to the selectively advantageous locus (i.e the blue bar goes off the scale).

3.5 Discussion

3.5.1 Summary of results

We have used a branching process to model scenarios where hybridisation is rare. A hybridisation event followed by some generations of backcrossing generates a highly unpredictable pattern of introgression at neutral molecular markers in the progeny. Variance in the degree of introgression across loci is high. If a number of such hybridisation events occur in a population of hybridising taxa, then the overall pattern of introgression is smoothed, and the relative variance across loci is reduced. In a theoretical scenario in which hybridisation is occurring at the rate of one event per generation and 30 diagnostic molecular markers are used to survey introgression, a selectively advantageous gene linked to one of the molecular markers with a recombination rate of 0.01, would need to confer a 75% increase in reproductive success to the carrier to be detected 80% of the time, with a 95% level of confidence after ten generations of backcrossing. In reality, however, the power to detect selection is even lower: it is hard to obtain large numbers of diagnostic markers, and once obtained difficult to be certain they are truly diagnostic. Generally, this issue will result in the need to eliminate individuals carrying only one introgressed allele from the analysis. Because selection must be very strong to be detectable in recent backcrosses (strong enough to be detected in the highly stochastic initial generations), this leads to a drastic reduction in the theoretical power of detecting selection.

3.5.2 Gene flow under rare hybridisation

The pattern of introgression generated by a single hybridisation event, followed for any number of generations of backcrossing, depends chiefly on the outcome of the events in the initial generations. In each reproductive event in the branching process, the passage of introgressing genes from one generation to the next is associated with a considerable amount

of stochasticity introduced through recombination and assortment during gametogenesis and through the variable reproductive success of the parent. In the initial generations, these reproductive events are few and only a few of the many possible outcomes of reproduction are played out. Providing introgression continues at all, potentially large numbers of reproductive events smooth reproductive stochasticity in later generations but, crucially, the outcome of these events has already been constrained by the outcome of random events in the early generations. The signature of stochasticity remains.

In a hybridising population, of course, we expect more than one hybridisation event. An increase in the number of hybridisation events smoothes out the stochasticity caused by the early generations of the branching process. This leads to a decrease in the relative variance in introgression across loci (Figure 6), making selection easier to detect. It should be emphasised here that it is not rate of hybridisation *per se*, but the total number of hybridisation events that have occurred in the population (rate x time) that is the important factor in smoothing the pattern overall pattern of introgression. Although of course, if the technique described here that relies on the pattern of linkage disequilibrium is used to detect selection then a high number of recent events must have occurred.

3.5.3 Numbers of markers

The large stochastic effect caused by the branching process can also be moderated by the use of increased numbers of marker loci (Figure 6). But increasing the markers has limited use because it is in the early generations where these markers are still present in the same genome, where chance has the biggest effect. This means that although the numbers of markers may be large, the number of independently acting portions of genome is far fewer. In any case, as the number of markers used is increased, a point is reached where they no longer recombine independently and a further increase in number will do little to remedy this problem. This point is reached when the number of markers is approximately equal to the

map length of the organism. This is similar to the problem encountered in QTL mapping of hybrid zones, where increasing the number of markers on a map does little to narrow the location of the QTL if the individuals surveyed have only undergone a few generations of recombination (Rieseberg & Buerkle 2002).

3.5.4 Types of markers

In this study we began by investigating hybridisation using wholly diagnostic markers. This serves to illustrate as a 'best case' scenario for detecting selection under rare hybridisation. A number of studies have relied on assuming that markers are truly diagnostic (Payseur *et al.* 2004; Szymura & Barton 1991), but generally it is hard to justify the absence of ancestral polymorphism in natural population. As we have seen, making allowances for ancestral polymorphism by excluding 'hybrids' in linkage equilibrium (Goodman *et al.* 1999), reduces the power to detect selection considerably. Methods that rely on the assumption of fixed allele frequency differences between populations (Rieseberg *et al.* 1999) or do not make assumptions about the diagnostic reliability of markers, but searched for patterns of F_{ST} will result in further loss of power, due to misclassification of hybrid individuals and a failure to 'observe' introgression events at some loci.

So far this study has ignored the additional loss of power that sampling from the population would add. Given these results, it seems unnecessary to investigate the further loss of power added by small sample sizes and cross generational sampling. It is clear from this study that even if whole genome information was available and the entire population was sampled, there comes a point where hybridisation is so rare, it is impossible to detect selection.

The important point this model emphasises is that even under neutral conditions, 'selection-like' patterns of introgression can be generated (Figure 10). For this reason, it is crucial to have some knowledge of the frequency of hybridisation events and age of hybridisation as no

method of surveying hybridising taxa is immune to this effect of stochasticity if hybridisation is rare.

3.5.5 How rare is rare?

The limitation to this model is that it only applies to scenarios where the conditions of introgression as a chain of successive backcrosses hold. In reality, as the number of hybridisation events increase, the probability of backcrossed individuals meeting rises and the conditions will eventually be broken. This will depend on the size of the hybridising population in question. For this reason, a shortcoming of this model is that it cannot determine a threshold above which detecting selection becomes feasible.

An alternative model of gene flow during hybridisation is that of the diffusion approximation (Nagylaki 1975) employed in cline analysis (Barton 1979; Barton & Gale 1993). Here mating between the two populations is assumed indiscriminate at the point of contact (individuals are equally likely to mate with any individual in the vicinity regardless of genotype). As the rate of hybridisation increases, the branching process will eventually converge on the diffusion model, but we have a poor idea of the intermediate scenarios.

Although hybridisation rates are quoted in some studies (Rieseberg *et al.* 1998), if they are known in others then, generally, little concern is given to them in the literature. An exception to this is the field of genetic risk assessment for transgenic crops (Hails & Morley 2005). As already stated, hybridisation can be rare but introgression high (Goodman *et al.* 1999) and if the rate of hybridisation and the history of a hybridising population are not known, this could lead to spurious reports of selection. In general, one might expect that 'risk' scenarios for stochastic patterns of introgression might be in hybridising taxa that are highly divergent and have recently come into contact – e.g. hybridisation between endemic and human translocated species.

It may, in fact, be possible to gain simple estimates for the number of hybridisation events in scenarios where hybridisation is relatively rare: if markers are highly polymorphic (for example Goodman *et al.* (1999)), then estimates could be devised based on the allelic variants present in the introgressed individuals compared to the pure population. The simplest assumption would be that each hybridisation event introduces an allele that is sampled as if independently from the pool in the parental population. If the parental population is highly polymorphic, then we can approximately expect each event to introduce a new allele. This provides a definite null hypothesis for the two cases of a single versus many hybridisation events contributing to the pattern of introgression. The number of introduced allelic variants would allow for rough estimates of the numbers of hybridisation events. Multiple polymorphic loci (microsatellites, SNP haplotypes) would provide independent tests.

3.5.6 Existing hybridisation studies in the light of these findings

Cline analysis

In studies of hybrid zones involving cline analysis such as the *Bombina* and *Mus* hybrid zones (Payseur *et al.* 2004; Szymura & Barton 1991), there is an implicit assumption that many hybridisation events have contributed to the zone: the structure of the cline would not be present if hybridisation was not common. Selectively advantageous alleles are expected to penetrate furthest across the hybrid zone and in the tail ends of the cline they will be in linkage equilibrium, which may lead to uncertainty about ancestral polymorphism (see above). However, denser maps might lead to greater certainty on allele origin as, even in the tails of the cline, introgressed regions carrying multiple markers might be present.

Frequency in introgressed markers and patterns of FST

The hybridisation rate in the *Helianthus annuus* x *H. petiolaris* hybrid zones is 0.04- 0.15 (Rieseberg *et al.* 1998) with around 5.6% pollen viability in the F1 (Ungerer *et al.* 1998) . The hybrid zones are thought to have originated in the early 1900's (Rieseberg *et al.* 1999). It is unclear whether this represents a high or low number of hybridisation events, but the fact that the patterns of introgression were consistent across three transects seems to imply that hybridisation is frequent enough for selection to be detectable (Rieseberg *et al.* 1999).

Experimental crosses are a good way of generating large numbers of hydisation events, such as in studies using mapped linked markers (Iris retroelement transposon display marker system) on hybridising species of Louisiana Irises to detect adaptive trait introgression (Bouck *et al.* 2005; Martin *et al.* 2006). The generation of a large number of crosses is also the reason for the success of Linkage Group Selection, a technique used in malaria research to map drug resistant phenotypes (Culleton *et al.* 2005; Martinelli *et al.* 2005). Resistant and non-resistant strains are crossed, which create thousands of independent recombinants (Culleton *et al.* 2005), and the progeny are grown for a number of backcross generations under the selective pressure of the drug. Genotyping the progeny at mapped linked markers leads to the identification of candidate genes for resistance. However, experimental crosses are most suited to investigate taxa that can be kept in laboratories or greenhouses and that have short generation times because otherwise it will not be possible to produce sufficient generations for fine scale QTL analysis.

Studies that have analysed the patterns of F_{st} in search of selected genes have involved studies of human populations (Akey *et al.* 2002) and differentiation in a cline of morphs of *Littorina saxalitis* (Wilding *et al.* 2001) which appear to represent a primary contact zone. Gene flow is high and ancient in both cases. Because gene flow is extensive, the large

number of markers involved in these surveys (especially Akey *et al.* (2002)), mean that the power to detect selection is high.

Discordant genealogies

This work clearly illustrates the point that there is high stochasticity in the fate of individual lineages. At any given locus, a genealogy discordant with the general pattern of genealogies may indicate that a hybridisation event has taken place in its history (although ancestral polymorphism is suspected if coalescence predates species divergence). However any single past hybridisation event might result in none or a number of introgressed loci in any one sampled individual (depending on chance and selection). Generally, however, if a single hybridisation event results in introgression, it will result in only a few short blocks of the introgressing genome persisting over time (Baird *et al.* 2003). Thus, the greater the number of historic hybridisation events, the greater the extent to which the genealogies can be expected to be randomised.

At any one locus a discordant genealogy may give confirmation of hybridisation, but even the fixation of introgressed genes is not, of course, proof of selection and could simply be a result of neutral introgression.

3.5.7 Conclusion

In this article we have used a branching process model to investigate rare hybridisation. The results of this study reveal that: (1) the pattern of introgression even of a large numbers of genetic markers will be highly stochastic if only a few hybridisation events have contributed to the hybridising population. (2) This pattern of neutral variation in introgression can have high enough variance that it could be mistaken for selection. (3) Therefore, even if strong selection is acting, it may not be possible to distinguish its effects from neutral variation.

Future work needs to involve modelling the transition between scenarios of rare and frequent hybridisation.

3.6 Appendix 1

For a mathematical explanation of generating functions see Grimmett & Stirzaker (2001) Section 5.1. Generating functions are an elegant mathematical tool for finding the distribution of numbers of copies up to around ten generations, but become inefficient after that. The probability function for the Poisson distribution for the number of introgressing alleles passed on Pj is associated with a generating function P[z]. This is defined as the expectation of zj:

$$\tilde{P}[z] = \sum_{j=0}^{\infty} e^{-\lambda} \frac{(\lambda)^j}{j!} z^j = e^{-(\lambda)(1-z)}$$

Conveniently, to obtain the distribution of alleles after t generations, φ_t, one simply has to apply the generating function to the (t-1) th generation.

$$\varphi_t = e^{-(\lambda)(1-z)}$$

$$\varphi_{t+1} = e^{-(\lambda)(1-(e^{-(\lambda)(1-z)}))}$$

…

The derivative at zero (z=0) gives the probability distribution:

$$\varphi(0) = P(j=0) \qquad \varphi'(0) = P(j=1) \qquad \varphi''(0) = 2!P(j=2) \qquad \varphi^{(h)}(0) = h!P(j=h)$$

Derivatives of the generating function at 1 (z=1) can be used to obtain the mean and higher moments of the distribution:

$$\varphi(1) = 1 \qquad \varphi'(1) = E[j] \qquad \varphi''(1) = E[j(j-1)] \qquad \varphi^{(h)}(1) = E[j..(j-h+1)]$$

3.7 Appendix 2

For multiple loci it is necessary to gain the distribution of the vector of genotype numbers, $\underline{n}$. The generation function depends on the dummy variable $\underline{z}$. We choose not to track $n_{000\ldots}$, because these will not be distinguishable. Let $\lambda_{X,Y}$ be the expected number of offspring of type Y from type X:

$$P_{t+1}[\underline{z}] = \tilde{P}_t[Exp[-\sum_Y \lambda_{X,Y}(1-z_Y)]]$$

$$\tilde{P}_0[\underline{z}] = z_{111\ldots}$$

In this recursion, z_X is replaced by $Exp[-\sum_Y \lambda_{X,Y}(1-z_Y)]$ in the argument of the generating function. The simplest case is no linkage, and a growth rate ϕ which would be 2 for steady state. Then, with n loci, the expected # of any of the 2^n genotypes is $\lambda_{X,Y} 2^{-n}\phi$. For example, with 2 loci:

$$P_{t+1}[\underline{z}] = \tilde{P}_t[Exp[-\sum_Y \lambda_{X,Y}(1-z_Y)]]$$

$$\tilde{P}_0[\underline{z}] = z_{11}$$

$$\tilde{P}_1[z] = Exp[-\frac{\phi}{4}(3-z_{11}-z_{01}-z_{10})]$$

$$\tilde{P}_2[z] = Exp[-\frac{\phi}{4}(3-Exp[-\frac{\phi}{4}(3-z_{11}-z_{01}-z_{10})]-Exp[-\frac{\phi}{2}(1-z_{01})]-Exp[-\frac{\phi}{2}(1-z_{10})])]$$

$$\tilde{P}_3[z] = Exp[-\frac{\phi}{4}(3-Exp[-\frac{\phi}{4}(3-Exp[-\frac{\phi}{4}(3-z_{11}-z_{01}-z_{10})]-Exp[-\frac{\phi}{2}(1-z_{01})]-Exp[-\frac{\phi}{2}(1-z_{10})])]-$$
$$Exp[-\frac{\phi}{2}(1-Exp[-\frac{\phi}{2}(1-z_{01})])]-Exp[-\frac{\phi}{2}(1-Exp[-\frac{\phi}{2}(1-z_{10})])])]$$

This gives the joint density of $\{z_{11}, z_{01}, z_{10}\}$ after three generations.

3.8 References

Akey JM, Zhang G, Zhang K, Jin L, Shriver MD (2002) Interrogating a high-density SNP map for signatures of natural selection. *Genome Research* **12**, 1805-1814.

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

Baird SJE, Barton NH, Etheridge AM (2003) The distribution of surviving blocks of an ancestral genome. *Theoretical Population Biology* **64**, 451-471.

Barton NH, Gale KS (1993) Genetic analysis of hybrid zones. In: *Hybrid zones and the evolutionary process* (ed. Harrison RG). Oxford University Press, Oxford.

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

Boecklen WJ, Howard DJ (1997) Genetic analysis of hybrid zones: Numbers of markers and power of resolution. *Ecology* **78**, 2611-2616.

Bouck A, Peeler R, Arnold ML, Wessler SR (2005) Genetic mapping of species boundaries in Louisiana irises using IRRE retrotransposon display markers. *Genetics* **171**, 1289-1303.

Corander J, Marttinen P (2006) Bayesian identification of admixture events using multilocus molecular markers. *Molecular Ecology* **15**, 2833-2843.

Culleton R, Martinelli A, Hunt P, Carter R (2005) Linkage group selection: Rapid gene discovery in malaria parasites. *Genome Research* **15**, 92-97.

Endler JA (1977) *Geographic Variation, Speciation, and Clines* Princeton Univ. Press. Princeton.

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.

Goodman SJ, Barton NH, Swanson G, Abernethy K, Pemberton JM (1999) Introgression through rare hybridisation: a genetic study of a hybrid zone between red and sika deer (genus *Cervus*), in Argyll, Scotland. *Genetics* **152**, 355-371.

Grimmett GR, Stirzaker DR (2001) *Probability and Random Processes*, 3rd edn. Oxford University Press.

Hails RS, Morley K (2005) Genes invading new populations: a risk assessment perspective. *Trends in Ecology & Evolution* **20**, 245-252.

Haldane JBS (1948) The theory of a cline. *J. Genet.* **48**, 277-284.

Kliman RM, Andolfatto P, Coyne JA*, et al.* (2000) The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics* **156**, 1913-1931.

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics* **74**, 175-195.

Machado CA, Kliman RM, Markert JA, Hey J (2002) Inferring the history of speciation from multilocus DNA sequence data: The case of *Drosophila pseudoobscura* and close relatives. *Molecular Biology and Evolution* **19**, 472-488.

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

Manel S, Gaggiotti OE, Waples RS (2005) Assignment methods: matching biological questions techniques with appropriate. *Trends in Ecology & Evolution* **20**, 136-142.

Martin NH, Bouck AC, Arnold ML (2006) Detecting adaptive trait introgression between *Iris fulva* and *I. brevicaulis* in highly selective field conditions. *Genetics* **172**, 2481-2489.

Martinelli A, Cheesman S, Hunt P, *et al.* (2005) A genetic approach to the de novo identification of targets of strain-specific immunity in malaria parasites. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 814-819.

Nagylaki T (1975) Conditions for the existence of clines. *Genetics* **80**, 595-615.

Patterson N, Richter DJ, Gnerre S, Lander ES, Reich D (2006) Genetic evidence for complex speciation of humans and chimpanzees. *Nature* **441**, 1103-1108.

Payseur BA, Krenz JG, Nachman MW (2004) Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution* **58**, 2064-2078.

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

Randler C (2002) Avian hybridization, mixed pairing and female choice. *Animal Behaviour* **63**, 103-119.

Rieseberg LH, Baird SJE, Desrochers AM (1998) Patterns of mating in wild sunflower hybrid zones. *Evolution* **52**, 713-726.

Rieseberg LH, Buerkle CA (2002) Genetic mapping in hybrid zones. *American Naturalist* **159**, S36-S50.

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

Rieseberg LH, Wood TE, Baack EJ (2006) The nature of plant species. *Nature* **440**, 524-527.

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.

Seehausen O (2004) Hybridization and adaptive radiation. *Trends In Ecology & Evolution* **19**, 198-207.

Szymura JM, Barton NH (1991) The genetic structure of the hybrid zone between the fire-bellied toads Bombina bombina and B. variegata: comparisons between transects and between loci. *Evolution* **45**, 237-261.

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.

Wilding CS, Butlin RK, Grahame J (2001) Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers. *Journal of Evolutionary Biology* **14**, 611-619.

Chapter 4

Investigating temporal changes in hybridisation and introgression in a predominantly bimodal hybridising population of invasive sika (Cervus nippon) and native red deer (Cervus elaphus) on the Kintyre Peninsula, Scotland.

H.V. Senn[1,2], N. H. Barton[3], S. J. Goodman[4], G. M. Swanson[5], K.A. Abernethy[6], J. M. Pemberton[1]

[1] Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JT, UK

[2] Macaulay Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK

[3]Institute of Science and Technology, Am Campus 1, Klosterneuburg, Austria A3400

[4]The Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK

[5] Christchuch, NZ

[6] School of Biological and Environmental Sciences, University of Stirling, Stirling, FK9 4LA, UK

Authors' Contributions:

HVS, GMS and KAA organised sampling. HVS and SJG carried out laboratory work. HVS wrote the manuscript and performed the analyses with contribution from NHB. JMP and NHB commented on multiple drafts of the manuscript. KAA also commented on the manuscript.

4.1 Abstract

Temporal studies of hybridisation between native and invasive species, being examples of species hybridising under recent secondary contact, provide a unique opportunity to investigate the processes involved in hybrid zone formation. Here we investigate the dynamic of hybridisation between native red deer (*Cervus elaphus*) and invasive Japanese sika (*Cervus nippon*) on the Kintyre Peninsula, Scotland over a 15 year period. Through analysis of a dataset of 1513 samples of deer at 20 microsatellite loci and a mtDNA marker, we investigated the extent to which the genetic structure of the hybridising populations had changed over time. We found no evidence that either the proportion of recent hybrids, or the levels of introgression had changed over the time of the study. Nevertheless, in one population where two species have been in contact since ~ 1970, 44% of individuals sampled during the study were hybrids. This suggests that hybridisation between these species can proceed fairly rapidly. Via analysis of the number of alleles that have introgressed from polymorphic red deer into the genetically homogenous sika population we reconstructed the haplotypes of red deer alleles introduced by backcrossing. For a large section of the Kintyre Peninsula it can be concluded that as few as 5 separate hybridisation events could reasonably account for the recently hybridised sika-like individuals in the dataset. We discuss how the low number of hybridisation events between red and sika deer might be shaping temporal and spatial patterns of hybridisation.

4.2 Introduction

When two populations meet and hybridise we can expect a variety of outcomes: hybrid speciation aside, the two populations might remain genetically distinct, they might merge or they might form stable clinal or mosaic hybrid zones. The theoretical framework for describing gene flow under hybridisation is well established (Barton & Gale 1993; Haldane 1948; Slatkin 1973): gene flow between the two populations is facilitated by dispersal and hybridisation, and is stemmed by selection that maintains divergence.

Precisely because of their age, ancient hybrid zones, in which two species have been in contact for a long time, are thought to be stable. For example, the chromosomal races of the grasshopper *Podisma pedestris* have been in contact along a ridge in the Alpes Maritimes, France since ice sheet retreat 5000 - 10000 years ago (Barton & Hewitt 1981). The 800m wide cline is maintained by endogenous selection against hybrids and is held in position by low population density, caused by inhospitable habitat along the ridge. There are numerous examples of postglacial hybrid zones (e.g. Butlin & Hewitt (1985), Hunt & Selander (1973) and Szymura & Barton (1986)), with diverse selective pressures maintaining them. It is interesting, however, to consider how many other pairs of populations have met and hybridised (e.g. following glacial retreat on mountain passes) but have not formed hybrid zones that have remained stable until the present day: ancient hybrid zones show us a snapshot of just one outcome of the meeting of two species.

In contrast to ancient hybrid zones, hybridisation between introduced and native species provides us with the opportunity to observe gene flow between two populations at an early stage. Here, it is not yet clear whether the hybridising populations will stabilise and form a hybrid zone, whether the two populations will merge, or indeed, whether they will be driven

apart by reinforcement. If hybridisation between the two populations is rare and has been occurring for a short time, we can expect the geographic pattern of introgression between the two populations to be highly stochastic (Chapter 2). On the other hand, if hybrids form freely and the organisms involved are highly mobile we might expect a rapid merging of the two populations, as with hybridisation between Africanised and European honey bees (*Apis mellifera*) in the south-western United States (Pinto *et al.* 2005).

The dynamics of gene flow between hybridising populations are interesting from both a practical and a theoretical standpoint. Hybridisation between non-native and endemic species is often of conservation concern, due to fears over loss of species integrity (e.g. in European wild cats (Randi *et al.* 2001)) or introgression of ecologically undesirable traits (e.g. from cultivated crops to their wild relatives (Sorensen *et al.* 2007)). Hybridisation of non-native and endemic species also provide us with unequivocal examples of secondary contact and allows us to investigate the early dynamics of species contact en route to any of the possible outcomes mentioned above.

A number of approaches have been taken to investigate gene flow between hybridising populations. In samples taken at a single point in time, movement of hybrid zones have been inferred from asymmetric introgression of markers: a moving hybrid zone should leave a tail of introgressed alleles in its wake. One example of this is a cottonwood hybrid zone situated on an altitudinal gradient in Utah, USA, in which *Populus angustifolia* is found above *P. fremontii* (Martinsen *et al.* 2001). Here the presence of introgressed *P. fremontii* mtDNA at high altitude and a trend of increasingly advanced backcrossed *P. angustifolia* with increasing altitude (but not *vice versa*), have been interpreted as evidence that the *P. angustifolia* population and the hybrid zone is gradually descending to lower altitude (Martinsen *et al.* 2001). A lizard hybrid zone (*Scelopurus cowlesi* x *S. tristichus*) that has moved over a period of 30 years does indeed show an asymmetric pattern of mtDNA

introgression (Leache & Cole 2007). But, generally, it is difficult to infer dynamic processes from samples taken at a single time point: asymmetrical patterns of introgression could also be due to positive selection on introgressing alleles, ancestral polymorphism or long-distance migration events of individuals of one species into the range of the other (extensively reviewed in Buggs (2007)).

Gene flow dynamics and the stability of hybridising populations have also been investigated through longitudinal studies, but few such studies exist (reviewed in Buggs (2007) and Strayer *et al.* (2006)). Some noteworthy exceptions are an 18 year biannual survey of morphological introgression in salamanders (Hairston *et al.* 1992), a 14 year annual study of allozyme variation in a ground cricket hybrid zone (Britch *et al.* 2001), a nine year annual survey of invasive crayfish hybridisation (Perry *et al.* 2001), an 11 year annual study of honeybee Africanization (Pinto *et al.* 2004), a 36 year long pedigree-based study of baboon hybridisation (Tung *et al.* 2008) and a study of an *Anartia* hybrid zone which was sampled 4 times in a 23 year period (Dasmahapatra *et al*. 2002). Additionally, some studies have resurveyed hybrid zones after a number of years (Blum 2002) or made comparisons with historical data (Szymura & Barton 1991).

Here we present a study spanning 10 to 15 years of the native red deer (*Cervus elaphus*) and introduced Japanese sika (*C. nippon*) on the Kintyre Peninsula in Scotland. The history and natural history of this system have been described in detail elsewhere (Abernethy 1994; Goodman *et al.* 1999; Chapter 2). In brief, Japanese sika were introduced near the south end of the peninsula at Carradale Estate in 1893 (Figure 1, see later). Then the population expanded and now sika range over a large part of the peninsula (from South Kintyre to West Loch Awe, Figure 1); sika (mostly males) are sighted occasionally across the rest of the study area (Forestry Commission Scotland personal communication). Hybridisation appears to be rare, estimated at ~1:500-1000 mating events (Goodman *et al.* 1999) and is followed

mainly by backcrossing into the parental populations (Goodman *et al.* 1999; Chapter 2). Although introgression is low on average (6.9%), at one site, West Loch Awe (Site 8, Figure 1), hybridisation is extensive (Chapter 2). Here 43% of individuals are hybrids and mating appears to be random (Chapter 2). Across the study area, mtDNA introgression is mostly from red into sika (97%), implying that hybridisation events occur between red deer females and sika males. Senn and Pemberton (Chapter 2) argued that because of the rarity of hybridisation events and the short time (<120 years) that the two species have been in contact, the total number of hybridisation events that have taken place on the peninsula is likely to be low. Thus, the spatial distribution of introgression, within the areas of overlap between the two species, is determined by the chance location of rare F1 hybridisation events. They argue additionally, that the chance build up of hybrids, perhaps due to a number of F1 events coinciding in space and time, is the most likely explanation for the hybrid swarm found at West Loch Awe. A localised build up of hybrids could lead to breakdown in assortative mating, which could kick-start a positive feedback mechanism resulting in the collapse of the two populations into a hybrid swarm.

When hybridisation is rare, introgression proceeds via backcrossing because F1s only have parental-type individuals to mate with (Goodman *et al.* 1999; Chapter 2). Providing backcrossed individuals only mate with individuals from the parental species that they are genetically closest to, then in each subsequent backcross generation the introgressed portion of the genome will be reduced by half. If we measure hybridisation at 20 loci, then after 4 generations of backcrossing we expect individuals to carry 1.25 introgressed alleles on average. Each new F1 event will introduce a set of introgressing alleles into the population, which will be dispersed by recombination over successive generations. As hybridisation continues, we expect to see a small fraction of alleles that derive from recent F1s, and so are still associated with each other, in strong linkage disequilibrium (LD). In addition, there will be an increasing frequency of introgressed alleles that are dispersed through the gene pool,

close to linkage equilibrium (LE). Given the phenotypic differences between red and sika, we expect selection to act on hybrids. The F1 and early backcrosses might gain fitness through heterosis (Whitlock *et al.* 2000), or lose fitness through genetic incompatibility. As mentioned above, if sufficient hybridisation events occur in the same area, assortative mating might break down, leading to collapse into a hybrid swarm. But even if the two taxa remain distinct, small blocks of genome will introgress independently; weaker selection may then act for or against individual alleles: in the long run, some may be fixed, while others may be eliminated.

In this study we assess gene flow over time between red deer and sika, through analysis of three sample sets collected in 1991/2 (Abernethy (1994) and Goodman *et al.* (1999), data reanalysed for this study), 1996/7 (previously unpublished) and 2006/7 (Chapter 2). We use nuclear and mtDNA markers to follow changes in the genetic structure of the hybridising populations. We also attempt to estimate the number of F1 hybrids that have contributed to recent introgression in sika through examination of the alleles that have introgressed from the highly polymorphic red deer population (it is not possible to do this the other way round because sika on Kintyre are genetically homogenous). By examining the haplotypes of sika likely to be from recent backcross generations we were able to reconstruct the probable haplotypes of the F1 hybrids that contributed to recent introgression. A study of gene flow in this system should shed light on the dynamics of hybridisation between populations that have recently come into contact, and should also provide us with a better understanding of the long term implications of hybridisation between Japanese sika and Scottish red deer.

4.3 Methods

4.3.1 Samples

The dataset consisted of 1513 samples from deer collected at three time intervals spanning 15 years, during the cull seasons 1991-2, 1996-7 and 2006-7. The study area covered the Kintyre and Cowal Peninsulas and stretched to Fort William in the north and Crianlarich in the northeast (Figure 1). Samples were collected from 27 commercial forestry sites by Forestry Commission Scotland rangers. Not all sites were sampled in each of the sampling episodes: in total 1276 samples came from sites that were sampled more than once, with sample sizes >20 individuals (Figure 1). Throughout, the purpose of the culls was to reduce overall deer numbers. Rangers were instructed to sample both species and no limit was placed on the number of a given sex or species gathered from each site.

1991-2 dataset:

This dataset consisted of 124 samples from deer (53 male, 66 female and 5 unknown) collected from 5 sites (Figure 1). These are a subset of the dataset previously analysed by Abernethy *et al.* (1994) and Goodman *et al.* (1999). Not all samples in the 1991-2 dataset could be reanalysed due to lack of time, and so samples from sites from which large sample sizes were obtained were chosen for analysis. The sample sizes from these sites are not the same as in the original papers because some samples were missing or were excluded due to poor quality of DNA extractions. Samples consisted of kidney tissue frozen to -70^{o}C and subsequently stored at -30^{o}C for fifteen years prior to extraction. DNA was extracted with the DNeasy Blood and Tissue Kit (QIAGEN) according to the manufacturer's instructions.

1996-7 dataset:

This dataset consisted of 654 samples from deer (84 male, 563 female and 7 unknown) collected from 21 sites (Figure 1). This sample set is previously unpublished. Tissue samples consisted of tongue tips frozen at -20^{o}C. DNA was extracted using either standard proteinase-K phenol-chloroform (Sambrook *et al.* 1989) or chelex (Walsh *et al.* 1991) protocols.

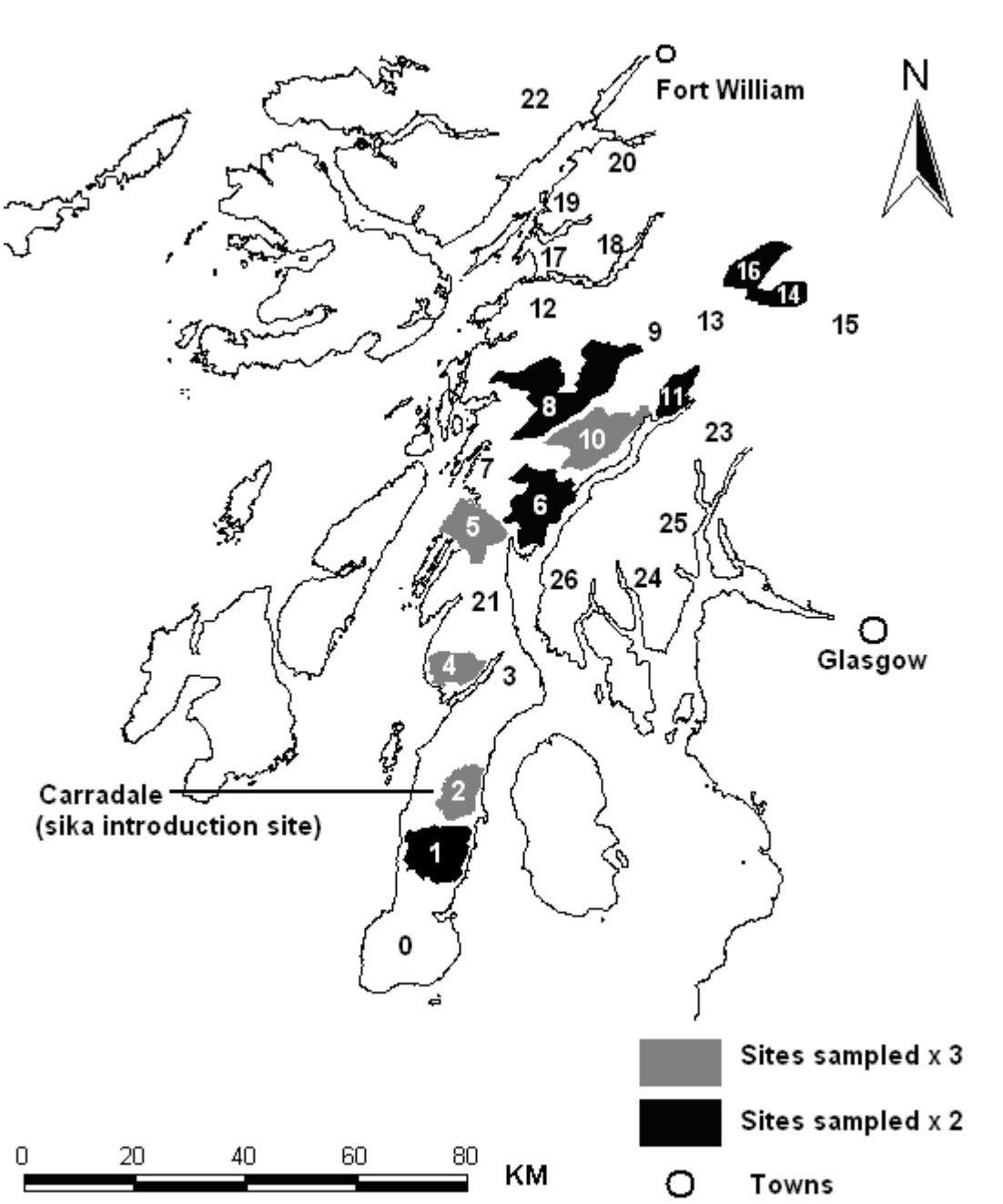


Figure 1: Map of the study area with sampling sites shown and numbered (0-26). For the sites shaded in grey, adequate sample sizes (>20 individuals) were collected in each of the three sampling periods (1991/2, 1996/7 & 2006/7). For the sites shaded in black, adequate sample sizes (>20 individuals) were collected for the sampling periods 1996/7 and 2006/7. The locations of remaining sites are numbered; here samples were only collected in one year or in multiple years but with small sample sizes. The names of the sample sites are: 0 (South Kintyre), 1(Lussa), 2 (Carradale), 3 (South Tarbert), 4 (Achaglachach), 5 (Knapdale), 6 (Kilmichael), 7(Ormaig), 8 (West Loch Awe), 9 (Collaig), 10 (Eredine/Birdfield), 11 (Shira), 12 (Oban), 13 (Succoth), 14 (Glen Lochy), 15 (Ben More), 16 (Glen Orchy), 17 (Barcaldine), 18 (Barrs), 19 (Appin), 20 (South Ballachulish), 21 (North Tarbert), 22 (Morvern), 23 (North Cowal), 24 (South Cowal), 25 (East Cowal), 26 (Glendaruel). In Goodman *et al.* (1999) sites 0, 2, 4, 5, 6 & 10, have the same names but are numbered differently.

2006-7 dataset:

This dataset consisted of 735 samples (312 female, 423 male) collected from 21 sites (Figure 1). This sample set is previously published (Chapter 2). Tissue samples consisted of an ear tip that was preserved directly in 100% ethanol. DNA was extracted with the DNeasy Blood and Tissue Kit (QIAGEN) according to the manufacturer's instructions.

4.3.2 Nuclear genotyping

All 1513 samples were genotyped at 20 nuclear loci (AGLA293, BM4006, BM6438, BM757, BOVIRBP, FCB193, IDVGA29, IDVGA55, INRA5, INRA6, INRA131, MM12, RM12, RM188, RM95, TGLA40, TGLA126, TGLA127, TGLA337 & UWCA47). The loci were selected for having no shared alleles in a test panel of 44 red and 44 sika of diverse geographic origins (Goodman *et al.* 1999). Of the 20 loci, 16 are mapped to red deer linkage groups (Slate *et al.* 2002) and these markers are either situated on separate linkage groups or at opposite ends of the same linkage group (47.7 -75.3 cM). Full details of the loci can be found in Chapter 2.

Genotyping of the 1991-2 and 2006-7 datasets was carried out using a multiplex PCR protocol followed by capillary array sequencing as detailed in Chapter 2. The 1996-7 dataset was genotyped using uniplex PCR protocols followed by polyacrylamide slab gel electrophoresis and autoradiography as described by Slate *et al.* (1998) and Goodman *et al.* (1999).

For information on how we verified that the different analysis methods used in these datasets were comparable see Supplementary Material, Section 4.6.1.

4.3.3 MtDNA genotyping

A mitochondrial DNA haplotype was assigned for all samples in the datasets using a 39-bp tandem repeat in the mitochondrial control region (Cook 1993). Red deer have a single repeat, whereas sika have multiple repeats. Length variation in repeats was assayed on 4% agarose gel stained with ethidum bromide (Goodman *et al.* 1999; Chapter 2). All samples from the 1991/2 dataset (Abernethy 1994; Goodman *et al.* 1999) were also rerun for the mtDNA marker for this study.

4.3.4 Assignment of genetic hybrid scores

Individuals were assigned an admixture proportion (Q) based on their microsatellite genotype using a Bayesian population clustering algorithm implemented in the software STRUCTURE 2.2 (Falush *et al.* 2007). STRUCTURE's model assumes that populations, characterised by sets of allele frequencies at each locus, are in Hardy-Weinberg and linkage equilibrium. Admixture is modelled by assuming each individual inherits some fraction of its genome from each population and is a priori equally likely to belong to any population. Given a specified number of populations, STRUCTURE uses a Markov Chain Monte Carlo (MCMC) algorithm to cluster individuals in the data set into the most likely arrangement under assumptions of the model. All individuals from each of the three data sets were included in one analysis. Individuals from sample sites that were only sampled at one time point were included in order to add power to the analysis. The most likely numbers of populations in the data set (K) was estimated by conducting 5 independent replicates of K=1-5. The model was run using a burn-in of $5x10^4$ and 10^6 Markov chain Monte Carlo (MCMC) steps, under the standard assumptions of the admixed ancestry model (with the parameter α inferred from the data, using a uniform prior) and the model of correlated allele frequency (λ=1) (see Pritchard *et al*. 2007). The presence of null alleles was estimated simultaneously (new to STRUCTURE 2.2) as in Chapter 2.

4.4 Analyses and results

4.4.1 Genetic diversity and assignment

Genetic diversity indices for the entire dataset are summarised in Table 1 of the Supplementary Material, Section 4.6.2. The number of alleles per locus ranged from 2 to 15 with a mean of 8.2. Mean allelic diversity was significantly lower in sika (3.05) than in red (7.45) (two-tailed Wilcoxon's signed rank test=150, $p<0.001$).

The results of the replicate STRUCTURE 2.2 simulations were highly consistent at each value of K. Division of the data into two clusters (K=2) captured the greatest proportion of the data structure (see Supplementary Material Figure 1 for further details). Here the estimate of an individual's proportion of ancestry obtained by STRUCTURE will be referred to in terms of proportion of red deer alleles (1=red, 0=sika) and will be called Q.

Individuals in the range $0.05 \leq Q \leq 0.95$ were considered to be recent hybrids (as in Chapter 2). Additionally individuals in the range $0.01 \leq Q < 0.05$ and $0.95 < Q \leq 0.99$ were considered to be sika-like and red-like animals and are referred to as "distant hybrids" throughout this paper.

The markers used in this study were highly differentiated and the majority of alleles could be identified as coming from red or sika populations (Supplementary Material Figure 2). For some analyses we assigned alleles to a population of origin as in Goodman *et al.* (1999). Red alleles were defined subjectively as any allele that was estimated as being present in the red parental populations but not in sika by the STRUCTURE analysis, or that was estimated as being at higher frequency in red than in sika by the STRUCTURE analysis and where the size of the allele fell within the size range of the other red alleles (and vice versa for sika alleles). Allele frequency distributions and designations can be found in Figure 2,

Supplementary Material. This system of designation allows for the calculation of a classic hybrid index [number of introgressed alleles / 2(number loci scored)]. When hybrid scores were compared to STRUCTURE's Q, the agreement was generally good, apart from some disagreement close to Q=0 and Q=1 (Supplementary Material Figure 3).

This disagreement is to be expected and before continuing it is worth considering why we find difference between STRUCTURE's Q (here called "hybrid score") and a classic hybrid index using diagnostic loci. A classic hybrid index relies on the accurate identification of alleles as coming from each population. If some doubt surrounding allelic origin exists, then to reduce the error in assignment of hybrids, only individuals with multiple apparently introgressed alleles can be considered as hybrids, providing hybridisation is rare (Goodman *et al.* 1999). However, in STRUCTURE, alleles are in effect weighted according to the extent to which they are in LD across the populations - if an individual carries an apparently introgressed allele that is in high LD across the population clusters, then it will be assigned a Q value that reflects a higher proportion of ancestry of the other species than if it carried an allele that is in lower LD. Low LD alleles might be alleles that are actually ancestrally shared polymorphisms or they might be alleles that have introgressed from one population into the other a number of generations ago. This accounts for some of the discrepancy we see between Q and hybrid score at values close to one and zero. Throughout this paper we define 'recent hybrids' as individuals with recent hybrid ancestry; these individuals can be distinguished from the background population because they are in LD (at the loci we are observing). In other words, they will have intermediate Q values and will be carrying a number of alleles derived from both populations. In practice, since LD is a continuous measure, for the majority of this paper we interpret these individuals in LD to have $0.05<Q<0.95$; although for some analyses we define then as carrying 4 or more apparently introgressed alleles (see later).

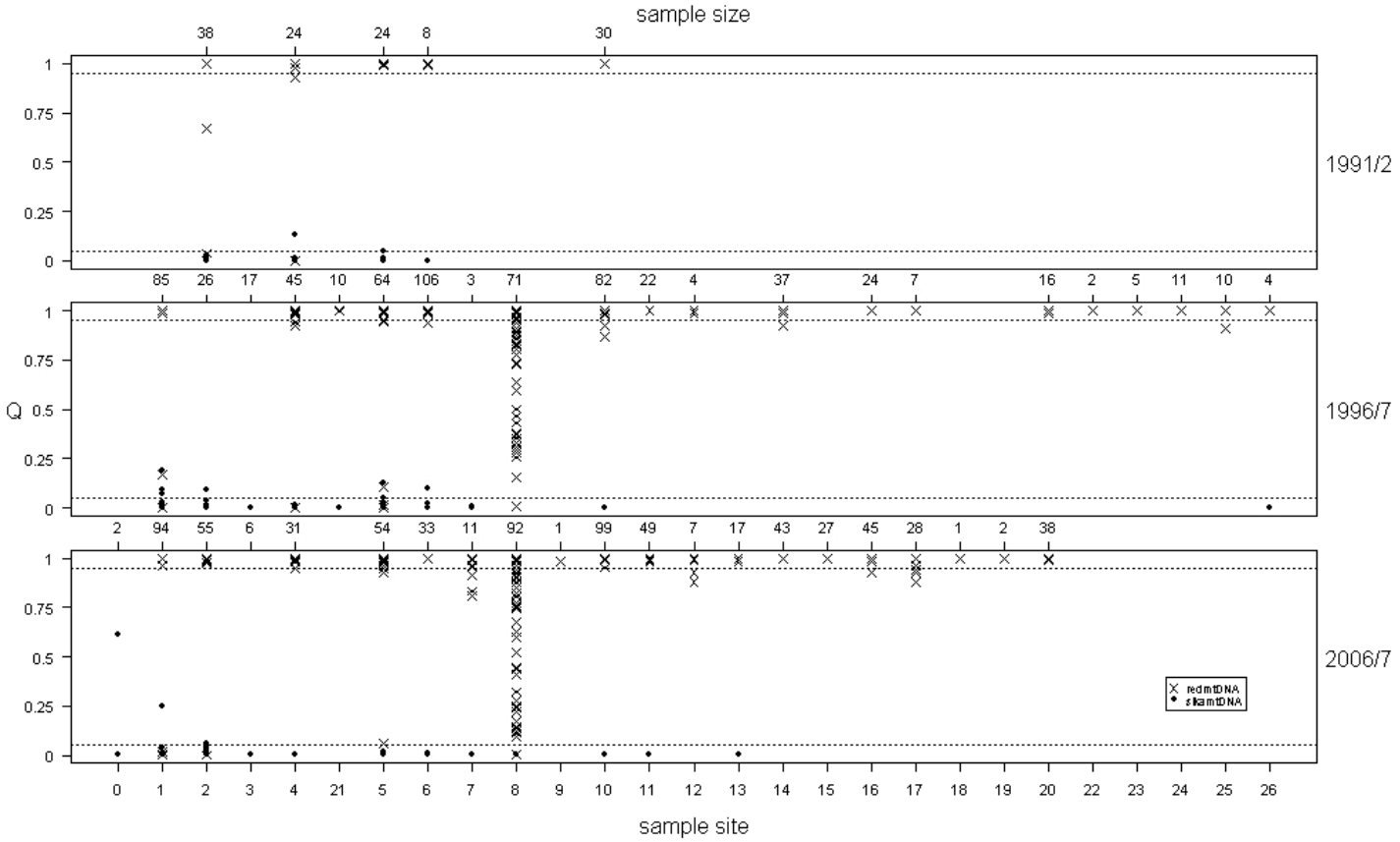


Figure 2: Distribution of Q by sample site for each of the three datasets. Individuals are also coded according to mtDNA haplotype. Note that many points overlap in the upper and lower regions of the graphs; sample sizes at each site are given above each plot.

The distribution of hybrid scores (Q) at each site and sample periods can be found in Figure 2. Generally, hybrid individuals had Q scores that indicated a low level of hybridisation (distant hybrids, or recent hybrid with values close to 0.05 or 0.95). These individuals had patterns of introgression typical of back-crossing (e.g. see later Table 4). As previously noted by Senn and Pemberton (Chapter 2), Site 8 (West Loch Awe) had a strikingly different pattern of hybridisation with hybrids possessing a broad range of Q scores and a pattern of introgression generated by mating between hybrids as well as backcrossing (e.g. see later Table 5). This pattern is found in the 1996/7 data set as well as the 2006/7 dataset (this site was not sampled in 1991/2). None of the individuals in the data set with Q values close to 0.5 had the pattern of allelic introgression expected for a F1 hybrid – in none were all, or even the majority, of markers heterozygous for 'red' and 'sika' alleles.

4.4.2 Changes in the proportion of recent hybrids and distant hybrids over time

We first examined the changes in the proportion of red (Q>0.5) and sika (Q<0.5) deer across the study area and over time. We performed logistic regression with binomial errors. The response variable (y) was a vector containing the number of red deer and sika deer found at a site (y= [number of red, number of sika]). We compared the fit of the most complex model (y ~ year of sampling + (year of sampling) ^2 + distance sampled from introduction point of sika, with interaction) with reduced models in a sequential manner to obtain the minimal adequate model. Only data from sites which had been sampled in two or more sampling episodes were included in the model (sites 1, 2, 4, 5, 6, 8, 10, 11, 14 & 16). Distance was calculated as the shortest distance in kilometres by land from the introduction point of sika.

There was a significant change in the proportion of deer species across the peninsula, with increasing composition of red deer moving northward from the introduction site, but there was no evidence that the cline in species composition had changed over time (Table 1, Figure 3).

Table 1: The results of logistic regression with binomial errors of the proportion of hybrids with distance from the sika introduction point at Carradale. No temporal trend was detected in any of the analyses.

response	estimate for distance	s.e	d.f	t	p
[Q>0.5,Q<0.5]	0.057	0.007	23	8.136	**<0.001**
all red , all sika					
[0.5<Q<0.95, Q>0.95]	0.007	0.018	22	0.372	0.713
red hybrids , red					
[0.99>Q>0.95, Q>0.99]	-0.024	0.010	22	-2.34	**0.029**
distant red , pure red					
As above excl. site 08	-0.038	0.008	20	-4.794	**<0.001**
[0.05>Q>0.5, Q<0.05]	0.050	0.017	17	2.964	**0.009***
sika hybrids , sika					
As above excl. site 08	-0.001	0.008	15	-0.155	0.879
[0.01<Q<0.05, Q<0.01]	0.001	0.007	17	0.081	0.936
distant sika, pure sika					

* The change in the proportion of recent sika hybrids with distance from the introduction site is only significant if site 08 is included; when site 08 is excluded the significance disappears.

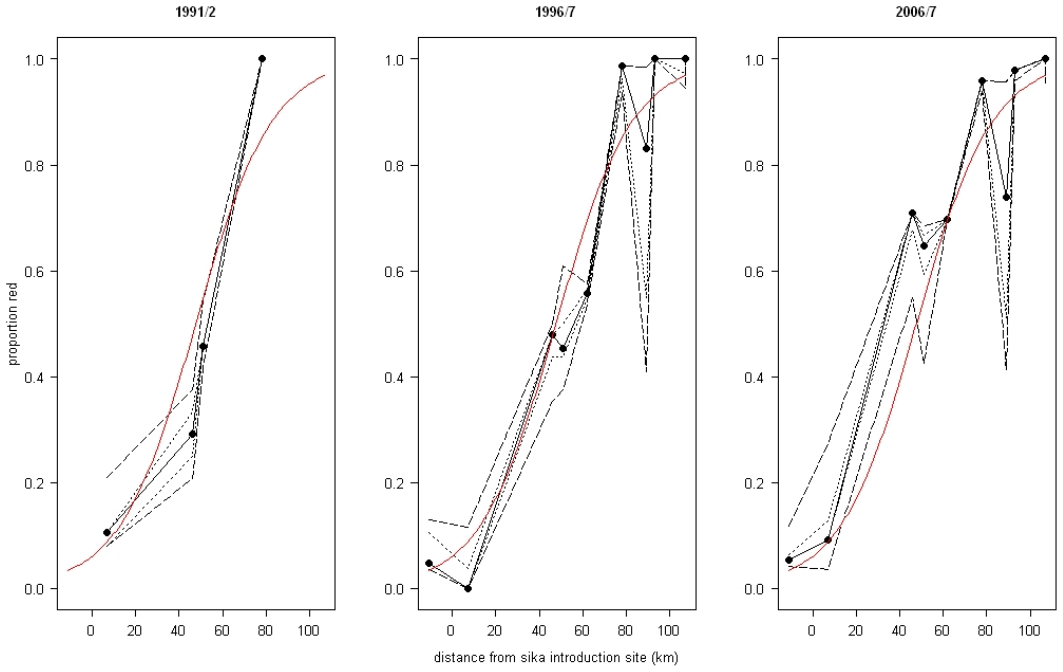


Figure 3: The proportion of red deer (Q>0.5) sampled across the peninsula (solid black line). Dotted lines bound the proportion of hybrids (0.05≤Q≤0.95) and dashed lines bound the proportion of distant hybrids (0.01≤Q≤0.99). The red curve is the fitted relationship for the model of proportion of red deer with distance (it is the same for all three time periods).

Similar logistic regression models were fitted to assess the change in proportion of hybrids over time. In order to assess temporal and spatial changes in recent hybrids, we fitted a vector [the number of recent red hybrids ($0.5<Q\leq0.95$), number of red deer ($Q>0.95$)] as a response variable against time of sampling and distance from the introduction point of sika (as above). In order to assess changes in the numbers of distant hybrids, another vector [numbers of distant hybrids ($0.95<Q\leq0.99$), number of pure deer ($Q>0.99$)] was fitted as a response as above. The same comparisons were made within sika (see Table 1).

Year of sampling (or the interaction with distance from introduction point) was not a significant variable in any of the models investigated. There was a significant decrease in the number of distant red hybrids moving away from the introduction site ($t= -2.3$, d.f.=22, $p=0.029$). There was also a significant increase in the number of recent sika hybrids moving away from the introduction site ($t=2.964$, d.f.=17, $p=0.009$). However, this relationship appeared to be driven by the high proportion hybrid individuals sampled at West Loch Awe. If data from West Loch Awe were removed from the model, the significance of distance disappeared (Table 1). There was still a significant decrease in distant red hybrids moving away from the introduction site if the data from West Loch Awe were removed from that model. These results are qualitatively the same as the findings of Goodman *et al.* (1999) for the 1991/2 data set (based on more samples than we analysed here, see methods). The proportion of hybrid classes found at each sample site are summarised in Figure 4.

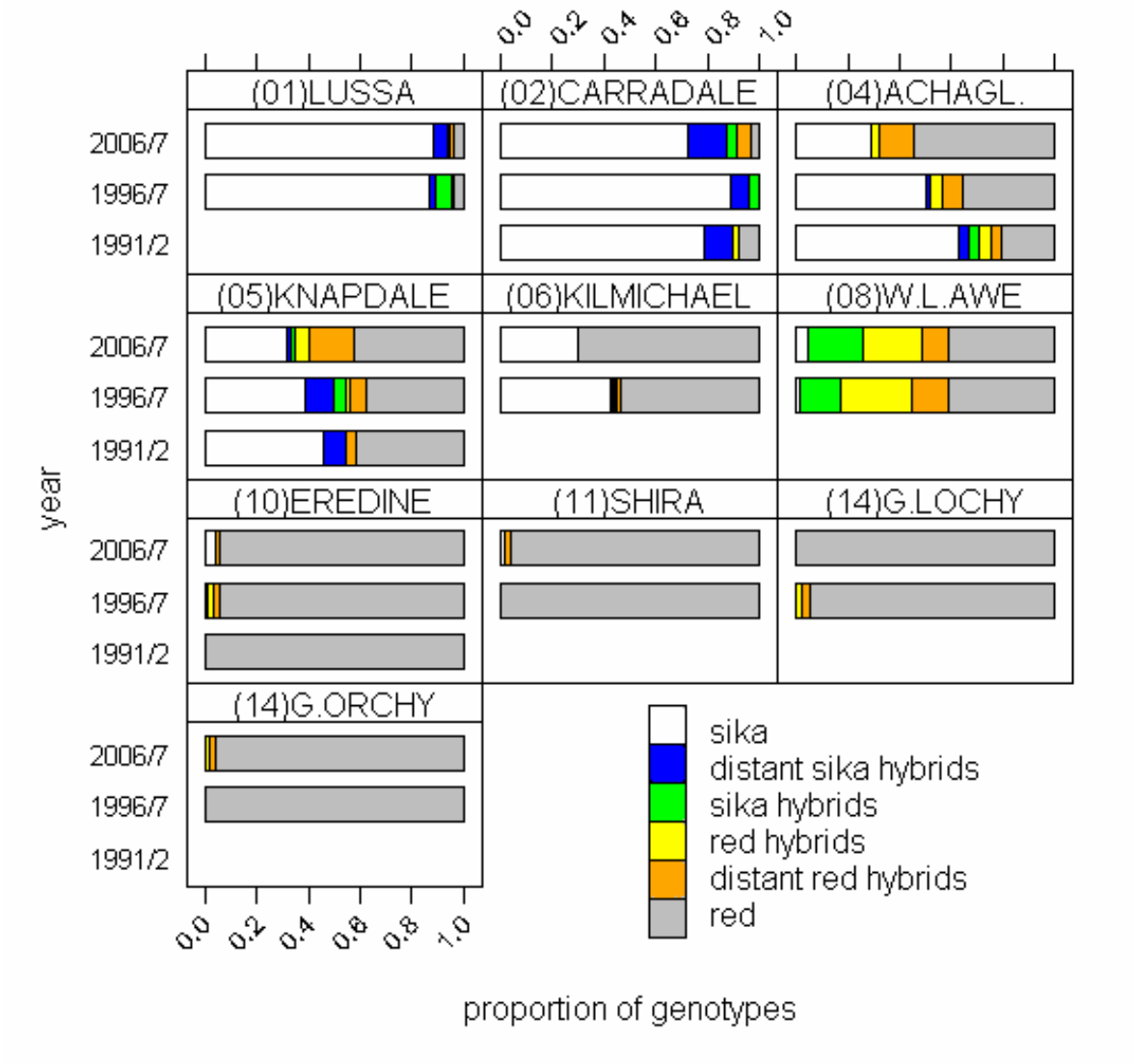


Figure 4: Proportion of animals in the genotype classes 0≤Q<0.01 (sika), 0.01≤Q<0.05 (distant sika hybrids), 0.05≤Q<0.5 (recent sika hybrids), 0.5<Q≤0.95 (recent red hybrids), 0.95<Q≤0.99 (distant red hybrids) and 0.99<Q≤1 (red deer) by population and sample period.

4.4.3 Spatial and temporal patterns of linkage disequilibrium

We tested for linkage disequilibrium in each species (sika Q<0.5, red Q>0.5) within each sample site and each time point as in Goodman *et al.* (1999) and examined the spatio-

temporal patterns of LD. Under LE, the number of introgressed alleles within each species should approximately follow the Poisson distribution, provided the alleles are rare (λ= number of introgressed alleles/ number of individuals). Using allelic designations as above, this gives a test for excess of individuals carrying multiple introgressed alleles (i.e recent hybrids, showing LD). The G-statistic ($G = 2 \sum$ Observed ln (Observed/Expected)) provides a summary of deviation from the Poisson distribution, but assessing its significance against the asymptotic $\chi 2$ would be inaccurate because the expected numbers of introgressed alleles are very small. Instead, we compared G against the distribution of G generated by randomisations of the dataset in which single locus diploid genotypes were shuffled across individuals 1000 times (Table 2). Simulations were carried out in R (version 2.8.1). For this test, alleles of unknown origin (see above) were reclassified as belonging to the parental species that was being examined (this is a conservative assumption).

We found significant LD in red deer in 9 samples: Carradale 1991/2 ($p<0.001$), Ormaig 2006/7 ($p<0.001$), West Loch Awe 1996/7 and 2006/7 ($p<0.001$), Eredine 1996/7 and 2006/7 ($p=0.006$, $p=0.003$), Oban 2006/7 ($p=0.001$) , Glen Orchy 2006/7 ($p= 0.017$) and Barcaldine 2006/7 ($p< 0.001$). However, at Carradale the LD originated from only one individual carrying 13 introgressed alleles. Interestingly, even though Eredine has high LD, there are no red deer with a high proportion of sika genes (Figure 4). One animal in 1996/7 carried 5 introgressed alleles and one had 3, but in the remainder of the data set no individual carried more than two alleles (see Table 2 for further details).

Table 2: Analysis of linkage disequilibrium (LD) in sika and red deer for each sample at which more than one individual was present.

		RED						SIKA					
sample site	**year**	**n**	**n_i**	**n_a**	**observed G**	**max G**	**p-value**	**n**	**n_i**	**n_a**	**observed G**	**max G**	**p-value**
Lussa	1996	4	3	5	3.034	10.648	0.287	81	22	46	64.009	20.139	**<0.001**
	2006	5	3	5	2.223	10.000	0.366	89	31	50	35.719	21.578	**<0.001**
Carradale	1991	4	1	13	35.961	12.845	**<0.001**	34	27	22	5.486	23.984	0.327
	1996							26	9	16	10.490	18.959	**0.061**
	2006	5	4	7	0.937	11.137	0.618	50	30	53	3.617	28.484	0.622
S. Tarbert	1996							17	3	3	0.564	26.662	-
	2006							6	0	0			
N.Tarbert	1996	8	2	3	1.501	14.462	0.532	2	0	0			
Achaglachach	1991	7	2	4	4.912	12.845	0.120	17	7	13	8.908	23.683	**0.079**
	1996	23	13	19	2.114	17.011	0.551	25	5	6	0.488	20.660	0.997
	2006	22	9	15	5.374	19.837	0.167	9	3	3	1.134	17.815	-
Knapdale	1991	11	2	3	1.978	20.910	0.503	13	7	10	1.124	17.670	0.902
	1996	29	11	14	1.163	20.784	0.833	35	14	35	22.342	21.634	**0.002**
	2006	35	22	37	5.195	19.784	0.188	19	13	20	2.596	16.700	0.676
Kilmichael	1991	5	2	2	0.935	12.845	-	3	2	3	0.795	13.022	0.804
	1996	59	11	14	3.283	22.873	0.467	47	17	26	12.534	21.190	**0.030**
	2006	23	2	3	3.245	19.982	0.389	10	5	6	0.649	24.129	0.972
Ormaig	1996							3	3	7	5.486	13.022	0.125
	2006	9	6	24	18.123	15.991	**<0.001**	2	0	0			
W. Loch Awe	1996	59	39	181	166.591	20.373	**<0.001**	12	12	145	23.763	17.167	**<0.001**
	2006	68	41	189	200.197	24.106	**<0.001**	24	23	187	45.277	21.052	**<0.001**
Eredine	1991	30	2	2	0.136	18.764	-						
	1996	81	14	21	17.523	23.293	**0.006**						
	2006	95	15	33	18.737	25.566	**0.003**	4	2	2	1.227	14.161	-
Shira	1996	22	5	5	1.234	17.926	-						
	2006	48	10	21	0.909	19.166	0.920						
Oban	1996	4	3	6	1.014	12.063	0.556						
	2006	7	3	10	13.456	18.037	**0.001**						
Succoth	2006	16	2	3	2.601	17.148	0.419						
Glen Lochy	1996	37	5	7	6.583	20.066	0.114						
	2006	43	13	17	2.346	25.434	0.570						
Ben More	2006	27	5	6	0.556	29.536	0.969						
Glen Orchy	1996	24	5	6	0.454	20.733	0.977						
	2006	45	3	6	12.513	21.448	**0.017**						
Barcaldine	1996	7	0	0	Indeterminate								
	2006	28	7	17	23.308	20.693	**0.001**						
Appin	2006	2	0	0	Indeterminate								
Ballachullish	1996	16	6	8	1.055	21.995	0.782						
	2006	38	4	4	0.437	20.540	-						
Morvern	1996	2	0	0	Indeterminate								
N.Cowal	1996	5	1	1	0.215	11.137	-						
S.Cowal	1996	11	6	8	1.071	14.188	0.754						
E.Cowal	1996	10	4	8	5.535	19.174	0.109						
Glendaruel	1996	3	1	1	0.378	8.705	-						

(n=sample size, n_i= number of introgressed individuals, n_a= number of introgressed alleles, max G and p-value were both derived by permutation test (1000 shufflings of the data). Observed G= Indeterminate indicates there were no individuals with introgressed alleles at that sample, P = - indicates there were no individuals with multiple introgressed alleles at that site.

In sika we found significant LD at 6 sites: Lussa 1996/7 and 2006/7 ($p<0.001$), Knapdale 1996/7 ($p=0.002$), Kilmichael 1996/7 ($p=0.030$) and West Loch Awe 1996/7 and 2006/7 ($p <0.001$). We also found marginally significant LD at Carradale 1996/7 ($p=0.061$), and Achaglachach 1991/2 ($p= 0.079$).

4.4.4 Changes in the number of introgressed alleles in red and sika

For this analysis we selected sites sampled at multiple time points, from which we had also collected at least 10 animals from each species at each time point (sites 1, 2, 4, 5, 6, 8, 10, 14 & 16). We compared the frequency of introgressed alleles within each species across the 20 scored loci using one-sided paired Wilcoxon's signed ranks tests (Ho= no increase in introgressed allele frequency over time). We found a very marginally significant increase in introgressed allele frequency in sika at Carradale between 1996/7 and 2006/7 ($V=32$, $p=0.059$) and 1991/2 and 2006/7 ($V=48$, $p=0.093$), but this significance would not survive a Bonferroni correction for multiple tests. Within red deer, we found a significant increase in introgressed allele frequency at Eredine between 1991/2 and 1996/7 ($V=0$, $p= 0.002$) and a marginally significant increase between 1991/2 and 2006/7 ($V=0$, $p=0.063$). Tables of the test results can be found in Supplementary Material Tables 2 and 3). There was a significant decrease in introgressed allele frequency in sika at West Loch Awe between 1996/7 and 2006/7 ($V=135.5$, $p=0.0017$), but here there was also an increase in pure sika ($0.01<Q$) from 1.4- 4.3% between the two time periods (Figure 4).

4.4.5 MtDNA introgression

MtDNA introgression was predominantly from red into sika (60 out of 61 cases, where sika is $Q<0.5$). The single instance of introgression from sika into red occurred at the south of the Peninsula (site 0) in the 2006/7 dataset (Chapter 2). MtDNA introgression from red into sika

occurred at 5 sites (Table 3). At West Loch Awe (site 8) all sika-like individuals (Q<0.5) carried red deer mtDNA in 1996/7 and in 2006/7 92% of sika-like individuals carried red deer mtDNA (Table 3). Apart from West Loch Awe where many individuals with intermediate scores were present, mtDNA introgressed individuals tended to have low Q scores (Figure 2). Note that two reported cases of introgression of sika mtDNA into red previously reported in the 1991/2 dataset (Abernethy 1994; Goodman *et al.* 1999) were discovered to be incorrect in the retyping of the samples. At sites 2, 4 and 5 the number of individuals found with mtDNA introgression in any one sample period was very low (0-4) and there is no evidence for any temporal trend in mtDNA introgression. Site 1 shows an increase in mtDNA introgression from 6.2% in 1996/7 to 13.5% in 2006/7. However this is not statistically significant ($\chi 2$ (Yates' correction) =1.77, d.f.=1, p= 0.183).

Table 3: Number of sika-like (Q<0.5) individuals sampled with red mtDNA (total number of sika-like individuals sampled).

	1991/2	***1996/7***	***2006/7***
1 Lussa	~	5 (81)	12 (89)
2 Carradale	1 (34)	0 (26)	1 (50)
4 Achaglachach	1 (17)	1 (25)	0 (9)
5 Knapdale	0 (13)	4 (35)	1 (19)
8 West Loch Awe	~	12 (12)	22 (24)

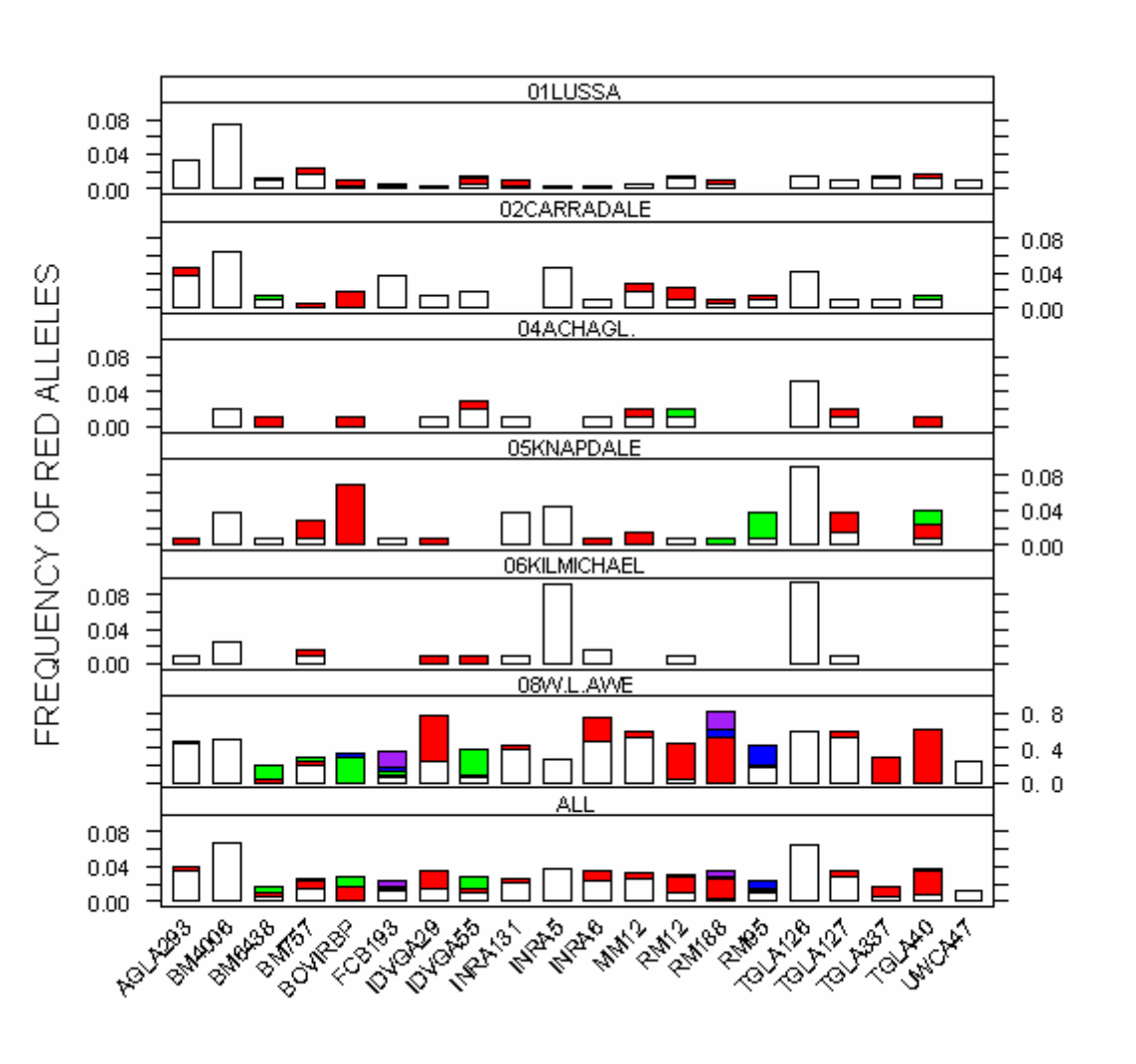
4.4.6 Estimating the number of hybridisation events contributing to introgression into sika

Since red deer are highly polymorphic at most loci (see Section 4.4.1) it should be possible to a) gain a rough estimate of the minimum possible number of hybridisation events that have contributed to the introgression of sika at each locus by counting the number of sizes of red alleles in the sika population and b) for populations in which introgression proceeds by backcrossing only (i.e. not West Loch Awe) it should be possible to reconstruct the

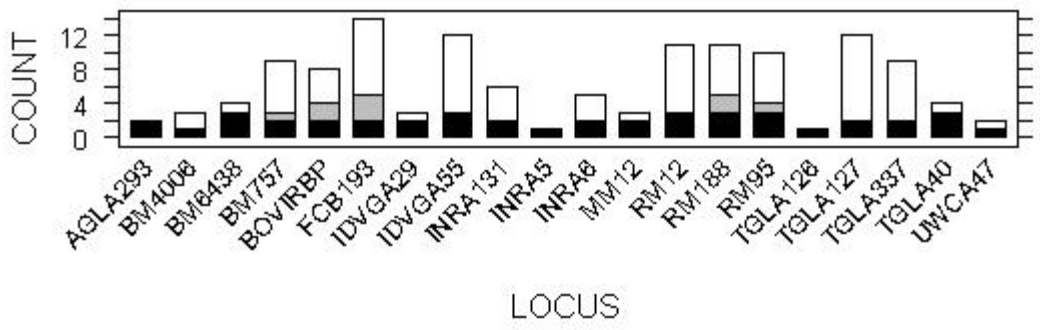
haplotypes of the red deer involved in F1 events, and thus gain an estimate for the number of F1 events.

Across the data set, all loci had introgressed alleles and the frequency of introgressed alleles was fairly even across all loci (Figure 5a). Frequency and sizes of introgressed red alleles in sika varied considerably with site, with some loci having introgressed alleles at some sites and not others. In some cases the same loci at different sites had introgressed alleles of different sizes (Figure 5a). Across the whole study area, but outwith West Loch Awe, the maximum number of red alleles found introgressed into sika at any locus was three. If West Loch Awe was included, this number rose to five (Figure 5b). If the data set was divided up into the three sample periods (Figure 5c), this revealed a heterogeneous pattern of introgression over time. Since the red alleles that could possibly introgress into sika exist at different frequencies (Supplementary Material, Figure 2) we verified that the total number of F1 events required to introduce the number of introgressed alleles found at each locus was indeed low. We did this by random sampling with replacement from the distribution of red allele frequencies at each locus estimated by the STRUCTURE analysis (10 000 replicates). The average number of F1 hybridisation events likely to have contributed to introgression at each locus can be found in Table 4 and Figure 4 of the Supplementary Material, it can be seen that although it cannot be assumed that each hybridisation event introduced a new allele into the sika population, the number of red alleles that we find in the sika population is still consistent with a very small number of introduction events (2- 17 depending on the locus).

5a)


01LUSSA
02CARRADALE
04ACHAGL.
05KNAPDALE
06KILMICHAEL
08W.L.AWE
ALL
FREQUENCY OF RED ALLELES
0.08
0.04
0.00
AGLA293
BM4006
BM6438
BM757
BOVIRBP
FCB193
IDVGA29
IDVGA55
INRA131
INRA5
INRA6
MM12
RM12
RM188
RM95
TGLA126
TGLA127
TGLA337
TGLA40
UWCA47

5b)


COUNT
12
8
4
0
AGLA293
BM4006
BM6438
BM757
BOVIRBP
FCB193
IDVGA29
IDVGA55
INRA131
INRA5
INRA6
MM12
RM12
RM188
RM95
TGLA126
TGLA127
TGLA337
TGLA40
UWCA47
LOCUS

5c)

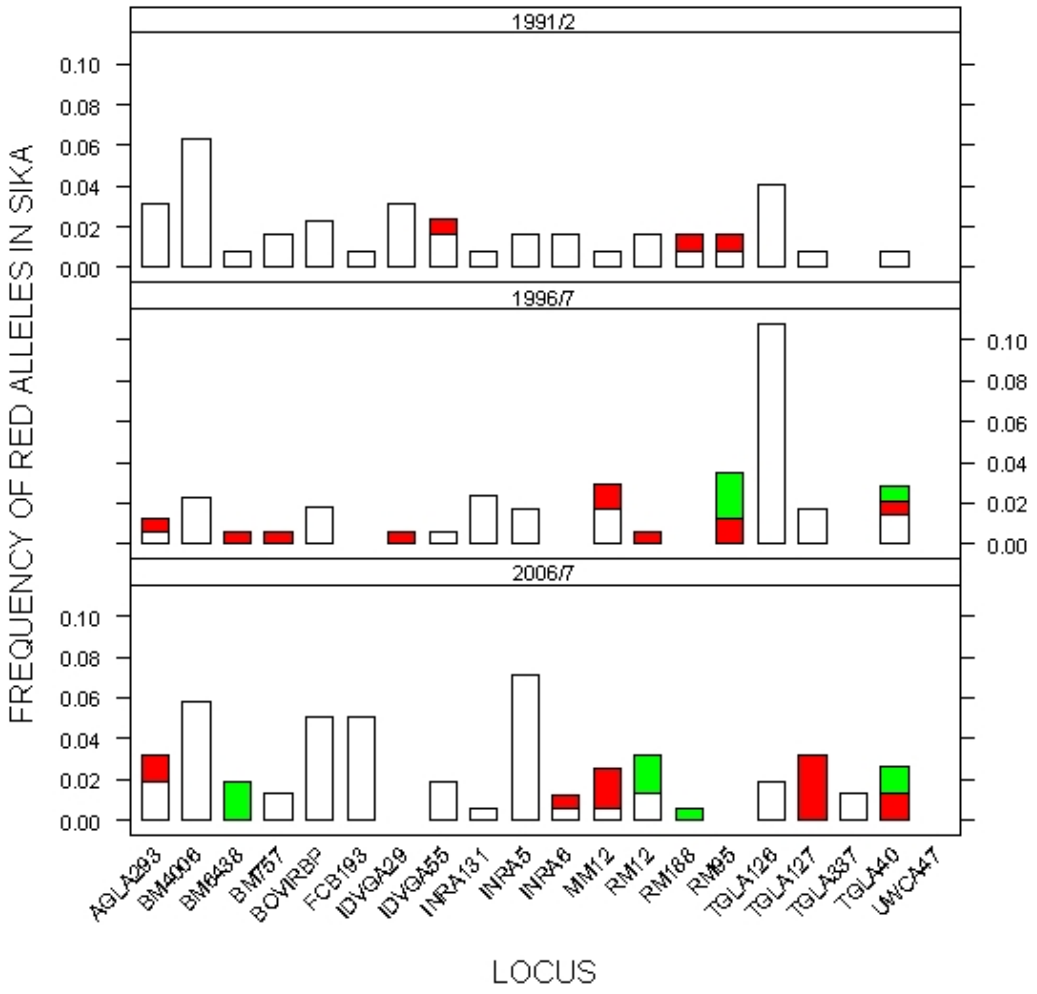


Figure 5: a) Frequency of introgressed alleles from red into sika (Q<0.5) at each locus by site (all three temporal data sets combined). Sample sizes for the sites are Lussa (n=170), Carradale (n=110), Achaglachach (n=51), Knapdale (n=67), Kilmichael (n=60), West Loch Awe (n=36). Unweighted allele frequencies across all sites are also included. Each allele at each locus is presented in a different colour: 1st allele (white), 2nd allele (red), 3rd allele (green), 4th allele (blue), 5th allele (purple). Across different sites the same allele at the same locus is represented by the same colour. Note that the scale on the y-axis is different at West Loch Awe (08W.L.AWE) from the other sites. b) Count of introgressed alleles from red into sika (Q<0.5) at each locus at all sites apart from West Loch Awe (in black) with count of additional alleles introgressed at West Loch Awe (in grey) in comparison to the additional number of alleles that could have introgressed from red (in white).c) Frequency of introgressed alleles from red into sika (Q<0.5) at each locus by year (data from sites 2, 4 and 5). Sample sizes for 1991/2 (n=64), 1996/7 (n=86), 2006/7 (n=78). Unweighted allele frequencies across all sites are also included. Each allele at each locus is presented in a different colour: 1st allele (white), 2nd allele (red), 3rd allele (green). Across different sample years the same allele at the same locus is represented by the same colour.

Next we examined whether it was possible to reconstruct haplotypes involved in the F1 hybridisation events. In the populations where introgression proceeds via backcrossing (i.e. not West Loch Awe), each F1 hybridisation event will introduce a haplotype of red alleles into the sika population which will be broken up as backcrossing proceeds. Hybrids that are recent backcrosses can be identified because they are likely to carry multiple introgressed alleles: examination of these individuals should allow us to reconstruct the haplotypes that have contributed to recent introgression within these populations. We first decided on a new definition of a recent hybrid: the distribution of the introduction of red alleles into sika at *n* loci at *x* th generation of backcrossing can be represented by the probability density function of the binomial distribution (2n, 1/(2(x+1)) Supplementary Material, Figure 5). Inspection of this distribution showed that sika-like individuals carrying 4 or more red alleles are very unlikely to be carrying them due to chance reassembly of previously introgressed red alleles that are now at (or close to) LE.

Table 4: Genotypes of sika-like animals with more than four introgressed red alleles from sites outwith West Loch Awe. Only the introgressed red alleles are shown. The animals are grouped into the most parsimonious groupings given genotype, population of origin, and sampling episode. This is just one of many of the groupings that is possible. The minimum number of groupings possible is 4 if population of origin is ignored.

Population	*Year*	*MtDNA*	*age*	*AGLA293*	*BM4006*	*BM6438*	*BM757*	*BOVIRBP*	*FCB193*	*IDVGA29*	*IDVGA55*	*INRA5*	*INRA6*	*INRA131*	*MM12*	*RM12*	*RM188*	*RM95*	*TGLA40*	*TGLA126*	*TGLA127*	*TGLA337*	*UWCA47*
1	1996/7	s	1	144	93	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	101	Ø	Ø	145	Ø
1	1996/7	s	0	Ø	93	Ø	162	Ø	Ø	Ø	195	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	231
1	1996/7	s	2	144	93	Ø	Ø	Ø	Ø	Ø	195	Ø	Ø	100	91	125	Ø	Ø	101	105	Ø	Ø	Ø
1	1996/7	r	2	144	93	Ø	162	Ø	Ø	Ø	Ø	Ø	Ø	100	91	Ø	Ø	Ø	101	Ø	Ø	145	Ø
1	1996/7	s	2	144	Ø	Ø	162	Ø	Ø	Ø	Ø	126	Ø	Ø	Ø	Ø	Ø	Ø	101	Ø	Ø	Ø	Ø
1 (1996/7) consensus				**144**	**93**	**Ø**	**162**	**Ø**	**Ø**	**Ø**	**195**	**126**	**Ø**	**100**	**91**	**125**	**Ø**	**Ø**	**101**	**105**	**Ø**	**145**	**231**
1	**2006/7**	**s**	**?**	**Ø**	**93**	**249**	**Ø**	**151**	**113**	**136**	**Ø**	**Ø**	**Ø**	**Ø**	**Ø**	**Ø**	**123**	**Ø**	**97**	**Ø**	**Ø**	**136**	**231/131**
2	1991/2	s	?	144	Ø	Ø	198	153	Ø	Ø	Ø	Ø	Ø	Ø	Ø	139	Ø	Ø	Ø	Ø	Ø	Ø	Ø
2	1996/7	s	0	Ø	Ø	Ø	Ø	Ø	Ø	Ø	195	Ø	Ø	Ø	91	Ø	Ø	128	Ø	105/105	Ø	Ø	Ø
2	2006/7	s	1	Ø	93	Ø	Ø	Ø	Ø	Ø	Ø	126	Ø	Ø	91	Ø	Ø	Ø	Ø	Ø	Ø	145	Ø
2	2006/7	s	0	144	Ø	Ø	Ø	153	103	Ø	195	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø
2 consensus				**144**	**93**	**Ø**	**198**	**153**	**103**	**Ø**	**195**	**126**	**Ø**	**Ø**	**91**	**139**		**128**	**Ø**	**105**	**Ø**	**145**	**Ø**
4	**1991/2**	**s**	**?**	**Ø**	**Ø**	**249**	**Ø**	**Ø**	**Ø**	**Ø**	**195/199**	**Ø**	**134**	**Ø**	**Ø**	**Ø**	**Ø**	**Ø**	**Ø**	**Ø**	**169**	**Ø**	**Ø**
5	1996/7	s	0	Ø	Ø	Ø	162	153	Ø	143	Ø	126	Ø	Ø	Ø	Ø	Ø	138	Ø	Ø	Ø	Ø	Ø
5	1996/7	r	0	Ø	93	Ø	Ø	153	Ø	Ø	Ø	126	Ø	Ø	89	Ø	Ø	Ø	Ø	105	Ø	Ø	Ø
5	1996/7	r	0	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	98	Ø	Ø	Ø	Ø	91/91	Ø	169	Ø	Ø
5	2006/7	r	4	Ø	93	251	Ø	153	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	105	Ø	Ø	Ø
6	1996/7	s	2	Ø	Ø	Ø	162	Ø	Ø	143	Ø	126	Ø	Ø	Ø	Ø	Ø	Ø	Ø	105/105	Ø	Ø	Ø
5&6 consensus				**Ø**	**93**	**251**	**162**	**153**	**Ø**	**143**	**Ø**	**126**	**Ø**	**98**	**89**	**Ø**	**Ø**	**138**	**91**	**105**	**169**	**Ø**	**Ø**

Ø means both alleles found at this locus were of sika origin. If one allele is presented (e.g.. 144) this means the genotype at this locus was for the red allele '144' and one sika allele. Genotypes with two alleles (e.g. 195/199) mean two red alleles were present at this locus. In some cases these genotypes are homozygous (e.g. 91/91). This could be because this animal carried two introgressed red alleles identical by state, or it may have an introgressed red allele and a null allele at this locus. For further discussion of null alleles at these loci see Chapter 2. The bold genotypes represent the five separate haplotypes represented in the data.

So for this section of the study we defined sika-like individuals carrying four or more introgressed alleles as recent hybrids and used these individuals to reconstruct the number of F1 hybridisation events contributing to recent introgression. These hybrid haplotypes could be grouped together in a number of ways; the smallest possible number of groupings was four, but the most parsimonious accounting for geography and the time of sampling is five (Table 4). A number of individuals have loci with two introgressed red alleles. This is either because one of the two alleles originates from a recent hybridisation event and the other originates from a much older hybridisation event, or because a null allele is present at this locus (this is only a possibility in the cases where the introgressed alleles are identical in state). The genotypes of the hybrids with four or more introgressed alleles at West Loch Awe can be found in Table 5. These individuals cannot be grouped by haplotypes, because a large number have loci at which two red alleles are present. This indicates that these animals have a more complicated pedigree than the backcrossed individuals on the rest of the peninsula (e.g. they are F2s, Bx1 * Bx1, Bx2*F1 etc). Here we do not expect that red deer alleles are only introduced into sika through F1 haplotypes, but enter the population through a variety of different hybrid matings.

Table 5a: Genotypes of sika-like animals with more than 4 (but with 20 or fewer) red alleles from West Loch Awe. Only the introgressed red alleles are shown.

Year	*mtDNA*	*age*	*AGLA293*	*BM4006*	*BM6438*	*BM757*	*BOVIRBP*	*FCB193*	*IDVGA29*	*IDVGA55*	*INRA5*	*INRA6*	*INRA131*	*MM12*	*RM12*	*RM188*	*RM95*	*TGLA40*	*TGLA126*	*TGLA127*	*TGLA337*	*UWCA47*
2006/7	r	2	144	Ø	Ø	Ø	Ø	124	143	Ø	Ø	126	Ø	Ø	125	Ø	Ø	Ø	Ø	Ø	97	Ø
	r	3	144	Ø	Ø	Ø	Ø	Ø	Ø	195	Ø	126	134/136	91	125	137	Ø	105	Ø	136	Ø	Ø
	r	3	144	93/93	Ø	Ø	147	10 3	143/143	197	98	Ø	134	91	125	127	132	Ø	178/178	Ø	97	231
	r	0	Ø	93	Ø	Ø	Ø	124	143	Ø	Ø	Ø	134	91	Ø	127	132	Ø	178	Ø	Ø	231
	r	3	Ø	93	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	91	Ø	127	Ø	Ø	178	136	97	Ø
	r	1	144	93	Ø	Ø	Ø	124	143	197	100	Ø	134	Ø	125	127	Ø	Ø	Ø	Ø	97	Ø
	r	3	Ø	93	Ø	183	147	Ø	136/143	Ø	98	Ø	136	89	125	129	132	Ø	169/178	Ø	97	231/231
	r	4	Ø	Ø	249	Ø	147	Ø	136	199	98	Ø	134/134	89	125	Ø	132	Ø	Ø	Ø	Ø	Ø
	r	4	Ø	Ø	Ø	Ø	Ø	Ø	Ø	197	Ø	Ø	134	Ø	Ø	Ø	Ø	Ø	178	Ø	97/97	Ø
	r	4	Ø	Ø	Ø	Ø	Ø	10 9	Ø	Ø	Ø	Ø	134	91	Ø	Ø	Ø	Ø	Ø	Ø	97	Ø
	r	1	144	Ø	Ø	Ø	147	Ø	136	Ø	98	Ø	Ø	Ø	Ø	129	Ø	105	178	136	Ø	231
	r	5	144	93/93	Ø	198	Ø	Ø	Ø	Ø	98	Ø	136	91	Ø	127/137	132	105	178	Ø	Ø	231
	r	3	144	93	Ø	Ø	147	Ø	143	Ø	98	Ø	134	Ø	Ø	137	132	105	178	Ø	Ø	Ø
	r	0	Ø	93	Ø	Ø	Ø	124	Ø	197	98	Ø	Ø	91	125	127	126/128	105	Ø	136	Ø	Ø
	r	3	Ø	93/93	253	162/162	Ø	113/124	143	195	Ø	126	134	Ø	125	127	132	105	Ø	136	Ø	231
	r	3	144	Ø	Ø	Ø	Ø	Ø	Ø	197	Ø	Ø	134/136	Ø	Ø	137	Ø	Ø	Ø	Ø	97/97	Ø
	r	0	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	Ø	129	Ø	105	Ø	136	97	231
	r	0	144	93	Ø	Ø	Ø	Ø	Ø	Ø	Ø	126	Ø	Ø	Ø	137	Ø	105	178	Ø	Ø	Ø
	r	0	128	Ø	Ø	Ø	Ø	Ø	143	Ø	Ø	Ø	136	Ø	Ø	Ø	Ø	Ø	Ø	Ø	97	Ø
	r	0	Ø	93	Ø	Ø	Ø	109	Ø	Ø	Ø	Ø	134/136	91	Ø	Ø	Ø	Ø	169	136	Ø	Ø
Total			**2**	**1**	**2**	**3**	**1**	**4**	**2**	**3**	**2**	**1**	**2**	**2**	**1**	**3**	**3**	**1**	**2**	**1**	**1**	**1**

Table 5b: Genotypes of sika-like animals with more than 4 (but with 20 or fewer) red alleles from West Loch Awe. Only the introgressed red alleles are shown.

Year	*mtDNA*	*age*	*AGLA293*	*BM4006*	*BM6438*	*BM757*	*BOVIRBP*	*FCB193*	*IDVGA29*	*IDVGA55*	*INRA5*	*INRA6*	*INRA131*	*MM12*	*RM12*	*RM188*	*RM95*	*TGLA40*	*TGLA126*	*TGLA127*	*TGLA337*	*UWCA47*
1996/7	r	8	144/144	Ø	253	Ø	147	Ø	136/143	197	98	Ø	136	91	125	127	128	105/105	178	136/136	97	Ø
	r	2	144	Ø	Ø	Ø	157	Ø	136	Ø	98/98	126	136	Ø	Ø	127	Ø	105	178	Ø	Ø	Ø
	r	3	Ø	Ø	253	162	147	Ø	136/143	197	Ø	Ø	134	91	125/125	127	128/128	Ø	178	136	Ø	Ø
	r	0	Ø	Ø	253	162	147	Ø	143/143	197	98	Ø	134	91	125	127	Ø	105/105	178	Ø	97	Ø
	r	1	144/144	Ø	Ø	Ø	Ø	Ø	143	197	Ø	126	134	91/91	125	127	Ø	10 5	178	Ø	Ø	Ø
	r	2	Ø	Ø	253	162	147	Ø	143/143	197	98	Ø	134	91	125	127	Ø	Ø	178	Ø	97	Ø
	r	0	Ø	93	Ø	162	147	Ø	143/143	197/197	Ø	126	134	89/91	125/125	127	Ø	105	178/178	136/136	97	Ø
	r	1	Ø	93	Ø	162	147	122	136/136	Ø	98	126	134/136	91	Ø	127/137	128/132	Ø	Ø	Ø	97	Ø
	r	3	Ø	93/93	Ø	198	Ø	103/124	143	Ø	Ø	126	136	91	Ø	127/137	Ø	Ø	Ø	Ø	97	231
	r	2	144	93	Ø	162	Ø	Ø	143	Ø	Ø	126	Ø	91	125	127	128	105/105	178/178	Ø	Ø	Ø
	r	1	144	Ø	253	Ø	157	Ø	136	197	98	Ø	Ø	Ø	Ø	127	Ø	105/105	178	Ø	97	Ø
	r	1	Ø	Ø	Ø	Ø	Ø	124	143	Ø	Ø	Ø	Ø	91	Ø	127/127	Ø	Ø	Ø	Ø	97	Ø
Total			**1**	**1**	**1**	**2**	**2**	**3**	**2**	**1**	**1**	**1**	**2**	**2**	**1**	**2**	**2**	**1**	**1**	**1**	**1**	**1**

4.5 Discussion

4.5.1 Summary

Our analysis using 20 unlinked loci and one mtDNA marker on a sample of 1513 individuals taken from populations of hybridising red and sika deer at three time periods spanning 15 years shows little change in the structure of the hybridising population over time. Broadly, the structure of the hybrid zone shows a clinal transition from sika deer to red deer as one moves away from the introduction point of sika on the south of the peninsula at Carradale (Figure 1). However, neither the distribution of recent sika and red hybrids, nor the distribution of distant sika hybrids shows clinal variation across the peninsula. The proportion of distant red hybrids does decrease with distance from the introduction point of sika, indicating that introgression into red deer has been occurring for longer on the south of the peninsula. The lack of clinal distribution of recent hybrids is highlighted by the contrast between the two sites, Eredine and West Loch Awe, sited approximately equally far from the introduction point of sika but with very different population structures (<1 % and 44% recent hybrids respectively, Figure 4).

There was no systematic change across the study area in the proportion of animals designated as recent hybrids or distant hybrids over the time period of the study (15 years). No significant increase was detected in the level of mtDNA introgression over time. The majority of mtDNA introgression across the whole data set (60/61 cases) was from red into sika implying that mating between sika stags and red deer hinds is the main cause of hybridisation. Little within-site increase in the frequency of introgressed alleles over time within red deer and sika was detected. Individual sites did however reveal some changes in the pattern of LD. Within red deer three sites beyond the main range of sika (Glen Orchy, Oban & Barcaldine) showed significant LD in the 2006/7 sampling but not in 1996/7.

Additionally at Eredine in 1991/2 no individuals carrying multiple introgressed alleles were present, but significant LD was found in 1996/7 (p=0.006) and 2006/7 (p=0.003). Within sika we found significant or marginally significant LD in 1996/7, but not in 1991/2 and 2006/7 at Carradale (p=0.061), Knapdale (p=0.002) and Kilmichael (p=0.030). We also found marginally significant LD at Achaglachach in 1991/2 (p=0.079), but not in subsequent samplings.

Counting the allelic variants introgressed from red deer into sika revealed that a minimum of five separate introductions of red alleles into sika (or three if West loch Awe is excluded) have occurred. Simulations from the allelic distributions of red deer confirm that the number is likely to be ≤17 (or ≤5 if West Loch Awe is excluded). The reconstruction of the haplotypes of recent sika backcrosses (not at West Loch Awe), reveals that four separate F1 events could explain their generation, although a more parsimonious arrangement which groups individuals based on geographic proximity and the time of sampling suggests that a minimum of five separate F1 events would be needed.

4.5.2 Changes in the structure of the hybridising populations

The 15 year period spanned by this study represents 3-5 generations of red and sika deer. 15 years also represents 13% of the ~115 years that sika have been present on the Kintyre Peninsula. We know that 11 sika were released initially and we know that currently the range of sika extends ~70km from the introduction site and that 6.9% of individuals across the study area are hybrids (Chapter 2). Considerable change in the structure of the hybridising populations has occurred within the last 115 years, so it is perhaps rather surprising that we don't detect much change in genetic structure of the hybridising population in the form of an increase in recent hybrids over the last 15 years. We do see a signal of recent expansion of introgression into red deer in some of the northern populations at Glen Orchy, Oban and Barcaldine (although some of the samples are quite small, see above). Barcaldine is so far

north that it is possible that the hybrids found here originate from a sika population that arose from a separate introduction to the north east of Fort William at Forth Augustus (Ratcliffe 1987), although we see no sign of introgression at more northerly sites. Introgression into red deer also appears to have increased at Eredine over the 15 years of the study. Although we only found two 2 individuals considered to be recent hybrids using our definition of Q (carrying 3 and 5 introgressed alleles) and these are in the 1996/7 data set, there is nevertheless a significant amount of LD caused by 8 individuals carrying two introgressed alleles in 2006/7. It is possible that these introgressed alleles originate from F1 events that occurred at Eredine in the past, or that they have diffused out from neighbouring West Loch Awe. Ormaig to the south west and Oban to the north of West Loch Awe also show signs of recent introgression into red, signalling that sika alleles may be spreading out from West Loch Awe into the surrounding red populations.

Within sika, the picture is somewhat different. At three sites in 1996/7 we see a number of recent hybrids with reasonably high Q values (Figure 2) in comparison to the other sample periods. We also see significant LD or marginally significant LD at those sites at this time (Carradale, p=0.061, Knapdale, p= 0.002, Kilmichael, p= 0.030). Provided this pattern is not sampling artefact, it appears that there must have been F1 hybridisations which led to the generation of recent hybrid backcrosses within sika between 1991/2 and 1996/7, but by 2006/7 introgressed alleles (which are present, see Table 2) had dispersed through the population via backcrossing into sika and were no longer in LD. At Achaglachach we see a similar decline in LD between 1991/2 (p= 0.079) and the later sampling periods, although LD is, in the main, caused by one individual carrying 5 introgressed alleles (Figure 2). At Lussa, on the other hand, we see highly significant LD (P<0.001) within sika at both sampling episodes (1996/7 & 2006/7) although it has decreased fairly considerably between the two periods (from G=64.0 to G=35.7, see Table 2).

An inherent and unavoidable problem with this study is that hybrids are rare. Whilst samples at sites involved in the temporal analyses rarely contained fewer than 20 individuals (Figure 2), this may not be sufficient given the rarity of hybridisation. If hybridisation is rare, not only are large samples needed to detect recent hybrids (and therefore to find statistically significant changes), but introgressing alleles in more numerous advanced backcross classes will also be at low frequency, making them hard to detect. It is also regrettable that the temporal data series is not complete across all sites, fewer sites were sampled in 1991/2, so that for some sites we cannot compare samples between which we would expect the greatest change.

4.5.3 West Loch Awe

As previously reported (Chapter 2), the structure of the hybridising population is clearly different at West Loch Awe from other regions on the peninsula. The proportion of recent hybrids ($0.05 \leq Q \leq 0.95$) does not vary much between the two sample periods (43.6- 44.6%) and nor does the proportion of pure red deer (~41%), but we do observe a shift over time from 51% to 65% sika-like individuals within the recent hybrids (Figure 4), although this is not statistically significant ($\chi 2=0.788$, d.f =1, $p=0.375$). Interestingly we sampled no distant sika hybrids ($0.01 \leq Q < 0.05$) at either time point, and indeed the proportion of pure sika at this site is very low (1.4- 4.3 % in both males and females). This suggests that sika have relatively recently arrived at this site and so perhaps more matings have occurred between hybrids and red deer than between hybrids and sika (because we do see distant red deer hybrids). This suggestion also supports the idea that the breakdown between the two species must have occurred fairly rapidly: sika are thought to have arrived at West Loch Awe some time between 1965 and 1975 (Ratcliffe 1987), and since the hybrid swarm was present in 1996 this gives a period of 20-30 years (~5-8 generations) for its formation. Rapid change has occurred across the peninsula even if we have not been able to observe it in this study.

Since hybridisation is currently extensive at West Loch Awe, but not at other sites, it might be tempting to suggest that introgression detected at other sites is as a result of gene flow out from West Loch Awe. The pattern of introgression of red alleles into sika clearly shows that this cannot be the case as alleles that are not found at West Loch Awe are found at other sites across the peninsula (Figure 5a). Differences in the distributions of introgressed red alleles can also be detected between the other sites (Lussa, Carradale, Achaglachach, Knapdale and Kilmichael). We conclude that F1 hybridisation events have occurred across most of the peninsula, not just at West Loch Awe and that variation in the timing and number of hybridisation events at each site is responsible for the variation in introgression that we see across the study area. For a discussion of the striking pattern of mtDNA introgression found at West Loch Awe see Chapters 2 and 6.

4.5.4 The number of hybridisation events

The results of the stochastic simulations show that the minimum number of hybridisation events contributing to introgression into sika at each locus is low (<17, Supplementary Material, Table 4). However, estimating the number of hybridisation events that have occurred overall, is more complex: After n generations approximately 20% of F1 hybridisation events followed by backcrossing leave no introgressed offspring at 20 surveyed loci (but may leave offspring introgressed elsewhere across the genome, Chapter 3). Other F1 hybridisation events followed by backcrossing may result in introgression at one or more loci. Thus the pattern of introgression seen across the study area could be due to the minimum number of events possible from the pattern of introgression at the most polymorphic loci (in this case 5, see Figure 5) or could be due to a larger number of hybridisation events each contributing to introgression at separate loci. Additionally, at West Loch Awe red alleles are being introduced into sika through a complex pattern of mating (not just through the generation of F1s followed by back crossing). We have shown that the recent backcross sika hybrids found outwith West loch Awe could be explained by as few as

5 separate F1 events (Table 4). A similar analysis is not possible for West Loch Awe (because we can not assume backcrossing) or the red deer population (because sika alleles are homogenous). However, it is clear that within the sika population overall, hybridisation has occurred few enough times that many alleles that are relatively common in red deer have not passed into sika (Figure 5b).

4.5.5 Expected changes in introgression

In the areas where red deer and sika have overlapped for decades, the lack of increase in the number of recent hybrids over time is not particularly surprising, as an increase (or decrease) in numbers would imply a change in the overall rate of hybridisation. More surprising is the lack of increase in introgressed allele frequencies in sika and red deer, since these are predicted to increase even if the hybridisation rate is constant. Yet there is no evidence to support the hypothesis that the proportion of recent hybrids in the population has decreased over time, so it seems unlikely that a slow-down in the rate of hybridisation is responsible for the observed pattern in introgression.

One possibility for the lack of increase in introgression over the time of the study is that negative selection is acting on introgressing alleles. This suggestion was previously raised by Goodman *et al*. (1999) in their analysis of the 1991/2 dataset where they estimated that the level of introgressed alleles in LE was 30-40% of that expected from the proportion of recent hybrids, if introgressed alleles were selectively neutral. They did this using various modelling assumptions about population growth rate and the reproductive success of backcross classes, and although introgression was lower that expected, the results were not statistically significant. Red and sika karyotypes differ by two Robertsonian translocations (2n red =68, 2n sika= 64), captive bred F1 hybrids show no evidence of nondisjunction (Herzog & Harrington 1991), so we would expect any negative selection to operate at the level of the gene. Anyhow, if negative selection was genome wide (or operating across a

large number of loci), its effects would be difficult to disentangle from the effects of low hybridisation rate since F1 hybrids would have very low fitness. In theory, it would be possible to attempt to estimate the levels of introgression expected under neutrality if the current hybridisation rate has remained constant over time. This way the effect of selection on the total levels of introgression and/or on the number of allelic types that have introgressed could be searched for by comparison of observed with expected levels. In practise, though this is not easy: the pattern of neutral introgression is highly stochastic if F1 hybridisation is rare (Chapter 3). This is because the fate of individual F1 events is highly stochastic, both because of random factors in the reproductive success of individuals and random assortment of introgressing alleles (Chapter 3). We have already established that, under rare hybridisation, it is not appropriate to look for evidence of selection on specific loci, using the differential pattern of introgression across loci (Chapter 3). However, stochastic factor could also act on the patterns of overall levels of introgression of alleles in LE. Provided introgression is neutral, a gradual build-up of small blocks of the red deer genome should be occurring in the sika population (and vice versa) (Baird *et al.* 2003), but it is possible that insufficient numbers of hybridisation events have occurred for this to translate into an overall increase in introgressed allele frequency at the specific loci we are observing. After all, we are only observing the effect of hybridisation on 20 loci in a ~2532cM length genome (Slate *et al.* 2002). Most alleles that introgress from one species into another will persist for a few generations before they go extinct even if they are neutral: stochastic simulations of generations of backcrossing at one neutral locus show that 80% of cases of introgression have gone extinct by 10 generations and even after only 3 generations of backcrossing around 60% have gone extinct, although some alleles, by chance, will introgress to high levels (Chapter 3). Attempting to disentangle the possible effects of selection and neutral processes on the overall levels of background introgression requires a modelling approach that could be the subject of a future study, but is beyond the scope of this paper.

4.5.6 Insights from other longitudinal studies

In the case of honeybee Africanization in Texas, USA (Pinto *et al.* 2004; Pinto *et al.* 2005), gene diversity and allelic richness increased rapidly in a feral honeybee population within 5 years of it coming into contact with invading *Apis mellifera* (Figure 1, Pinto *et al.* (2005)). The formation of a hybrid swarm between native crayfish *Orconectes propinquus* and introduced rusty crayfish *O. rusticus* in Trout Lake, Wisconsin, USA over a 20 year period appears to have occurred for complex reasons: *O. rusticus* is competitively superior to the native species, but hybridisation is slowing the extirpation of *O. propinquus* genes because putative F1 (based on 2 diagnostic allozyme loci) have high fitness and are competitively superior to either parental species (Perry *et al.* 2001). In general, we might expect that positive selection on hybrids, non-assortative mating, high dispersal, and short generation time might all be factors that would result in the rapid formation of a hybrid swarm. However, even species with quite long generation times can form hybrid swarms rapidly: Natural immigration of a small number (~24) of yellow baboons (*Papio cynocephalus*) into a group of anubis baboons (*P. anubis*) in Amboseli National Park, Kenya has resulted in a shift from 0% to 31.3% hybrid composition of the population in the space of 36 years (Tung *et al*. 2008). Hybridisation was initiated through immigration of *P. cynocephalus* males although a number of females of this species also arrived in the 1980s. This lends support to the idea that the hybrid swarm at West Loch Awe (43% hybrids) could have formed within the space of 20- 30 years (see above) as a result of immigration of a small number of sika males into the area (Chapter 2). Tung *et al* (2008) also observe that whilst the number of hybrids had increased over the time of the study, the average hybrid score has been decreasing. One of the explanations they give for this is the stochastic nature of the influx of *P. cynocephalus* into the study area because of a large stretch of waterless land which separates the two populations. When F1 hybridisation rates are low, the progress of hybridisation appears to be unpredictable on the evolutionary short timescales in which longitudinal studies such as this one are attempting to observe change.

4.5.7 Synthesis

The conclusion that we draw from this study supports previous findings (Chapter 2) that the structure of the hybrid zone is dominated by the stochastic timing of hybridisation events. Hybridisation seems to occur anywhere where the two species come into contact. Since mating is predominantly between sika stags and red deer hinds, even red deer populations on the leading edge of sika expansion are likely to become hybridised because sika males range in advance of females. In fact, it is possible that red deer populations in which sika have recently arrived are more prone to hybridisation precisely because sika females are absent. We see in this study that the populations of red deer on the far edges of the sika range all show signs of recent introgression and the predominantly red deer population at Eredine appears to have become hybridised within the course of this study. The pattern of introgression also suggests that F1 hybridisation is occurring across the peninsula, and not just at West Loch Awe. Hybridisation can lead to extensive introgression as at West Loch Awe, but across the rest of the peninsula, although introgression into both species is pervasive, it seems to be at a fairly low level considering the time the two species have been in contact. It is difficult to tell whether this is due to an extremely low number of hybridisation events contributing to introgression, or due to negative selection on introgressing alleles. A hybrid swarm like that at West Loch Awe appears to be able to form within a short space of time (5-8 generations). The most likely explanation seems to be that if a number of hybridisation events coincide within a small area, this leads to a localised area of animals in which assortative mating begins to breakdown, resulting in collapse into a hybrid swarm via positive feedback. Elsewhere, linkage disequilibrium fluctuates with the timing of F1 events: isolated hybridisation events give rise to a chain of backcrossed animals that gradually disperse the foreign genes throughout the population. However, even if negative selection is acting on some loci over time, there should be a gradual build of introgression at neutral and positively selected loci. This in turn might lead to an increase in the hybridisation rate because of a breakdown in assortative mating. Migration of hybridised

animals will also diffuse introgressed alleles throughout the populations. In this study we have not sampled the entirety of the peninsula and it is possible that sites adjoining West Loch Awe are experiencing increased introgression as a result of the presence of the hybrid swarm. At sites closest to West Loch Awe (Oban, Ormaig and Eredine) the red deer populations all show signs of linkage disequilibrium. In the long run (over thousands of years) we expect that the structure of these hybridising populations will be shaped by selection. If there is negative selection against introgressing alleles, then the two populations may form a broad hybrid zone in which the two species remain distinct at many loci (and phenotypic traits), but this study suggests this is unlikely to occur without extensive gene exchange between the two.

4.6 Supplementary material

4.6.1 Amalgamation of databases

During allele scoring, alleles were not given the same sizes using the polyacrylamide gel electrophoresis and capillary array sequencing methods. A shift in peak size between different genotyping methods is a common phenomenon (Taubert & Bradley 2008). In order to translate allele sizes obtained using the polyacrylamide scoring method (1996-7 dataset) set to capillary array sequencing sizes, a panel of 95 samples from the 1996-7 data set was re-genotyped using the multiplex PCR and capillary array genotyping methods used on the 1991-2 and 2006-7 datasets (Chapter 2). A panel of animals was chosen to represent as many allele sizes as possible, although it was not possible to include a few rare alleles. Corresponding genotypes were matched up and translation was carried out by eye. Usually, a constant shift in allele size was observed between the two methods. For example, all allele peaks at locus BM6438 were 2 base pairs longer using the polyacrylamide method. At some loci, the size shift changed in magnitude across the range of allele sizes. Translation of allele

sizes was generally unambiguous. At three loci, however, some allele peaks were combined into one bin because of inconsistent sizing between the two methods. At TGLA126 peaks 100 and 101 (105 and 106 on polyacrylamide) were binned into one bin called '100'. At UWCA47 peaks 229 and 231 (only ever 233 on polyacrylamide) were binned into '231'. At BM757 all alleles with a size greater than 198 (greater than 201 on polyacrylamide) were binned into '198'. In all three cases, previously observed allele frequencies at these loci designated binned alleles unambiguously from the same taxon of origin (Chapter 2). Therefore, this data processing should not affect the detection of hybridisation. With a small number of rare alleles, a guess had to be made at the size they would translate to; this should not affect the quality of the data since these alleles were only present in the 1996-7 dataset. All data sets had originally also been typed at two additional loci (FSHB and RME25), but these were dropped for this comparative study as allele sizes did not translate consistently. Error rates using the capillary array sequencer method were found to be low (0.3-0.8%) across all loci in the 2006-7 dataset (Chapter 2). Once the 1996-7 data set had been translated, allele frequencies in this data set were compared to those in the 2006-7 dataset to ensure no major error had occurred in the translation process.

4.6.2 Supplementary tables

Table 1s: Diversity indices

Locus	N_r	K_r	H_{Or}	H_{Er}	$Null_r$	N_s	K_s	H_{Os}	H_{Es}	$Null_s$
AGLA293	957	3	0.327	0.449	**0.131**	532	3	0.06	0.085	0.005
BM4006	966	4	0.312	0.32	0.006	531	2	0.117	0.128	0.018
BM6438	958	4	0.543	0.594	0.034	526	3	0.392	0.408	0.031
BM757	969	11*	0.676	0.71	0.014	530	4	0.074	0.088	0.002
BOVIRBP	927	8	0.648	0.725	**0.049**	526	2	0.053	0.072	0.003
FCB193	950	15	0.78	0.816	0.013	531	2	0.047	0.059	0.001
IDVGA29	963	3	0.396	0.458	0.039	532	2	0.041	0.075	0.003
IDVGA55	968	13	0.743	0.767	0.003	532	4	0.182	0.215	0.005
INRA5	938	2	0.022	0.028	**0.044**	530	2	0.072	0.09	0.003
INRA6	967	6	0.455	0.495	0.039	531	3	0.051	0.073	0.002
INRA131	955	6	0.595	0.622	0.005	530	2	0.049	0.061	0.001
MM12	966	4	0.381	0.431	**0.041**	531	3	0.066	0.087	0.002
RM12	965	12	0.773	0.814	0.011	532	3	0.06	0.075	0.001
RM188	942	13	0.73	0.737	0.005	515	6	0.526	0.579	0.004
RM95	966	11	0.777	0.807	0.007	531	4	0.041	0.056	0.001
TGLA40	960	5	0.575	0.624	0.029	517	6	0.077	0.097	0.005
TGLA126	952	2	0.014	0.014	0.013	514	2*	0.06	0.125	**0.087**
TGLA127	969	14	0.69	0.705	0.016	532	3	0.18	0.216	0.003
TGLA337	932	11	0.678	0.801	**0.083**	526	4	0.148	0.17	0.033
UWCA47	969	2*	0.025	0.047	0.027	532	1	0.015	0.037	0.002

Nr and Ns, number of samples typed at each locus in phenotypic red and sika; kr and ks, number of alleles considered to be in the red and sika populations according to the STRUCTURE analysis, these values were obtained from the allele frequency data; HO r and HO s , observed heterozygosity in phenotypic red (n= 970) and sika (n=532) ; HE r and HE s, expected heterozygosity in phenotypic red and sika; Nullr, estimated null allele frequency in red; Nulls, estimated null allele frequency in sika. Both the null estimates were obtained with STRUCTURE 2.2. Values in bold indicate null frequency > 0.04; Heterozygosity estimates were calculated with CERVUS (Kalinowski *et al.* 2007). * In these cases some alleles were binned.

Table 2s: Test statistics (and significance prior to Bonferroni correction) for one-sided Wilcoxon's signed rank test comparing the frequency of red alleles in sika deer (Q<0.5) between temporal datasets. (Ho= no increase in gene diversity over time).

	1991/2 vs 1996/7	***1996/7 vs 2006/7***	***1991/2 vs 2006/7***
1 Lussa		95 (0.661)	
2 Carradale	65 (0.615)	32 (0.0585)	48 (0.093)
4 Achaglachach		58 (0.991)	
5 Knapdale	40 (0.226)	63 (0.571)	45 (0.213)
6 Kilmichael		23(0.527)	
8 West Loch Awe		135.5 (0.999)	

Table 3s: Test statistics (and significance prior to Bonferroni correction) for one-sided Wilcoxon's signed rank test comparing the frequency of sika alleles in red deer (Q>0.5) between temporal datasets. (Ho= no increase in gene diversity over time).

	1991/2 vs 1996/7	**1996/7 vs 2006/7**	**1991/2 vs 2006/7**
5 Knapdale		29 (0.072)	
6 Kilmichael		15 (0.999)	
8 West Loch Awe		113 (0.764)	
10 Eredine	0 (0.002)	36 (0.951)	0 (0.063)
11 Shira		2 (0.375)	
14 Glen Lochy		10 (0.844)	
16 Glen Orchy		5(0.375)	
20 Ballachulish		7 (0.813)	

Table 4s: The minimum number and most likely number of F1 hybridisation events required to produce the observed pattern of red allele introgression into sika at each locus across the whole study area excluding West Loch Awe (values including West Loch Awe in brackets)

Locus*	**Minimum number of F1 hybridisation events****	**Most likely number of F1 hybridisation events*****
AGLA293	2	4+
BM4006	1	1-4
BOVIRBP	2	2-4
FCB193	2 (5)	2-3 (7-11)
IDVGA29	2	2+
IDVGA55	3	3-6
INRA6	2	2-17
INRA131	2	2-5
MM12	2	2+
RM12	3	3-5
RM95	3 (4)	3-5 (5-8)
RM188	3 (5)	3-6 (9-14)
TGLA127	2	2-4
TGLA337	2	2-3
TGLA40	3	5+

* Loci BM6438, BM757, UWCA47, INRA5, TGLA126 excluded because red alleles were binned at this locus or because only one red allele was found.

**Minimum number of F1 hybridisation events is the number of red alleles found in sika

***Most likely number of hybridisation events is calculated from the simulation results (see supplementary material Figure5s). The range is obtained by reading off the number of alleles sampled (x-axis) using values 0.5 above and below the number of introgressed red alleles sizes found (on the y-axis).

4.6.3 Supplementary figures

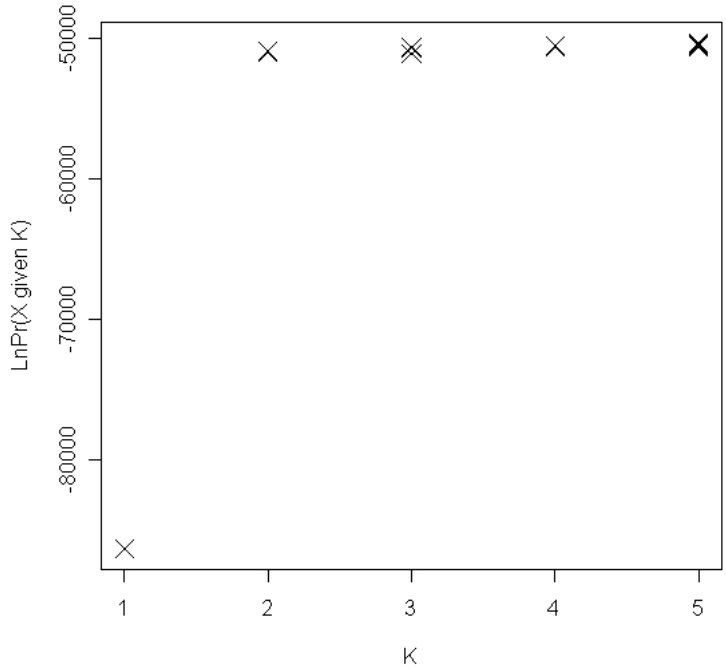


Figure 1s: The estimated probability of the data given K (ln (Pr (X | K))) in 5 independent runs of the simulation over the range of K=1-5. The model was run using a burn-in of 5x104 followed by 106 Markov chain Monte Carlo (MCMC) steps, under the admixed ancestry model and the model of correlated allele frequency, using the option for estimating null allele frequency at all loci. Note that in some cases, a few data points lie on top of each other.

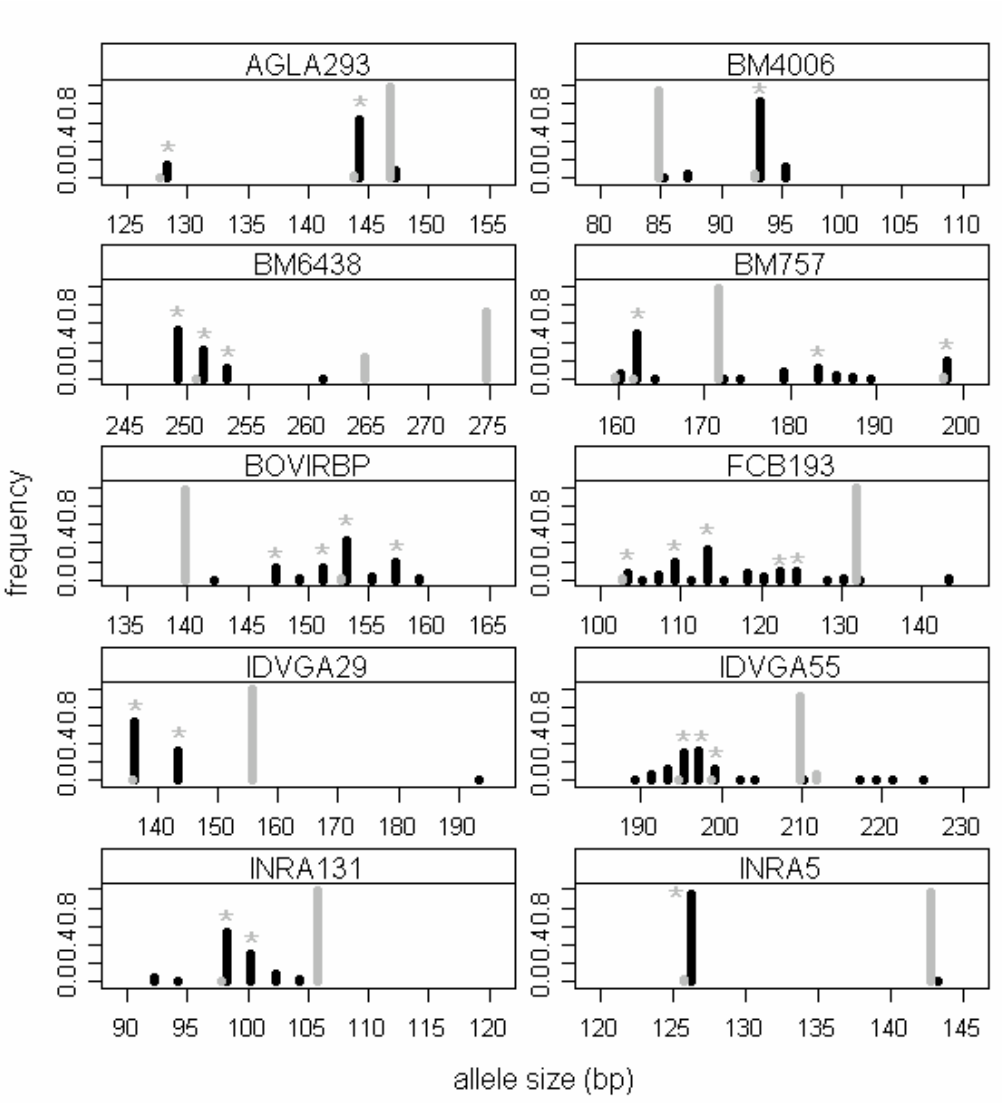


Figure 2s: (see over)

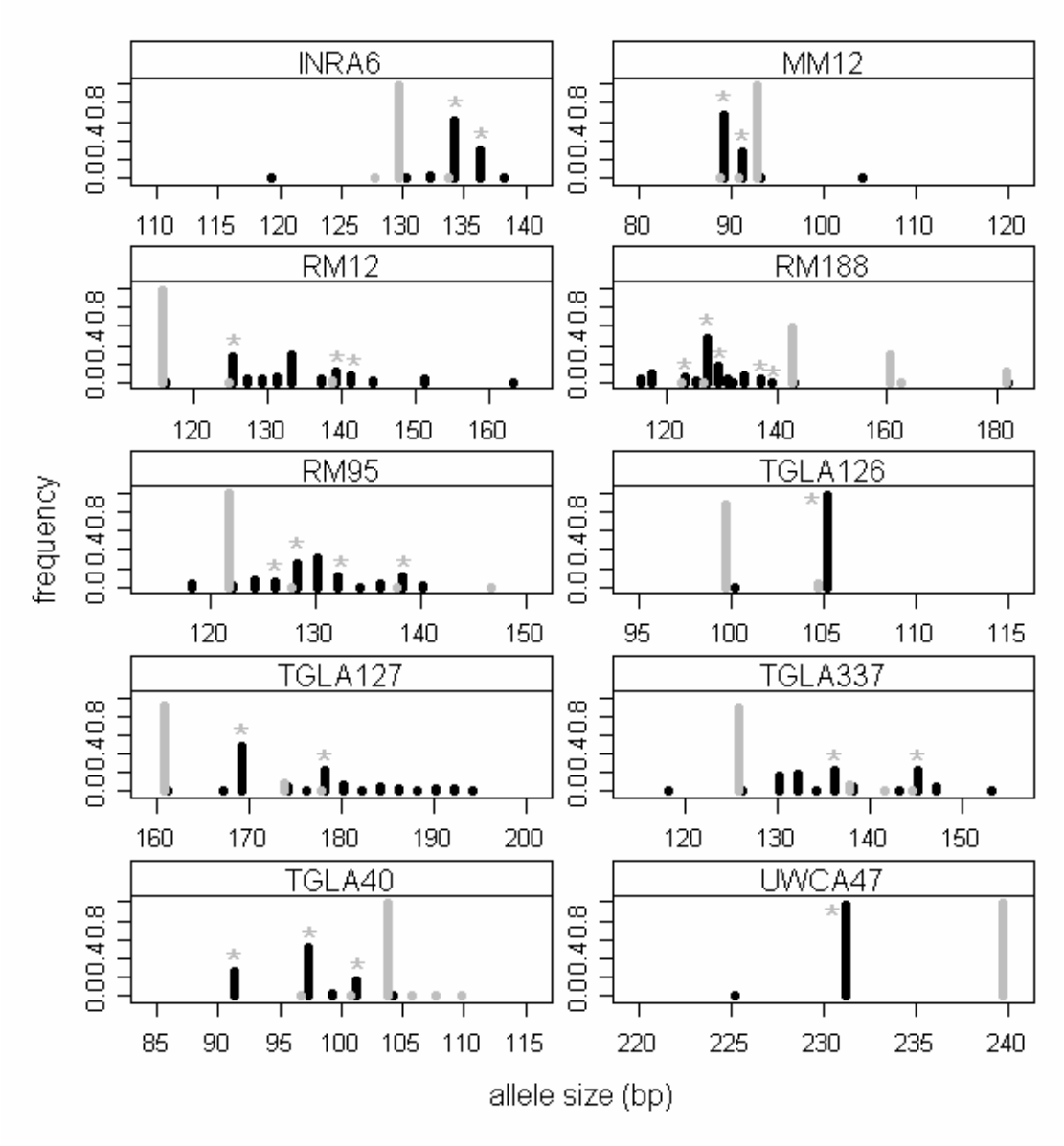


Figure 2s: Estimated allele frequencies for all loci in parental sika (in grey) and red (in black) populations as estimated by STRUCTURE. Bars with asterisk (*) above them indicate alleles that have probably introgressed from red into sika. Some asterisked alleles have been estimated to be present in both parental populations by the STRUCTURE analysis. These are, however, likely to be introgressed due to their low frequency in sika and the size of the alleles (see text).

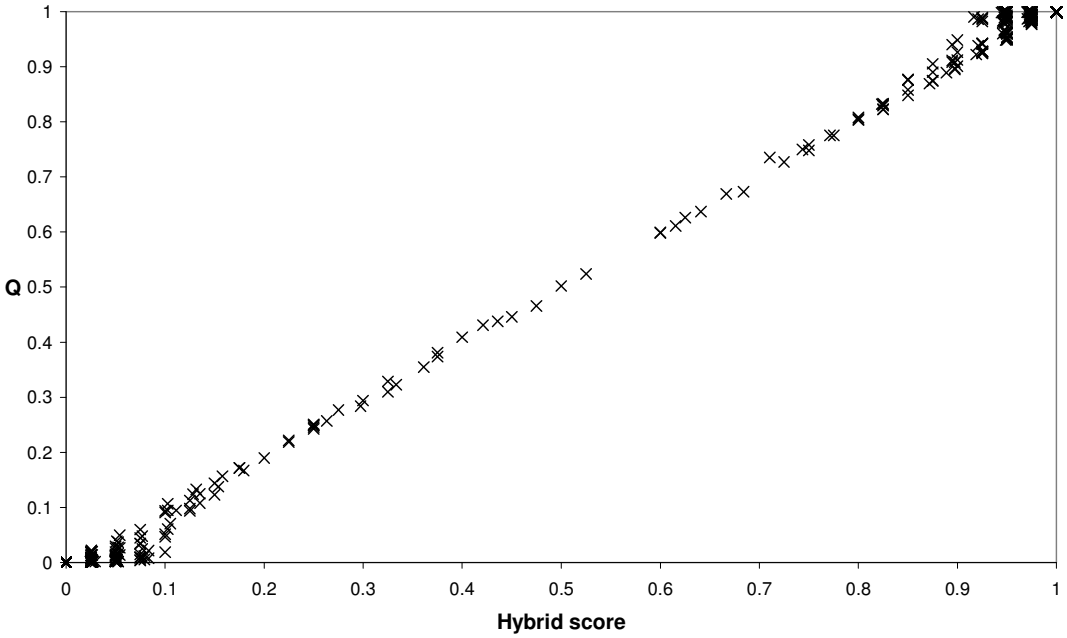


Figure 3s: STRUCTURE's Q plotted against a hybrid score generated using allele assignments. There is some disagreement between the scores close to Q=0 and Q=1, probably because a number of red alleles are close to LE across the whole dataset and so a hybrid score based on assigning alleles inflates the estimate of hybridisation slightly; this should not affect any of the results in the analysis of number of F1's. Two alleles could not be assigned to red or sika populations (TGLA127-174 & TGLA337-138), so any individuals carrying these alleles simply had that allele omitted from the calculations of hybrid score.

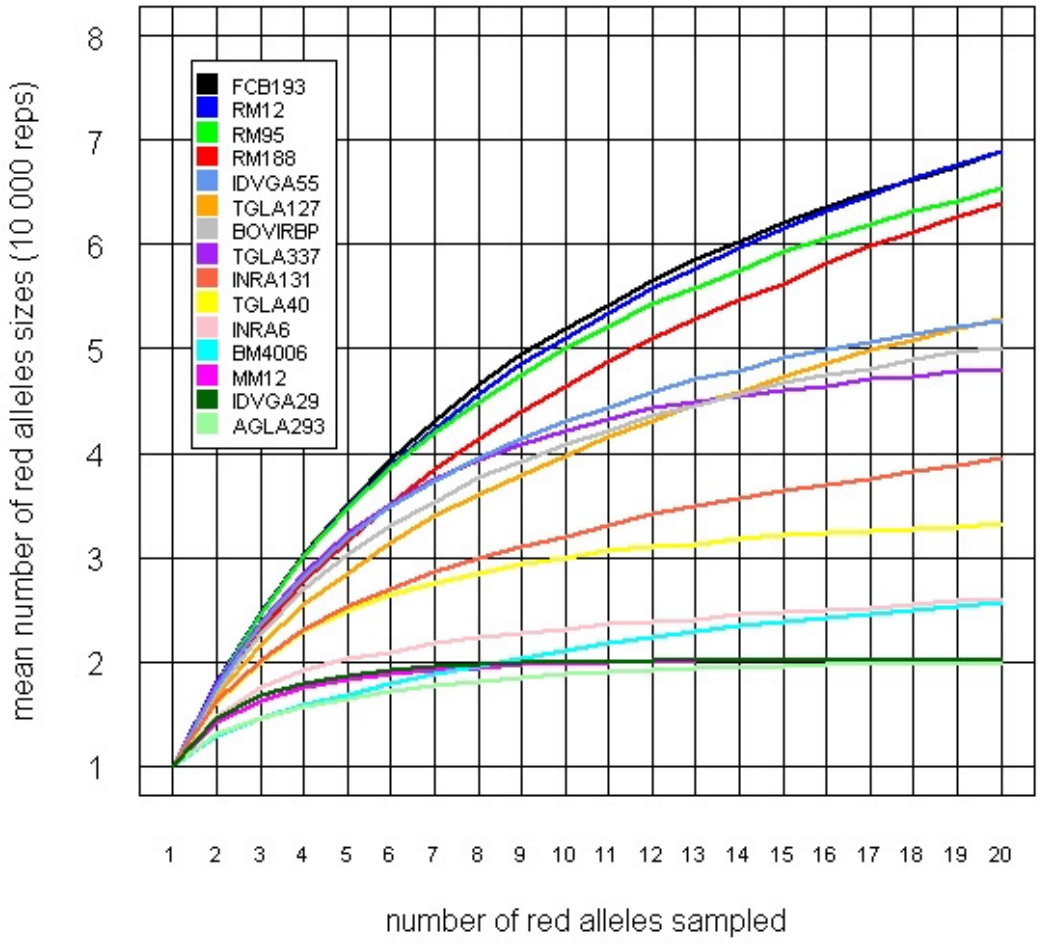


Figure 4s: Mean number of allele sizes picked in 10 000 replicates of sampling with replacement from distributions of red allele frequencies. The number of alleles picked in each sampling event was varied from 1-20. Each allele picked represents one F1 hybridisation event. Sampling was carried out from distributions of alleles at various loci (see key). Loci at which alleles had been binned (BM6438, TGLA126, UWCA47 and BM757) or in which only one red allele existed (INRA5) were excluded from this analysis. For ease of viewing, despite the y-axis being categorical, the results at each locus have been joined up into a continuous line. The loci in the key are ordered in the same way as the lines on the right hand side of the graph. For example, if we find 4 red alleles introgressed into sika FCB193, this would imply an average of successful 6 F1 hybridisation events had lead to introgression.

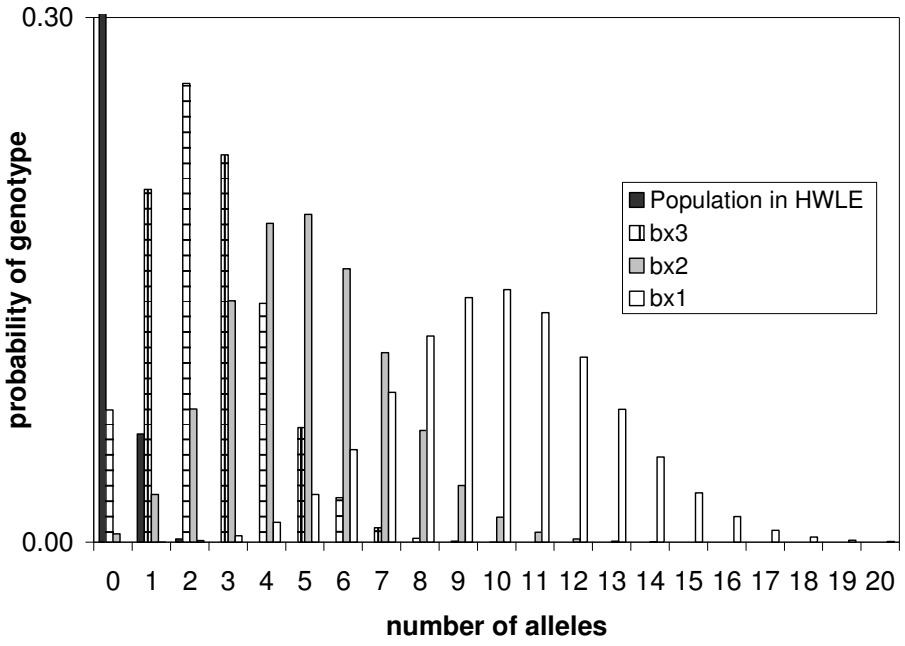


Figure 5s: The probability distribution of the number of red alleles at 20 diploid marker loci in successive backcrosses into sika. Reading right to left, the graph shows the distribution of the number of introgressed alleles found in Bx1 individuals (averaging ¼, or 10 red alleles), Bx2 (averaging 1/8, or 5 red alleles) and Bx3 (averaging 1/16, or 2.5 red alleles). The left most distribution is for a population at Hardy-Weinberg and LE with an allele frequency of introgressed red alleles of p=1.65% (average introgression of red alleles in sika). Individuals carrying 4 or more introgressed alleles are very unlikely to come from a population in Hardy-Weinberg LE and are approximately equally likely to come from a Bx2 or a Bx3. These probability distributions are simply generated from the binomial distribution where n= 2(number of loci) and p= the average proportion of introgressed alleles.

4.7 References

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

Baird SJE, Barton NH, Etheridge AM (2003) The distribution of surviving blocks of an ancestral genome. *Theoretical Population Biology* **64**, 451-471.

Barton NH, Gale KS (1993) Genetic analysis of hybrid zones. In: *Hybrid zones and the evolutionary process* (ed. Harrison RG), Oxford University Press, Oxford.

Barton NH, Hewitt GM (1981) A chromosomal cline in the grasshopper *Podisma pedestris. Evolution* **35**, 1008-1018.

Blum MJ (2002) Rapid movement of a *Heliconius* hybrid zone: Evidence for phase III of Wright's shifting balance theory? *Evolution* **56**, 1992-1998.

Britch SC, Cain ML, Howard DJ (2001) Spatio-temporal dynamics of the *Allonemobius fasciatus - A. socius* mosaic hybrid zone: a 14-year perspective. *Molecular Ecology* **10**, 627-638.

Buggs RJA (2007) Empirical study of hybrid zone movement. *Heredity* **99**, 301-312.

Butlin RK, Hewitt GM (1985) A hybrid zone between *Chorthippus parallelus parallelus* and *Chorthippus parallelus erythropus* (Orthoptera: Acrididae): behavioural characters. *Biological Journal of the Linnean Society* **26**, 287-299.

Cook CE (1993) *Historical Biogeography and Phylogeny of Sika Deer (Cervus nippon) in East Asia*, unpublished PhD thesis, University of California

Dasmahapatra KK, Blum MJ, Aiello A*, et al.* (2002) Inferences from a rapidly moving hybrid zone. *Evolution* **56**, 741-753.

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.

Goodman SJ, Barton NH, Swanson G, Abernethy K, Pemberton JM (1999) Introgression through rare hybridisation: a genetic study of a hybrid zone between red and sika deer (genus *Cervus*), in Argyll, Scotland. *Genetics* **152**, 355-371.

Hairston NG, Wiley RH, Smith CK, Kneidel KA (1992) The dynamics of 2 hybrid zones in Appalachian salamanders of the genus *plethodon*. *Evolution* **46**, 930-938.

Haldane JBS (1948) The theory of a cline. *J. Genet.* **48**, 277-284.

Herzog S, Harington R (1991) The role of hybridization in the karyotype evolution of deer (Cervidae; Artiodactyla; Mammalia). *Theoretical and Applied Genetics* **82**, 425-429.

Hunt GW, Selander RK (1973) Biochemical genetics of hybridisation in European house mice. *Heredity* **31**, 11-33.

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.

Leache AD, Cole CJ (2007) Hybridization between multiple fence lizard lineages in an ecotone: locally discordant variation in mitochondrial DNA, chromosomes, and morphology. *Molecular Ecology* **16**, 1035-1054.

Martinsen, G. D., Whitham, T. G., Turek, R. J. & Keim, P. 2001. Hybrid populations selectively filter gene introgression between species. *Evolution* **55**: 1325-1335.

Perry WL, Feder JL, Dwyer G, Lodge DM (2001) Hybrid zone dynamics and species replacement between *Orconectes* crayfishes in a northern Wisconsin lake. *Evolution* **55**, 1153-1166.

Pinto MA, Rubink WL, Coulson RN, Patton JC, Johnston JS (2004) Temporal pattern of Africanization in a feral honeybee population from Texas inferred from mitochondrial DNA. *Evolution* **58**, 1047-1055.

Pinto MA, Rubink WL, Patton JC, Coulson RN, Johnston JS (2005) Africanization in the United States: Replacement of feral European honeybees (*Apis mellifera* L.) by an African hybrid swarm. *Genetics* **170**, 1653-1665.
Pritchard JK, Wen X, Falush D (2007) Documentation for structure software: Version 2.2.
Randi E, Pierpaoli M, Beaumont M, Ragni B, Sforzi A (2001) Genetic identification of wild and domestic cats (*Felis silvestris*) and their hybrids using Bayesian clustering methods. *Molecular Biology And Evolution* **18**, 1679-1693.
Ratcliffe PR (1987) Distribution and current status of Sika Deer, *Cervus nippon*, in Great Britain. *Mammal Review* **17**, 39-58.
Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual* Cold Spring Harbour Laboratory Press, New York.
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.
Slate J, Van Stijn TC, Anderson RM*, et al.* (2002) A deer (subfamily cervinae) genetic linkage map and the evolution of ruminant Genomes. *Genetics* **160**, 1587-1597.
Slatkin M (1973) Gene flow and selection in a cline. *Genetics* **75**, 733-756.
Sorensen BS, Kiaer LP, Jorgensen RB, Hauser TP (2007) The temporal development in a hybridizing population of wild and cultivated chicory (*Cichorium intybus* L.). *Molecular Ecology* **16**, 3292-3298.
Strayer DL, Eviner VT, Jeschke JM, Pace ML (2006) Understanding the long-term effects of species invasions. *Trends in Ecology & Evolution* **21**, 645-651.
Szymura JM, Barton NH (1986) Genetic analysis of a hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*, near Cracow in Southern Poland. *Evolution* **40**, 1141-1159.
Szymura JM, Barton NH (1991) The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. *Evolution* **45**, 237-261.
Taubert H, Bradley DG (2008) COMBI.PL: computer program to combine data sets with inconsistent microsatellite marker allele size information. *Molecular Ecology Resources* **8**, 572-574.
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.
Walsh PS, Metzger DA, Higuchi R (1991) Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* **10**, 506-513.
Whitlock MC, Ingvarsson PK, Hatfield T (2000) Local drift load and the heterosis of interconnected populations. *Heredity* **84**, 452-457.

Chapter 5

Phenotypic correlates of hybridisation between red and sika deer (genus *Cervus*).

H. V. Senn[1,2], G. M. Swanson[3] , S. J. Goodman[4] , N. H. Barton[5] & J. M. Pemberton[1].

[1] Institute of Evolutionary Biology, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JT, UK

[2] Macaulay Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK

[3] Christchurch, NZ

[4]The Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

[5] Institute of Science and Technology, Am Campus 1, Klosterneuburg, Austria A3400

Authors' Contributions:

HVS collected genetic data, performed the statistical analysis and wrote the manuscript. GMS collected phenotypic measurements and commented on the manuscript. SJG collected genetic data. NHB and JMP supervised the project and commented on multiple drafts of the manuscript.

5.1 Abstract

1. Hybridisation with an invasive species has the potential to alter the phenotype and hence the ecology of the native counterpart.
2. Here we use data from populations of native red deer (*Cervus elaphus*) and invasive sika deer (*Cervus nippon*) in Scotland to assess the extent to which hybridisation between them is causing phenotypic change. We do this by regression of phenotypic traits against genetic hybrid scores.
3. Hybridisation is causing changes in the body weight of sika-like deer and red-like females; it is causing changes in jaw length in the sika-like female population and is causing changes in incisor arcade breadth in females of both populations.
4. There is no evidence that hybridisation is causing changes in the kidney fat or pregnancy rates of either population.
5. Increased phenotypic similarity between the two species is likely to lead to further hybridisation. The ecological consequences of this are difficult to predict.

Key-words: hybridisation, invasive species, phenotypic variation

Running title: Phenotypic correlates of hybridisation

5.2 Introduction

Hybridisation blurs the boundaries between native and introduced species. A large number of factors affect the success and spread of an alien, including its mode of dispersal, reproductive strategy, capacity for phenotypic plasticity, ability to utilise local resources, its response to competitors, the response of endemic predators, the presence of indirect biotic effects (e.g. pathogens) and indirect abiotic effects (e.g. climatic fluctuations) and so forth (for example reviewed in Hastings *et al.* (2005) and White *et al.* (2006)). The presence of a native counterpart with which the invader hybridises adds to the complexity of the situation because interspecific gene flow is accompanied by the possibility of the exchange of selectively advantageous genes between native and invader, the break up of locally adapted genotypes in the native species causing loss of fitness, or the complete merger of the two species into a hybrid swarm (Rhymer and Simberloff 1996). Under hybridisation, competition between the invader and native occurs not just at the level of the species, but also at the level of the gene.

If the native and invader have distinct phenotypes, their hybrids will have a mixture of genes derived from the two parental populations and hence, a mixture of phenotypes. The existence of intermediate hybrid phenotypes may facilitate further gene flow between the two populations creating a positive feedback that ultimately results in complete merger of the two populations. On the other hand, if a phenotypic trait is under divergent selection, so that hybrid phenotypes have lower fitness than either parental type populations, then negative selection against an intermediate phenotype will stem gene flow between the two.

Here we investigate the role hybridisation might have on the future ecology of a native species and an already successful invader. Sika deer (*Cervus nippon*) were introduced to

Britain at numerous locations around 1900 (Ratcliffe 1987) and populations have successfully expanded, especially in Scotland where sika now exist across around 40% of the country (Ward 2005). Originally from Japan, sika are strongly genetically differentiated from the native red deer (*C. elaphus*) with which they hybridise (Goodman *et al.* 1999). The two species differ markedly in appearance (Chapter 2, Table 1). For example, red deer are larger than sika, typically around 30cm higher at the shoulder and whilst red stags can grow antlers of 12 points or more, sika antlers rarely exceed 8 points.

Since hybridisation between red deer and sika deer is likely to result in increasing phenotypic similarity of the two populations, various consequences of hybridisation can be theorised. Hybridisation of the two species could alter their nutritional ecology because of changes in body size and dentition (Bell 1971; Geist 1974; Gordon & Illius 1988; Jarman 1974). Additionally, hybridisation is likely to have management consequences, because in the absence of any substantial predation, deer populations require considerable management effort to maintain numbers at levels acceptable for agriculture, forestry and conservation (Clutton-Brock & Albon 1992; Cote *et al.* 2004). Sika deer have higher fertility rates than red deer (Chadwick *et al.* 1996); are thought to be highly destructive to forestry (McLean 1993; Husheer *et al.* 2006; Swanson and Putman 2008); anecdotally, they are harder to control, because of high alertness, preference for dense forest habitat and propensity to change their behaviour in response to culling pressure (Ratcliffe 1987; McLean 1993; Bartos *et al.* 1998; Husheer *et al.* 2006) and in Scotland they also show high resistance to infection by *Elaphostrongylus spp*., lungworms common in red deer (Bohm *et al.* 2006). Changes in the appearance of red deer due to hybridisation with sika could also lead to a decrease in the trophy value of red deer to the stalking industry. In the long run, of course, it is possible that sika could out-compete red deer completely or that the two species could merge into a hybrid swarm across their range. However, even if the species coexist in a form

close to their original state, hybridisation may still result in the introgression of some "sika-traits" into red (and *vice versa*) if they are selectively advantageous.

The existence of hybrid swarms between red deer and sika deer have been documented in County Wicklow, Ireland (Harrington 1973; McDevitt *et al.* in review) and in the Lake District (Lowe & Gardiner 1974), although neither a recent, nor a genetic analysis of this population exists. However, until recently the populations of mainland British red deer were thought to be relatively free from phenotypic signs of introgression from sika deer, because reports of phenotypic hybrids are rare and because evidence from existing genetic studies indicated that although hybridisation between the two species was occurring, the majority of genetic hybrids found had very low levels of introgression (Goodman et al. 1999; Diaz *et al.* 2006). In previous studies we discovered, through genetic testing of animals culled across the Kintyre Peninsula, at 20-22 microsatellite loci and one mtDNA marker, that a hybrid swarm has existed unreported in one region of the peninsula (to the west of Loch Awe) since at least 1996 (Chapter 2; Chapter 4). Although rangers that culled the deer used in our studies rarely report genetic hybrids as phenotypic hybrids (Chapter 2), the existence of the hybrid swarm on the Kintyre Peninsula raises questions about what impact hybridisation is having or might have on the future ecology of British red deer.

Whilst hybridisation between these two species has been the focus of a number of studies, these have generally focused on either phenotypic (Lowe and Gardiner 1975; Bartos *et al.* 1981; Bartos and Zirovnicky 1982; Putman and Hunt 1993) or genotypic (Abernethy 1994a; Goodman *et al.* 1999; Diaz *et al.* 2006; McDevitt *et al.* in review; Chapter 2; Chapter 4) consequences of hybridisation and have not compared the two (although see Harrington (1979; 1982) for studies combining phenotypic and immuno-electrophoretic measurements). In this study we investigate the association between *Hybrid Scores* calculated at 20 neutral microsatellite markers and various quantitative phenotypic traits measured in a study

population of red deer that is hybridising with Japanese sika deer on the Kintyre Peninsula, Scotland. These traits are: Carcass weight, kidney fat (a measure of condition in deer; Mitchell *et al.* 1976), jaw length (a proxy for skeletal size; Suttie & Mitchell 1983) and incisor arcade breadth. We also examine pregnancy rates. We assume that these traits are under polygenic control (i.e. involving many loci) and investigate whether there is any gross association of hybrid score with phenotype.

Provided that these quantitative traits are under the control of a large number of alleles with additive effects, and there is no significant selection acting on them, then the trait mean is expected to change linearly with hybrid score**.** This expectation is borne out in**,** for example**,** the belly patterning, mating call, and skeletal proportions in a hybrid zone between the toads *Bombina variegata* and *B. bombina* (Nurnberger *et al.* 1995). The reason why we expect to find a relationship between phenotypic traits and neutral nuclear molecular markers (that do not themselves control the traits) is because, in hybrid populations, the gene pools are incompletely mixed. Incomplete mixing means that alleles at molecular markers and alleles controlling quantitative phenotypic traits show covariance (they are in linkage disequilibrium). Covariance of marker alleles and phenotypic traits is strongest in individuals with the most recent hybrid ancestry (the most intermediate hybrids) but decays over subsequent generations as assortment and recombination break up associations between the marker alleles and the alleles controlling phenotype.

In this study we aim to ascertain the relationship between hybridism and phenotype in red-sika hybrids. If hybridisation is resulting in an increased similarity in phenotype between red deer and sika deer, then we expect this not only to lead to further hybridisation as the two species become more similar but also to have long-term ecological consequences - these will be discussed.

5.3 Methods

5.3.1 Study populations

All data used were collected on the Kintyre Peninsula and areas north to Fort William and east to Crianlarich by Forestry Commission Scotland Rangers in the cull seasons 1996-7 and 2006-7. The study population is described in detail in Chapter 2. Phenotypic measurements were taken from culled animals by the rangers or by GMS. Hybrid scores were obtained by analysis of tissue samples collected from the animals by the rangers as detailed in Chapter 2. Briefly, each sample was analysed at 20 highly differentiated microsatellite loci and individual hybrid scores (Q) were assigned using the Bayesian clustering program Structure 2.2 (Pritchard *et al.* 2000; Falush *et al.* 2003). The analysis revealed that approximately 7% of the population could be considered recent hybrids (defined as $0.05 \leq Q \leq 0.95$). No F1 (first generation) hybrids were found**,** and an examination of the genotypes revealed that the majority of recent hybrids were backcrosses (of various generations). This implies that although F1 hybridisation is rare, F1 hybrids are fertile and so, necessarily, mate with individuals from the parental population resulting in backcrossed progeny. However, at one population (West Loch Awe) the pattern of hybridism is different from most of the study area. Here 43% of individuals are hybrids and, because the density of hybrids is high, these individuals are not just backcrosses, but also the progeny from mating between hybrids (see Chapters 2 & 4 for further details). Remaining individuals that are not defined as recent hybrids are either sika-like ($Q<0.05$) or red-like ($Q>0.95$). These individuals are either hybrids with a small portion of introgressed genes and therefore have distant hybrid ancestry or, especially with Q scores close to 1 or 0 may, in fact, have no hybrid ancestors although we cannot be certain of this (Chapter 2). After 5 generations of backcrossing we would expect a hybridised animal to carry 0.625 introgressed markers on average, so we cannot distinguish many advanced backcrossed individuals from non-hybridised individuals. Not

withstanding this, for convenience, we refer to individuals with no apparent signs of introgression (Q=1, Q=0) as ‘pure red’ or ‘pure sika’.

5.3.2 Response variables

Carcass weight - The weight in kilograms of the animal at death following removal of the head, internal organs and lower legs.

Kidney fat - Rangers were asked to remove kidneys with any surrounding fat. In moderate and low condition animals, this task was straightforward as the kidney and fat come away as a single unit. However, in some very good condition deer, the perinephric fat is attached continuously to the fat stores running along the spine and, in these cases, the point at which the kidney broke away from the wall determined the sample. The kidney plus fat was weighed and then the fat was peeled off from the kidney and the kidney was weighed separately to determine kidney fat weight. Kidney fat is a commonly used indicator of condition in ungulates (e.g. Mitchell *et al.* (1976) and Albon *et al.* (1986)).

Jaw length - Jaw length was measured on cleaned jaws from the posterior margin of the alveolus of the fourth incisiform to the process angularis to the nearest mm. The right and left jaw were measured and the mean of the two measurements was taken. Jaw length is a proxy for skeletal size (Suttie & Mitchell 1983).

Incisor arcade breadth- was defined as the distance between the outermost points of the incisiform canines on the arcade and was measured to the nearest 0.1mm. Incisor arcade breadth is thought to dictate diet selectivity in ruminants (Gordon & Illius 1988).

Pregnancy - If the animal was female, then its pregnancy status was recorded as a binary variable (pregnant or not pregnant). This was scored by the rangers following inspection of the uterus.

5.3.3 Explanatory variables

Hybrid score (Q) - 'Q-hat' value generated by Structure 2.2 analysis of microsatellite variation at 20 loci (see above). It is a continuous variable on the scale of 0-1 where 0= pure sika, 1= pure red. For ease, this value will be referred to as Q. Hybridisation can only be interpreted as having a significant effect on phenotype, if hybrid score is significant in the analyses carried out in red-like and sika-like animals separately (see below).

Age - The estimated age in days of the animal at death. This is calculated by combining ranger estimated age of the animal in years (estimated using tooth eruption) with the day of the year shot. Day of year shot is calculated from June 1st (i.e. a birth date of June 1st is assumed for all animals).

Pregnancy - As above

Population - The forestry block in which the animal was shot. The populations are the same as those in Chapter 4.

Sampling - Sampling episode, either 1996/7 or 2006/7.

5.3.4 Statistical analysis

Linear mixed-effect models (LME) were fitted to the data using the Restricted Maximum Likelihood (REML) method implemented in the `nmle` library of the Statistical package R version 2.8.1. Mixed-effect logistic regression for binary response variables was performed in the Hierarchical Generalized Linear Models (HGLM) package of GENSTAT version 11.1. All explanatory variables were centred on their mean prior to inclusion in the analysis, so that in the presence of interactions, the coefficients for linear variables were evaluated at the mean level of the interacting term. All second order interactions between relevant linear terms were fitted in the maximal model. The significance of fixed-effect terms in the model were evaluated through t-statistics for each term. Non-significant fixed effects were excluded from the model in a sequential manner until only those with p-values < 0.05 remained. The

significance of change in log-likelihood (deviance) between the new and old model was evaluated against the χ^2 distribution, at the exclusion of each term. Once a minimal adequate model was obtained, its reliability was double checked through sequential addition of the term to the null model.

Our strategy throughout this paper was first to find the best model that describes the relationship between genotype and phenotype across the dataset of red and sika deer and then to use this model to ask whether there is a significant relationship between hybrid score and phenotype **within** each species separately. Since all response variables investigated here showed a difference in mean value between pure red and pure sika, we expected the effect of hybrid score on the response variables to be significant**,** even if intermediate hybrids were absent from the dataset. In order to investigate the effect of hybrid score on hybrid individuals in the dataset, we refitted the models generated above to a subset of the data. These subsets were either red-like animals (hybrid score > 0.5) or sika-like animal (hybrid score < 0.5). The same terms were kept in each model and the significance of the effect of hybrid score on the response variable was assessed. Crucially, we emphasise again here, hybridisation can only be interpreted to have a significant effect on a response variable if the relationship is found, **within** red and/or sika deer (i.e. within the refitted models).

Model of carcass weight

In order to model the effect of hybridisation on weight, we fitted a model to datasets of males (n= 432, from 20 populations) and females (n= 728, from 20 populations) separately. Weight was log transformed prior to inclusion in the model, because residuals showed heteroscedasticity with weight as a non-transformed response variable. Both datasets contained individuals sampled in 1996/7 and 2006/7. The maximal model fitted included *age*, *hybrid score*, *sampling* and in females *pregnancy* as fixed effects. *Population* was fitted

as a random effect (see above for explanations of covariates). Age was fitted as a linear and quadratic term to account for life time growth. To account for seasonal weight variation sin(x) + cos(x) and sin(2x) + cos(2x) terms for age were also fitted, where x= 2π *age in days/365 (Winfree 1980). Hybrid score was fitted as a linear, quadratic and cubic term to test for non-linear relationship of phenotype with genotype.

Model of kidney fat

The kidney fat dataset consisted of 414 measurements taken from female deer (11 populations, sampled in 1996/7). Prior to inclusion in the model, kidney fat was log-transformed, because it was not normally distributed. The maximal model fitted included *age, hybrid score* and *pregnancy,* in exactly the same way as above. *Population* was fitted as a random effect. We did not fit *carcass weight* as a covariate because it was correlated with age.

Model of jaw length

The jaw length dataset consisted of measurements taken from 410 female deer, sampled from 11 populations in 1996/7 only. The maximal model fitted included *age* and *hybrid score,* in exactly the same way as above. Sin(x) and cos(x) terms for age were not fitted as we do not expect seasonal fluctuations in the jaw length. *Population* was fitted as a random effect.

Model of incisor arcade breadth

The incisor arcade breadth dataset consisted of measurements taken from 424 female deer sampled from 11 populations in 1996/7 only. The maximal model fitted the same way as for jaw length. *Population* was fitted as a random effect.

Model of pregnancy

The pregnancy dataset consisted of 728 individuals from 20 populations, sampled in 1996/7 and 2006/7. Pregnancy was fitted as a binary response variable. In the maximal model, terms for *hybrid score* and *age* were fitted in exactly the same way as for the carcass weight models. *Sampling* episode was also fitted as a factor and *population* was fitted as a random effect. We did not fit carcass weight as a covariate because it is correlated with age.

5.4 Results

5.4.1 Weight

The final model of log weight in males, fitted across the whole dataset, contained hybrid score, age as linear, quadratic and sin(x)+ cos(x) terms and the interaction between age with (sin(x) + cos(x)). In females, the model included hybrid score, a linear, quadratic and sin(x) + cos(x) terms, pregnancy and hybrid score by age interaction. Log weight increased linearly with hybrid score. On the 1st of January, a three year old pure red deer male at West Loch Awe (hybrid score =1) was on average 28.2kg heavier than its pure sika counterpart (hybrid score =0) (Table 1). A non-pregnant pure red female in the same category was on average 21.4 kg heavier than its non-pregnant pure sika female counterpart (Table 2). In females, there was a significant increase in the differential of log weight between red deer and sika deer with age (Table). See Figure 1 for plots of the raw data and fitted models.

Table 1: Estimates of random and fixed effects in a linear mixed model for log male carcass weight (n= 432). All explanatory variables have been centred on their mean. DF(fixed effects) =405

Random effects	*SD*	*Variance component*		
Population	0.096	23.299		
Residual	0.174			
Fixed effects	*Estimate*	*SE*	*t-value*	*p-value*
Intercept	3.572	0.026		
Hybrid Score*	0.657	0.027	24.103	<0.001
Age	5.17×10^{-4}	1.86×10^{-5}	27.713	<0.001
Sin (2πAge/365)	0.108	0.013	8.199	<0.001
Cos (2πAge/365)	0.020	0.014	1.483	0.139
Age^2	-1.58×10^{-7}	1.6×10^{-8}	-9.797	<0.001
Age:Sin (2πAge/365)	5.6×10^{-5}	2.07×10^{-5}	2.722	0.0068
Age:Cos (2πAge/365)	-2.9×10^{-5}	2.62×10^{-5}	-1.111	0.267

*Hybridisation can only be interpreted as having a significant effect on phenotype if hybrid score is significant in the models fitted separately to each species (see Table 7).

Table 2: Estimates of random and fixed effects in a linear mixed model for female carcass weight (n= 728). All explanatory variables have been centred on their mean. DF(fixed effects)=699

Random effects	*SD*	*Variance component*		
Population	0.073	15.735		
Residual	0.169			
Fixed effects	*Estimate*	*SE*	*t-value*	*p-value*
Intercept	3.361	0.023		
Hybrid Score*	0.609	0.020	31.191	<0.001
Age	0.000	1.18×10^{-5}	28.285	<0.001
Sin (2πAge/365)	0.093	0.017	5.506	<0.001
Cos (2πAge/365)	-0.022	0.026	-0.850	0.396
Age^2	1.51×10^{-5}	9×10^{-9}	-17.077	<0.001
Pregnant (Yes)	0.092	0.017	5.463	<0.001
Age:Sin (2πAge/365)	-2×10^{-6}	2.28×10^{-5}	-0.083	0.934
Age:Cos (2πAge/365)	8.9×10^{-5}	3.49×10^{-5}	2.560	0.011
Hybrid Score:Age	5.5×10^{-5}	1.68×10^{-5}	3.282	0.001

*Hybridisation can only be interpreted as having a significant effect on phenotype if hybrid score is significant in the models fitted separately to each species (see Table 7).

When the model was fitted separately to red and sika, hybrid score was significant in sika (males: t= 2.813, p=0.033; females: t=3.090, p=0.002), and red deer females (t=2.382, p=0.018). There was no significant relationship between hybrid score and weight within red deer males (Table 7). The estimates (with standard errors) for the slope of hybrid score in the models in which red deer and sika deer were examined separately fall within the estimates derived from the entire dataset (Compare Tables 1 & 2 with Table 7).

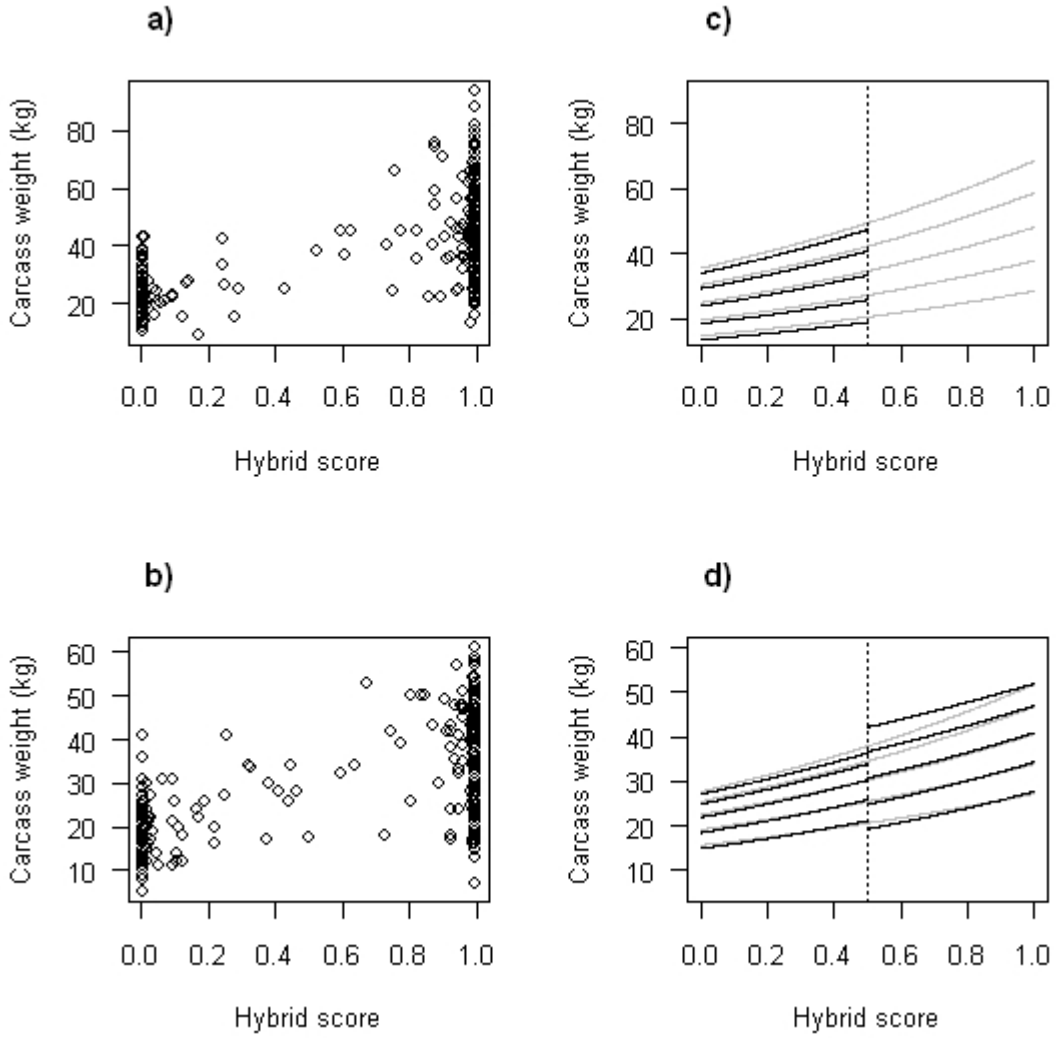


Figure 1: Raw carcass weight data plotted against hybrid score (0= pure sika, 1= pure red) for the dataset of a) males and b) females. Model generated carcass weights estimates for c) males and d) non-pregnant female individuals on 1st September aged 0,1,2,3 and 4 years (bottom to top) at West Loch Awe. Black lines are the predictions from the model fitted to the separate datasets. Only significant relationships are shown. For comparison, grey lines show predictions from the model fitted across the whole dataset.

5.4.2 Kidney fat

The final model of ln (kidney fat), fitted across the whole dataset, contained hybrid score, age as linear, quadratic, sin(x) and cos(x) terms, pregnancy and age by (sin(x)+cos(x)) interactions. In females, log kidney fat decreased linearly with hybrid score (Table 3). On average a pregnant pure sika female from West Loch Awe aged three on the 1st of January is predicted to have 123.39 g of kidney fat, 37.88 g more than a pure red female in the same category. However, hybrid score was not significant in either of the models that were fitted to the red-like or sika-like data separately (Table 7). This indicates that the relationship is being driven by the difference between pure red deer and pure sika and we have no evidence to suggest that the hybrids in this dataset differ significantly in kidney fat from the pure species they are genetically closest to. See Figure 2 for a plot of the raw kidney fat data.

Table 3: Estimates of random and fixed effects in a linear mixed model for female log kidney fat (n=414). All explanatory variables have been centred on their mean. DF(fixed effects)=395

Random effects	*SD*	*Variance component*		
Population	0.187	5.624		
Residual	0.765			
Fixed effects	*Estimate*	*SE*	*t-value*	*p-value*
Intercept	3.633	0.104		
Hybrid Score*	-0.367	0.098	-3.752	<0.001
Age	0.001	7.7×10^{-5}	8.628	<0.001
Sin (2πAge/365)	-0.605	0.388	-1.558	0.120
Cos (2πAge/365)	-0.865	0.237	-3.650	<0.001
Age^2	-2.75×10^{-7}	1×10^{-7}	-5.355	<0.001
Pregnant (yes)	0.573	0.097	5.889	<0.001
Age:Sin(2πAge/365)	4.18×10^{-4}	4.27×10^{-4}	0.978	0.329
Age:Cos(2πAge/365)	0.001	3.32×10^{-4}	2.432	0.016

*Hybridisation can only be interpreted as having a significant effect on phenotype if hybrid score is significant in the models fitted separately to each species (Table 7).

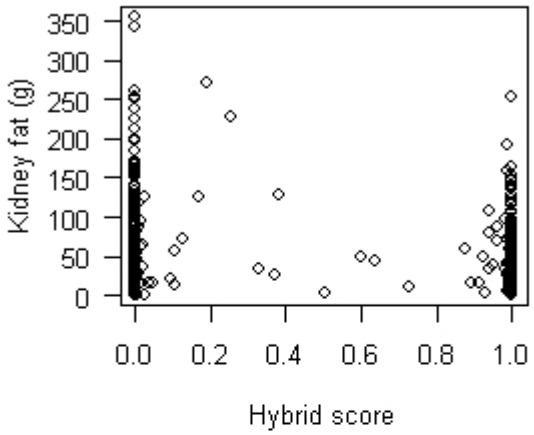


Figure 2: Raw kidney fat data plotted against hybrid score (0= pure sika,1= pure red) for all individuals in the dataset (all are females).

5.4.3 Jaw length

The final model of jaw length contained hybrid score, age as a linear and quadratic term and an age by hybrid score interaction. In females**,** jaw length increased linearly with hybrid score (Figure 3, Table 4). Our models predicted that, on average, a pure sika female aged 3 on the 1st of January had a jaw length of 17.3 cm whilst a red deer in the same category had a jaw length of 23.1cm. A significant effect of hybrid score on jaw length was found when the model was fitted to sika separately ($t= 5.86$, $d.f =136$ $p<0.001$) but the effect was only marginally significant when fitted to red deer ($t=1.94$, $d.f. =169$, $p= 0.055$). We found a highly significant interaction of age with hybrid score, indicating that the pattern of jaw growth varies between the two phenotypic classes. The estimates (with standard errors) for the slope of hybrid score for the jaw length models in which red deer and sika deer were examined separately fall within the estimates derived from the entire dataset (Compare Tables 4 with Table 7).

Table 4: Estimates of random and fixed effects in a linear mixed model for female jaw length (n= 410). All explanatory variables have been centred on their mean. DF (fixed effects)=395.

Random effects	*SD*	*Variance component*		
Population	0.211	3.054		
Residual	1.189			
Fixed effects	*Estimate*	*SE*	*t-value*	*p-value*
Intercept	19.789	0.109		
Hybrid Score*	5.692	0.141	40.348	<0.001
Age	2.48×10^{-3}	9.33×10^{-5}	26.568	<0.001
Age^2	-1×10^{-6}	7×10^{-8}	-14.442	<0.001
Age: Hybrid Score	9.99×10^{-4}	1.38×10^{-4}	7.227	<0.001

*Hybridisation can only be interpreted as having a significant effect on phenotype if hybrid score is significant in the models fitted separately to each species (see Table 7).

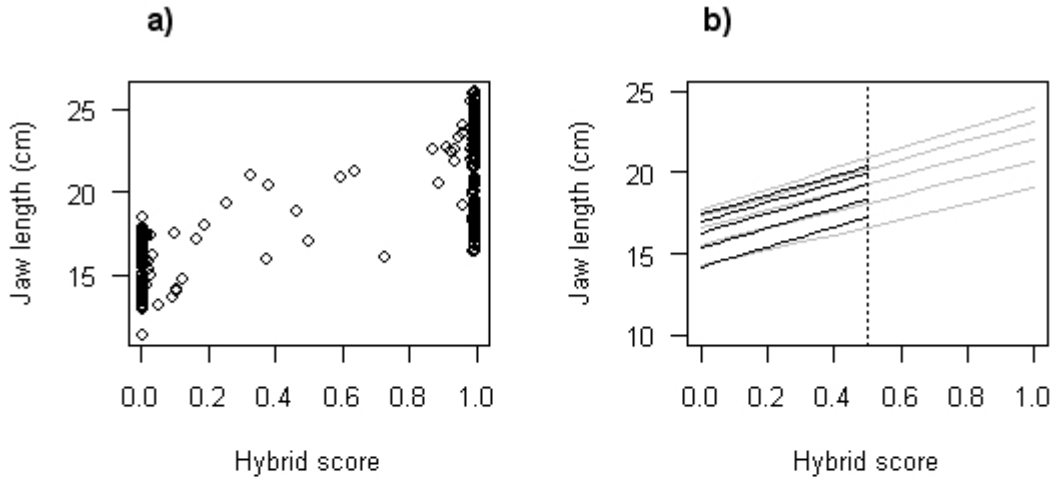


Figure 3: a) Raw data for jaw length by hybrid score (0= pure sika, 1= pure red) for all individuals in the dataset (all females). b) Model generated jaw length estimates for individuals on 1st January aged 0,1,2,3 and 4 years (bottom to top) at West Loch Awe. Black lines are predictions from the models fitted to the separate datasets. Only the significant relationship for the sika-like dataset is shown. For comparison, grey lines show predictions from the model fitted across the whole dataset.

5.4.4 Incisor arcade breadth

The final model of incisor arcade breadth contained hybrid score, age as a linear and quadratic term and an age by hybrid score interaction. In females**,** incisor arcade breadth increased linearly with hybrid score (Table 5; Figure 4). Our models predicted that on average a pure sika female aged 3 on 1st of January had an incisor arcade breadth of 27.3mm whilst a red deer in the same category had an incisor arcade breadth of 40.0 mm. Incisor arcade breadth also varied significantly with hybrid score within red and sika datasets (Table 7). We found a highly significant interaction of age with hybrid score, indicating that the pattern incisor arcade growth varies between the two phenotypic classes. The estimates (with standard errors) for the slope of hybrid score for jaw length in the models in which red deer and sika deer were examined separately do not fall within the estimates derived from the entire dataset (Compare Tables 5 with Table 7). The slope is steeper for red deer (17.940±5.375 SE) and shallower for sika deer (8.560±1.976SE) than when the relationship is fitted across both species (12.372 ±0.279SE).

Table 5: Estimates of random and fixed effects in a linear mixed model for female incisor arcade breadth (n= 424). All explanatory variables have been centred on their mean. DF(fixed effects)= 409.

Random effects	*SD*	*Variance component*		
Population	0.193	0.525		
Residual	2.656			
Fixed effects	*Estimate*	*SE*	*t-value*	*p-value*
Intercept	33.080	0.192		
Hybrid Score*	12.372	0.279	44.391	<0.001
Age	0.005	2.06×10^{-4}	26.605	<0.001
Age^2	2.61×10^{-6}	1.6×10^{-7}	-16.419	<0.001
Age: Hybrid Score	0.002	3.06×10^{-4}	5.921	<0.001

*Hybridisation can only be interpreted as having a significant effect on phenotype if hybrid score is significant in the models fitted separately to each species (Table 7).

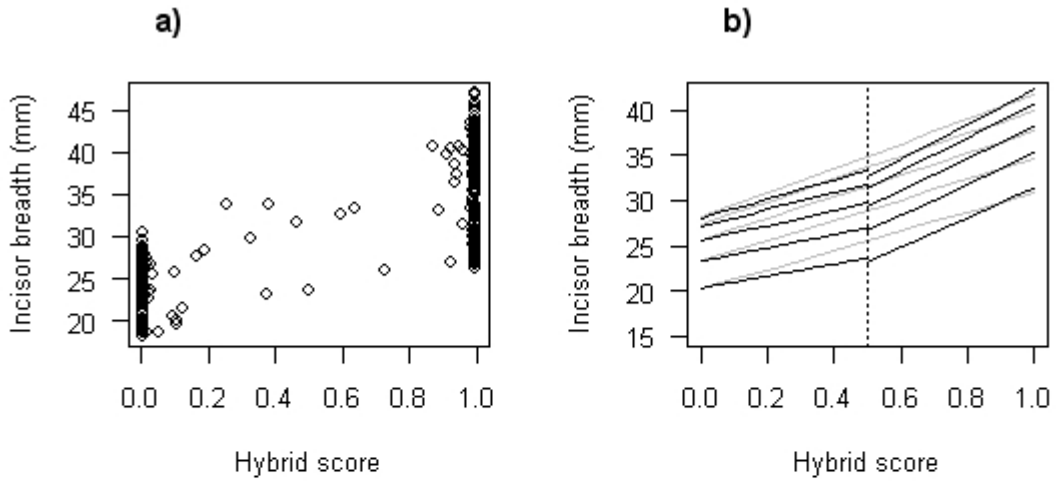


Figure 4: Raw data for incisor arcade breadth (0= pure sika, 1= pure red) for all individuals in the dataset (all are females).b) Model generated incisor arcade breadth estimates for individuals on 1st January aged 0, 1, 2, 3 and 4 years (bottom to top) at West Loch Awe. Black lines are predictions from the model fitted to the separate datasets. Only the significant relationship for the sika-like dataset is shown. For comparison, grey lines show predictions from the model fitted across the whole dataset.

5.4.5 Pregnancy

The final model included hybrid score, age as a linear and quadratic term and sine(x) + cosine(x) terms. There was a significant decrease in the probability of pregnancy with increasing hybrid score (Table 6; Figure 5). However, there was no significant effect of hybrid score when the model was refitted to red and sika datasets separately (Table 7). This indicates that the relationship is being driven by the difference between pure red deer and pure sika and we have no evidence to suggest that the hybrids in this dataset differ significantly in pregnancy rates from the pure species they are genetically closest to.

Table 6: Estimates of random and fixed effects in a linear mixed model with binomial error structure for pregnancy (n= 728). All explanatory variables have been centred on their mean. DF (fixed effects)=722.

Random effects	*Variance component*	*SE*		
Population	0.223	0.161		
Fixed effects	*Estimate*	*SE*	*t-value*	*p-value*
Intercept	0.298	0.178		
Hybrid Score*	-0.762	0.254	-3.002	0.0027
Age	0.002	1.86×10^{-4}	13.233	<0.001
Sin (2πAge/365)	-1.247	0.385	-3.238	0.0012
Cos (2πAge/365)	-2.504	0.298	-8.418	<0.001
Age^2	-1.2×10^{-6}	1.26×10^{-7}	-9.628	<0.001

*Hybridisation can only be interpreted as having a significant effect on phenotype if hybrid score is significant in the models fitted separately to each species (Table 7).

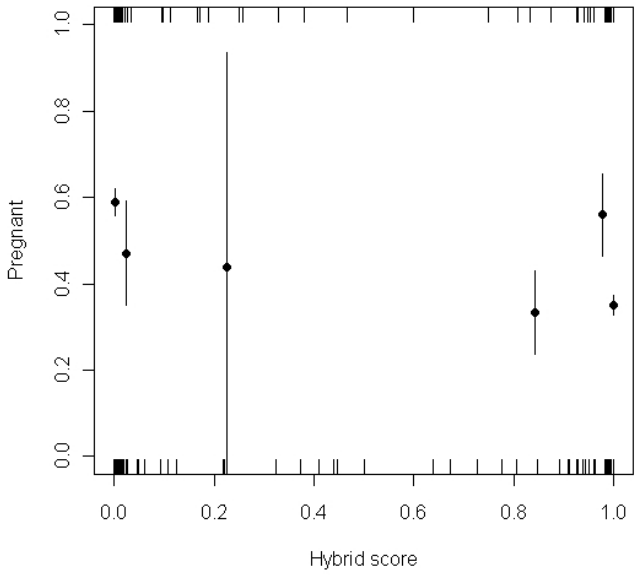


Figure 5: Rug plot of pregnancy status against hybrid score (top and bottom of graph). Additional points are probabilities of pregnancy (with standard errors of a binomial proportion) for average values in the bins 0<Q<0.01, 0.01≤Q<0.05, 0.05≤Q<0.5, 0.5<Q≤0.95, 0.95<Q<1. The sample sizes for each bin were 142, 8, 11, 9, 14 and 138 respectively.

Table 7: Models refitted to data from red (Hybrid score >0. 5) or sika (Hybrid score <0.5) only. All explanatory variables were re-centred on their mean prior to inclusion in the model.

Trait	*n*	*n(Hybrid)**	*Slope for hybrid score*	*DF*	*t-value*	*p-value*
♂Weight (Red)	295	27	-0.201± 0.175SE	269	-1.150	0.251
♀Weight (Red)	445	27	0.463±0.194SE	417	2.382	**0.018**
♂Weight (Sika)	137	15	0.665± 0.308SE	119	2.813	**0.033**
♀Weight (Sika)	283	26	0.635± 0.205SE	264	3.090	**0.002**
♀ Log kidney fat (Red)	213	13	0.913 ± 0.967SE	202	0.945	0.346
♀ Log kidney fat (Sika)	196	13	1.323± 1.043SE	178	1.269	0.206
♀ Jaw length (Red)	181	5	5.161 ± 2.667 SE	169	1.935	0.055
♀ Jaw length (Sika)	149	10	6.381±1.08SE	136	5.863	**<0.001**
♀ Incisor width (Red)	225	12	17.940 ± 5.3754SE	213	3.337	**0.001**
♀ Incisor width (Sika)	199	14	8.560±1.976SE	185	4.333	**<0.001**
Pregnancy (Red)	445	27	1.841 ± 2.733SE	439	0.674	0.500
Pregnancy (Sika)	283	26	-2.812 ± 2.423SE	268	-1.160	0.246

* Number of individuals Q≤0.95 in red deer dataset and number of individuals where Q≥0.05 in sika deer dataset.

5.5 Discussion

5.5.1 Summary

We found that pure red deer from the Kintyre Peninsula had significantly higher weight, jaw length and incisor arcade breadth than pure introduced Japanese sika from the same area. We also found that sika had significantly higher kidney fat than red deer and higher pregnancy rates. Despite these differences in phenotype between red and sika deer, not all measured traits varied significantly with hybrid score within red deer (Q>0.5) and sika deer (Q<0.5). Thus, we found no evidence that hybridisation has led to changes in the amount of female kidney fat or changed pregnancy rates within the two species, but we did find evidence that hybridisation is leading to an increase in the carcass weight of sika-like males and females and to an increase in incisor arcade breadth and jaw length of sika-like females. We also

found evidence that hybridisation is resulting in a decrease in weight and incisor arcade breadth of red deer-like females.

5.5.2 Adequacy of dataset

F1 hybridisation between red deer and sika deer is rare but may be followed by substantial introgression: we find no F1 hybrids in our dataset (n = 1513) despite finding 111 recent hybrids (Chapter 2). In one population (West Loch Awe) 43% of individuals are defined as genetic hybrids ($0.05 \leq Q \leq 0.95$) (Chapter 2), but usually, the proportion of hybrids in the population is much lower (~1%, see Figure 5, Chapter 4). So**,** despite fairly large sample sizes (n = 410-670, or 134-287 for the split analyses) only 2.7-11.1% of individuals in the datasets used were hybrids (Table 7). This has two consequences.

First, in the analyses split by species (Table 7), it is difficult to ascertain whether the lack of relationship found between hybrid score and the phenotypic traits (jaw length, kidney fat and male weight in red deer and kidney fat in sika deer) is due to a genuinely low effect of hybrid score or because sample sizes of hybrids are too small (i.e. a lack of power). For example, for jaw length, it is probable that small sample sizes are at fault: jaw length in red deer varied marginally significantly with hybrid score ($t= 1.935$, $p=0.055$) but the dataset only has 2.7% of individuals considered to be recent hybrids, in comparison to the jaw dataset for sika ($t= 5.863$, $p<0.001$) in which the incidence of recent hybridism is 6.3% (Table 7). But sample sizes *per se*, are not really the main issue – the extent to which the response variable is under genetic as opposed to environmental control (its heritability) is important. The higher the heritability of a trait, the more power there will be to detect the effect of hybridisation in a given sample size. Skeletal measures (such as jaw length and incisor arcade breadth) probably show less environmental variability than condition dependent measures that fluctuate over an individuals life time such as weight, kidney fat and

pregnancy. This may explain why we find more significant relationships with hybrid score and higher levels of significance for the skeletal measures than the other measures (Table 7).

Second, the linear relationship found between the response variables and hybrid score should be viewed as a rough estimate because there may not be enough intermediate hybrids to distinguish between linear and quadratic or higher order relationships. Non-linear relationships might be produced by heterosis, inbreeding depression, dominance or epistatic effects (i.e. non-additive genetic variance) or by strong selection on the trait in question. The large number of individuals with hybrid scores near 0 or 1 will exert a high degree of influence on the fitted relationship and in the extreme case of a dataset of both parental species and no hybrids, there would necessarily be a linear fit through the means of the two parental populations. Log weight and jaw length estimates for hybrid score, where significant, agreed in the separately fitted models with those found for the whole dataset. This lends support to the idea of a linear relationship between the trait and hybrid index. For incisor arcade breadth, the estimates do not agree and this points towards the possibility of a more complex relationship between trait and hybrid index **-** a relationship that may or may not emerge using a larger dataset.

Apart from environmental variation of the phenotypic trait in question, additional variation is introduced because individuals with identical hybrid scores show genetic variation. For example, each 1st generation backcross into red deer possesses a unique random sample of sika genes that make up ~25% of its genome. Unless the phenotypic trait in question is under the additive control of a large number of genes of reasonably similar effects, there will not be a strong correlation between the proportion of genome that is introgressed and genetic effect on phenotype. Additionally, even if the relationship between the proportion of genome introgressed and the genetic effect on phenotype is good, the proxy measure of hybrid score introduces extra variation because it may not accurately reflect the true underlying

proportion of introgressed genotype. The more markers used to generate the hybrid score, the better this relationship is expected to be (Boecklen & Howard 1997; Vaha & Primmer 2006). So it may be hard to detect the average relation between genotype and phenotype and may be misleading even to try to do so. However, because we recover a significant relationship between hybrid scores and some phenotypic traits, despite these objections, this suggests that this approach is reasonable. Other studies have also carried out the regression of hybrid scores against phenotypic traits successfully despite having used smaller numbers of markers (Charpentier *et al.*, 2008; Nurnberger *et al.*, 1995).

5.5.3 Adequacy of models

Kidney fat is commonly used as an indicator of condition in deer (e.g. Mitchell *et al.* (1976) and Albon *et al.* (1986)). However, it is often incorporated into a kidney fat index (KFI) which is obtained by dividing the weight of the kidneys plus fat by the weight of the kidneys (Riney 1955). We did not employ this index here because the use of a ratio as a response variable in which the relationship between the numerator and the denominator is not isometric (linear and passing through the origin) can lead to spurious significant results and KFI has been shown to suffer from this problem (Kronmal 1993; Serrano et al. 2008). Since we might reasonably expect kidney fat to vary with kidney size, we also tested the inclusion of kidney weight as covariate in the model. The inclusion of kidney weight caused a significant increase in deviance (χ^2 =5.2847, p= 0.0215) but did not affect the significance of hybrid score (t=-4.63, p<0.001). This implies that kidney weight explains some additional variation in kidney fat weight, but that the relationship between kidney fat and hybrid score is not just being driven by differences in kidney weight between the two species. Because of a strong correlation of kidney weight with age, and because inclusion of kidney weight as a covariate devalues the predictive power of the model, we dropped kidney weight from the final model and only modelled kidney fat (Table 3).

Out of necessity we fitted pregnancy as a factor when examining changes in weight and kidney fat, but one might expect that how far into gestation a female is will influence both these traits. All females in the dataset were shot between October and April and the majority were shot in February and March. The lack of sampling at the time when no females are expected to be pregnant (summer) may explain the fact that we found no pregnancy by season interaction. Given the period of sampling, we believe it is reasonable to fit pregnancy as a factor**,** since fecundity is known to be positively associated with weight and kidney fat in red deer females (Mitchell & Brown 1974; Albon *et al.* 1986), meaning that pregnant females are likely to be heavier than non-pregnant females even in early pregnancy. We fitted pregnancy in the models of kidney fat and carcass weight, not so much as a predictive tool, but to minimise any potential spurious influences its omission might have upon the interpretation of hybrid score.

5.5.4 Hybridisation in an ecological context

In the long run, the phenotypic and genetic outcome of hybridisation between red deer and sika deer will be determined by selection. Crosses between two species may produce individuals with phenotypic traits that differ significantly from the mean of either parental population, but if these traits confer a fitness disadvantage**,** we do not expect them to alter the trait mean within either population over time, because the introgressing genes responsible will be eliminated by selection. *A priori* we might expect the effect of hybridisation to have different selective outcomes between the two sexes: in red deer, male reproductive success is correlated with absolute body size, antler size and shape and age - bigger, older (but not very old) males with big antlers sire more calves (Clutton-Brock *et al.* 1988; Kruuk *et al.* 2002); in red deer females' reproductive success is correlated with body condition - heavy but skeletally small females have a higher chance of being pregnant (Albon *et al.* 1986). We might therefore expect hybridisation to be selectively disadvantageous for red deer males and sika females, but selectively advantageous for sika males and red deer females

(Abernethy 1994b). In this study we have demonstrated that the extent of hybridisation has an effect on weight in sika and, at least, red deer females. Jaw length, which is a good proxy for skeletal size (Suttie & Mitchell 1983) also varies significantly with the extent of hybridisation within red and sika deer. Since we observe changes in body weight and size within red and sika, it is possible that selection could act on this variation in the directions mentioned above.

Body size is also likely to be an important factor in the nutritional ecology of red and sika deer populations. The Jarman-Bell principle states that relative energy requirements in herbivores decrease with increasing body size (energy requirement~ weight$^{0.75}$), while rumen volume is isometric with size (rumen volume ~ weight). This means that large herbivores should be able to survive on lower-quality diets than smaller ones (Bell 1971; Geist 1974; Jarman 1974). This is thought to be one of the reasons diet selection differs between the sexes of sexually dimorphic ungulates such as red deer e.g. (Clutton-Brock *et al.* 1982; Staines *et al.* 1982; Clutton-Brock *et al.* 1987) and between ungulate species of different sizes (Bell 1971). Red and sika deer on the Kintyre peninsula show differences in diet selection and habitat use both between species and sexes, so hybridisation is also likely to affect competitive interactions between the two species (Abernethy 1994b; Chadwick *et al.* 1996).

A consequence of evolved differences in dietary need due to body size (or evolved differences in body size due to dietary availability) is that dental morphology is likely to come under different selective pressures as food choice moves between grazing of large volumes of low quality material (in large animals) and selective browsing of higher quality material (in smaller animals). In particular, small animals should require dentition (and mouth morphology) that enables them to feed with a higher degree of selectivity than larger animals. Incisor arcade breadth is thought to directly reflect the degree of selectivity an

animal can exert because the incisors are used to cut herbage, so grazing ruminants have wider and flatter incisor arcades than similar sized browsers (Gordon & Illius 1988). In this study we find that variation in incisor arcade breadth is correlated with hybrid score in both red and sika females (Table 7), so we might expect a shift in the selective pressure exerted by their usual habitat on these populations, resulting in changes in fitness or changes in food selection behaviour of the deer. However, a recent comparison of incisor arcade breadth across ungulate species found no evidence for variation in arcade breadth with feeding style, after body size and the effect of phylogeny had been controlled for**,** and suggests that the evolutionary relationship is between body size and selectivity and that incisor arcade breadth varies incidentally with body size and has no additional evolutionary relationship with diet (Perez-Barberia & Gordon 2001).

In general, the long term consequences of hybridisation on phenotype will be a result of a complex interaction between genotype and habitat, ecological competition between the two species, the extent of spread of sika deer and sexually antagonistic selection (see above) against a background of erosion of the genetic differences between the two species at an increasing number of loci. This complex scenario means that selective outcomes are hard to predict, especially because of uncertainty over the causal direction of selective interactions (e.g. will change in dental morphology select for differences in browsing behaviour, or does browsing behaviour select for changes in dental morphology?). In the short term, however, the consequences of hybridisation are fairly clear: although pregnancy rates differ between red deer and sika deer, we find no evidence that hybrids have lower pregnancy rates than either of the pure parent species. We know that F1 hybrids must be fertile at least occasionally because the dataset contains backcrosses and other hybrid classes (Chapters 2 & 4) and here we find no compelling evidence that hybridised females are less likely to be pregnant than the parental species although, of course, the offspring might still be less viable (Tables 6 & 7). Since hybridisation between the two species is occurring and, in one

population, 44% of individuals are hybrids, the potential for extensive gene flow between the two species exists. We have shown here that this gene flow is accompanied by changes in weight, skeletal size and jaw morphology. Evidence from mitochondrial DNA introgression suggests that in the vast majority of cases hybridisation takes place between red deer females and sika deer males (Chapters 2 & 4), so we expect hybridisation to erode the size differential between these pairings, facilitating further hybridisation. In the short term, red and sika deer have become phenotypically more like each other through hybridisation and are expected to become more like each other in other parts of Britain if hybridisation occurs. Over evolutionary time scales we are unable to say whether selection will maintain some of the species differences through the formation of a hybrid zone or species reinforcement, but this is unlikely to occur without substantial genotypic and phenotypic introgression. From a practical point of view, the presence of sika in mainland Britain is likely to alter the ecology and appearance of red deer and further studies involving greater numbers of hybrids and other traits of economic interest (i.e. antler morphology) are merited.

5.6 Acknowledgements

HVS thanks Stephen Senn and Jarrod Hadfield for statistical advice. The authors thank the Forestry Commission Scotland rangers for all their help with providing the samples of red and sika deer and Steve Albon for helpful comments on the manuscript. This project was funded through a NERC studentship to HVS which was case partnered by the Macaulay Institute.

5.7 References

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

Abernethy K (1994b) *The Introduction of Sika Deer (Cervus nippon nippon) to Scotland*. University of Edinburgh, unpublished PhD thesis

Albon SD, Mitchell B, Huby BJ & Brown D (1986) Fertility in female red deer (*Cervus elaphus*) - the effects of body-composition, age and reproductive status. *Journal of Zoology* **209,** 447-460.

Bartos L, Herrmann H, Siler J, Losos S & Mikes J (1998) Variation of mating systems of introduced sika deer. *Revue D Ecologie-La Terre Et La Vi,* **53,** 337-345.

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

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

Bell RHV (1971) Grazing ecosystem in Serengeti. *Scientific American,* **225,** 86-&.

Boecklen WJ, Howard DJ (1997) Genetic analysis of hybrid zones: numbers of markers and power of resolution. *Ecology* **78**, 2611-2616.

Bohm M, White PCL, Daniels MJ, Allcroft DJ, Munro R & Hutchings MR (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.

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.

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.

Clutton-Brock, T. H., Albon, S. D. & Guinness, F. E. (1988) Reproductive success in male and female red deer. *Reproductive Success* (eds TH Clutton-Brock). The University of Chicago Press, Chicago.

Clutton-Brock TH, Guinness FE & Albon SD (1982) *Red Deer: Behaviour and Ecology of Two Sexes,* 1st edn. The University of Chicago Press.

Clutton-Brock TH & Albon SD (1992) Trial and error in the highlands. *Nature* **358,** 11-12.

Clutton-Brock TH, Iason GR, & Guinness FE (1987) Sexual segregation and density-related changes in habitat use in male and female red deer (*Cervus elaphus*). *Journal of Zoology* **211,** 275-289.

Cote SD, Rooney TP, Tremblay JP, Dussault C & Waller DM (2004) Ecological impacts of deer overabundance. *Annual Review of Ecology Evolution and Systematics* **35,** 113-147.

Diaz 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.

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.

Geist V (1974) Relationship of social evolution and ecology in ungulates. *American Zoologist* **14,** 205-220.

Goodman SJ, Barton NH, Swanson G, Abernethy K & Pemberton JM (1999) Introgression through rare hybridisation: a genetic study of a hybrid zone between red and sika deer (genus *Cervus*), in Argyll, Scotland. *Genetics* **152,** 355-371.

Gordon IJ & Illius AW (1988) Incisor arcade structure and diet selection in ruminants. *Functional Ecology* **2,** 15-22.
Harrington R (1973) Hybridisation among deer and its implication for conservation. *Irish Forestry* **30,** 64-78.
Harrington R (1979) *Some Aspects of the Biology and Taxonomy of the Deer of the County Wicklow Region, Ireland,* unpublished PhD thesis, National University of Ireland.
Harrington R (1982) Hybridization of red deer (*Cervus elaphus* L. 1758) and Japanese siak deer (*Cervus nippon* Temminck, 1838). *Transactions of the International Congress of Game Biology* **14,** 559-571.
Hastings A, Cuddington K, Davies KF, Dugaw CJ, Elmendorf S, Freestone A, Harrison S, Holland M, Lambrinos J, Malvadkar U, Melbourne B A, Moore K, Taylor C & Thomson D (2005) The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters* **8,** 91-101.
Husheer SW, Allen RB & Robertson AW (2006) Suppression of regeneration in New Zealand mountain beech forests is dependent on species of introduced deer. *Biological Invasions* **8,** 823-834.
Jarman P J (1974) Social-organization of antelope in relation to their ecology. *Behaviour,* **48,** 215
Kronmal R A (1993) Spurious correlation and the fallacy of the ratio standard revisited. *Journal of the Royal Statistical Society Series a-Statistics in Society* **156,** 379-392.
Kruuk LEB, Slate J, Pemberton JM, Brotherstone S, Guinness F & Clutton-Brock TH (2002) *Antler size in red deer: heritability and selection but no evolution. Evolution* **56,** 1683-1695.
Lowe VPW, Gardiner AS (1974) A re-examination of the subspecies of red deer (*Cervus elaphus*) with particular reference to the stocks in Britain. *Journal of Zoology, London* **174**, 185-201.
Lowe VPW & Gardiner A S (1975) Hybridisation between Red deer (*Cervus elaphus*) and Sika deer (*Cervus nippon*) with particular reference to stocks in N.W. England. *J. Zool. Lond.* **177,** 553-566.
McDevitt AD, Edward JE, O'Toole P, O'Sullivan P, O'Reilly C & Carden RF (in review) Genetic structure of, and hybridisation between red (*Cervus elaphus*) and sika (*Cervus nippon*) deer in Ireland. *Mammalian Biology*.
McLean C (1993) *Sika Deer Control*. Red Deer Commision.
Mitchell B & Brown D (1974) The effects of age and body size on fertility in female red deer (*Cervus elaphus* L.). *Proceedings of the International Congress of Game Biology* **11,** 89-98.
Mitchell B, McCowan D & Nicholson IA (1976) Annual cycles of body weight and condition in Scottish Red deer, *Cervus elaphus*. *Journal the Zoological Society, London* **180,** 107-127.
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.
Perez-Barberia FJ & Gordon IJ (2001) Relationships between oral morphology and feeding style in the Ungulata: a phylogenetically controlled evaluation. *Proceedings of the Royal Society of London Series B-Biological Sciences* **268,** 1023-1032.
Pritchard JK, Stephens M & Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155,** 945-959.
Putman RJ & Hunt EJ (1993) Hybridisation between red and sika deer in Britain. *Deer* **9,** 104-108.
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 hybridisation and introgression *Annual Review of Ecology and Systematics* **27,** 83-109.

Riney T (1955) Evaluating condition of free-ranging red deer (*Cervus elaphus*). with special reference to New Zealand. *New Zealand Journal of Science and Technology Series E* **36,** 429-463.

Serrano E, Alpizar-Jara R, Morellet N & Hewison AJM (2008) A half a century of measuring ungulate body condition using indices: is it time for a change? *European Journal of Wildlife Research* **54,** 675-680.

Staines BW, Crisp JM & Parish T (1982) Differences in the quality of food eaten by red deer (*Cervus elaphus*) stags and hinds in winter. *Journal of Applied Ecology* **19,** 65-77.

Suttie JM & Mitchell B (1983) Jaw length and hind foot length as measures of skeletal development of red deer (*Cervus elaphus*). *Journal of Zoology* **200,** 431-434.

Swanson GM & Putman RJ (2008) Sika deer in the British Isles. *Sika Deer: Biology, Conservation and Management* (eds DR McCullough, S Takatsuki & K Kaji). Springer, Japan.

Vaha J-P, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15**, 63-72.

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

White EM, Wilson JC & Clarke AR (2006) Biotic indirect effects: a neglected concept in invasion biology. *Diversity and Distributions* **12,** 443-455.

Winfree AT (1980) *The Geometry of Biological Time*. Springer-Verlag, New York.

Chapter 6

Conclusions

6.1 General Conclusions

There are many well-studied examples of ancient stable hybrid zones such as those of *Bombina, Heliconius, Iris, Mus* and *Gryllus* mentioned in the introduction. Numerous studies have surveyed these hybrid zones, assessed the balance of dispersal and selection maintaining the zones, elucidated the nature of selection operating in the zone and searched for the genes responsible. Studies of hybridisation between introduced and native species are in comparative infancy, yet these studies provide a unique angle on the process of hybridisation because they provide the opportunity to observe hybridisation between two populations at an early stage. In this thesis I have presented the most recent chapter in a 15 year study of hybridisation between red and sika deer on the Kintyre Peninsula, Scotland (Abernethy 1994a; Abernethy 1994b; Goodman *et al.* 1999; Goodman *et al.* 2001; Swanson 2000). This study adds to the small number of existing longitudinal genetic studies of hybridisation between invasive and native species (Perry *et al.* 2001; Pinto *et al.* 2004; McGinnity *et al.* 2003). This study also complements existing well-studied examples of mammalian hybridisation which have focused on hybridisation between domestic species and their wild relatives (wild cat x domestic cat (Beaumont *et al.* 2001; Driscoll *et al.* 2007; Lecis *et al.* 2006; Randi *et al.* 2001), dog x wolf (Lucchini *et al.* 2004; Randi & Lucchini 2002; Verardi *et al.* 2006; Anderson *et al.* 2009), bison x cow (Halbert & Derr 2007; Halbert & Derr 2008; Halbert *et al.* 2005; Polziehn *et al.* 1995; Ward *et al.* 1999)) or have studied natural hybrid zones (*Mus* (Dod *et al.* 1993; Hunt & Selander 1973; Payseur *et al.* 2004; Payseur & Nachman 2005; Raufaste *et al.* 2005; Smadja & Ganem 2005; Vanlerberghe *et al.*

1986) or naturally arising hybrid populations (Baboon (Charpentier *et al.* 2008; Tung *et al.* 2008)).

Hybridisation between red and sika deer on the Kintyre Peninsula demonstrates that only a small number of F1 hybridisation events can lead to substantial introgression (Goodman *et al.* 1999; Chapter 4). This finding is of interest from the point of view of the control of introduced species, perhaps particularly from the perspective of transgene escape from genetically modified plants into wild relatives (Hails & Morley 2005; Pineyro-Nelson *et al.* 2009); hybridisation need only be very rare for the risk of introgression to be substantial. Similarly, isolated hybridisation events have the potential to leave a lasting evolutionary legacy which would be detected in the form of discordant genealogies. The simulation results in Chapter 3 also highlight that even the fixation of an allele with discordant genealogy (due to hybridisation, not incomplete lineage sorting) can simply be the result of neutral introgression and is not in itself proof that the allele was selectively advantageous. The simulation results additionally reveal the need for high levels of recombination between the two parental populations if QTL mapping of natural hybrid zones (Rieseberg & Buerkle 2002) is to be fruitful. This means that QTL mapping is only suitable in natural hybrid populations where hybridisation between the two parental species is frequent. If doubt exists about the suitability of a mapping population then it can be resolved by verifying whether patterns of introgression are concordant in multiple independent hybrid populations and thus consistent with the selective, as opposed to the neutral, scenario (e.g. Rieseberg *et al.* (1998)).

In Chapter 4 I also present a method for estimating the number of hybridisation events that have occurred in a population by using information from diagnostic, but polymorphic, nuclear markers to reconstruct ancestral haplotypes. This is, to my knowledge, the first study to attempt such an analysis, although various studies have used the number of

polymorphisms at nuclear and mtDNA markers (Ross & Shoemaker 2008) or just mtDNA markers (Jackson *et al.* 2008; Murray-McIntosh *et al.* 1998) to estimate sizes of founder populations.

The rapid formation of hybrid swarms between endemic and introduced species has been documented over time scales of 5-10 years before (e.g. Childs *et al.* 1996; Pinto *et al.* 2005). The red-sika hybrid swarm at West Loch Awe provides another example of rapid formation. This swarm probably formed over 20-30 years (Chapter 4). An interesting feature of red-sika hybridisation is that hybrid swarm formation appears to be sporadic. We cannot be certain of the reason why a hybrid swarms exist at West Loch Awe and not at other sites across the peninsula. However, as discussed in Chapters 2 and 4, this could simply be due to a chance build up of hybrids, leading to breakdown in assortative mating and collapse of the population into a hybrid swarm via positive feedback. If this is the case, rarely hybridising populations like red and sika deer might be able to coexist for long periods with relatively little gene flow, but then undergo rapid and unpredictable periods of hybridisation.

6.2 Management Conclusions

The discovery of a hybrid swarm on the Kintyre Peninsula (Chapters 2 & 4) opens up the possibility that the two species could eventually merge across their entire range in Britain. Extensive hybridisation between red deer and sika deer found in county Wicklow, Ireland (Harrington 1973; McDevitt *et al.* in review) is thought to have arisen due to hybridisation of the deer prior to their release from Powerscourt Park (Powerscourt 1884). The swarm documented here is the first genetically verified example of breakdown in assortative mating between the two species arising in the wild. Since this hybrid swarm was largely unreported,

even by the rangers involved in culling the deer (Chapter 2), it seems possible that other such swarms exist undetected across Britain; and indeed Germany, Austria and the Czech Republic where both species are found (Bartos *et al.* 1981; Pitra & Lutz 2005). Phenotypic evidence (Chapter 5) suggests that hybridisation is narrowing the morphological gap between the two species. Since, we also find no suggestion of reduced pregnancy rates in hybrid animals, it can be suggested that increasing similarity between the two populations may facilitate further hybridisation (Chapter 5). The long term phenotypic and ecological consequences of hybridisation are hard to predict, for reasons discussed in Chapter 5, but it is certain that the red deer population, where in contact with sika, is under real threat of genetic and phenotypic introgression.

The key to understanding how to manage red-sika hybridisation may lie in understanding the role of sika stags in the process. Sika males have been reported to travel large distances (>100km) from their main range and lone males are nearly always on the leading edge of sika range expansion (Livingstone 2001). Ratcliffe (1987) suggests that colonising stags can precede the arrival of hinds in an area by up to 10 years and an examination of the Forestry Commission Scotland cull records for sites on the leading edge of the distribution of sika (Shira, Oban and Succoth) supports this observation since 16 sika stags, and no sika hinds, have been culled at these sites in the last ten years (I did not have access to records further back). The virtually complete introgression of red mtDNA into sika-like animals at West Loch Awe excludes the possibility that pure sika females could have been involved in hybridisation here since all hybrid animals must have an unbroken line of red mtDNA-carrying female ancestry. This pattern of introgression supports the scenario that pioneering stags entering areas devoid of sika females are mating with locally available red deer hinds. MtDNA introgression at other sites in the study area is predominantly in the same direction, although these examples do not necessarily support the pioneering sika stag theory since the examples come from sites where both species have large populations (Lussa, Carradale,

Knapdale and Achaglachach, Figure 2, Chapter 4). Additionally some mtDNA introgressed individuals are introgressed at many loci (Table 4, Chapter 4) meaning that they are the result of recent hybridisation. Nevertheless, we would expect it to be difficult to find evidence of red mtDNA introgression on the leading edges of the sika range, since the hybrids generated (excepting such cases as West Loch Awe) are presumably most likely to backcross into the red population leaving no discordant mtDNA. I conclude that, although hybridisation between pioneering stags and red deer hinds appears to be a compelling explanation for the situation at West loch Awe, and it is highly likely that pioneering stags are involved in hybridisation events elsewhere, the pattern of introgression found across the peninsula does not exclude the possibility that such hybrid pairings do not also take place in areas where sika females are present. Indeed, since there is also a single example of sika mtDNA introgression in the 2006/7 dataset (Chapter 2), it is apparent that hybridisation can also take place between red deer stags and sika hinds.

Whether or not the threat posed to red deer by introgression is of serious concern is open to debate. In Scotland, I have hypothesised that hybridisation could lead to ecological and economic problems both because of trait introgression from sika into red deer and because the loss of phenotypic integrity of red deer might have negative consequences for the stalking industry (Chapter 5). As yet, there is no evidence to suggest that these concerns are a reality. But following a precautionary principle might be sensible - by the time we know the answers to these questions the situation will probably be irretrievable. Based on the knowledge gained through the course of this study I shall now discuss management recommendations directed at specific areas of Scotland:

6.2.1 West Loch Awe

The situation at West Loch Awe is clearly the biggest threat to the genetic integrity of red deer in Scotland discovered so far. The population is apparently reasonably contained

(Figure 5, Chapter 2, although see arguments in Chapter 4), but an attempt at eradication through increased culling would be difficult because of the dense plantation forestry which the deer inhabit. In fact, heavy culling might be ill advised as this could scatter hybrids into new areas. A novel suggestion would be an attempt at artificially lowering the reproductive rate of the hybrid population *in situ*, thus hopefully containing the spread of hybrids. Most wildlife contraceptives involve injection or implantation, but feed containing Melengestrol Acetate has been successfully employed as a contraceptive in captive Cervids including sika (Raphael *et al.* 2003). However, the frequency of treatment required and the effects on non-target species mean that its application is generally not deemed suitable in the wild (Patton *et al.* 2007). It is perhaps conceivable that targeted delivery mechanisms of oral contraceptives could be developed (e.g. specialised feeding boxes), but objection over the safety of such agents entering the human food chain via the consumption of venison might still be a serious objection to such an action being taken.

The most realistic option might be a combination of fencing and culling. In theory, it might be possible to make use of the fact that Loch Awe acts as a serious barrier to deer spreading to the east and to erect fencing in strategic places to create a zone of containment. If it was possible to fence/ improve fencing along the A85 from near Ben Cruachan to Connel, perhaps skirting around Taynuilt and rejoining the road later, this would act as a serious barrier to migration of hybrids and sika northwards. This project would require around 20 km of fencing and would be an expensive operation; however, such fencing might bring the added benefit of reducing deer vehicle collisions along the A85. A second fence, from the southern end of Loch Awe westward to the sea could also be considered (Ford to Ardfern ~6km). Containing movement of deer southwards might be less important since the red deer to the south on Kintyre are already fairly hybridised. Once fencing (at least to the north) had been erected, a campaign to cull heavily within the fenced area could be conducted with fewer worries about the spread of hybrids into previously unaffected areas. This stratergy

offers little hope of completely eliminating introgression into red deer populations northeast of the fenced off area, because we have already found evidence of some introgression in these areas. Nevertheless, there might be merit in reducing the density of the population of very intermediate animals found at West Loch Awe, as the presence of these animals is likely be a much more potent source of further introgression into red deer, than the presence of sika, or the presence of red deer carrying a small proportion of sika genes. There is obviously no easy solution to this issue, but a working group with the right combination of technical and practical knowledge might be able to produce a realistic strategy that could at least be considered as an option.

6.2.2 Mainland of Scotland

It seems the genetic situation on mainland Scotland may be irretrievable; on Kintyre we have discovered introgression into red deer in all places where they overlap with sika and in some places where sika are apparently rare (e.g. Barcaldine and Glen Lochy Figure 2, Chapter 4). Since sika overlap on 40% of the red deer range in mainland Scotland (Livingstone 2001; Ward 2005), we have no reason to expect the picture will be different elsewhere. Since pioneering sika stags are almost certainly involved in hybridisation, even red deer populations far from the main range of sika are at risk of introgression. Complete genetic containment of sika to particular areas of the mainland is probably unfeasible due to the extremely large areas involved of land involved. Nevertheless, whilst the situation in West Loch Awe is concerning, because it highlights that substantial introgression could occur or is occurring between the two species elsewhere, there are still clearly phenotypically and ecologically distinct populations of red deer in many areas of Scotland, regardless of their genetic status. Whilst I have demonstrated here that sika clearly pose a genetic threat to red deer, even their presence without hybridisation creates economic problems through their damage to forestry (Swanson & Putman 2008; Perez-Espona *et al.* in press). Thus, conservation effort on mainland Scotland should not be centred on worries about genetic

purity *per se,* but should focus, where possible, on controlling the spread of sika populations into new areas, even if it is unrealistic to expect that introgression from sika into red can be kept out completely. A number of estates are actively protecting sika ranging onto their land in the hope that populations will take hold - sika stags command higher prices than red deer stags on stalking lets (personal communication N. Rowantree). On the mainland, public bodies (Scottish National Heritage, Deer Commission for Scotland and National Park Authorities) should put resources into preventing the active encouragement of sika spread and increasing the control of sika in areas on the fringe of their range, as opposed to investigating genetic purity. Sensible strategies might include educating land owners, especially on reasonably isolated pieces of land at the edge of the sika distribution (e.g. Mull, Skye, Morvern, estates in mountainous regions), about the possible risks of introgression posed by sika, and in particular sika stags on their land. This campaign would highlight the fact that the arrival of sika stags in an area does not necessarily signal that a self-sustaining sika population exists yet and that these sika stags may be particularly likely to hybridise with the local red deer populations. Photographic evidence from the population of hybrids at West Loch Awe (which look intermediate, personal observation by the author and Josephine Pemberton), and more detailed phenotypic studies might aid such a campaign. These estates should be encouraged to be vigilant to the arrival of sika stags and to have an active policy of shooting them and any animals of hybrid appearance. The support for culling pioneering stags exists on paper (DCS 2008), but communication with, and regular encouragement of, relevant landowners to pursue this policy is needed. It may be that some genetic testing is necessary to provide the impetus for further action, but the danger is that the failure to find evidence of introgression in relatively small samples taken from target red deer populations will provide a false sense of security and delay action that could be taken anyway.

6.2.3 Islands

On the islands, the situation is different, and a genuine chance to preserve the genetic purity of *Cervus elaphus* exists, even if some population are not purely Scottish due to introductions of red deer from abroad (e.g. Nussey et al. (2006)). The Island Refugia Policy which makes it illegal to introduce any deer of the genus *Cervus* into the wild on the islands of the Outer Hebrides and the islands of Arran, Islay, Jura and Rum [Wildlife & Countryside act 1981 (variation of schedule 9) Order 1999], has been a vital step. Additional legislation is needed to ensure similar laws apply to the introduction of deer for farming. Further work should be carried out by the relevant governmental agencies to raise awareness of the law on this matter and also to instigate a framework (e.g. a website) through which potential sika sighting and possible illegal introductions could be systematically reported. The status of the islands as refugia for red deer should, hopefully, be of economic benefit to the island communities who could use it as a selling point for stalking tourism. Regular genetic testing (perhaps every 10 years) should be one, but not the only, way to monitor the island populations. Over time, the relative cost of testing should decline, but testing strategies should still be considered carefully. Since it is possible for sika or hybridised red deer to swim from the mainland, it would be most sensible to concentrate efforts on areas highest at risk from these events (e.g. Jura and Arran) and to combine testing strategies with information gathered locally about possible sightings. Asking deer managers in areas where sika are most likely to arrive from the mainland to be vigilant and to report any sightings is also important.

6.3 Future research

MtDNA haplotype sequencing of the datasets used in this study, assuming red deer haplotypes are as polymorphic in the rest of Scotland (Perez-Espona *et al.* 2009), would

provide additional lines of information on the number of hybridisation events that have contributed to introgression into sika. Evidence from mtDNA introgression reveals that hybridisation is predominantly between red deer hinds and sika stags, but a Y-chromosome marker that distinguishes between the two species (Lewis 2003) might have some use in determining whether or not male hybrids are involved in backcrossing into red deer populations and should further the understanding of the mating patterns causing the strikingly high levels of mtDNA introgression at West Loch Awe. However, a Y-chromosome marker may have limited power, firstly because it reduces sample sizes by around 50% and secondly because male deer have highly skewed reproductive success (Clutton-Brock *et al.* 1988) and *a priori* we might expect hybrid males to be relatively unsuccessful in reproduction compared with females

Is the hybrid swarm at West Loch Awe an isolated incident or the inevitable outcome of hybridisation in the long term? Surveys of other areas in which hybridisation is suspected might help answer this question. In particular, it would be interesting to survey the southern Lake District where hybridisation was reported on the basis of craniological measurement in a study by Lowe and Gardiner (1975). Other populations of interest might be the Scottish Borders, Easter Ross and Sutherland which were previously shown to have introgression at a number of loci in red and sika when a survey using 11 microsatellite loci was carried out (Swanson 2000). Indeed reanalysis of these samples collected during Swanson's thesis with the multiplex panel developed here might be fruitful and might also form the basis for a parallel temporal study to that on Kintyre if the populations were sampled again. On Kintyre, surveys of private woodland in the areas surrounding West Loch Awe would establish how contained the hybrid swarm is, and it would also be interesting to survey the very southern end of the peninsula more thoroughly, since the population of red deer there originated from a deer farm escape and are reportedly hybridising with sika (personal communication Kevin McKillop). Although I only tested two samples from this site one of them was a highly

hybridised red-like animal and was the only sample in the entire dataset to carry introgressed sika mtDNA (Figure 2, Chapter 4). Investigation of this site might shed further light on the factors that control the direction of hybridisation between these two species.

There are various documented introductions of wapiti (*Cervus canadenisis*) into Scotland (Whitehead 1964). The 22-multiplex marker panel developed in Chapter 2 could be trialled on Wapiti samples, since successful inclusion of wapiti samples in STRUCTURE (Pritchard *et al.* 2000) analyses would allow current and future samples of red deer to be simultaneously assayed for wapiti as well as sika introgression.

With 20-22 markers (Chapters 2 & 4) we have the power to detect introgression to around four backcross generations. Additional suitable microsatellite markers should not be hard to find due to the many microsatellites isolated in cattle and sheep that amplify in deer (Perez-Espona *et al.* 2008; Slate *et al.* 1998) I have shown that analysis of the microsatellite data with the Bayesian software STRUCTURE 2.2 (Falush *et al.* 2007; Pritchard *et al.* 2000) provides a workable method to simultaneously estimate allele frequencies in the parental populations, assign hybrid scores and also estimate and account for the incidence of null alleles in this system. Whilst the 20-22 markers used in this study were extremely differentiated, additional markers, even if they showed lower levels of differentiation would still add power (Vaha & Primmer 2006). Alternatively sets of hundreds of SNPs could be developed (Aitken *et al.* 2004; Bouck &Vision 2007), which would provide a very efficient and reliable method of screening for hybridism (Schlotterer 2004). Increasing the numbers of markers increases linkage between markers and if distances between markers are known, linkage can be accounted for when assessing admixture, providing a powerful method for detecting more distant hybridisation events (Falush *et al.* 2007). Providing appropriate reference data was available, admixture analysis with linked markers would, for example, make it possible to ascertain the relatedness of introduced sika populations in the British

Isles (e.g. did sika in Kintyre originally come via Powerscourt?) and investigate whether sika were previously hybridised with red deer or indeed any other cervid species prior to their arrival on the Kintyre peninsula (as has been suggested for many of the sika that came from Powerscourt after documented hybridisation in the park (Powerscourt 1884)). Closely linked SNP haplotypes could be used in a manner analogous to microsatellite markers, to investigate the number of introgression events into the sika population and the number of individuals involved in founding the sika population (Leblois & Slatkin 2007); although it is still possible that diversity at SNP haplotypes will be too low (restricted by the bottleneck sika have passed through) to investigate the number of introgression events into red deer. Ultimately, a large scale study of many separate populations (perhaps in collaboration with the groups that have samples from Ireland and Southern England (Diaz *et al.* 2006; McDevitt *et al.* in review)), with a large number of markers, would shed light both on the history of introductions and the factors involved in hybridisation.

6.4 References

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

Abernethy K (1994b) *The Introduction of Sika Deer (Cervus nippon nippon) to Scotland*, unpublished PhD thesis, University of Edinburgh.

Aitken N, Smith S, Schwarz C, Morin PA (2004) Single nucleotide polymorphism (SNP) discovery in mammals: a targeted-gene approach. *Molecular Ecology* **13**, 1423-1431.

Anderson TM, vonHoldt BM, Candille SI*, et al.* (2009) Molecular and evolutionary history of melanism in North American gray wolves. *Science* **323**, 1339-1343.

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

Beaumont M, Barratt EM, Gottelli D*, et al.* (2001) Genetic diversity and introgression in the Scottish wildcat. *Molecular Ecology* **10**, 319-336.

Bouck A, Vision T (2007) The molecular ecologist's guide to expressed sequence tags. *Molecular Ecology* **16**, 907-924.

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.

Clutton-Brock TH, Albon SD, Guinness FE (1988) Reproductive success in male and female red deer. In: *Reproductive Success* (ed. Clutton-Brock TH). The University of Chicago Press, Chicago.

DCS (2008) Scotland's Wild Deer a National Approach (ed. Scotland DCS). The Scottish Government.

Diaz 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.

Dod B, Jermiin LS, Boursot P*, et al.* (1993) Counterselection on sex-chromosomes in the *Mus musculus* European hybrid zone. *Journal of Evolutionary Biology* **6**, 529-546.

Driscoll CA, Menotti-Raymond M, Roca AL*, et al.* (2007) The Near Eastern origin of cat domestication. *Science* **317**, 519-523.

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.

Goodman SJ, Barton NH, Swanson G, Abernethy K, Pemberton JM (1999) Introgression through rare hybridisation: 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.

Hails RS, Morley K (2005) Genes invading new populations: a risk assessment perspective. *Trends in Ecology & Evolution* **20**, 245-252.

Halbert ND, Derr JN (2007) A comprehensive evaluation of cattle introgression into US federal bison herds. *Journal of Heredity* **98**, 1-12.

Halbert ND, Derr JN (2008) Patterns of genetic variation in US federal bison herds. *Molecular Ecology* **17**, 4963-4977.

Halbert ND, Ward TJ, Schnabel RD, Taylor JF, Derr JN (2005) Conservation genomics: disequilibrium mapping of domestic cattle chromosomal segments in North American bison populations. *Molecular Ecology* **14**, 2343-2362.

Harrington R (1973) Hybridisation among deer and its implication for conservation. *Irish Forestry* **30**, 64-78.

Hunt GW, Selander RK (1973) Biochemical genetics of hybridisation in European house mice. *Heredity* **31**, 11-33.

Jackson JA, Patenaude NJ, Carroll EL, Baker CS (2008) How few whales were there after whaling? Inference from contemporary mtDNA diversity. *Molecular Ecology* **17**, 236-251.

Leblois R, Slatkin M (2007) Estimating the number of founder lineages from haplotypes of closely linked SNPs. *Molecular Ecology* **16**, 2237-2245.

Lecis R, Pierpaoli M, Biro ZS*, et al.* (2006) Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. *Molecular Ecology* **15**, 119-131.

Lewis S (2003) *Y-chromosome Markers for the Study of Population History in Deer*, unpublished honours thesis, Edinburgh University.

Livingstone SR (2001) *The Application of GIS to the Spread of Introduced Japanese Sika Deer (Cervus nippon) in Scotland.*, unpublished MSc thesis, University of Edinburgh.

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

Lucchini V, Galov A, Randi E (2004) Evidence of genetic distinction and long-term population decline in wolves (*Canis lupus*) in the Italian Apennines. *Molecular Ecology* **13**, 523-536.

McDevitt AD, Edward JE, O'Toole P*, et al.* (in review) Genetic structure of, and hybridisation between red (*Cervus elaphus*) and sika (*Cervus nippon*) deer in Ireland. *Mammalian Biology*.

McGinnity P, Prodohl P, Ferguson K*, et al.* (2003) Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proceedings of the Royal Society of London Series B-Biological Sciences* **270**, 2443-2450.

Murray-McIntosh RP, Scrimshaw BJ, Hatfield PJ, Penny D (1998) Testing migration patterns and estimating founding population size in Polynesia by using human mtDNA sequences. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 9047-9052.

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.

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

Payseur BA, Krenz JG, Nachman MW (2004) Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution* **58**, 2064-2078.

Payseur BA, Nachman MW (2005) The genomics of speciation: investigating the molecular correlates of X chromosome introgression across the hybrid zone between *Mus domesticus* and *Mus musculus*. *Biological Journal of the Linnean Society* **84**, 523-534.

Perez-Espona S, Perez-Barberia FJ, McLeoad JE, Jiggins CD, Gordon IJ, Pemberton JM (2008) Landscape features affect gene flow in Scottish Highland red deer.*Molecular Ecology* **17,** 981-996.

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

Perez-Espona S, Pemberton JM, Putman R (in press) Red and sika deer in the British Isles, current management issues and management policy. *Mammalian Biology*.

Perry WL, Feder JL, Dwyer G, Lodge DM (2001) Hybrid zone dynamics and species replacement between *Orconectes* crayfishes in a northern Wisconsin lake. *Evolution* **55**, 1153-1166.

Pineyro-Nelson A, Van Heerwaarden J, Perales HR*, et al.* (2009) Transgenes in Mexican maize: molecular evidence and methodological considerations for GMO detection in landrace populations. *Molecular Ecology* **18**, 750-761.

Pinto MA, Rubink WL, Coulson RN, Patton JC, Johnston JS (2004) Temporal pattern of Africanization in a feral honeybee population from Texas inferred from mitochondrial DNA. *Evolution* **58**, 1047-1055.

Pinto MA, Rubink WL, Patton JC, Coulson RN, Johnston JS (2005) Africanization in the United States: replacement of feral European honeybees (*Apis mellifera* L.) by an African hybrid swarm. *Genetics* **170**, 1653-1665.

Pitra C, Lutz W (2005) Population genetic structure and the effect of founder events on the genetic variability of introduced sika deer, *Cervus nippon*, in Germany and Austria (vol 51, pg 95, 2005). *European Journal of Wildlife Research* **51**, 295-295.

Polziehn RO, Strobeck C, Sheraton J, Beech R (1995) Bovine mtDNA discovered in North American bison populations. *Conservation Biology* **9**, 1638-1643.
Powerscourt V (1884) On the acclimatization of the Japanese deer at Powerscourt. *Proceedings of the Zoological Society of London* **1884**, 207-209.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.
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.
Randi E, Lucchini V (2002) Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. *Conservation Genetics* **3**, 31-45.
Randi E, Pierpaoli M, Beaumont M, Ragni B, Sforzi A (2001) Genetic identification of wild and domestic cats (*Felis silvestris*) and their hybrids using Bayesian clustering methods. *Molecular Biology and Evolution* **18**, 1679-1693.
Ratcliffe PR (1987) Distribution and current status of Sika Deer, *Cervus nippon*, in Great Britain. *Mammal Review* **17**, 39-58.
Raufaste N, Orth A, Belkhir K*, et al.* (2005) Inferences of selection and migration in the Danish house mouse hybrid zone. *Biological Journal of the Linnean Society* **84**, 593-616.
Rieseberg LH, Baird SJE, Desrochers AM (1998) Patterns of mating in wild sunflower hybrid zones. *Evolution* **52**, 713-726.
Rieseberg LH, Buerkle CA (2002) Genetic mapping in hybrid zones. *American Naturalist* **159**, S36-S50.
Ross KG, Shoemaker DD (2008) Estimation of the number of founders of an invasive pest insect population: the fire ant *Solenopsis invicta* in the USA. *Proceedings of the Royal Society B-Biological Sciences* **275**, 2231-2240.
Schlotterer C (2004) The evolution of molecular markers - just a matter of fashion? *Nature Reviews Genetics* **5**, 63-69.
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.
Smadja C, Ganem G (2005) Asymmetrical reproductive character displacement in the house mouse. *Journal of Evolutionary Biology* **18**, 1485-1493.
Swanson GM (2000) *The genetic a phenotypic consequences of translocations of deer (Genus Cervus) in Scotland*, unpublished PhD thesis, University of Edinburgh.
Swanson GM, Putman RJ (2008) Sika Deer in the British Isles. In: *Sika Deer: Biology, Conservation and Management* (eds. McCullough DR, Takatsuki S, Kaji K). Springer, Japan.
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.
Vaha J-P, Primmer CR (2006) Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. *Molecular Ecology* **15**, 63-72.
Vanlerberghe F, Dod B, Boursot P, Bellis M, Bonhomme F (1986) Absence of Y-chromosome introgression across the hybrid zone between *Mus musculus domesticus* and *Mus musculus musculus*. *Genetical Research* **48**, 191-197.
Verardi A, Lucchini V, Randi E (2006) Detecting introgressive hybridization between free-ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium analysis. *Molecular Ecology* **15**, 2845-2855.
Ward AI (2005) Expanding ranges of wild and feral deer in Great Britain. *Mammal Review* **35**, 165-173.

Ward TJ, Bielawski JP, Davis SK, Templeton JW, Derr JN (1999) Identification of domestic cattle hybrids in wild cattle and bison species: a general approach using mtDNA markers and the parametric bootstrap. *Animal Conservation* **2**, 51-57.

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

Appendices

Data for Chapter 2

Table A1: Allele frequencies in red and sika populations estimated by STRUCTURE 2.2, under the admixture model, with null alleles estimated at all loci.

Locus	Allele	Sika	Red
AGLA293	128	0.005	0.143
	144	0.016	0.609
	147	0.970	0.084
	Null	0.008	0.163
BM4006	85	0.946	0.002
	87	0.000	0.053
	93	0.053	0.830
	95	0.000	0.109
	Null	0.001	0.007
BM6438	249	0.000	0.549
	251	0.009	0.324
	253	0.000	0.117
	261	0.000	0.004
	265	0.241	0.000
	275	0.747	0.000
	Null	0.002	0.005
BM757	160	0.014	0.055
	162	0.005	0.480
	164	0.000	0.003
	172	0.973	0.007
	174	0.000	0.007
	179	0.000	0.076
	183	0.000	0.101
	185	0.000	0.047
	187	0.000	0.008
	189	0.000	0.003
	198	0.007	0.077
	200	0.000	0.095
	202	0.000	0.031
	Null	0.001	0.010
BOVIRBP	140	0.974	0.001
	147	0.000	0.114
	149	0.000	0.016
	151	0.000	0.126
	153	0.022	0.428
	155	0.000	0.048
	157	0.000	0.188
	159	0.000	0.019
	Null	0.004	0.060

FCB193	103	0.021	0.097
	107	0.000	0.060
	109	0.000	0.206
	111	0.000	0.003
	113	0.000	0.293
	115	0.000	0.001
	118	0.000	0.069
	120	0.000	0.038
	122	0.000	0.093
	124	0.000	0.116
	128	0.000	0.003
	130	0.000	0.007
	132	0.977	0.001
	143	0.000	0.010
	Null	0.001	0.004
FSHB	180	0.989	0.004
	182	0.000	0.001
	184	0.000	0.013
	185	0.000	0.292
	188	0.000	0.102
	189	0.000	0.113
	191	0.000	0.058
	192	0.000	0.039
	194	0.002	0.046
	197	0.000	0.009
	198	0.000	0.091
	199	0.000	0.025
	201	0.000	0.008
	202	0.000	0.037
	203	0.000	0.027
	205	0.007	0.076
	206	0.000	0.030
	207	0.000	0.017
	210	0.000	0.003
	Null	0.001	0.008
IDVGA29	136	0.001	0.683
	143	0.001	0.306
	156	0.997	0.001
	Null	0.002	0.011
IDVGA55	191	0.000	0.070
	193	0.000	0.128
	195	0.005	0.299
	197	0.000	0.329
	199	0.005	0.123
	202	0.000	0.025
	204	0.000	0.006
	210	0.923	0.004
	212	0.066	0.000
	217	0.000	0.004

	219	0.000	0.006
	221	0.000	0.001
	Null	0.001	0.004
INRA5	126	0.036	0.985
	143	0.961	0.001
	Null	0.003	0.014
INRA6	128	0.002	0.000
	130	0.987	0.001
	132	0.000	0.033
	134	0.008	0.651
	136	0.002	0.304
	138	0.000	0.006
	Null	0.001	0.005
INRA131	92	0.000	0.050
	94	0.000	0.006
	98	0.007	0.542
	100	0.000	0.287
	102	0.000	0.084
	104	0.000	0.024
	106	0.992	0.000
	Null	0.001	0.005
MM12	89	0.001	0.686
	91	0.009	0.266
	93	0.988	0.004
	104	0.000	0.001
	Null	0.001	0.043
RM12	116	0.980	0.008
	125	0.005	0.278
	127	0.000	0.047
	129	0.000	0.049
	131	0.000	0.053
	133	0.000	0.281
	137	0.000	0.028
	139	0.014	0.119
	141	0.000	0.064
	144	0.000	0.023
	151	0.000	0.037
	Null	0.001	0.013
RM188	115	0.000	0.016
	117	0.000	0.093
	123	0.000	0.065
	125	0.000	0.017
	127	0.001	0.469
	129	0.000	0.176
	131	0.000	0.042
	132	0.000	0.001
	134	0.000	0.065
	137	0.001	0.039
	139	0.000	0.004

	143	0.567	0.003
	161	0.295	0.000
	163	0.002	0.000
	182	0.133	0.000
	Null	0.001	0.010
RM95	118	0.000	0.034
	122	0.996	0.006
	124	0.000	0.083
	126	0.000	0.049
	128	0.000	0.213
	130	0.000	0.329
	132	0.000	0.122
	134	0.000	0.003
	136	0.000	0.031
	138	0.000	0.108
	140	0.000	0.013
	147	0.002	0.000
	Null	0.001	0.009
RME25	151	0.000	0.042
	155	0.000	0.032
	159	0.000	0.004
	168	0.014	0.762
	170	0.011	0.131
	193	0.971	0.001
	207	0.000	0.011
	Null	0.004	0.017
TGLA40	91	0.000	0.253
	97	0.017	0.537
	99	0.000	0.023
	101	0.002	0.168
	104	0.974	0.003
	106	0.002	0.000
	108	0.002	0.000
	Null	0.002	0.015
TGLA126	100	0.402	0.000
	101	0.576	0.002
	105	0.019	0.983
	Null	0.003	0.015
TGLA127	161	0.877	0.001
	167	0.000	0.008
	169	0.000	0.464
	174	0.095	0.045
	178	0.025	0.237
	180	0.000	0.064
	184	0.000	0.065
	186	0.000	0.038
	190	0.000	0.032
	192	0.000	0.031
	Null	0.001	0.014

TGLA337	126	0.948	0.003
	130	0.000	0.165
	132	0.000	0.177
	134	0.000	0.002
	136	0.000	0.247
	138	0.043	0.041
	145	0.007	0.222
	147	0.000	0.065
	Null	0.001	0.078
UWCA47	225	0.000	0.008
	229	0.000	0.057
	231	0.001	0.880
	240	0.995	0.000
	Null	0.004	0.054

Data for Chapter 4

Table A2: Translation of alleles sizes between genotyping methods, allele frequencies and allele origins for loci used in Chapter 4

locus	Senn allele size	Goodman allele size	STRUCTURE estimated frequency in sika	STRUCTURE estimated frequency in red	likely population of origin
AGLA293	128	126	0.003	0.146	red
AGLA293	144	142	0.016	0.636	red
AGLA293	147	144	0.977	0.087	sika
AGLA293	null		0.005	0.131	-
BM4006	85	89	0.939	0.003	sika
BM4006	87	91	0	0.05	red
BM4006	93	97	0.043	0.826	red
BM4006	95	99	0	0.115	red
BM4006	null		0.018	0.006	-
BM4638	249	251	0	0.52	red
BM4638	251	253	0.003	0.318	red
BM4638	253	255	0	0.121	red
BM4638	261	263	0	0.007	red
BM4638	265	267	0.239	0	sika
BM4638	275	277	0.727	0	sika
BM4638	null		0.031	0.034	-
BM757	160	163	0.011	0.055	?
BM757	162	165	0.002	0.481	red
BM757	164	?	0	0.003	red
BM757	172	175	0.976	0.003	sika
BM757	174	179	0	0.007	red
BM757	179	183	0	0.071	red

BM757	183	187	0	0.106	red
BM757	185	189	0	0.045	red
BM757	187	191	0	0.009	red
BM757	189	?	0	0.003	red
BM757	198	201/203	0.009	0.203	red
BM757	null		0.002	0.014	-
BOVIRBP	140	144	0.985	0	sika
BOVIRBP	142	?	0	0.001	red
BOVIRBP	147	150	0	0.128	red
BOVIRBP	149	152	0	0.016	red
BOVIRBP	151	154	0	0.136	red
BOVIRBP	153	156	0.012	0.427	red
BOVIRBP	155	158	0	0.038	red
BOVIRBP	157	160	0	0.189	red
BOVIRBP	159	?	0	0.015	red
BOVIRBP	null		0.003	0.049	-
FCB193	103	101	0.009	0.081	red
FCB193	105	?	0	0.001	red
FCB193	107	105	0	0.05	red
FCB193	109	107	0	0.2	red
FCB193	111	109	0	0.002	red
FCB193	113	111	0	0.329	red
FCB193	115	113	0	0.001	red
FCB193	118	115	0	0.075	red
FCB193	120	117	0	0.029	red
FCB193	122	119	0	0.099	red
FCB193	124	121	0	0.101	red
FCB193	128	?	0	0.002	red
FCB193	130	?	0	0.007	red
FCB193	132	129	0.99	0.003	sika
FCB193	143	139	0	0.009	red
FCB193	null		0.001	0.013	-
IDVGA29	136	139	0.003	0.64	red
IDVGA29	143	147	0	0.32	red
IDVGA29	156	159	0.994	0	sika
IDVGA29	193	?	0	0.001	red
IDVGA29	null		0.003	0.039	-
IDVGA55	189	?	0	0.001	red
IDVGA55	191	193	0	0.074	red
IDVGA55	193	195	0	0.132	red
IDVGA55	195	197	0.002	0.311	red
IDVGA55	197	199	0	0.327	red
IDVGA55	199	201	0.007	0.118	red
IDVGA55	202	205	0	0.018	red
IDVGA55	204	207	0	0.005	red
IDVGA55	210	213	0.91	0.001	sika
IDVGA55	212	?	0.075	0	sika
IDVGA55	217	?	0	0.005	red
IDVGA55	219	?	0	0.004	red

IDVGA55	221	?	0	0.001	red
IDVGA55	225	?	0	0.001	red
IDVGA55	null		0.005	0.003	-
INRA131	92	93	0	0.041	red
INRA131	94	95	0	0.005	red
INRA131	98	99	0.006	0.528	red
INRA131	100	101	0	0.306	red
INRA131	102	103	0	0.086	red
INRA131	104	105	0	0.028	red
INRA131	106	107	0.993	0	sika
INRA131	null		0.001	0.005	-
INRA5	126	128	0.025	0.954	red
INRA5	143	144	0.972	0.002	sika
INRA5	null		0.003	0.044	-
INRA6	119	118	0	0.001	red
INRA6	128	?	0.001	0	sika
INRA6	130	128	0.994	0.001	sika
INRA6	132	130	0	0.03	red
INRA6	134	132	0.003	0.618	red
INRA6	136	134	0	0.306	red
INRA6	138	136	0	0.005	red
INRA6	null		0.002	0.039	-
MM12	89	89	0.006	0.673	red
MM12	91	91	0.006	0.282	red
MM12	93	93	0.987	0.004	sika
MM12	104	?	0	0.001	red
MM12	null		0.002	0.041	-
RM12	116	115	0.987	0.006	sika
RM12	125	123	0.001	0.267	red
RM12	127	125	0	0.044	red
RM12	129	127	0	0.046	red
RM12	131	129	0	0.048	red
RM12	133	131	0	0.298	red
RM12	137	133	0	0.03	red
RM12	139	137	0.01	0.113	red
RM12	141	139	0	0.081	red
RM12	144	141	0	0.018	red
RM12	151	147	0	0.036	red
RM12	163	?	0	0.001	red
RM12	null		0.001	0.011	-
RM188	115	120	0	0.032	red
RM188	117	122	0	0.09	red
RM188	123	126	0.002	0.057	red
RM188	125	128	0	0.021	red
RM188	127	130	0.001	0.468	red
RM188	129	132	0	0.168	red
RM188	131	134	0	0.043	red
RM188	132	136?	0	0.002	red
RM188	134	138	0	0.068	red

RM188	137	140	0	0.04	red
RM188	139	142	0	0.005	red
RM188	143	146	0.59	0.001	sika
RM188	161	162	0.291	0	sika
RM188	163	?	0.001	0	sika
RM188	182	182	0.11	0.001	sika
RM188	null		0.004	0.005	-
RM95	118	122	0	0.034	red
RM95	122	126	0.995	0.004	sika
RM95	124	128	0	0.081	red
RM95	126	130	0	0.052	red
RM95	128	132	0.001	0.242	red
RM95	130	134	0	0.316	red
RM95	132	136	0	0.108	red
RM95	134	?	0	0.002	red
RM95	136	140	0	0.033	red
RM95	138	142	0.002	0.109	red
RM95	140	144	0	0.013	red
RM95	147	?	0.001	0	sika
RM95	null		0.001	0.007	-
TGLA126	100	105	0.882	0.001	sika
TGLA126	105	109	0.032	0.986	red
TGLA126	null		0.087	0.013	-
TGLA127	161	157	0.905	0.002	sika
TGLA127	167	163	0	0.006	red
TGLA127	169	165	0	0.479	red
TGLA127	174	169	0.08	0.047	?
TGLA127	176	171	0	0.003	red
TGLA127	178	173	0.011	0.218	red
TGLA127	180	175	0	0.074	red
TGLA127	182	?	0	0.001	red
TGLA127	184	179	0	0.053	red
TGLA127	186	181	0	0.037	red
TGLA127	188	?	0	0.001	red
TGLA127	190	185	0	0.035	red
TGLA127	192	187	0	0.029	red
TGLA127	194	?	0	0.001	red
TGLA127	null		0.003	0.016	-
TGLA337	118	?	0	0.001	red
TGLA337	126	128	0.901	0.002	sika
TGLA337	130	132	0	0.171	red
TGLA337	132	134	0	0.19	red
TGLA337	134	?	0	0.003	red
TGLA337	136	138	0	0.234	red
TGLA337	138	140	0.062	0.047	?
TGLA337	142	?	0.001	0	sika
TGLA337	143	?	0	0.001	red
TGLA337	145	146	0.003	0.218	red
TGLA337	147	?	0	0.05	red

TGLA337	153	?	0	0.001	red
TGLA337	null		0.033	0.083	-
TGLA40	91	94	0	0.261	red
TGLA40	97	100	0.007	0.518	red
TGLA40	99	102	0	0.02	red
TGLA40	101	104	0.002	0.167	red
TGLA40	104	108	0.982	0.004	sika
TGLA40	106	?	0.002	0	sika
TGLA40	108	?	0.001	0	sika
TGLA40	110	?	0.001	0	sika
TGLA40	null		0.005	0.029	-
UWCA47	225	?	0	0.005	red
UWCA47	231	233	0	0.968	red
UWCA47	240	239	0.997	0	sika
UWCA47	null		0.002	0.027	-

Sample Details

Table A3: Sample details

Sample ID*	**Sample site**	**Sex**	**Age**	**MtDNA genotype**	**Q-hat (20 loci)**
HS0001	08WAWE	M	2	R	0.999
HS0002	01LUSA	M	2	S	0.001
HS0003	06KILM	M	1	R	0.999
HS0004	01LUSA	M	1	S	0.001
HS0005	04WTAR	M	2	S	0.003
HS0006	08WAWE	M	1	R	0.999
HS0007	01LUSA	M		S	0.001
HS0008	03STAR	M	4	S	0.001
HS0009	01LUSA	M		S	0.002
HS0010	01LUSA	M		S	0.001
HS0011	06KILM	M		R	0.999
HS0012	10ERDN	M	1	R	0.999
HS0013	02CARA	M	1	S	0.001
HS0014	05KNAP	M	1	R	0.999
HS0015	08WAWE	M	2	R	0.999
HS0017	10ERDN	M	1	R	0.999
HS0018	01LUSA	M		S	0.25
HS0019	04WTAR	M	3	R	0.999
HS0020	02CARA	M	1	S	0.052
HS0021	04WTAR	M	2	R	0.98
HS0022	10ERDN	M	2	R	0.999
HS0023	05KNAP	M	2	S	0.001
HS0025	02CARA	M	2	S	0.001
HS0026	08WAWE	M	3	R	0.626
HS0027	10ERDN	M	3	R	0.999
HS0028	02CARA	M	1	S	0.002

HS0029	10ERDN	M	2	R	0.999
HS0030	05KNAP	M	2	R	0.978
HS0031	06KILM	M	2	S	0.001
HS0032	10ERDN	M	3	R	0.999
HS0033	10ERDN	M	1	R	0.999
HS0034	06KILM	M	1	S	0.001
HS0035	02CARA	M	2	S	0.001
HS0036	01LUSA	M	2	S	0.001
HS0037	04WTAR	M	3	R	0.999
HS0038	08WAWE	M	3	R	0.876
HS0039	01LUSA	M		S	0.001
HS0040	11SHRA	M	3	R	0.999
HS0041	08WAWE	M	1	R	0.999
HS0042	02CARA	M	1	S	0.002
HS0043	11SHRA	M	3	R	0.999
HS0044	10ERDN	M	2	R	0.999
HS0045	01LUSA	M	1	S	0.003
HS0046	09COLL	M	1	R	0.981
HS0047	10ERDN	M	2	R	0.999
HS0048	01LUSA	M	2	R	0.001
HS0049	05KNAP	M	1	R	0.989
HS0050	02CARA	M	1	S	0.001
HS0051	11SHRA	M		R	0.999
HS0052	01LUSA	M	2	S	0.001
HS0053	10ERDN	M	2	R	0.999
HS0054	05KNAP	M	2	R	0.993
HS0055	11SHRA	M	2	R	0.999
HS0057	06KILM	M	2	R	0.999
HS0058	07ORMG	M	1	R	0.999
HS0059	10ERDN	M	2	R	0.999
HS0060	05KNAP	M	1	R	0.999
HS0062	10ERDN	M	3	R	0.999
HS0063	10ERDN	M	2	R	0.999
HS0064	10ERDN	M	1	R	0.999
HS0065	10ERDN	M	1	R	0.999
HS0066	01LUSA	M	2	S	0.001
HS0067	04WTAR	M	1	S	0.001
HS0068	10ERDN	M	2	R	0.999
HS0069	04WTAR	M	1	R	0.948
HS0070	01LUSA	M	2	S	0.001
HS0071	08WAWE	M	3	R	0.999
HS0072	06KILM	M	2	R	0.999
HS0073	05KNAP	M	2	R	0.999
HS0074	10ERDN	M	1	R	0.999
HS0075	11SHRA	M		R	0.999
HS0076	10ERDN	M	3	R	0.999
HS0077	01LUSA	M	4	S	0.002
HS0078	10ERDN	M	2	R	0.999
HS0079	02CARA	M	2	S	0.003

HS0080	11SHRA	M	1	R	0.999
HS0081	11SHRA	M		R	0.999
HS0082	01LUSA	M	1	S	0.039
HS0083	02CARA	M	1	S	0.001
HS0084	04WTAR	M	1	R	0.999
HS0085	04WTAR	M	1	R	0.999
HS0086	11SHRA	M		R	0.999
HS0087	11SHRA	M	3	R	0.999
HS0088	08WAWE	M	2	R	0.144
HS0089	07ORMG	M	3	R	0.963
HS0090	10ERDN	M	1	R	0.999
HS0091	08WAWE	M	3	R	0.999
HS0092	05KNAP	M	1	R	0.979
HS0093	02CARA	M	1	S	0.001
HS0094	10ERDN	M	1	R	0.999
HS0095	11SHRA	M	2	R	0.999
HS0096	08WAWE	M	3	R	0.999
HS0097	08WAWE	M	1	R	0.003
HS0098	04WTAR	M		S	0.002
HS0099	07ORMG	M	1	R	0.953
HS0100	02CARA	M	1	S	0.001
HS0101	06KILM	M		R	0.999
HS0102	02CARA	M	2	S	0.034
HS0103	08WAWE	M	5	R	0.988
HS0104	10ERDN	M	1	R	0.999
HS0105	06KILM	M	2	S	0.001
HS0106	10ERDN	M	1	R	0.999
HS0107	11SHRA	M	1	R	0.999
HS0108	10ERDN	M	1	R	0.999
HS0109	06KILM	M	2	R	0.999
HS0110	05KNAP	M	7	S	0.002
HS0111	10ERDN	M	1	R	0.999
HS0112	10ERDN	M	4	R	0.999
HS0113	06KILM	M	2	R	0.999
HS0114	10ERDN	M	1	R	0.999
HS0115	06KILM	M		R	0.999
HS0116	06KILM	M		R	0.999
HS0117	04WTAR	M	1	R	0.999
HS0118	05KNAP	M	1	R	0.961
HS0119	06KILM	M	1	R	0.999
HS0120	10ERDN	M	2	R	0.999
HS0121	04WTAR	M	3	R	0.999
HS0122	05KNAP	M	2	R	0.999
HS0123	11SHRA	M	2	R	0.999
HS0124	10ERDN	M	1	R	0.999
HS0125	04WTAR	M	3	R	0.999
HS0126	10ERDN	M	1	R	0.999
HS0127	06KILM	M	2	S	0.001
HS0128	08WAWE	M	2	S	0.001

HS0129	11SHRA	M	2	R	0.999
HS0130	08WAWE	M	3	R	0.999
HS0131	06KILM	M	1	R	0.999
HS0132	05KNAP	M	2	R	0.999
HS0133	02CARA	M	1	S	0.001
HS0134	05KNAP	M		R	0.989
HS0135	05KNAP	M	1	R	0.993
HS0136	08WAWE	M	4	R	0.984
HS0137	06KILM	M	2	R	0.999
HS0138	01LUSA	M		S	0.001
HS0139	02CARA	M	2	R	0.988
HS0140	01LUSA	M	2	R	0.961
HS0141	02CARA	M	1	S	0.001
HS0142	01LUSA	M	1	S	0.001
HS0143	02CARA	M	1	S	0.001
HS0145	01LUSA	M		S	0.002
HS0146	01LUSA	M	6	S	0.002
HS0147	02CARA	M	2	S	0.002
HS0148	01LUSA	F		S	0.001
HS0149	01LUSA	M	1	S	0.008
HS0150	01LUSA	M		S	0.001
HS0151	02CARA	M	0	S	0.001
HS0152	01LUSA	M	2	R	0.002
HS0153	02CARA	F	2	S	0.001
HS0154	02CARA	F		S	0.001
HS0155	02CARA	M	2	R	0.98
HS0156	01LUSA	F	3	S	0.001
HS0157	02CARA	M	1	S	0.002
HS0158	02CARA	F	3	S	0.003
HS0159	02CARA	F	3	S	0.026
HS0160	01LUSA	M		S	0.001
HS0161	01LUSA	M	5	S	0.001
HS0162	01LUSA	M		R	0.999
HS0163	01LUSA	M		S	0.001
HS0165	02CARA	F	2	S	0.001
HS0166	01LUSA	M		S	0.001
HS0167	01LUSA	M	4	R	0.999
HS0168	01LUSA	M		R	0.001
HS0169	02CARA	M		S	0.001
HS0171	01LUSA	M	2	S	0.001
HS0172	02CARA	M	4	S	0.001
HS0173	02CARA	F	0	S	0.025
HS0174	02CARA	F	1	R	0.999
HS0175	02CARA	M		S	0.002
HS0176	00DALB	M		S	0.001
HS0177	01LUSA	F	2	S	0.001
HS0178	01LUSA	F	2	R	0.001
HS0179	02CARA	F	0	S	0.004
HS0180	01LUSA	M		S	0.001

HS0181	02CARA	M	3	S	0.001
HS0182	02CARA	F	0	S	0.005
HS0183	01LUSA	F		R	0.001
HS0184	01LUSA	M	1	S	0.001
HS0185	02CARA	M	1	S	0.001
HS0188	02CARA	F	3	S	0.002
HS0189	02CARA	M		S	0.012
HS0190	01LUSA	M	3	S	0.001
HS0191	01LUSA	M	2	S	0.001
HS0192	02CARA	M	3	R	0.999
HS0193	01LUSA	F	1	S	0.001
HS0194	02CARA	M	1	S	0.001
HS0195	01LUSA	M	1	S	0.001
HS0196	01LUSA	M		S	0.001
HS0197	01LUSA	M	0	S	0.002
HS0198	01LUSA	M	1	S	0.001
HS0199	02CARA	M	3	R	0.977
HS0200	02CARA	F	2	S	0.019
HS0201	01LUSA	F	3	S	0.001
HS0203	02CARA	M		S	0.001
HS0204	01LUSA	F	3	S	0.001
HS0205	01LUSA	F	2	S	0.001
HS0206	08WAWE	M	3	R	0.246
HS0208	10ERDN	F	2	R	0.991
HS0209	10ERDN	M	1	R	0.999
HS0210	05KNAP	M	4	R	0.999
HS0211	10ERDN	M	1	R	0.999
HS0212	07ORMG	M	5	S	0.001
HS0213	06KILM	M	1	R	0.999
HS0214	11SHRA	M	3	R	0.999
HS0215	08WAWE	F	0	R	0.982
HS0216	08WAWE	F	3	R	0.446
HS0217	05KNAP	M	2	R	0.999
HS0218	10ERDN	M	6	R	0.999
HS0221	10ERDN	F	3	R	0.999
HS0222	10ERDN	M	4	R	0.999
HS0223	05KNAP	M	2	R	0.999
HS0224	08WAWE	F	3	R	0.999
HS0225	05KNAP	M	8	R	0.999
HS0226	08WAWE	M	2	R	0.999
HS0228	08WAWE	F	0	R	0.222
HS0229	08WAWE	M	0	R	0.995
HS0230	04WTAR	M	5	R	0.999
HS0231	06KILM	M	1	R	0.999
HS0232	10ERDN	F	3	R	0.999
HS0233	05KNAP	M	4	S	0.002
HS0234	10ERDN	F	3	R	0.999
HS0235	05KNAP	M	3	S	0.002
HS0236	06KILM	M	3	R	0.999

HS0237	10ERDN	F	2	R	0.999
HS0238	11SHRA	M		S	0.001
HS0239	11SHRA	M	3	R	0.999
HS0240	10ERDN	F	7	R	0.999
HS0241	08WAWE	M	3	R	0.138
HS0242	10ERDN	M	1	R	0.999
HS0243	08WAWE	M	2	R	0.985
HS0244	08WAWE	F	3	R	0.775
HS0245	08WAWE	F	3	R	0.999
HS0246	06KILM	M	1	R	0.999
HS0247	05KNAP	M	2	R	0.993
HS0248	11SHRA	F	8	R	0.999
HS0249	08WAWE	M	3	R	0.999
HS0250	05KNAP	M	1	S	0.001
HS0251	05KNAP	M	4	S	0.001
HS0252	05KNAP	M	6	R	0.999
HS0253	08WAWE	M	6	R	0.995
HS0254	08WAWE	M	2	R	0.942
HS0255	10ERDN	F	3	R	0.999
HS0256	05KNAP	M	7	S	0.01
HS0258	10ERDN	F	4	R	0.999
HS0259	10ERDN	F	0	R	0.999
HS0260	05KNAP	M	3	S	0.006
HS0261	10ERDN	F	2	R	0.999
HS0262	06KILM	M	1	S	0.005
HS0263	10ERDN	M	3	R	0.999
HS0264	05KNAP	M	3	S	0.001
HS0265	08WAWE	M	4	R	0.758
HS0266	10ERDN	F	2	R	0.999
HS0267	08WAWE	M	1	R	0.249
HS0268	08WAWE	M	7	R	0.901
HS0269	07ORMG	M	3	R	0.993
HS0270	06KILM	M	3	R	0.999
HS0271	08WAWE	F	3	R	0.409
HS0272	07ORMG	M	4	R	0.999
HS0273	08WAWE	M	2	R	0.999
HS0274	05KNAP	F	4	R	0.061
HS0275	08WAWE	M	1	R	0.924
HS0276	10ERDN	M	2	R	0.999
HS0277	06KILM	M	1	S	0.004
HS0278	10ERDN	F	4	R	0.999
HS0280	05KNAP	M	3	R	0.999
HS0282	20BALA	M		R	0.999
HS0283	20BALA	M	4	R	0.999
HS0284	20BALA	M	6	R	0.999
HS0285	12SOBN	M		R	0.999
HS0286	20BALA	M	1	R	0.999
HS0287	19APIN	F	0	R	0.999
HS0288	20BALA	M	2	R	0.999

HS0289	17NOBN	F	1	R	0.999
HS0290	20BALA	M	9	R	0.999
HS0291	17NOBN	M	0	R	0.999
HS0292	12SOBN	M	1	R	0.926
HS0293	19APIN	M	3	R	0.999
HS0294	20BALA	F	4	R	0.999
HS0295	20BALA	M		R	0.999
HS0296	20BALA	M	3	R	0.999
HS0297	20BALA	M	3	R	0.995
HS0298	20BALA	M	1	R	0.999
HS0299	17NOBN	M	2	R	0.999
HS0300	20BALA	M		R	0.999
HS0302	20BALA	M	3	R	0.999
HS0303	17NOBN	M	4	R	0.999
HS0304	20BALA	F		R	0.999
HS0307	17NOBN	M	2	R	0.999
HS0308	14LOCH	M	4	R	0.999
HS0309	13SUCC	F		R	0.999
HS0310	20BALA	F	1	R	0.999
HS0311	20BALA	M	1	R	0.999
HS0312	20BALA	M		R	0.999
HS0313	17NOBN	M	2	R	0.877
HS0314	20BALA	M		R	0.999
HS0315	14LOCH	M	7	R	0.999
HS0316	14LOCH	M	2	R	0.999
HS0317	17NOBN	M	2	R	0.999
HS0319	16ORCH	F	5	R	0.999
HS0320	17NOBN	M	2	R	0.999
HS0321	13SUCC	M	3	R	0.999
HS0322	15BENM	M	3	R	0.999
HS0323	15BENM	M	4	R	0.999
HS0324	15BENM	M	5	R	0.999
HS0325	16ORCH	M	3	R	0.986
HS0326	14LOCH	F	4	R	0.999
HS0327	14LOCH	M	0	R	0.999
HS0328	13SUCC	F	6	R	0.999
HS0329	17NOBN	M	2	R	0.999
HS0330	12SOBN	M	3	R	0.999
HS0331	20BALA	M	4	R	0.999
HS0332	17NOBN	M	1	R	0.965
HS0333	16ORCH	M	6	R	0.999
HS0334	20BALA	M	1	R	0.999
HS0335	17NOBN	M	3	R	0.999
HS0337	20BALA	M	4	R	0.999
HS0338	12SOBN	F	4	R	0.999
HS0339	15BENM	M	5	R	0.999
HS0340	17NOBN	M	2	R	0.942
HS0341	20BALA	M	1	R	0.999
HS0342	20BALA	F		R	0.999

HS0343	12SOBN	F	1	R	0.999
HS0344	14LOCH	M	4	R	0.999
HS0345	20BALA	M		R	0.999
HS0346	15BENM	M	4	R	0.999
HS0347	15BENM	F	3	R	0.999
HS0348	14LOCH	F	5	R	0.999
HS0349	15BENM	M	0	R	0.999
HS0350	20BALA	M		R	0.999
HS0351	14LOCH	F	0	R	0.999
HS0352	14LOCH	M	0	R	0.999
HS0353	14LOCH	F	7	R	0.999
HS0354	14LOCH	F	5	R	0.999
HS0356	15BENM	M	3	R	0.999
HS0357	16ORCH	M	7	R	0.999
HS0358	16ORCH	M	3	R	0.999
HS0359	16ORCH	M	3	R	0.999
HS0360	14LOCH	F	0	R	0.999
HS0361	16ORCH	F	3	R	0.999
HS0362	16ORCH	F	0	R	0.999
HS0363	17NOBN	M	2	R	0.999
HS0364	14LOCH	F	0	R	0.999
HS0365	16ORCH	F	4	R	0.999
HS0367	16ORCH	F	2	R	0.999
HS0368	16ORCH	M	5	R	0.999
HS0369	16ORCH	M	3	R	0.999
HS0370	16ORCH	M	8	R	0.999
HS0371	16ORCH	F	0	R	0.999
HS0372	16ORCH	M	0	R	0.999
HS0373	16ORCH	F	4	R	0.999
HS0374	16ORCH	F	1	R	0.999
HS0375	16ORCH	M	3	R	0.999
HS0376	16ORCH	F	4	R	0.999
HS0377	12SOBN	M	5	R	0.876
HS0378	16ORCH	F	0	R	0.999
HS0379	16ORCH	F	4	R	0.999
HS0380	16ORCH	F	3	R	0.999
HS0381	16ORCH	M	5	R	0.999
HS0382	16ORCH	F	3	R	0.999
HS0383	16ORCH	F	2	R	0.999
HS0384	16ORCH	F	0	R	0.999
HS0385	16ORCH	M	0	R	0.999
HS0386	16ORCH	M	3	R	0.999
HS0387	16ORCH	F	0	R	0.927
HS0388	16ORCH	M	4	R	0.999
HS0389	16ORCH	M	0	R	0.999
HS0390	12SOBN	F	0	R	0.991
HS0391	14LOCH	F	4	R	0.999
HS0392	14LOCH	M	3	R	0.999
HS0394	16ORCH	F	3	R	0.999

HS0395	16ORCH	M	2	R	0.999
HS0396	16ORCH	F	2	R	0.999
HS0397	16ORCH	M	6	R	0.999
HS0398	16ORCH	F	2	R	0.999
HS0399	16ORCH	F	3	R	0.999
HS0400	11SHRA	F	3	R	0.993
HS0402	08WAWE	M	5	R	0.875
HS0403	01LUSA	F	5	S	0.001
HS0404	03STAR	F	5	S	0.001
HS0405	11SHRA	F	1	R	0.999
HS0406	01LUSA	M	4	S	0.001
HS0407	08WAWE	F	4	R	0.251
HS0409	10ERDN	F	2	R	0.999
HS0410	02CARA	M	1	S	0.06
HS0411	02CARA	M	1	S	0.01
HS0412	01LUSA	F	3	S	0.001
HS0413	03STAR	F	3	S	0.001
HS0414	01LUSA	F	3	S	0.004
HS0416	01LUSA	F	1	S	0.001
HS0417	03STAR	F	0	S	0.001
HS0418	05KNAP	F	3	R	0.999
HS0419	08WAWE	F	2	R	0.999
HS0420	08WAWE	F	3	R	0.928
HS0421	03STAR	F	4	S	0.001
HS0422	02CARA	F	3	S	0.002
HS0423	01LUSA	F	7	R	0.001
HS0424	10ERDN	F	3	R	0.999
HS0428	11SHRA	F	0	R	0.999
HS0430	08WAWE	F	4	R	0.113
HS0431	08WAWE	F	4	R	0.095
HS0435	08WAWE	F	8	R	0.99
HS0436	11SHRA	F	0	R	0.999
HS0437	10ERDN	M	0	R	0.999
HS0438	02CARA	F	2	S	0.001
HS0439	07ORMG	F	5	R	0.833
HS0440	08WAWE	F	5	R	0.999
HS0441	08WAWE	F	0	R	0.805
HS0442	08WAWE	F	0	R	0.999
HS0443	11SHRA	F	1	R	0.999
HS0444	07ORMG	M	0	R	0.912
HS0445	10ERDN	F	5	R	0.952
HS0446	10ERDN	F	3	R	0.999
HS0447	11SHRA	F	7	R	0.999
HS0448	01LUSA	M	0	S	0.001
HS0449	11SHRA	F	5	R	0.999
HS0450	08WAWE	F	0	R	0.003
HS0451	01LUSA	F	1	S	0.001
HS0452	10ERDN	M	7	S	0.002
HS0453	08WAWE	F	4	R	0.673

HS0454	10ERDN	F	5	R	0.999
HS0455	11SHRA	F	0	R	0.999
HS0456	08WAWE	M	2	S	0.003
HS0457	08WAWE	F	4	R	0.999
HS0459	01LUSA	M	0	S	0.012
HS0461	01LUSA	M	1	S	0.001
HS0462	08WAWE	M	0	R	0.995
HS0463	05KNAP	F	1	R	0.986
HS0464	11SHRA	F	2	R	0.999
HS0465	10ERDN	F	3	R	0.999
HS0466	08WAWE	M	3	R	0.999
HS0467	08WAWE	F	3	R	0.995
HS0468	10ERDN	F	0	R	0.999
HS0469	10ERDN	M	0	R	0.999
HS0471	04WTAR	F	5	R	0.999
HS0472	10ERDN	F	0	R	0.999
HS0473	04WTAR	M	3	R	0.999
HS0475	01LUSA	F	2	S	0.007
HS0476	00DALB	M	1	S	0.611
HS0477	04WTAR	F	1	R	0.999
HS0478	04WTAR	F	1	R	0.993
HS0479	10ERDN	M	0	R	0.999
HS0480	08WAWE	F	3	R	0.748
HS0481	01LUSA	F	4	S	0.001
HS0482	07ORMG	M	4	S	0.001
HS0483	10ERDN	F	3	R	0.999
HS0484	08WAWE	M	0	R	0.999
HS0486	03STAR	F	4	S	0.001
HS0487	11SHRA	F	3	R	0.999
HS0489	01LUSA	F	5	R	0.999
HS0490	01LUSA	F	2	S	0.001
HS0491	10ERDN	F	2	R	0.999
HS0492	02CARA	F	5	S	0.001
HS0493	04WTAR	F	1	R	0.951
HS0494	08WAWE	F	2	R	0.999
HS0496	04WTAR	M	2	R	0.999
HS0497	08WAWE	M	1	R	0.822
HS0498	06KILM	F	5	R	0.999
HS0499	05KNAP	M	2	S	0.001
HS0501	05KNAP	M	1	S	0.001
HS0502	08WAWE	M	4	R	0.993
HS0503	01LUSA	F	6	S	0.001
HS0504	11SHRA	F	5	R	0.999
HS0505	08WAWE	F	3	R	0.999
HS0506	01LUSA	M	1	S	0.001
HS0507	08WAWE	F	1	R	0.219
HS0508	10ERDN	F	1	R	0.999
HS0510	04WTAR	M	1	R	0.999
HS0511	06KILM	M	0	S	0.002

HS0512	10ERDN	F	0	R	0.999
HS0514	05KNAP	F	0	R	0.949
HS0516	05KNAP	F	7	R	0.949
HS0517	02CARA	M	1	R	0.001
HS0520	04WTAR	M	1	R	0.985
HS0522	04WTAR	F	2	R	0.999
HS0523	04WTAR	F	3	R	0.984
HS0524	01LUSA	M	0	S	0.001
HS0525	11SHRA	F	1	R	0.999
HS0526	10ERDN	F	3	R	0.999
HS0528	08WAWE	F	5	R	0.848
HS0529	07ORMG	F	6	R	0.808
HS0531	11SHRA	F	5	R	0.999
HS0532	01LUSA	F	6	S	0.001
HS0533	11SHRA	F	8	R	0.999
HS0534	01LUSA	F	0	S	0.013
HS0535	01LUSA	F	0	S	0.005
HS0536	06KILM	F	1	S	0.008
HS0537	10ERDN	F	0	R	0.999
HS0538	05KNAP	M	0	S	0.002
HS0539	10ERDN	F	0	R	0.999
HS0540	05KNAP	F	2	R	0.999
HS0541	10ERDN	F	6	R	0.999
HS0544	08WAWE	F	5	R	0.323
HS0545	05KNAP	M	3	S	0.019
HS0546	01LUSA	F	2	S	0.001
HS0547	08WAWE	F	3	R	0.999
HS0548	05KNAP	F	2	R	0.999
HS0549	05KNAP	F	4	R	0.999
HS0550	06KILM	M	2	S	0.002
HS0551	06KILM	F	4	R	0.999
HS0552	08WAWE	F	2	R	0.999
HS0553	11SHRA	F	0	R	0.999
HS0554	02CARA	F	0	S	0.047
HS0555	10ERDN	F	3	R	0.999
HS0556	01LUSA	F	2	S	0.001
HS0557	05KNAP	M	1	S	0.001
HS0559	11SHRA	F	2	R	0.999
HS0560	10ERDN	F	0	R	0.999
HS0561	10ERDN	F	3	R	0.999
HS0562	05KNAP	F	5	S	0.002
HS0563	11SHRA	F	3	R	0.999
HS0564	10ERDN	F	1	S	0.001
HS0565	01LUSA	M	1	S	0.001
HS0566	11SHRA	F	2	R	0.999
HS0567	05KNAP	M	1	R	0.999
HS0568	10ERDN	F	4	R	0.999
HS0569	02CARA	M	3	S	0.013
HS0570	02CARA	M	4	S	0.001

HS0571	10ERDN	M	0	R	0.999
HS0572	01LUSA	F	4	S	0.014
HS0573	10ERDN	F	3	R	0.999
HS0574	05KNAP	M	3	S	0.004
HS0575	01LUSA	F	2	S	0.001
HS0576	11SHRA	F	3	R	0.999
HS0577	01LUSA	F	1	S	0.001
HS0578	08WAWE	F	3	R	0.908
HS0579	08WAWE	M	1	R	0.905
HS0580	10ERDN	F	2	R	0.999
HS0582	05KNAP	F	3	R	0.999
HS0584	05KNAP	F	2	S	0.001
HS0585	04WTAR	M	4	S	0.001
HS0586	10ERDN	F	3	R	0.999
HS0587	07ORMG	M	0	R	0.999
HS0588	08WAWE	F	3	R	0.999
HS0589	01LUSA	F	2	S	0.001
HS0591	11SHRA	F	1	R	0.999
HS0592	06KILM	F	0	R	0.999
HS0593	10ERDN	F	3	R	0.999
HS0594	01LUSA	F	3	S	0.001
HS0595	08WAWE	M	1	R	0.599
HS0597	10ERDN	F	2	R	0.999
HS0598	11SHRA	F	0	R	0.999
HS0599	11SHRA	F	8	R	0.999
HS0600	05KNAP	F	3	S	0.005
HS0602	11SHRA	F	8	R	0.999
HS0603	10ERDN	M	0	R	0.999
HS0604	11SHRA	F	2	R	0.999
HS0606	02CARA	M	1	S	0.003
HS0607	02CARA	M	4	S	0.048
HS0608	08WAWE	M	3	R	0.991
HS0609	04WTAR	F	4	R	0.999
HS0610	10ERDN	F	0	R	0.999
HS0611	10ERDN	F	2	R	0.999
HS0613	08WAWE	F	5	R	0.999
HS0614	08WAWE	M	3	R	0.243
HS0615	11SHRA	F	1	R	0.985
HS0616	08WAWE	M	2	R	0.524
HS0617	05KNAP	F	3	R	0.961
HS0618	11SHRA	F	3	R	0.993
HS0620	10ERDN	F	1	R	0.999
HS0621	01LUSA	F	2	R	0.001
HS0622	10ERDN	M		R	0.99
HS0623	02CARA	F	0	S	0.001
HS0624	08WAWE	M	0	R	0.277
HS0627	11SHRA	F	0	R	0.993
HS0628	05KNAP	F	3	R	0.999
HS0629	01LUSA	F	6	R	0.001

HS0630	01LUSA	M	1	S	0.001
HS0632	06KILM	M	3	R	0.999
HS0633	08WAWE	M	0	R	0.989
HS0634	04WTAR	M	5	S	0.001
HS0635	01LUSA	M	0	R	0.019
HS0636	08WAWE	M	0	R	0.75
HS0638	08WAWE	F	3	R	0.999
HS0641	01LUSA	F	1	R	0.001
HS0642	08WAWE	M	6	R	0.999
HS0643	10ERDN	F	1	R	0.999
HS0644	08WAWE	F	3	R	0.438
HS0645	11SHRA	F	10	R	0.999
HS0646	01LUSA	M	2	R	0.999
HS0647	08WAWE	F	3	R	0.172
HS0648	02CARA	F	5	S	0.003
HS0649	01LUSA	F	1	S	0.002
HS0650	05KNAP	F	1	R	0.99
HS0651	02CARA	F	0	S	0.002
HS0653	08WAWE	F	0	R	0.124
HS0654	10ERDN	F	2	R	0.999
HS0655	01LUSA	M	0	S	0.001
HS0656	10ERDN	F	0	S	0.002
HS0659	06KILM	F	4	R	0.999
HS0660	01LUSA	F	7	R	0.001
HS0662	01LUSA	M	0	S	0.001
HS0664	05KNAP	F	0	R	0.999
HS0665	04WTAR	M	1	R	0.999
HS0667	06KILM	F	0	R	0.999
HS0668	08WAWE	M	0	R	0.123
HS0669	10ERDN	M	1	R	0.999
HS0671	11SHRA	F	1	R	0.999
HS0672	01LUSA	M	0	R	0.001
HS0674	05KNAP	M	10	R	0.984
HS0675	05KNAP	F	0	R	0.925
HS0676	01LUSA	M	0	S	0.006
HS0677	08WAWE	F		R	0.926
HS0678	08WAWE	M	2	R	0.987
HS0679	10ERDN	F	2	R	0.999
HS0680	10ERDN	F	1	R	0.999
HS0681	04WTAR	M	3	S	0.001
HS0682	05KNAP	F	0	R	0.999
HS0684	01LUSA	F	6	S	0.006
HS0685	04WTAR	F	3	S	0.004
HS0686	08WAWE	F	2	R	0.999
HS0687	08WAWE	F	3	R	0.963
HS0688	08WAWE	F	0	R	0.094
HS0689	04WTAR	F	2	S	0.001
HS0690	01LUSA	F	5	S	0.001
HS0691	06KILM	F	0	S	0.001

HS0693	04WTAR	M	1	S	0.001
HS0694	10ERDN	M	3	S	0.001
HS0695	11SHRA	F	3	R	0.999
HS0698	08WAWE	M	3	R	0.999
HS0699	08WAWE	M	0	R	0.941
HS0700	08WAWE	F	2	R	0.96
HS0701	01LUSA	F	1	S	0.001
HS0702	10ERDN	F	0	R	0.999
HS0703	01LUSA	F	5	S	0.001
HS0705	10ERDN	F	3	R	0.999
HS0706	11SHRA	F	2	R	0.999
HS0707	01LUSA	F	5	S	0.001
HS0710	10ERDN	F	6	R	0.999
HS0711	10ERDN	F	6	R	0.999
HS0712	08WAWE	M	2	R	0.993
HS0713	11SHRA	F	5	R	0.999
HS0714	10ERDN	F	2	R	0.999
HS0715	08WAWE	M	0	R	0.172
HS0716	01LUSA	F	2	S	0.001
HS0718	10ERDN	F	3	R	0.999
HS0719	08WAWE	M	2	R	0.999
HS0720	10ERDN	M	9	R	0.999
HS0721	08WAWE	F		R	0.999
HS0724	15BENM	M	0	R	0.999
HS0725	15BENM	F	0	R	0.999
HS0726	15BENM	M	4	R	0.999
HS0727	15BENM	F	0	R	0.999
HS0728	15BENM	F	2	R	0.999
HS0729	14LOCH	F	2	R	0.999
HS0730	14LOCH	M	2	R	0.999
HS0731	14LOCH	F	6	R	0.999
HS0733	13SUCC	M	4	R	0.985
HS0734	14LOCH	M	4	R	0.999
HS0735	15BENM	F	4	R	0.999
HS0736	14LOCH	F	0	R	0.999
HS0737	13SUCC	M	2	R	0.999
HS0738	14LOCH	M	0	R	0.999
HS0739	15BENM	F	5	R	0.999
HS0740	13SUCC	M	5	R	0.999
HS0741	15BENM	F	2	R	0.999
HS0742	14LOCH	M	2	R	0.999
HS0743	14LOCH	F	3	R	0.999
HS0744	14LOCH	F	0	R	0.999
HS0745	15BENM	F	5	R	0.999
HS0748	15BENM	F	5	R	0.999
HS0751	14LOCH	F	5	R	0.999
HS0752	15BENM	M	2	R	0.999
HS0753	14LOCH	F	4	R	0.999
HS0754	16ORCH	F	2	R	0.999

HS0755	14LOCH	M	2	R	0.999
HS0756	15BENM	F	5	R	0.999
HS0757	14LOCH	M	0	R	0.999
HS0759	15BENM	F	3	R	0.999
HS0760	14LOCH	F	5	R	0.999
HS0761	14LOCH	F	4	R	0.999
HS0762	15BENM	M	4	R	0.999
HS0763	14LOCH	F	6	R	0.999
HS0764	13SUCC	M	4	R	0.999
HS0765	20BALA	M	3	R	0.999
HS0766	20BALA	F	3	R	0.999
HS0767	20BALA	F	6	R	0.999
HS0768	18BARR	F	0	R	0.999
HS0769	20BALA	M	2	R	0.999
HS0770	17NOBN	M	8	R	0.999
HS0771	15BENM	M	3	R	0.999
HS0772	20BALA	M		R	0.999
HS0773	20BALA	M		R	0.999
HS0774	17NOBN	M		R	0.999
HS0775	17NOBN	M		R	0.999
HS0776	13SUCC	F	2	R	0.999
HS0777	13SUCC	F	4	R	0.999
HS0778	15BENM	M	2	R	0.999
HS0780	17NOBN	M	2	R	0.999
HS0781	14LOCH	F	5	R	0.999
HS0782	14LOCH	M	2	R	0.999
HS0784	14LOCH	F	0	R	0.999
HS0785	14LOCH	M	4	R	0.999
HS0787	16ORCH	F		R	0.999
HS0788	14LOCH	M	0	R	0.999
HS0789	14LOCH	F	1	R	0.999
HS0790	17NOBN	F	4	R	0.999
HS0791	20BALA	M	0	R	0.999
HS0792	20BALA	M	2	R	0.999
HS0793	20BALA	F	2	R	0.999
HS0794	13SUCC	F	0	R	0.999
HS0795	20BALA	F	0	R	0.999
HS0796	15BENM	F	5	R	0.999
HS0797	14LOCH	F	3	R	0.999
HS0798	16ORCH	F	5	R	0.999
HS0799	16ORCH	F	3	R	0.999
HS0800	20BALA	M	3	R	0.999
HS0801	17NOBN	M	1	R	0.999
HS0802	17NOBN	M	5	R	0.999
HS0803	13SUCC	F	5	R	0.999
HS0804	13SUCC	F	3	R	0.999
HS0805	15BENM	M	4	R	0.999
HS0807	13SUCC	M	4	R	0.999
HS0808	14LOCH	F	0	R	0.999

HS0809	16ORCH	M	0	R	0.999
HS0810	17NOBN	F	4	R	0.999
HS0811	13SUCC	M	4	S	0.002
HS0812	17NOBN	M	0	R	0.999
HS0813	13SUCC	M	3	R	0.999
HS0814	15BENM	M	5	R	0.999
HS0815	17NOBN	F	2	R	0.927
HS0818	20BALA	M	2	R	0.999
HS0819	16ORCH	F	3	R	0.999
HS0820	20BALA	M	3	R	0.993
HS0821	17NOBN	M	3	R	0.999
HS0822	13SUCC	M	0	R	0.999
HS0824	17NOBN	M	0	R	0.999
HS0825	14LOCH	M	0	R	0.999
HS0827	14LOCH	M	2	R	0.999
HS0828	13SUCC	M	2	R	0.999
HS0831	16ORCH	F	0	R	0.999
HS0832	16ORCH	M	0	R	0.999
HS0833	14LOCH	M	6	R	0.999
HS0834	17NOBN	M	3	R	0.999
HS0835	14LOCH	M	5	R	0.999
HS0836	17NOBN	M	7	R	0.999
HS0837	17NOBN	F	5	R	0.999
KA008	06KILM	F		S	0.002
KA009	06KILM	M		S	0.001
KA010	06KILM	M		S	0.002
KA011	02CARA	M		S	0.001
KA012	02CARA	F		S	0.001
KA017	06KILM	F		R	0.999
KA019	06KILM	M		R	0.993
KA020	06KILM	F		R	0.999
KA021	06KILM	F		R	0.999
KA022	06KILM	F		R	0.993
KA023	02CARA	F		R	0.999
KA026	02CARA	M		R	0.999
KA029	02CARA	M		R	0.999
KA032	02CARA	M		S	0.001
KA033	02CARA	M		S	0.009
KA034	02CARA	F		S	0.001
KA035	02CARA	F		S	0.002
KA036	02CARA	F		S	0.001
KA037	02CARA	F		S	0.001
KA038	02CARA	F		S	0.002
KA039	02CARA	M		S	0.019
KA041	02CARA	M		S	0.001
KA043	02CARA	F		S	0.001
KA044	02CARA	M		S	0.009
KA045	02CARA	F		S	0.001
KA046	02CARA	M		S	0.002

KA047	02CARA	F		S	0.001
KA048	02CARA	F		S	0.02
KA049	02CARA	F		S	0.001
KA050	02CARA	M		S	0.001
KA053	05KNAP	M		R	0.999
KA057	05KNAP	M		S	0.001
KA059	05KNAP	F		R	0.999
KA060	05KNAP	M		R	0.999
KA061	02CARA	M		S	0.002
KA062	02CARA	M		S	0.001
KA081	02CARA	F		S	0.001
KA082	02CARA	F		S	0.002
KA084	02CARA	M		R	0.669
KA085	02CARA	M		S	0.004
KA086	02CARA	F		S	0.001
KA087	02CARA	M		S	0.002
KA088	02CARA	F		S	0.03
KA092	02CARA	F		S	0.001
KA094	02CARA	M		S	0.001
KA095	02CARA	F		S	0.001
KA096	02CARA	F		R	0.037
KA097	02CARA	F		S	0.001
KA098	02CARA	F		S	0.001
KA100	02CARA	M		S	0.001
KA101	10ERDN	M		R	0.999
KA105	10ERDN	F		R	0.999
KA107	10ERDN	M		R	0.999
KA109	10ERDN	F		R	0.999
KA121	10ERDN	F		R	0.999
KA122	10ERDN	F		R	0.999
KA123	10ERDN	M		R	0.999
KA124	10ERDN	F		R	0.999
KA125	10ERDN	M		R	0.999
KA126	10ERDN	F		R	0.999
KA127	10ERDN	M		R	0.999
KA128	10ERDN	F		R	0.999
KA129	10ERDN	F		R	0.999
KA130	10ERDN	M		R	0.999
KA131	10ERDN	F		R	0.999
KA132	10ERDN	F		R	0.999
KA133	10ERDN	M		R	0.999
KA134	10ERDN	M		R	0.999
KA135	10ERDN	F		R	0.999
KA136	10ERDN	F		R	0.999
KA139	10ERDN	M		R	0.999
KA140	10ERDN	F		R	0.999
KA141	10ERDN	F		R	0.999
KA143	10ERDN	M		R	0.999
KA144	10ERDN	M		R	0.999

KA145	10ERDN	F		R	0.999
KA146	10ERDN	F		R	0.999
KA147	10ERDN	F		R	0.999
KA148	10ERDN	F		R	0.999
KA150	10ERDN	F		R	0.999
KA351	05KNAP	F		R	0.999
KA362	04WTAR	M		R	0.927
KA363	05KNAP	M		S	0.002
KA364	05KNAP	F		R	0.988
KA365	05KNAP	x		R	0.993
KA369	04WTAR	M		S	0.001
KA371	05KNAP	x		S	0.002
KA372	05KNAP	x		S	0.002
KA375	05KNAP	M		S	0.001
KA376	05KNAP	F		S	0.003
KA377	04WTAR	M		S	0.001
KA378	04WTAR	M		S	0.014
KA379	04WTAR	M		R	0.999
KA380	04WTAR	F		S	0.001
KA381	04WTAR	M		S	0.002
KA383	04WTAR	M		R	0.999
KA384	04WTAR	F		S	0.001
KA385	04WTAR	F		S	0.004
KA386	05KNAP	F		S	0.001
KA387	05KNAP	M		S	0.001
KA388	05KNAP	F		S	0.001
KA389	05KNAP	F		R	0.999
KA391	04WTAR	x		S	0.001
KA392	04WTAR	M		R	0.999
KA393	05KNAP	F		R	0.999
KA394	05KNAP	F		S	0.013
KA395	05KNAP	F		R	0.999
KA396	05KNAP	F		R	0.999
KA398	04WTAR	F		S	0.01
KA399	04WTAR	M		S	0.001
KA400	04WTAR	M		S	0.133
KA411	05KNAP	M		R	0.999
KA413	05KNAP	F		S	0.002
KA414	05KNAP	F		S	0.045
KA415	05KNAP	F		S	0.001
KA416	04WTAR	F		R	0.001
KA417	04WTAR	x		R	0.999
KA418	04WTAR	M		R	0.999
KA419	04WTAR	M		S	0.001
KA420	04WTAR	M		R	0.98
KA422	04WTAR	M		S	0.001
KA427	04WTAR	F		S	0.003
KA428	04WTAR	F		S	0.001
KA429	04WTAR	M		S	0.001

SGCOW002	23SCOW	F	1.5	R	0.999
SGCOW006	24ECOW	F	10.5	R	0.999
SGCOW007	23SCOW	F	4	R	0.999
SGCOW010	24ECOW	F	8.5	R	0.999
SGCOW018	23SCOW	F	3	R	0.999
SGCOW023	25DRUL	F	5	R	0.999
SGCOW037	22NCOW	F		R	0.999
SGCOW044	22NCOW	F		R	0.999
SGCOW045	22NCOW	F		R	0.999
SGCOW047	23SCOW	F	1.5	R	0.999
SGCOW051	23SCOW	F	1.5	R	0.999
SGCOW053	22NCOW	F		R	0.999
SGCOW060	22NCOW	F		R	0.999
SGCOW064	23SCOW	F	6	R	0.999
SGCOW065	24ECOW	F	5.5	R	0.908
SGCOW068	25DRUL	F	0.5	R	0.999
SGCOW071	23SCOW	F	4	R	0.999
SGCOW075	24ECOW	F	4.5	R	0.999
SGCOW076	24ECOW	F	2.5	R	0.999
SGCOW078	24ECOW	F	4.5	R	0.999
SGCOW079	24ECOW	F		R	0.999
SGCOW080	24ECOW	F	1.5	R	0.999
SGCOW081	24ECOW	F	1.5	R	0.999
SGCOW083	23SCOW	F	4	R	0.999
SGCOW091	23SCOW	F	2	R	0.999
SGCOW092	24ECOW	F	5.5	R	0.999
SGCOW094	25DRUL	F	4	R	0.999
SGCOW095	23SCOW	F	4	R	0.999
SGCOW129	23SCOW	F	2	R	0.999
SGCOW139	25DRUL	F		S	0.001
SGLGP001	06KILM	F	2	R	0.999
SGLGP002	01LUSA	F	1	S	0.001
SGLGP003	10ERDN	M	3	R	0.999
SGLGP004	10ERDN	M	2	R	0.999
SGLGP005	06KILM	F	2	R	0.999
SGLGP007	01LUSA	F	3	S	0.004
SGLGP008	10ERDN	F	2	S	0.001
SGLGP009	05KNAP	F	0	S	0.027
SGLGP010	08WAWE	F	0	R	0.637
SGLGP011	01LUSA	F	2	S	0.001
SGLGP012	10ERDN	F	6	R	0.999
SGLGP013	10ERDN	M	1	R	0.999
SGLGP015	06KILM	F	0	R	0.999
SGLGP016	11SHRA	F	2	R	0.999
SGLGP017	06KILM	F	0	R	0.999
SGLGP018	10ERDN	F	7	R	0.999
SGLGP019	01LUSA	F	4	S	0.001
SGLGP020	01LUSA	F	2	S	0.001
SGLGP021	10ERDN	F	2	R	0.999

SGLGP022	01LUSA	F	2	R	0.999
SGLGP023	08WAWE	F	0	R	0.962
SGLGP024	10ERDN	F	0	R	0.999
SGLGP025	10ERDN	F	2	R	0.999
SGLGP026	06KILM	F	5	R	0.999
SGLGP027	08WAWE	F	2	R	0.96
SGLGP028	10ERDN	M	2	R	0.999
SGLGP029	10ERDN	M	2	R	0.999
SGLGP030	10ERDN	F	2	R	0.999
SGLGP031	06KILM	F	2	R	0.999
SGLGP032	08WAWE	F	8	R	0.466
SGLGP033	08WAWE	F	2	R	0.999
SGLGP034	11SHRA	F	0	R	0.999
SGLGP035	05KNAP	F	2	S	0.001
SGLGP037	10ERDN	F	2	R	0.999
SGLGP038	05KNAP	F	3	R	0.999
SGLGP039	11SHRA	F	2	R	0.999
SGLGP040	10ERDN	F		R	0.999
SGLGP041	06KILM	F	0	S	0.002
SGLGP042	01LUSA	F	1	S	0.001
SGLGP043	06KILM	F	3	R	0.999
SGLGP044	10ERDN	F	2	R	0.999
SGLGP045	08WAWE	M	2	R	0.823
SGLGP046	08WAWE	F	2	R	0.257
SGLGP047	10ERDN	F	2	R	0.999
SGLGP048	06KILM	F	2	S	0.001
SGLGP049	01LUSA	F	3	S	0.002
SGLGP050	10ERDN	F	2	R	0.999
SGLGP051	11SHRA	F	0	R	0.999
SGLGP052	10ERDN	F	3	R	0.999
SGLGP053	10ERDN	F	2	R	0.999
SGLGP054	05KNAP	F	1	R	0.993
SGLGP055	10ERDN	F	2	R	0.999
SGLGP057	06KILM	F	0	R	0.999
SGLGP058	10ERDN	M	3	R	0.982
SGLGP059	06KILM	F	2	R	0.938
SGLGP060	10ERDN	F	2	R	0.999
SGLGP061	06KILM	F	4	R	0.999
SGLGP062	11SHRA	F	2	R	0.999
SGLGP063	01LUSA	F	2	S	0.001
SGLGP064	05KNAP	F	0	R	0.999
SGLGP065	06KILM	F	2	S	0.001
SGLGP067	05KNAP	F	0	R	0.999
SGLGP068	10ERDN	F	5	R	0.999
SGLGP069	01LUSA	F	2	S	0.001
SGLGP070	11SHRA	F	2	R	0.999
SGLGP071	06KILM	F	0	R	0.999
SGLGP072	01LUSA	F	2	S	0.001
SGLGP073	06KILM	F	5	S	0.002

SGLGP074	06KILM	F	2	R	0.999
SGLGP075	08WAWE	F	3	R	0.874
SGLGP076	06KILM	F	1	R	0.999
SGLGP077	10ERDN	F	0	R	0.999
SGLGP078	06KILM	F	3	R	0.999
SGLGP079	01LUSA	F	2	S	0.001
SGLGP080	10ERDN	F	4	R	0.999
SGLGP082	11SHRA	F	2	R	0.999
SGLGP083	05KNAP	F	3	S	0.012
SGLGP084	10ERDN	F	4	R	0.999
SGLGP085	10ERDN	M	2	R	0.869
SGLGP086	08WAWE	M	1	R	0.95
SGLGP087	05KNAP	F		R	0.999
SGLGP088	08WAWE	F	0	R	0.989
SGLGP089	11SHRA	F	0	R	0.999
SGLGP090	08WAWE	F	4	R	0.999
SGLGP091	03STAR	F	2	S	0.001
SGLGP092	05KNAP	F	2	S	0.002
SGLGP093	02CARA	F	4	S	0.004
SGLGP094	04WTAR	F	0	R	0.999
SGLGP095	05KNAP	F	2	S	0.001
SGLGP096	01LUSA	F	2	S	0.001
SGLGP097	01LUSA		0	S	0.002
SGLGP099	02CARA	F	2	S	0.001
SGLGP100	02CARA	F	3	S	0.001
SGLGP101	01LUSA	M	7	S	0.001
SGLGP102	03STAR	F	0	S	0.001
SGLGP103	05KNAP	F	0	S	0.001
SGLGP104	10ERDN	F	0	R	0.999
SGLGP105	10ERDN	F	3	R	0.999
SGLGP106	05KNAP	F	0	R	0.999
SGLGP107	05KNAP	F	0	R	0.999
SGLGP108	01LUSA	F		S	0.001
SGLGP109	05KNAP	F	2	S	0.001
SGLGP110	01LUSA	F	5	S	0.001
SGLGP111	01LUSA	F	2	S	0.001
SGLGP112	11SHRA	F	0	R	0.999
SGLGP114	06KILM	F	2	R	0.999
SGLGP115	05KNAP	F	2	R	0.999
SGLGP116	10ERDN	F	2	R	0.999
SGLGP117	05KNAP	F	4	R	0.999
SGLGP118	10ERDN	F	4	R	0.999
SGLGP119	01LUSA	M	4	S	0.091
SGLGP120	11SHRA	F	2	R	0.999
SGLGP121	05KNAP	F	2	R	0.999
SGLGP122	06KILM	F	5	R	0.999
SGLGP123	10ERDN	F	2	R	0.999
SGLGP124	01LUSA	M	2	S	0.001
SGLGP125	06KILM	F	2	R	0.999

SGLGP126	05KNAP	F	0	S	0.001
SGLGP127	06KILM	F	0	R	0.999
SGLGP128	10ERDN	M	1	R	0.999
SGLGP129	10ERDN	F	0	R	0.999
SGLGP130	08WAWE	F	3	R	0.381
SGLGP131	05KNAP	F	2	S	0.001
SGLGP132	10ERDN	F	4	R	0.999
SGLGP133	01LUSA	F	2	R	0.001
SGLGP134	10ERDN	F	2	R	0.999
SGLGP135	10ERDN	F	4	R	0.999
SGLGP136	10ERDN	F	2	R	0.999
SGLGP137	05KNAP	F	0	R	0.108
SGLGP138	01LUSA		6	R	0.986
SGLGP139	10ERDN	F	0	R	0.999
SGLGP140	10ERDN	F	2	R	0.999
SGLGP141	11SHRA	F	0	R	0.999
SGLGP142	04WTAR	F	2	R	0.99
SGLGP143	08WAWE	F	5	R	0.999
SGLGP144	02CARA	F	2	S	0.001
SGLGP145	04WTAR	F	0	R	0.999
SGLGP146	10ERDN	M	2	R	0.999
SGLGP147	05KNAP	F	2	S	0.001
SGLGP148	10ERDN	F	4	R	0.999
SGLGP149	08WAWE	M	1	R	0.999
SGLGP150	045NTAR	F	0	R	0.999
SGLGP151	10ERDN	F	0	R	0.999
SGLGP152	05KNAP	F	0	S	0.125
SGLGP153	04WTAR	F	2	R	0.999
SGLGP154	02CARA	F	2	S	0.001
SGLGP155	045NTAR	F	2	R	0.999
SGLGP157	01LUSA	F	1	S	0.001
SGLGP158	06KILM	F	2	S	0.001
SGLGP159	05KNAP	F	2	S	0.001
SGLGP160	10ERDN	F	0	R	0.999
SGLGP161	06KILM	F	0	R	0.999
SGLGP162	01LUSA	F	1	S	0.001
SGLGP163	06KILM	F		S	0.001
SGLGP164	10ERDN	F	2	R	0.999
SGLGP165	10ERDN	F	2	R	0.999
SGLGP166	06KILM	F	0	R	0.993
SGLGP167	10ERDN	F	0	R	0.999
SGLGP168	06KILM	F	0	R	0.999
SGLGP169	03STAR	F	2	S	0.001
SGLGP170	06KILM	F	2	S	0.001
SGLGP171	10ERDN	M	4	R	0.999
SGLGP172	10ERDN	F	0	R	0.999
SGLGP173	10ERDN	F	0	R	0.999
SGLGP174	01LUSA	F	2	S	0.01
SGLGP175	05KNAP			R	0.999

SGLGP176	06KILM	F	0	S	0.002
SGLGP177	10ERDN	M	1	R	0.999
SGLGP178	11SHRA	F	2	R	0.999
SGLGP179	10ERDN	F	1	R	0.999
SGLGP180	10ERDN	F	0	R	0.999
SGLGP181	06KILM	F	0	R	0.999
SGLGP182	06KILM	F	2	R	0.999
SGLGP183	04WTAR	F	3	R	0.983
SGLGP184	05KNAP	F	2	R	0.002
SGLGP185	06KILM	F	2	R	0.999
SGLGP186	05KNAP	F	0	S	0.001
SGLGP187	10ERDN	F	0	R	0.999
SGLGP188	06KILM	F	3	R	0.999
SGLGP189	05KNAP	F	0	S	0.001
SGLGP190	06KILM	F		S	0.001
SGLGP191	10ERDN	F	0	R	0.999
SGLGP192	06KILM	F	0	R	0.999
SGLGP193	06KILM	F	0	R	0.999
SGLGP194	10ERDN	F	0	R	0.999
SGLGP195	06KILM	F	0	S	0.001
SGLGP196	06KILM	F	2	S	0.001
SGLGP197	10ERDN	F	2	R	0.999
SGLGP198	06KILM	F	2	S	0.002
SGLGP199	01LUSA	F	2	R	0.999
SGLGP200	045NTAR	F	2	R	0.999
SGLGP201	06KILM	F	5	R	0.999
SGLGP202	10ERDN	F	2	R	0.999
SGLGP203	10ERDN	F	2	R	0.999
SGLGP204	08WAWE	F	2	R	0.999
SGLGP205	10ERDN	F	0	R	0.999
SGLGP206	04WTAR	F	0	S	0.002
SGLGP207	06KILM	M	2	S	0.001
SGLGP208	10ERDN	F	1	R	0.999
SGLGP209	10ERDN	F	0	R	0.999
SGLGP210	01LUSA	F	2	S	0.028
SGLGP211	02CARA	F	0	S	0.001
SGLGP212	03STAR	F	0	S	0.001
SGLGP213	08WAWE	M	1	R	0.96
SGLGP214	01LUSA	F	2	S	0.001
SGLGP215	11SHRA	F	2	R	0.999
SGLGP216	10ERDN	M	1	R	0.999
SGLGP217	04WTAR	F	0	R	0.999
SGLGP218	01LUSA	F	0	R	0.999
SGLGP219	06KILM	F	0	S	0.001
SGLGP220	06KILM	F	2	S	0.001
SGLGP221	01LUSA	F	2	S	0.19
SGLGP222	05KNAP	F	2	R	0.999
SGLGP223	05KNAP	F	2	S	0.001
SGLGP224	10ERDN	F	2	R	0.999

SGLGP225	01LUSA	M	1	R	0.002
SGLGP226	10ERDN	F	2	R	0.999
SGLGP228	10ERDN	F	0	R	0.999
SGLGP230	04WTAR	F	0	R	0.999
SGLGP231	06KILM	F	2	S	0.001
SGLGP232	10ERDN	F	1	R	0.999
SGLGP233	02CARA	F	2	S	0.001
SGLGP234	08WAWE	F	3	R	0.999
SGLGP235	11SHRA	F	0	R	0.999
SGLGP236	01LUSA	M	2	S	0.001
SGLGP237	02CARA	F	2	S	0.002
SGLGP238	04WTAR	F	2	S	0.001
SGLGP240	06KILM	F	0	S	0.001
SGLGP241	05KNAP	F	2	R	0.999
SGLGP242	01LUSA	F	2	S	0.001
SGLGP243	02CARA	F	2	S	0.014
SGLGP244	08WAWE	F	2	R	0.999
SGLGP245	04WTAR	F	0	R	0.927
SGLGP246	01LUSA	M	2	S	0.001
SGLGP247	10ERDN	F	0	R	0.999
SGLGP248	05KNAP	F	2	S	0.001
SGLGP249	11SHRA	F	2	R	0.999
SGLGP250	11SHRA	F	2	R	0.999
SGLGP251	08WAWE	F	5	R	0.999
SGLGP252	03STAR	F	0	S	0.004
SGLGP253	01LUSA	F	2	S	0.001
SGLGP254	10ERDN	F	2	R	0.999
SGLGP255	03STAR	F	0	S	0.001
SGLGP256	04WTAR	F	3	S	0.002
SGLGP257	08WAWE	F	4	R	0.989
SGLGP258	06KILM	F	0	R	0.999
SGLGP259	05KNAP	F	0	S	0.001
SGLGP260	02CARA	F	0	S	0.001
SGLGP261	01LUSA	F	2	S	0.001
SGLGP262	10ERDN	F	0	R	0.999
SGLGP263	04WTAR	F	2	R	0.999
SGLGP264	05KNAP	F	5	R	0.999
SGLGP265	10ERDN	F	5	R	0.999
SGLGP266	01LUSA	M	2	S	0.001
SGLGP267	10ERDN	F	4	R	0.999
SGLGP268	06KILM	F	2	S	0.098
SGLGP269	06KILM	F	2	R	0.987
SGLGP270	01LUSA	F	2	S	0.001
SGLGP271	01LUSA	F	2	S	0.001
SGLGP272	11SHRA	F	0	R	0.999
SGLGP273	10ERDN	F	4	R	0.999
SGLGP274	01LUSA	F	2	S	0.001
SGLGP275	01LUSA	F	2	S	0.001
SGLGP276	01LUSA	F		S	0.001

SGLGP277	03STAR	F	2	S	0.005
SGLGP278	01LUSA	F	2	S	0.001
SGLGP280	02CARA	F	2	S	0.001
SGLGP281	01LUSA	F	0	S	0.001
SGLGP282	02CARA	F	0	S	0.002
SGLGP283	06KILM	F	2	S	0.001
SGLGP284	10ERDN	F	3	R	0.999
SGLGP285	06KILM	M	1	R	0.999
SGLGP286	08WAWE	F	0	R	0.374
SGLGP287	10ERDN	F	1	R	0.999
SGLGP288	01LUSA	F	0	S	0.001
SGLGP289	08WAWE	F	2	R	0.993
SGLGP290	08WAWE	F	5	R	0.99
SGLGP291	08WAWE	F	1	R	0.911
SGLGP292	01LUSA	F	2	S	0.001
SGLGP293	10ERDN	F	4	R	0.988
SGLGP294	06KILM	F	0	R	0.999
SGLGP295	01LUSA	F	0	S	0.001
SGLGP296	01LUSA	F	5	S	0.001
SGLGP297	02CARA	F	2	S	0.004
SGLGP298	01LUSA	F	2	S	0.001
SGLGP299	06KILM	F	2	R	0.999
SGLGP300	01LUSA	F	0	S	0.001
SGLGP301	01LUSA	F	2	R	0.002
SGLGP302	01LUSA	F	2	S	0.002
SGLGP303	06KILM	F	6	R	0.999
SGLGP304	04WTAR	F	2	R	0.942
SGLGP305	10ERDN	F	0	R	0.999
SGLGP306	10ERDN	F	0	R	0.999
SGLGP307	01LUSA	F	0	S	0.001
SGLGP308	02CARA	F	2	S	0.001
SGLGP309	06KILM	F	2	S	0.001
SGLGP310	11SHRA	F	2	R	0.999
SGLGP311	05KNAP	F	0	R	0.999
SGLGP312	05KNAP	F	0	R	0.989
SGLGP313	10ERDN	F	2	R	0.999
SGLGP314	06KILM	F	2	R	0.999
SGLGP315	02CARA	F	0	S	0.094
SGLGP316	04WTAR	F	0	R	0.999
SGLGP317	01LUSA	F	2	R	0.001
SGLGP318	03STAR	F	2	S	0.004
SGLGP319	01LUSA	F	2	S	0.001
SGLGP320	08WAWE	F		R	0.895
SGLGP321	02CARA	F	2	S	0.035
SGLGP322	05KNAP	F		S	0.004
SGLGP323	04WTAR	F	2	S	0.001
SGLGP324	03STAR	F	0	S	0.001
SGLGP325	06KILM	F	0	R	0.999
SGLGP326	11SHRA	F	2	R	0.999

SGLGP327	06KILM	F	0	S	0.002
SGLGP328	05KNAP	F	2	S	0.015
SGLGP329	04WTAR	F	2	S	0.001
SGLGP330	05KNAP	F	0	R	0.988
SGLGP331	045NTAR	F	2	R	0.999
SGLGP332	04WTAR	F	0	S	0.001
SGLGP333	04WTAR	F	0	S	0.001
SGLGP334	04WTAR	F	0	R	0.999
SGLGP335	05KNAP	F	0	S	0.001
SGLGP336	06KILM	F	2	S	0.004
SGLGP337	05KNAP	F	2	S	0.009
SGLGP338	11SHRA	F	2	R	0.999
SGLGP339	02CARA	F	2	S	0.002
SGLGP340	01LUSA	M	2	S	0.001
SGLGP341	04WTAR	F	2	R	0.999
SGLGP342	05KNAP	F	2	S	0.022
SGLGP343	01LUSA	M	2	S	0.095
SGLGP344	05KNAP	F	2	R	0.999
SGLGP345	045NTAR	F		R	0.999
SGLGP346	02CARA	F	0	S	0.001
SGLGP347	03STAR	F	2	S	0.001
SGLGP348	02CARA	F	2	S	0.001
SGLGP349	045NTAR	F	2	R	0.999
SGLGP350	06KILM	F		R	0.993
SGLGP351	045NTAR	F	0	R	0.998
SGLGP352	04WTAR	M	1	R	0.988
SGLGP353	06KILM	F	2	S	0.001
SGLGP354	05KNAP	F	2	S	0.001
SGLGP355	01LUSA	M	1	S	0.071
SGLGP356	06KILM	F	2	R	0.999
SGLGP357	04WTAR	F	2	R	0.983
SGLGP358	05KNAP	F	2	S	0.001
SGLGP359	11SHRA	F	2	R	0.999
SGLGP360	06KILM	F	0	S	0.001
SGLGP361	01LUSA	F	2	S	0.001
SGLGP362	05KNAP	F	2	S	0.016
SGLGP363	04WTAR	F	2	S	0.001
SGLGP364	06KILM	F	2	S	0.001
SGLGP365	045NTAR	F	0	S	0.001
SGLGP366	04WTAR	F	0	R	0.999
SGLGP367	045NTAR	F	2	S	0.001
SGLGP368	03STAR	F	2	S	0.001
SGLGP369	08WAWE	F	0	R	0.999
SGLGP370	06KILM	F	0	S	0.002
SGLGP371	06KILM	M	2	S	0.002
SGLGP372	05KNAP		1	R	0.999
SGLGP373	10ERDN	F	2	R	0.925
SGLGP374	06KILM	F	0	R	0.999
SGLGP375	08WAWE	F	2	R	0.89

SGLGP376	06KILM	F	2	S	0.001
SGLGP377	01LUSA	F	2	S	0.001
SGLGP378	06KILM	F	2	R	0.999
SGLGP380	06KILM	F	0	R	0.999
SGLGP381	03STAR	F	2	S	0.001
SGLGP382	08WAWE	F	3	R	0.006
SGLGP384	08WAWE	F	0	R	0.995
SGLGP385	06KILM	F	2	S	0.001
SGLGP386	01LUSA	F	2	S	0.001
SGLGP387	04WTAR	F	2	S	0.001
SGLGP388	01LUSA	F	2	S	0.001
SGLGP389	04WTAR	F	0	S	0.001
SGLGP390	06KILM	F	0	R	0.999
SGLGP391	06KILM	F	2	S	0.002
SGLGP392	05KNAP	F	0	S	0.05
SGLGP393	06KILM	F		S	0.022
SGLGP394	06KILM	M	0	S	0.002
SGLGP395	06KILM	F	5	S	0.001
SGLGP396	08WAWE	F	2	R	0.99
SGLGP398	04WTAR	F	2	S	0.001
SGLGP399	06KILM	F	2	R	0.999
SGLGP400	08WAWE	F	2	R	0.94
SGLGP401	08WAWE	M	1	R	0.294
SGLGP402	02CARA	F	2	S	0.001
SGLGP403	05KNAP	F	0	R	0.999
SGLGP404	06KILM	F	0	R	0.999
SGLGP405	06KILM	F	2	R	0.999
SGLGP406	07ORMG	F	3	S	0.008
SGLGP407	06KILM	F	2	R	0.999
SGLGP408	01LUSA	F	2	S	0.022
SGLGP409	06KILM	M	2	S	0.001
SGLGP410	04WTAR	F	2	R	0.992
SGLGP411	04WTAR	F	1	R	0.993
SGLGP412	04WTAR	F	2	R	0.003
SGLGP413	04WTAR	F	2	S	0.001
SGLGP414	01LUSA	M	1	S	0.001
SGLGP415	06KILM	F	0	S	0.002
SGLGP416	06KILM	F	2	R	0.999
SGLGP417	06KILM	F	2	R	0.999
SGLGP418	06KILM	M	2	S	0.001
SGLGP419	08WAWE	F	2	R	0.329
SGLGP420	08WAWE	F	0	R	0.727
SGLGP421	04WTAR	F	2	S	0.001
SGLGP422	01LUSA	F	2	S	0.001
SGLGP423	06KILM	F	0	R	0.999
SGLGP424	06KILM	F	2	R	0.999
SGLGP425	06KILM	F	6	S	0.001
SGLGP426	04WTAR	F	2	R	0.999
SGLGP427	08WAWE	F	3	R	0.999

SGLGP428	01LUSA	F	2	R	0.167
SGLGP429	05KNAP		2	R	0.954
SGLGP430	045NTAR	F	2	R	0.999
SGLGP431	04WTAR	F	0	S	0.001
SGLGP432	04WTAR	F	1	S	0.001
SGLGP433	06KILM	F	2	R	0.999
SGLGP434	06KILM	F	2	R	0.999
SGLGP435	08WAWE	F	3	R	0.999
SGLGP436	08WAWE	F	2	R	0.995
SGLGP437	05KNAP	F	2	R	0.999
SGLGP438	03STAR	F	2	S	0.001
SGLGP439	04WTAR	F	0	R	0.999
SGLGP440	01LUSA	F	2	S	0.001
SGLGP441	06KILM	F		S	0.001
SGLGP442	01LUSA	F	2	S	0.001
SGLGP443	08WAWE	F		R	0.999
SGLGP444	04WTAR	F	2	S	0.001
SGLGP445	01LUSA	F	2	S	0.001
SGLGP446	01LUSA	F	2	S	0.001
SGLGP447	04WTAR	F	0	S	0.018
SGLGP448	10ERDN	F	4	R	0.999
SGLGP449	04WTAR	F	2	R	0.999
SGLGP450	06KILM	F	0	R	0.999
SGLGP451	08WAWE	F	3	R	0.598
SGLGP452	03STAR	F	2	S	0.001
SGLGP453	04WTAR	F	2	S	0.001
SGLGP454	03STAR	F	2	S	0.001
SGLGP455	04WTAR	F	0	S	0.001
SGLGP456	08WAWE	F	0	R	0.502
SGLGP457	01LUSA	M	1	S	0.001
SGLGP458	01LUSA	F	2	S	0.001
SGLGP459	01LUSA	F	2	S	0.001
SGLGP460	01LUSA	F	2	S	0.001
SGLGP461	01LUSA	M	1	S	0.001
SGLGP462	06KILM	F	2	R	0.999
SGLGP463	01LUSA	M	1	S	0.001
SGLGP464	05KNAP	M	1	S	0.001
SGLGP465	01LUSA	F	2	S	0.001
SGLGP466	04WTAR	F	2	R	0.999
SGLGP467	05KNAP	M	1	R	0.015
SGLGP468	01LUSA	M	2	S	0.002
SGLGP469	08WAWE	M?	2	R	0.999
SGLGP470	08WAWE	F	2	R	0.999
SGLGP471	07ORMG	F	2	S	0.006
SGLGP472	05KNAP	F	2	R	0.948
SGLGP473	10ERDN	F	2	R	0.999
SGLGP474	08WAWE	M?	2	R	0.999
SGLGP475	08WAWE	F	8	R	0.999
SGLGP476	05KNAP	M	1	R	0.999

SGLGP477	08WAWE	F	4	R	0.96
SGLGP478	04WTAR	F	0	S	0.001
SGLGP479	04WTAR	F	2	R	0.999
SGLGP480	01LUSA	F	2	S	0.001
SGLGP481	01LUSA	M	2	S	0.001
SGLGP483	04WTAR	F	0	S	0.001
SGLGP484	04WTAR	F	0	S	0.001
SGLGP485	04WTAR	F	2	S	0.001
SGLGP486	03STAR	F	2	S	0.001
SGLGP487	06KILM	F	1	S	0.002
SGLGP488	08WAWE	F	2	R	0.992
SGLGP489	06KILM	F	2	R	0.999
SGLGP490	06KILM	F	0	R	0.999
SGLGP491	05KNAP	F	0	R	0.999
SGLGP492	06KILM	F	0	R	0.999
SGLGP493	01LUSA	F	2	S	0.001
SGLGP494	11SHRA	F	2	R	0.999
SGLGP495	07ORMG	F		S	0.002
SGLGP496	02CARA	F	2	S	0.001
SGLGP497	01LUSA	F	3	S	0.001
SGLGP498	02CARA	F	2	S	0.001
SGLGP499	03STAR	F		S	0.001
SGLGP500	01LUSA	F	2	S	0.001
SGLGP501	02CARA	F	2	S	0.001
SGLGP502	02CARA	F	2	S	0.001
SGLGP503	02CARA	F	2	S	0.001
SGLGP504	02CARA	F	2	S	0.001
SGLGP505	08WAWE	F	5	R	0.999
SGLGP506	05KNAP	F	2	R	0.999
SGLGP507	01LUSA	M	2	S	0.001
SGLGP508	05KNAP	F	0	R	0.999
SGLGP509	01LUSA	M	1	S	0.001
SGLGP510	05KNAP	F		S	0.001
SGLGP511	05KNAP	F	2	R	0.999
SGLGP512	06KILM	F	2	R	0.999
SGLGP513	05KNAP	F	2	S	0.001
SGLGP514	06KILM	F	2	S	0.002
SGLGP515	04WTAR	F	2	S	0.002
SGLGP516	05KNAP	F	0	R	0.107
SGLGP517	06KILM	F	2	S	0.001
SGLGP518	05KNAP	F	2	R	0.987
SGLGP519	06KILM	F	0	S	0.002
SGLGP520	06KILM	F	2	S	0.001
SGLGP521	06KILM	F	2	R	0.999
SGLGP522	06KILM	F	2	S	0.001
SGLGP523	05KNAP	F	2	S	0.001
SGLGP524	06KILM	M	1	S	0.001
SGLGP525	06KILM	F	1	S	0.001
SGLGP526	06KILM	F	0	S	0.001

SGLGP527	05KNAP	F	0	S	0.001
SGLGP528	08WAWE	M	0	R	0.858
SGLGP529	08WAWE	M	0	R	0.999
SGLGP530	08WAWE	F	3	R	0.967
SGLGP531	08WAWE	M	0	R	0.992
SGLGP532	08WAWE	F	3	R	0.999
SGLGP533	08WAWE	F	6	R	0.803
SGLGP534	08WAWE	F	1	R	0.831
SGLGP535	08WAWE	F	3	R	0.832
SGLGP536	08WAWE	F	4	R	0.829
SGLGP538	08WAWE	M	1	R	0.735
SGLGP540	08WAWE	M	3	R	0.999
SGLGP541	08WAWE	M	1	R	0.431
SGLGP542	08WAWE	F	3	R	0.355
SGLGP543	08WAWE	M	1	R	0.96
SGLGP544	08WAWE	F	2	R	0.31
SGLGP545	08WAWE	M	1	R	0.999
SGLGP546	08WAWE	M	1	R	0.775
SGLGP547	08WAWE	F	1	R	0.284
SGLGP548	08WAWE	M	0	R	0.897
SGLGP549	08WAWE	F	1	R	0.157
SGLGP551	08WAWE	F	4	R	0.999
SGLGP552	08WAWE	F	0	R	0.889
SGLGP553	08WAWE	M	0	R	0.999
SGLGP554	08WAWE	F	4	R	0.999
SGLOR009	21MORV	F	8	R	0.999
SGLOR012	14LOCH	F	3	R	0.999
SGLOR015	12SOBN	F	5	R	0.989
SGLOR020	20BALA	F	0	R	0.999
SGLOR021	16ORCH	F	3	R	0.999
SGLOR022	20BALA	F	3	R	0.999
SGLOR023	17NOBN	M	5	R	0.999
SGLOR024	17NOBN	M	4	R	0.999
SGLOR027	20BALA	F	3	R	0.999
SGLOR028	14LOCH	F	7	R	0.999
SGLOR029	21MORV	F	9	R	0.999
SGLOR033	16ORCH	F	2	R	0.999
SGLOR034	14LOCH	F	0	R	0.999
SGLOR035	16ORCH	F	3	R	0.999
SGLOR059	16ORCH	F	3	R	0.999
SGLOR066	14LOCH	M	3	R	0.999
SGLOR068	14LOCH	F	3	R	0.999
SGLOR075	16ORCH	M	7	R	0.999
SGLOR076	14LOCH	F	3	R	0.999
SGLOR081	20BALA	F	3	R	0.999
SGLOR086	20BALA	F	3	R	0.985
SGLOR093	20BALA	F	3	R	0.999
SGLOR096	14LOCH	F	3	R	0.922
SGLOR097	16ORCH	F	3	R	0.999

SGLOR099	14LOCH	F	0	R	0.999
SGLOR100	20BALA	M	2	R	0.999
SGLOR101	16ORCH	M	3	R	0.999
SGLOR102	14LOCH	F	3	R	0.999
SGLOR103	20BALA	F	2	R	0.999
SGLOR104	17NOBN	F	4	R	0.999
SGLOR105	16ORCH	F	0	R	0.999
SGLOR107	16ORCH	F	3	R	0.999
SGLOR108	12SOBN	M	6	R	0.989
SGLOR110	16ORCH	M	3	R	0.999
SGLOR112	14LOCH	M	3	R	0.999
SGLOR113	16ORCH	F	3	R	0.999
SGLOR114	16ORCH	M	3	R	0.999
SGLOR115	14LOCH	F	2	R	0.999
SGLOR117	14LOCH	M	3	R	0.999
SGLOR118	20BALA	F	3	R	0.999
SGLOR119	14LOCH	F	3	R	0.999
SGLOR120	16ORCH	M	3	R	0.999
SGLOR121	20BALA	F	0	R	0.999
SGLOR124	14LOCH	F	3	R	0.999
SGLOR126	20BALA	F	0	R	0.999
SGLOR128	14LOCH	F	3	R	0.999
SGLOR129	16ORCH	M	3	R	0.999
SGLOR131	14LOCH	F	3	R	0.999
SGLOR133	12SOBN	F	4	R	0.999
SGLOR136	17NOBN	M	6	R	0.999
SGLOR137	14LOCH	F	3	R	0.999
SGLOR138	16ORCH	F	0	R	0.999
SGLOR140	16ORCH	F	3	R	0.999
SGLOR141	14LOCH	F	3	R	0.999
SGLOR142	14LOCH	M	3	R	0.999
SGLOR143	16ORCH	M	5	R	0.999
SGLOR144	14LOCH	F	3	R	0.999
SGLOR146	14LOCH	F	7	R	0.999
SGLOR147	14LOCH	M	3	R	0.999
SGLOR148	14LOCH	F	3	R	0.999
SGLOR149	16ORCH	M	3	R	0.999
SGLOR150	14LOCH	F	3	R	0.999
SGLOR151	14LOCH	F	3	R	0.999
SGLOR153	14LOCH	F	3	R	0.999
SGLOR154	14LOCH	F	3	R	0.987
SGLOR155	20BALA	M	2	R	0.999
SGLOR158	14LOCH	F	3	R	0.999
SGLOR160	16ORCH	M	3	R	0.999
SGLOR161	20BALA	M	2	R	0.999
SGLOR166	16ORCH	F	3	R	0.999
SGLOR170	14LOCH	M	3	R	0.999
SGLOR171	14LOCH	F	3	R	0.999
SGLOR177	12SOBN	M	2	R	0.999

SGLOR181	17NOBN	M	6	R	0.999
SGLOR182	14LOCH	F	3	R	0.999
SGLOR183	16ORCH	F	3	R	0.999
SGLOR193	20BALA	F	0	R	0.999
SGLOR216	20BALA	F	3	R	0.999
SGLOR219	16ORCH	F	3	R	0.999
SGLOR223	14LOCH	F	3	R	0.999
SGLOR224	14LOCH	F	3	R	0.999
SGLOR225	14LOCH	F	3	R	0.999
SGLOR226	16ORCH	M	3	R	0.999
SGLOR227	20BALA	F	0	R	0.999
SGLOR229	16ORCH	M	5	R	0.999
SGLOR232	17NOBN	F	5	R	0.999
SGLOR233	17NOBN	F	5	R	0.999
SGLOR234	14LOCH	M	7	R	0.999
SGLOR241	14LOCH	F	3	R	0.999
SGLOR242	14LOCH	F	3	R	0.999

***HS (Helen Senn) prefix denotes samples taken in 2006/7, SG (Simon Goodman) prefix denotes samples taken in 1996/7 and KA (Kate Abernethy) prefix denotes samples taken in 1991/2.**