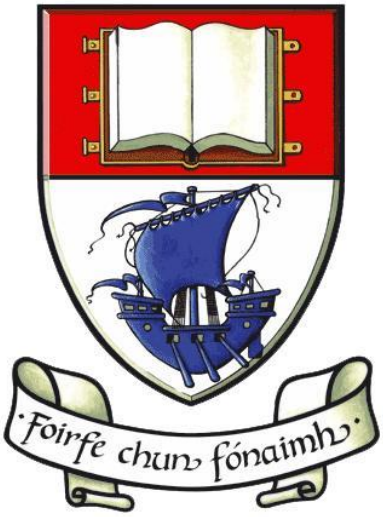

The DNA Toolbox:

Development and optimisation of molecular techniques to identify red squirrels (*Sciurus vulgaris*)

A thesis submitted to

Waterford Institute of Technology

by

Denise Bridget O'Meara

2014

In partial fulfilment of the requirements for the Degree of Doctor of Philosophy

Under the supervision of Dr. Catherine O'Reilly and Dr. Peter Turner

Department of Chemical & Life Sciences

Declaration

No element of the work described in this thesis has been previously submitted for a degree at any institute. The work described in this thesis has been performed by the author, and contributions from others are clearly outlined as collaborations or referenced in the text.

_____

Denise B. O'Meara

Abstract

The red squirrel (*Sciurus vulgaris*) has seen a population decline in Ireland and Britain over the last 100 years since the introduction of the North American grey squirrel (*Sciurus carolineneis*). The decline has been attributed to a number of factors including habitat loss and direct competition with the invasive grey squirrel. As a result, there is a growing need to monitor the distribution of both species for conservation management. Squirrels can be surveyed non-invasively using hair-tubes, and the species can be identified by microscopic identification of individual hairs. To increase the information obtained from non-invasive surveys a suite of DNA tools to identify both squirrel species from non-invasive DNA samples were developed in this study. This involved the design of real-time PCR assays designed to amplify short DNA fragments of red and grey squirrel DNA. A molecular dietary study of pine marten (*Martes martes*) scats using the species-specific assays to detect squirrel DNA in the pine marten diet was also conducted. Methods to determine the gender of squirrels were investigated by testing the utility of a range of Y-chromosome markers. A panel of previously published microsatellite DNA markers were screened and optimised to identify individual red squirrels. The techniques were applied to a hair-tube study in County Waterford, South East Ireland, where individual red squirrels were identified. Additional samples were collected from subpopulations throughout Ireland, and a remnant red squirrel population in mid Wales where the genetic variability of the populations were assessed using microsatellite genotyping and mitochondrial DNA sequencing. The combination of these tools can be used for non-invasive monitoring and translocation projects of the red squirrel.

Acknowledgments

I am very grateful to Catherine O'Reilly and Peter Turner for giving me the opportunity to pursue a PhD. This project was originally funded as an MSc project (Strand 1 – National Development Fund), and I've been very fortunate to have been supported both financially and academically throughout my time at WIT. I've also been fortunate to have had the opportunity to work on the Mammals in a Sustainable Environment (MISE) project at WIT, and to continue researching red squirrels while also committing my time to that project. This thesis would not have been possible for me to finish without having the luxury of time, finance and the experience that my role at the MISE project has provided me with. The support I received from both Catherine and Pete has often gone well beyond what I believe is expected of an academic supervisor.

The work in this project has benefitted from collaborations with Emma Sheehy and Colin Lawton (National University of Ireland, Galway), Declan O'Mahony (Ecological Management Group), Andrew Harrington (MISE, Waterford County Council), Jenny MacPherson (MISE, The Vincent Wildlife Trust) and Huw Denman (SelectFor).

I would like to thank the following people for access to their property and for the collection of samples including: Seán and Síle Murphy and Rosemary Higby for providing access to their property for hair-tube surveys. I would also like to thank Michael Desmond and Sarah Rubalcava for conducting hair tube surveys in Counties Waterford and Dublin in their own time. The following people provided red and grey squirrel samples that were used throughout this study included: Andrew Harrington, Ann FitzPatrick, Catherine Waters, Colin Lawton, Daniel Buckley, Daniel McAfee, David O'Neill, David Tosh, Emma Sheehy, Eugene Burke, Finton Malone, Ger O'Brien, Huw Denman, Jacinta Mullins, Jamie Gorrell, Jenny MacPherson, Juliet Robinson, Mark Ferryman, Paddy Sleeman, Padraig Whelan, Richard Wales, Ruth Carden, Siobhan Moran, Will Carr, anonymous senders and senders of incidental road kill samples.

I am indebted to the help and support I received over the years in the WIT lab from Lee Coffey, David O'Neill, Siobhan Moran and Jacinta Mullins. The technicians, Karen, Walter and Aidan have also been helpful and kind. I have continued to have a great friendship with Jacinta after she left WIT in 2010, and I have relied on her help and expertise throughout the project. David has been especially good in terms of his lab support and discussing improvements to molecular approaches, but also for his friendship over the years. I've appreciated the banter and friendship from Ciara, John, Edel and Emma in BO1, and I have also enjoyed the lunch time conversations with Maeve, Gemma, Laura and Erica.

Many people (other than Catherine and Pete) were burdened with the task of reading, advising, and providing comments on various chapters of this thesis including Allan McDevitt, Jacinta Mullins, David O'Neill and Andrew Harrington. I really appreciate all the advice, help, discussions, amendments, support, and the time you gave to do so.

I would like to thank Mam, Dad and Julie for all their support and encouragement. I had a shaky start to life with childhood cancer from 1988-1990, but as a result I was always made feel I could achieve anything since I had beaten cancer. That belief has worked well in my life, and to this day I still think I can achieve almost anything. Mam and Dad always said “go away and study whatever subjects you want, as long as you are willing to work at them”, and they always maintained that “education is no load to bear”. I’ve always studied what I enjoyed doing and I think that is some of the best advice you could ever give anyone. I also have my parents to thank for fostering my love of nature through farming.

Finally, poor Andrew has been the long suffering partner in all of this. He has silently taken over my share of the domestic duties, has cooked countless meals, walked the dogs and fed the cats, while I was working late. I feel very lucky to be in such a loving relationship with someone who understands the time that research requires, and who shares my passion for nature.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Squirrels of the world

Squirrels form the Sciuridae family and are part of the order Rodentia, all of which are believed to have ascended from a common ancestor that lived around 35 million years ago (Thorington et al. 2012). The earliest squirrel fossil was found in North America and was dated to 36 million years ago. This squirrel, *Douglasssciurus jeffersoni* might be the earliest known ancestor for the Sciuridae (Thorington et al. 2012). The family Sciuridae comprises 273 species and 50 genera divided in two subfamilies: Sciurinae, the tree and ground squirrels, and Pteromyinae, the flying squirrels (Steppan et al. 2004).

The Sciuridae are found throughout the world, with absences in southern South America, Madagascar and Australasia (Steppan et al. 2004). Steppan et al. (2004) examined the phylogeny of the Sciuridae and found that flying squirrels are most closely related to holarctic tree squirrels, which consists of North American and Northern Eurasian tree squirrels (Fig 1.1). African tree squirrels formed part of a subfamily from Southern Asian tree squirrels (Fig 1.1).

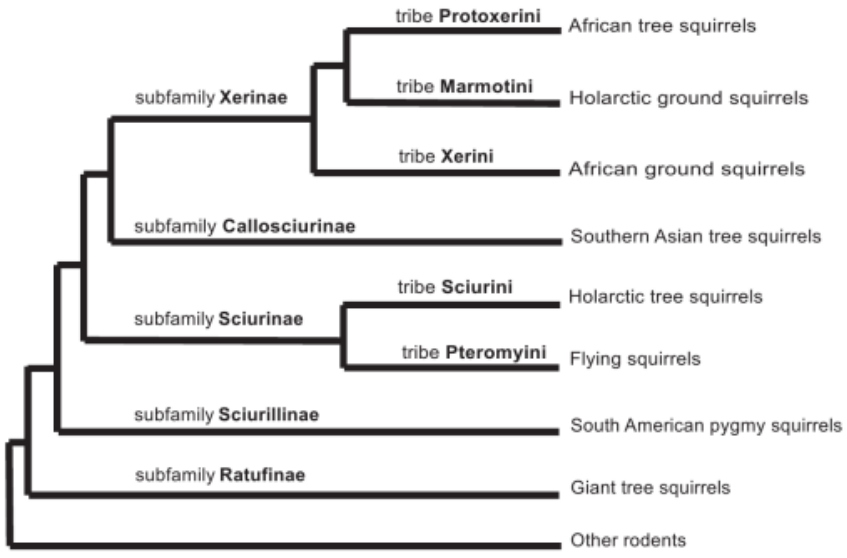

Fig 1.1: Phylogeny of the Sciuridae from Thorington et al. (2012) and based on research by Steppan et al. (2004).

The role of squirrels in ecosystem functioning is well documented in terms of seed and fungal spore dispersal, and planting and pollination of seeds in a variety of habitats (Thorington et al. 2012). Squirrels are an important prey item for a range of predators, and have many uses by people. For instance, grey squirrels (*Sciurus carolinensis*) are an important game species in their home range in North America, and *Callosciurus* are an important food source in Asia. The pelts of squirrels have been and continue to be used by people, especially in Eurasia (Thorington et al. 2012). In some countries, the conservation, and the encouragement of recreational observers has led to the establishment of squirrel associated industries including eco-tourism activities, a demand for nest boxes and feeders, and an illegal pet trade (Thorington et al. 2012). Squirrels are also associated with negative economic impacts, such as bark stripping, predation of crops, damage to electrical wires through gnawing, burrowing into human structures, and they are important vectors of diseases, including human diseases such as the plague (Thorington et al. 2012).

When considering the ecological importance of the Sciuridae at a global level, it is important to understand the overall conservation status of the Sciuridae. The largest threats to the Sciuridae are habitat loss and fragmentation, overexploitation and hunting pressures, the effects of climate change, and invasive species (Thorington et al. 2012). Squirrels rely on mature forests that produce a large quantity of seed, and have adequate areas for seed storage and nest building (Koprowski 2005). Clearcutting directly removes and degrades habitat which decreases squirrel activity, and potentially increases their vulnerability to risks of predation and extinction. Thinning also reduces squirrel density, and red squirrels were found to avoid areas that were open, or that contained increased shrub cover (Gurnell et al. 2002; Koprowski 2005). Thorington et al. (2012) report that worldwide deforestation rates are at 0.33% annually and grasslands are being annually reduced by 0.08%, resulting in a combined habitat loss of 4% per decade. The results of habitat fragmentation can change genetic variation and alter population densities and affect the overall biodiversity of the species. As a result, about one fifth of squirrel species across the world are considered at risk or are so poorly studied that there is a lack of knowledge regarding their conservation status (Fig 1.2). Only 35% of all species are considered

to have stable populations, with the remainder decreasing or are of unknown status (Thorington et al. 2012).

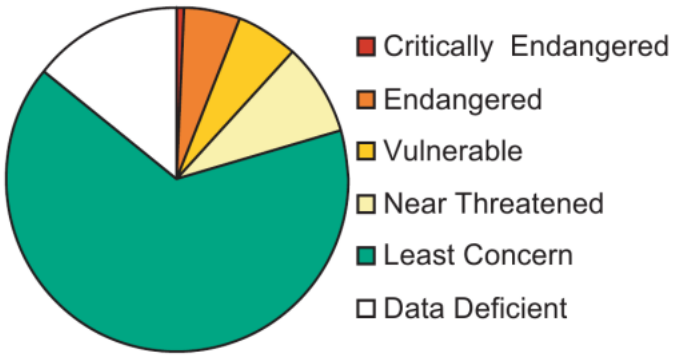

Fig 1.2: IUCN (International Union for the Conservation of Nature) conservation assessment of the Sciuridae (Thorington et al. 2012).

1.2 Holarctic tree squirrels

Within the Sciuridae there are three genera; *Sciurus* (n = 28), *Tamiasciurus* (n = 3), and *Glaucomys* (n =2) species (Koprowski 2005). An example of the high diversity of Holaractic tree squirrels can be seen in the Madrean Archpeligo in North America, an area between Arizona and New Mexico, encompassing an area of basin and range topography that includes a diverse range of island habitats that are isolated due to lower desert elevations. As a result of isolation, the area has one of the highest numbers of squirrel species in the world, with five *Sciurus* and *Tamiasciurus* species and an additional two introduced *Sciurus* species. There are also 16 subspecies that evolved due to isolation in the region. The future of all endemic species in the region is under threat (Koprowski 2005).

1.3 The red squirrel (*Sciurus vulgaris*)

The Eurasian red squirrel (*Sciurus vulgaris*) (Fig 1.3) is distributed in forested areas across the Palaearctic from Ireland, Britain, Spain and Portugal to Eastern Russia, Japan, Mongolia and China (O' Teangana et al. 2000; Thorington et al. 2012) (IUCN Red list 2008) (Fig 1.4). There are at least 22 documented subspecies of *Sciurus*

vulgaris, but according to Thorington et al. (2012) there is a need for broader studies across the species massive range.

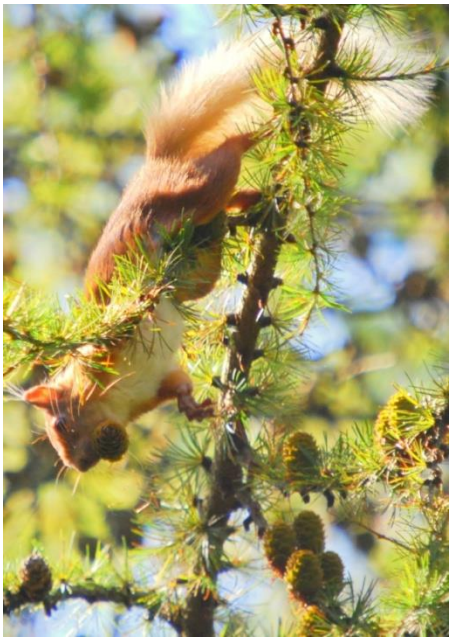

Fig 1.3: Red squirrel (*Sciurus vulgaris*) photographed in the Comeragh Mountains Co. Waterford (Photographed by Maurice Flynn).

The habitat requirements of the species vary throughout the range, but it is mostly found to be abundant in large areas of coniferous woodland, although it is also found in deciduous and mixed woodlands. The species can occur in subalpine forests and can even occur in urban environments (Bertolino 2008; Thorington et al. 2012). Red squirrels live in nests called dreys that tend to be spherical and positioned close to the tree trunk or situated between a forked branch. Squirrels may alternate between different dreys for day rests and night time sleeping (Gurnell et al. 2008).

According to the IUCN, the red squirrel is listed as being of 'least concern', however, the overall trend of the population is in a state of decrease, and if the red squirrel in the British Isles was to be considered by its sub-species title (*S.v. leucurus*), then it would be considered critically endangered. The species is also considered to be near threatened in Mongolia and Croatia, and is of vulnerable status in Poland (Thorington et al. 2012). It is the only native tree squirrel in Europe (Bertolino 2008), and has the largest habitat range of all squirrels in the world (Thorington et al. 2012). Its colour variations include dark red, black, brown, grey and bluish, and the underneath varies

from white to cream, with the tail ranging from the same colour as the dorsum to paler colours in summer (Thorington et al. 2012). The ear tufts also vary with season, being more pronounced in winter than summer.

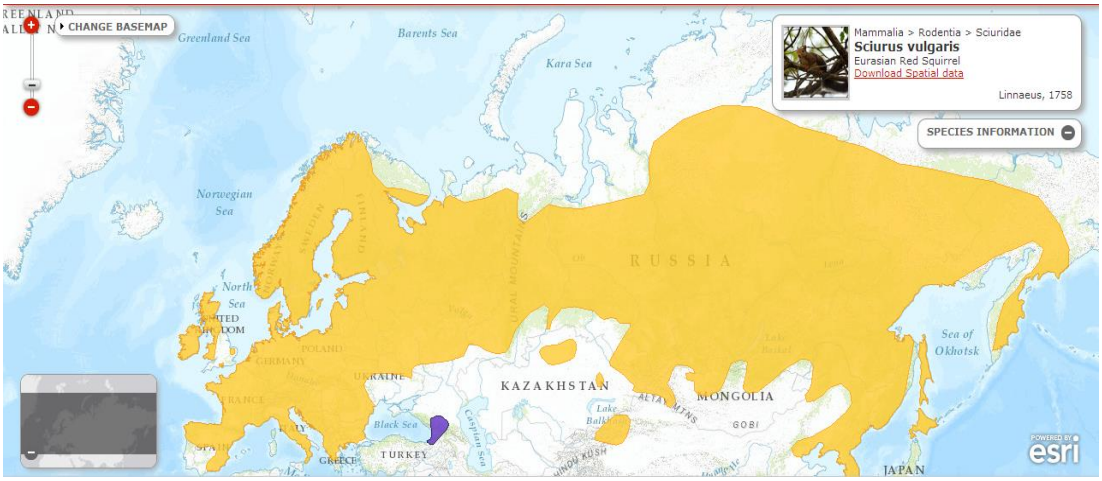

Fig 1.4: Current distribution of red squirrels in their extant (yellow) and introduced (purple) habitats across Eurasia (Accessed from the IUCN 2008 data on 23rd of August 2013).

1.3.1 The red squirrel in Western Europe

The earliest radiocarbon dated bones of red squirrel (after the last glaciation which eneded around 12,000 years ago) are from a site on the Atlantic coast of Norway. This was dated as being between 11,500 and 10,000 years old (Stewart and Lister 2001). Stewart and Lister (2001) believe this is evidence that the pine line was this far north, or possibly that there was an oceanic northern refugium. Remains of the red squirrel were also found in Hungary and were estimated to be from the middle Pleistocene, and later records from France place fossils as late Pleistocene (more exact dates were not made available in the study). A radiocarbon dated fossil from Germany was dated to be about 8,000 years old (Lurz et al. 2005).

1.3.2 The red squirrel in the British Isles

The red squirrel in the British Isles (Britain and Ireland) represents the most westerly point of its Eurasian distribution. It is not known precisely when the red squirrel arrived in Britain, but the oldest bones found in Britain are from the Isle of Wight, and they were radiocarbon dated and estimated to be about 4,500 years old (Lurz et al. 2005). The fossil record is poorly defined for the red squirrel in mainland Britain, and where it does appear in caves, it is associated with forest floras. Shorten (1954) wrote that the bone fragments of a squirrel species, possibly an ancient relative of the modern red squirrel, and ossilized chewed fir cones (possibly chewed by squirrels) were found in "bone caves" (sites that contain relics of ancient mammals) in Norfol,k and were dated to a period in the late Pliocene. Shorten also documents a premolar that was found in similar cave deposits from that time which were identified as *Sciurus whitei*. This squirrel species is not well defined according to Reumer and van den Hoek Ostende (2003).

To help explain the poor fossil record for squirrels, Reumer and van den Hoek Ostende (2003) suggested that as squirrels are good tree climbers they are a difficult species for predators to catch and subsequently, this may explain why they do not regularly occur in bone caves. In fact, squirrels in general are poorly represented in the fossil record across the world, despite their wide distribution (Thorington et al. 2012).

The red squirrel and people

The red squirrel does not feature extensively in Irish folklore. The wolf (*Canis lupis*) and wild pig (*Sus scrofa*) feature extensively in Irish place names (Hickey 2011), while the red squirrel rarely does. One exception is the place Carrickreagh, or Carraig na hIorua, in Donegal, that means "rock of the squirrels". One of the earliest mentions of the red squirrel in Ireland was in Brehon Laws, where it was classed as a cat, which possibly means that they were kept as pets. Pine marten (*Martes martes*) were also classed in the same category, and a famous example of a pet marten was the one kept by Queen Maeve of Ulster in Irish Mythology. Squirrels were highly regarded for

their fur, known as 'vair', and its use was restricted to nobility during the Middle Ages (Mac Coitir 2010).

The fur of red squirrels from the Russian dominions was favoured over European varieties due to the bluish/grey colour. It was said that 15,000,000 were captured annually in Russia, and 3,000,000 of those were destined for the British market. The white underbelly was used to line coats. The tails were made into boas for foreign markets and the tails were also made into artists' pencils (Royal Commission, London 1851). Squirrel hats became popular in central Europe in the 1550s and were favoured by the less wealthy as the red squirrel fur from Western Europe was cheaper than prized beaver fur which was worn by the wealthy at this time. The demand for squirrel fur declined at end of the 14th Century as it became quite common, and then it began to be replaced for a demand for mustelid fur (Hoffman 2010). Around this time fur trading markets opened up on an international scale, as many of the European animals became depleted during Medieval times.

A recent history of the red squirrel in Ireland

Le Fanu (1922) cites the "rates outward" book that documented thousands of red squirrel skin exports in the 16th Century. These exports then ceased and it is presumed that a lack of habitat and over exploitation caused the proposed red squirrel extinction in the 17th Century. It is not fully understood if red squirrels became extinct at this point in Irish history, or if they became so rare that the export business became no longer viable. The cessation of squirrel tail exports also coincided with a depression in the export market (Mac Coitir 2010). Shorten (1954) added a further caveat to the Irish export record however that there might be a danger in assuming that all the skins were in fact native to Ireland. Shorten provided an example of red squirrel skins that were listed as an export item from England to Scotland, but the skins had actually originated in Scandinavia and Russia. It is now generally accepted that the red squirrel was present in Ireland prior to its reintroduction in the 1800s, but it is not known how long red squirrels had been present in Ireland prior to that, again confounded by a lack of archaeological evidence.

The first documented reintroduction of the red squirrel in Ireland refers to dates between 1815 and 1825 in Glenmore Estate, Ashford, Co. Wicklow. Other introductions are also referred to by Barrington 1880 (Table 1.1).

Table 1.1 Documented introductions of red squirrels in Ireland from Barrington (1880).

Location	Person responsible for release	Number and Source	Year
Glenmore Estate, Ashford, Co. Wicklow	J. Synge	A single pair, bred in the house and released	1815-1825
Castle Howard Estate, Co. Wicklow	Alice Howard, Countess of Wicklow	Unknown	1800s
Lucan, Co. Dublin	J. Shackleton	10 squirrels, purchased from a Dublin dealer	1876
Brownes' Hill, Co. Carlow	Colonel Bruen	Multiple squirrels	1850
Birr Castle, Co. Offaly	Lord Roose	Multiple squirrels from Sussex and Yorkshire	1864
Castle Taylor, Garbally, Co. Galway	Lord and Lady Glancarthy	Two - four pairs from London	1833
Roscommon	Stable boy	Translocated the Garbally squirrels	1874
Castleforbes, Co. Longford	unknown	Unknown	Common in 1836
Multyfarnham, Co. Westmeath	Nugent	Unknown	1800s
Ballyarr, Ramelton, Co. Donegal	Lord Hill	Kept as pets, possibly released/escaped	Present in 1876
Moneyglass, Toome, Co. Antrim	Jones	Multiple squirrels from England	Around 1874
Ravendale Park, Newry, Co. Down.	Lord Clermont	Two separate releases of 10 and 15 squirrels	1851

In 1880, most of the south of Ireland did not contain any red squirrels. There were no records of red squirrels in Counties Cork, Waterford and Kerry (Fig 1.5 [a]). Interestingly, most of the records of introductions occurred in Anglo-Irish estates, with many of the owners serving in the British parliament which provided relatively easy access to the British squirrel population. Barrington wrote in 1915 that red squirrels had arrived in Waterford and were so common in Portlaw that they were described as a pest. By 1910, red squirrels had been documented in every county in Ireland (Shorten 1954) (Fig 1.5 [b]).

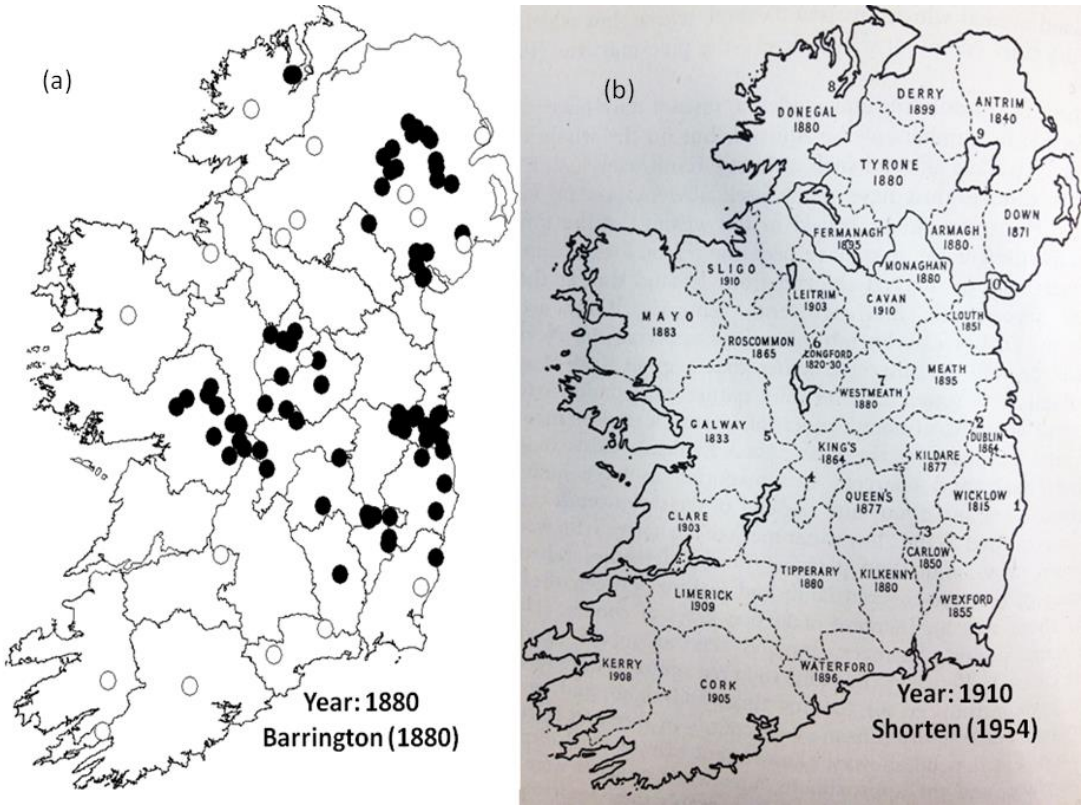

Fig 1.5: (a) Locations of red squirrel records from Barrington (1880). *Black circles* represent documented presence and *white circles* documented absence. Map created by A Harrington. (b) Every county in Ireland had records of red squirrels by 1910. The first year documented per county is shown (Shorten 1954).

A recent history of the red squirrel in Britain

The red squirrel in Scotland is also thought to have become extinct during the 18th Century (Shorten 1954). Although, Harvie-Brown (1881) thought that there may have been a possibility that red squirrels survived in the Scottish highlands. Some of the

reasons suggested for the decline relate to the destruction of woodland habitat for building and to fuel iron smelting furnaces. Reintroductions into Scotland occurred from 1772 to 1872. Most of the reintroductions were reported to have come from England or Wales, and some from squirrels from Norway or Sweden were introduced into Perthshire (Shorten 1954).

Shorten (1954) found it difficult to find direct evidence for the decline of the red squirrel in Wales and England in the 18th Century. There was an increase in industry at the time and woodlands were neglected during this period. There was a documented scarcity of wood and a high industrial demand for wood meant that it had to be imported to support the industry demand. As a result, Shorten believes that the red squirrels in England and Wales must have suffered similar declines to Ireland and Scotland during this time.

Conifer species were introduced during the end of the 18th Century and between 1810 - 1823 over 50,000 acres were planted. By 1890, red squirrels were beginning to become abundant throughout the British Isles. Shorten (1954) cites a note from London dated to 1837, that reports that 20,000 red squirrels were sold at London markets annually. Some of these red squirrels were reported to have been imported from France, while many more were brought in by labourers from the local areas in the south of England. Shorten cautioned however that this could be an exaggeration. It is clear however, that the demand for red squirrels in Britain at this time was high and that imports from Continental Europe were occurring to supplement the market demand. It is also probable that red squirrels were being bred in captivity to support the market. This is also a likely origin for many of the Irish red squirrels, both through dealers, and those travelling between London and Ireland.

1.4 The North American grey squirrel (*Sciurus carolinensis*)

The grey squirrel (*Sciurus carolinensis*) (Fig 1.6) is native to Eastern North America, and is found from Saskatchewan in Canada to Texas (Fig 1.7). It was introduced into Britain (Scotland, Wales and England), Ireland, Italy, South Africa, Australia, Mexico and various locations in Western Canada and North America including California,

Montana, Oregon and Washington. Similar introductions took place in Quebec, New Brunswick, British Columbia, Manitoba, Nova Scotia and Ontario (Thorington et al. 2012). Introductions to Mexico, Australia, and Hawaii failed to establish (Thorington et al. 2012).

Fig 1.6: Grey squirrel (*Sciurus carolinensis*) (photographed by Brian Power).

Grey squirrels are relatively large and weigh 500 to 710 g (Thorrington et al. 2012). The squirrel has a grizzled pale to slate grey colour that sometimes exhibits a reddish or cinnamon colour (Thorington et al. 2012). Melanism tends to be more common in northern populations of grey squirrels, in their natural distribution, but black squirrels have also been seen in Britain. Albino squirrels are also found in Arizona and Midwestern USA (Thorington et al. 2012).

The grey squirrel can be found in hardwood, mixed forests and urban areas with few trees (Thorington et al. 2012). The diet of the grey squirrel in North America consists of seeds, fruit, nuts, insects, grains, fungi, bones and cannibalism (Bertolino 2008). Thorington et al. (2012) also report that grey squirrels will feed on nestlings and eggs. No cases of bark damage by grey squirrels have been reported in the United States despite huge economic losses in Europe as a result of bark stripping (Huxley 2003; Mayle and Broome 2013).

The grey squirrel has many more natural predators in its native range than it does in Britain and Ireland. In the USA, martens, both fishers (*Martes pennanti*) and the American marten (*Martes americana*) will predate on the grey squirrel (Hales et al. 2004). The number of predators in Britain and Ireland are markedly reduced and may be a factor that facilitates the invasion process. The only potential predator of grey squirrels in Ireland is the pine marten (*Martes martes*) and possibly foxes (*Vulpes vulpes*) and cats (*Felis silvestris catus*) (Carey et al. 2007).

Grey squirrels are a vector for the squirrel poxvirus which causes a lethal disease in red squirrels (McInnes et al. 2013). The grey squirrel population has developed immunity to the disease (Huxley 2003). The first occurrence of the virus in Ireland was documented in 2012 (McInnes et al. 2013). However, it is unknown if the virus was previously present in Ireland, but had gone undetected.

Fig 1.7: Native range of the grey squirrel (Sciurus carolinensis) in North America and Canada (Accessed from the IUCN 2008 data on 23rd of August 2013).

1.4.1 The grey squirrels in the British Isles

Shorten (1954) wrote that "the pastime of importing, keeping, releasing, and distributing the American squirrel was in vogue between 1876 and 1929". Many documented introductions of the grey squirrel took place throughout Britain in the 19th Century (Reynolds 1985). Shorten (1954) detailed a number of the most important introductions in Britain (Table 1.2). The importation of grey squirrels was

finally banned in Britain in 1938 (Huxley 2003). Despite the British ban, imports to Italy also occurred in Piedmont, Stupinigi, Genoa, and Trecate (although some were removed) (Huxley 2003), but an increased number of populations have been observed in recent years (Martinoli et al. 2010).

Table 1.2: Documented introductions and releases of the grey squirrel throughout Britain and Ireland (captured from Shorten 1954).

NO.	LOCALITY	DATE	SOURCE	NO.	RESULT
1	Henbury, Cheshire	1876	U.S.A.	4	Increased
2	Bushey, Middlesex	1889	U.S.A.	5	Died out
3	Woburn, Bedfordshire	1890	U.S.A.	10	Increased
4	Loch Long, Scotland	1892	Canada	3	Increased
5	Benenden, Kent	..	..	..	Increased
6	Nuneham, Oxfordshire	..	Woburn	..	Increased
7	Richmond, Surrey	1902	U.S.A.	100	Increased
8	Wrexham, Denbighshire	1903	Woburn	5	Increased
9	Lyme, Cheshire	1903–4	—	25	Increased
10	Regent's Park, London	1905–7	Woburn	91	Increased
11	Malton, Yorkshire	1906	Woburn	36	Increased
12	Cliveden, Buckinghamshire	..	..	..	Increased.
13	Kew Gardens, London	1908	Woburn	4	Increased
14	Farnham Royal Buckinghamshire	1908	U.S.A.	..	Increased
14a	Farnham Royal, Buckinghamshire	1909	U.S.A.	5	Increased
15	Frimley, Surrey	1910	U.S.A.	8	Increased
16	Dunham, Cheshire	1910	..	2	Increased
17	Sandling, Kent	1910	..	..	Increased
18	Bramhall, Cheshire	1911–12	Woburn	5	Uncertain
19	Birmingham	1912	..	..	Increased
20	Castle Forbes, Ireland	1913	Woburn	8	Increased
21	Bedale, Yorkshire	1913	..	..	Increased
22	Bingley, Yorkshire	1914	London	14	Slight increase
23	Darlington	1914–15	..	..	Increased
24	Exeter	1915	..	4	Increased
25	Stanwick, Northamptonshire	1918	..	2	Increased
26	Dunfermline, Scotland	1919	..	..	Increased
27	Bournemouth, Hampshire	..	London	6	Increased
28	Hebden Bridge, Yorkshire	1921	..	8	Slight increase
29	Edinburgh	..	Zoo	..	Occasional
30	Aberdare, Glamorgan	1922	London	..	Slight increase
31	Needwood Forest, Staffordshire	1929	B'mouth	2	Uncertain

Also said to have

1.5 The decline of the red squirrel in association with the spread of the grey squirrel

Red squirrels reached record abundance in Britain in the late 1800s, and reports of damage in plantations were beginning to be documented throughout Britain (Harvie-Brown 1881; Shorten 1954). Red squirrels were subsequently controlled to reduce their numbers (Shorten 1954). This period was followed by a population crash between 1900 – 1920 and it occurred throughout the British Isles, including Jersey where the grey squirrel was never introduced. Despite the decline being associated with the introduction of the grey squirrel, Shorten (1954) reported that some of the dead red squirrels from this period showed signs of coccidiosis (a parasitic disease of the intestinal tract), but other diseases may have also been prevalent at the time. By the 1930s red squirrels began to recover again, but a decline during the Second World War also occurred due to a high demand for timber. Fluctuations in red squirrel populations have been reported throughout their distribution and records from countries where skins were exported such as Russia and Finland exhibited ten year cycles (Shorten 1954). Barrington (1915) also noticed the red squirrel decline in Ireland, and areas where the red squirrel had been plentiful in 1880 now had fewer squirrels. Barrington (1915) pointed to reasons including habitat loss, poor weather, disease, inter-breeding, predation, and a lack of food. Middleton (1929) also reported fluctuations in the Irish red squirrel population, where numbers might be plentiful one year, and very scarce the next. These trends were noticed in Wexford, Wicklow and Fermanagh. Middleton believed these trends were similar to the British situation and predicted that a disease might be the cause.

The Irish squirrel survey in 2007 showed that the grey squirrel was present across most of the eastern half of Ireland, and the red squirrel strongholds were reserved to the south and west of Ireland (Fig 1.8) (Carey et al. 2007). More recent work is now showing a general increase in red squirrels and a decrease in grey squirrels at a regional scale. The grey squirrel is now absent or rare in Counties Laois, Offaly, Longford, Westmeath and North Tipperary. The red squirrel has successfully recolonised Laois, Offaly and North Tipperrary. However, the grey squirrel is continuing to move south and south west in the Munster (South Ireland) region (Sheehy 2013). Presently, red squirrels are distributed in parts of Wales, Northern

England and Scotland, while grey squirrels are distributed across most of Britain (Fig 1.9).

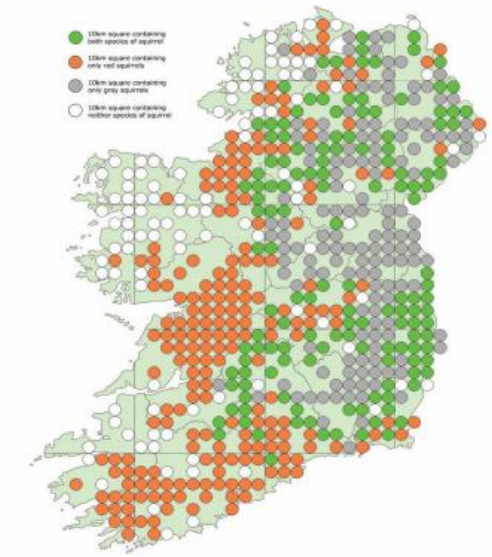

Fig 1.8: The most recent red and grey squirrel survey of Ireland from Carey et al. (2007). *Red circles* red squirrel only, *grey circles* grey squirrel only, *green circles* both species present and *white circles* both species absent.

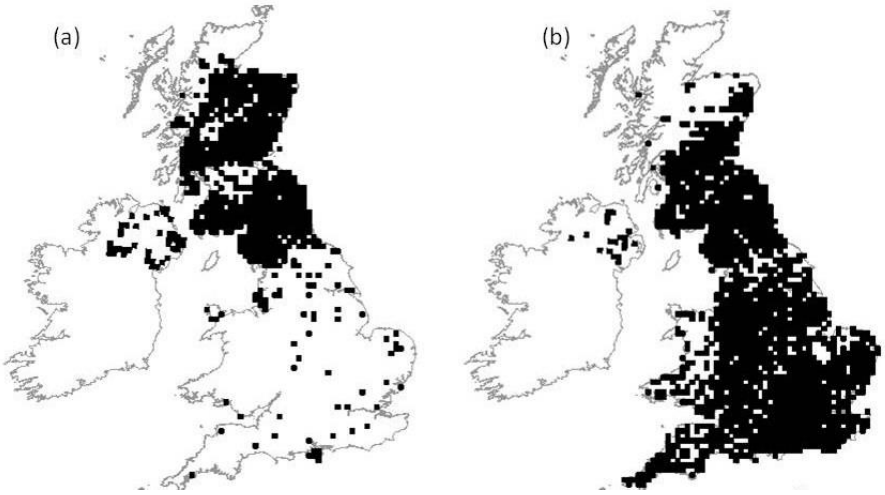

Fig 1.9: Red (a) and grey (b) squirrel distribution in Britain and Northern Ireland. 10 km records from 2007-2013, NBN Gateway, accessed August 25th 2013.

1.5.1 The pine marten

The pine marten (*Martes martes*) (Fig 1.10) is a natural predator of red squirrels in Ireland, and the species occurs in the diet at low frequencies (Lynch and McCann 2007). The pine marten is absent in England (but well distributed in Scotland), and is only recently recovering in Ireland (Fig 1.11) (O'Mahony et al. 2012; Sheehy et al. 2014). Game keepers in Ireland and Scotland have anecdotally claimed that the pine marten may be playing a role in the biological control of the grey squirrel, by preferentially preying upon the larger, less arboreal squirrel (Carey et al 2007; Caryl 2008). The claim that pine marten may be influencing the presence or absence of squirrels is difficult to prove as previous analysis of pine marten scats in Britain and Ireland revealed that the red squirrel formed a low percentage in the diet (Halliwell 1997; Lynch and McCann 2007; Caryl et al. 2012). A method that could rapidly screen a larger sample size for species occurring at low frequency, such as squirrels, could help address this question.

Fig 1.10: Pine marten (*Martes martes*) photographed by Maurice Flynn in Kilmeadan, Co. Waterford.

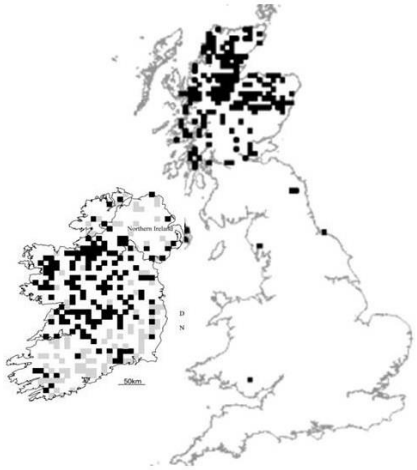

Fig 1.11: Pine marten distribution in Britain and Ireland. *Black squares* represent positive sites in both countries. British data is from NBN gateway for records from 2007 – present, and the Irish map is from O'Mahony et al. (2012). *Grey squares* (in Ireland) represent negative sites.

1.6 Protecting the remaining red squirrel populations

Red squirrels are fully protected under the Wildlife and Countryside Act 1981 in the UK, under Schedule 5 of the Wildlife (Northern Ireland) Order 1985, and it is protected under the 5th Schedule of the Wildlife Act 1976 and the Wildlife (Amendment) Act 2000 in the Republic of Ireland. The species is also on schedule III of the Berne Convention (ACTION: NPWS, EHS).

A number of initiatives have been put in place in Britain to help protect the red squirrel including the creation of red squirrel conservation groups that largely rely on volunteers to help monitor and carry out practical conservation work. Reintroductions and translocations have taken place in both Britain and Ireland. Red squirrels have also been translocated to sites in the west of Ireland, Derryclare, Co. Galway where grey squirrels are not found (Poole and Lawton 2009). Red squirrels were also translocated to Anglesey, an island of the north Welsh coastline, to boost the genetic diversity of a remnant population (Ogden et al. 2005).

Grey squirrels can be controlled and killed by cage trapping in Britain, which is now in place following legislation laid down by the Wildlife and Countryside Act 1981. There have been a number of localised extermination efforts, but these are largely unsuccessful in the long-term without a nationwide strategy (Mayle and Broom 2013). A well known exception is the Isle of Anglesey, where grey squirrels were successfully eradicated, and red squirrels were restored to the island (Ogden et al. 2005). Islands, or areas connected by bridges, may be more successfully managed as the sea acts as useful buffer for the zone. The public often regard the grey squirrel as an attractive species which makes it difficult to gain public support for their removal. Some efforts have also been made to remove grey squirrels in Ireland, and recommendations were made but were never implemented nationally (Carey 2008).

1.6.1 Methods for surveying squirrels

The monitoring of red squirrels is quite difficult. The most viable method to gain population specific information involves trapping, a labour expensive technique that requires high levels of training and licensing (Poole and Lawton 2009; Waters 2012). Other less invasive techniques include indirect monitoring strategies (Gurnell et al. 2009). Some of these techniques include looking for feeding signs, such as stripped pine cones, counting squirrel dreys and hair-tube surveys (Fig 1.12) (Gurnell et al. 2009). Gurrnell et al. (2004) found that most of these techniques were good for establishing the presence of squirrels (with the exception of drey counts), but that these indirect methods were poor to moderate for estimating abundance.

Fig 1.12: How to find signs of squirrel activity (a) squirrel drey, (b) stripped pine cones, (c) hazel nuts split in two and (d) hair-tubes to collect hair, and remote cameras.

Hair-tube studies have shown to be the most accurate form of indirect surveying and a number of studies have used this strategy. Finnegan et al. (2007) compared trapping and hair-tube surveys in sites in Ireland, and found that hair-tube surveys were an effective method for monitoring both red and grey squirrels, but found that variable success rates made it difficult to use hair-tubes to assess abundance levels.

One of the main drawbacks of this appraoch is that it cannot reliably distinguish between red and grey squirrels. Even in the case of hair samples, similar colour variation can exist between both red and grey squirrels, and microscopic analysis of the hair samples is needed to accurately identify the species. Furthermore, population specific questions such as the number of individuals living within the woodland are difficult to quantify using such techniques, and the use of DNA techniques may possibly address some of these gaps. Finally, there has been some debate about the possibility of disease transfer in sites that contain both species of squirrel, as the presence of hair-tubes may be inadvertently bringing the two species

within close proximity of one another and hence increasing the risk of disease transfer.

1.7 Previous genetic studies of squirrels

The majority of previous genetic studies relating to red squirrels have focused on the origins of the species, relying on a combination of geography and genetics called phylogeography (Avise 1987). The role of the Last Glacial Maximum (that ended around 12,000 years ago), and how it impacted the colonisation of plants and animals in Europe is a subject that continues to intrigue scientists, but the main line of thought is that plants and animals remained in glacial refugium in Southern Europe (Iberia, Italy and the Balkans) during the last glaciation, and then proceeded to move northwards into the rest of Europe, following the glacial retreat. The migration routes that these plants and animals took can be inferred using DNA, especially mitochondrial DNA, as it can be used to infer historical migration. The post glacial route is different for every species, and each one has to be considered independently (Hewitt 1999; 2000). This is further complicated in the case of the red squirrel due to the large numbers of translocations that have historically taken place, especially in Britain and Ireland, and the possibility that red squirrels were not restricted to glacial refugia during the last glaciation (Grill et al. 2009), as red squirrels would have survived in pine forests that were not restricted to glacial refugia (Stuart and Lister 2001).

1.7.1 Genetic studies of the red squirrel in Britain and Ireland

England

Barratt et al. (1999) examined the mitochondrial D-loop DNA (D-loop is a non-coding sequence that contains information essential for the initiation of transcription and DNA replication) of 207 red squirrels in Britain and mainland Europe and found 26 different mitochondrial haplotypes, and when analysed for relatedness found that the British haplotypes tended to cluster with various haplotypes from mainland Europe, but there was a general lack of geographic structure in the samples (Fig 1.13). Haplotypes are haploid DNA sequences defined by a set of associated nucleotide

substitutions (Beebee and Rowe 2008), and can be analysed as characters of a populations the study was limited by a small sample size from Central Europe i.e. Germany, and consequently a lack of sampling across Europe limited the inferrences that could be made regarding the origins of British squirrels in this particular study.

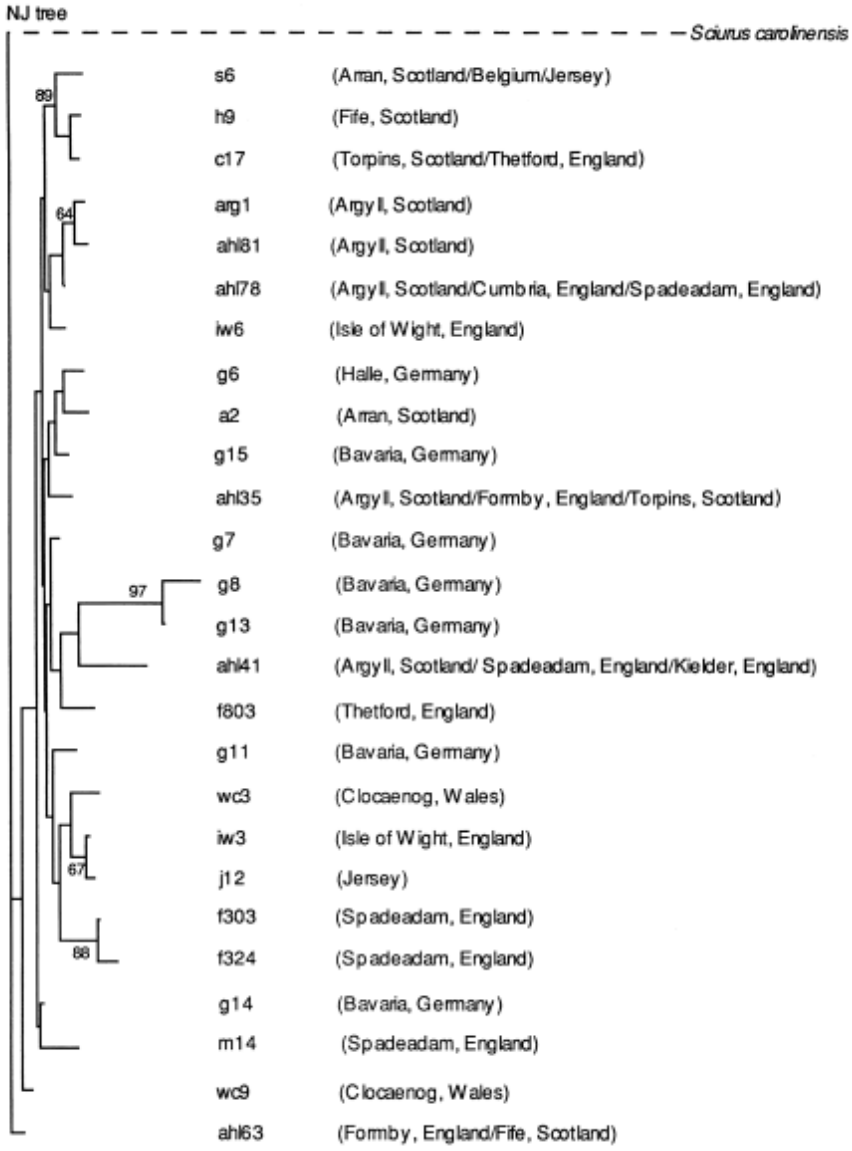

Fig 1.13: Mitochondrial D-loop haplotypes from Barratt et al. (1999) displayed in a neighbour-joining tree to illustrate geographical relationship between haplotypes.

Hale et al. (2004) investigated the genetic footprint of red squirrels in Britain, and compared the samples to others collected in Europe (Fig 1.14). From 182 individuals, a total of 28 different mitochondrial haplotypes were found, with 17 of those haplotypes recorded in Britain and only seven of the haplotypes recorded in more

than one region. Twelve haplotypes were recorded in the European countries analysed in this study.

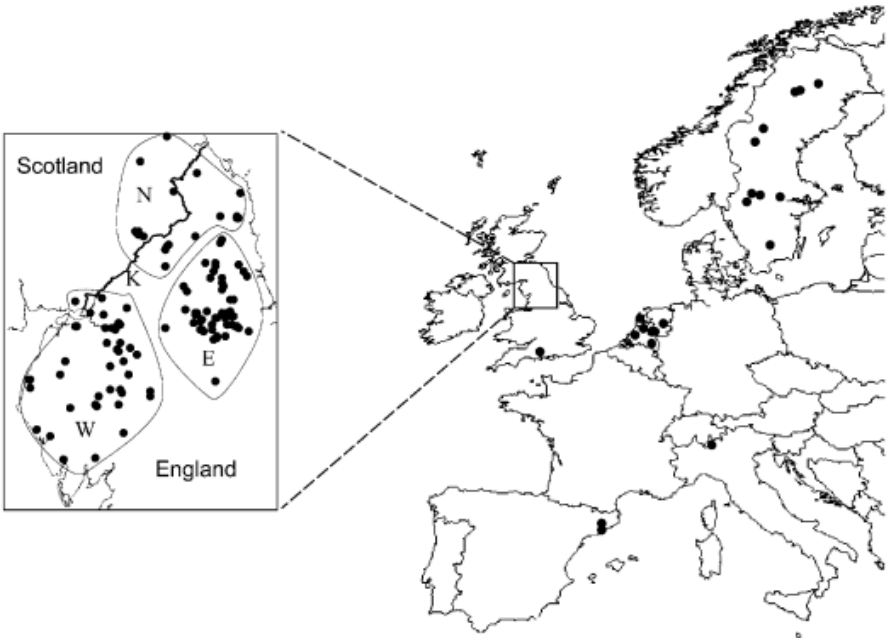

Fig 1.14: Distribution of red squirrel samples analysed by Hale et al. (2004).

Similar to Barratt et al. (1999), Hale et al. (2004) found that there was a general lack of geographic structure within Britain, and there were high levels of differentiation between each of the sites that suggested a lack of contemporary geneflow, but this was also appearant within Europe, there also appeared to be a lack of structure (Fig 1.15). Within the phylogenetic tree, H20 (Italy) groped with a Swedish haplotype (H18). H1 found in Britain was introduced from Sweden, and was found in 68% of British samples. H14 was found in both the Netherlands and in Northumberland, Britain. Hale et al. (2004) concluded that North East British squirrels primarily consisted of recent Continental European ancestry, and a large proportion of the squirrels exhibited a Swedish haplotype (H1). Hale suggested that this particular haplotype may have a selective advantage over British or Continental European red squirrels, especially in areas of conifer plantations. This study was also limited by biased sampling efforts, as the majority of the British samples originated in Northern England and Southern Scotland, with a small number of samples also being used as respresentatives from Central and Northern Europe.

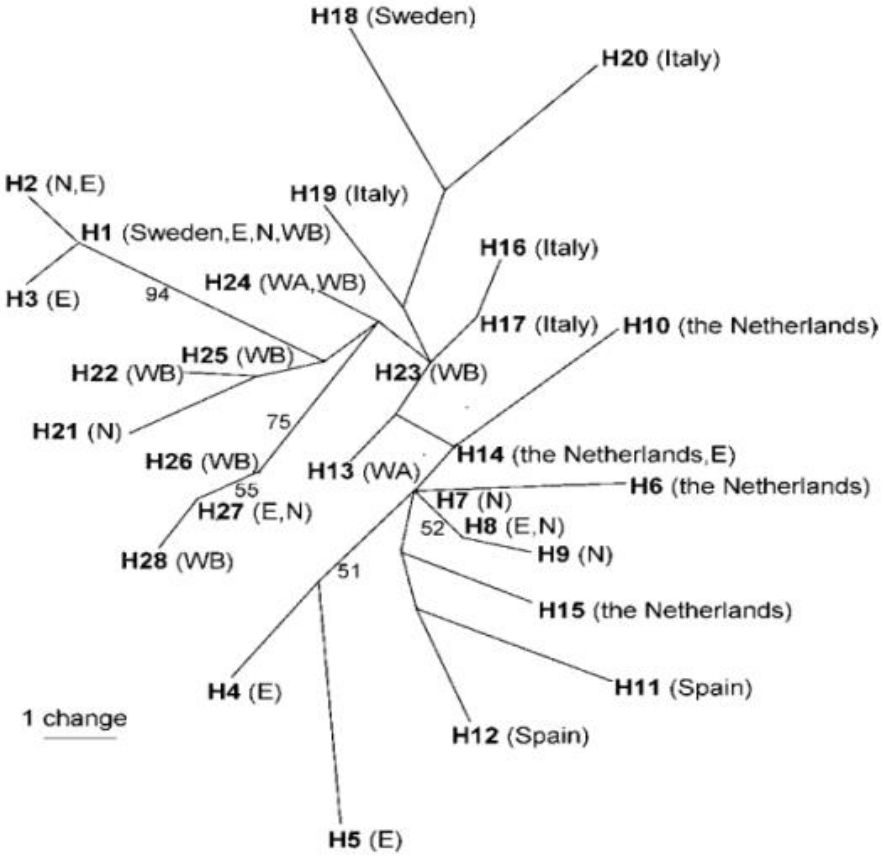

Fig 1.15: Phylogenetic tree of red squirrel mitochondrial DNA haplotypes in Britain and Europe. E = UK Eeastern region, N = UK northern region, WA = UK western region pre-1980, WB = UK western region post-1980.

Wales

The population of red squirrels in Pentraeth, Anglesey, Wales recovered to 95 individuals from about 40 individuals following extensive trapping and removal of grey squirrels in the preceding years (Ogden et al. 2005). This was facilitated by the fact that Anglesey is an island, connected to the mainland by only a bridge, and therefore possible to monitor. The study by Ogden et al. (2005) aimed to investigate the genetic diversity of this group and use the results to select individuals to relocate to another part of the island (Newborough), where red squirrels previously occurred prior to eradication due to competition with the grey squirrel. The proposed reintroduction project involved the release of captive breeding pairs into the wild. The study used mitochondrial D-loop DNA sequencing and microsatellite genotyping

to construct an applied conservation strategy for management of the red squirrel population.

Ogden et al. (2005) sampled 64 individual red squirrel samples; 48 from Pentraeth; one Mid Wales; one roadkill from 1983 in Anglesey and 12 candidate captive red squirrel candidates proposed for the Newborough reintroduction that had originated from Cumbria.

In the mitochondrial DNA analysis, three new haplotypes were found in Wales (ANG1, ANG2, MW1) and two new haplotypes were also recorded in the captive red squirrels that were proposed for the reintroduction into Newborough (WMZ1 and WMZ2). Haplotypes WC3 and WC9 were previously recorded by Barratt et al. (1999) and were found in Clocaenog, Wales. WC3 showed close homology to the mid Wales haplotype (MW1), while WC9 appeared more divergent, and more similar to Northern England haplotypes. Ogden et al. (2005) compared the Welsh sequences to those previously published by Barratt et al. (1999) and found little structure within the phylogenetic analysis, but did find that the mid Wales, Anglesey, Jersey and Isle of Wight haplotypes clustered together, suggesting some level of phylogeographic structure (Fig 1.16).

Using the microsatellite data, Ogden found low levels of genetic variability in the Pentraeth population, with three of the six loci monomorphic (exhibiting no variability), two loci displayed two alleles and only one locus displayed three alleles. Ogden et al. (2005) suggested that the low levels of genetic variability were as a result of a genetic bottleneck in the population following the recent expansion of the population. The captive and wild caught squirrels subsequently selected for reintroduction in to Newborough were more genetically diverse than the Pentraeth population, and the squirrels finally chosen were selected on the basis that they contained different alleles to the Pentraeth population.

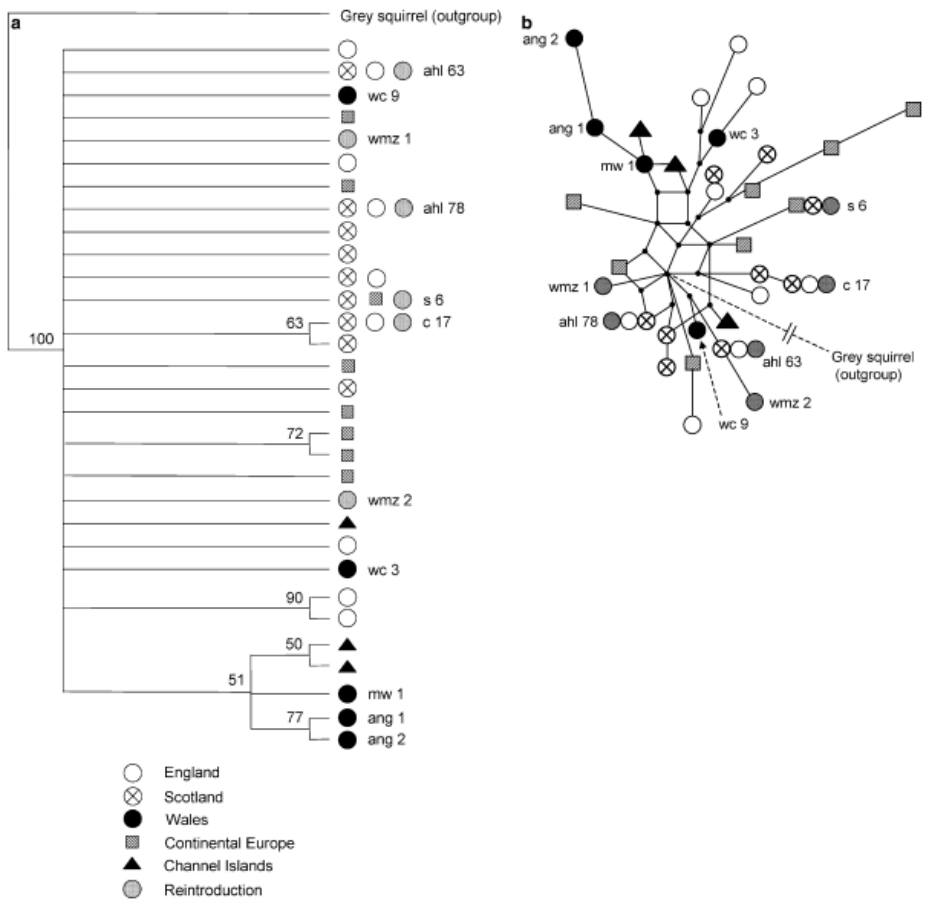

Fig 1.16: Red squirrel phylogenetic analysis from Ogden et al. (2005) comparing new mitochondrial DNA D-loop haplotypes found in that study with those previously published by Barratt et al. (1999).

Jersey

Simpson et al. (2013) investigated the origins of the red squirrels on the Jersey Island using mitochondrial D-loop DNA and microsatellite data. Simpson used samples from Jersey, the Isle of Wight, and France for comparative analysis. Only two different mitochondrial DNA haplotypes were found on Jersey, with eight variable sites between the sequences, indicating quite divergent and different origins for the squirrel population. JE120 was predominantly of British origin, while JE162 was predominantly of European origin, probably French. Phylogeny analysis revealed that many of the haplotypes were found to cluster with geographically distant haplotypes such as IOW6, IOW3, IOW9, IOW26 from the Isle of Wight that were found scattered

throughout the phylogeny. Simpson et al. (2013) showed that this provided substantial evidence for the historical movement of red squirrels by people.

Simpson et al. (2013) successfully genotyped 169 individuals and found some evidence for deviations from Hardy-Weinberg equilibrium (allele and genotype frequencies in a population stay constant over time in the absence of evolutionary influences). In this case, the deviations at some loci were not repeated across all populations and Simpson et al. (2013) suggested that this may be due to population specific effects, especially in islands where multiple introductions from different locations can cause substructuring in the area and result in deviations from Hardy-Weinberg Equilibrium, also known as the 'Wahlund Effect'. Microsatellite analysis of the Jersey data also confirmed the presence of two genetic clusters on the island that subsequently supported the introduction of both a French and British introduction on the Island.

Ireland

Finnegan et al. (2008) investigated the genetic origins of red squirrels in Ireland using mitochondrial D-loop DNA. Finnegan sampled 101 samples; 87 contemporary Irish samples and 10 museum specimens of Irish origin and four museum samples from England. The mitochondrial DNA analysis revealed a total of 26 novel haplotypes. One of the museum specimens from England and Ireland shared the same haplotype, and one English museum specimen was recorded in contemporary Irish samples (IEGB1). Only one previously described haploype (H15), from Hale et al. 2004) was found in the Irish population. Using a network analysis (a method to organise related sequences), Finngean et al. (2008) found a general lack of geographic clustering in the population, but found that a large number of haplotypes appeared related to museum specimens (Fig 1.17). The H15 haplotype previously recorded in both England and the Netherlands was shown to occur in Co. Wicklow, and two other haplotypes IE3 and IE4 appear related to this haplotype. Tests for population expansion showed that the Irish population was under going a population expansion and that this was due to relatively recent translocations and possible genetic bottlenecks.

Finnegan concluded that the Irish population was composed of a mixture of introduced and native haplotypes that have diverged since their original introduction, and highlighted the importance of the Irish population for conserving the original British haplotypes.

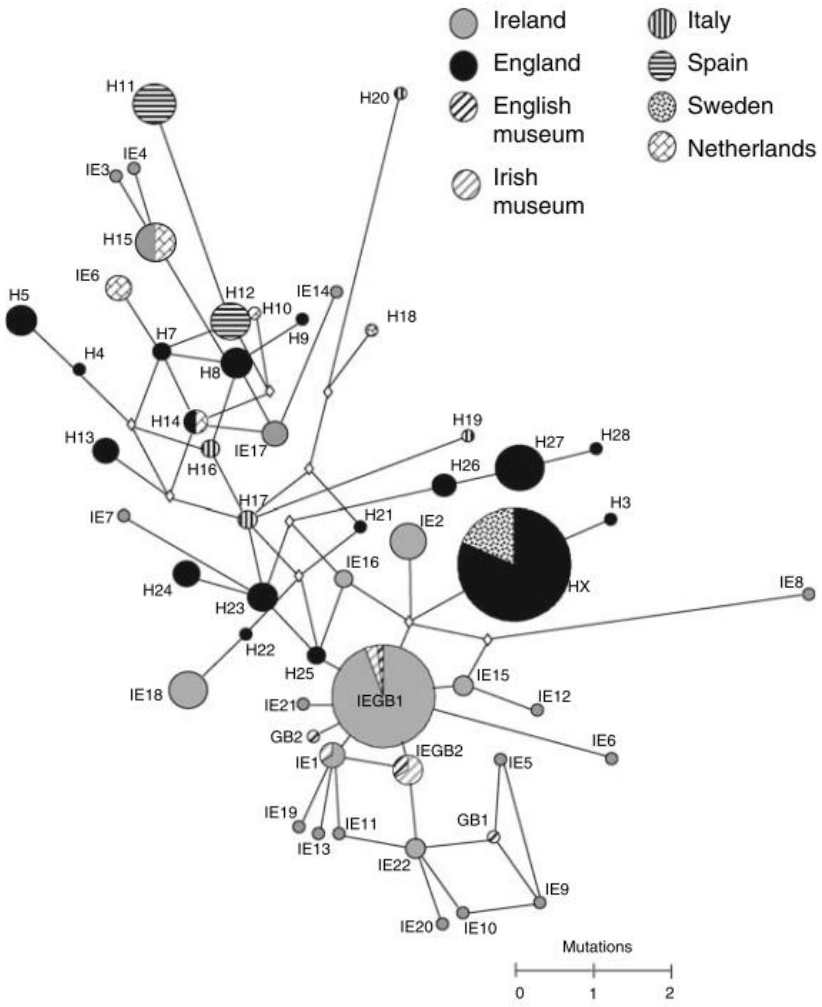

Fig 1.17: Comparison of Mitochondrial D-loop DNA haplotypes from Ireland (Finnegan et al. 2008) and compared to previous haplotypes published by Hale et al. (2004).

1.7.2 Phylogeography of the red squirrel in Europe

Grill et al. (2009) used a combination of mitochondrial DNA and microsatellite DNA to assess the genetic diversity of the European red squirrel, and assessed the likelihood of a red squirrel refuge within Italy during the Last Glacial Maximum. Grill et al. (2009) obtained samples from 17 putative subspecies of *Sciurus vulgaris*

samples across Europe. In the D-loop fragment of DNA (252 bp), 66 different mitochondrial DNA haplotypes were identified from 144 analysed individuals, indicating a high genetic diversity in the species. Each geographical area analysed contained unique haplotypes, and only two haplotypes found in this study had been previously reported. Those were two Italian haplotypes (H19 and H20) that had been previously recorded by Hale et al. (2004), and corresponded to haplotypes (SV76, SV16) in Grill et al. (2009). A neighbour-joining analysis revealed little geographic structure of the observed haplotypes (Fig 1.18). Individuals from Portugal and Calabria were the only groups to show some evidence of recent demographic expansion. The microsatellite data revealed that squirrels from Iberia formed one genetic population, with squirrels from Calabria forming a second, and the remaining animals from Central Italy, Central Europe, the Balkans, Russia and China forming a third. Grill et al. (2009) concluded that microsatellite data were more congruent with geography than the mitochondrial data. In summary, the Calabrian group was differentiated using microsatellite DNA analysis, but not as much as in mitochondrial DNA, while the Iberian group was clearly differentiated in microsatellites but not in mitochondrial DNA. Grill et al. (2009) explained that the contrasting results from the two markers could have been caused by differences in sex related dispersal, or that microsatellite data represent present rather than historical gene flow.

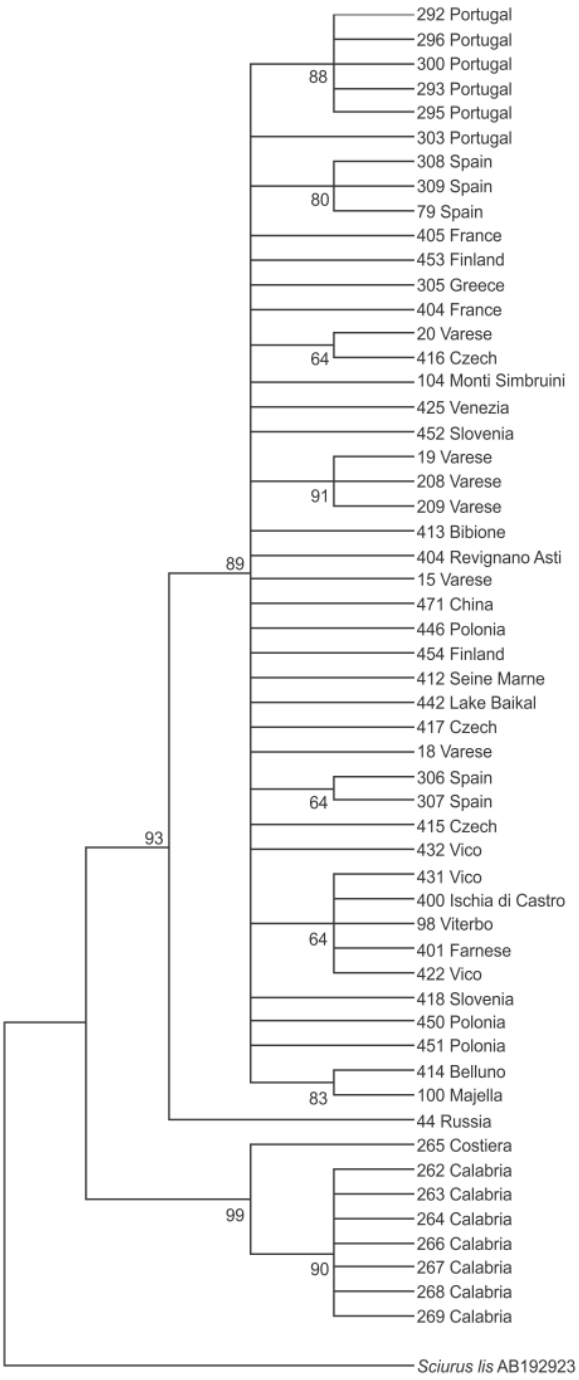

Fig 1.18: Neighbour joining tree of red squirrel D-loop mitochondrial DNA haplotypes from across Eurasia showing little geographic structure (Grill et al. 2009).

Dozières et al. (2012) conducted an extensive study of the genetic variation of the red squirrel in France using a section of the mitochondrial DNA D-loop and they compared the results to Grill et al. (2009). The first part of the study involved the

assessment of the mitochondrial DNA haplotypes present in France and to compare them to haplotypes previously recorded in Europe. Grill et al. (2009) found no differentiation of the French red squirrel population from the rest of the European sample, but Dozières suggested that this may be a consequence of the small sample (n = 6) used to represent the French population.

Dozières et al. (2012) collected red squirrel tissue samples from throughout France, that mostly came from road kill (n =110). Mitochondrial D-loop DNA was amplified using the primers designed by Trizio et al. (2005), amplifying a section of 516 bp of DNA. The results showed that there were 71 different haplotypes found in the French population, and these haplotypes (as in Grill et al. 2009) did not generally conform to geography.

1.7.3 The use of genetic analysis to assess fragmentation of red squirrel populations

Hale et al. (2001a) assessed the impacts of defragmentation that has increased tree cover between Northern England and Scotland using microsatellite analysis with just four microsatellite loci Scv3, Scv8, Scv9, and Scv10 (Hale et al. 2001b). The study used museum specimens, some of which were over 100 years old to assess levels of contemporary and historical genetic diversity. The study found a startling change in genetic composition over the last 100 years. Despite the use of a very small microsatellite panel there was evidence for rapid introgression of northern genes into western populations which led the authors to conclude that human movements into the region and the deforestation of woodlands caused the initial substructuring to happen. The study highlighted the importance of habitat patches to create ecological corridors and aid dispersal.

Trizio et al. (2005) analysed the effects of isolation and habitat fragmentation on the red squirrel in the Italian Alps across seven sites. This study used a combination of microsatellite and mitochondrial DNA data. The study sites included a range of diverse habitats within the Italian Alps including mature, secondary montane and sub-Alpine mixed conifer forests across altitudes ranging from 1,100 to 2,100 m

above sea level. Sample collection involved live trapping of individual squirrels and 30-50 hairs were pulled from each animal, or alternatively tissue samples were collected by taking a small sample from the ear. For the mitochondrial DNA analysis, Trizio et al. (2005) combined samples from their study with sequences published by Barratt et al. (1998). Twenty three unique mitochondrial DNA haplotypes were found from 70 individuals. However, these haplotypes were evolutionarily very similar, and out of 378 bp, only 34 sites were informative. Trizio et al (2005) found that the sequence analysis revealed a lack of geographical structure and found some admixture with the British samples analysed in the study. The microsatellite data showed that the red squirrels exhibited high levels of genetic diversity, with allele numbers ranging from four to 10 per locus. High levels of population substructure were found within the Italian Alps, possibly caused by geographic isolation.

1.7.4 Red squirrel microsatellites

Todd (2000) developed five microsatellite loci for amplification with red squirrel DNA; Rsu1, Rsu3, Rsu4, Rsu5 and Rsu6, all of which were found to be polymorphic with red squirrel DNA from 163 individual red squirrels from Belgium and Germany. Within Belgium and Germany, the loci were shown to have a range of alleles from four at Rsu6 to eight at Rsu4.

Hale et al. (2001a) developed a total of 21 microsatellite loci for red squirrels, and were also generally polymorphic with the grey squirrel. The loci included Scv1, Scv3, Scv4, Scv6, Scv8, Scv9, Scv10, Scv12, Scv13, Scv14, Scv15, Scv16, Scv18, Sc19, Scv20, Scv23, Scv24, Scv25, Scv27, Scv31 and Scv32. The number of alleles per locus ranged from two at Scv12 and Scv18 to 12 at Scv3. The samples obtained came from three sources in Northern England, and amplification success ranged from 37 individuals amplifying at Scv14 and Scv24 to as few as 14 individuals amplifying at Scv6 and Scv16.

Gunn et al. (2005) developed microsatellite loci for the North American red squirrel (*Tamiasciurus hudsonicus*) and found that 11 of the loci were also polymorphic in the Eurasian red squirrel. Those loci included: Thu03, Thu14, Thu23, Thu25, Thu31, Thu33, Thu37, Thu41, Thu42, Thu50 and Thu55. The Eurasian red squirrel samples

originated from Britain, and the number of alleles ranged from two at Thu3, Thu23 and Thu31 to eight alleles at Thu14. Twelve of the loci were also found to cross amplify with North American grey squirrel DNA.

Painter et al. (2004) designed six microsatellite loci for the Siberian flying squirrel (*Pteromys volans*), and tested the cross species amplification with red squirrel DNA from five individuals. Only one marker (PvolE6) amplified with the five samples, although the number of alleles was not assessed.

Other microsatellite loci have been developed for other squirrel species including Fike et al. (2009) who developed 26 microsatellite markers for the North American fox squirrel (*Sciurus niger*), and cross-species testing revealed that 22 of those loci were also polymorphic in the grey squirrel. Shibata et al. (2003) designed 11 microsatellite loci for the Japanese Squirrel (*Sciurus Lis*), but the authors did not test those loci with DNA from other squirrel species.

1.8 Non-invasive genetics

Non-invasive genetics involves the use of a set of molecular tools that can be used to remotely study a species in the wild without needing to observe, trap or disturb the animal. The technique is especially useful as a means to study rare, endangered and elusive species and has aided the conservation and management of many species since the 1990s (Beja-Pereira et al. 2009). The technique was first introduced in the 1990s as a way to collect samples from brown bears (*Ursus arctos*), and since then the general idea has been technically advanced through the development of better forensic techniques to deal with samples of low quality and quantity DNA (Taberlet and Bouvet 1992; Waits and Paetkau 2005). The source of DNA used in non-invasive studies includes scats, urine, hair, feathers, saliva and environmental DNA (Beja-Pereira et al. 2009). The most frequently addressed question in these studies is the correct identification of the species, and the identification of unique individuals from genotyping. Non-invasive genetics has been used to address other population specific questions such as gene flow, genetic bottlenecks and the presence of genetic structure (Beja-Pereira et al. 2009).

1.8.1 Species identification

The first step in the analysis of non-invasive DNA samples is to ensure the target species DNA is present in the sample. One of the most general approaches is the use of primers that will amplify a range of species, such as the approach taken by Statham et al. (2005) (Fig 1.19). In that study, Statham et al. (2005) amplified mitochondrial D-loop PCR products for the fox (*Vulpes vulpes*), pine marten, mink (*Neovison vison*) and stoat (*Mustela erminia*) using the sample primer pair.

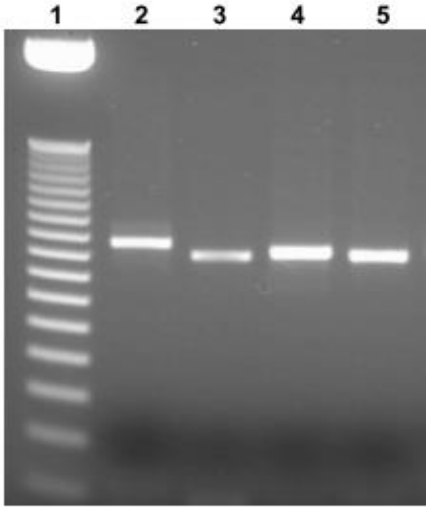

Fig 1.19: PCR products amplified with primers LRCB1 and H16498M from Statham et al. (2005). *Lane 2* fox, *lane 3* pine marten, *lane 4* mink, and *lane 5* stoat. A 50 bp DNA ladder is provided for sizing, with PCR products randing from 450 – 500 bp.

RFLP Assays

Typically, studies have relied on conventional PCR techniques for species-identification of the target DNA. Examples include mitochondrial DNA RFLP (Restriction Fragment Length Polymorphism) assays, whereby a number of similar target sequences are co-amplified using the same primers. The assays are developed by aligning sequences of the same region across target species and selecting primers that will co-amplify all targets. The next step involves locating restriction enzyme cleavage sites within the selected primer sequence, used to identify species-

diagnostic enzyme combinations for RFLP screening (Ruiz- González et al. 2008). Following the restriction enzyme digestion of PCR products, the products can be separated and visualised on agarose gels stained with ethidium bromide (Fig 1.20).

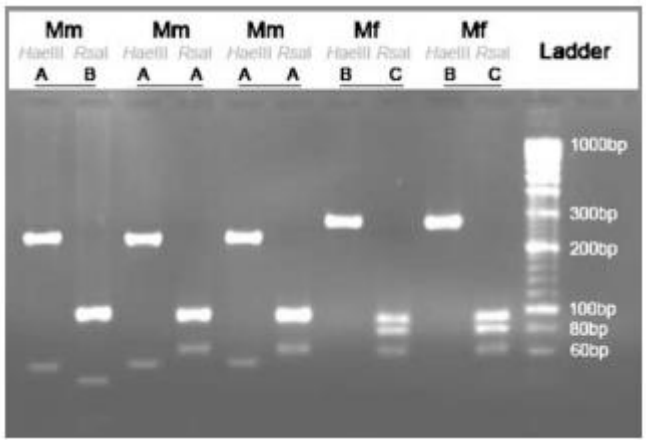

Fig 1.20: Example of an agarose gel showing ethidium bromide stained bands and diagnostic restriction enzyme patterns (AB; AA; BC) generated using restriction enzymes for *Martes martes* (Mm) and *Martes foina* (Mf) mitochondrial D-loop PCR products (Ruiz-González et al. 2007).

Positive or negative amplification

Another strategy relies on positive or negative amplification of the target DNA sequence using species-specific primers. This approach can be advanced by multiplexing multiple primer pairs, thus targeting multiple species. The target species can subsequently be discriminated by size fragment differences when the PCR products are visualised on agarose gels. Fernandes et al. (2008) used this approach to target Iberian carnivore species, with the assays designed to amplify a short region of cytochrome b DNA that could be applied to non-invasive samples. Each primer pair was designed to specifically amplify the targeted species of interest (Fig 1.21). Similarly, Meganathan et al. (2010) designed a multiplex PCR system to discriminate three species of crocodile by amplifying species-specific mitochondrial DNA amplicons.

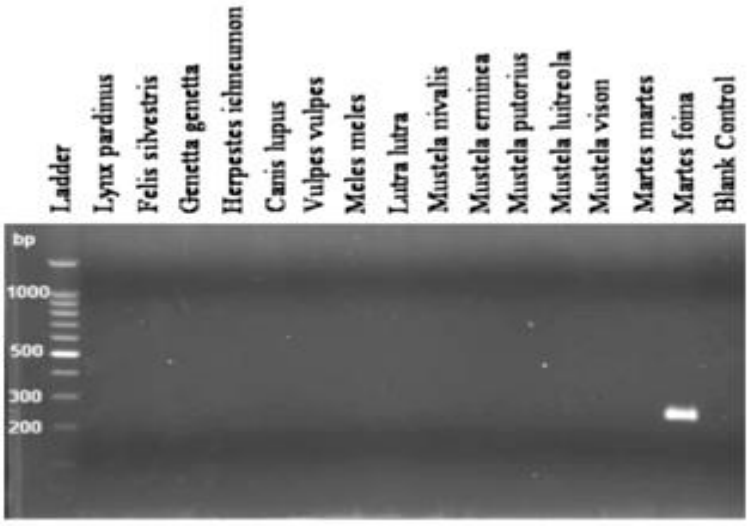

Fig 1.21: An example of species-specific PCR where primers specifically designed to amplify DNA from stone marten (*Martes foina*) only amplify DNA from that species and no cross-species reactivity was seen with the DNA from other species' DNA (Fernandes et al. 2008).

Real-time PCR

As conventional PCR detection of target species requires a relatively large quantity of DNA template, there may be occurrences of false negatives as a result of having a sample with degraded or low quality DNA (Matejusová et al. 2013). An alternative strategy relies on the use of real-time PCR for species detection. The assays are typically designed to amplify ~100 bp fragment of target DNA, and due to the small size amplicon, amplification success of low or degraded DNA samples can be increased (Matejusová et al. 2013). The technique has been successfully applied to species identification of carnivores, small mammals and fish (Moran et al. 2008; O'Reilly et al. 2008; Mullins et al. 2010; Baerwald et al. 2011; Matejusová et al. 2013; O'Neill et al. 2013).

Real-time PCR chemistries

Two main chemistries are used in this project for the detection of target DNA: intercalating dyes (SYBR® Green) and fluorogenic probes (TaqMan® MGB). The difference between these chemistries is sequence specificity; intercalating dyes will

bind to any double-stranded DNA in the SYBR® Green 1 assay, while the TaqMan® probes are specific to the target DNA sequence (Fig 1.22). In the SYBR® Green 1 assay, the intercalating dye is added to the DNA sample where it binds to the double stranded DNA present, and as the PCR continues the increase in the fluorescence is proportional to the increased copies of the generated targeted amplicon. The confirmation of specific amplification is then confirmed with melt-curve analysis which calculates the Tm of the amplified DNA based on the loss of fluorescence due to the separation of the two DNA strands (Mullins 2010). The TaqMan® probe principle relies on the 5′–3′ exonuclease activity of Taq polymerase that cleaves a dual-labelled probe as it hybridizes to the complementary target sequence. The resulting fluorescence signal emits a quantitative measurement of the increase of the amplicon during the exponential stages of the PCR. The addition of the TaqMan® probe increases the specificity of the detection (Cao and Shocky 2012) (Fig 1.22).

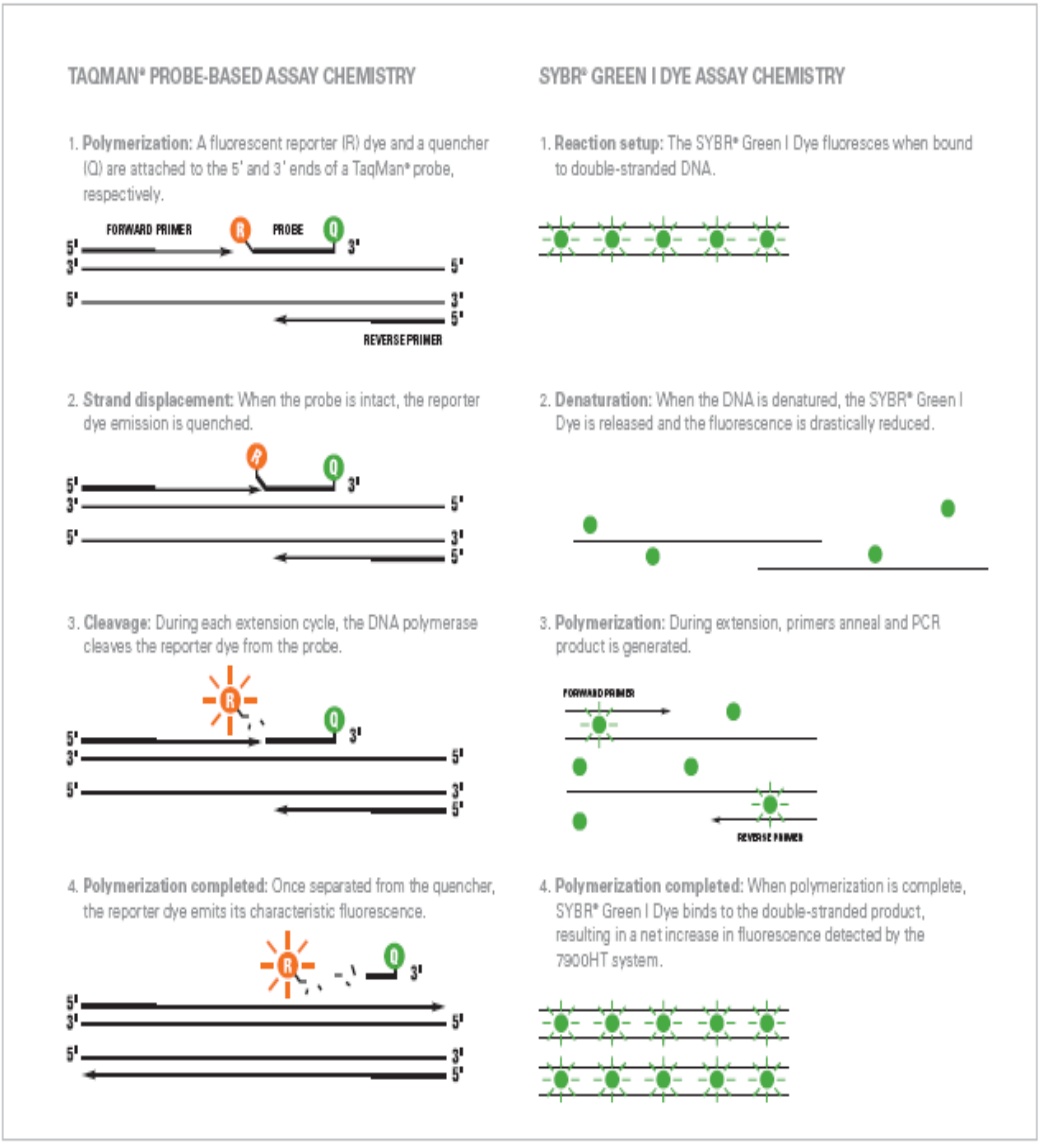

Fig 1.22: Comparison of TaqMan® and SYBR® Green based detection workflows (Applied Biosystems).

A positive result is detected by accumulation of a fluorescent signal during the real-time PCR reaction (Fig 1.23). The corresponding Ct (cycle threshold) is described by the number of cycles required for the fluorescent signal to exceed the threshold or background fluorescence (Applied Biosystems). The resulting Ct values are inversely proportional to the amount of target DNA in the sample (i.e. the lower the Ct value the greater the amount of target DNA in the sample) (Applied Biosystems). In general terms, positive Ct values $\leq$ 30 indicate an abundant quantity of target DNA, while positive Ct values from 31 - 37 indicate a moderate quantity of target DNA. Ct values greater than this value can be discounted on the basis that the target DNA quantity is of low quality. This might suggest that the target DNA had potentially

come from a source of field or laboratory contamination. The Tm is described as the melting temperature of the PCR product and is monitored during melt curve analysis to confirm specific DNA amplification; deviations from the described Tm of the target species may indicate non-specific amplification such as primer-dimer formation, and would result in a significantly altered Tm (Moran et al. 2008).

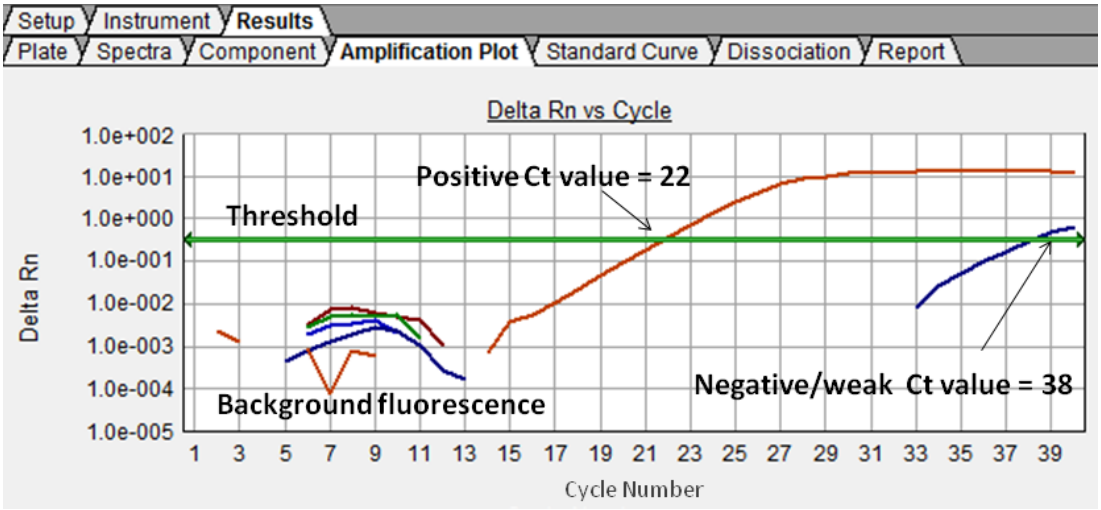

Fig 1.23: An example of a real-time PCR plot showing a positive result (above the threshold), a weak/negative result and the presence of background fluorescence (noise). Published as supplementary material in O'Meara et al. (2014).

Studies that use real-time PCR for species detection

Real-time PCR can be used for the detection of DNA from non-invasively collected samples due to its high sensitivity, small product size (70 - 150 bp) and high throughput (Beja-Pereira et al. 2009). Real-time PCR has been used to identify samples to species from non-invasively collected samples (Moran et al. 2008; O'Reilly et al. 2008; Mullins et al. 2010; O'Meara et al. 2012; O'Neill et al. 2013).

O'Reilly et al. (2008) used real-time PCR to detect pine marten and fox using DNA extracted from scats using SYBR® Green 1 and TaqMan® MGB probe real-time polymerase detection. Primers and probes were designed by downloading the targeted available sequences from Genbank and aligning them. In this case, fox and pine marten mitochondrial D-loop sequences were used to compare the differences

and similarities between the sequences. The primers and probes were subsequently designed using Primer Express software (Applied Biosystems). O'Reilly et al. (2008) found a region that had five bp differences between the two species, and focused in on this section to design a probe (Fig 1.24). In the SYBR® Green 1 real-time PCR reaction, the forward and reverse primers were used without the addition of the species-specific probe, which amplified both fox and pine marten DNA. The species could be distinguished post-PCR by analysing the Tm values of each reaction as both species (fox and pine marten) produced significantly different Tm values from one another. The TaqMan® real-time PCR reaction included the addition of two species-specific probes, one designed specifically to target the fox DNA sequence, and the other specifically to the pine marten DNA. The probes are fluorescently labelled so that only the target species is amplified, and no post real-time PCR analysis is required.

O'Reilly et al. (2008) developed the assays by initially testing the assays with DNA extracted from samples from of known species. The assays were secondly tested with DNA extracted from field collected scat samples.

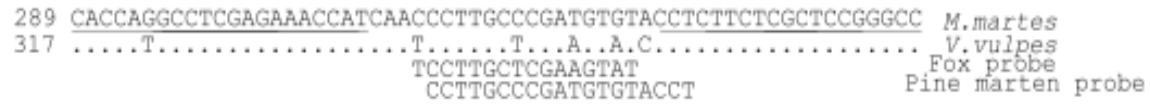

Fig 1.24: DNA alignment of a short region of the mitochondrial D-loop sequences of fox and pine marten showing the position of primers (underlined) and probes used in O'Reilly et al. (2008).

DNA Barcoding for species identification

One of the limitations of the species identification techniques discussed thus far is that there is a lack of standardisation in the approach, such as differences in the targeted DNA sequences. This can be overcome using a standard section of short sequence (< 250 bp) diagnostic mitochondrial DNA sequences, called minimalist DNA barcodes or mini-barcodes and can be a powerful tool for identifying degraded or non-invasive DNA (Chaves et al. 2012). The proposed gene target for mammals is the

cytochrome c oxidase 1 gene (COI), a section of the mitochondrial DNA which represents a 658 bp portion of DNA due to its high variability (Fig 1.25). The overall aim of using a standardised gene for species-identification was to enable a standard approach to biodiversity inventory, conservation and the detection of illegal movement of wildlife (Allendorf and Luikart 2007).

Chaves et al. (2012) sequenced sections of the COI, ATP6 and cytochrome b genes to assess their use for identification of 66 species of carnivores (Fig 1.24). The resulting primer pairs targeted less than 200 bp of sequence and the primers amplified 86% (ATP6), 85% (COI) and 70% (cytb) of the species panel, showing some potential for the use of these segments as standard markers for carnivores. The study also highlighted problems with the use of "universal" or group specific primers due to the high variability of the targeted sequence region which makes it difficult to design primers that will amplify entire groups.

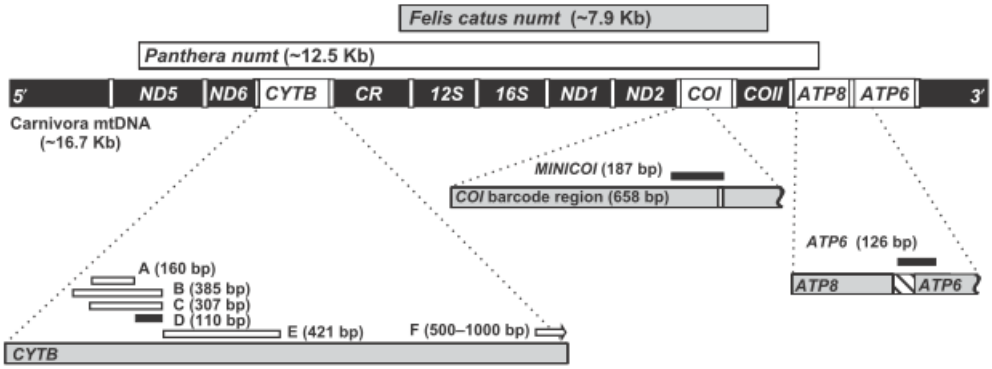

Fig 1.25: Schematic view of a linearised mitochondrial DNA molecule showing the relative positions of most coding and noncoding regions (Chaves et al. 2012).

1.8.2: Molecular dietary studies

Molecular analysis as a tool to study the diet of species has grown as a research field over the last number of years. Deagle et al. (2005) published one of the pioneering studies in the field by investigating prey DNA in the scats of steller sea lions (*Eumetopias jubatus*). Until then, most molecular dietary studies concentrated on invertebrates, where the entire organism was used for the DNA extraction (destructive sampling) (Deagle et al. 2005). There are two general approaches in

molecular dietary analysis. The first involves the use of group specific primers and the second involves non-specific primers that will amplify a broad range of species, the former requiring some level of knowledge of the species diet (Valentini et al. 2009).

Some molecular dietary studies now employ next generation sequencing (NGS) 'massively parallel sequencing' to genetically identify components in the diet. NGS differs from conventional DNA sequencing as it works on the basis of parallel sequencing by yielding shorter reads (~21 to ~400 bp), but yields millions of reads. Unlike conventional or Sanger based sequencing methods that require DNA amplification (the final sequence is representative of a model population of DNA templates), sequencing can now be performed from single DNA molecules using the next generation sequencing technologies (Reis-Filho 2009).

Shehzad et al. (2012) employed next generation sequencing (NGS) to genetically identify the diet of the leopard cat (*Prionailurus bengalensis*). The authors used traditional DNA extraction techniques and then used a universal PCR primer set to target vertebrate DNA, targeting an amplicon of 100 bp of the 12S rRNA gene. The authors also used a blocking oligonucleotide that was designed specifically to the leopard cat sequence. Shehzad et al. (2012) ran the PCR products on a next generation sequencer and identified mammalian, avian and amphibian and fish DNA. The technique could be applied to other species and could also be used on older platforms.

Real-time PCR for molecular dietary studies

Murray et al. (2011) studied the diet of the little penguin (*Eudyptula minor*) using real-time PCR to target four prevalent fish species in the diet using species-specific primers. Murray et al. (2011) found that the results obtained from real-time PCR were highly comparable with the results obtained using an NGS approach. Bowles et al. (2011) also successfully applied real-time PCR to a dietary study of the stellar sea lion to assess for the presence of herring (*Clupea palasii*), eulachon (*Thaleichthyes pacificus*), squid (*Loligo opalescens*) and rosethorn rockfish (*Sebastes*

helvomaculatus). Real-time PCR can therefore be successfully applied to study the target specific species of interest in the diet.

1.8.3 Molecular sex determination

Attaining the knowledge of the sex of individuals in a population aids the study of breeding and social systems, evolutionary ecology, and provides crucial data for management and conservation programs (Rosel 2003). One of the drawbacks of traditionally getting this information is that the animal needs to be trapped, which as previously outlined is time and labour demanding. An alternative approach may be the use of molecular sexing techniques that can be applied to tissue, hair or non-invasive sources of DNA such as faecal DNA of the species.

The two main approaches for molecular sex determination using polymerase chain reaction (PCR) rely on amplification of fragments specific to the Y-chromosome or amplification of homologous fragments from both X- and Y-chromosomes (Aasen and Medrano 1990; Pomp et al. 1995; Shaw et al. 2003).

Zinc Finger Protein (ZF)

The zinc finger (ZF) protein is present on the X- and Y-chromosome, with the Zinc finger X (ZFX) and the Zinc finger Y (ZFY) DNA sequences showing a high degree of similarity to one another (Palsboll et al. 1992; Statham 2005). PCR amplification of this region will result in the generation of sex specific PCR products.

The method described by Shaw et al. (2003) used primers developed by Cathey et al. (1998) to target a ZF intron. Males and females can be identified by visualising the post PCR products using gel electrophoresis, where females are identified by the presence of a single PCR product (ZFX) and males are identified by the presence of a second size differentiated band (ZFY) (Fig 1.26 [a]). Aasen and Medrano (1990) relied on the co-amplification of ZFX and ZFY, a conserved exon spanning introns that can be separated by RFLP analysis (Fig 1.26 [b]) or direct sequencing of the PCR products, where females are detected by visualisation of single peaked sequence chromatograms, while males are identified by the presence of double peaks in the sequence chromatogram (Fig 1.26 [c]).

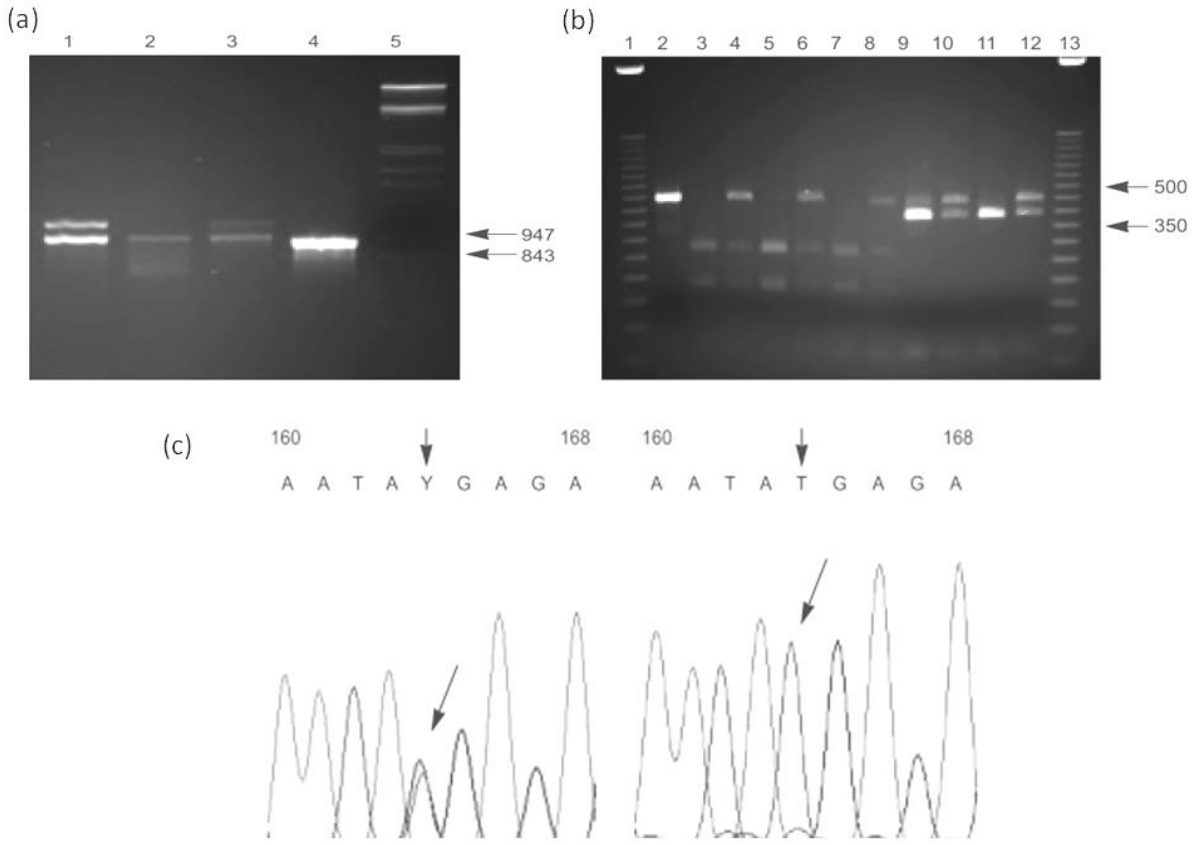

Fig 1.26: Methods of sex identification using ZFX/ZFY (Statham et al. 2007). (a) An example of sex identification using primers, LGL331 and LGL335 to amplify ZFX/ZFY products. *Lane 1 and 3* contain male pine marten (ZFX and ZFY visible) and *lane 2 and 4* female pine marten (only ZFX visible). (b) An example of sex identification using the primers P1-5EZ and P2-3EZ to amplify ZFX/ZFY product. *Lanes 3 and 4* female and male pine marten, *Lanes 5 and 6* female and male mink, *lanes 7 and 8* female and male badger, *lanes 9 and 10* male and female stoat, and *lanes 11 and 12* male and female otter. The PCR products were restricted. (c) DNA sequencing chromatograms of ZFX/ZFY PCR products generated using P1-5EZ and P2-3EZ primers. The Y arrow denotes the presence of a double peak representing the ZFX and ZFY SNPs in a male stoat, and the T arrow shows the absence of the second peak (ZFY) in a female stoat.

While the ZF gene has been shown to be very useful for the sex determination of a range of a range of mammal species, previous studies that have attempted to amplify the ZF gene in the Sciuridae have found that the region may not be sex specific

(Ermakov et al. 2006). Ermakov et al. (2006) reported that both male and female ground squirrels (*Spermophilus spp*.) produced the same size PCR products when amplified with ZFY specific primers (Fig 1.27 a). This led Ermakov to conclude that the ZFY region was not exclusively attributed to the Y-chromosome of ground squirrels, and that the region mostly likely occurred on the ground squirrel autosome. The region has not however been specifically tested with either the red squirrel or the grey squirrel.

Sex determining region of the Y chromosome (SRY)

The SRY is known as the testis determining factor, and is an example of a region that is specific to the Y-chromosome (Statham et al. 2007). One of the problems with amplification of just a male specific region could result in a false negative result (Durnin et al. 2007), and measures to circumvent this include the amplification of another PCR product (present in both male and females targets) to act as an internal control (Durnin et al. 2007). This can introduce additional problems, especially when dealing with non-invasive or low quality DNA samples, as some genes are more likely to amplify such as mitochondrial DNA, which is more abundant than nuclear DNA (used in sex determination) (O'Neill et al. 2013). The SRY gene has been commonly used for sex discrimination in a wide selection of fauna including a variety of carnivores and rodents such as the North American beaver (*Castor canadensis*) and the primitive Point Arena mountain beaver (*Aplodontia rufa nigra*) (Kurose et al. 2005; Goldberg et al. 2011; Pilgrim et al. 2012).

A number of studies have used the SRY gene for sex determination of the Sciuridae. Ermakov et al. (2006) encountered difficulties when attempting to amplify PCR products using SRY primers in ground squirrel species such as non-specific binding that occurred in both sexes of ground squirrel (Fig 1.27 b). Kusahara et al. (2006) used conserved SRY primers to successfully sex Pallas's tree squirrel (*Callosciurus erythraeus*), but Gorrell et al. (2012) found that these primers failed to differentiate the sex of squirrels in their study species, one of which included the North American grey squirrel (*S. carolinensis*). Bryja and Konecny (2003) also obtained mixed results for the use of the SRY gene as a method for the gender determination of the red

squirrel and found that two sets of SRY primers, developed by Pomp et al. (1995) and Sánchez et al. (1996) provided conflicting results. Bryja and Konecny (2003) used primers developed by Pomp et al. (1995) and found that they successfully amplified with male red squirrel DNA samples, but found that SRY primers developed by Sánchez et al. (1996) failed to generate reliable male specific PCR products.

Minor histocompatibility complex antigen (SMCY)

The SMCY gene (minor histocompatibility complex antigen) was previously found to be male specific in four species of ground squirrel (Ermakov et al. 2006) (Fig 1.27 c). Gorrell et al. (2012) also successfully designed conventional PCR assays based on this gene for the North American grey squirrel and the North American red squirrel (*Tamiasciurus hudsonicus*) (Fig 1.26). Chang et al. (2011) focused on the SMCY gene to derive Y-chromosomal sequence from a variety of Asian striped squirrels, genus *Tamiops*. The study used primers targeting SCMY7 a Y-chromosome conserved anchored tagged sequence developed by Hellborg and Ellegren (2006). The SMCY gene is unusual in that it is also present in the Marsupial Y and is the only Y-chromosomal gene to be found in a single copy in both mice (*Mus mus*) and humans (Agulnik et al. 1999). Gorrell et al. (2012) also successfully designed conventional PCR assays based on this gene for the North American grey squirrel and the North American red squirrel (*Tamiasciurus hudsonicus*) (Fig 1.28).

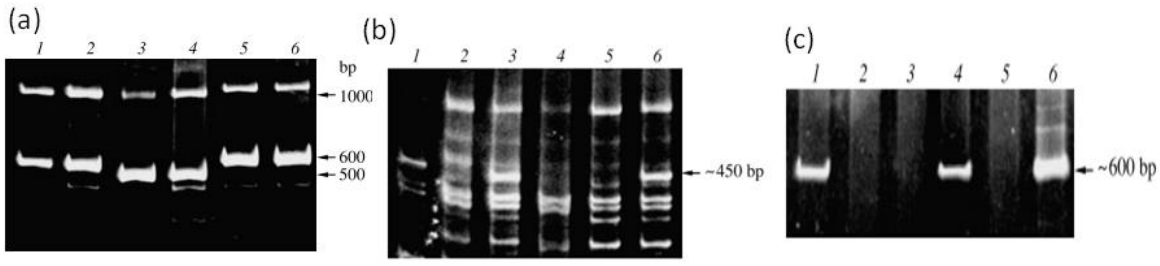

Fig 1.27: Gel images from Ermakov et al. (2006). (a) The amplification of both male and female ground squirrels with ZFY primers. (b) Non-specific binding of male and female ground squirrels with SRY primers. (c) Amplification of male ground squirrels with SMCY primers. No amplification was observed with DNA from females.

DEAD box Y-linked gene (DBY)

The polypeptide 3, Y-linked DBY gene has also been successfully used to sex a number of species using the conserved DBY primers developed by Hellborg and Ellegren (2003). Assays have been developed for the naked mole rat (*Heterocephalus glaber*), felids and an insectivore, the Iberian desman (*Galemys pyrenaicus*) using the DBY gene (Sugimoto et al. 2006; Katsushima et al. 2010; Vidal et al. 2010). The assay can be used in conjunction with the DBX (a homologous gene on the X-chromosome) gene or another mitochondrial or nuclear gene, used as an internal control. It has not previously been tested with any Sciurids.

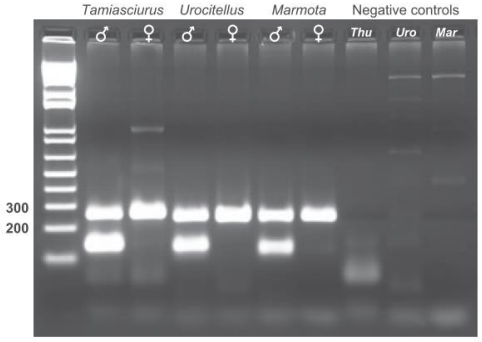

Fig 1.28: The amplification of SMCY products and a microsatellite positive control in three North American Sciurid species including a tree squirrel, ground squirrel and marmot. Males are identified by the presence of a second, smaller band. The primers used to generate the *Tamiasciurus* SMCY gene were also shown to amplify the North American grey squirrel (Gorrell et al. 2012).

Peculiarities in the squirrel Y-chromosome

Li et al. (2004) studied the organisation of the North American grey squirrel karyotype through the process of cross chromosome painting. Li et al. (2004) defined whole genome homologies between humans and representative squirrel species including two tree squirrels the grey squirrel and *Callosciurus erythraeus*, a flying squirrel (*Petaurista albiventer*) and a chipmunk (*Tamias sibiricus*) through a process of cross-chromosome painting. The results showed that the X-chromosome was homologous in all species, consistent with the evolution of the mammalian genome. However,

large blocks in the genome of the grey squirrel, (unpainted with human chromosome specific probes), showed cross-hybridization signals with painting probes specific for the grey squirrel Y-chromosome and other grey squirrel chromosomes (Fig 1.29). Additional analysis suggested the presence of DNA repeats in these heterochromatin blocks. This suggests that there may be sections of DNA present on the squirrel Y-chromosome that are also distributed across other chromosomes.

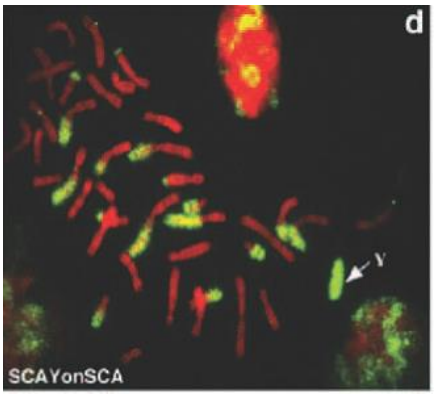

Fig 1.29: Hybridization pattern of SCA Y probe showing cross-hybridization with other heterochromatic regions. Areas of hybridiziation are shown in green (Li et al. 2004).

The use of real-time PCR for sex determination

Real-time PCR has been successfully used to determine the sex of non-invasive samples including mammals and birds (Mullins et al. 2010; Ahlering et al. 2011; Brubaker et al. 2011; Matejusová et al. 2013; O'Neill et al. 2013). Mullins et al. (2010) targeted X- and Y-chromosome specific sequences of the Zinc Finger gene. X- and Y-specific primers and probes were designed to amplify 79 and 75 bp amplicons respectively. The probes were labelled with two different reporter dyes, VIC or 6-FAM to enable multiplexing of the assays, with the X-chromosome target acting as an internal control for the reaction. Using the sex specific DNA sequences generated by Statham et al. (2007), O'Neill et al. (2013) designed ZFX and ZFY real-time PCR assays to amplify short sections of DNA sequence from the X- and Y- chromosome

using DNA extracted from otter spraints (faeces). In this case, the ZFX assay also acted as an internal control for the assay, as the ZFX was present in both male and female samples.

1.8.4 Individual identification

Identifying individual animals from non-invasive samples such as hair and faeces has become popular since the 1990s, and has enabled the identification of a wide range of species (Broquet et al. 2007; Beja-Pereira et al. 2009). To address some of the shortcomings experienced in non-invasive genetic identification studies such as allelic dropout and the presence of null alleles a number of strategies have been developed to ensure the accuracy of the data. For instance, it is necessary to replicate all amplifications when using non-invasive samples and a consensus genotype should be generated to accurately identify individuals. In addition, microsatellite markers can be redesigned to amplify PCR products < 200 bp which helps avoid allelic dropout. Such an approach is commonly seen in non-invasive studies where amplicons are between 200-300 bp (Beja-Pereira et al. 2009). Other studies also use SNPs (Single nucleotide polymorphism), that can experience a higher amplification success rate with non-invasive samples, but larger numbers of SNPs are generally needed than microsatellites due to the lower levels of heterozygosity in SNP markers that can increase the genotyping error (Beja-Pereira et al. 2009). Despite the potential for problems with the technique, non-invasive genotyping of individuals and the analysis of the subsequent data is becoming one of the most time effective and accurate methods for wildlife monitoring (Waits and Paetkau 2005).

Mullins et al. (2010) used a non-invasive genetic approach to census and monitor the genetic diversity of a population of pine marten in a wood in Co. Waterford. Mullins used a series of techniques to identify pine marten from scat and hair samples, using real-time PCR assays to differentiate pine marten from fox samples, and developed sex identification assays targeting the X- and Y-chromosome for the gender identification of pine marten. Finally, a panel of microsatellite markers were optimised, some redesigned, to identify individual pine marten. Mullins et al. (2010) were able to apply their suite of molecular tools to the hair samples that had been

collected using hair-tubes (Fig 1.30), a non-invasive approach to genetically tag individuals.

Mullins et al. (2010) initially screened 20 microsatellite loci for amplification with pine marten DNA samples. The loci screened included loci originally developed for amplification with other mustelid species: wolverine (*Gulo gulo*), otter, badger (*Meles meles*), stoat, European mink (*Mustela lutreola*) and American mink.

The hair-tube technique used by Mullins et al. (2010) successfully removed 114 hair samples from pine marten, 53 of these samples were selected for genotyping polymorphisms with the panel of eight microsatellite loci, and 94% of samples were amplified with all loci, and a low genotyping error (less than 2%) was associated with the samples. Full genotypes were obtained for 49 samples, which corresponded to seven different individuals, with the number of recaptures ranging from one to 18.

Fig 1.30: Hair-tube design used in Mullins et al. (2010) to collect pine marten hair non-invasively. The top inset is a hair-patch containing pine marten hair. The lower inset is a view into the hair-tube showing the placement of the sticky patch.

1.8.4.1 Other methods for the identification of individuals

SNP genotyping

Single nucleotide polymorphisms (SNPs) are the most abundant polymorphism in the genome as there is about one SNP for every 500 bp in many wildlife populations (Allendorf and Luikart 2007). An example of a SNP might be a G replaced by a C in different individuals at a particular position in the in a DNA sequence. Typically, these SNPs only have two alleles and thousands of them can be used to assess allele frequencies across populations. These alleleic patterns tend to cluster across populations, making them informative to infer population differentiation (Helyar et al. 2011). Microsatellites on the other hand can have a higher number of alleles, and therefore fewer loci are needed for population studies, but the segregation signal of SNPs can be very strong, meaning that fewer are needed (Heylar et al. 2011) (Fig 1.31). Allendorf and Luikart (2007) use the example of two humans who will differ from each other at up to several million SNPs, resulting in the creation of a differentiation signal. In 2007, Allendorf and Luikart predicted that SNPs would likely replace the use of microsatellite markers in conservation genetic studies, but to date this has still not happened.

Some of the challenges associated with the use of SNPs relate to the data analysis such as parentage, linkage and relatedness (Heylar et al. 2011). SNPs are useful for inferring population substructure, but it has been difficult to develop software that can cope with large data sets. SNP marker development is also associated with ascertainment bias. This is where SNPs are developed from a small number of samples representing a species, and typically these samples all come from one population, meaning the derived SNPs are not representative of the entire population. When the SNP panel is then applied to a wider population representative, there may be more SNPs present that are missed by the panel and thus result in an underestimation of the genetic diversity (Allendorf and Luikart 2007; Heylar et al. 2011).

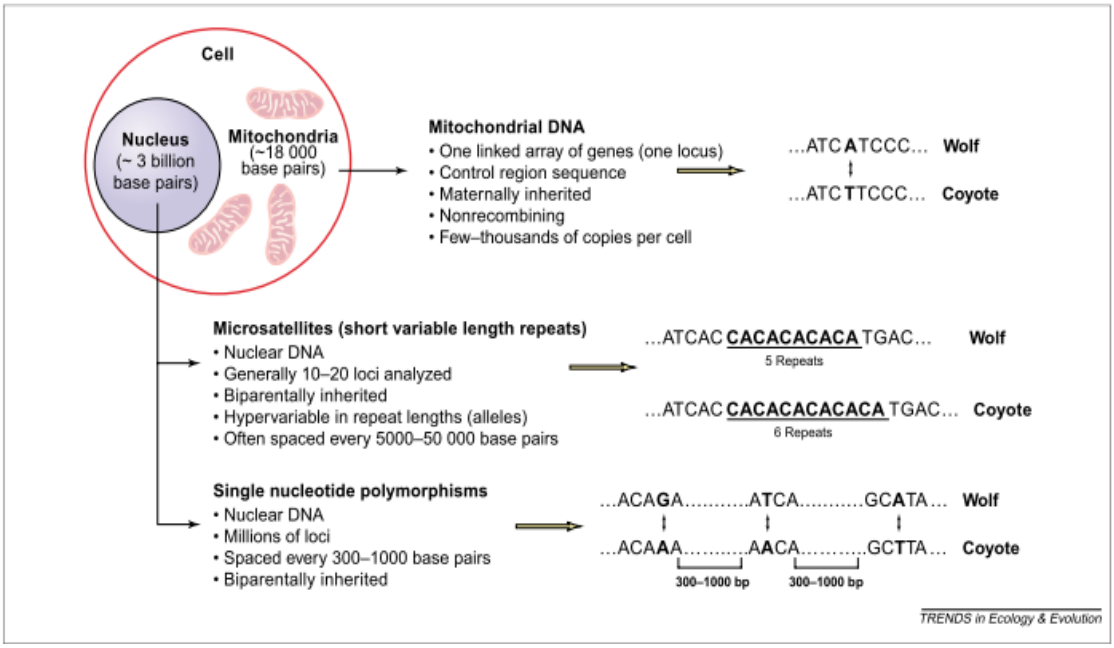

Fig 1.31: Comparison of the characteristics of mitochondrial DNA, microsatellites and single nucleotide polymorphism (SNPs) as genetic markers (Morin et al. 2004).

Morin et al. (2004) reviewed methods relating to the discovery of SNPs in non-model species (species other than mice or humans). The method relies on the amplification of sections of DNA sequences across multiple individuals and subsequently aligning the sequences and screening for polymorphisms for SNP design. Primers can be designed based on conserved segments of the genome between multiple species such as mouse and human to amplify orthologous gene regions. Around 50 SNPs are needed for population genetic studies, and SNPs occur approximately every 200 – 500 bp, meaning about 70 – 100 independent genome sections of 500 – 800 bp of sequence are needed. As it is sometimes difficult to amplify conserved primers in all species, an alternative approach relies on methods based on sequencing of amplified fragment length polymorphisms (AFLPs) (Morin et al. 2004). The process has been enhanced by the use of next generation sequencing technologies that has improved the speed and accuracy of SNP discovery (Helyar et al. 2011).

RAD-tag sequencing

As with the previously described RFLP (restriction fragment length polymorphisms), the introduction of restriction-site associated DNA (RAD) tags can be used to identify thousands of SNPs, but across a reduced representation of the genome of the

species, thus reducing the cost and time involved. The technique is also known as one of the methods that allow genotyping by sequencing (Baird et al. 2008; Heylar et al. 2011). The addition of RAD tags allow sequencing of nucleotides adjacent to restriction sites and subsequent detection of SNPs. The required number of markers can then be generated based on the choice of the restriction enzyme used, and additional restriction enzymes can allow for more SNPs to be detected. The method can allow pooled populations to be genotyped for bulk segregation analysis (to infer associations) and also multiplexing of individuals for fine-scale mapping (Baird et al. 2008).

Emerson et al. (2010) applied a RAD tag method to investigate the phylogeography of the pitcher plant mosquito (*Wyeomyia smithi*). The authors used RAD tag technology to isolate SNPs that are fixed within populations and variable among populations to determine the phylogeographic pattern of the species. Without any prior investment in genomic resources, Emerson et al. (2010) was able to resolve fine-scale genetic structure in a species. Similar results could have been achieved if the entire COI gene had been sequenced across all individuals. However, the use of RAD tags to achieve a similar result, produced over 3,000 SNPs across the nuclear genome (over 13,000 loci), all within two lanes of a single run of an Illumina sequencer. This required the use of only one restriction enzyme for identifying population structure among closely related populations, a technique that could be applied to any species, making it particularly useful for species that researchers have little genetic background information for.

1.8.4.2 Marker choice

Although, biases have been associated with SNP markers, similar challenges can also arise with the use of microsatellite markers. For instance the use of microsatellite markers developed for one species, but applied to another, can result in an underestimation of genetic diversity. Researchers also tend to select the most variable markers for population genetic studies, and this may inflate inferences of genetic diversity and differentiation between populations. More recent work has suggested that researchers need to carefully select their markers based on their

research question. Relatedness studies need the most variable markers, while studies assessing the genetic diversity of a population should select markers that include a range of allelic diversity (Queiros et al. 2013).

1.9 Outline of thesis

- The main objective of this project was to develop genetic methods that could be used to identify red squirrels and grey squirrels from non-invasive sources such as hair.
- The species identification assays were subsequently applied to a molecular dietary study, where both squirrels were genetically identified from pine marten scats, along with other small mammal species.
- DNA based methods for sex determinationwere investigated for use with the red squirrel and grey squirrel.
- Microsatellite markers were redesigned and optimised to identify individual red squirrels from low quantity DNA sources. The genetic toolbox was applied to a red squirrel hair-tube study in Co. Waterford.
- The red squirrel population in Co. Waterford was compared to other individuals in Ireland using microsatellite and mitochondrial DNA analysis to investigate levels of genetic diversity.
- The genetic techniques developed in this study were used to assess the genetic diversity of a remnant population of red squirrels in mid Wales.

Chapter 2

Development of species identification assays for the detection of red squirrel (*Sciurus vulgaris*) and the grey squirrel (*Sciurus carolinensis*)

Publications associated with this work

O'Meara DB, Turner PD, Coffey L, O'Reilly C (2012) TaqMan® assays for species identification of the red squirrel (*Sciurus vulgaris*) and the grey squirrel (*Sciurus carolinensis*). *Conservation Genetics Resources*, **4**, 603-604.

2.1 Introduction

In Section 1.8.1 methods for species identification were discussed. However, many of those methods rely on good quality DNA such as that found in tissue samples. The target DNA source in this study was non-invasive DNA obtained from low quality and quality DNA sources such as hair. Therefore, DNA methods needed to be developed that will accuractly and efficiently amplify low quantity and quality DNA, a method that has not previously been developed for squirrels.

The development of molecular techniques suitable for the species identification of red and grey squirrels is described in this chapter. The first method relied on the use of real-time PCR, employing both species-specific SYBR® Green 1 real-time PCR detection and the use of species-specific TaqMan® probe real-time PCR detection. The technique relies on the development of species-specific primers for both detection strategies and the addition of TaqMan® probes for the latter. The second method involved the development of a conventional PCR assay, a triple primer PCR assay consisting of one common forward primer and two reverse primers. PCR products can be DNA sequenced for species verification and can be used to resolve the mitochondrial DNA gene haplotype. Potential sources of squirrel DNA were examined including hair, feeding remains, and a small number of squirrel faecal samples.

2.2 Materials and Methods

2.2.1 Sample collection

For assay development, tissue samples from known species (red squirrel n = 5, grey squirrel n = 5) were initially used and tested to ensure that the assays were species-specific. A panel of 25 plucked hair samples from both species of squirrels was used to assess the applicability of the assays with use hair samples. To ensure that the assays were species specific, other potential non-invasive targets including rat (*Rattus norvegicus*), wood mouse (*Apodemus* sylvaticus) and pine marten were amplified with the primers sets. The assays were then applied to a hair-tube study conducted by E. Sheehy in the Irish Midlands (Co. Laois and Co. Offaly) to identify the species from 68 hair samples (Sheehy 2013).

In addition, maize kernels and chewed pine cones used to bait hair-tubes for red squirrels were collected and investigated for their potential as an alternative source of environmental DNA. Faecal samples from known red squirrels were also tested. Faecal samples were obtained from a live-trapping programme carried out by the Mammal Research Group, NUI, Galway.

2.2.2 DNA extraction

DNA was extracted from tissue or muscle using ZR Genomic DNA™ – Tissue Microprep (Zymo Research, cat. no. D3051) according to the Solid Tissue protocol with Zymo-Spin™ II columns (Zymo Research, cat. no. D3041).

Hair DNA extraction

DNA was extracted from plucked hair following a modified Chelex-100 protocol (Walsh et al. 1991; Statham 2005). Normally, 10 -20 hairs, but sometimes as few as three hairs were transferred into a microcentrifuge tube. All scissors and tweezers used in the extractions were alcohol flamed between samples to prevent cross contamination. Hair samples were digested for at least three hours at 56 °C in 200 µl of a digestion buffer containing 20 mg Proteinase K (AppliChem GmbH cat. no. A7932, 0500), 40 mM dithiothreitol (Molekula, cat. no. 578517) and 5% (w/v) Chelex-

100 resin (Bio-Rad, cat. no. 142-1253). Samples were incubated in a Biosan TS-100 shaking incubator. After digestion, the samples were heated to 100 °C for 8 minutes and centrifuged at 13,000 xg for 3 minutes. The supernatant was removed and stored at -20 °C.

Food remains

Genomic DNA was extracted from the surface of chewed pine cones by placing a cone in a universal container with 500 µl of a wash buffer solution (NaCl 5M, Tris 1M pH 7.8-8, EDTA 0.5 5M and SDS 20%). The container was rotated for one hour to wash any genomic DNA from the surface of the cone. The solution was centrifuged at 12,000 xg for 5 min. The supernatant was removed and purified using Zymo DNA Clean and Concentrator-25^{TM} kit (Zymo Research, cat. no. D40006) according to the manufacturer's instructions for the short protocol.

Genomic DNA was isolated from maize kernels by placing a kernel in a microcentrifuge tube with 200 µl of wash buffer (as described above) and rotated for one hour. The kernel was removed and solution was centrifuged at 12,000 xg to settle any debris. The supernatant was removed and purified using Zymo DNA Clean and Concentrator-25^{TM} kit (Zymo Research).

Faecal pellets

Approximately 0.2 g of squirrel faecal pellets was taken from each pellet using a disposable forceps. The sample was first placed in one ml STAR buffer (Stool Transport and Recovery buffer) (Roche, cat. no. 03335208001) and vortexed to mix. The mixture was then allowed to stand for 30 minuties. The solution was centrifuged for one minute at 10,000 xg. DNA was then isolated from 150 µl of the supernatant using the Zymo ZR Genomic DNA kit according to the Solid Tissue protocol with Zymo-Spin^{TM} II columns.

2.2.3 Primer Design

Design of real-time PCR assays

Real-time PCR assays for red and grey squirrel were designed to amplify short sections of mitochondrial DNA D-loop DNA (~ 100 bp products). Two assays were initially developed to target species-specific single nucleotide polymorphisms (SNPs) in the mitochondrial DNA control region. Species assignment was based on either melt-curve analysis (SYBR® Green 1) or TaqMan® probe fluorescence. The primers and probes were designed to target a short region of mitochondrial D-loop DNA using haplotypes AF111001–AF111027 (Barratt et al. 1999) and AM412650–AM412675 (Finnegan et al. 2008). Sequences were aligned using MegAlign 5.05 (DNASTAR) and conserved species-specific regions were identified and used to design primers and probes using Primer Express 2 software (Applied Biosystems), targeting species-specific nucleotide polymorphisms towards the 3' ends of the primers and in the middle of the TaqMan® probes to enhance the specificity and sensitivity of the amplification.

2.2.4 PCR

2.2.4.1 Real-time PCR

Real-time PCR protocol

PCR for the SYBR® Green 1 assays consisted of 5 µl 2 X Power SYBR® Green PCR Master Mix (Applied Biosystems, cat. no. 4364346), 0.2 µM of each primer and 1 µl DNA in a total volume of 10 µl. Negative controls contained molecular grade water instead of DNA. PCR for the TaqMan® probe assays consisted of 5 µl of TaqMan® MGB Universal PCR Mastermix (Applied Biosystems, cat. no. 4304437), 0.2 µM of each primer, 0.2 µM probe and 1 µl DNA in a total volume of 10 µl. Negative controls contained molecular grade water instead of DNA.

All PCR reactions were carried out in an ABI 7300 real-time PCR system with MicroAmp Optical 96-well reaction plates (Applied Biosystems, cat. no. N8010560), and sealed with MicroAmp Optical Adhesive Film (Applied Biosystems, cat. no.

4311971) and amplifications were performed using an Applied Biosystems, 7300 Real-Time PCR System (Applied Biosystems, cat. no. 4351103) with the following profile 2 minuties at 50 °C, 10 minutes at 95 °C and then 40 cycles of 15 s at 95 °C, 60 s at 60 °C. Fluorescence was recorded following each cycle and the Ct values were used to describe the reactions. The Ct value is the cycle at which the fluorescence passes a predetermined threshold level, representing the accumulation of amplified PCR product (Moran et al. 2008). A dissociation step of 15 s at 95 °C, 30 s at 60 °C and 15 s at 95 °C was added for melt curve analysis (SYBR® Green 1 only) to confirm specific amplification. The dissociation curve identified the melting temperature of the DNA (dissociation of the double strand of DNA). The dissociation step gradually increases the temperature of the plate, and as the amplicon melts, the decrease in the SYBR® Green 1 dye is recorded.

Real-time PCR Assay sensitivity and specificity

The sensitivity of the SYBR® Green 1 and TaqMan® MGB probe systems were tested on both red and grey squirrel tissue DNA prior to hair DNA . To test the sensitivity of the assay a 10-fold dilution series was performed. Standard curves were generated for each real-time PCR assay. The R^2 value and the gradient of the curve were calculated in Excel to assess the efficiency of the assays.

The species identification assays were tested against non-target species for cross-species reactivity. This was to ensure that the assays correctly identified the correct target. Each assay was tested with non-targets including non-target squirrel species, wood mouse (*Apodemus sylvaticus*), rat (*Rattus norvegicus*) and pine marten.

Real-time PCR assay – applications

The real-time PCR assays were tested on hair from known squirrels (trapped squirrels), and from a hair-tube study. DNA was extracted from maize kernels used to bait red squirrel hair-tubes and pine cones found in the same area and subjected to real-time PCR analysis to test for the presence of squirrel DNA. 1 µl of DNA extract was used in the PCR reactions. In the hair-tube study conducted by E. Sheehy the area also contained pine marten, and the non-squirrel DNA extracts were also tested

for the presence of pine marten DNA using the primers designed by Mullins et al. (2010): PM3F (5′-CTTGCCCCATGCATATAAGCA-3) and PM2-Rev (5′-GCCTGGTGATTAAGCTCGTGAT-3) and a species specific TaqMan® probe (5′-6-FAM-CGTGCACCTCACTTAG-3′), following the same real-time PCR protocol already described for squirrel TaqMan® probe assays.

2.2.4.2 Conventional PCR

Conventional PCR Protocol

PCR consisted of 5 µl of GoTaq® Hot Start Green Master Mix (Promega cat. no. M5123), 0.2 µM of each primer and 1 µl of DNA in a total volume of 10 µl. Negative controls contained water instead of DNA. PCR was carried out in a 2720 Thermal Cycler (Applied Biosystems). The PCR programme consisted of 95 °C for 5 minutes followed by 30 cycles of 94 °C for 30 s, 57 °C for 30 s, and 72 °C for 30 s followed by 5 minute final extension at 72 °C.

Conventional PCR Assay Specificity

PCR assays were initially tested on known tissue DNA prior to testing with DNA extracted from hairs. To test the primers for cross-species reactivity, red and grey squirrel samples along with DNA samples from rat, wood mouse and pine marten were used to test for cross-species reactivity of the primers (n = 5). PCR products were separated and visualised on an agarose gel stained with ethidium bromide (containing 0.25 µg/ml). The gel was prepared by dissolving 2 g of agarose (Sigma, cat. no. A9539) in 150 ml 1 X TAE buffer (40 mM Tris base, 20 mM glacial acetic acid, 1 mM EDTA, pH 8.0) at approximately 60 °C and 2 µl of ethidium bromide (Sigma, cat. no. E1510) was added for fragment visualisation under UV light using GeneSnap V6.10 image analysis system (SynGene).

DNA sequencing

A subset of the PCR products (n = 10 from both species) were purified using DNA Clean and Concentrator-5 Kit (ZYMO Research), used according to the manufacturer's instructions and sequenced in both directions with BigDye Terminator Cycle Sequencing Kit 3.1 (Applied Biosystems, part no. 4337454). The sequencing reactions were prepared in MicroAmp™ optical 8-tube strips (Applied Biosystems, cat. no. 4316567), and contained 1 µl of 2.5 x Ready Reaction Premix, 3.5 µl of 5 x BigDye sequencing buffer, 3.2 pmol of primer (SvSc_F), and 2 µl of purified PCR product and water in a total volume of 20 µl. The sealed tubes (caps: Applied Biosystems, part no. 4323032) were placed in a 2720 Thermal Cycler (Applied Biosystems). The extension reaction protocol was as follows: 96 °C for 1 minute and 25 cycles of 96 °C for 10 s, 50 °C for 5 s and 60 °C for 4 minute. An EDTA and ethanol precipitation clean up step was then used as follows: 5 µl of 125 mM EDTA was added to the bottom of each tube. This was followed by 60 µl of 100 % ethanol to each tube. The tubes were then resealed, inverted four times, and allowed to sit at room temperature for 15 minutes. The tubes were positioned into a plate adapter and centrifuged at 3000 xg at 4 °C for 30 minutes, and afterwards, the ethanol was removed by inverting the plate and centrifuging for less than 10 seconds to allow the ethanol to be collected on tissue paper. This was followed by the addition of 60 µl of 70% ethanol to each well, the plate was resealed and centrifuged 3000 xg at 4 °C for 30 minutes for 15 minutes. The plate was then inverted as before to remove the ethanol. The tubes were then allowed to air dry for 30 minutes at room temperature. The final step involved the addition of 20 µl of Hi-Di™ formamide (Applied Biosystems, cat. no. 4311320) to each reaction tube. The samples were denatured for 3 minutes at 96 °C and then slowly cooled to 4 °C.

Sequences were obtained by running the products on an ABI310™ automated DNA sequencer (Applied Biosystems), using a short capillary with POP-4™ Polymer (cat. no. 402838). Nucleotide sequences were analysed using the BLAST software at http://www.ncbi.nlm.nih.gov/BLAST/ from the GenBank (NCBI) database (Altschul et al. 1990).

2.3 Results

2.3.1 Real-time PCR

For the species identification of red squirrel DNA, a short section of mitochondrial D-loop sequence was targeted (74 bp) in a SYBR® Green 1 real-time PCR assay using the primers Red_F (5'-TGTGAGTATTAATGTGCATGCTT-3') and Red_R (5'-CATAGAACATATCATGTTTAATCAACA-3'). For TaqMan® probe real-time PCR assays, the same forward and reverse primers were used in conjunction with a fluorescently labelled oligonucleotide probe. The probe was designed to be short (17 bp), Red_P (5'-6FAM-AGCATGTGGTGGAGGTT-MGB-3') (Fig 2.1 [a]). For the identification of grey squirrel DNA, a section of mitochondrial D-loop DNA was also targeted (150 bp) using the primers Grey_F (5'-GGGGAATGTAAAATTGAAGGG-3') and Grey_R (5'-TTGGTCCAGTACAATAAATGTAAGAA-3') (Fig 2.1 [b]). A second reverse primer was later designed to reduce the overall amplicon to 94 bp by redesigning the reverse primer next to the probe Grey_R2 (5'-ACATTAATATTACATCCCCATGCA-3') (Fig 2.1 [c]). The probe Grey_P (5'-VIC- AGTATCTATGGACATGCTTATAT-3'), was labelled with a different label (Fig 1.2 [b, c]). All assays were designed to include as many SNPs as possible between each species. Primers were purchased from Eurofins MWG Operon and probes from Applied Biosystems.

Fig 2.1: Mitochondrial DNA D-loop region sequence alignment for TaqMan® probe and SYBR® Green 1 assays for red (a) and grey (b, c) squirrels. Sequence accession numbers used here are AF111011.1 and AF111027.1 (Barratt et al. 1999). Primer and probes are underlined and differences between the species are highlighted. The reverse primer of the grey squirrel assay was redesigned in (c). The reverse primer in (c) is positioned directly next to the probe and is differentiated from the probe in the bold text.

(a) Red squirrel real-time PCR assays

Red TGTGAGTATTAATGTGCATGCTTATAAGCATGTGGTGGAGGTTTTAATGTTGATTAAACATGATATGTTCTATG 409
Grey TATGAGTATCTATGGACATGCTTATATGCATGGGGATGTAATATTAATGTTGATTAAACATACTATGTACTATG 338

(b) Grey squirrel real-time PCR assays

Red TAGGGCTTTAAATGTGGAGGGTTAGATCCTAATGTTTTATGTACTATGTGAGTATTAATGTGCATGCTTATAAGCATGTGGTGGAGGTTTTAATGTTGATTAAACATGATATGTTCTATGTACTTTCTTATTATTATGTACTGTACCATTTAA 440
Grey GGGGGAATGTAAAATTGAAGGGAGGACTGTATGTTTTATGTACTATATGAGTATCTATGGACATGCTTATATGCATGGGGATGTAATATTAATGTTGATTAAACATACTATGTACTATGTACTTCTTACATTTATTGTACTGGACCATT-AA 365

(c) Reduced size grey squirrel real-time PCR assays

Red TAGGGCTTTAAATGTGGAGGGTTAGATCCTAATGTTTTATGTACTATGTGAGTATTAATGTGCATGCTTATAAGCATGTGGTGGAGGTTTTAATG 385
Grey GGGGGAATGTAAAATTGAAGGGAGGACTGTATGTTTTATGTACTATATGAGTATCTATGGACATGCTTATA**TGCATGGGGATGTAATATTAATG** 315

Real-time PCR Assay Sensitivity

The sensitivity of the two SYBR® Green 1 and TaqMan® probe systems were tested on both red and grey squirrel tissue DNA. To test the sensitivity of the assay a 10-fold dilution series was performed on tissue DNA (Fig 2.2). Standard curves were generated for each real-time PCR assay (Fig 2.3), by plotting the Ct value against the log of the DNA concentration. The R^2 value and the gradient of the curve were calculated in Excel to assess the efficiency of the assay. The R^2 values ranged from 0.98 – 0.99 for both the SYBR® Green 1 and TaqMan® assays indicating that the assays were 98-99% efficient. The higher the efficiency of the assays at optimal conditions, the more efficient the assays will perform at low target DNA. Applied Biosystems recommend an efficiency level between 90-110% for the assays to perform at an acceptable level.

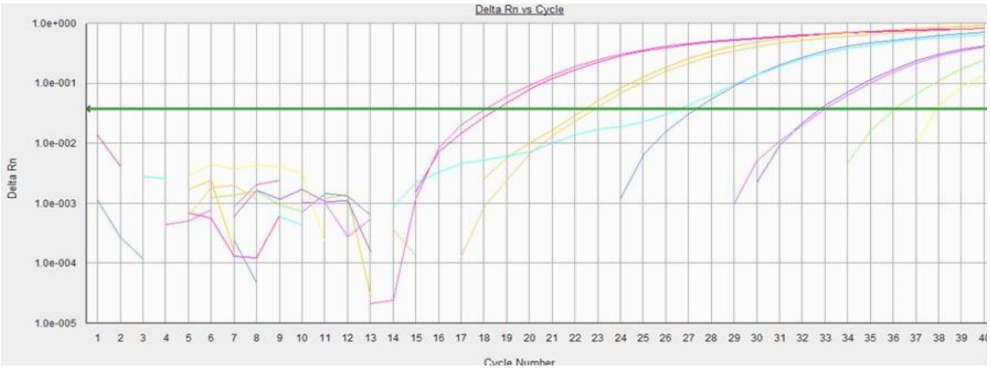

Fig 2.2: Example of amplification of a 10-fold dilution series for the red squirrel TaqMan® probe assay, starting at 4 ng of DNA.

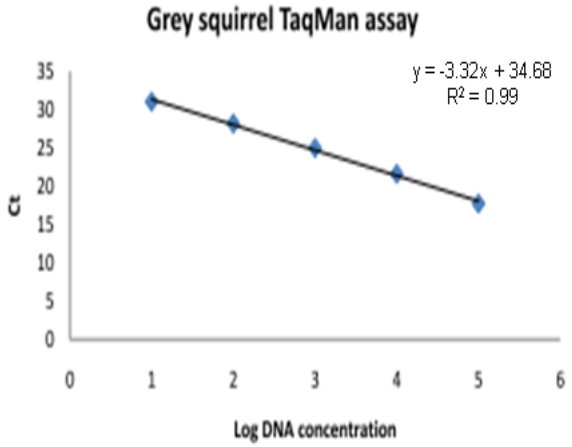

Fig 2.3: Example of a standard curve generated from a dilution series of grey squirrel DNA using the TaqMan® probe real-time PCR assay designed to target the species.

Real-time PCR assay specificity

Using red and grey squirrel tissue DNA samples with approximately 4-5 ng of DNA, the real-time PCR assays were found to be species-specific and there was no cross-species amplification. Ct values of the red and grey squirrel SYBR® Green 1 and TaqMan® probe real-time PCR assays are described in Table 2.1.

Table 2.1: Specificity of species-specific real-time PCR assays, showing Ct values of specific amplification determined by SYBR® Green 1 and TaqMan® probe chemistries. *U* = no fluorescence detected. No amplification was detected with wood mouse, brown rat or pine marten DNA.

	Species-specific Assay			
	Grey squirrel SYBR®	**Grey squirrel TaqMan®**	**Red squirrel SYBR®**	**Red squirrel TaqMan®**
Species	TM (°C) = 71.6 ±0.4		TM (°C) 71.9 ± 0.9	
Grey squirrel	16.97± 0.19	17.14 ± 0.05	U	U
Grey squirrel	16.33± 0.52	17.86 ± 0.43	U	U
Grey squirrel	16.89 ± 0.35	17.91 ± 0.43	U	U
Grey squirrel	16.35 ± 0.69	17.53 ± 0.09	U	U
Grey squirrel	16.0 ± 0.04	16.14 ± 0.67	U	U
Red squirrel	U	U	14.27 ± 0.14	14.435 ± 1.29
Red squirrel	U	U	15.27 ± 0.02	15.09 ± 0.63
Red squirrel	U	U	16.42 ± 0.30	16.81 ± 1.42
Red squirrel	U	U	17.48 ± 1.42	17.16 ± .87
Red squirrel	U	U	17.63 ± 0.28	17.54 ± 0.51
Wood mouse	U	U	U	U
Brown rat	U	U	U	U
Pine marten	U	U	U	U

Real-time PCR assay – applications

All of the assays successfully amplified DNA extracted from hair samples of known individuals using a panel of 25 hair samples from both species (Table 2.2). The samples contained DNA of varying quantity and quality, which is reflected across the range of Ct values. An example of a red squirrel dissociation curve and the corresponding amplification plot of a serially diluted hair sample are provided in Fig 2.4.

Table 2.2: A selection of known red and grey squirrel hair samples amplified with SYBR® Green 1 and TaqMan® probe assays.

RS hair sample	RS SYBR® Ct (mean ± SD) *TM (°C) 71.89 ± 0.18*	RS TaqMan® Ct (mean ± SD)	GS hair sample	GS SYBR® Ct (mean ± SD) *TM (°C) 71.6 ± 0.4*	GS TaqMan® Ct (mean ± SD)
RSH1	25.5 ± 0.71	26.14 ± 0.66	GSH1	20.65 ± 1.41	23.64 ± 0.26
RSH2	26.07 ± 0.54	25.56 ± 1.04	GSH2	21.85 ± 1.14	24.12 ± 1.29
RSH3	27.07 ± 0.67	30.87 ± 0.21	GSH3	22.12 ± 0.91	24.82 ± 1.41
RSH4	27.91 ± 0.89	28.5 ± 0.08	GSH4	22.34 ± 0.02	24.83 ± 0.74
RSH5	27.99 ± 0.43	27.82 ± 1.36	GSH5	23.07 ± 0.14	25.91 ± 1.67
RSH6	28.44 ± 0.28	26.91 ± 0.48	GSH6	23.61 ± 0.09	26.83 ± 0.67
RSH7	28.8 ± 0.54	28.15 ± 0.56	GSH7	24.13 ± 0.26	26.83 ± 0.71
RSH8	29.49 ± 0.71	31.04 ± 0.38	GSH8	24.22 ± 0.99	31.87 ± 1.81
RSH9	29.6 ± 0.47	31.55 ± 0.3	GSH9	24.88 ± 0.4	26.45 ± 0.61
RSH10	30.17 ± 0.13	28.99 ± 0.01	GSH10	25.18 ±0.65	28.18 ± 0.45
RSH11	30.55 ± 0.33	30.73 ± 0.13	GSH11	25.45 ± 0.7	27.36 ± 0.89
RSH12	30.64 ± 0.52	29.62 ± 0.21	GSH12	25.82 ± 0.7	28.45 ± 0.33
RSH13	30.8 ± 0.65	29.03 ± 0.11	GSH13	27.43 ± 0.30	25.63 ± 0.47
RSH14	30.97 ± 0.05	35.64 ± 1.42	GSH14	27.63 ± 0.22	29.45 ± 1.25
RSH15	31.05 ± 0.28	30.91 ± 1.75	GSH15	27.63 ± 0.81	28.46 ± 0.48
RSH16	31.76 ± 0.88	31.38 ± 0.16	GSH16	27.85 ± 0.26	29.96 ± 0.78
RSH17	31.81 ± 0.13	30.54 ± 0.04	GSH17	28.35 ± 0.18	29.45 ± 1.36
RSH18	32.26 ± 0.33	31.61 ± 0.19	GSH18	28.45 ± 1.15	30.33 ± 1
RSH19	32.45 ± 0.25	32.77 ± 2.08	GSH19	28.49 ± 1.19	31.45 ± 0.72
RSH20	33 ± 0.37	30.04 ± 0.25	GSH20	28.49 ±.019	31.56 ± 1.23
RSH21	33 ± 0.06	33.09 ± 0.43	GSH21	29.07 ± 0.07	32.79 ± 0.99
RSH22	33.02 ± 0.03	30.87 ± 0.21	GSH22	30.18 ± 3.2	33.46 ± 0.1
RSH23	33.26 ± 0.1	31.27 ± 0.26	GSH23	30.56 ± 0.86	34.78 ± 0.17
RSH24	33.3 ± 1.2	31.96 ± 0.53	GSH24	30.82 ± 1.03	31.48 ± 0.89
RSH25	34.16 ± 0.37	30.38 ± 0.13	GSH25	31.83 ± 0.55	33.45 ± 1.02

(a) (b)

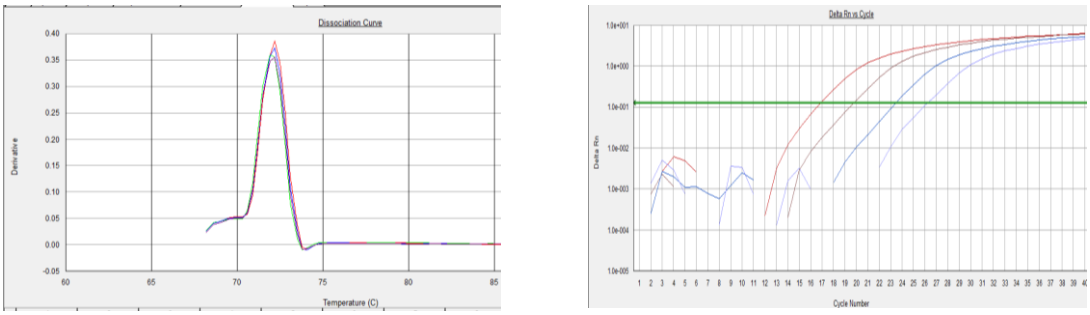

Fig 2.4: Dissociation and amplification curves for the SYBR® Green 1 real-time PCR assay specific for red squirrel. (a) Shows the dissociation curve of a hair DNA sample that was amplified across a 10-fold dilution. (b) The amplification curve for the red squirrel DNA extract from 10^{-1} – 10^{-4}.

Hair-tube Study

In total, 68 DNA samples were extracted from the hair collected by E. Sheehy in the Irish Midlands. Of the 68 samples, 20 samples were identified as red squirrel with the red squirrel TaqMan® probe assay, and no grey squirrel DNA was detected. As another species was likely present in some of the samples, the extracts were also tested for the presence of pine marten DNA (known to be present at the site), and 37 of the samples were subsequently identified as containing pine marten DNA. Two samples contained both red squirrel and pine marten DNA, and 13 samples were not identified to species (Table 2.3). The average red squirrel Ct values was 26.99 ± 2.5 and the average pine marten result was 26.67 ± 3.29.

Table 2.3: Species identification of samples collected from red squirrel hair-tubes in the Irish Midlands using TaqMan® probe real-time PCR assays for red squirrel (RS), grey squirrel (GS), pine marten (PM) and species (Sp).

Sample code:	RS Ct	GS Ct	PM Ct	Sp	Sample code:	RS Ct	GS Ct	PM Ct	Sp
ADSQ12	U	U	26.44	**PM**	ADSQ03	28.41	U	U	**RS**
ADSQ14	U	U	24.46	**PM**	ADSQ10	27.8	U	U	**RS**
ADSQ15	U	U	25.4	**PM**	BCSQ01	24.49	U	U	**RS**
ADSQ24	U	U	34.31	**PM**	BCSQ12	32.61	U	U	**RS**
ADSQ29	U	U	32.73	**PM**	BCSQ28	28.48	U	U	**RS**
BKPM05	U	U	25.18	**PM**	BCSQO2	24.32	U	U	**RS**
BKSQ02	U	U	25.71	**PM**	CVSQ06	29.06	U	U	**RS**
BKSQ05	U	U	24.29	**PM**	CVSQ09	25.45	U	U	**RS**
BKSQ10	U	U	31.8	**PM**	CVSQ10	27.34	U	U	**RS**
BKSQ27	U	U	24.76	**PM**	EMSQ12	23.76	U	U	**RS**
BKSQ28	U	U	25.72	**PM**	EMSQ13	27.58	U	U	**RS**
CVSQ03	U	U	28.73	**PM**	EMSQ14	28.56	U	U	**RS**
CVSQ05	U	U	25.16	**PM**	GHSQ02	26.74	U	U	**RS**
CVSQ07	U	U	28.43	**PM**	GHSQ04	25.59	U	U	**RS**
CVSQ11	U	U	27.93	**PM**	GHSQ08	26.43	U	U	**RS**
CVSQ13	U	U	24.19	**PM**	GHSQ14	26.28	U	U	**RS**
CVSQ14	U	U	23.52	**PM**	GHSQ18	26.61	U	U	**RS**
CVSQ16	U	U	32.51	**PM**	GHSQ20	26.36	U	U	**RS**
CVSQ18	U	U	32.57	**PM**	ADSQ13	U	U	U	**U**
CVSQ20	U	U	26.59	**PM**	ADSQ16	U	U	U	**U**
CVSQ22	U	U	25.05	**PM**	ADSQ17	U	U	U	**U**
CVSQ23	U	U	30.17	**PM**	ADSQ20	U	U	U	**U**
CVSQ24	U	U	26.77	**PM**	ADSQ22	U	U	U	**U**
EMSQ03	U	U	23.05	**PM**	BCSQ08	U	U	U	**U**
EMSQ04	U	U	24.48	**PM**	BKSQ08	U	U	U	**U**
EMSQ06	U	U	27.61	**PM**	BKSQ15	U	U	U	**U**
EMSQ08	U	U	22.96	**PM**	BKSQ20	U	U	U	**U**
EMSQ10	U	U	20.25	**PM**	BKSQ21	U	U	U	**U**
EMSQ15	U	U	28.89	**PM**	CVSQ04	U	U	U	**U**
EMSQ25	U	U	26.12	**PM**	EMSQ11	U	U	U	**U**
EMSQ28	U	U	23.97	**PM**	GHSQ24	U	U	U	**U**
EMSQ30	U	U	27.41	**PM**	ADSQ11	32.74	U	28.88	**RS/PM**
GHSQ03	U	U	27.99	**PM**	GHSQ06	31.14	U	26.4	**RS/PM**
GHSQ05	U	U	21.45	**PM**					
GHSQ13	U	U	26.73	**PM**					

Feeding remains

Maize bait

In total, 32 chewed maize samples were successfully extracted and amplified using the red squirrel SYBR® green 1 assay. The Ct values were high indicating low quantities of DNA. Lower Ct values were found for other non-invasive sampling methods such as hair DNA. The average Ct value for maize kernels was 36.53 ± 1.80. Ct values were high but the successful Ct values were consistent across replicates and no contamination was found in the no-template controls. The dissociation curve also indicated that the DNA was correctly amplified (Fig 2.5). However, only 21.88% of samples produced a Ct value less than 36 showing that the general quality and quantity of the DNA was very low.

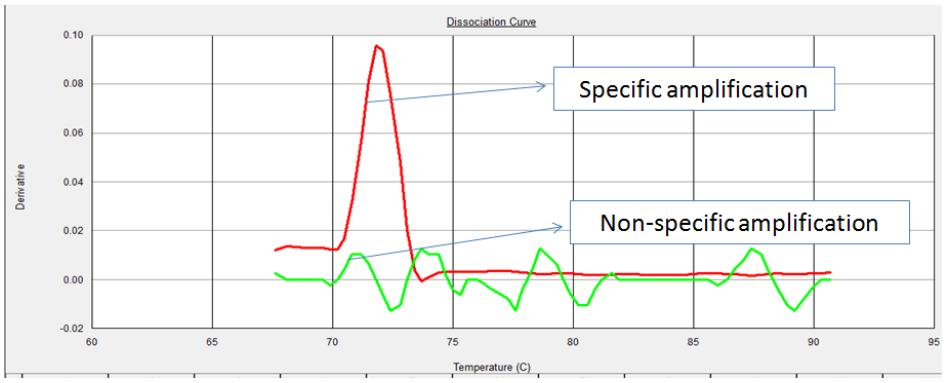

Fig 2.5: Dissociation curve for the red squirrel assay using DNA extraction from the surface of a maize kernel. The dissociation curve provides an example of specific and non-specific amplification.

DNA was extracted from the surface of 15 pine cones and produced mixed results using the red squirrel SYBR® Green 1 assay. The extracts that were successfully amplified had higher Ct values than other non-invasive sampling methods such as hair. Samples were amplified using a 10-fold dilution of the DNA extract. Inhibition was found to occur in many samples and replicates were not reliable, as Ct values were found to decrease with reduced concentration of DNA. Only four cones produced consistent results with Ct values ranging from 32.38 ± 1.37 at a dilution of 1/100 and 28.14 ± 1.89 at a dilution of 1/1000. No contamination occurred in the no-template controls. Pine cones produced low quantity and quality DNA, with

unreliable reproducibility. This may have been caused by extraction of the pine cone resin, which was co-extracted with the DNA and led to inhibition in the real-time PCR reaction.

Squirrel faecal samples

Squirrel faecal samples were obtained during a squirrel live trapping monitoring programme and represent an alternative method to hair samples for genetic analysis. DNA was extracted from five samples and were successfully amplified with the red squirrel assays and worked best using a dilution factor 1/100. Ct values ranged from 27.45 ± 1.41; Tm°C 71.85 ± 0.21 to 32.67 ± 1; Tm°C 71.4 ± 0.0.

2.3.3 Conventional PCR assay

Design of conventional PCR assays

PCR primers were chosen to amplify mitochondrial D-loop products in both red and grey squirrels. The forward primer was designed to amplify both red and grey squirrel SvSc_F (5' TTCACGGAGGTAGGTAGATTAAGA-3'), while species specific reverse primers were chosen for both species. The red squirrel reverse primer Svul_R (5' TTGATGTCTATGTAATTCGTGCAT-3') was designed to target 350 bp of sequence, and the grey reverse primer Scar_R (5'-TGGTCCAGTACAATAAATGTAAGAAG) was designed to target 300 bp of sequence. A common primer was chosen as this region of the sequence was similar between both species, and the region where the reverse was chosen was more variable. The primer Scar_R, is a one base pair shift to the reverse primer (Grey_R) which was redesigned to match the melting temperature of the common primer in this assay (SvSc_F). Available sequences from the Genbank database were aligned and the primers and mismatches between the species were highlighted, where the sequences overlapped. The technique successfully distinguished between both species and produced clear species-specific bands using gel electrophoresis (Fig 2.6). The assay was shown to be template specific (when tested with both tissue and hair DNA) and did not cross amplify with the other species including rat, wood mouse or pine marten. The assays were capable of amplifying DNA from both tissue and hair

samples. The DNA sequence analysis validated that the assay correctly identified the correct species.

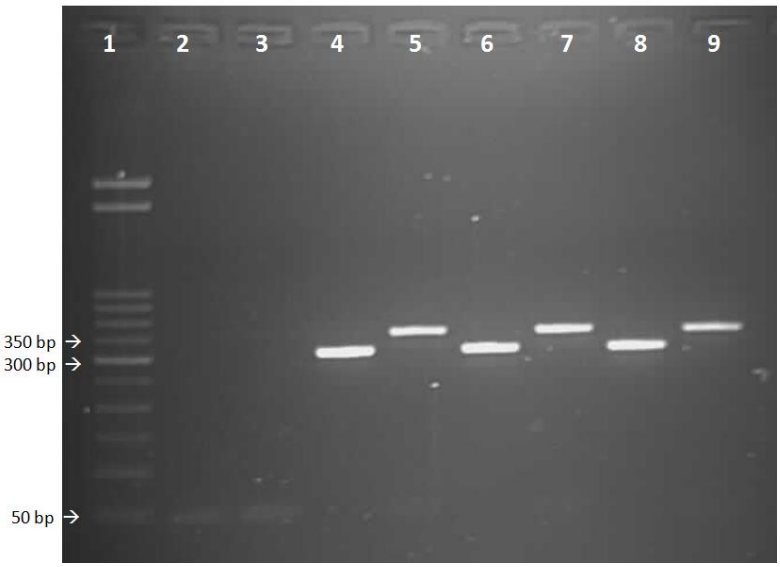

Fig 2.6: **Gel image of red and grey squirrel amplification illustrating the 50 bp size variation between the grey squirrel product (300 bp) and red squirrel (350 bp). *Lane 1* contains a 50 bp size standard. *Lanes 2* and *3* contain negative controls and *lanes 4, 6, 8* contain grey and red squirrel PCR products are in *lanes 5, 7* and *9*.**

2.4 Discussion

Real-Time PCR assays developed in this study successfully distinguished between red and grey squirrels using DNA extracted from tissue and hair, and were shown to be species-specific. The assays have successfully been tested on samples collected throughout Britain and Ireland, and were shown to amplify DNA from hair samples from throughout the region (O'Meara et al. 2012). The red squirrel assays were also successfully applied to other sources of red squirrel DNA including feeding remains and faeces.

The hair-tube study conducted by Sheehy (2013) in the Irish midlands showed that in areas where pine marten were plentiful (Sheehy et al. 2014), there was an increased possibility of pine marten entering the hair-tubes targeting squirrels. However, the combination of the red squirrel real-time PCR assay (this study) and the pine marten assay designed by Mullins et al. (2010) successfully differentiated the species in this situation.

The DNA extracted from feeding remains provided mixed results for monitoring the presence of red squirrels. DNA extracted from maize bait provided more reliable results than DNA extracted from chewed/stripped pine cones. As mentioned, resins were extracted with the DNA in the case of the pine cones, which could not be avoided during the DNA extraction process. Squirrel DNA extracted from maize bait may be a more promising area for further research into the uses of environmental DNA to survey squirrels, as less PCR inhibition was encountered. It is also difficult to tell how long a chewed pine cone had been exposed to environmental elements, which could lower the quality and quantity of squirrel DNA present. New improvements in the area of DNA extraction kits such as the DNeasy® mericon Food Kit (QIAGEN), designed to extract DNA from ingredients used in processed food could be targeted for future research. Nichols et al. (2013) successfully used this kit to extract ungulate DNA from browsed twigs, and such kits could be worth investigating for their application to squirrel feeding remains. In this study, we were interested in obtaining good quality DNA extracts that could be later used to identify individual

squirrels, and as better results were obtained from hair samples, DNA extracted from feeding remains was not pursued any further in this study.

The advantages of this real-time PCR including precision, sensitivity and specificity make the technique particularly useful for identifying samples from non-invasive sources (O'Neill et al. 2013). In this study, the technique was shown to be particularly useful for the identification of squirrels using DNA samples extracted from hair, faecal and feeding remains. The Ct values obtained from the species identification assays confirm the species and also provide an assessment of the quality and quantity of the target DNA extracted from the sample. The results showed that hair and faecal samples provided the best source of DNA, but it is very difficult to collect faecal samples of squirrels. This method would also be useful in identifying the species prior to genotyping and subsequent genetic population analysis of non-invasively collected material.

The conventional PCR assays amplified hair and tissue DNA well which makes this method useful for species identification and DNA sequencing of hair-tube samples. Furthermore, the area of the mitochondrial D-loop used for primer design in this study has also been used in a number of genetic studies of the red squirrel (Barratt et al. 1999; Finnegan et al. 2008). The region also contains a number of geographically specific haplotypes making this assay particularly informative for use in population analysis for both red and grey squirrels. While, there are currently primers available for the identification of haplotypes of red squirrels, the assays described in this study, amplify a shorter amplicon (350 bp) which is useful for non-invasive studies where low quantity and quality DNA is typically used.

2.5 Conclusion

- Real-time PCR assays were designed for the species identification of red squirrels and grey squirrels
- All assays were shown to be species-specific and only amplifed the target species

- Assays were successfully applied to sources of non-invasive squirrel DNA: hair samples, feeding remains and faecal samples
- Conventional PCR assays were designed to target both red and grey squirrels. These primers are suitable for generating PCR products for downstream sequence analysis

Chapter 3

The non-invasive identification of mammalian DNA in pine marten (*Martes martes*) Scats

Publications associated with this work

O'Meara DB, Turner PD, Coffey L, O'Reilly C (2012) TaqMan® assays for species identification of the red squirrel (*Sciurus vulgaris*) and the grey squirrel (*Sciurus carolinensis*). *Conservation Genetics Resources*, **4**, 603-604.

O'Meara DB, Sheehy E, Turner PD, O'Mahony D, Harrington AP, Denman H, Lawton C, MacPherson J, O'Reilly C (2014) Non-invasive multi-species monitoring: real-time PCR detection of small mammal and squirrel prey DNA in pine marten (*Martes martes*) scats. *Acta Theriologica*, **59**, 111 – 117.

Sheehy E, O'Meara DB, O'Reilly C, Smart A, Lawton C (2014) A non-invasive approach to determining pine marten abundance and predation. *European Journal of Wildlife Research*, **60**, 223 -236.

3.1 Introduction

Hard-part analysis of a large number of predator scats over a large scale to detect prey items may not be feasible taking constraints such as labour, time, specialist training and overall cost into consideration. Molecular dietary analysis has emerged as a reliable tool to identify prey items in the diet over the last decade (Deagle et al. 2005; Murray et al. 2011; Shehzad et al. 2012). A variety of molecular methods have been successfully used to identify prey items in the diet (Section 1.8.2), and real-time PCR has been shown to be a reliable method to detect prey items when background information about the predators' diet is known (Murray et al. 2011).

Over the last 100 years three mammalian species have been introduced into Ireland; the North American grey squirrel, the bank vole (*Myodes glareolus*) and the greater white-toothed shrew (*Crocidura russula*). The grey squirrel has competed with the indigenous red squirrel population, and the combined presence of the bank vole and greater white-toothed shrew has been shown to affect the abundance of the indigenous wood mouse and pygmy shrew (*Sorex minutus*) (Montgomery et al. 2012). These species may all be connected in the diet of the pine marten, and molecular analysis of scats collected across a wide distribution area could be a suitable method to assess for the presence of all species.

In this study, the first aim was to investigate if the real-time PCR technique developed for species identification of red squirrels and grey squirrels (Section 2.3.1) could be used to detect squirrel DNA in pine marten scats. Secondly, small mammal data generated from pine marten scats was also included in the analysis of this study to inform the distribution of small mammals and to compare the results to the current known distribution of the species. The third aim was to assess how the results obtained from this approach compared to previous dietary studies of Irish pine marten.

3.2 Materials and methods

3.2.1 Sample collection

Feeding trial

The feeding trial took place in Carmarthenshire, Wales. Grey squirrels were fed to a captive pine marten on 11 days over a 28 day feeding trial. Other dietary items included chicken, fruit and nuts. A total of 23 scats were collected during the feeding trial. The feeding of the pine marten, scat collection and data recording was conducted by H. Denman and J. MacPherson. Scats were collected and stored at -20 °C in sealed polythene bags to prevent cross contamination.

Field samples

Two hundred and seventy one samples were collected between March 2010 and October 2011 in the Irish Midlands, Laois and Offaly (Sheehy 2013), with 209 of those samples genetically identified as pine marten, and 84 samples from Portlaw Co. Waterford in August 2010 (P. Turner). The presence of red squirrels and the absence of grey squirrels were known at both these sites. One hundred and sixty eight samples were collected across Ireland between June and September 2005-2007 as part of the National Pine Marten Survey (NPMS) (O'Mahony et al. 2012). Background information on squirrel distribution at woodland level was not available at these sites. Scats were collected and stored as before. The field collection of samples involved the pooling of samples from three different sample collections to test for the presence of red and grey squirrel DNA (Table 3.1).

Table 3.1: Genetically identified pine marten scats used in this study.

Site	Collected by	No. of scats
Midlands	Sheehy (2013)	209
Waterford	P. Turner	84
NPMS	O'Mahony et al. (2012)	168
		Total: 461

3.2.2 DNA analysis

DNA had previously been extracted from scats from O'Mahony et al. (2012), and DNA was extracted from the Portlaw, Co. Waterford scats by C. O'Reilly. DNA was extracted from the scats collected by E. Sheehy following the protocol described for faecal DNA extraction from squirrel faecal pellets (Section 2.2.2).

Pine marten DNA extracts were identified by targeting a species-specific fragment of the mitochondrial DNA control region using a pine marten specific 5' nuclease assay (Mullins et al. 2010) (Table 3.2). Samples from the Midlands that failed to be detected as pine marten were also tested for fox DNA using a specific 5' nuclease assay (Mullins et al. 2010; Table 3.2). All pine marten samples (Table 3.1) were subsequently tested with species-specific real-time PCR assays designed to detect red and grey squirrels (Section 2.2.4.1; Table 3.2). A subset of samples (n = 252), were tested for the presence of small mammal DNA (bank vole, pygmy shrew and wood mouse) using real-time PCR assays targeting small mammal DNA (Moran et al. 2008; Table 3.2). This subset of samples which included the samples from O'Mahony et al. (2012) and the samples from Portlaw Co. Waterford was chosen as it represented the widest geographical area. This analysis was conducted by C. O'Reilly. Real-time PCR primers for the greater white-toothed shrew were designed by C. O'Reilly to target a short region of mitochondrial CytB gene using haplotypes AY918341-AY918400 (Brändli et al. 2005).

Real-time PCR with the SYBR® Green 1 and TaqMan® MBG Probe systems were as described in Section 2.2.4.1, except the SYBR® Green 1 Master Ready Mix (Applied Biosystems) was replaced with FastStart Universal SYBR® Green Master (Rox) (Roche, cat. no. 04913914001) and TaqMan® Universal PCR Master Mix (Applied Biosystems) was replaced with FastStart Universal Probe Master (Rox) (Roche, cat. no. 04914058001).

As the greater-white toothed shrew assay had not been previously published and confirmed, a 482 bp fragment of the CytB gene was amplified using PCR primers

CrocF482 (5'-CGCTTCTTCGCATTTCACTTT-3') and CrocR (Table 3.2), (designed by C. O'Reilly). The PCR and conditions were as described for the conventional PCR in Section 2.2.4.2. PCR products were resolved, cleaned and DNA sequence analysis was performed as in Section 2.2.4.2.

Table 3.2: Details of real-time PCR primer and probes used in this study.

Target species	Assay type	Forward and Reverse Primers (5'-3')	Probe (for TaqMan® assays)	Source
Pine marten	TaqMan®	PM3F: CTTGCCCCATGCATATAAGCA PMREV2: GCCTGGTGATTAAGCTCGTGAT	6-FAM: CGTGCACCTCACTTAG	Mullins et al. (2010)
Fox	TaqMan®	PM-FOR: CACCATGCCTGCATGCCTCGAGAAACCAT PM-REV: GGCCCGGAGCGAGCGAGAAGAGG	VIC: CTTGCTCGAAGTAT	O'Reilly et al. (2008)
Red squirrel	TaqMan®	RedF: TGTGAGTATTAATGTGCATGCTT RedR: CATAGAACATATCATGTTTAATCAACA	6-FAM: AGCATGTGGTGGAGGTT	Chapter2; O'Meara et al. (2012)
Grey squirrel	SYBR® Green	GreyF: GGGGAATGTAAAATTGAAGGG GreyR: TTGGTCCAGTACAATAAATGTAAGAA	VIC: AGTATCTATGGACATGCTTATAT	Chapter 2; O'Meara et al. (2012)
Bank vole	SYBR® Green	CG(95): ACTTTGGCTCCCTACTTGGCCTT CG(266): ATGGAGGCTCCGTTTGCATGCA		Moran et al. (2008)
Greater white-toothed shrew	SYBR® Green	CrocF: ATAAGCCAATGCATATTCTGAATTTTAG CrocR: CATGTTAATGTAAAGAGGTCCGCTAC		Designed by COR, Published in O'Meara et al. (2013)
Pygmy shrew	SYBR® Green	SM(241): TGAGGCGCAACAGTTATTACCA SM(544): GGAATGCGAAGAATCGAGTGA		Moran et al. (2008)
Wood mouse	SYBR® Green	AS(96): CTTCGGCTCATTGCTAGGAAT AS(264): TGAGGCTCCGTTTGCATGTA		Moran et al. (2008)

3.3 Results

3.3.1 Feeding trial

The feeding routine, scat collection and detection of pine marten and grey squirrel DNA is provided in Table 3.3. The 23 pine marten scats collected from the feeding trial all tested positive for pine marten DNA with Ct values ranging from 22.48 - 28.15. Twelve of those samples tested positive for grey squirrel yielding Ct values ranging from 29.38 - 36.79 (Table 3.3). All the positive grey squirrel results were replicated for verification. There was a positive correlation between the number of scats collected within one day of feeding with grey squirrel, and that scat containing grey squirrel DNA ($r = 0.78$, 14 d.f., $P = < 0.001$).

Table 3.3: Feeding trial with a captive pine marten over 28 days. The diet consisted of grey squirrel, chicken, fruit and nuts. Pine marten and grey squirrel DNA was detected in scats using real-time PCR.

Day	Diet	Scat collected	Pine marten DNA detected	Grey squirrel detected
1	Grey squirrel			
2	Chicken			
3	Chicken			
4	Chicken			
5	Grey squirrel			
6	Grey squirrel			
7	Grey squirrel	S1, S2	S1, S2	S1, S2
8	Grey squirrel	S3	S3	
9	Grey squirrel			
10	Fruit & nuts	S4	S4	S4
11	Grey squirrel	S5	S5	S5
12	Fruit & nuts	S6, S7, S8	S6, S7, S8	S8
13	Grey squirrel			
14	Grey squirrel	S9, S10	S9, S10	S9
15	Grey squirrel	S11, S12	S11, S12	S11, S12
16	Fruit & nuts	S13, S14, S15	S13, S14, S15	S13, S14, S15
17	Grey squirrel			
18	Fruit & nuts	S16	S16	S16
19	Fruit & nuts			
20	Fruit & nuts	S17	S17	
21	Fruit & nuts			
22	Chicken			
23	Chicken			
24	Chicken			
25	Chicken	S18	S18	
26	Chicken	S19, S20	S19, S20	
27	Chicken	S21, S22	S21, S22	
28	Chicken	S23	S23	

Table 3.4: Scat samples collected from pine marten feeding trial with the resulting pine marten and grey squirrel Ct values. U = unidentified.

Scat sample	Pine marten Ct	Grey squirrel Ct
S1	25.28	29.38
S2	28.15	36.52
S3	25.86	U
S4	24.02	32.64
S5	26.03	32.84
S6	24.59	U
S7	27.81	U
S8	27.69	35.78
S9	25.25	36.99
S10	23.38	U
S11	25.41	36.7
S12	22.48	34.56
S13	26.17	36.34
S14	23.38	35.56
S15	23.7	36.79
S16	27.9	35.48
S17	24.09	U
S18	25.49	U
S19	25.86	U
S20	23.45	U
S21	25.46	U
S22	27.48	U
S23	26.57	U

3.3.2 Field Study

3.3.2.1 DNA identification of scats

In the field study, scat samples collected at all sites were firstly tested for the presence of pine marten DNA using real-time PCR and samples with a Ct value < 33 were deemed positive for pine marten and to have an adequate quantity of DNA for subsequent analysis (Fig 3.1). In total, DNA was extracted from 271 samples; 209 samples were identified as pine marten and 26 were identified as fox. The Ct values corresponding to fox identification ranged from 13 – 34. A table of fox Ct values is provided in Table 3.5.

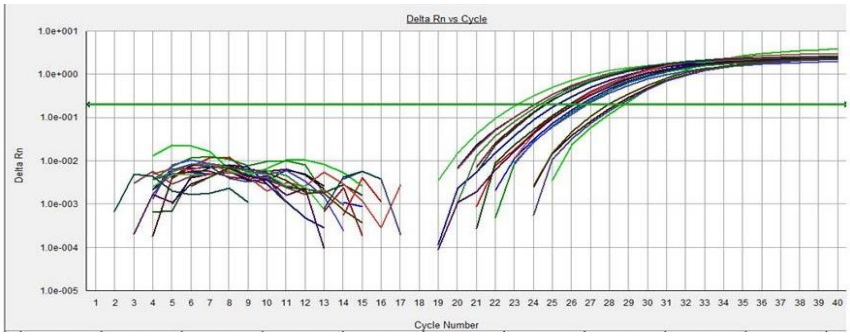

Fig 3.1: Example of the real-time PCR amplification of pine marten DNA

Table 3.5 Ct values of fox samples collected from the Midlands.

Scat Extract (Sheehy 2013)	Fox Ct	Scat Extract (Sheehy 2013)	Fox Ct
SCT219	13.18	SCT185	30
SCT224	17.42	SCT047	30.19
SCT203	21.53	SCT018	30.56
SCT215	23.15	SCT013	30.73
SCT022	25.09	SCT037	30.9
SCT257	25.45	SCT114	30.97
SCT073	26.71	SCT164	32.41
SCT214	27.11	SCT017	33.09
SCT186	27.32	SCT015	33.55
SCT191	27.53	SCT204	33.93
SCT223	28.67	SCT120	33.95
SCT077	29.34	SCT181	33.95
SCT033	29.65	SCT072	34.04

3.3.2.2 Squirrel DNA detection

Red squirrel DNA was detected in samples from all sites: Midlands (n = 7), Waterford (n = 1) and the NPMS (n = 1). Ct values ranged from 21-34. Grey squirrel DNA was not detected in any of the field collected samples. The corresponding Ct values for both pine marten and red squirrel detection from the Midlands site are displayed in Table 3.6, as both the pine marten and red squirrel real-time PCR was conducted for this study. Samples that tested positive for red squirrel DNA were replicated four times for verification. To assess if there was a relationship between the pine marten Ct value and the Ct values of red squirrel and grey squirrel detected in the pine marten

scats, a non-parametric Spearman's rank correlation was conducted using the data in Tables 3.4 and 3.6. A correlation was found between the increasing pine marten Ct value and the squirrel Ct value ($r = 0.52$, 17 d.f., $P = 0.022$), suggesting that if the pine marten DNA quality was of poor quality, resulting in a high Ct value, the detection of good quantity and quality prey DNA would be less likely. An example of the replicated detection of red squirrel DNA in a pine marten scat (SCT089) is displayed in Fig 3.2.

Table 3.6: Detection of pine marten and red squirrel DNA from pine marten scats collected from the Midlands, and the corresponding Ct values.

Scat Sample (Sheehy 2013)	Pine marten Ct	Red squirrel Ct
SCT059	24.58	31.33
SCT089	20.25	21.49
SCT138	19.04	22.03
SCT161	20.31	23.44
SCT233	24.73	28.49
SCT241	18.43	31.01
SCT265	32.47	34.26

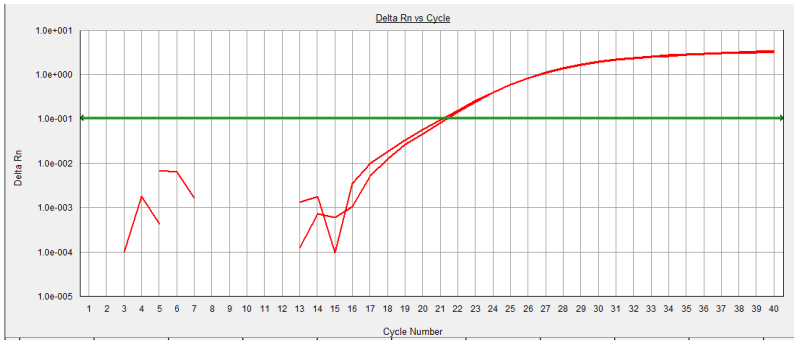

Fig 3.2: Real-time PCR detection of red squirrel DNA in DNA extracted from a pine marten scat (SCT089).

3.3.2.3 Small Mammal DNA Detection

Scat samples from Waterford and the NPMS were tested for the presence of small mammal DNA. Positive Ct values for the small mammal assays ranged from wood

mouse; 22-37, pygmy shrew; 22-36 and bank vole; 22-34 (Fig 3.3). The melt curve analysis revealed species-specific amplification of these samples and the Tm values were within the described range by Moran et al. (2008). The greater white-toothed shrew assay which was newly designed by C. O'Reilly was found to be species-specific (i.e it did not amplify with pine marten DNA), and did not cross amplify non-target DNA including the pygmy shrew, wood mouse, brown rat and bank vole. The Ct values ranged from 22-34, and the corresponding Tm value averaged 74 ± 0.2 °C.

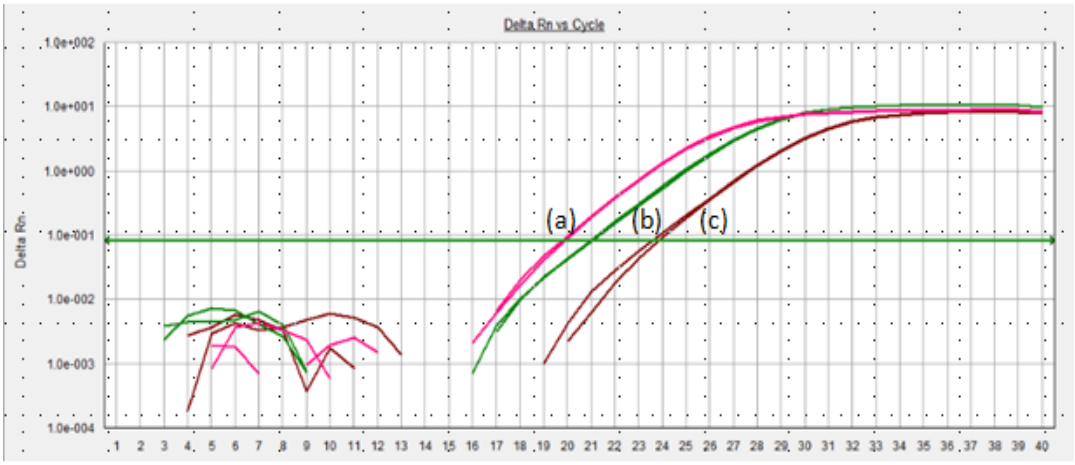

Fig 3.3: Real-time PCR detection of (a) pygmy shrew, (b) bank vole and (c) wood mouse from pine marten scats. Examples shown are replicated data.

Frequency of small mammal detection

Small mammal DNA was detected in the diet of pine marten. Wood mouse was found to be the most common small mammal occurring in 14.7% of the scats sampled. This was followed by pygmy shrew (13.9%), bank vole (3.2%), greater white-toothed shrew (2.4%), and red squirrel (2.1%). Grey squirrel did not occur in the field collected samples. When the results were mapped, wood mouse was found in 25 of the 10-km^2 sites, pygmy shrew in 26 sites, bank vole in five sites, and greater white-toothed shrew in six sites (Fig 3.4). The squirrel data was not mapped due to the low frequency detection, and most of the positive samples occurred in the Midlands where intensive scat collections took place in sites where pine marten and red squirrels were known to occur (Sheehy et al. 2014).

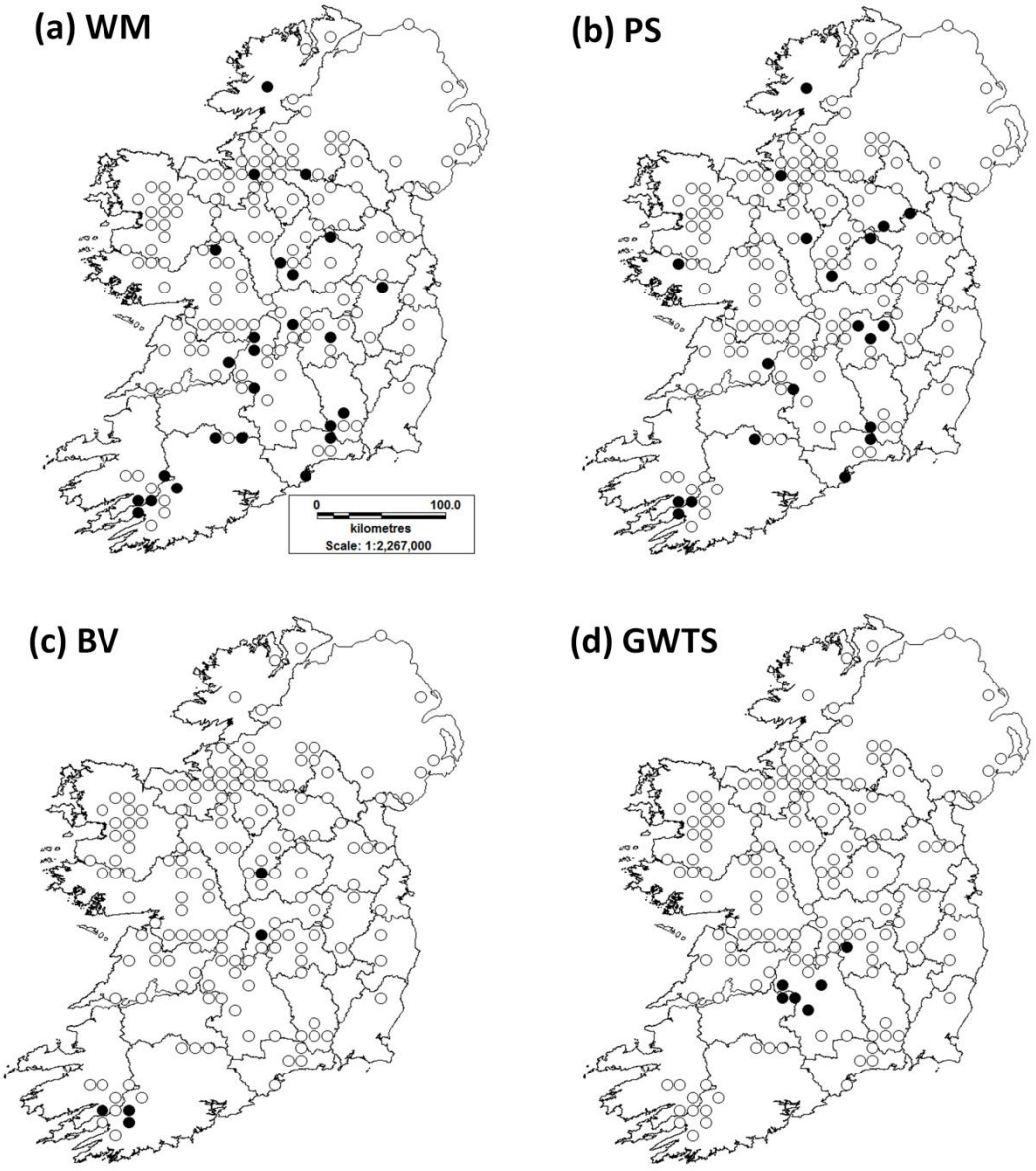

Fig 3.4: Distribution of small mammals (a) *WM* wood mouse, (b) *PS* pygmy shrew, (c) *BV* bank vole and (d) *GWTS* greater white-toothed shrew (O'Meara et al. 2014). Dark circles represent positive sites and white circles represent negative sites.

3.4 Discussion

In this study, a protocol was developed for the detection of squirrels and small mammal DNA in the pine marten diet. This was achieved using a panel of species-specific real-time PCR assays. This was the first study to use a molecular appraoch to test DNA extracted from pine marten scats to test for the presence of target species in the pine marten diet. The method was firstly tested using a feeding trial and then applied to field samples to confirm the utility of the approach. The results showed that low quality DNA scat samples were also likely to yield low quality prey DNA, which should be taken into consideration for future studies.

3.4.1 DNA identification of scats

From the Midlands site, 86.71% of scats were identified to species, with a total of 9.59% of the putative pine marten scats identified as fox. Combining the fox samples with the samples that failed to be identified as either fox or pine marten provided a total of 22.89%. If dietary analysis had been conducted on the entire data set, this could have potentially introduced a large error into the data. However, it is possible that some of the samples that failed to be identified as pine marten or fox simply contained poor quantity and quality DNA that could not be identified to species. Previous pine marten dietary studies conducted in Ireland by Lynch and McCann (2007) and Warner and O'Sullivan (1982) both identified scats by relying on the smell only, a method even expert field pine marten surveyors agree can be misleading (Ruiz-González et al. 2007; Balestrieri et al. 2011; Caryl et al. 2012).

3.4.2 Squirrel detection

The grey squirrel feeding trial demonstrated that grey squirrel DNA could be detected in the diet of the predator. Although grey squirrel DNA was detected in only 52% of the overall scats collected, grey squirrel DNA was detected in 75% of the scats collected within one day of feeding of grey squirrel. The higher Ct value, in comparison to the host DNA also represents the lower quantity and quality of DNA found in the prey DNA, which has been observed in similar studies (Matejusová et al. 2008). There was a positive correlation between the quantity of squirrel DNA

detected in scats and the DNA quantity of pine marten DNA, and this means that poor quality pine marten DNA samples are unlikely to yield good quality prey DNA.

There was no evidence of grey squirrel DNA in any of the field survey scats collected, despite some potential for overlap between pine marten and grey squirrel in the sample collection from O'Mahony et al. (2012). This shows that the grey squirrel does not frequently occur in the pine marten diet. However, the technique developed in this study was subsequently applied to a molecular dietary analysis of pine marten scats collected in the east of Ireland, and grey squirrel DNA was detected in approximately 10% of the scats tested (Sheehy et al. 2014).

The occurrence of red squirrel at just over 2% in scats was higher than previous pine marten hard part analysis of scats (Warner and O'Sullivan 1982; Halliwell 1997; Lynch and McCann 2007). Lynch and McCann (2007) found that red squirrel was found in less than 0.5% of scats. Interestingly, Warner and O'Sullivan (1982) found that the consumption of red squirrel by pine marten in Ireland was just over 1%, but red squirrel was only detected in winter months from November to February. Warner and O'Sullivan (1982) collected 125 scats over the winter period where red squirrel occurred in 6% of scats. Interestingly, Sheehy et al. (2014) found that squirrel consumption increased in early summer. Caryl et al. (2012) found no evidence of red squirrel predation by pine marten in Scotland, despite having four times the sample size and having personal observations of red squirrel in the survey site. In that particular site, field vole accounted for 30% of the pine marten diet. Typically, red squirrel predation by pine marten appeared to be only common in Northern Russia and Sweden, when other prey items were limited (De Marinis and Masetti 1995). It may be possible that the consumption of squirrels in the pine marten diet in Ireland is a little higher than in Britain due to the absence of field voles in Ireland, but further research is needed.

3.4.3 Small mammal detection

The bank vole and greater white-toothed shrew were recorded within the described distribution of the species (Tosh et al. 2008; Montgomery et al. 2012), although the

greater-white toothed shrew was also detected in Co. Laois, where it had not previously been recorded. Tosh et al. (2008) first identified the greater-white toothed shrew in owl pellets collected in Co. Limerick and Co. Tipperary. The samples from Waterford (n = 85) were not found to contain any bank vole (despite the known presence of bank vole in the area).

The frequency of occurrence of small mammals detected in this study was higher than some of the previously reported hard-part analysis studies (Warner and O'Sullivan 1982; Lynch and McCann 2007). There are a number of factors that may have contributed to this. The small bone remains from pygmy shrews for instance are difficult to manually identify, therefore the DNA method developed in this study may have provided a more accurate method of identification of the species. A hard part analysis study in Poland reported that intact skulls from small mammals are rarely found in carnivore scats, making identification of small bone remains more difficult (Krauze-Gryz et al. 2012).

3.4.4 Comparison to European studies

Previous dietary studies of marten species across their range have shown that small mammals, especially voles, are the most important food group in the diet (Lynch and McCann 2007; Caryl et al. 2012). De Marinis and Masetti (1995) showed that the wood mouse was a favoured small mammal in pine marten diet in only a few countries, mostly in the Mediterranean region. On island systems such as that in Minorca and Ireland, other prey items, especially fruit items, can amount to 30% of the diet, possibly as a result of the absence of field voles and formerly bank voles (Warner and O'Sullivan 1982, Lynch and McCann 2007). In this study, over 60% of scats did not contain any of the target species' DNA, indicating that other prey items such as insects and fruit were heavily supplementing the pine marten diet.

In most studies, shrews and insectivores are grouped into a single category, but collectively can account for as much as 24% of the diet in studies from Switzerland, Poland and France (De Marinis and Masetti 1995). Caryl et al. (2012) have shown that the pine marten in Scotland predominantly feed on the field vole (39%), but when

fruit becomes more plentiful in summer months, the pine marten becomes more of a generalist feeder, similar to the Irish pine marten. In Poland and Italy, the preferred mammal prey items of the pine marten are wood mice and bank voles (Rosellini et al. 2008; Balestrieri et al. 2011).

Zhou et al. (2011) described the general preference of the *Martes* species across the Holarctic region for fruit in more southerly countries, with the consumption of small mammals increasing in higher latitudes, which places the Irish pine marten as an exception to this generalisation. In addition, Zalewski (2004) also found that pine marten in Mediterranean regions were more likely to consume wood mouse than in northern areas, where field and bank vole consumption was more common. Other Irish species have however shown similar trends, and both the barn owl (*Tyto alba*) and short eared owl (*Asio flammeus*) have also shown preference for wood mouse over other small mammals in Ireland (Cullen and Smiddy 2012).

3.4.5 Comparison of molecular and traditional hard part analysis techniques

Sheehy et al. (2014) compared both molecular and traditional hard-part analysis of pine marten scats. The results showed that the methods were comparable for most of the assays used in this study, but found that the wood mouse was underestimated with the molecular analysis. There may be several explanations for this which could be related to the DNA extraction protocol used in this study. A small proportion (0.2 g) of the overall pine marten scat was taken, and it is possible that if the DNA of the prey item such as wood mouse was in low quantities, or not evenly distributed in the scat, the DNA extraction may have failed to co-extract the prey DNA. King et al. (2008) suggest multiple DNA extractions from different locations in the scats which may help detect prey items occurring at low or degraded quantities. Similarly, the storage of the scats prior to extraction may affect the results, with King et al. (2008) suggesting that scats could be stored in silica or ethanol in addition to freezing. Finally, although predator DNA could be detected, degradation of the prey DNA may cause it to go undetected in some cases due to the low quantities of prey DNA present in scats. King et al. (2008) recommend that traditional hard part analysis also be used in studies where more detailed and accurate results are needed. However, Sheehy et

al. (2014) found that the results were mostly comparable and gave an accurate estimate of pine marten diet and therefore, the method described in this study is a useful method to screen samples for mammalian prey DNA, while being aware of the benefits and the limitations of the technique.

3.5 Conclusion

- The DNA from two squirrel species, the red squirrel and the grey squirrel were successfully identified in DNA extracted from pine marten scats.
- The occurrence of mammal DNA in the diet of pine marten in this study was found to be broadly comparable with other Irish studies, with the exception of the pygmy shrew and red squirrel (detected more often in this study) which may be due to the increased detection ability of the suite of assays developed in this study (capable of identifying single hairs), and bias in the sample collection.
- Grey squirrel DNA was not detected in any of the field collected samples. However, it was detected in pine marten scats from Co. Wicklow in Sheehy et al. (2014). It is possible that grey squirrels and pine marten are less likely to overlap in distribution as has been suggested by Sheehy and Lawton (2014), but further research is needed across a broader area.
- In terms of limitations of this study, only mammal prey was identified, which is only a proportion of the overall constituents of the pine marten diet in Ireland. Other prey items such as invertebrates, reptiles, birds and fruit have been found to be equally important food items in the diet of the Irish pine marten.
- Future work could develop other assays for the identification of such food groups using real-time PCR, or a next generation sequencing approach could be developed to target other taxa.

Chapter 4

The search for sex specific markers for the red squirrel (*Sciurus vulgaris*) and the grey squirrel (*Sciurus carolinensis*)

4.1 Introduction

Non-invasive DNA assays facilitate the study of elusive, rare and endangered species, and are useful for estimates of population parameters. Sex determination of non-invasive samples provides further important information regarding the gender ratio and social organisation of the population. However, such assays remain largely unavailable for the red squirrel, although such studies have been quite limited for Sciurids in general as there are few studies relating to the sex identification of the Sciuridae. The ZFX/ZFY has not previously been successfully used for the sex identification of squirrels. Ermakov et al. (2006) found that ZFY primers amplified both sexes, leading Ermakov et al. (2006) to suggest that the region was not confined to the Y-chromosomes in the Sciuridae. Some studies have had success using the SRY gene for sex identification of the red squirrel,Bryja and Konecny (2003) used SRY primers developed by Pomp et al. (1995) to amplify a short section of the SRY gene in red squirrels, But Ermakov et al. (2006) found SRY primers amplified multiple fragments of the SRY in Sciurids and Gorrell et al. (2012) found that non-specific amplification also occurred. However, both Ermakov et al. (2006) and Gorrell et al. (2012) did find the SMCY gene was sex specific in the Sciuridae. Gorrell et al. (2012) also designed short internal primers for sex identification of tree squirrels, including the North American grey squirrel, and these were shown to be sex specific.

The aim of this study was to assess the suitability of a number of genes for sex identification in red squirrels and grey squirrels including the ZFX/ZFY, SRY, DBY, and SMCY. Regions of the genes were then DNA sequenced, and assays for real-time and conventional PCR for sex determination of squirrels were designed for use with tissue DNA and non-invasive samples.

4.2 Materials and Methods

4.2.1 Sample collection and DNA extraction

Tissue and hair DNA samples from known sex red squirrel and grey squirrel (n = 5 of each) were initially used to test published primers for ZFX/Y, SRY, DBY and SMCY (Table 4.1).

4.2.2 PCR

Primer pairs were used to generate PCR products for the corresponding genes from Table 4.1. The PCR consisted of 5 µl of GoTaq® Hot Start Green Master Mix (Promega), 0.2 µM of each primer and 1 µl of DNA in a total volume of 10 µl. Negative controls contained water in place of DNA. The PCR programmes were as published by the corresponding authors in Table 4.1, except in the case of the primers developed by Aasen and Medrano et al. (1990), where the protocol described by Statham et al. (2007) was used. PCR products were separated, visualised, cleaned and DNA sequenced as per Section 2.2.4.2. Real-time PCR for the SYBR® Green 1 assays was as described in Section 2.2.4.1. The number of cycles was increased from 40 – 50 to reflect the decreased quantity of nuclear DNA compared to the mitochondrial DNA.

The purified PCR products for primer pairs LGL331/LGL335 and DBY7F/R could not be directly sequenced and required cloning due to the amplification of multiple products. The purified PCR product was cloned using the Qiagen PCR Cloning Kit (cat. no. 231222) following the manufacturer's instructions.

4.2.3 Real-time PCR

Real-time PCR with the TaqMan® labelled probes consisted of 5 µl of TaqMan® Universal PCR Mastermix, 0.25 µl of the Custom TaqMan® Assay SNP Genotyping Mastermix (cat. no. 4332077) (supplied at a 40X concentration, Applied Biosystems) and 1 µl DNA in a total volume of 10 µl. The PCR program used was 2 minutes at 50 °C, 10 minutes at 95 °C, and 50 cycles of 15 s at 95°C and 1 minute at 60 °C.

Table 4.1: Previously published primers used in this study.

Target gene	Primers	Primer sequence (5'-3')	Authors
ZFX/ZFY	P1-5EZ /P2-3EZ	F:ATAATCACA TGGAGAGCCACAAGC R:GCACTTCTTTGGTATCTGAGAAAGT	Aasen and Medrano (1990)
ZFX/ZFY	LGL331/LGL335	F: CCA ATCATGCAAGGATAGAC R: CCAATCATGCAAGGATAGAC	Cathey et al. (1998)
SRY	SRYA-5/SRYA-3	F: TGAACGCATTCATGGTGT GGT R: AATCTCTGTGCCTCCTGGAA	Pomp et al. (1995)
SRY	SRY593 /SRY764	F: AAGCGACCCATGAACGCATT R: GTATTTCTCTCTGCATGG	Palsboll et al. (1992)
DBY7	DBY7F/R	F: GGTCCAGGAGARGCTTTGAA R: CAGCCAATTCTCTTGTTGGG	Hellborg and Ellegren (2003)
DBY8	DBY8F/R	F: CCCCAACAAGAGAATTGGCT R: CAGCACCACCATAKACTACA	Hellborg and Ellegren (2003)
SMCY7	SMCY7F/R	F: TGGAGGTGCCCRAARTGTA R: AACTCTGCAAASTRTACTCCT	Hellborg and Ellegren (2003)
SMC8	SMCY8D/SMCY9R	F: ATGCTCTCGTGGGGATGAAG R: ACAGGCATGTTGAAGTAGTC	Ermakov et al. (2006)
SMC8 (internal)	ThuSMCF/R	F: CCTTCCTGACATTCCTAGAG R: TTTAGTTTCCCTTGGTTCAG	Gorrell et al. (2012)

4.2.4 DNA Sequence Analysis

Nucleotide sequences were analysed using the BLAST software at http://www.ncbi.nlm.nih.gov/BLAST/ from the GenBank (NCBI) database (Altschul et al. 1990). DNAStar's SeqMan 5.05 software was used to compare forward and reverse sequences and was also to ensure base calling accuracy prior to building contigs, which is a sequence constructed based on overlapping sequences. The consensus sequences were then exported to MEGA V.5.05 (Tamura et al. 2011) for multiple alignments and compared to sequences on GenBank (based on the BLAST analysis).

4.3 Results

4.3.1 ZFX/ZFY (P1-5EZ and P2-3EZ primers)

The primers P1-5EZ and P2-3EZ (Aasen and Medrano 1990) were used to amplify a 447 bp product in male and female red and grey squirrels (Fig. 4.1). This primer pair were expected to have amplified both the ZFX and the ZFY together in the case of male samples, and just the ZFX in female samples.

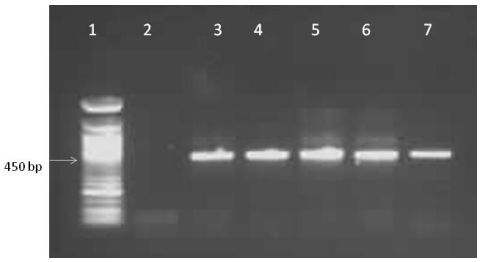

Fig 4.1: Amplification with P1-5EZ and P2-3EZ primers. *Lane 1* 50 bp ladder, *lane 2* negative control, *lanes 3* and *4* male red squirrel, *lane 5* male grey squirrel, *lane 6* female red squirrel and *lane 7* female grey squirrel.

The PCR products were sequenced for DNA sequence analysis. Double peaks were recorded in the sequencing chromatograms derived from male red squirrels (Fig 4.2). These second peaks correspond to the ZFY region, a conflict between co-amplified regions from both the ZFX and ZFY which has been reported in other species (Statham 2005). There were some sequence variations between both red squirrels and grey squirrels, which were later avoided for primer design so that the same assay could be used for both species.

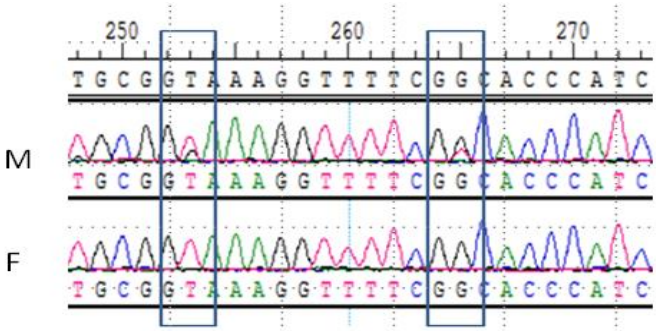

Fig 4.2: DNA sequencing chromatogram of ZFX/ZFY PCR products generated using P1-5EZ and P2-3EZ primers using DNA from male (M) and female (F) red squirrels. The first box shows the presence of a double peak at 'T', where a 'G' can also be seen in the male sequence, and is absent in the female sequence. The second box shows the presence of a double peak at 'G', where a 'T' can also be seen in the male sequence and is absent in the female sequence.

By removing the base seen in the ZFX sequence from the double peak in the ZFX/ZFY mixed sequence it was possible to create the ZFY sequence. The two sequences were subsequently created (ZFX and ZFY) of both red squirrels and grey squirrels to allow subsequent primer design for both species. The sequences were aligned using a Clustal W alignment (Fig 4.3). Primer and probes were designed using the Custom TaqMan® SNP Genotyping Assay Service. The assay was designed with a common forward and reverse primer along with two specific probes, one designed to amplify the ZFX and the other for the ZFY (Fig 4.3).

```
ZFX     AAAGGAGCCAACAAAATGCACAAGTGTAAATTCTGTGAATATGAGACAGCTGAACAAGGG 60
ZFY     AAAGGAGCCAACAAAATGCACAAGTGTAAATTCTGTGAATATGAGACAGCTGAACAAGGG 60

ZFX     TTGTTGAATCGCCACCTTTTGGCAGTCCACAGCAAGAACTTTCCTCAAATTTGTGTNGAG 120
ZFY     TTGTTGAATCGCCACCTTTTGGCGGTCCACAGCAAGAACTTTCCTCATATTTGTGTNGAG 120

ZFX     TGCGGTAAAGGTTTNCGNCATCCATCNGAGCTCAAAAAGCACATGCGAATCCATACTGGN 180
ZFY     TGCGGGAAAGGTTTNCGNCACCCGTCNGAGCTCAAAAAGCACATGCGCATCCATACTGGN 180

ZFX     GAGAAGCCTACCANTGCCAGTACTGNGAATATAGGTCTGCAGACTCTTCTAACTTGAAAA 240
ZFY     GAGAAGCCTACCANTGCCAGTACTGNGAATATAGGTCTGCAGACTCTTCTAACTTGAAAA 240

ZFX     NNCATGTAAAAACTTTGTTGAATCGCCACCTTTTGG 276
ZFY     NNCATGTAAAAACT---------- 254
```

Fig 4.3: Alignment of ZFX and ZFY sequences conserved between red squirrel and the grey squirrel. *N* is used where a SNP occurred in either the red squirrel or the grey squirrel. The forward and reverse primers were conserved between both sexes, and the probes contained a SNP (yellow), each specific to the ZFX or the ZFY.

The real-time PCR results revealed that the ZFX probe amplified both red and grey squirrel DNA. However, the ZFY probe failed to work with either species (Fig 4.4). However, both the ZFX and ZFY assays amplified the ZFX of pine marten. Further sequence analysis of the assay revealed that the intended ZFY squirrel primers were conserved in the pine marten ZFX, with only one bp substitution in the reverse primer (Fig. 4.5).

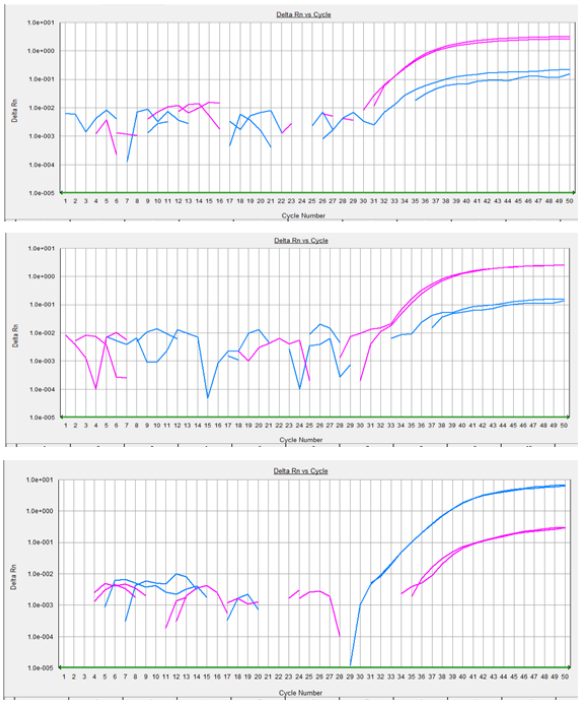

Fig 4.4: Example of real-time PCR amplification with the Custom TaqMan® SNP Genotyping Assay designed to target ZFX/ZFY sequences. The ZFY failed to amplify either sex, but amplified the ZFX in a pine marten. Blue = ZFY assay amplification, Pink = ZFX assay amplification.

```
Squirrel_ZFY    GTGTAAATTCTGTGAATATGAGACAGCTGAACAAGGGTTGTTGAATCGCCACCTTTTGGC 83
Squirrel_ZFX    GTGTAAATTCTGTGAATATGAGACAGCTGAACAAGGGTTGTTGAATCGCCACCTTTTGGC 83
Marten_Zfx      GTGTAAATTCTGTGAATACGAGACAGCCGAACAAGGGTTGTTGAATCGCCACCTTTTGGC 180
Squirrel_ZFY    GGTCCACAGCAAGAACTTTCCTCATATTTGTGTNGAGTGCGGGAAAGGTTTNCGNCACCC 143
Squirrel_ZFX    AGTCCACAGCAAGAACTTTCCTCAAATTTGTGTNGAGTGCGGTAAAGGTTTNCGNCATCC 143
Marten_Zfx      GGTCCACAGCAAGAACTTTCCTCATATCTGTGTGGAGTGCGGGAAAGGTTTCCGTCACCC 240
Squirrel_ZFY    GTCNGAGCTCAAAAAGCACATGCGCATCCATACCGGNGAGAAGCC-TACCANTGCCAGTA 202
Squirrel_ZFX    ATCNGAGCTCAAAAAGCACATGCGAATCCATACTGGNGAGAAGCC-TACCANTGCCAGTA 202
Marten_Zfx      GTCAGAGCTCAAAAAGCACATGCGCATCCATACCGGGGAGAAGCCGTACCAGTGCCAGTA 300
Squirrel_ZFY    CTGNGAATATAGGTCTGCAGACTCTTCTAACTTGAAAANNCATGTAAAAACT-------- 254
Squirrel_ZFX    CTGNGAATATAGGTCTGCAGACTCTTCTAACTTGAAAANNCATGTAAAAACT-------- 254
Marten_Zfx      CTGCGAGTATAGGTCCGCAGACTCTTCTAACTTGAAACACATGTAAAAACGAAGCATAG 360
```

Fig 4.5: Clustal W alignment of squirrel ZFX/ZFY with the pine marten ZFX. The alignment confirms that the intended target for the squirrel ZFY is conserved on the pine marten ZFX.

The results of the ZFX/ZFY assay show that the assay was functional as it amplified the ZFX in squirrels, and the ZFY assay amplified the pine marten ZFX.However, this shows that there was a problem in terms of the design of the ZFY assay as it did not amplify the intended target.

4.3.1.2 ZFX/ZFY (LGL-331 and LGL-335 primers)

The primers LGL-331 and LGL-335 (Cathay et al. 1998) were also used to generate ~1000 bp portions of the ZFX/ZFY genes in squirrels. This primer pair were expected to generate a single product in females (ZFX) and two products in males (ZFX/ZFY). Two products were found in the amplified squirrel samples, one at 500 bp and a second product at 900 bp (Fig 4.6). The males all produced two products, and the females produced one, although a faint second female product seen at SV04. However, repeat PCRs showed that only the 900 bp fragment could be reliably obtained and the generation of the 500 bp product did not appear to be consistent with gender identification.

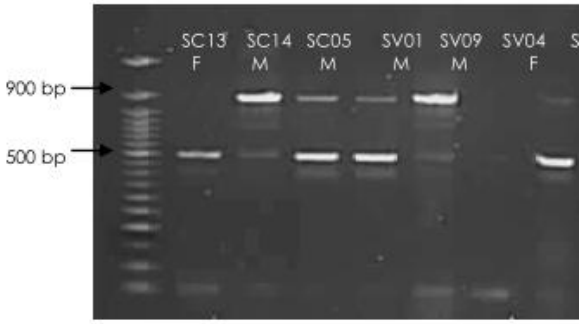

Fig 4.6: Amplification of male and female squirrels with LGL-331 and LGL-335 (Cathay et al. 1998).

The PCR products from Fig 4.6 were subsequently cloned and sequenced. The DNA Blast analysis (Altschul et al. 1990) revealed that the 900 bp product showed closest homology to both ZFX and ZFY sequences on Genbank, while the 500 bp product showed very poor homology to other sequences on Genbank with only about 150 bp showing homology to the ZFX region in other species. A table of closest identity hits (Table 4.2) shows that the 900 bp product showed a high degree of overlap (93%) homology (97%) to the ZFX region of other mammals, but the 500 bp product only showed 29-31% overlap and 84% homology to the ZFX.

Table 4.2: BLAST search results for squirrel ZF sequences derived from the LGL-331/LGL-335 primers.

	500 bp product homology				
	Species	**Common name**	**Gene target**	**Query coverage**	**Identity**
1	*Bos Taurus*	cow	ZFX	31%	84%
2	*Lynx rufus*	bobcat	ZFX	31%	82%
3	*Cervus elaphus*	red deer	ZFX	31%	82%
4	*Halichoerus grypus*	grey seal	ZFX	27%	84%
5	*Zalophus californianus*	Californian sea lion	ZFX	29%	82%
6	*Leptonychotes weddellii*	Weddell seal	ZFX	29%	82%
7	*Ommatophoca rossii*	Ross seal	ZFX	29%	82%
8	*Lobodon carcinophaga*	crabeater seal	ZFX	29%	82%
9	*Phoca vitulina*	harbour seal	ZFX	29%	81%
	908 bp product homology				
	Species	**Common name**	**Gene target**	**Query coverage**	**Identity**
1	*Spermophilus tridecemlineatus*	thirteen-lined ground squirrel	ZFX	97%	95%
2	*Ailuropoda melanoleuca*	giant panda	ZFX	97%	94%
3	*Cavia porcellus*	guinea pig	ZFX	97%	93%
4	*Tursiops truncatus*	common dolphin	ZFX	97%	93%
5	*Orcinus orca*	killer whale	ZFX	97%	93%
6	*Saimiri boliviensis*	blacked-capped squirrel monkey	ZFX	97%	93%
7	*Sus scrofa*	pig	ZFX	97%	93%
8	*Heterocephalus glaber*	naked mole rat	ZFX	97%	93%
9	*Mustela putorius furo*	ferret	ZFX	97%	93%

Further sequence analysis revealed that the region of the 500 bp product that showed homology to the ZFX sequences in Table 4.2, also overlapped with the larger 900 bp product in this study (Fig 4.7).

```
900 bp      ACAAGAGTGAGCCCGTGGCCAATAGTATTCAGAGCACATACTAAATTTTTGAAGAAGCCA 776
500 bp      --------GAGCATATGCCCAGTAATATTCAGAAAACACT---AATGTTTGAAAAAAGTA 393

900 bp      AAACAGAACTTGATTTGATCCCTTATGATCTTTTCTTTTCCTTTTTTAGCAATAATTATT 836
500 bp      AAACAGAACAGGGTTTGAACCCTTGTGTTTTTTTCTTTTCCTTTCATAGCAATAAT---- 449
```

Fig 4.7: Clustal W alignment showing the region of overlap between the 500 bp and the 900 bp sequence overlap.

Based on the sequences derived from the LGL-331/LGL-335 primers, new primers were designed to target the internal region of both these sequences. Two forward primers were designed specific to both sequences while a common reverse primer was used targeting a 600 bp product from the larger region and a 450 bp product

from the shorter sequence region. The aim of this was to investigate if the second sequence was sex specific.

```
GAAAAGGAATTCCTCAAATATGTTGCAAGCAGTAACTTTTTGTAAAAAAAAAAAAAATTA 300
------------------------------------------------------------


----------------------------------TTAGACCTGATTCCAGACAGTACCAA 26
GTATATTTAAAGAACCTAATTATGTCAGCATAAAGCAGGCATAATTT-ACAGAGCAGCAG 359

ACAGGTGAGAACACATGAATTCCACAGCACTG-CAAGCTCTGCTGGGCAACTCTCAGAGG 85
TCTTAAGTCCTTCAAAGCATGCGACATTCCTGTCTATCCTTGTATAATGACTTGCATACT 419

ACA---CTCATTCATCCACAGAGG-----ACATTCATT--CATTAATTACCTGTTGTAAT 135
AAAAATTTTATAGATATATTTAGGTTTTAATGTAAATTAGCATAAATTGTTTTTTTGAAT 479

AACCAAATATAGAAATGCTAA--ATAAACTCTGAGTTTTTTGTTTTTTTTTTTAACAGAT 193
AACCAAA-GTAGGAATTTTGATTATAAAACTTTACTTCTATGAATAATTTTCCC-CAGAT 537

TTCTAAGTTGT---AGTTTCATCACTACTTAGT----------TGTTCA---ATTT 233
TTCTGAATTGATAAGGTTTAACCACTGGTTAATAGAAGTTACATATTTATTCATTAATTT 597

 --TAAACTGGAAGGATTGAGATTAGTATGGTCATAAGCACACAGATTTTTGTTTTTATA 291
TGTTAAACTGAAGGGAGTGAGGTTGGTACAGTTCTACTTA-GTGAAGTTTCCAGACCAGG 656

GGATTGCAATGTCTCTA--GTTGAAAAGCGGTGAAAAAAAAGTTCTTATTTTACAG---- 345
TTCTCATTATATCTCTTATGTTGAAAAATGATAAAAGGGATGTTAGTAGCAGATAGCAAA 716

--------GAGCATATGCCCAGTAATATTCAGAAAACACT---AATGTTTGAAAAAAGTA 394
ACAAGAGTGAGCCCGTGGCCAATAGTATTCAGAGCACATACTAAATTTTTGAAGAAGCCA 776

AAACAGAACAGGGTTTGAACCCTTGTGTTTTTTTCTTTTCCTTTCATAGCAATAATTATT 454
AAACAGAACTTGATTTGATCCCTTATGATCTTTTCTTTTCCTTTTTTAGCAATAATTATT 836

GGCCCTGATGGACATCCCTTGACTGTCTATCCTTGCATGATTTGAA 500
GGCCCTGATGGACATCCCTTGACTGTCTATCCTTGCATGATTGGA- 881
```

Fig 4.8: Clustal W alignment of internal PCR primers for the amplification of the two sequences generated by the LGL-3111/LGL-335 primers. The position of two sequence specific forward primers and the common reverse primer is shown.

The conventional PCR primers were subsequently tested with a panel of known sex tissue samples, where it was found that the larger product sometimes did not amplify (possibly due to the larger size) (Fig 4.9), and replicated data showed that both primer

sets amplified both sexes. There was also a product found at the 50 bp line which was assumed to be primer dimer, but due to the intensity of the band, it may have been amplification of a second PCR product.

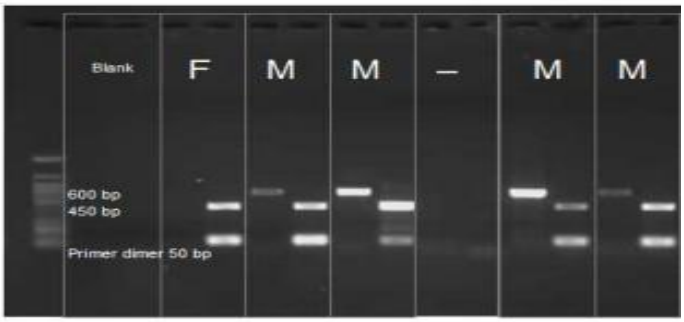

Fig 4.9: Amplification of male and female red squirrels with the internal primers (Fig 4.8) designed from the sequences derived from the LGL-331/LGL335 primers. The optimised PCR protocol consisted of denaturation 95 °C for 5 min, 40 cycles of 95 °C for 60 s, 56 °C for 60 s, 72 °C for 50 s, final extension 72 °C for 5 min.

A panel of real-time PCR primers were also designed based on the sequences and all primers amplified both sexes, again showing that this gene region was unsuitable for sex differentiation in squirrels (Fig 4.10 [a, b]). All primer pairs amplified both sexes. The primers were shown to be working properly, and Ct values were found to increase as the DNA template decreased using a dilution series (Fig 4.10 [c]).

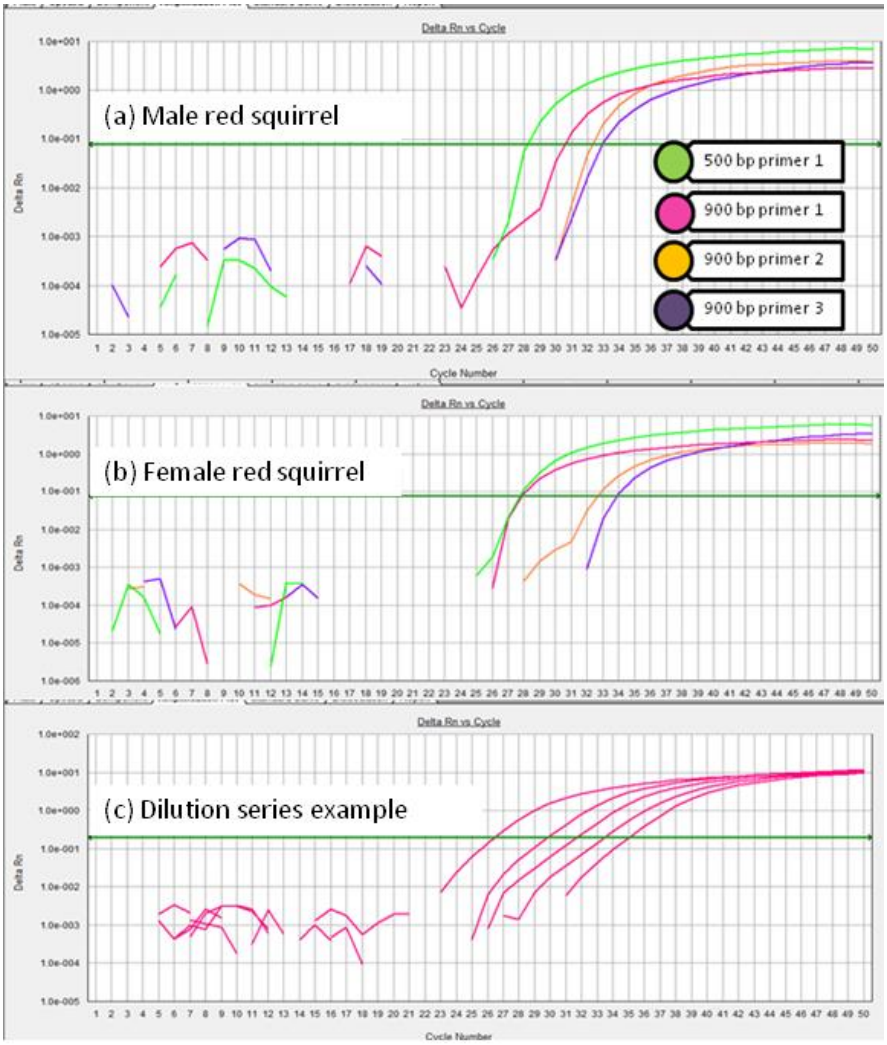

Fig 4.10: Real-time PCR amplification of (a) a male red squirrel (b) a female red squirrel with primers designed from the LGL-331/LGL-335 sequences. All primers amplified both sexes. (c) An example of a dilution series of a tissue sample with a starting template of 4 ng of DNA.

4.3.2 SRY primers

Two sets of SRY primers were used with the squirrel DNA panel. The primers designed by Pomp et al. (1995) (originally designed to sex pig embryos) were expected to yield a 157 bp product, but over 700 bp was produced in this study, and multiple bands were present which may have been difficult to sequence, and were not used for further analysis (Fig 4.11). Similar non-specific binding was also experienced using the other SRY primer sets (Table 4.1). The SRY gene was not further persued for sex identification of squirrels.

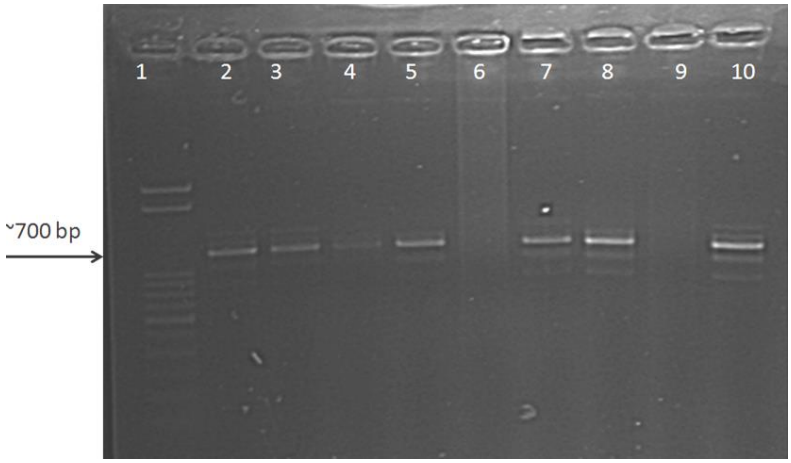

Fig 4.11: Example of SRY amplification with the primers SRYA-5/SRYA-3. Amplification was seen in both male and female samples. Samples 1-5 are female and 6 -10 are male.

4.3.3 DBY primers

The DBY assays did not reliably amplify male-specific DNA sequences from both species and were also non-specific as DNA from known females were also amplified. The primers DBY7 sometimes produced an amplicon of ~350 bp in both squirrel species and a second band was sometimes visible at 100 bp. The expected amplicon size was 250-700 bp for DBY7 and 185-200 bp for DBY8 (Hellborg and Ellegren 2003). In addition, when using grey squirrel DNA, the assay produced a second band (Fig 4.12).

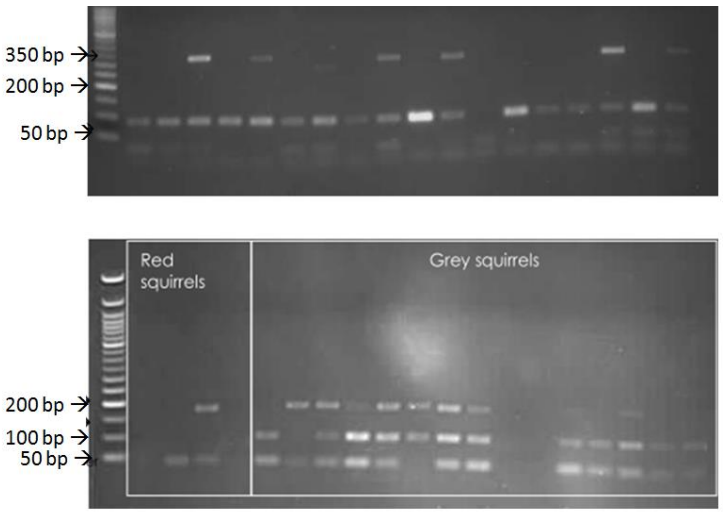

Fig 4.12: Amplification with (a) DBY 7 and (b) DBY 8 primers.

The BLAST analysis of grey squirrel DBY sequence revealed the closest identity (87%) to crab eater seal (*Lobodon carcinophaga*). The neighbour joining tree (Fig. 4.13) revealed that the DBY sequence of the grey squirrelwas quite genetically divergent to other species on the Genbank database. On the upper branch of the tree, three cervid species group together, and on the lower section of the tree two shrew species: *Crocidura russula* and *Crocidura leucodon* grouped together along side another group of three pinnipeds. Despite the availability of *Mus* and other rodent DBY sequences, the sequences used to construct the tree were the closest hits to grey squirrel.

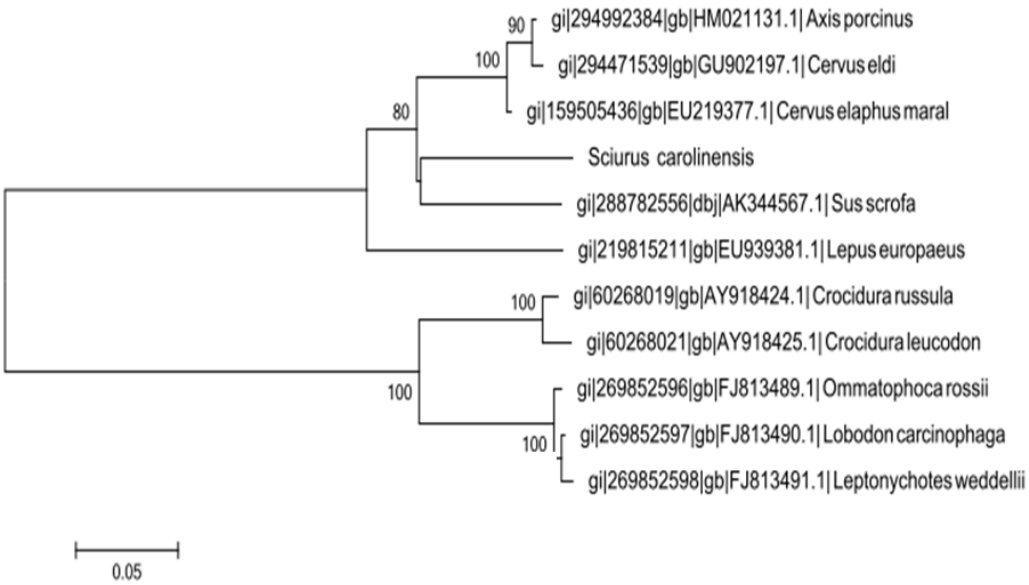

Fig 4.13: Neighbour joining tree of DBY sequences neighbour joining trees was constructed using 1000 bootstraps and genetic distance was computed using the *P*-distance method (Nei and Kumar 2000)**.**

Real-time PCR assays were designed based on the sequences derived from the larger amplified fragment, thought to be male specific (Table 4.3). The primer pair were tested with male and female squirrels, and were not found to be sex specific (Fig 4.14).

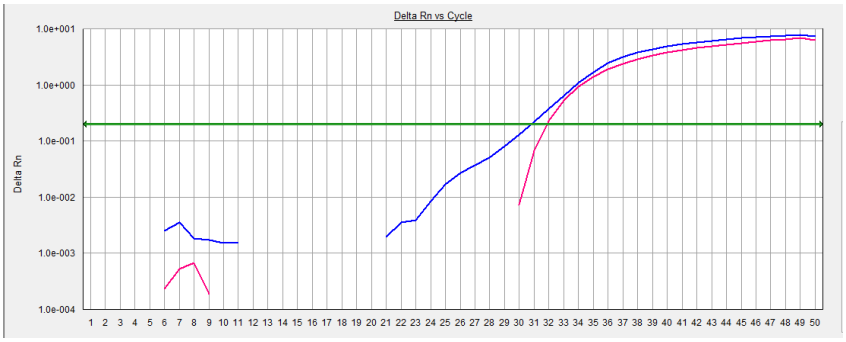

Fig 4.14: Example of amplification of DBY real-time PCR assay with a female squirrel (blue). The pink line is a ZFX positive control. The real-time PCR DBY primers were not found to be sex specific.

4.3.4 SMCY primers

All the SMCY primers amplified products of the expected size in males of both squirrels, with 500 bp for SCMY7 (Fig 4.15), and 600 bp for SCMY8/9.

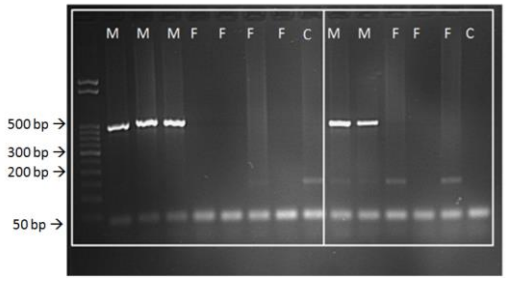

Fig 4.15: Amplification with the SMCY7 primers (Hellborg and Ellegren 2003). The left hand side of the image shows male red squirrel amplification, and the right hand side of the image shows male grey squirrel amplification. The primers were sex specific with both species.

The SMCY red and grey squirrel sequences (from this study) were truncated to 491 bp to compare with other published Sciurid sequences (Ermakov et al. 2006; Chang et al. 2011; Gorrell et al. 2012; Ermakov et al. 2006). The neighbour joining tree (Fig. 4.16) showed that the SMCY sequence derived from both red and grey squirrels were most closely related to North American red squirrel (*Tamiasciurus hudsonicus*). The tree grouped the Asian squirrels in one group, the ground squirrels in the second group and red squirrel, grey squirrel and *T. hudsonicus* in the third group. The tree

clearly separated the three known groups of squirrels to Sciuridae, *Tamiops* and *Spermophilus* or *Urocitellus*.

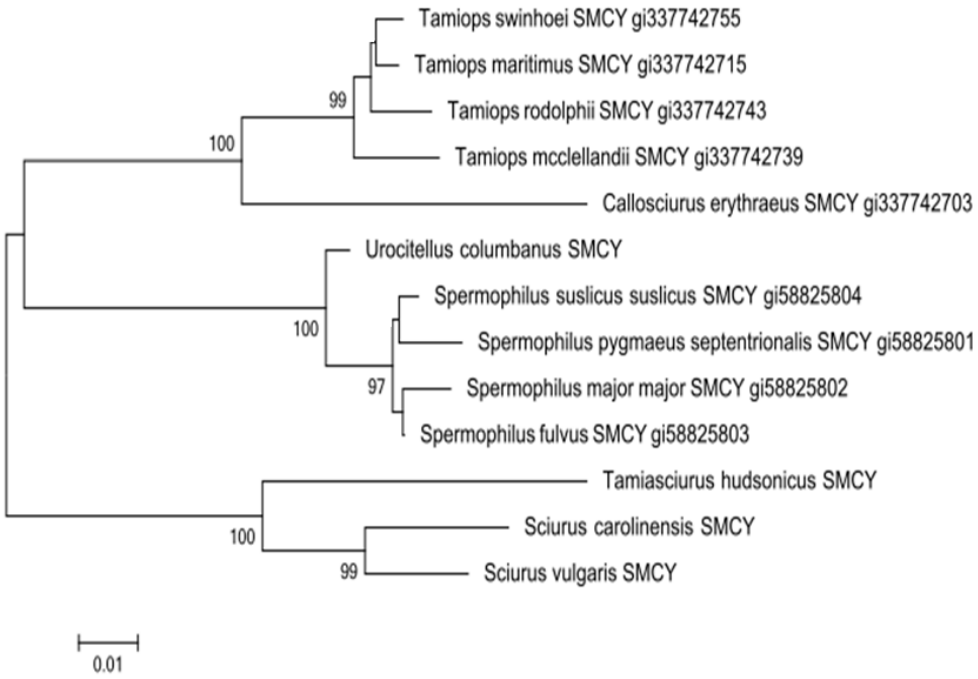

Fig 4.16: Neighbour joining tree of SMCY sequences of Sciurid species constructed using 1000 bootstraps and genetic distance was computed using the *P*-distance method (Nei and Kumar 2000)**.**

The internal primers designed by Gorrell et al. (2012) to amplify the SMCY of male *Tamiasciurus hundsonicus*, and had been shown to also amplify male grey squirrels were used in this study in a real-time PCR SYBR® Green 1 reaction with the red and grey squirrel sample panel. The results showed that the primers successfully amplified male grey squirrel samples, with some late amplification with the female grey squirrel samples. However, the primers amplified both sexes of red squirrels, and were not suitable for sex differentiation in that species.

4.3.4.1 SMCY Real-Time PCR primer design

Two sets of real-time PCR assays (to increase the target area covered by the assay) were designed for sex identification of both red squirrel and grey squirrel (Table 4.3).

The internal primers ThuSMCYF and ThuSMCYR (Gorrell et al. 2012)(Table 4.1) were successfully used for sex identification of grey squirrels using real-time PCR in this study. The Ct values increased with decreased DNA template, indicating that the assay was working and targeting the intended sequence. There was however late amplification in female samples (> 30 Ct). Both the forward and reverse primers only differed by a single base pair from the sequence obtained from the red squirrel, suggesting that if the region was sex specific in grey squirrels, it should also be sex specific in red squirrels (Fig. 4.17). However, this was not observed in this study, as the assay amplified both sexes of red squirrel.

Fig 4.17: SMCY alignment of red and grey squirrel sequences with the internal primers ThuSMCYF and ThuSMCYR (1) designed by Gorrell et al. (2012). 2 and 3 show the positions of real-time PCR primers designed in this study (SQ1SMCY and SQ2SMCY).

The real-time PCR assays SQ1SMCY and SQ2SMCY designed in this study amplified both species and both sexes of red and grey squirrel (Fig 4.18). The standard curve exhibiting an R^2 value of 0.999 indicating that the amplification was accurate as a result of an observed linear relationship between a dilution series of genomic DNA concentration and Ct value over the range tested. The internal assays based on the sequence in Gorrell et al. (2012) did not reliably amplify either sex.

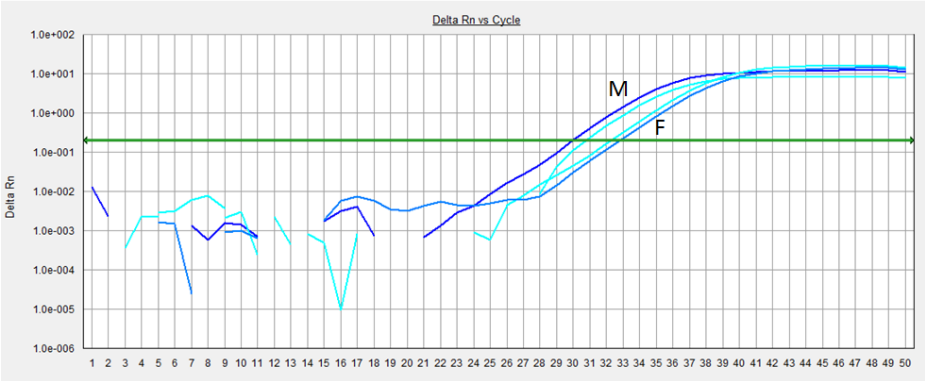

Fig 4.18: Real-time PCR amplification with SQ1SMCY and SQ2SMCY primers. Both primer sets amplified both sexes.

To further investigate why internal sections of the SMCY were amplifying both male and female squirrels, the forward primer from SQ1 and the reverse primer from SQ2 were used to amplify an ~250 bp product and the primer ThuSMCYF (Gorrell et al. 2012) was also used in conjunction with SQ2SMCYR (Fig. 4.19). All combinations were shown to amplify male and female squirrels from both species. The region was sequenced to verify if it was the SMCX target by alignment with SCMY sequences The sequence derived from females in this study was identical to that of the SMCY region from male squirrels indicating that at least a partial region of the sequence is not sex specific.

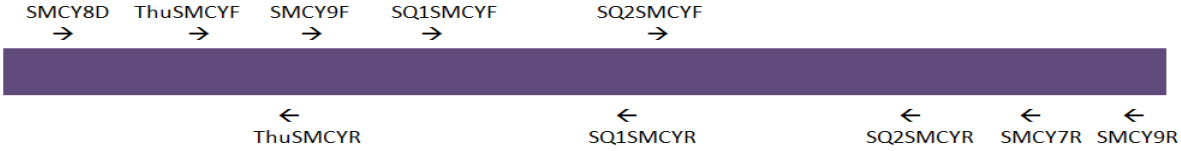

Fig 4.19: Schematic diagram of the SMCY gene region with the arrows broadly indicating primer binding sites and direction of extension.

Table 4.3: Primers used to amplify sex gene targets, and the subsequent primers designed for the sex identification of squirrels using real-time PCR (qPCR) and conventional PCR (cPCR).

Gene Target	Primers	PCR system	Primer name:	New Primers (5'-3')	Probe (5'-3')
ZFX	P1-5EZ/ P2-3EZ	qPCR	ZFX1F/ZFX1R	F: GGCATTGGTACGGCTTCTCAC R: TGGCATTGGTACGGCTTCTCA	
ZFY	P1-5EZ/ P2-3EZ	qPCR	ZFY1F/ZFY1R	F:TTGGCACTGGTATGGGTTCTCTC R: CCTCATATTTGTGTGGAATGTGGG	
ZF intron-exon region	P1-5EZ/ P2-3EZ	cPCR	SQZF/SQZR	F: TGGCCCTGATGGACATCCC R: GTGAGTAAACAAAGCCCCAG	
ZFX/ZFY	P1-5EZ / P2-3EZ	qPCR	SQZF/SQZR	F:GTCCACAGCAAGAACTTTCCTCA R:TCAAGTTAGAAGAGTCTGCAGACCTA	SQZFX: 6VIC-CACATGCGAATCCAT SQZFY: FAM-CACATGCGCATCCAT
ZFX (900 bp)	LG-331/ LG-335	cPCR	P12.1F/R	F: GTGAGGGCGCACGAGTTC R: CCCCTGTGGATACATTCTAGTTTCC	
ZFX (900 bp)	LG-331/ LG-335	qPCR	P12.2F/R	F:ATTCCTCAAATATGTTGCAAGCAGTA R: GGAATGTCGCATGCTTTGAA	
ZFX (900 bp)	LG-331/ LG-335	qPCR	P12.3F/R	F: GGTGTCATGATGGCATTTTAGC R: ATCCATGTTTGATCCATTTCC	
Possible ZFY (500 bp)	LG-331/ LG-335	cPCR	Carp9F/R	F:TGCAAGCTCTGCTGGGCAACTCTCAGAGG R: AGGGATGTCCATCAGGGCCAA	
ZFX	LG-331/ LG-335	cPCR	Carp12F/R	F: ATTCCTCAAATATGTTGCAAGCAGTA R: AGGGATGTCCATCAGGGCCAA	
DBY	DBY7-8	qPCR	DBYF/R	F:GGAAGCCAGAAAAGTAAGTATGCAT R: CGGTATGAAAACTATAAGCAAGAAATTTG	
SMCY	SMCY7	qPCR	SQ1SMCYF/R	F: CACTAGCCATGAGATGGTAGACCTT R: TTTGCTTCCATAGTTGAAGTCATCTAA	
SMCY	SMCY7	qPCR	SQ2SMCYF/R	F: GCTACCTTTGTTTTGCTAAAATACTTTCT R: CCTGTTCAAAACCAAAAGCTTCA	

4.4 Discussion

The aim of this work was to investigate the utility of previously published sex determining markers with red and grey squirrels in order to develop a sex determining assay for use with non-invasive samples. The ZFX/ZFY system is one of the most frequently used assays in non-invasive studies (Statham et al. 2007; Mullins et al. 2010; Pilgrim et al. 2012). However, the results in this study have confirmed that while the ZFX can be consistently and accurately reproduced in both red and grey squirrels, the ZFY could not. Results from the LGL331 and LGL335 primers (Cathey et al. 1998) indicated that the ZFY was not present in squirrels, despite a second fragment sometimes amplifying with the ZFX sequence. Sequence analysis revealed that this region was not sex-specific. Similar difficulties were experienced with the P1-5EZ and P2-3EZ primers (Aasen and Medrano 1990) also targeting the ZFX/ZFY region despite the appearance of double peaks in the chromatogram for known males, a diagnostic feature that has been used previously for sex determination in other species (Statham et al. 2007).

This is not the first study to report problems amplifying the ZFY in Sciurids. Ermakov et al. (2006) reported that the region was absent from the Y-chromosome in ground squirrels, or that the region occurred in autosomes, meaning that the region was not sex-specific. Those observations appear consistent with the results from this study and highlight the importance and need for careful interpretation of the results prior to sex determination. However, in this study the ZFX assays that were designed are conserved in both red and grey squirrels.

The SRY region again proved to be not suitable for sex determination in either squirrel species. Kusahara et al. (2006) successfully designed male specific assays for the Pallas's squirrel. However, Gorrell et al. (2012) also found that the method was not reliable in Scuirids. Ermakov et al. (2006) found that the SRY gene was unsuitable for sex determination in ground squirrels as the primers they used amplified multiple PCR products. The multiple banding was consistent with the results from the other SRY primers used with squirrels in this study. The SRY region may be repeated on

autosomes in the squirrel genome as the patterning was non-sex specific and remain suitable for use with another Y-specific marker for sex determination.

The DBY region has also been successfully used for sex determination in a number of studies covering a variety of species including the Florida manatee (*Trichechus manatus latirostris*) and the naked mole rat (*Heterocephalus glaber*) (Tringali et al. 2007; Katsushima et al. 2010). The primers DBY7 and DBY8 were not sex-specific in red and grey squirrels in this study. The sequences obtained from the DBY region confirmed that the region was comparable to other available DBY sequences from other mammals on the Genbank database. However, the region had not been previously investigated for its utility for gender determination of squirrels, and the results from this study have shown that this region may not be specific to the Y-chromosome of squirrels.

The SMCY gene promised to be the most suitable gene for sex determination in both squirrel species, as it has been successfully used for other Sciurids (Gorrell et al. 2012). However, work from this study has generated some interesting but conflicting results. The primers targeting the SMCY7 region (Hellborg and Ellegren 2003) were successfully used to generate male-specific sequence from male samples of red and grey squirrel. However, subsequent development of internal real-time PCR primers for sex specific amplification of red and grey squirrels generated products from both male and female squirrels, revealing further complexities within the squirrel Y-chromosome.

The internal SMCY primers designed by Gorrell et al. (2012) did distinguish the sex of grey squirrel (using real-time PCR), but late amplification occurred in female samples which may lead to some ambiguity when applying the assay to non-invasive samples. The primers designed by Gorrell et al. (2012) were not sex specific in the red squirrel, despite only 1 bp changes in both the forward and reverse primers from the red squirrel sequence obtained in this study. This study raises questions about the general use of the SMCY region for sex determination in all of the Sciurid species and requires further investigation prior to its application to other species. The results from this study suggest that while the primers developed by Gorrell et al. (2012) are

sex specific in the described species, but the assays may not be suitable for sex identification of other closely related species. Furthermore, difficulties were experienced in this study when applying primers designed to target regions within the SMCY gene for sex-specific amplification, as both sexes were amplified. This study suggests that further research is needed as there may be undocumented complexities within the region.

The grey squirrel chromosome painting study conducted by Li et al. (2004) showed that the Y-chromosome probe hybridized with other sections of the grey squirrel genome. Li et al. (2004) found this result 'surprising', and suggested that sections of DNA from the squirrel Y-chromosome were replicated across other chromosomes in the grey squirrel genome. The results from this study support this, as the primers designed to target short DNA sequences on the squirrel Y-chromosome may have amplified targets on autosomes, which may help explain why so many of the assays designed in this study were not found to be sex specific. This is the first study to attempt to use real-time PCR for sex identification of the Sciuridae, and it appears that the sensitivity of real-time PCR has shown the difficulties of designing accurate sex identification assays in red and grey squirrels.

It is possible that a number of these assays are not sex-specific in a wider range of fauna as McDevitt (*pers comm*.) also found that the DBY region amplified in female pygmy shrews (*Sorex minutus*). As researchers generally only publish the positive results from their research, it is difficult to assess if the problems surrounding this study are in fact more commonly experienced. As this work has clearly shown ambiguity in the sex specificity of genes currently used for sex differentiation in the Sciuridae, a wider review of molecular sex differentiation and accuracy is needed in the field of conservation genetics, and especially for its application to the Sciuridae. The wider implications of this study have also shown that the use of Y-chromosome markers for evolutionary studies may also have inaccuracies, as certain regions of the Y gene may not be Y-chromosome specific such as the SMCY region where a ~300 bp was found to occurr in both male and female squirrels, thus creating problems for the inferences of the evolution of the parental line of a species.

4.5 Conclusion

- ZFY, SRY and DBY largely proved unsuccessful for the identification of male squirrels
- ZFY and DBY primers designed in this study amplified both sexes
- Some SMCY primers were shown to be suitable for sex differentiation of red squirrels and grey squirrels when amplyfing larger DNA fragments
- Overall, the internal primers designed to target short sections of the squirrel Y-chromosome in this study were found to amplify both sexes, suggesting that the DNA targets were resent on other chromosomes other than the intended Y-chromosome target and requires further investigation

Chapter 5

Non invasive genetic sampling and individual identification of the red squirrel (*Sciurus vulgaris*)

Chapter 5: Non invasive genetic sampling and individual identification of the red squirrel (Sciurus vulgaris)

5.1 Introduction

Hair-tubes have been successfully used to survey red squirrels in Britain and Ireland (Gurnell et al. 2001; Finnegan et al. 2007). The method relies on a PVC tube that has a removable sticky block to capture hairs from red and grey squirrels as they enter the tube. Finnegan et al. (2007) successfully used hair-tubes to assess the distribution of both red and grey squirrels across six sites in both conifer and broad-leaved woodland near Co. Dublin, Ireland. Morphological techniques were to identify the hairs to species (Finnegan et al. 2007). The addition of a DNA approach adapted to work with hair-tube samples to identify the species and individual from hair-tubes could be a useful way to add value to the current survey strategy as it is less labour intensive and the DNA can be used for additional analysis such as population genetics.

The aim of this work was to develop a non-invasive genetic method to census red squirrels. This firstly involved designing a hair-tube to remove hair from squirrels. Secondly, a panel of nine microsatellite loci were chosen to individual red squirrels from hair-tube samples. The technique was then applied to a hair-tube survey.

5.2 Materials and Methods

5.2.1 Study area and sample collection

Hair-tube surveys took place in conifer and mixed woodland sites between 2009 and 2011 in County Waterford, Ireland (Fig 5.1). One grey squirrel site was selected in County Kilkenny (KD). Combined, these sites represented over 600 ha of conifer and mixed woodland sites that were surveyed for the presence of red squirrels. Sites were chosen after an initial assessment of the woodland by looking for evidence of feeding signs including split hazel nuts and stripped pine cones (Fig 1.12; Strachan 1998).

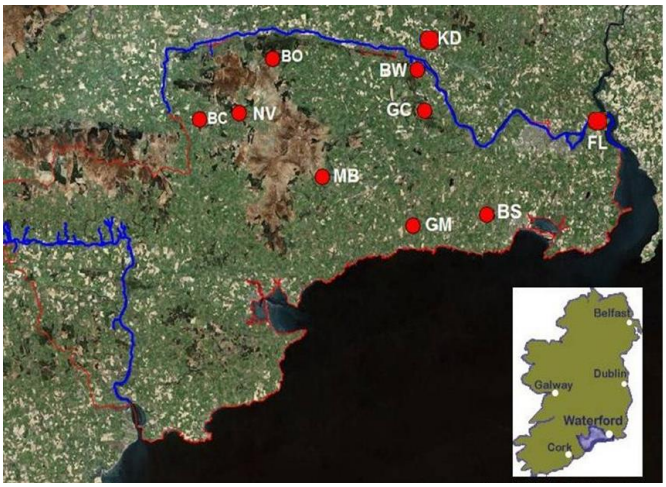

Fig 5.1: Study area in Co. Waterford (South East Ireland) showing the location of selected study sites: Brownswood (BW), Nire Valley (NV), Ballymacabry (BC), Boola Bridge (BO), Guilcagh (GC), Gardenmorris (GM), Mahon Bridge (MB), Ballyscanlon (BS), Faithlegg (FL) and Kildalton (KD). The Rivers Suir (East) and Colligan (West) are also mapped.

Hair Sampling

The hair-tubes were made from 300 mm lengths of 70 mm diameter PVC piping (Fig 5.2). Hair patches were made using corrugated card, based on the design used in Mullins et al. (2010). The card was covered in double sided sticky tape and a small section of mouse glue patch (10 x 14 mm) and was placed on the sticky patch. Patches were positioned inside the tube close to the wire (Fig 5.2 [a]). Patches were placed in both ends of the tubes. Tubes were tied to the tree using wire or cable ties, and

positioned at chest height and a minimum of 30 m apart. At least four baiting sessions took place every 7 – 14 days. Bait consisted of maize, peanuts and hazelnuts. Patches with hair were removed and placed in a 30 ml airtight universal sterile container (Sparks Lab Supplies, Dublin, Ireland), and frozen at -20 °C until DNA extraction. Hair patches remained in the tubes if there was no evidence of hair, and removed hair patches were replaced.

Tubes in the Nire Valley (NV), Ballymacabry (BC) Boola Bridge (BO) were managed by a volunteer (M. Desmond). The volunteer received training to check the tubes, rebait the tubes and was shown how to store and record the samples. Project officers (A. Harrington) from the Mammals in a Sustainable Environment (MISE) provided regular support and went to the field with the volunteer.

A remote camera (Stealth Cam® LLC, TX, USA) was placed on a tree at the Mahon Bridge (MB, Fig 5.1) study site to observe the red squirrels using the hair-tubes. The landowner at Mahon Bridge (MB) (S. Murphy) assisted with the remote camera and regularly checked and moved it to other locations when necessary.

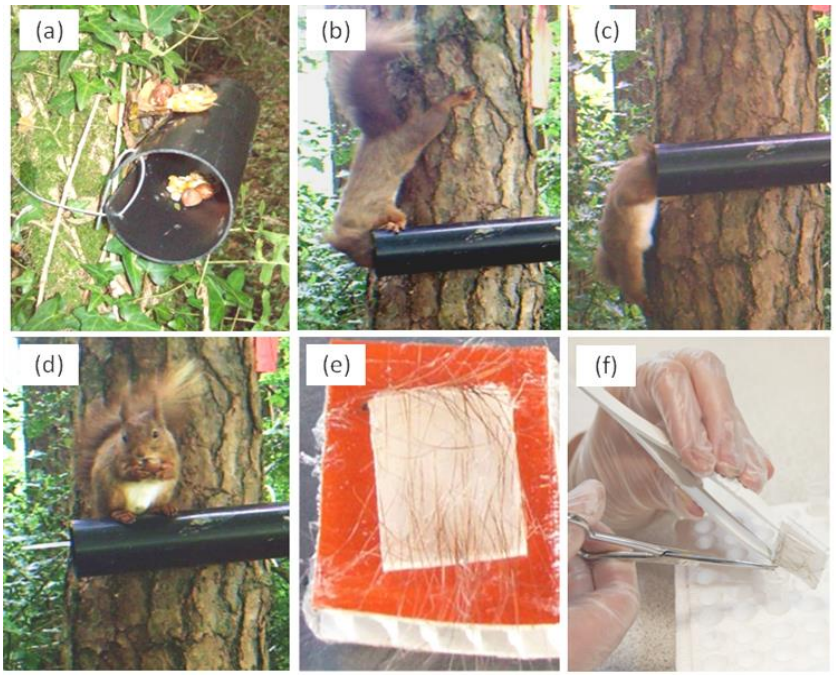

Fig 5.2: **Remote collection of squirrel hair using hair-tubes. (a) Baited hair-tube wired to a tree. Possible entry routes into the tube include (b) above or (c) below. (d) A squirrel that removed the bait is seen sitting on the hair-tube. (e) Hair-patch with hair after being removed from the hair-tube. (f) Removing hair from the hair-patch using forceps and scissors in the lab.**

5.2.2 DNA Protocols

DNA extraction and species identification

DNA was extracted from the hair samples collected in 2010 using the Chelex method as described in Section 2.2.2. DNA was extracted from hair samples collected during 2011 as described in Section 4.2.1. DNA extracts were identified to species following the protocol for red and grey squirrel identification (Section 2.2.4.1). A number of samples that failed to amplify as either squirrel were tested for pine marten DNA using the primers by Mullins et al. (2010) and described in Section 2.3.

Nuclear DNA screening

To test if extracts contained a sufficient quantity of nuclear DNA for microsatellite genotyping, the Custom TaqMan® SNP Genotyping Assay designed to target the ZFX sequence (Fig 4.3) (previous chapter), was used to screen the DNA extractions to assess the quantity and quality of nuclear DNA. Real-time PCR was as described in Section 4.3. Samples that produced a Ct value less than 35 were subsequently selected for downstream genotyping.

5.2.3 Microsatellite analysis

Twelve microsatellite markers (as detailed in Table 5.1) were initially screened for use with red squirrel hair samples. Microsatellites with small amplicons with 2 bp repeat motifs (where possible) were selected for screening. Broquet et al. (2007) showed that shorter motifs had the highest rate of amplication success and a lower

assosociated error rate in comparison to longer motifs. All microsatellites were initially tested with red squirrel tissue samples standardised to 4 ng of DNA per reaction, and amplified in singleplex reactions. The 10 µl PCR reaction consisted of 5 µl GoTaq Hot Start Green Master Mix (Promega), 0.2 µM of each primer, 1 µl (Chelex extracted DNA) to 4 µl (Zymo Kit extracted DNA), in a total volume of 10 µl. The initial PCR conditions consisted of 94 °C for 2 min, 40 cycles of: denaturation at 94 °C for 45 s, annealing at 52–58 °C for 45 s and extension at 72 °C for 110 s, and a final elongation at 72 °C for 10 min, derived from Grill et al. (2009).

The PCR products were diluted (1:20) in water prior to fragment analysis, and 1 µl was added to 15 µl of HiDi Formamide with 0.15 µl size standard (GS500 LIZ™ cat. no. 4322682). Fragment analysis was carried out on an ABI PRISM® 310 Genetic Analyser with 4% polyacrylamide (POP-4® polymer, cat. no. 402838), in a 47 cm x 50 µm capillary, under default run conditions. Alleles were scored using the GeneMapper software version 3.7 (Applied Biosystems).

Table 5.1: Microsatellite loci tested in this study for amplification with red squirrels.

Locus	Source	Target Species	Primer (5'-3')	Repeat Motif	Expected Size Range:
Scv3	Hale et al. (2001a)	*S. vulgaris*	F: TTGGCTCATGGTTTCAGAGA R: CCCCTCACTTCCTCCATTTC	$[GA]_{26}$	134-183
Scv4	Hale et al. (2001a)	*S. vulgaris*	F: CTGGAGATGGAGTGAGTGAGG R: CCAGGAATCCTCTTGAATGC	$[GT]_{23}$	199-215
Scv6	Hale et al. (2001a)	*S. vulgaris*	F: GCAATCCTTGTCCTTGCATT R: TGAGTCATTGGATGAAAACCA	$[TG]_{22}$	185-201
Scv8	Hale et al. (2001a)	*S. vulgaris*	F: TCTTATGGCCAGCCTGTCTT R: GGAATGGAGGTGGTTGGTAA	$[AC]_{17+3+2+5}$	192-200
Scv20	Hale et al. (2001a)	*S. vulgaris*	F: ATAGCCAGGGTTGACACCAG R: CATGTGCCATGGTTTGTGAT	$[TG]_{3+5+11}$	170-190
Scv23	Hale et al. (2001a)	*S. vulgaris*	F: AAACACCTGAGACAGGCAAC R: GTGTTTGGCAATGTCAGCTC	$[GT]_{10}$	151-167
Scv31	Hale et al. (2001a)	*S. vulgaris*	F: CCAAGTTCCAGACCAACCTC R: TCGGGTCTCTAAGGAGATGG	$[AG]_{29}$	179-201
Rsu1	Todd (2000)	*S. vulgaris*	F: CTGGGTTCACTGACTTCTCC R: CACTCTCAGAGGCCAAAGTC	$[GGAT]_{13}$	172-196
Rsu4	Todd (2000)	*S. vulgaris*	F: CAATCCTCCCATCCTGCTGC R: TAGGCAGTCAGATAGGTGGG	$[ATCC]_{21}$	256-284
Rsu5	Todd (2000)	*S. vulgaris*	F: CCCAGTCTACATTAAAGGGC R: GCCTATACACTATAATTGACTG	$[GT]_{10}$	123-143
Lis3	Shibata et al. (2003)	*S. lis*	F:CCAACAGTTGCTGTCGCTCC R: GTTTCTTCAAATCCTGAACCGGCCCC	$[GT]_{11}$	140-160
Lis12	Shibata et al. (2003)	*S. lis*	F :GTTAACCTGTGCTAGCCATGGAAC R: GTTTCTTGTGGTCCATCTCCATAGACTTCC	$[CT]_{21}$	155-211

Microsatellite Data analysis of non-invasively collected hair samples

To control and assess the level of genotyping errors, the software GIMLET, version 1.3.4 (Valière 2002) was used to assess the replicated data for the presence of errors including the presence of allelic drop out and false alleles. GIMLET was also used to calculate the percentage of positive PCRs in the dataset. Secondly, MICRO-CHECKER, version 2.2.3 (van Oosterhout et al. 2004) was used to identify possible genotyping errors, including the presence of null alleles, large allele dropout and scoring errors due to stutter peaks (using default settings). Unique genotypes were identified with GENALEX, version 6.5b (Peakall and Smouse 2006). GENALEX was also used to assess

for deviations from Hardy–Weinberg equilibrium (HWE), observed (H_O) and expected (H_E) heterozygosities, the number of alleles (A) and range of allele size, and was also used to estimate the probability of identity (PID).

5.2.4 Mitochondrial DNA analysis

The primers SvSc_F and Svul_R (Section 2.3.2) were used to amplify a 350 bp of the mitochondrial D-Loop DNA from a subset of 12 individuals selected across the sites. PCR, gel electrophoresis and DNA sequencing was as described in Section 2.2.4.2 . Nucleotide sequences were analysed using the BLAST software at http://www.ncbi.nlm.nih.gov/BLAST/ from the GenBank (NCBI) database (Altschul et al. 1990). Sequences were compared by multiple alignments in MEGA V.5.05 (Tamura et al. 2011) and haplotypes were identified using ARLEQUIN (version 3.5; Excoffier and Lischer 2010).

5.3 Results

5.3.1 Non-invasive genetic identification of squirrels

Fig 5.3 outlines the general strategy that was followed for identification of red squirrels from hair samples.

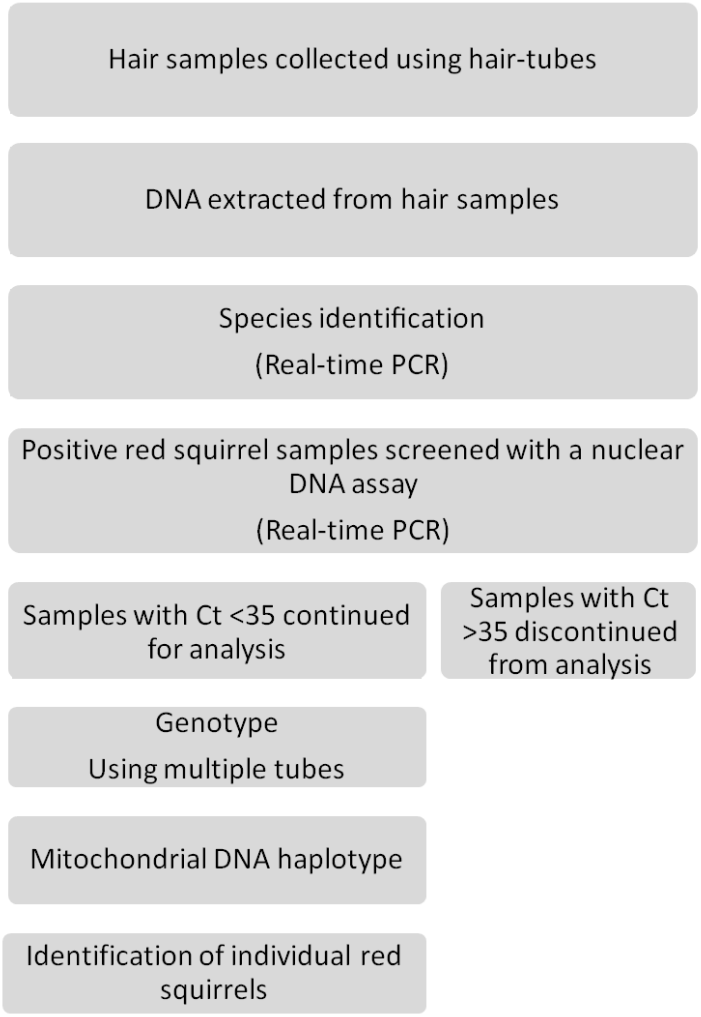

Fig 5.3: Flowchart diagram showing the various stages involved in the identification of red squirrels from hair samples.

Hair-patches that contained hair (n = 110) were removed from hair-tubes across the study area and the number of hairs found on the hair patches varied from 1-2 hairs to over 30. The number of patches removed per site ranged from one at Boola Bridge to 21 at Gardenmorris, Brownswood and Ballyscanlon (Table 5.2).

There was a significant increase in the number of samples collected across all sites as the number of visits and rebaiting sessions increased (R^2 = 0.97, P > 0.001). The trail camera positioned at the site in Mahon Bridge provided video footage of red squirrels entering the hair-tubes and removing the bait. Videos recorded from the site revealed that red squirrels shake the hair tubes to remove the bait, and attempted to avoid entering the tube, thus reducing the potential for contact between the hair patch and the animal (Fig 5.4).

Fig 5.4: Video stills of a red squirrel at Mahon Bridge using a hair tube. The squirrel was recorded shaking the tube to attempt to remove the bait without entering the tube.

Species Identification

The real-time PCR species identification test revealed that 86 of the 110 patches contained red squirrel DNA, with average real-time PCR Ct values per site ranging 21.5 ± 1.27 at Ballymacabry to 34.42 ± 2.26 at Faithlegg. Eight samples from Gardenmorris were identified as pine marten (28.41 ± 1.60), and four samples from Kildalton were identified as grey squirrel (22.46 ± 2.51). The samples that failed to amplify as red squirrel, grey squirrel or pine marten (n = 12) could not be identified to species as those samples contained a non-target species or contained DNA from a non-target species, and were discarded from further analysis (Table 5.2 (a)). Most of those samples had only 1-2 hairs, demonstrating that patches with fewer than 5-10 hairs are generally unsuitable for DNA analysis. The positive red squirrel samples were subsequently amplified with a nuclear DNA assay to assess the quantity and quality of nuclear DNA per sample. The samples with Ct < 35 (Table 5.2 (b)) were subsequently selected for amplification with the microsatellite panel. This Ct value was used as a screening step to select samples with good quality and quantity DNA.

Table 5.2: (a) Summary of real-time PCR results for species identification of red squirrel (RS), grey squirrel (GS), pine marten (PM) and non identified samples (ND). (b) Real-time PCR results for positive red squirrel samples amplified with the ZFX assay.

(a)	Total hair patches	RS Ct		PM Ct		GS Ct		ND
Site	**N**	**N**		**N**		**N**		**N**
GM	21	10	29.62 ± 4.95	8	28.41 ± 1.60	0		3
BS	21	19	25.25 ± 3.70	0		0		2
BW	21	16	26.46 ± 3.38	0		0		5
GC	3	3	26.26 ± 5.17	0		0		0
KD	4	0				4	22.46 ± 2.51	0
NV	16	16	31.25 ± 2.93	0		0		0
FL	4	2	34.42 ± 2.26			0		2
MB	17	17	21.87 ± 5.16	0		0		0
BO	1	1	22.51	0		0		0
BC	2	2	21.5 ± 1.27	0		0		0
Totals	**110**	**86**		**8**		**4**		**12**

(b)		ZFX Ct
Site	**N**	**Ct < 35**
GM	3	32.79 ± 0.44
BS	12	32.56 ± 1.52
BW	12	32.38 ± 1.34
GC	2	32.32 ± 1.34
KD	0	
NV	16	31.01 ± 3.91
FL	0	
MB	11	31.40 ± 1.43
BO	1	30.45
BC	2	32.10 ± 0.66
Total	**58**	

GM Gardenmorris, *BS* Ballyscanlon, *BW* Brownswood, *GC* Guilcagh, *KD* Kildalton, *NV* Nire Valley, *FL* Faithlegg, *MB* Mahon Bridge, *BO* Boola Bridge, *BC* Ballymacabry. *Neg samples* included those that failed to amplify as red squirrel due to low quality DNA.

5.3.2 Microsatellite Optimisation

Amplification

Twelve loci were screened for optimisation with the red squirrel. Attempts were made to optimise the PCR conditions for each locus by varying the annealing conditions between 52 °C and 58 °C. The PCR products were initially visualised on a 2% agarose gel, where it was observed that the annealing temperature varied across the loci. In an effort to optimise the microsatellite panel to amplify using the same PCR profile, different annealing temperatures were used Fig 5.5. Rsu1 and Rsu4 were difficult to amplify and SCV23 failed to amplify using a variety of annealing temperatures.

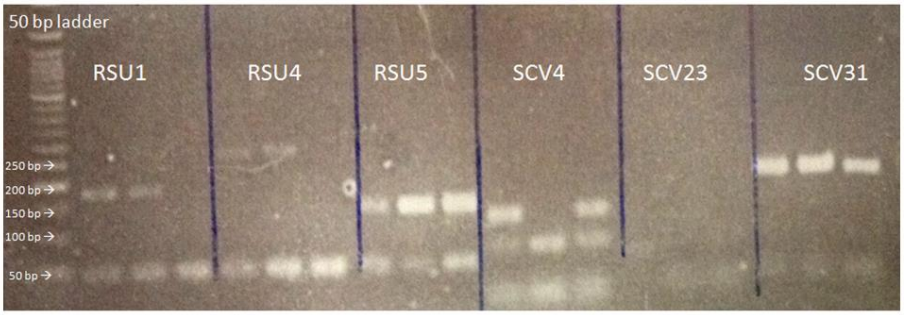

Fig 5.5: Example of amplification of six microsatellites using a 58°C annealing temperature using three DNA hair samples.

Two microsatellites that were previously developed by Shibata et al. (2006) for amplification with *Sciurus lis* were also tested for cross-utility with the red squirrel in this study. Both markers amplified well with red squirrel DNA and amplified across a range of annealing temperatures (Fig 5.6).

Fig 5.6: Amplification of Scv3 (Hale et al. 2001a) Lis3 and Lis12 (Shibata et al. 2006) with red squirrel hair DNA.

A dropdown PCR protocol adapted from Hellborg and Ellegren (2003) was finally used to amplify the majority of microsatellites in the panel. This involved one cycle of 95 °C for 10 min, followed by 20 cycles of 95 °C for 30 s and a touchdown from 60 °C to 55 °C for 1 minute decreasing by 0.5 °C per cycle, and then 72 °C for 1.50 minutes. This was followed by 20 cycles of 95 °C for 30 s, 45 °C for 1 minute and 72 °C for 1.5 minutes, and a final extension of 72 °C for 10 minutes. This protocol was subsequently used as the majority of the loci amplified under these conditions. Scv23, failed to amplify at any of the PCR conditions and was discontinued from further analysis. Rsu1 and Rsu4 produced multiple PCR products when they were amplified using the touchdown PCR protocol and were discontinued from further analysis.

Design of mini-microsatellites

Where possible, the reverse primers were redesigned to sit closer to the repeat region. This was done to reduce the overall size of the amplicon and to increase the specificity of the primer pair to work in conjunction with non-invasive samples. The sequences were downloaded from Genbank for all the microsatellites used, and the reverse primers were redesigned. Reverse primers were selected by designing a new primer that had a similar melting temperature to the forward primer. Mini microsatellites were successfully designed for Scv4, Scv31, and Scv8 and (Fig 5.7; Table 5.3). The original allele sizes for those primers were Scv4 (199-215); Scv31 (128-208) and Scv8 (192-200) (Hale et al. 2001a). After redesign in this study, the product sizes were reduced to the follow allele size range Scv4 (114-148); Scv31 (162-166); and Scv8 (162-166).

>gi|12958255|gb|AF309452.1| Sciurus vulgaris clone Scv4 microsatellite sequence
AAAATACCTGGAGATGGAGTGAGTGAGGAGTGAGTGTGTATATGTATGTGTGTGTGTGTGTGTGTGTG
F-Primer (Hale et al. 2001a)
TGTGTGTGTGTGTGTGTGTGTGTATTCCCAAATGATTCCTGTGCAAACTGAAGTTTGTTACTCATTAATC
Mini-R-Primer (This Study)
TAAGGAGTTGAAACATTTCCATCTGCACTATTTGGATAGTCATCTACAGATGAGACAAAATTGCTTGCAT
TCAAGAGGATTCCTGGATC
R-Primer (Hale et al. 2001a)

Fig 5.7: Scv4 microsatellite sequence showing the position of the original forward and reverse primers (grey colour) from Hale et al. (2001a). The reverse primer was redesigned in this study (green colour) to sit next to the repeat region (TG). This reduced the amplicon from the original 214 bp to 111 bp.

Multiplex PCR

The nine microsatellite panel were successfully combined into three groups for multiplex PCR (Table 5.3). The programme AutoDimer (Vallone and Butler 2004) showed that there was some risk for the formation of primer dimer and hairpin structures if Scv3 and Scv4 were placed in the same multiplex reaction (Fig 5.8).

```
3'NGGAGTGAGTGAGGTAGAGGTC-5'
  x|||||||xx||||||x|||
5'CCCCTCACTTCCTCCATTTC-3'
```

Fig 5.8: Illustrating the possibility of Scv3 and Scv4 binding to one another in a multiplex reaction.

Preliminary multiplex reactions contained equal quantities of primers (Fig 5.9). By looking at the intensity of the PCR bands, primer concentrations were varied to allow equal amplification of the products. This was done through a process of increasing or decreasing the concentration of one of the microsatellites in the multiplex (Fig 5.10). The final primer concentration primer labels are in Table 5.3.

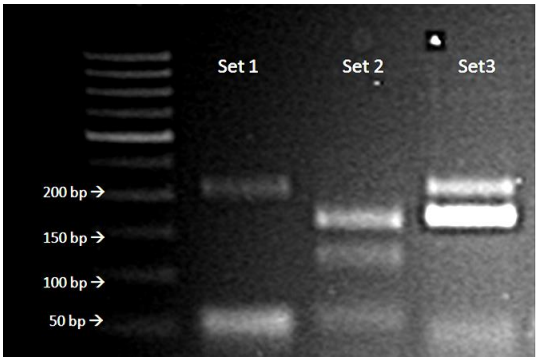

Fig 5.9: Optimisation of microsatellite multiplexes. Primer concentrations were increased/decreased based on band intensity. Set 1 – 3 are outlined in Table 5.3.

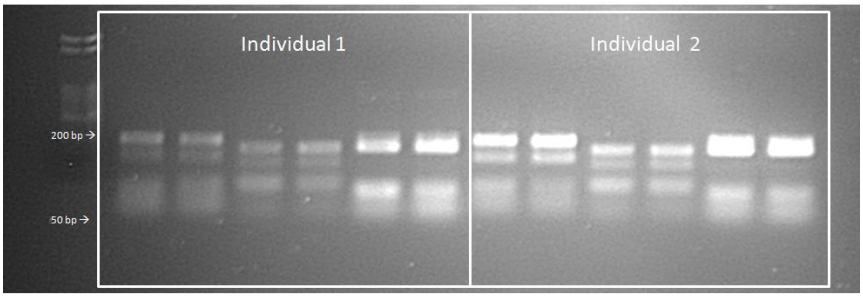

Fig 5.10: An example of the optimised multiplex PCR for two individual hair DNA samples. Both individuals were amplified with the three multiplex reactions in duplicate.

The PCR products were then diluted (1 µl of pooled PCR mix into 19 µl of water) prior to fragment analysis (Fig 5.11). The three multiplexes were later pooled for combined analysis to save resources

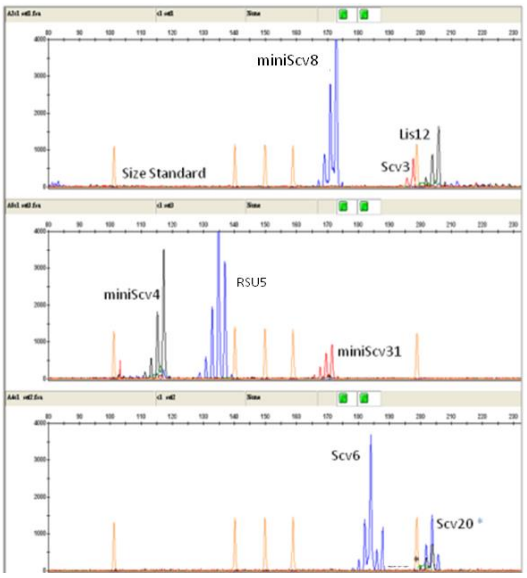

Fig 5.11: Microsatellite loci from the three multiplexes. Lis3 is missing from the third multiplex.

The success rate for PCR amplification with non-invasive samples ranged from 0.97 with miniScv4 to 0.67 with Scv20, the smallest and largest sized amplicons in this study respectively.

Table 5.3: Final microsatellite panel including redesigned reverse 'mini' primers. The microsatellites were divided into three multiplex reactions. The final primer concentration (conc µM) and the labels are also provided.

Locus	Source	Primer (5'-3')	Multiplex	Conc µm	Label
Scv3	Hale et al. (2001a)	F: TTGGCTCATGGTTTCAGAGA R: GTTTCTTCCCCTCACTTCCTCCATTTC	1	0.2	FAM
Lis12	Shibata et al. (2006)	F: GTTAACCTGTGCTAGCCATGGAAC R: GTTTCTTGTGGTCCATCTCCATAGACTTCC	1	0.3	ATTO 550
Rsu5	Todd et al. (2000)	F: CCCAGTCTACATTAAAGGGC R: GCCTATACACTATAATTGACTG	1	0.18	FAM
miniScv4	Hale et al. (2001a)	F: CTGGAGATGGAGTGAGTGAGG miniR: GTTTCTTGCACAGGAATCATTTGGGA	2	0.15	ATTO 551
miniScv31	Hale et al. (2001a)	F: CCAAGTTCCAGACCAACCTC miniR: GTTTCTTGAACAGAAGTGATATGAGGCCAG	2	0.15	ATTO 550
miniScv8	Hale et al. (2001a)	F: TCTTATGGCCAGCCTGTCTT miniR: TGTGTTTGGGTGTCTGTGTGTT	2	0.2	FAM
Scv6	Hale et al. (2001a)	F: GCAATCCTTGTCCTTGCATT R: ATGTAAAGTCTCAGGGTAGGAGGA	3	0.25	FAM
Scv20	Hale et al. (2001a)	F:ATAGCCAGGGTTGACACCAG R: CATGTGCCATGGTTTGTGAT	3	0.3	FAM
Lis3	Shibata et al. (2006)	F: CCAACAGTTGCTGTCGCTCC R: GTTTCTTCAAATCCTGAACCGGCCCC	3	0.1	ATTO 551

5.3.3 Red squirrels in Co. Waterford

In total, 30 individual red squirrels were non-invasively identified in this study Table 5.4. The number of individual red squirrels identified in Co. Waterford ranged from nine in Ballyscanlon to one each in the Nire Valley and Boola Bridge. Replicated genotypes were recorded at four sites with the number of replicated genotypes ranging from one recapture to 16 recaptures.

Low levels of allelic dropout were identified at RSu5 and miniScv20 (0.01 each). The overall dropout rate across all loci and PCRs was 0.03. No incidences of false alleles were recorded. PCR success ranged from 0.67 at miniScv20 to 0.97 at miniScv4, with an average of 0.85 across all loci. Summary statistics are provided in Table 5.5. The number of alleles per locus ranged from two at Rsu5 to four at Lis12, miniScv31, miniScv8, Scv6, Scv20 and Lis3, and averaged 3.6 alleles per loci. Observed heterozygosity ranged from 0.286 at Lis12 to 0.800 at Scv6, and averaged 0.532 across all loci. Expected levels of heterozygosity ranged from 0.312 at Lis12 to 0.588

at Scv6, with an average of 0.482 across all loci. Two loci Scv6 and miniScv31 significantly deviated from Hardy-Weinberg equilibrium (P = 0.026). Probability of identity averaged 3.5 x 10^{-1} across loci, with a cumulative PI of 4.7 x 10^{-5}. Probability of identity for related individuals averaged 6.0 x 10^{-1}, with a cumulative PIsib of 8.8 x 10^{-3}.

Table 5.4: Microsatellite genotypes of the unique individuals identified from each woodland Boola Bridge (BO), Ballymacabry (BC), Brownswood (BR), Ballyscanlon (Bs), Gardenmorris (GM) Guilcagh (GC), Mahon Bridge (MB) and Nire Valley (NV). The loci are displayed horizontally and the individual samples are displayed vertically.

Loci	Rsu5	Lis12	Scv3	miniScv4	miniScv31	miniScv8	Scv6	Scv20	Lis3	Rep
B01	136 / 138	205 / 205	185 / 185	116 / 116	167 / 167	162 / 164	176 / 178	213 / 219	175 / 175	1
BC1	136 / 138	205 / 207	185 / 185	0 / 0	167 / 167	162 / 164	176 / 178	213 / 213	173 / 175	1
BC2	136 / 136	203 / 205	185 / 185	0 / 0	167 / 173	164 / 164	176 / 178	213 / 213	173 / 175	1
Br1	136 / 138	205 / 205	185 / 185	114 / 116	0 / 0	166 / 168	176 / 178	213 / 213	173 / 175	1
Br2	136 / 136	205 / 207	185 / 189	114 / 116	0 / 0	166 / 168	176 / 180	213 / 215	173 / 175	2
Br3	136 / 138	203 / 205	185 / 185	114 / 116	173 / 173	164 / 166	176 / 178	213 / 213	173 / 175	2
Br4	136 / 138	205 / 205	185 / 189	114 / 116	173 / 173	166 / 168	176 / 178	213 / 213	173 / 175	3
Br5	136 / 138	205 / 205	185 / 185	114 / 116	173 / 173	0 / 0	176 / 178	213 / 213	173 / 173	1
Br6	136 / 138	205 / 205	185 / 185	114 / 116	173 / 173	166 / 168	176 / 178	213 / 215	173 / 175	1
Br7	136 / 138	205 / 205	185 / 185	114 / 116	0 / 0	166 / 168	178 / 180	213 / 213	173 / 173	1
By1	0 / 0	205 / 205	185 / 189	116 / 118	0 / 0	164 / 164	178 / 178	213 / 215	175 / 175	1
By2	136 / 138	205 / 205	185 / 189	116 / 116	173 / 173	162 / 164	178 / 178	213 / 213	175 / 175	1
By3	136 / 138	203 / 203	185 / 189	116 / 116	173 / 173	164 / 164	176 / 178	213 / 213	175 / 175	1
By4	136 / 138	205 / 205	185 / 187	116 / 116	173 / 173	162 / 164	176 / 178	213 / 215	173 / 175	1
By5	138 / 138	205 / 205	185 / 189	116 / 116	171 / 173	164 / 164	176 / 178	213 / 215	173 / 175	1
By6	136 / 138	205 / 205	185 / 189	116 / 116	0 / 0	164 / 164	178 / 178	213 / 215	175 / 175	1
By7	136 / 138	205 / 205	185 / 189	116 / 116	0 / 0	164 / 164	176 / 178	213 / 213	173 / 175	1
By8	136 / 138	0 / 0	185 / 189	0 / 0	0 / 0	164 / 164	176 / 178	213 / 215	175 / 175	1
By9	138 / 138	205 / 209	185 / 185	114 / 116	173 / 173	162 / 164	180 / 180	213 / 213	175 / 179	3
G1	136 / 138	205 / 205	185 / 187	116 / 116	173 / 175	162 / 164	176 / 178	213 / 213	173 / 175	1
G2	136 / 138	205 / 205	185 / 187	114 / 116	171 / 173	162 / 164	176 / 178	213 / 213	173 / 175	1
Gu1	136 / 138	205 / 205	0 / 0	114 / 116	173 / 173	162 / 166	176 / 180	213 / 215	173 / 179	1
Gu2	138 / 138	203 / 205	185 / 185	114 / 116	173 / 175	162 / 164	176 / 178	213 / 213	173 / 175	1
MB1	138 / 138	205 / 207	185 / 185	116 / 116	173 / 173	162 / 164	176 / 178	215 / 217	173 / 175	2
MB2	138 / 138	205 / 205	185 / 185	116 / 116	173 / 173	162 / 162	176 / 178	215 / 215	171 / 175	3
MB3	138 / 138	205 / 205	185 / 185	114 / 114	173 / 173	162 / 166	176 / 178	213 / 215	173 / 175	1
MB4	138 / 138	205 / 205	189 / 189	116 / 116	173 / 175	0 / 0	172 / 176	0 / 0	173 / 175	2
MB5	138 / 138	0 / 0	185 / 189	114 / 114	171 / 173	168 / 168	176 / 176	213 / 215	175 / 175	1
MB6	136 / 136	205 / 205	185 / 185	114 / 116	0 / 0	164 / 164	176 / 176	213 / 215	171 / 175	2
NV1	136 / 136	205 / 207	185 / 185	116 / 116	171 / 173	162 / 164	176 / 178	213 / 215	171 / 175	16

Table 5.5 Descriptive statistics for the red squirrels identified in Co. Waterford. Number of samples amplified per locus (N), number of alleles per locus (A), observed heterozygosity (H_O), expected heterozygosity (H_E), Hardy-Weinberg expectations (HWE), probability of identity per locus (PI) and probability of identity for siblings (PIsib). * denotes deviations from Hardy-Weinberg equilibrium.

	Rsu5	Lis12	Scv3	miniScv4	miniscv31	miniscv8	Scv6	Scv20	Lis3	Average
N	29	28	29	27	22	28	30	29	30	**28.0**
A	2	4	3	3	4	4	4	4	4	**3.6**
H_O	0.586	0.286	0.448	0.481	0.364	0.643	0.800	0.483	0.700	**0.532**
H_E	0.490	0.312	0.405	0.442	0.445	0.669	0.588	0.433	0.553	**0.482**
HWE	0.293	0.695	0.751	0.931	**0.026***	0.456	**0.026***	0.690	0.507	
PI	3.8×10^{-1}	4.9×10^{-1}	$4. \times 10^{-1}$	4.0×10^{-1}	3.4×10^{-1}	1.6×10^{-1}	2.6×10^{-1}	3.9×10^{-1}	2.8×10^{-1}	**3.5×10^{-1}**
Pisib	6.0×10^{-1}	7.2×10^{-1}	6.5×10^{-1}	6.4×10^{-1}	6.1×10^{-1}	4.6×10^{-1}	5.2×10^{-1}	6.3×10^{-1}	5.4×10^{-1}	**6.0×10^{-1}**

5.3.4 Mitochondrial DNA results

Three mitochondrial DNA haplotypes were found in the Co. Waterford study area from the 12 samples used to generate sequences. No sequence was generated for the Gardenmorrs site. One haplotype (IE16) was previously published by Finnegan et al. (2008), and was recorded at four sites in this study. Two novel haplotypes were found in the county, with IE23 found at three sites, and IE24 found at one site (Fig 5.12). IE24 was resequenced for verification. The new names follow on from the haplotypes named by Finnegan et al. (2008). The relationship of these haplotypes will be investigated in Chapter 6.

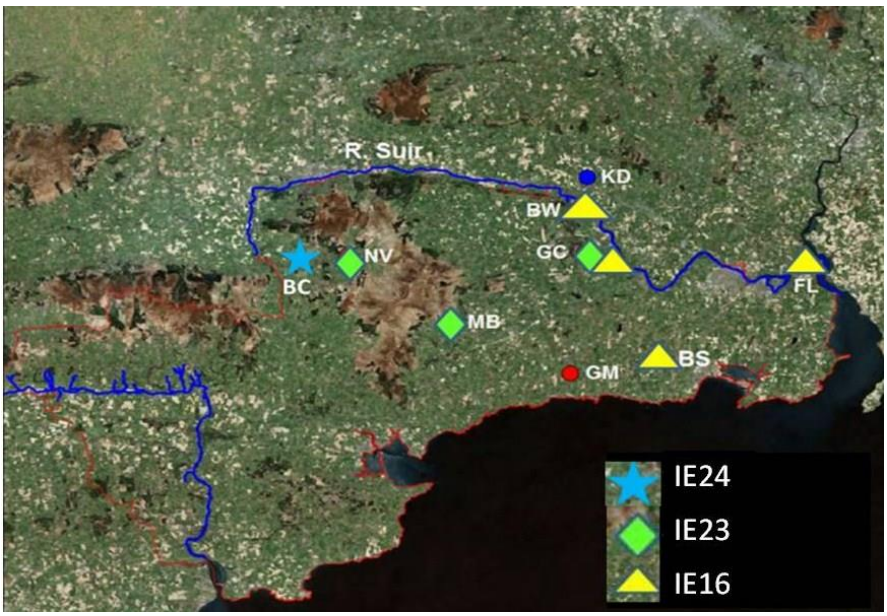

Fig 5.12: Map of haplotype distribution of remotely collected red squirrel individuals in Co. Waterford. ***Three individuals from Mahon Bridge had the IE23 haplotype.***

5.4 Discussion

5.4.1 Hair-tubes for the non-invasive genetic study of red squirrels

This is the first study that used hair-tube samples collected from red squirrels for genetic analysis. Trizio et al. (2005) trapped individual squirrels and successfully used plucked hair for subsequent analysis. Trizio et al. (2005) were able to ensure that at least 30 hairs were plucked per animal. In this current study, it was not possible to control the number of hairs that were removed from each animal, and subsequently some of the samples had as few as one to two hairs. An additional problem encountered with the hair-tubes was that squirrels took a period of time (latency) to use the tubes. Using trail cameras it was observed that the squirrels tended to avoid entering the tubes by shaking the food out. This suggests that further optimisation of the hair-tube design is needed. Adjustments might include securing the bait inside the tube, or the use of a mesh tube that squirrels might be more likely to enter.

The sticky patch design used in this study was initially developed by Mullins et al. (2010) and was originally designed to collect hairs from pine marten for DNA analysis. During the course of this study it was found that the hair collected from the red squirrel was shorter and finer than that of the pine marten, and less hair was generally collected. The pine marten tubes were baited with chicken that was wired into the tubes, requiring significant effort on the part of the pine marten to remove the chicken, causing the pine marten to push against the tube which ultimately resulted in more hair being removed from the pine maten. Adapting the tubes to increase the surface area that squirrel hairs could be captured from may also help to overcome this issue such as the sticky tape snare developed by Henry et al. (2011) that was successfully used to survey pikas (*Ochotona princeps*) and remotely remove hairs for genetic analysis (Henry et al. 2012) (Fig 5.13). The sticky tape hair snare could be used at the entrance of the tube in addition to the use of a hair patch to capture more hairs from red squirrels. This may allow additional root follicles to be remotely plucked from red squirrels, as the number of root follicles has previously been shown to be important for genotyping (Trizio et al. 2005).

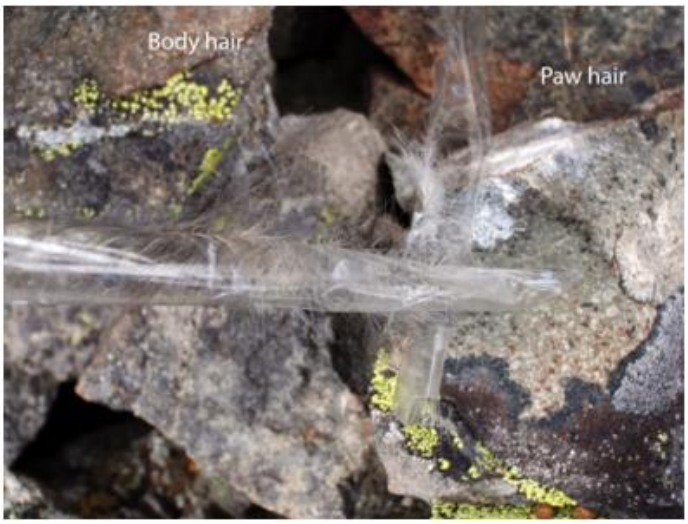

Fig 5.13: A sticky tape snare (containing pika hair) at the entrance of a pika burrow (Henry et al. 2011).

The labour involved in non-invasive studies also needs to be considered. This study required up to four baiting periods in order to obtain hair samples. Trapping is more labour intensive, but it may be more worthwhile in terms of gaining an adequate estimate of the red squirrel population in a particular study site. However, some of the benefits of hair-tubes are that there is no requirement for licensing and volunteers can be easily trained to conduct hair-tube studies. This type of approach is also beneficial to the squirrel as it avoids disturbing the animals, or encouraging some individuals to become 'trap happy'.

5.4.2 Non-invasive genetic identification of red squirrels

DNA Extraction procedures

In this study, two types of DNA extractions were used. The results showed that the DNA extraction kit provided DNA extracts with a higher success rate that (a) passed the screening stage and were taken to the genotyping step, (b) provided a genotype. DNA extraction kits should be used in future studies, especially when the starting template may be poor in comparison to other species. It should be noted that by using the DNA extraction kit, we were able to remove the PCR inhibitors in the DNA

extraction and this allowed us to add four times the starting template in comparison to the Chelex extraction.

Microsatellite Optimisation

This is the first study to optimise microsatellite primers for use with remotely obtained hair samples from squirrels. Two microsatellite primers originally developed for *Sciurus lis*, and both loci were included in the microsatellite panel in this study. Both markers produced slightly lower levels of variability, but were generally easy to score and worked well in the multiplex set up, making them a useful addition to the microsatellite panel.

The genotyping error rate reported in this study is on the lower end of estimates reported in a review of non-invasive studies by Broquet et al. (2007). However, more recent studies such as Henry et al. (2012) also reported low levels of errors, possibly as a result of advancements in the use of DNA extraction kits, as evidenced in this study, but also the additional use of a novel screening stage to select the best DNAs. The approach taken in this study, where the results from the ZFX assay were used as a screening protocol was quite strict, and this may have helped improve the results. A similar approach was taken in non-invasive studies for pine marten and otters (Mullins et al. 2010; O'Neill et al. 2013; Sheehy et al. 2014).

The final microsatellite panel used in this study had enough statistical power to identify individuals and differentiate siblings, as the probability of identification was lower than the 0.001 threshold described by Mills et al. (2000). Two loci (miniScv31 and Scv6) showed some evidence for deviations from Hardy-Weinberg equilibrium (P = 0.03). Levels of genetic diversity indicated that the population of red squirrels in Co. Waterford appear relatively healthy, with low to medium levels of heterozygosity and genetic variability across the population. In the case of miniScv31, there was a higher level of expected heterozygosity compared to observed heterozygosity, while in the case of Scv6, there was a higher level of observed heterozygosity compared to the expected. In the case of miniScv31, the presence of null alleles in the data may have caused the deviation. Inbreeding can also cause deviations from Hardy-

Weinberg equilibrium, as can the presence of population substructure. The later is possible in this case, as three different mitochondrial haplotypes were found in this population. IE16 and IE23 were closely related, with only 1 bp difference. This haplotype was previously found in Limerick and Galway (Finnegan et al. 2008), but IE24 was more closely related to IE14, previously recorded in Co. Wexford. This haplotype was only recorded in the North of the Comeragh Mountains, and is also closely related to haplotypes found in Wales (Ogden et al. 2005), Jersey, the Isle of Wight (Simpson et al. 2013) and France (Dozieres et al. 2012). This will be discussed further in Chapter 6.

There was no evidence of grey squirrels in any of the Co. Waterford woodlands samples, although grey squirrels do occur in Kildalton, Co. Kilkenny. Both pine marten and red squirrels were identified from hair-tubes at Gardenmorris, Co. Waterford. Recent studies by Sheehy et al. (2013) and Sheehy (2013) have shown that the presence of pine marten may have a positive impact on the red squirrel population, as grey squirrels are less likely to occur in high numbers where pine marten are abundant. Previous work by Mullins et al. (2010) and on-going work at WIT have shown that pine marten also occur in other sites sampled in this study including Brownswood, Guilcagh and Faithlegg. The occurrence of pine marten in the Nire Valley, Boola Bridge and Ballymacabry has not been investigated. Future studies that intend to use hair-tubes to survey red squirrels should consider the presence of pine marten as it has been shown in this work and in Chapter 2 that pine marten will use red squirrel hair-tubes.

5.5 Conclusion

- This study has successfully designed a suitable method to detect, identify and assess the genetic diversity of a red squirrel population using non-invasive genetic methods.
- The techniques are suitable for use with the hair-tube collected samples, opportunistically collected road kill samples and hair plucked from squirrels during trapping studies.

- Such techniques have useful applications across Ireland and Britain for conservation planning, management and mitigation, and have useful applications for red squirrel volunteer groups, as well as conservation managers.
- Using these techniques, both remnant squirrel populations and potential reintroduction stock can be firstly genetically screened for their suitability as reintroduction or translocation stock.

Chapter 6:

Tracing the genetic heritage of the red squirrel (*Sciurus vulgaris*) in Ireland

Introduction 6.1

The red squirrel in Ireland has a relatively recent history in the country since it was reintroduced from Britain in the 1800s on multiple occasions (Section 1.2.5; Barrington 1880). The British stock is also quite mixed as translocations and documented reintroductions took place throughout Scotland, Wales and England (Shorten 1954). Some of the stock used to repopulate Britain included animals translocated from Scandinavia, and Continental European from countries such as France (Shorten 1954).

Finnegan et al. (2008) found a high number of mitochondrial D-loop haplotypes (n = 26) in the Irish population, most of which were unique to Ireland. Exceptions included H15, a haplotype that had been previously recorded by Hale et al (2004) and was believed to have originated in central Europe, and two haplotypes that were recorded in museum samples from both Britain and Ireland. Many of the contemporary DNA haplotypes recorded in Ireland appear to be closely related to the haplotypes shared between museum samples in Britain and Ireland, and this led Finnegan et al. (2008) to believe that the Irish red squirrels were predominantly composed of translocated individuals from Britain.

Despite a number of genetic studies of red squirrels in Ireland, Britain, Jersey, Isle of Wight, France, Italy and across Europe, we still do not fully understand the phylogeographic origins of the species within Europe (Barratt et al. 1998; Hale et al. 2001; Hale et al. 2004; Ogden et al. 2005; Trizio et al. 2005; Finnegan et al. 2008; Grill et al. 2009; Dozierèz et al. 2012; Simpson et al. 2013). Although, the low divergence rate between red squirrels from Europe, Russia and China, suggest a historical migration from the East to the West, and only the Southern Italian red squirrels exhibit genetic differentiation (Grill et al. 2009). The previous studies have found that the red squirrel is genetically quite diverse, with high levels of admixture or mixing occurring across their distribution, some of which may be attributed to human translocations (Section 1.2).

Using samples from Chapter 5, and additional samples collected in Ireland, the aim of this chapter is to use microsatellite data to investigate the levels of genetic diversity, gene flow, and genetic structure to assess if a contemporary genetic signature of small, random, human mediated red squirrel introductions from multiple locations in the 1800s is still evident in the Irish red squirrel population today. Secondly, given the proposed relatively recent history of red squirrels in Ireland, inferences will be made about the origins of red squirrels in each area based on comparisons of mitochondrial DNA variability between Irish, British and central European red squirrels. The combination of these factors will then be discussed in relation to the genetic history, and conservation management of the red squirrel in Ireland.

6.2 Materials and Methods

6.2.1 Sample collection

Samples used in this study are provided in Table 6.1. The Wexford samples came from red squirrels translocated to Killiney, Co. Dublin for a reintroduction project, and the Galway samples also originated from a translocated population in Portumna, Co. Galway, but have been breeding in the site for a number of years (Poole and Lawton 2009). The additional samples were used for statistical analysis and to increase the geographic spread of the samples. All samples were screened with nine microsatellite loci as described in Section 5.3.2 using the final microsatellite panel in Table 5.3.

Table 6.1: Hair sample collection details.

Location	Collectors	Sample Type	No
Waterford	D. O'Meara, P. Turner, A. Harrington, M. Desmond	Hair-tube	30
Midlands (Laois + Offaly)	E. Sheehy	Hair-tube	3
Wexford, the Raven	W. Carr	Plucked hair	14
Galway (Derryclare)	C. Waters + C. Lawton	Plucked hair	11
Northern Ireland	D. Tosh	Plucked hair	11
Dublin (Carrickgollogan)	S. Rubalcava	Hair-tube	4
Cork (Glengarriff)	P. Sleeman	Plucked hair	4
		Total:	**77**

6.2.2 Microsatellite data analysis

To account for genotyping errors, GIMLET version 1.3.4 (Valière 2002) was used to assess the replicated data for the presence of errors and to calculate the percentage of positive PCRs in the dataset. MICRO-CHECKER version 2.2.3 (van Oosterhout et al. 2004) was used to identify possible genotyping errors including the presence of null alleles (an allele that is present in a sample, but did not amplify), large allele dropout (one or more alleles are not present) and scoring errors due to stutter peak (using default settings), as errors present in the dataset would affect subsequent analysis.

Descriptive statistics

Populations were divided geographically according to Table 6.1 and groups with at least 11 individuals were described as a population and assessed for descriptive statistics. Observed (H_O) and expected (H_E) heterozygosities, the number of alleles and deviations from Hardy–Weinberg Equilibrium (HWE) were assessed using GENALEX, version 6.5b (Peakall and Smouse 2006). The measure of the amount of heterozygosity across loci can be used as a general indicator of the amount of genetic variability. Heterozygosity is the proportion of individuals in a population that are heterozygous at a particular locus (Beebee and Rowe 2008). Due to small sample size, allelic richness (A_R), an index of genetic diversity was assessed using FSTAT, version 2.9.3.2 (Goudet 2001). Deviations from Hardy-Weinberg equilibrium can indicate inbreeding, population stratification, and genotyping errors (Wiggenton et al. 2005). The Hardy-Weinberg equilibrium law states that allele frequencies should remain unchanged from one generation to the next, while assuming no migration or mutation (Beebee and Rowe 2008). Genotypic linkage disequilibrium was assessed using GENEPOP, version 4.2 (Rousset, 2009). The non-random associations of alleles at different loci is called gametic phase disequilibrium or, linkage disequilibrium.

Genetic differentiation

"F-statistics are a measure of the deficit of heterozygotes relative to expected Hardy-Weinberg proportions in the specified base population" (Allendorf and Luikark 2007). F_{ST} is a measure of allele frequency divergence among subpopulations (Allendorf and Luikark 2007). Calculating Weir and Cockerham's F_{ST} is very useful because it is unbiased with respect to sample size. This is especially useful for analysing the data in this study, as some of the samples came from geographically isolated areas, but were represented by few individuals. The method of Weir and Cockerham (1984) uses an ANOVA approach to estimate within- and among-population variance components (Meirmans and Hedrick 2011). The programme FSTAT Version 2.9.3 (Goudet 1995) was used to estimate pair-wise F_{ST} statistics, and the significance was

tested using 1000 randomisations and applying a Bonferroni correction. The Bonferroni correction is an 'add-on' statistic that accounts for the possibility of false positives being generated when multiple pairwise comparisons are conducted on the same data set, and adjusts the *P* value according to the number of tests that have been conducted.

Analysing the presence of genetic structure

To investigate population substructure within the red squirrel dataset, the number of genetic clusters present was assessed with the programme STRUCTURE version 2.3.1 (Pritchard et al. 2000; Falush et al. 2003). The programme uses a clustering algorithm to group genetic data, and can be used to detect migrants, or admixed individuals, as well as detect cryptic population structure (Hubisz et al. 2009). The programme is provided with a prior number of clusters (*K*), and assumes that the data is within Hardy-Weinberg and linkage equilibrium. The programme then assesses the allele frequencies and population membership of each individual to the putative *K*, using Markov chain Monte Carlo (MCMC) to iterate the data and assign individuals to clusters across the assigned prior parameters (Hubisz et al. 2009). The Bayesian clustering algorithm implemented in the programme was used to analyse the data with default settings with a burn-in period of 10,000, followed by 100,000 replicates with no prior population information. The likelihood of *K* equals 1 to 10, with each *K*-value replicated three times to assess the most likely number of inferred populations. The most likely *K* was assessed by implementing the delta-*K* method (Evanno et al. (2005), using STRUCTURE Harvester (Earl and van Holdt 2012). The method employs an ad-hoc statistic (delta-*K*) that is based on the rate of change log probability of data between successive *K* values, but is sensitive to the type of marker or the number of individuals sampled from a putative population. Bayesian statistics generate multiple 'posterior probabilities' or a posterior probability distribution by providing a probability that an individual's genotype belongs to a particular population and can give an estimate for a number of populations present in the dataset (Beebee and Rowe 2008).

The relationship between individual and population based genotypes using Principle coordinate analysis

To further investigate the presence of genetic structure, a principal coordinate analysis (PCoA) in GENALEX, version 6.5b (Peakall and Smouse 2006) was employed to further examine the genetic relationships among the samples. This multivariate approach was chosen to complement the STRUCTURE analysis as that might be affected by the presence of related individuals in the dataset. PCoA has shown to be more informative regarding genetic distances among major groups (Kassa et al. 2012). The PCoA was performed for both individuals, and for populations. PCoA is based on genetic distance (Nei 1978). It allows the plotting of the major patterns within a multivariate dataset e.g. multiple loci and multiple samples. The analysis option in the PCoA menu finds the relationship between the distance matrix elements based on their first three principal coordinates (Blyton and Flanagan 2012). The genetic distance matrix was also analyzed via Analysis of Molecular Variance (AMOVA). This analysis was also conducted in GENALEX. The AMOVA made it possible to compare genetic variation between and within the populations to determine if there was population differentiation present in the data.

Is Isolation by Distance causing genetic structure?

Isolation-by-distance (IBD) (Wright 1943) refers to a pattern of decreasing genetic relatedness between individuals with increasing geographic distance. The IBD model assumes that organisms have the ability to move freely between two points within dispersal distance, but gene flow may be restricted in the presence of physical barriers to dispersal such as habitat fragmentation.

Mantel's test for isolation by distance was run on the samples using GENALEX, with Nei's standard genetic distance (Nei 1972, 1978) and with 999 permutations. This distance measure was shown to perform well in tests of isolation by distance, and showed relatively low variance, high linearity, and a slope that approaches zero at low geographical distances (Paetkau et al. 1997). If geographic distance was causing genetic differentiation in the data set then the Mantel test between genetic distance

(G) and geographic distance (D) should be significant ($P < 0.05$). For missing data, GENALEX interpolates the average genetic distance (calculated across all non-missing pairwise individual distances) at that locus for the relevant pairwise population contrast. Latitude and longitude data was provided for each sample, and the programme returns the distance in Km.

Testing for relatedness within populations

It is often difficult to obtain reliable estimates of relatedness using microsatellite data as it is often associated with large errors. In addition, computer simulated data has indicated that a high numbers of loci are required to obtain reliable estimates (e.g. > 30 microsatellite loci). However, even with a few loci it appears that informative estimates of average pairwise relatedness within groups can be achieved (Blouin 2003). To assess if the genetically defined populations (STRUCTURE, PCoA) might be affected by the presence of related individuals, GENALEX, version 6.5b (Peakall and Smouse 2006) was used to estimate relatedness using the Pairwise RELATIONSHIP coefficient 'r' from Lynch & Ritland (1999). R has an expected value of 0 for unrelated individuals, 0.25 for half siblings and 0.5 for full siblings or parent–offspring relationships. The mean across all populations was then randomised and bootstrapped (999 permutations) to iteratively calculate relatedness and account for errors.

6.2.3 Mitochondrial DNA analysis

A subset of samples were sequenced in this study: Waterford (n = 12), Cork (n = 4), Wexford (n = 14) the Midlands (n =3), Dublin (n = 4), and Antrim (n = 10). The PCR protocol was as described in Section (2.2.4.1), using the primers from Section (2.3).

Mitochondrial DNA data analysis

Nucleotide sequences were analysed using the Basic Local Alignment Search Tool BLAST) software at http://www.ncbi.nlm.nih.gov/BLAST/ from the GenBank National Centre Biotechnology Information (NCBI) database (Altschul et al. 1990). DNA

sequences were compared by multiple alignments in MEGA V.5.05 (Tamura et al. 2011) and DNA sequences were assigned to haplotypes using ARLEQUIN version 3.5 (Excoffier and Lischer, 2010).

To combine the mitochondrial DNA data from this study with the previously published dataset by Finnegan et al. (2008), sequences were truncated to 263 bp. DNA sequences were compared by multiple alignments in MEGA V.5.05 (Tamura et al. 2011) and haplotypes were described as before.

Testing for population expansion

DNASP version 5.10.01 (Librado and Rozas 2009) was used to estimate nucleotide diversity across the pooled mitochondrial DNA dataset. To assess if the Irish red squirrel population was undergoing a genetic expansion (which might be expected if the red squirrels were relatively recently introduced), deviations from a population expansion model were investigated using Fu's (1997) F_S statistic and Tajima's (1989) *D* statistic. Fu's F_s determines whether mutations are selectively neutral and significant negative estimates can indicate population expansion (Fu, 1997). Tajimas *D* statistic investigates if a sequence is evolving randomly, or if it is evolving under selection, which might indicate demographic expansion. Harpending's raggedness index (Hri) tests were also conducted to assess whether the data deviated significantly from a population expansion model (Harpending 1994). The probability was determined using the coalescent simulation tool with 1000 replicates.

Testing for relationships between mitochondrial DNA haplotypes

A median-joining network was constructed using the median algorithm of Bandelt et al. (1999) in NETWORK version 4.6 (http://www.fluxus-engineering.com). Sequences were aligned with MEGA version 5.0 (Tamura et al. 2007).. Simulation studies have demonstrated that this method provides reliable estimates of the true genealogy (Cassens et al. 2005).

In addition, a neighbour joining tree was constructed in MEGA using haplotypes previously recorded in Ireland, Jersey, Isle of Wight and Wales (Ogden et al. 2005; Finnegan et al. 2008; Simpson et al. 2013). The analysis was conducted using 1000 bootstraps and genetic distance was computed using the *P*-distance method (Nei and Kumar 2000). To compare the DNA sequences from this study with previously published studies from Britain and Ireland (Barratt et al. 1998; Hale et al. 2004; Ogden et al. 2005; Finnegan et al. 2008) and with sequences from across Europe (Grill et al. 2009), sequences were truncated to 190 bp.

Haplotypes from this study were split by geographic area and combined with haplotypes previously published by Finnegan et al. (2008) to compute pair-wise ϕ_{ST} values between populations. This analysis takes into account divergence between haplotypes. The haplotype recorded in Wexford was combined with the Waterford population, as only one haplotype was recorded there. Haplotypes previously recorded in Killarney, Co. Kerry were also included. Pair-wise ϕ_{ST} were calculated using DNASP using 10000 permutations, and the resulting population affinities were exported into MEGA and visualised using a MDS (multidimensional scaling plot) as implemented in XLSTAT (XLstat v.7.5.2, http://www.xlstat.com/en/home/). To account for geographic biases, estimates of evolutionary divergence between all pairs of sequences, including the haplotypes from this study and those from Finnegan et al. (2008). The number of base differences per site from between sequences were calculated between all pairs of haplotypes in MEGA5 using 1000 bootstraps. The resulting affinities were again visualised using an MDS plot.

6.3 Results

6.3.1 Microsatellite Data Analysis

All loci and data were used for analysis following use of the programme MICROCHECKER as despite some evidence of null alleles there was no systematic presence of null alleles, large allele dropout or stutter at any locus. Individuals with two or more null alleles were however removed from the population assignment tests in STRUCTURE. A representative population was selected from Co. Waterford (n = 18), which included samples from all sites analysed in Chapter 5.

Descriptive statistics

Descriptive statistics were calculated for subpopulations with 11 or more individuals. The average number of alleles ranged from 2.7 in the Waterford population to 3.2 in the other populations (Galway, Wexford and Northern Ireland). Allelic richness also referred to as allelic diversity or mean number of alleles per locus ranged from 1.8 in the Waterford population to 2.2 in the Galway population. Average levels of expected heterozygosity ranged from 0.40 in the Waterford population to 0.55 in the Galway population. Average observed levels of heterozygosity ranged from 0.38 in the Waterford population to 0.60 in the Northern Ireland population. No loci significantly deviated from Hardy-Weinberg equilibrium across all populations, but all populations except Northern Ireland contained some loci that had deviations. The Galway and Wexford populations contained the highest number of loci with Hardy-Weinberg deviations (Table 6.2).

Table 6.2: Descriptive statistics for microsatellite data for subpopulations.

Pop	Locus	N	N_A	A_R	H_O	H_E	HWE
WAT	Rsu5	18	2	1.9	0.50	0.49	0.227
	Lis12	17	3	1.3	0.18	0.26	0.112
	Scv3	17	3	1.8	0.18	0.26	0.794
	miniScv4	16	2	1.8	0.50	0.47	0.931
	miniScv31	15	4	1.9	0.47	0.55	**0.037**
	miniscv8	18	2	1.9	0.39	0.31	0.082
	Scv6	18	3	1.9	0.50	0.50	0.513
	Scv20	14	3	2.2	0.50	0.57	0.221
	Lis3	17	2	1.3	0.24	0.21	0.648
	Average	16.7	2.7	1.8	0.383	0.401	0.396
GAL	Rsu5	11	6	3.1	0.45	0.8	**0.004**
	Lis12	7	2	1.9	0.43	0.46	0.860
	Scv3	10	2	1.7	0.3	0.38	0.527
	miniScv4	9	3	2.0	0.22	0.44	0.095
	miniScv31	11	4	2.5	0.1	0.72	**0.001**
	miniscv8	11	3	2.2	0.82	0.57	**0.000**
	Scv6	11	5	2.6	0.91	0.66	**0.001**
	Scv20	10	2	1.9	0.3	0.5	0.213
	Lis3	11	2	1.9	0.82	0.5	**0.035**
	Average	10.1	3.2	2.2	0.48	0.55	0.193
WX	Rsu5	14	3	2.2	0.71	0.55	0.517
	Lis12	14	3	1.4	0.21	0.2	0.977
	Scv3	14	3	1.5	0.14	0.26	**0.003**
	miniScv4	14	2	1.6	0.36	0.29	0.416
	miniScv31	14	7	2.9	0.71	0.74	0.412
	miniscv8	13	3	2.2	1	0.57	**0.005**
	Scv6	14	3	2.3	1	0.61	**0.011**
	Scv20	13	2	1.6	0.38	0.31	0.391
	Lis3	14	3	1.8	0.5	0.39	0.670
	Average	13.8	3.2	1.9	0.56	0.43	0.378
NI	Rsu5	11	3	1.9	0.55	0.42	0.671
	Lis12	11	3	1.5	0.27	0.24	0.965
	Scv3	11	3	2.2	0.68	0.7	0.537
	miniScv4	10	2	1.7	0.33	0.28	0.729
	miniScv31	10	5	2.5	0.7	0.65	0.540
	miniscv8	11	3	2.4	0.73	0.64	0.624
	Scv6	11	3	2.5	0.91	0.66	0.226
	Scv20	11	4	2.2	0.45	0.55	0.888
	Lis3	11	3	2.1	0.82	0.54	0.203
	Average	10	3.2	2.1	0.6	0.52	0.598

N_A number of alleles, A_R allelic richness, H_O observed heterozygosity, H_E expected heterozygosity, and *HWE* Hardy-Weinberg Equilibrium (P = 0.05), with significant values in bold.

Genetic Differentiation

The pairwise F_{ST} estimates showed that subpopulations were highly differentiated from one another, which may be an artefact of sampling isolated populations (Table 6.3). It is likely that the other populations are also significantly differentiated, but the results may be affected by small sample sizes. The overall F_{ST} = 0.237, 99% c.l = 0.158, 0.324 was found to be significantly different from zero.

Table 6.3: Pairwise F_{ST} rates between subpopulations, and corresponding significance after Bonferroni correction (*P* = 0.002).

	WD	CK	GAL	WX	MDS	NI
DUB	0.2278	0.2567	0.060	0.2967	0.0408	0.1894
WD		**0.257**	**0.216**	**0.343**	0.182	**0.2868**
CK			**0.174**	**0.397**	**0.258**	0.3335
GAL				0.199	0.105	0.2151
WX					**0.234**	**0.2597**
MDS						0.0762

Bold values are significant. *WD* Waterford, *CK* Cork, *GAL* Galway, *WX* Wexford, *MDS* Midlands, and *NI* Northern Ireland.

Genetic Structure

A total of 58 individuals were analysed for assignment, removing samples that failed to amplify at two or more loci. The STRUCTURE programme found that *K* = 4 was the mostly likely number of genetic clusters in the population, but there was also a second peak at *K* = 7. This was observed when the mean of the log of posterior probability was graphed. Using the Evanno et al. (2005) method implemented in STRUCTURE HARVESTER (Earl and van Holt 2012), there was not a clear divide between *K* = 3 and *K* = 4 as the most appropriate number of populations in the red squirrel population (Fig 6.1 [a, b]). *K* = 4 also showed the low deviance from the mean, and appeared more uniform in comparison to the other *K* values.

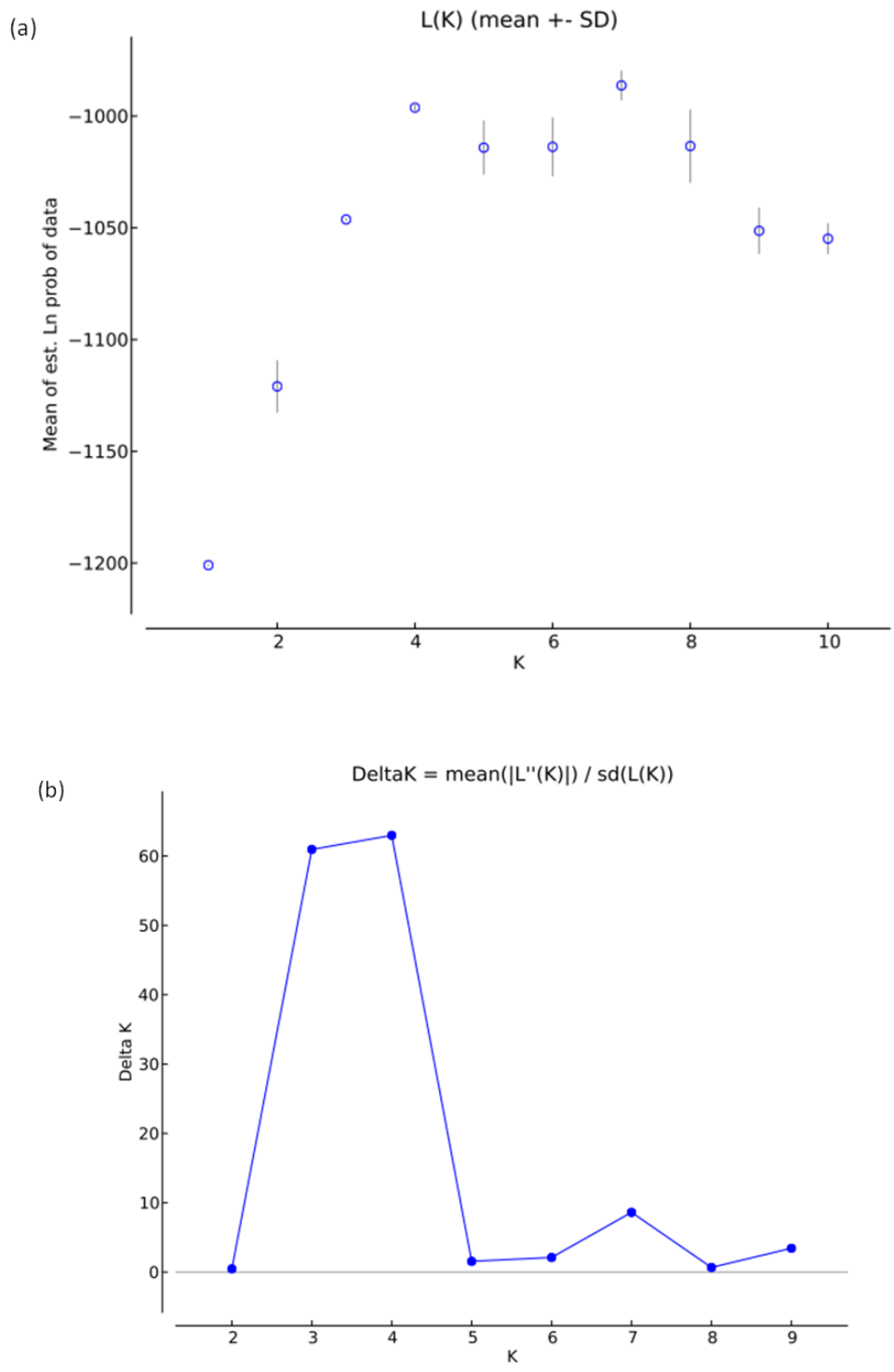

Fig 6.1 (a): Description of L(K) (±SD) across the three replicates and (b) Delta *K* values computed by the software STRUCTURE (Pritchard et al. 2000) following Evanno et al. (2005) at *K* = 1-10.

The STRUCTURE results showed that there was a high level of genetic substructuring within the red squirrel data. The most likely *K* values (*K* = 2, 3 and 4) are displayed in Fig 6.2. At *K* = 2, the Galway and Wexford samples were disctinct from the remaining populations. At *K* = 3, Dublin, Galway, the Midlands and some individuals from Northern Ireland clustered together. The Waterford, Cork and some individuals from Northern Ireland and Galway clustered together, and the Wexford samples mainly clustered together. At *K* = 4, Dublin and Galway clustered together, Waterford and Cork clustered in a second group, Wexford mainly grouped separately, and the Northern Ireland population formed a fourth cluster, but there was admixture with other putative clusters.

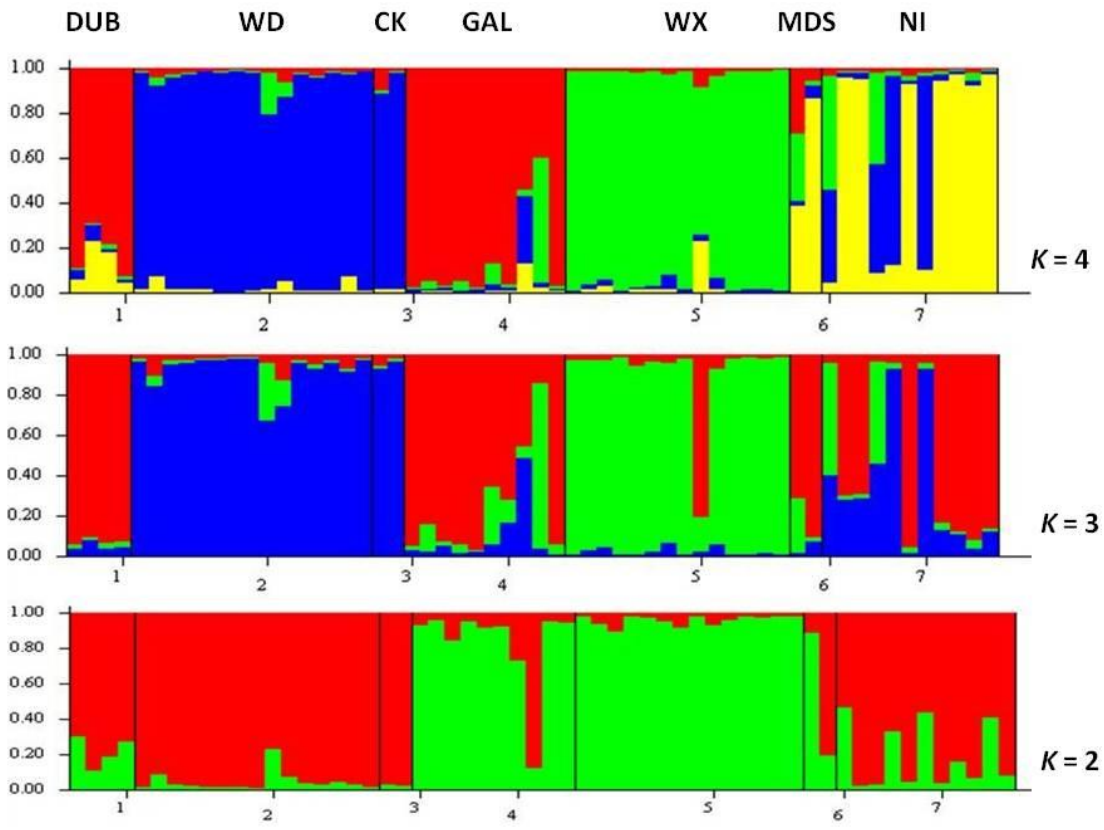

Fig 6.2: STRUCTURE plots for *K* = 2, *K* = 3 and *K* = 4.

Principal Coordinate Analysis

In the Principal Coordinate Analysis (PCoA) (Fig 6.3), Wexford individuals were largely resolved from the other individuals, but there was one sample from Waterford and some individuals from Galway also clustered in that quadrant, which agrees with the STRUCTURE output for *K* = 2, and where the remaining populations are generally unresolved. The Waterford individuals mostly occurred in opposite quadrant along axis 1, but with some individual clustering in axis 2. Axis 2 was largely mixed. The second axis clustered the Northern Ireland individuals together, but there was some admixture with individuals from Galway, Dublin, Midlands and some of the Waterford individuals. Some of the Waterford samples also exhibited some evidence of clustering. However, the majority of the Galway, Dublin midlands and many of the Waterford samples remained unresolved. The PCoA analysis showed high levels of admixture similar to the STRUCTURE results, but there was a higher level of geographic correlation.

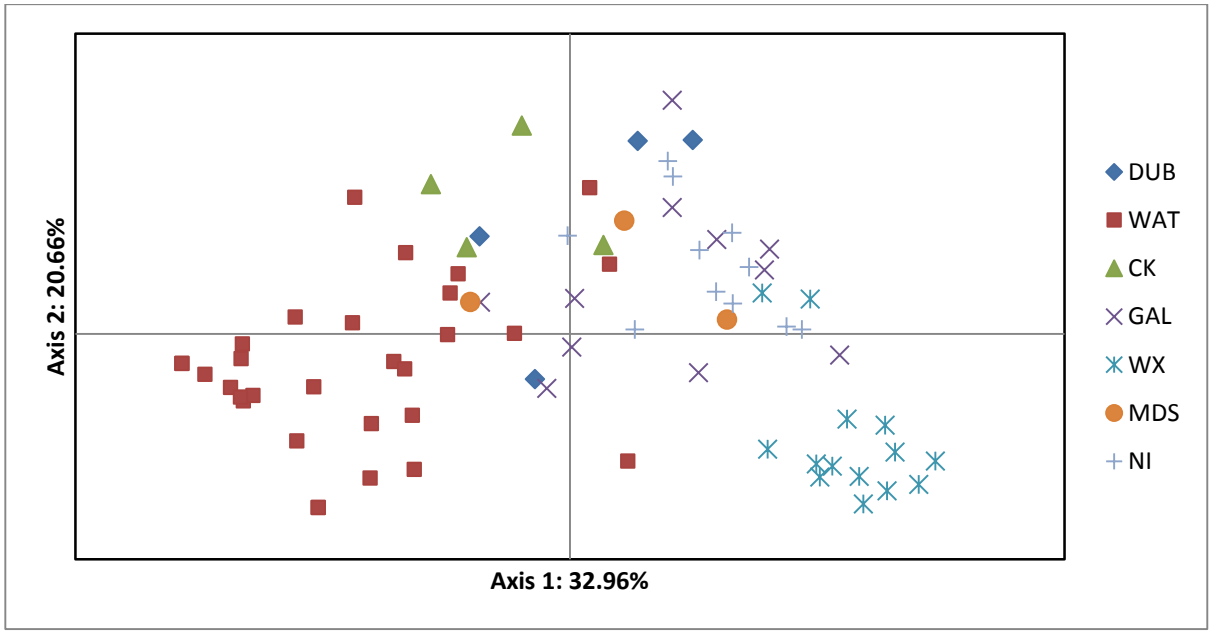

Fig 6.3: Principal coordinate analysis of red squirrel samples from Ireland, across nine loci and explaining 53.62% of the genetic variation *DUB* Dublin, *WD* Waterford, *CK* Cork, *GAL* Galway, *WX* Wexford, *MDS* Midlands, *NI* Northern Ireland.

An additional PCoA analysis was conducted to visualise the clustering of geographic populations (Fig 6.4). This analysis accounted for 77.32% of the overall genetic variation. The analysis clustered the Waterford and Cork samples closely, Dublin and Galway grouped in the same quadrant and Wexford showed some similarity to the Midlands and Northern Ireland. The AMOVA results showed that 67% of molecular variation occurred within the populations and 33% occurred between the populations. The results from the population based PCoA were in agreement with the STRUCTURE output for *K* = 3 and 4.

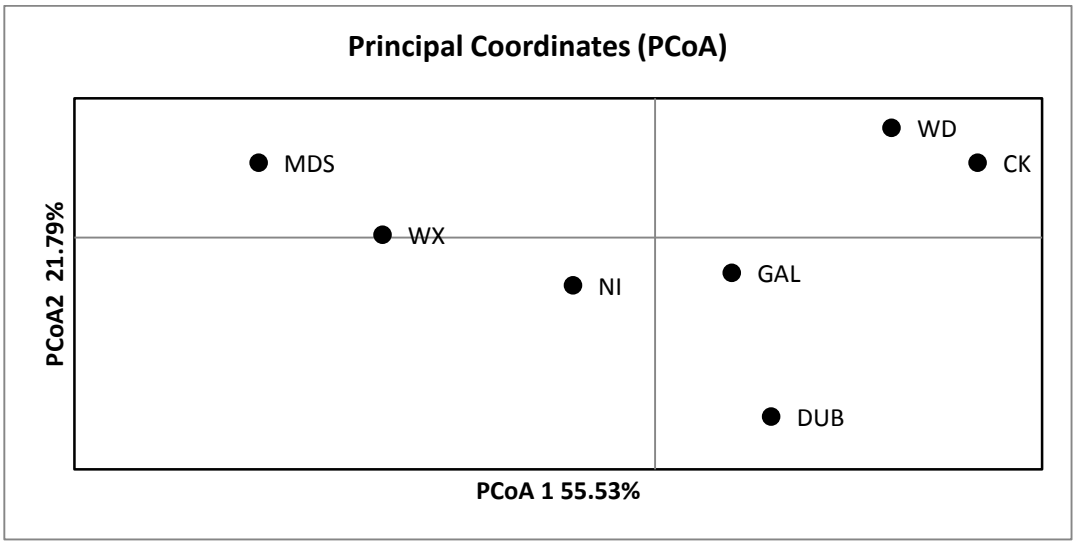

Fig 6.4: Principal coordinate analysis of red squirrel subpopulations. Grouping the individuals by population explained 77.32% of the total genetic variation.

Isolation By Distance (IBD)

The result of Mantel's test revealed a significant pattern of isolation by distance (IBD) (R^2 = 0.1015; *P* = < 0.001) (Fig 6.5), which was expected as samples were used from a broad geographical area, without sampling between the regions. Within the Waterford population (where samples were collected across the county), there was also evidence of IBD (R^2 = 0.043; *P* = 0.05), illustrating that IBD was also occurring at local population level.

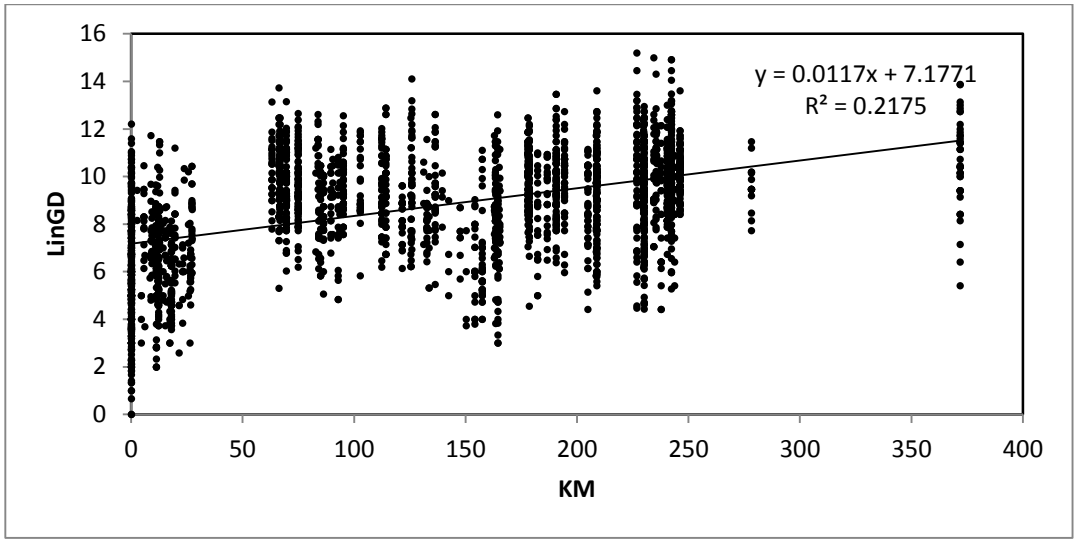

Fig 6.5: Mantel test indicating observed pattern of isolation by distance (RXY* = 0.466; *P* = 0.001). *Rxy = correlation coefficient of Mantel test. P (rxy-rand > = rxy-data) = probability of positive autocorrelation (one tailed).

In addition, the parameter F_{ST} (pairwise F_{ST}) was standardized to $F_{ST}/(1-F_{ST})$ and plotted against the distance between populations (Fig 6.6). Populations that showed the least divergence can be found under the trend line and include Dublin/Midlands, Dublin/Midlands and Dublin/Galway, Midlands/Northern Ireland. Populations that showed the highest divergence from one another can be found above the trend line: Waterford/Wexford, Cork/Wexford, Cork/Northern Ireland. All populations were differentiated from one another, but there is not a straight-line correlation between geographic populations using this generalised approach. The percentage of molecular variance across all subpopulations showed that 64% of the variance was represented by individuals, 21% of variance was shared among individuals and 15% was shared among populations**,** and again this would contribute to IBD and high levels of genetic structuring.

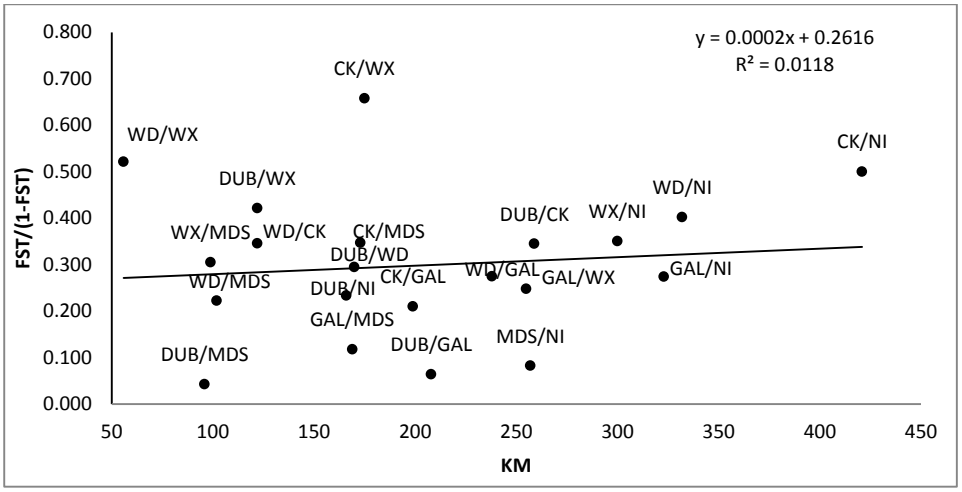

Fig 6.6: Comparison of genetic distance (Log F_{ST}) versus distance (KM). Solid line represents trendline calculated from all population pairs. Coefficient of determination (R^2) was 0.0118, and was not found to be significant (student T test). Distance is defined as the minimum roadway in km. F_{ST} / (1 − F_{ST}) was used as a measure of genetic distance.

Relatedness

The pair-wise relatedness estimation across populations revealed that all populations contained some evidence of related individuals (Fig 6.7). The Dublin, Waterford and Wexford populations contained the highest number of related individuals, and Galway contained the fewest number of related individuals. This can be attributed to the sampling effort, which was concentrated in smaller areas in Dublin, Galway and Wexford.

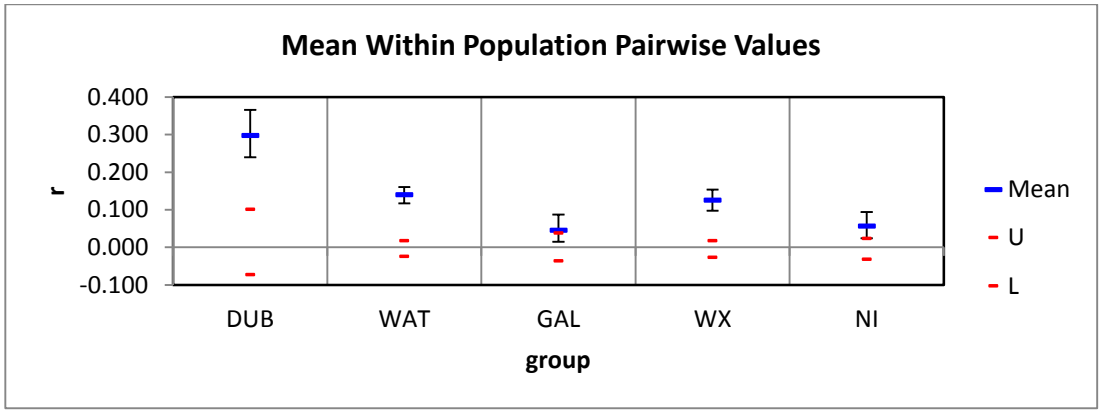

Fig 6.7: The average relatedness (r) figures across all populations. *U* upper and *L* lower 95% confidence limits. P(rand >= data) = 0.001 for Waterford and Wexford; 0.002 for Dublin; 0.004 for Northern Ireland and 0.01 for Galway. Probability, P(rand>=data), for average r is based on permutation across the full data set.

6.3.2 Mitochondrial DNA Analysis

Across the 263 bp of sequence, 226 sites were monomorphic, 37 polymorphic (including a total of 46 mutations), 11 sites included singletons and 26 sites were parsimony informative with an overall nucleotide diversity of 0.02929 ± 0.09491 and gene diversity of 0.995 ± 0.00007. A total of six haplotypes were unique to this study. F_S and *D* values deviated from zero, (-0.0.34170; $P < 0.0001$, -1.0.05132; $P > 0.1$). Tau was estimated at 7.00. Harpending's raggedness index was 0.04; $P < 0.0001$. The sum of squared deviations from mismatch analysis was not significant (0.113; $P > 0.1$). Despite some of the analysis indicating that the data was conforming to a model of demographic expansion, the combined analysis did not indicate this. However a significant F_S and non-significant *D* supports population growth and or range expansion (Weckworth et al. 2012). A table of all haplotypes from this study and those previously recorded by Finnegan et al. (2008) are provided in Table 6.4.

Table 6.4: Locations of recorded mitochondrial DNA haplotypes in Ireland. Haplotypes recorded in this study are displayed in bold.

Haplotype	Location	Source
IE1	Dublin	Finnegan et al. (2008)
IE2	**Wicklow, Dublin**	**Finnegan et al. (2008), this study**
IE3	Wicklow	Finnegan et al. (2008)
IE4	Wicklow	Finnegan et al. (2008)
IE5	Wicklow	Finnegan et al. (2008)
IE6	Wicklow	Finnegan et al. (2008)
IE7	Antrim	Finnegan et al. (2008)
IE8	Kerry	Finnegan et al. (2008)
IE9	Kerry	Finnegan et al. (2008)
IE10	Galway	Finnegan et al. (2008)
IE11	Galway	Finnegan et al. (2008)
IE12	Galway	Finnegan et al. (2008)
IE13	Galway	Finnegan et al. (2008)
IE14	**Wexford**	**Finnegan et al. (2008), this study**
IE15	Wicklow, Kerry	Finnegan et al. (2008)
IE16	**Limerick, Galway, Waterford**	**Finnegan et al. (2008), this study**
IE17	Down, Wicklow	Finnegan et al. (2008)
IE18	**Waterford, Cork**	**Finnegan et al. (2008), this study**
IE19	Dublin	Finnegan et al. (2008)
IE20	Kilkenny	Finnegan et al. (2008)
IE21	Wicklow	Finnegan et al. (2008)
IE22	Dublin, Galway	Finnegan et al. (2008)
IE23	**Waterford**	**This study**
IE24	**Waterford**	**This study**
IE25	**Antrim**	**This study**
IE26	**Antrim**	**This study**
IE27	**Cork**	**This study**
IEGB1	Cork, Kerry, Galway, Wicklow, Tipperary, England	Finnegan et al. (2008)
IEGB2	Kilkenny, Laois, Wicklow, England	Finnegan et al. (2008)
H15	Wicklow, the Netherlands	Hale et al. (2004), Finnegan et al. (2008)

The mitochondrial DNA haplotypes tended to be unique to each sampling area, and this was reflected by a high level of variability in the haplotype NETWORK (Fig 6.8). Samples from Antrim, Northern Ireland included two haplotypes, (IE25 and IE26) both similar to IE7, a haplotype previously recorded in Northern Ireland (Finnegan et al. 2008). Samples from Cork contained the haplotype IE18 and IE27, both closely related. IE18 was previously recorded in Cork and Lismore (West Waterford), and this

haplotype showed close similarity to a haplotype observed in Northern Ireland, possibly representing a similar colonization history, or introductions from the same source. In Waterford, three haplotypes were present IE23 and IE16, which were similar to one another, and one other haplotype (IE24) which was more similar to IE14 found in Wexford. IE14 was previously described as a Central European haplotype (Finnegan et al. 2008). IE16 was previously described in Limerick. Samples from Carrickgollogan Co. Dublin contained the haplotype IE2.

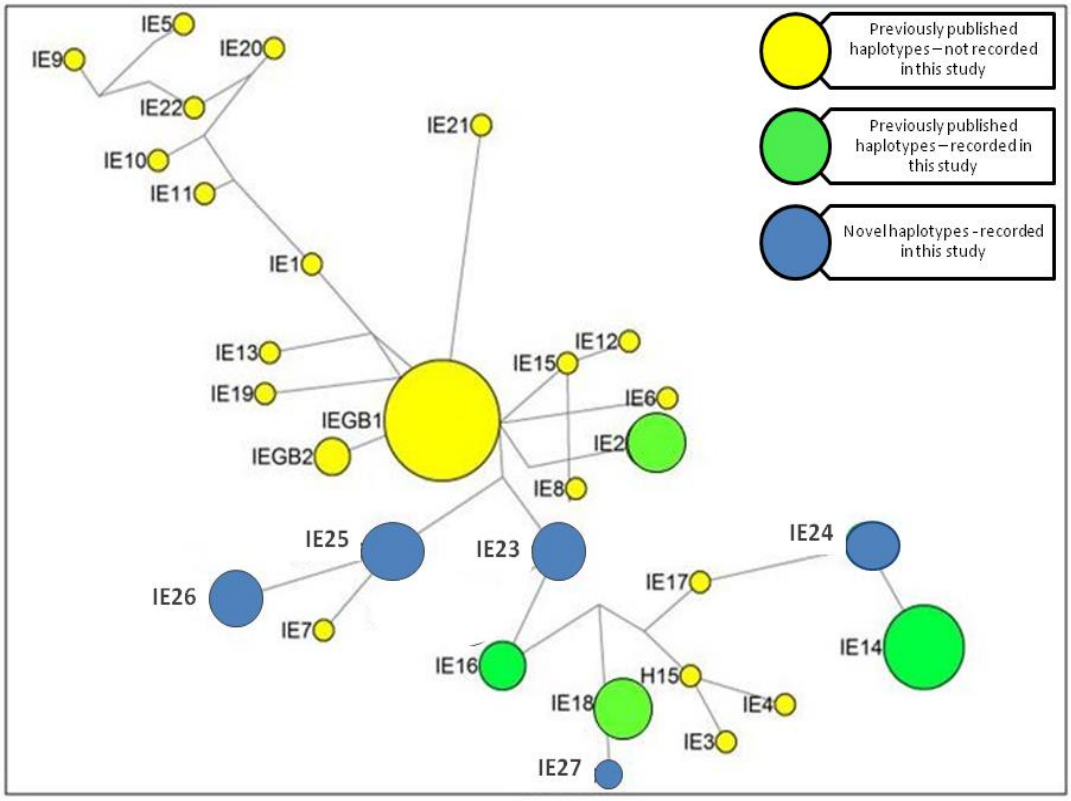

Fig 6.8: Mitochondrial DNA median-joining network. Constructed from 263 bp of control region mitochondrial DNA. Samples from this study are in green. Previously published haplotypes by Finnegan et al. (2008). Northern Ireland (IE25, IE26), Waterford (IE23, IE24, IE16), Wexford (IE14), Dublin (IE2) Cork (IE18, IE27).

The locations of the sequenced samples from this study are mapped in Fig 6.9., with closely related haplotypes, those differing by only one to two single bases coloured the same. Only Counties Wexford and Waterford shared similar haplotyes (IE14, IE24).

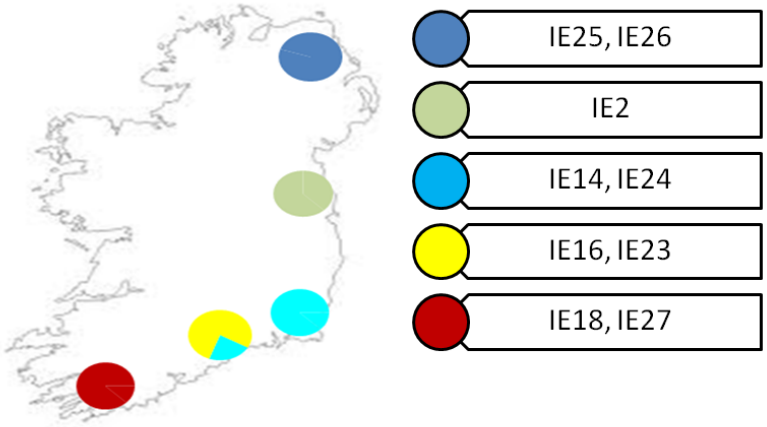

Fig 6.9: Distribution of red squirrel haplotypes observed in this study. Haplotype relatedness can be seen in Fig 6.6. Haplotypes IE23, 24, 25, 26 and 27 are new haplotypes.

To infer phylogenetic relationships between the populations surveyed in this study, a non-metric multidimensional scaling (MDS) analysis based on the ϕ_{st} distance matrix calculated was performed. The plot shows an affinity between haplotypes from Cork and Northern Ireland, Galway and Waterford/Wexford and Dublin and the Midlands (Fig 6.10). The analysis also included haplotypes previously recorded in Killarney National Park, and that population also clustered with the Midlands and Dublin population.

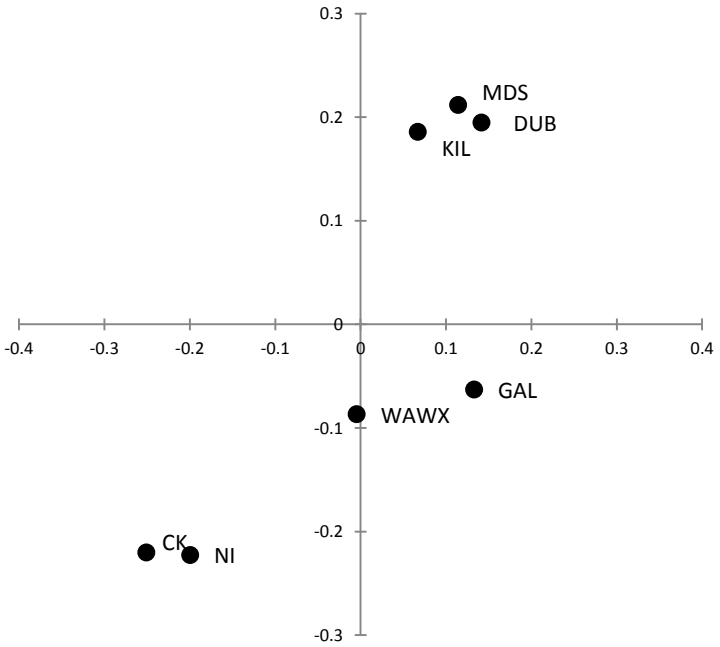

Fig 6.10: Two-dimensional MDS plot of pair-wise relatedness values from red squirrel haplotypes from each of the geographical areas surveyed. The stress value for the MDS plot is 0.130. CK Cork, *NI* Northern Ireland, *WAWX* Waterford/Wexford, *GAL* Galway, *MDS* Midlands, *DUB* Dublin, *KIL* Kerry.

As the NETWORK analysis had shown that geographic areas contained red squirrels from divergent origins, estimates of evolutionary divergence were also calculated between all pairs of haplotypes. The results correlated well with previous knowledge of divergent haplotypes and provided insights into at least four haplotype groups (Fig 6.11). IE14 and IE7 clustered on the left upper axis, while the majority of haplotypes (including those from Waterford, Northern Ireland and Cork) clustered on the right upper axis. Another cluster of haplotypes included IE1, IE11, IE20 and IE22 previously recorded by Finnegan et al. (2008) in Kilkenny and Galway. IE8 previously recorded in Killarney National Park (Finnegan et al. 2008) did not show any similarity to the other haplotypes.

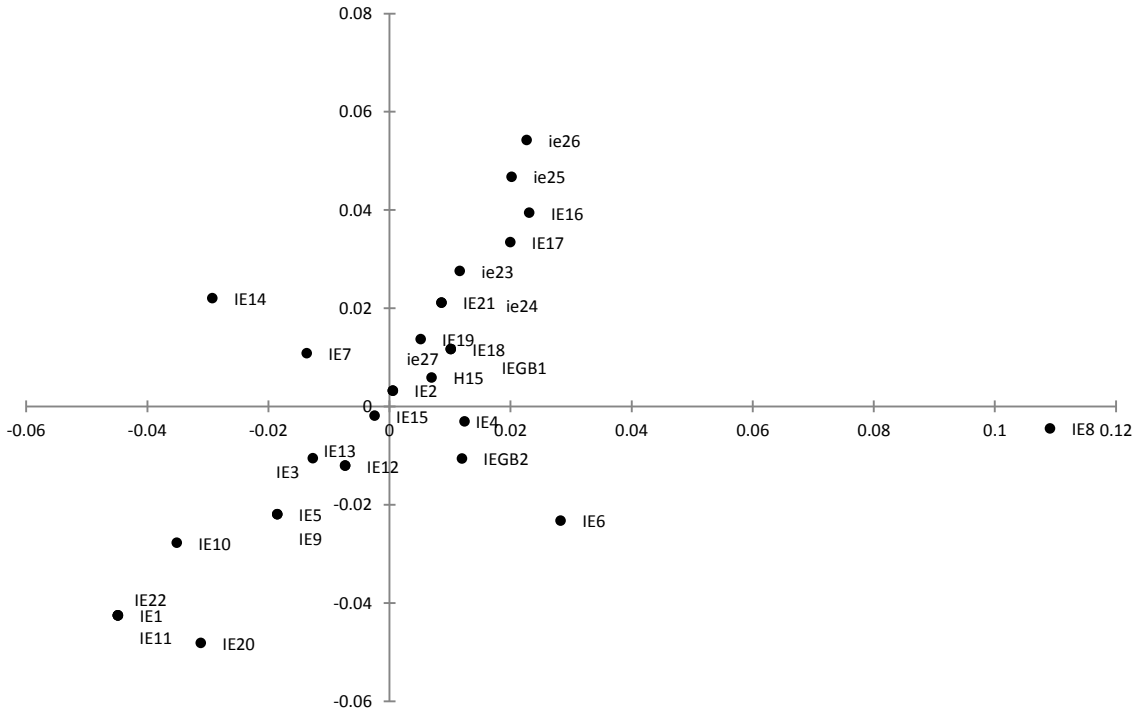

Fig 6.11: MDS plot of estimates of evolutionary divergence between haplotypes. Kruskal's stress = 0.243. *IE* haplotypes from Finnegan et al. (2008), *ie* haplotypes from this study.

Comparison of sequences from Wales, Jersey and the Isle of Wight

The analysis that compared Irish, Jersey, Isle of Wight and Welsh haploypes showed that the haplotypes were mostly divergent, as shown in the phylogeny and NETWORK analysis (Fig 6.12 [a, b]). However, there were some similarities that showed that some of the Irish haplotypes share a relatively recent ancestry with those on the Jersey and the Isle of Wight. IE14 (Wexford), IE17 (Northern Ireland and Wicklow), IE24 (Waterford) and IOW6 (Isle of Wight) all grouped together in the phylogeny analysis. Shared ancestry was also found between JE162 (Jersey), IOW3 (Isle of Wight) and IE26 (Northern Ireland), and a final group consisting of IE3, IE4, H15 (all found in Wicklow) and JE120 (Jersey). Similar groupings were also apparent in the median-joining NETWORK diagram, with additional similarities showing between MIDW (Mid Wales) and IE7 (Northern Ireland). IE23 and IE16 (both recorded in Limerick and Waterford) also shared some similarities with WC3 (Wales). The shared

ancestry between Ireland, Jersey, Isle of Wight and Britain, appears to occur mostly with red squirrels in the east of Ireland.

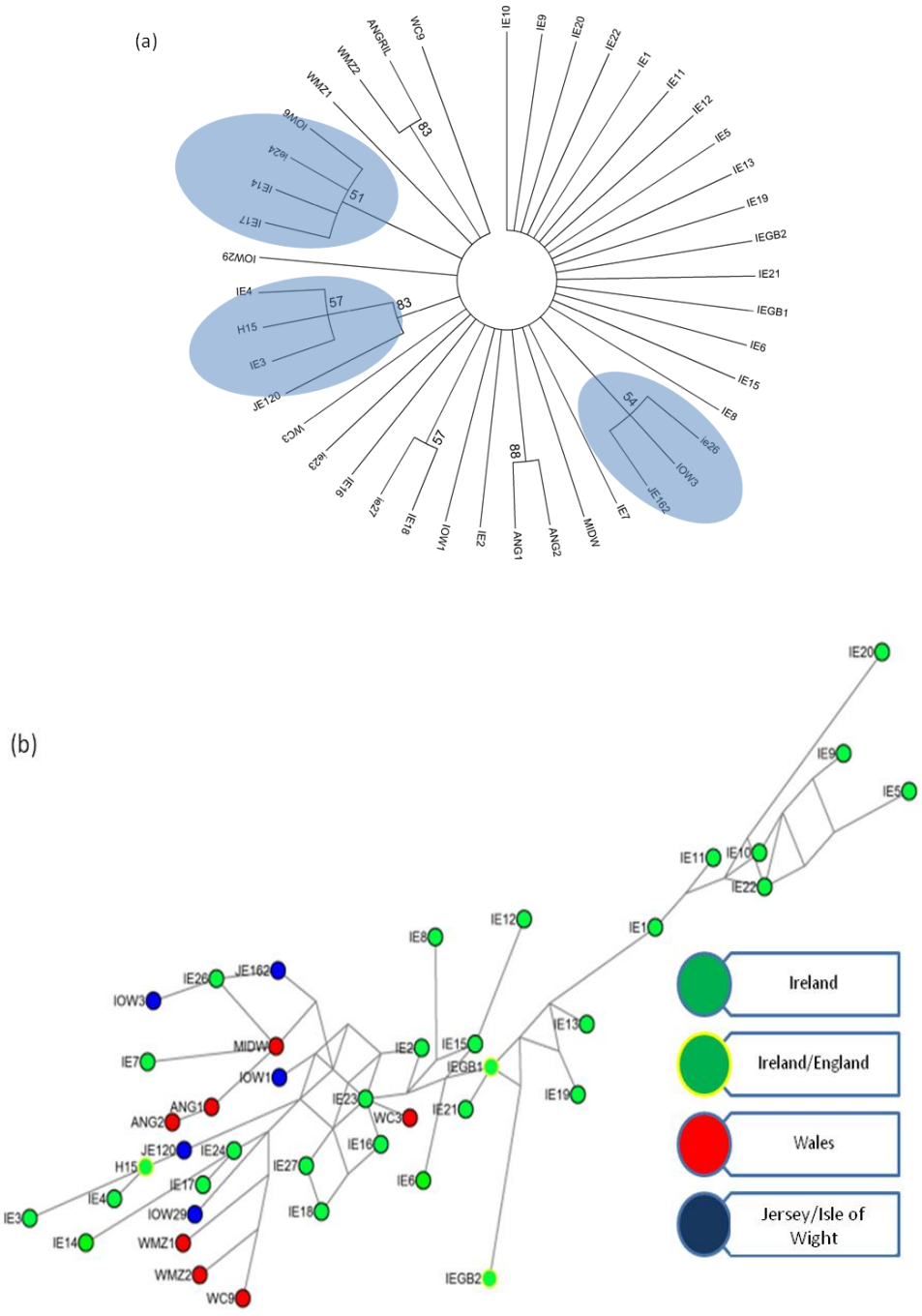

Fig 6.12 (a) Neighbour-joining tree of red squirrel haplotypes from Ireland, Wales and Jersey/the Isle of Wight. Low bootstrap support values > 0.50 are provided at branches. (b) Median-joining network of red squirrel haplotypes from Ireland, Wales and the Jersey and the Isle of Wight.

In terms of the larger European context, the median-joining NETWORK revealed a general lack of phylogeographic structure of red squirrel populations across Europe (Fig 6.13). The Irish samples (green) mostly clustered together in the NETWORK and shared most similarity with British samples (yellow). Some of the Irish and British haplotypes appeared some what distinct from other haplotypes in Europe, which may represent the original British haplotypes prior to the reintroduction of Continental European red squirrels. This however, needs further analysis and exploration. However, one cluster of haplotypes that included samples from Wexford and Waterford (IE14 and IE24), grouped more closely with samples from Continental Europe, Italy and Britain, likely representing some of the introduced stock from Continental Europe. The British haplotypes were distributed throughout the NETWORK and there was a lack of evident structure. Samples from Iberia (orange) shared haplotypes with Italy, mainland Europe and Finland (grey). The Finnish samples were also found dispersed throughout the NETWORK. The cluster was largely unresolved. The British haplotypes (yellow) were distributed throughout the Network. Some clustering between Irish and British haplotypes occurred, but the British haplotypes were found intermixed throughout Europe, demonstrating a lack of geographic structure. This is further investigated in Section 7.3.1.

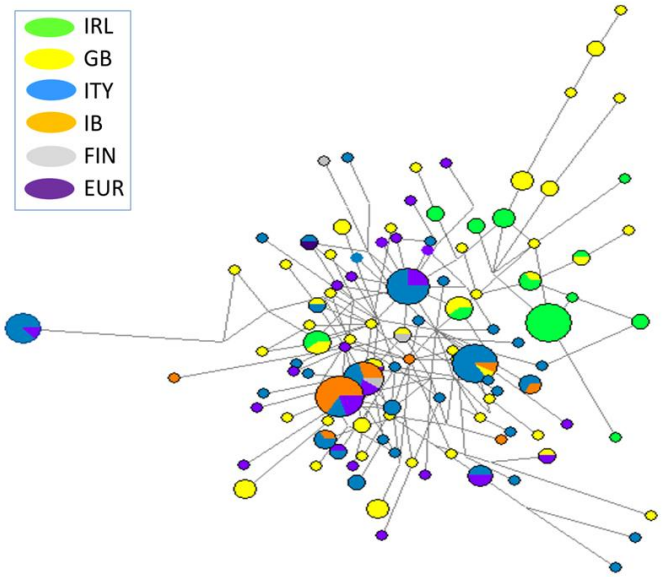

Fig 6.13: Median-joining NETWORK diagram of European haplotypes (190 bp). Pie chart size corresponds to increased haplotype number due to truncated sequences.

6.4 Discussion

Microsatellite data was used in this study to compare levels of genetic diversity in the sampled populations, and to investigate the presence of cryptic genetic structure that could be caused by isolation due to random and small translocations that took place in Ireland in the mid 1800s. To further investigate the causes of genetic structure, isolation-by-distance and relatedness was also investigated in each of the subpopulations. The mitochondrial DNA data was then used to assess the historical footprint of the red squirrel in Ireland, and the combined analysis will now be discussed.

6.4.1 Microsatellite variation

On average, a low number of alleles were present across the Irish red squirrel population. The Galway and Wexford populations had the highest number of deviations from Hardy-Weinberg Equilibrium (HWE). For the Galway population, this may be best explained by the 'Wahlund effect', which is the presence of genetic substructuring within a population. This might be expected as the Galway population is composed of recently translocated individuals that have since successfully bred in the region (Poole and Lawton 2009). Finnegan et al. (2008) also showed that his population contained a high number of divergent haplotypes. Ogden et al. (2005) suggested that deviations from HWE are often accompanied by population declines and inbreeding, which is to be expected in a recently translocated population. The populations in Galway and Wexford are isolated so there is a low probability of natural recruitment into the area. However, the genetic diversity is higher than what was recorded in Anglesey, Wales by Ogden et al. (2005), where red squirrels have recently recovered from very low numbers, but no physical effects have been attributed to low genetically variability to date. Overall, the genetic diversity in the red squirrel population was relatively low, although much of the genetic variation was found within the sub populations and there are region specific alleles. Other populations that also contained deviations from HWE may be attributed to population substructuring as sampling had taken place from different woodlands

within Northern Ireland. The mean number of alleles was lower than what was found in Italy and Britain (Hale et al. 2001; Trizio et al. 2005).

F_{ST}-based population differentiation was significant at both local and regional scales, and this was also supported by the structuring analysis (both STRUCTURE and PCoA analysis) that indicated strong genetic differentiation at most populations. The results however should be interpreted with caution as the sample sizes representing different localities were small within Ireland, and there was also evidence of relatedness within the sampling groups that could contribute to differentiation.

During the evolution of a gene flow–drift equilibrium system, isolation by distance (IBD) should be first seen over small geographical distances, then over larger ones (Slatkin 1993; Trizio et al. 2005). The Mantel test conducted across Waterford, and again across the rest of Ireland, showed that there was significant evidence of IBD across the data set. Trizio et al. (2005) also found that red squirrels in Italy showed evidence of IBD at both a local and regional level and attributed this to the species' sensitivity to barriers of dispersal such as habitat fragmentation. This was evident in the Waterford population, and that area consists of highly fragmented woodlands (Chapter 5). Patterns of IBD have also been shown to strongly bias genetic clustering programmes such as STRUCTURE, and this provides further evidence that the results of STRUCTURE should be viewed and interpreted with a level of caution (Frantz et al. 2010).

The results of the STRUCTURE do however make geographical sense as the Waterford and Cork red squirrels cluster in one group, Dublin, the Midlands and Galway form a second group, Wexford forms a third cluster, and the Northern Irish squirrels form a fourth cluster. The population substructure in Wexford could be attributed to the high levels of relatedness within the sample group, and that particular population is also geographically isolated in the Raven, Co. Wexford. The admixture in Northern Ireland may be attributed to sampling from multiple areas within the North, and it is likely that increased sampling may yield further insights there.

Previous studies have shown that barriers to dispersal can cause genetic substructuring caused by genetic drift, and that can result in the loss of genetic diversity and cause local extinctions (Hitchings and Beebee 1997). The Glanville fritillary (*Melitaea cinxia*) was shown to have a greater probability of extinction when populations were small and isolated with little immigration (Hanski et al. 1995; Hitchings and Beebee 1997). This study has also shown that red squirrels appear to be sensitive to the affects of habitat loss and habitat fragmentation such as what has been previously described in Italy (Trizio et al. 2005). The understanding of the presence of genetic structure in Ireland is important for the future conservation of this species.

6.4.2 Mitochondrial DNA variation

The mitochondrial DNA haplotypes found in this study appear to be highly divergent. It is likely that a large number of haploypes are missing from the NETWORK diagram, and more sampling is needed to better assess the number of haplotypes present in the country. The pair-wise relatedness statistics showed that Cork and Northern Ireland shared some similarity to one another. This supports anthropogenic introductions into the area. There were also shared haplotypes between Waterford and Galway, again suggesting historical movement of red squirrels between these areas. One of the haplotypes IE16 found in both these populations was recorded in Limerick by Finnegan et al. (2008). The MDS plot of the haplotype analysis also showed similarity between these regions, and may help explain some of the admixture found within the Northern Ireland red squirrels.

The disparities between the microsatellite and mitochondrial data in this study may also be further explored. For instance, the Wexford population appears differentiated at the microsatellite analysis from other groups. However, the mitochondrial DNA analysis revealed that similar haplotypes are present in Waterford, Wexford and in Northern Ireland. It is likely that this pattern reflects small populations that are affected by isolation and genetic drift, which appears to be evident in this study. The effects of this are likely further exacerbated by multiple

introductions from divergent origins. Fragmented populations reduce the possibility of gene flow and thus reduce the possibility of a rescue effect and can lead to local extinctions (Wentworth et al. 2012; Mondol et al. 2013).

6.4.3 Tracing the Irish red squirrel genetic footprint

The analysis with additional haplotypes from Wales, Jersey and the Isle of Wight revealed a number of interesting trends. A number of haplotypes from the East of Ireland appear to share a common ancestry with red squirrels in Wales, Jersey and the Isle of Wight. Simpson et al. (2013) investigated the origins of the red squirrel in the Isle of Wight and Jersey. It was found that there were two introductions into Jersey, one from mainland Britain and the other from France that occurred approximately 120 years ago. Jersey and the Isle of Wight haplotypes JE162 and IOW3 were shown to be of English origin by Simpson et al. (2013), and this group of haplotypes shared similarity with IE26 found in Northern Ireland in this study. This suggests that it is also likely that this Irish haplotype originated in England. The Jersey haplotype (JE120) that showed similarity to Irish haplotypes IE3, IE4, H15 (Wicklow), was shown to have originated in France by Simpson et al. (2013). This indicates a possible French introduction in Ireland, either directly from France or an introduction of French translocated red squirrels in Britain into Ireland. The final group that contained similarities between Ireland and the Isle of Wight contained IOW6, IE14, IE17 and IE24 (East Ireland). Simpson et al. (2013) showed that IOW6 found in the Isle of Wight shared an ancestry with red squirrels in Britain, the Netherlands, France and Ireland.

Simpson et al. (2013) were able to decipher the mixed origins of red squirrels in Jersey that were introduced in 1894, as a phylogeographic footprint of the introductions was retained in both the mitochondrial and microsatellite DNA. The DNA results confirmed historical records of introductions of red squirrels into Jersey from both England and France. Evidence of similar introductions are also evident in the Irish red squirrel population, particularly in the east of Ireland, and it is likely that

multiple introductions also took place into the region, probably from the same sources as those used to populate Jersey.

Reintroductions of the red squirrel started to take place in the 1800s (Barrington 1880). Reintroductions were reported to have originated in Britain, and it is likely that reintroductions also took place from continental Europe into Britain (Shorten 1954). The combination of a high number of divergent mitochondrial DNA haplotypes and the presence of genetic clustering in the microsatellite data provide contemporary evidence for multiple historical introductions of the red squirrel in Ireland.

Further evidence for the presence of Continental European red squirrels in Ireland can be seen through morphological variation. Traditionally, the British and Irish red squirrels were known to have a bleached tail (Finnegan et al. 2007), although only a portion of the Irish red squirrel population is thought to still have this characteristic. A black red squirrel was recorded in Down and Midlothian in 1950 (Shorten 1954). The black colour is more common in Central Europe. Black morphs have also been reported more recently in Cork in 2012 (David Kiely pers comm.).

A haplotype found in Killarney (IE8) by Finnegan et al. (2008) exhibited little similarity to British or Irish red squirrels, and Finnegan et al. (2008) suggested that this area may have contained red squirrels that may have represented a remnant red squirrel population in Ireland. Alternatively, they represent mitochondrial DNA haplotypes not sampled elsewhere, or haplotypes that have become extinct in parts of Britain, which was shown in the case of the pygmy shrew (McDevitt et al. 2011).

6.4.4 Using genetic knowledge to aid conservation

The red squirrels sampled from Co. Wexford in this study have since been translocated into Co. Dublin. Genetic screening prior to this introduction would have shown that this population of red squirrels are genetically related, and contain a mitochondrial DNA haplotype that is not common in the Dublin area, and would appear to have different origins than the red squirrel mitochondrial DNA haplotypes

previously recorded in Dublin and Wicklow by Finnegan et al. (2008). In this study, it was shown that IE2 was the common DNA haplotype in the Carrickgollogan area of Dublin, suggesting that red squirrels from Wicklow may have been a better choice for the area. However, given the relatively recent history of red squirrels in Ireland since the mid 1800s, the origin of the red squirrels may be less important. In addition, as that current population in Dublin had already declined, introducing new alleles might be beneficial to the area, which was the conservation management approach taken by Ogden et al. (2005) in Anglesey.

Simpson et al. (2013) showed that red squirrels are likely to retain a genetic footprint of their origins for over 120 years, the genetic background of the population should be firstly assessed to ensure that the population is compatible with the remnant population present, as patterns of genetic structure caused by multiple divergent introductions are likely to persist in the population. Simpson et al. (2013) found a correlation between the mixing of divergent origin red squirrels and the occurrence of amyloidosis, a disease that can have genetic links. It is possible that 'outbreeding', where two individuals from geographically and genetically distinct populations breed could result in a reduced immunity in the population (Goldberg et al. 2005). It is therefore important to understand the original origins of red squirrels as specific ecotypes (defined as populations that have particular adaptations to environmental conditions) may have important roles in the long-term survival of the species (Weckworth et al. 2012). This may be especially important in the red squirrel as over 40 subspecies have been previously described (Barratt et al. 1999), but it is far more likely that these 'subspecies' are in fact ecotypes that have adapted to their current environment. It is also possible that historic declines of the red squirrel (Shorten 1954) were as a result of the translocation of specific ecotypes that were ill-adapted to their new environment.

6.4.5 Future work and limitations of this study

Due to the high levels of divergence between mitochondrial DNA sequences within Ireland and across Europe, there is a need to sequence longer sections of the mitochondrial DNA to aid the wider understanding of red squirrel phylogeography. This became especially evident when the sequence length was truncated for wider European analysis (only 190 bp used in this study). It is likely that the high number of red squirrel mitochondrial DNA haplotypes present across Europe are caused by habitat fragmentation and low migration of females, resulting in locality-specific haplotypes with high levels of intraspecies diversity.

6.5 Conclusion

- It is likely that the pattern of genetic structure evident throughout Ireland is as a direct result of multiple introductions into Ireland over the last 200 years.
- It appears that a lack of gene flow has led to high levels of isolation of the populations. Future conservation projects should consider the genetic history of each red squirrel population prior to relocation or translocation given the high levels of genetic structure found in Ireland, as genetic evidence of translocations are likely to persist in the genome for hundreds of years, and there is currently a poor understanding of the potential consequences to the red squirrels.
- There are some genetic links between squirrels on Jersey, Isle of Wight, Wales and the east of Ireland, and management efforts could be communicated between the relevant conservation groups, as there may be scope for the establishment of a combined breeding programme. However, disease transmission such as squirrel para-pox virus and amyloidosis need to be considered.

Chapter 7:

Application of genetic techniques to a remnant red squirrel (*Sciurus vulgaris*) population in Mid Wales

7.1 Introduction

The red squirrel population in Wales mostly consists of small isolated population pockets that are threatened by the presence of the grey squirrel (Ogden et al. 2005). The island of Anglesey in North Wales was used as an area for reintroduction of the red squirrel as the island could be protected from the invasion of the grey squirrel with a regular trapping programme. Ogden et al. (2005) assessed the current levels of genetic diversity of the population on the island and found that there had been a loss of genetic diversity on the island over the last 20 years. Ogden et al. (2005) analysed squirrels from Mid Wales (n = 2) where the haplotype "MIDW" was recorded. Ogden et al. (2005) also found the Mid Wales haplotype was closely related to a haplotype found in Clocaenog, North Wales "WC3". Ogden et al. (2005) found that there was a homology between the Mid Wales and Anglesey haplotypes, which also showed a similarity haplotypes in Jersey and the Isle of Wight. Despite a lack of overall genetic structure in the European red squirrel population (Grill et al. 2009), Ogden et al (2005) found a limited level of genetic structure within the Welsh population and to conserve this structure, it was suggested that Mid Wales could be a good source for candidate squirrels for further reintroductions into Anglesey.

Within Mid Wales, there has also been an increase in conservation efforts to help protect the remaining red squirrel population. The Mid Wales Red Squirrel Project (MWRSP) is a partnership between the following organisations: Carmarthenshire, Ceredigion and Powys County Councils, The Wildlife Trust of South and West Wales, Brecknock Wildlife Trust, CCW, Forestry Commission Wales, UPM Tilhill, the MISE project, and private forester Huw Denman (Wilberforce 2012, unpublished report). In 2005, the Wildlife DNA Services received 32 red squirrel hair samples from the Mid Wales area (Ogden and McEwing 2005). 17 of those samples yielded sufficient quantities of DNA to amplify mitochondrial DNA. Three haplotypes were found in the area, ANG1, WC3 and MIDW. The most common haplotype was WC3 which was recorded in 11 samples, followed by ANG1, recorded in five samples and MIDW, recorded in one sample. Ogden and McEwing (2005) suggested that the Welsh red squirrel haplotypes possibly represent ancestral red squirrel stock and are important

to conserve, and management is required to prevent both outbreeding and inbreeding of the population. Of the five haplotypes previously described in Wales, four are described as being closely related to each other, with three of these haplotypes found within the Mid Wales population. The other closely related Welsh haplotype (ANG2) has not been recorded in an extant squirrel, but was found in an old specimen dating to over 20 years ago.

In this study, it was aimed to assess the genetic diversity of red squirrels that have been trapped in the Mid Wales area from 2010 - 2012 using mitochondrial DNA analysis to compare haplotypes from this study with that of Ogden and McEwing (2005). Secondly, microsatellite analysis was used to provide an assessment of the levels of genetic diversity in the Mid Wales red squirrel population. The results from the work were used to infer the origins of the Mid Wales population and to make suggestions for possible breeding and management programmes between other areas of the British Isles where there are red squirrel populations with similar ancestry.

7.2 Materials and Methods

7.2.1 Sample collection and mitochondrial DNA haplotype determination

Trapping of red squirrels took place in Mid Wales from 2008 - 2012 by H. Denman, P. Harries and J. MacPherson (Table 7.1). Samples collected in 2012 were immediately frozen, while samples collected prior to that were stored dry at room temperature. DNA was extracted using the ZR Genomic DNA Kit (Section 2.2.2).

Nucleotide sequencing to resolve the mitochondrial DNA haplotype was conducted as described in Section 2.2.4.2. The region that was chosen in this study overlaps with previous genetic studies of the British red squirrel population (Barratt et al. 1998) and can also be compared to the previous haplotype results obtained from Wales (Ogden et al. 2005; Ogden and McEwing 2005). The primers designed by Ogden et al. (2005) to resolve the mitochondrial DNA haplotype of Welsh squirrels amplified 280 bp and were also designed on the sequences previously described by Barratt et al. (1998), but encompassed 100% of the described haplotypes (Ogden et al. 2005).

7.2.2.1 Mitochondrial DNA Data analysis

Nucleotide sequences were analysed using the BLAST software at http://www.ncbi.nlm.nih.gov/BLAST/ from the GenBank (NCBI) database (Altschul et al. 1990). Sequences were compared by multiple alignments in MEGA V.5.05 (Tamura et al. 2011) and haplotypes were resolved using ARLEQUIN (version 3.5; Excoffier and Lischer 2010). To combine the mitochondrial DNA data from this study with the previously published Welsh dataset by Ogden et al. (2005), and two additional Welsh haplotypes (WC3 and WC9), first published by Barratt et al. (1998), and six haplotypes from Jersey and the Isle of Wight (Simpson et al. 2013). Sequences were truncated to 260 bp. Sequences were compared by multiple alignments in MEGA V.5.05 (Tamura et al. 2011) and haplotypes were resolved using ARLEQUIN (version 3.5; Excoffier and Lischer 2010). A neighbour joining tree was constructed with 1000 bootstraps in MEGA. To compare the haplotypes to previously published haplotypes from Britain and Ireland (Barratt et al. 1998; Hale et al. 2004; Finnegan et al. 2008), sequences were truncated to 186 bp and a median-joining network was constructed

using the median algorithm of Bandelt et al. (1999) in NETWORK, version 4.6 (http://www.fluxus-engineering.com). A number of network diagrams were constructed, the first included haplotypes from Barratt et al. (1998), Hale et al. (2004) and Ogden et al. (2005). Hale et al. (2004) included samples from Britain, Sweden, the Netherlands, Spain and Italy. Due to the large number of haplotypes present in the dataset, the second network diagram only included haplotypes from Finnegan et al. (2008) in addition to the Welsh and island haplotypes (Barratt et al. 1998; Hale et al. 2004; Ogden et al.2005; Finnegan et al. 2008; Dozières et al. 2012; Simpson et al. 2013), to investigate if there were any affinities between those regions.

7.2.2 Microsatellite Amplification

Microsatellite amplification and multiplexes were as described in Section 5.2.3.

7.2.2.1 Microsatellite Data Analysis

Unique individuals, observed (H_O) and expected (H_E) heterozygosities, the number of alleles (A) and allelic richness per locus per sample (A_R) were assessed as described in Section 6.2.2. BOTTLENECK, version 1.2.02 (Piry et al. 1999) was used to evaluate whether any population had undergone a recent bottleneck.

7.3 Results

7.3.1 Mitochondrial DNA

The haplotypes found in this study were WC3 and MIDW, but WC3 was more common than MIDW (Table 7.1). MIDW was previously described in this area by Ogden and McEwing (2005). The neighbour joining tree constructed using sequences from Wales and Jersey and the Isle of Wight, revealed a close homology between the Welsh hap WC3 and some of the Isle of Wight and Jersey haplotypes (Fig. 7.1).

Table 7.1: **Sampling information including the date, location, grid reference and sex of the trapped squirrel.**

Sampling Date	Location	Grid Reference	Sex	DNA Extraction Code	Haplotype
16/01/2008	Cwm Cae'r Odyn	SN740482	M	dom181011.26	WC3
21/01/2008	Cwm Cae'r Odyn	SN740482	M	dom181011.27	WC3
10/12/2008	Nant-ystalwyn, Dôlgoch	SN803575	F	dom181011.28	WC3
10/12/2008	Nant-ystalwyn, Dôlgoch	SN803575	F	dom181011.29	WC3
03/06/2009	Cwm Cae'r Odyn	SN738489	?	dom181011.30	WC3
05/06/2009	Cwm Cae'r Odyn	SN738489	?	dom181011.35	WC3
03/03/2010	Bryn Mawr	SN659499	?	dom181011.32	WC3
16/06/2012	Llethr Gwinau	SN720469	F	ps150812.12	WC3
16/06/2012	Llethr Gwinau	SN720469	F	PS150812.11	WC3
18/06/2012	Bryn Arau Duon	SN720469	F	PS150812.9	WC3
21/06/2012	Llethr Gwinau	SN721469	F	PS150812.10	WC3
23/10/2012	Cwm Cae'r Odyn	SN 7415 4795	M	dom051112.2	WC3
23/10/2012	Cwm Cae'r Odyn	SN 7415 4791	F	dom051112.3	MIDW
24/10/2012	Llethr Gwinau	SN 7216 4677	F	dom051112.5	MIDW
24/10/2012	Cwm Cae'r Odyn	SN 7437 4789	F	dom051112.4	WC3
24/10/2012	Cwm Cae'r Odyn	SN 7437 4789	M	dom051112.6	WC3
25/10/2012	Cwm Cae'r Odyn	SN 7415 4795	M	dom051112.7	WC3
25/10/2012	Cwm Cae'r Odyn	SN 7437 4789	M	dom051112.8	failed
26/10/2012	Cwm Cae'r Odyn	SN 7437 4789	M	dom051112.1	WC3
30/10/2012	Cwm Cae'r Odyn	SN 7437 4789	M	dom201112.2	WC3
30/10/2012	Cwm Cae'r Odyn	SN 7437 4789	M	dom201112.1	failed
31/10/2012	Cwm Cae'r Odyn	SN 7415 4956	F	dom201112.4	WC3
01/11/2012	Cwm Gwenffrwd	SN 7415 4795	M	dom201112.3	WC3

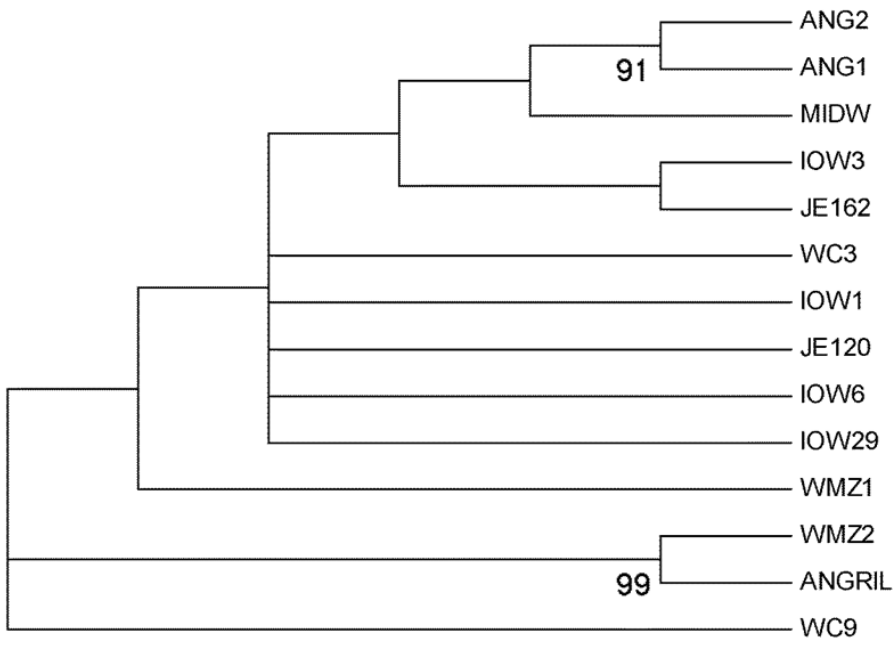

Fig 7.1: **Neighbour joining tree of Welsh, Isle of Wight and Jersey and red squirrel haplotypes.**

The first median joining network was constructed using sequences from Wales, Jersey and the Isle of Wight and haplotypes from Hale et al. (2004) that included some samples from Sweden, Germany, Italy and the Netherlands (Fig. 7.2). The network placed the Welsh haplotype WC3 on the same branch as haplotypes H26, H27 and H28, all of which originated in Cumbria, North England. The other Welsh haplotype found in this study, MIDW grouped with haplotypes from Anglesey, Ang1 and Ang2 and also included haplotypes from Jersey and the Isle of Wight (JE162 and IOW3). Simpson et al. (2013) described those haplotypes as being predominantly English. These two strands were connected by haplotypes H23, H24 and H25, from Cumbria, North England. Interestingly, H23 was directly connected to H16, H17 and H19, all Italian haplotypes. The Isle of Wight haplotype (IOW1) was placed next to H19, again indicating a possible Italian/Central European origin for this group. A Swedish haplotype H18 was also found on this branch, a haplotype also shown to have anthropogenic origins from Central Europe. Welsh hapotypes (WC9, WMZ1,

WMZ2) grouped together with an Isle of Wight haplotype (IOW29) and were connected to H3, H13 and stemmed from H14. H3 was previously recorded in the east of England and H13 was found near Glasgow. However, H14 was recorded in the Netherlands. The Isle of Wight haplotype (IOW6) also stemmed from the Dutch haplotype, which appears to be of Italian/French origin in the network. Two Spanish haps (H11 and H12) were loosely grouped with haplotype H6, (the Netherlands) and H7, H8 and H9 (North England and Scotland).

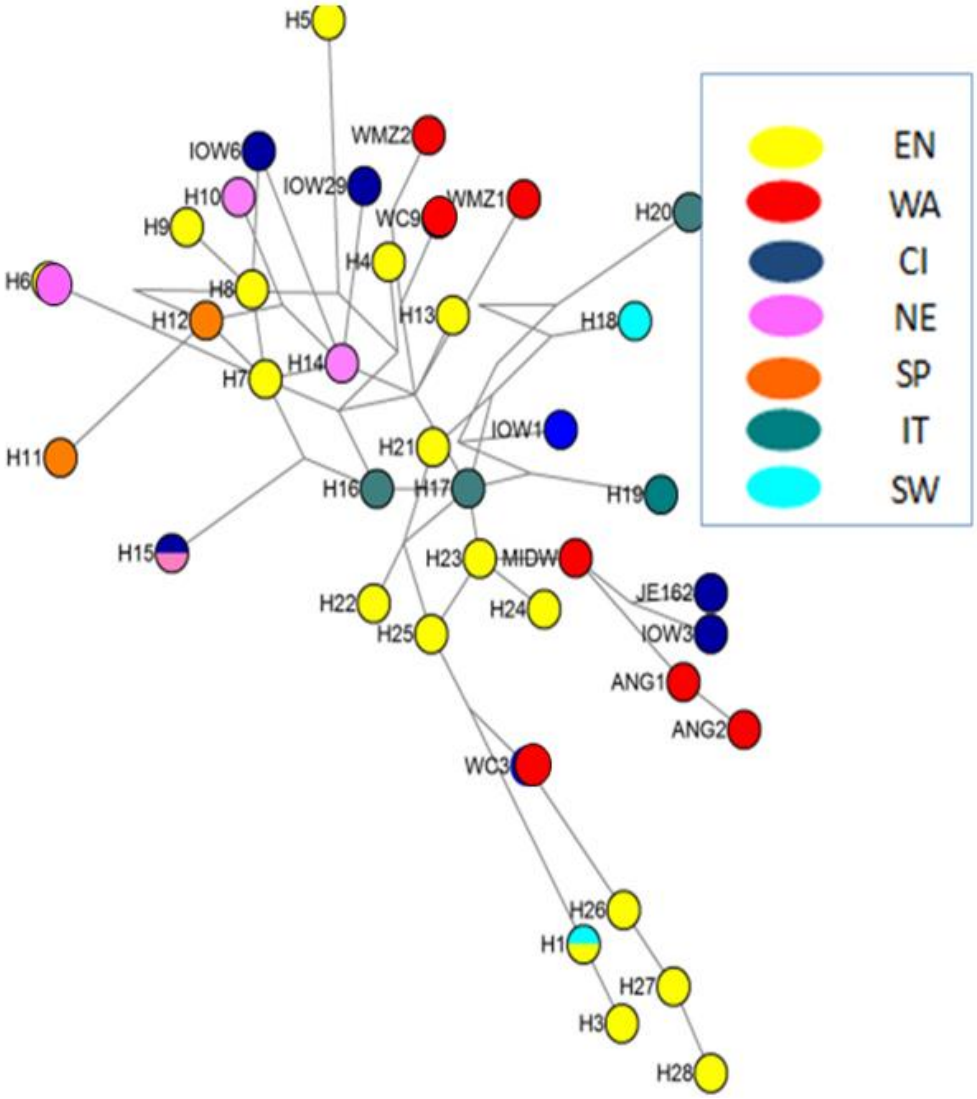

Fig 7.2: **Median joining network of haplotypes from England (EN), Wales (WA), the Channel Islands (CI), the Netherlands (NE), Spain (SP), Italy (IT) and Sweden (SW).**

Finally, the third network (Fig. 7.3) included haplotypes from Wales, Jersey, the Isle of Wight, Ireland, England, the Netherlands, Italy, Spain, France and Sweden, highlighting the mixed origins of the group.

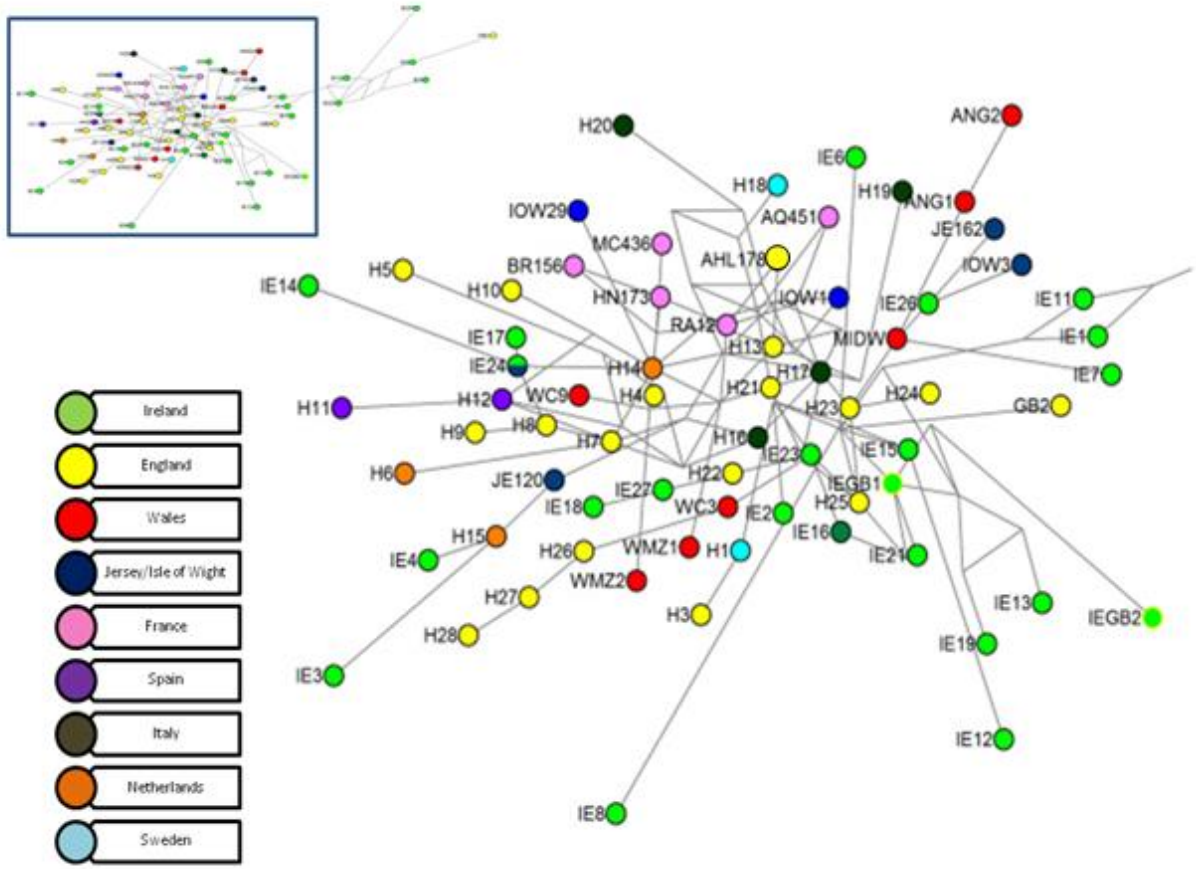

Fig. 7.3: Network diagram of European haplotypes across 186 bp of mitochondrial DNA. ***The haplotypes from the Netherlands (H6 and H15) are orange. AHL178 is a Scottish haplotype.***

7.3.2 Microsatellite Data

As hair samples were used for the source of DNA in this study, a genotype was generated for each sample using two independent PCRs. Where a replicate could not be obtained for a loci, no score was recorded for that sample, or it was repeated. Samples collected and analysed in 2012 produced better quality DNA than samples collected in the previous years. Reliable genotypes were not generated from samples collected before 2012. The number of alleles (A) ranged from 2 – 6 at Scv3 and

miniScv31 and levels of allelic richness (A_R) ranged from 1.84 (miniScv4) to 5.54 (Rsu5). Levels of expected heterozygosity (H_E) ranged from 0.111 at Scv3 to 0.760 at Rsu5 and observed heterozygosity (H_O) values ranged from 0.118 at mini*Scv3* to 1.0 at mini*Scv6* (Table 7.2). Tests for Hardy-Weinberg Equilibrium (HWE) revealed that miniScv4 and miniScv6 showed significant deviations from HWE $P = < 0.001$. Despite the deviations from HWE, there was no evidence of a genetic bottleneck in the population. 16 individuals were identified with at least seven microsatellites. One individual was sampled twice (Table 7.1).

Table 7.2: Descriptive statistics for the Mid Wales red squirrel population

	Rsu5	Lis12	Scv3	miniScv4	miniScv31	miniScv8	Scv6	Scv20	lis3	Mean
N	17	10	17	14	17	17	17	16	17	15.8
A	6	4	2	5	6	3	3	4	3	4.0
A_R	5.4	4.0	1.8	4.9	5.0	3.0	2.6	3.6	2.9	3.7
A_S	185-217	201-207	136-138	112-120	163-181	162-166	176-180	167-217	171-175	
H_O	0.765	0.400	0.118	0.429	0.529	0.706	1.000	0.313	0.824	0.565
H_E	0.760	0.510	0.111	0.724	0.619	0.656	0.528	0.576	0.542	0.558
HWE	0.0165	0.2882	**P<0.001**	0.0070	0.1375	0.1375	**P<0.001**	0.0023	0.1426	

Summary statistics across nine microsatellite markers: *A* = No. of alleles, A_R = Allelic richness, A_s = Allele size range, H_E = Expected Heterozygosity, H_O = Observed Homozygosity, *HWE* = Hardy-Weinberg Equilibrium.

7.4 Discussion

This study has provided evidence of a small but viable red squirrel population in Mid Wales. The mitochondrial DNA analysis was used to identify two haplotypes found in the Mid Wales area, and by analysing the data with previously described haplotypes from the British Isles and Continental Europe, inferences about the history of the Mid Wales red squirrel population could be made. The Mid Wales and Anglesey red squirrels appear to have originated from multiple sources, probably due to multiple introductions into and within Britain (see Chapter 1). There was a connection between some of the Isle of Wight, Jersey, Ireland and Central European haplotypes. The haplotype WC3 also shared some similarities with haplotypes from Cumbria. WC3 is the dominant haplotype in the area and it may be possible that individuals with this haplotype have a selective advantage in the area. The last genetic study that took place in the Mid Wales area by Ogden and McEwing (2005) recorded an additional haplotyoe (ANG1), which was not recorded during this study.

As conservation efforts for red squirrels are taking place throughout Britain, a breeding programme between these various interest groups, especially between Mid Wales, Anglesey and the Isle of Wight and Jersey could be considered. Red squirrels with similar ancestry in Ireland include IE14, which includes squirrels found in the Raven, Co. Wexford that have also been used for a reintroduction project in Co. Dublin, and indicating a potential source for cross-border conservation efforts. Ogden et al. (2005) and Ogden and McEwing (2005) suggested that the Mid Wales population could be used to support the conservation of the Anglesey population, but as this study has shown that there is a viable population of red squirrels in the Mid Wales area, this population is worth monitoring more intensively to evaluate the density of red squirrels in the area. Furthermore, grey squirrel control efforts are on going in the area (Huw Denman, pers comm.) and it may be worth supplementing the current population with additional individuals given the good quality and substantial habitat in the region.

Although the phylogeographic structure of red squirrels in Europe is still largely unresolved due to the large number of introductions that took place in Britain,

Ireland and Europe, there are some local clusters of similar haplotypes that may have shared a common history. For instance, some of the Welsh, Jersey and Isle of Wight haplotypes appear to be similar to samples from the Netherlands, Italy and France (Hale et al. 2004; Grill et al. 2009; Simpson et al. 2013).

The microsatellite data indicated a low to medium level of heterozygosity in the population, with two loci showing significant deviations from HWE. Those two loci also exhibited higher levels of expected versus observed heterozgosity which could be caused by a bottleneck, despite the bottleneck analysis showing that there was no significant evidence of one. In addition, the number of observed alleles were low at these loci, another factor that is observed in populations that have undergone a genetic bottleneck, as rare alleles are the first to be lost in the population resulting in as few as 2-3 alleles (Marshall et al. 2009). However, another possibility of HWE deviations may have been caused by the Wahlund effect that leads to population substructure (Simpson et al. 2013). We have some evidence for this as two different mitochondrial haplotypes were found within the population, and these haplotypes appear to have come from different populations and this may be reflected in the contemporary nuclear DNA results.

7.5 Conclusion

- The Mid Wales red squirrel population has an interesting genetic heritage from multiple sources and conservation efforts should consider the divergent origins of the population.
- Breeding programmes between Mid Wales, Anglesey, Jersey, the Isle of Wight and Southeast Ireland could be considered.

Chapter 8

Summary and future work

8.1 Summary

The loss of woodland cover combined with hunting pressures for the export of skins in the 15th and 16th Centuries resulted in the depletion of the Irish red squirrel population. There is some debate as to how long the red squirrel had been in Ireland prior to this due to a poor representation in the archaeological record. It is clear however, that the contemporary population of squirrels in Ireland originated from introductions that took place from at least ten documented introductions in the 1800s (Barrington 1880). Some of the sources of the reintroductions have been attributed to Britain, but within Britain, introductions also took place from France and other Continental European countries. It was therefore not surprising that previous mitochondrial DNA studies of the red squirrel in Ireland showed that the squirrel exhibited high levels of genetic variability and a large number of mitochondrial DNA haplotypes, which supported the historical introductions (Finnegan et al. 2008).

The grey squirrel was first introduced into Ireland just over 100 years ago and since then has successfully spread throughout the east of Ireland. The presence of the grey squirrel has been attributed to the decline of the red squirrel in the areas where the species overlap. Other factors that have contributed to the red squirrel decline include habitat fragmentation and loss, disease, and weather conditions. As a result of the decline of the red squirrel, there is a need to monitor both species in order to carefully plan the conservation and management of both species.

Methods to survey for squirrels include searching for feeding signs such as stripped pine cones, and using hair-tubes to remotely collect hair samples that can be identified to species using microscopy. In order to increase the information yielded from non-invasive squirrel surveys, a molecular toolbox was developed in this study. This firstly involved the design of species identification assays. This was achieved by developing real-time PCR assays that successfully identified and distinguished both red and grey squirrels. The assays were designed to amplify DNA from poor and or low quality sources such as feeding signs and hair. The assays were also successfully applied to a pine marten molecular dietary study, where it was shown that squirrel

DNA could be detected in the faeces of the species, after it had predated on the target species. This showed that it was possible to detect the presence of both red and grey squirrels from multiple sources, without physically observing or trapping the animals.

In order to further investigate the population dynamics of the red squirrel, methods for molecular sex identification were investigated. The most successful gene for sex identification in this study was the SMCY gene (present on the Y-chromosome) when a large PCR product was amplified ~600 bp. Internal primers designed to amplify much shorter products (~150 bp) by Gorrel et al. (2012) successfully amplified male grey squirrels, but amplified both sexes of red squirrel. In this study, real-time PCR primers were designed to detect male squirrels, but both sexes amplified with the primers. A similar pattern was observed with other other primers developed in this study, targeting other Y-chromosome genes. This may be best explained using the work by Li et al. (2004) that showed that sections of the Y-chromosome were distributed across the squirrel genome. It is likely that the short sequence regions targeted for primer design in this study were amplifying sections of DNA that may have been located throughout the genome. This study has shown the difficulties of developing sex specific markers for species whose genome and evolution is still not properly understood.

To identify individual red squirrels, microsatellite markers were screened, redesigned, and optimised to amplify DNA extracted from hair samples. A panel of nine markers were finally selected that could distinguish individuals. In order to reduce the workload and the number of poor quality samples genotyped, a screening method was introduced based on the ZFX assay from Chapter 4. This assay was used to select DNA samples with a good quality and quantity nuclear DNA content. The panel of microsatellites was then used on the selected samples to identify individual red squirrels from a hair-tube study conducted in Co. Waterford. The mitochondrial D-loop DNA haplotype was also obtained.

In Chapter 6, the individual red squirrels from Co. Waterford were analysed in combination with sub-populations from throughout Ireland to compare the

microsatellite DNA and mitochondrial DNA variation. The microsatellite data showed that all sub-populations were genetically distinct from one another. This was evident using measures of genetic distance, and Bayesian clustering analysis that showed the presence of multiple genetic populations. It was shown that an underlying factor in the genetic distinctiveness was the presence of isolation-by-distance, a term used to describe genetic differentiation caused by geographic separation over generations. This was especially evident in this study and may have been more pronounced by the method of sampling used in this study that selected geographically separated individuals, without sampling intermediate individuals. The mitochondrial DNA analysis also showed that each area contained unique mitochondrial DNA haplotypes, again supporting the separation of each of these subgroups. Both the microsatellite and mitochondrial DNA analysis support the historical data that isolated introductions, from divergent sources took place in Ireland.

The DNA toolbox was finally applied to a red squirrel population in Mid Wales. The analysis showed the presence of two mitochondrial DNA haplotypes that showed a similar genetic history to red squirrels in the Isle of Wight and Jersey. Red squirrel breeding programmes could be established with conservation groups in the core areas of Anglesey, the off shore Islands, and the south of Ireland, to increase the genetic diversity of the current populations.

8.2 Future Work

- DNA extraction kits used for processed food could be tested with squirrel feeding signs to investigate if a higher quality of DNA could be yielded.
- The real-time PCR approach used in this study to detect mammalian prey items could be compared to a next-generation sequencing approach to detect additional prey groups in the pine marten diet. The techniques could also be applied to other species such as foxes or mink, and could be adapted to other countries where additional mammalian prey items are present in the carnivore diet.
- The accurate sex identification of squirrels, especially from non-invasive sources warrants further research. The complexities within the squirrel

genome are poorly understood and could be further investigated using a rad-tag sequencing approach that could be used to identify sex associated regions. Such approaches have recently been used for creating sex linked markers in other species (Anderson et al. 2012; Palaiokostas et al. 2013).

- The molecular techniques optimised to identify individual red squirrels in this study were shown to work successfully with hair samples. However, some problems were encountered with the general use of hair-tubes as success rates were quite low, and substantial effort to collect samples was required. Tubes need to be further optimised to increase the amount of hair collected, and to encourage squirrels to enter the tubes more frequently.
- The genetic analysis of red squirrels in Ireland showed that there are high levels of localised genetic diversity and variability, but the data showed that the populations exhibit high levels of genetic differentiation. This is caused by a combination of factors including the historical introductions of the species, and the sampling approach used in this study. However, future conservation efforts should be considered to increase gene flow. The genetic tools developed in this study are useful for translocation programmes to assess the genetic background of remnant populations and the proposed reintroduction populations to inform the conservation management of the species.
- This study has shown that the red squirrel has a very high number of mitochondrial DNA haplotypes. As a consequence, there is a risk that the species identification assays, especially the TaqMan® probe assay will not be conserved across all populations. Using the SYBR® Green 1 assay may be a better alternative when background information of the population is not known. Another alternative is to resesign the probe sequence following DNA sequence analysis of the target study population.
- In some areas of Ireland, such as the Irish Midlands, a decline has recently been documented in the grey squirrel population, while the red squirrel population has increased (Sheehy 2013). The genetic toolbox developed in this study could be adapted to study the genetics of the grey squirrel in

Ireland to investigate if a genetic bottleneck has occurred in the population that may be contributing to the species decline.

- Red squirrel conservation efforts are on-going in Mid Wales, and further genetic analysis will be conducted on samples collected in 2013 and 2014 to inform longterm conservation management programmes.

References

Aasen E, Medrano JF (1990) Amplification of the ZFX and ZFY genes for sex identification in humans, cattle, sheep and goats. *Bio-Technology*, **8**, 1279-1281.

Ahlering MA, Hailer F, Roberts MT, Foley C (2011) A Simple and accurate method to sex savannah, forest and Asian elephants using noninvasive sampling techniques. *Molecular Ecology Resources*, **11**, 831-834.

Allendorf FW, Luikart G (2007*) Conservation and the genetics of populations*. Blackwell Publishing.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389-3402.

Anderson JL, Rodríguez Marí A, Braasch I, Amores A, Hohenlohe P, Betzel P, Postlethwait JH (2012) Multiple Sex-Associated Regions and a Putative Sex Chromosome in Zebrafish Revealed by RAD Mapping and Population Genomics. *PLoS One*, **7**, e40701.

Avise JC, Arnold J, Ball RM, Bermingham JrE, Lamb T, Neigel JE, Reed CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics,* **18**, 489–522.

Baerwald MR, Schumer G, Schreier BM, May B (2011) TaqMan® assays for the genetic identification of delta smelt (*Hypomesus transpacificus*) and wakasagi smelt (*Hypomesus nipponensis*). *Molecular Ecology Resources*, **11**, 784–787.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Zachary AL, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping ssing dequenced RAD markers. *PLoS One*, **3**, e3376.

Balestrieri A, Remonti L, Ruiz-González A, Vergara M, Capelli E, Gómez-Moliner B, Prigioni C (2010) Food habits of genetically identified pine marten (*Martes martes*) expanding in agricultural lowlands (NW Italy). *Acta Theriologica*, **56**, 199-207.

Bandelt HJ, Forster P, Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, **16**, 37–48.

Barratt EM, Gurnell J, Malarky G, Deville R and Bruford MW (1999) Genetic structure of fragmented populations of red squirrel (*Sciurus vulgaris*) in the UK. *Molecular Ecology*, **8**, 55-63.

Barrington R (1915) Decrease of the squirrel. *The Irish Naturalists' Journal*, **24**, 42.

Barrington RM (1880) On the introduction of the squirrel into Ireland. *Scientific Proceedings of the Royal Dublin Society*, **2**, 615-631.

Beebee T, Rowe G (2008) *An introduction to Molecular Ecology*. Second Edition. Oxford University Press.

Beja-Pereira A, Oliveira R, Alves PC, Schwartz MK, Luikart G (2009) Advancing ecological understandings through technological transformations in noninvasive genetics. *Molecular Ecology Resources*, **9**, 1279-301.

Bertolino S (2008) Introduction of the American grey squirrel (*Sciurus carolinensis*) in Europe, a case study in biological invasion. *Current Science*, **95**, 903-906.

Blouin MS (2003) DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology and Evolution*, **28**, 503-511.

Bylton MDJ, Flannagan NS (2012) A comprehensive guide to GENALEX. Austrailian National University. http://biology.anu.edu.au/GenAlEx/

Brändli L, Lawson Handley L, Vogel P, Perrin N (2005) Evolutionary history of the greater white-toothed shrew (*Crocidura russula*) inferred from analysis of mtDNA, Y, and X chromosome markers. *Molecular Phylogeny and Evolution*, **37**, 832-844.

Broquet T, Ménard N, Petit E (2007) Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. *Conservation Genetics*, **8**, 249-260.

Brubaker JL, Karouna-Renier NK, Chen Y, Jenko K, Sprague DT, Henry PFP (2011) A noninvasive, direct real-time PCR method for sex determination in multiple avian species. *Molecular Ecology Resources*, **11**, 415-417.

Bryja J, Konecny A (2003) Fast sex identification in wild mammals using PCR amplification of the SRY gene. *Folia Zoologica*, **52**, 269-274.

Cao H, Shockey JM (2012) Comparison of TaqMan® and SYBR® Green qPCR methods for quantitative gene expression in tung tree tissues. *Journal of Agricultural and Food Chemistry*, **60**, 12296–12303.

Carey M (2008) Securing Ireland's broadleaf forests and conservation of native red squirrels through integrated management of grey squirrels. (CRISIS). http://www.agriculture.gov.ie/forestservice/publications/projectreports/

Carey M, Hamilton G, Poole A, Lawton C (2007) *The Irish squirrel survey 2007*. COFORD, Dublin.

Caryl FM (2008) *Pine marten diet and habitat use within a managed coniferous forest*. PhD thesis, University of Stirling, Sterling, Scotland.

Caryl FM, Raynor R, Quine CP, Park KJ (2012) The seasonal diet of British pine marten determined from genetically identified scats. *Journal of Zoology*, **288**, 252-259.

Cassens I, Mardulyn P, Milinkovitch MC (2005) Evaluating intraspecific 'network' construction methods using simulated sequence data: do existing algorithms outperform the Global Maximum Parsimony approach? *Systematic Biology*, **54**, 363-372.

Cathey JC, Bickham JW, Patton JC (1998) Introgressive hybridization and nonconcordant evolutionary history of maternal and paternal lineages in North American deer. *Evolution*, **52**, 1224-1229.

Chang SW, Oshida T, Endo H, Nguyen ST, Dang CN, Nguyen DX, Jiang X, Li ZJ, Lin LK (2011) Ancient hybridization and underestimated species diversity in Asian striped squirrels (genus Tamiops): inference from paternal, maternal and biparental markers. *Journal of Zoology*, **285**, 128-138.

Chaves PB, Graeff VG, Lion MB, Oliveira LR, Eizirik E (2012) DNA barcoding meets molecular scatology: short mtDNA sequences for standardized species assignment of carnivore noninvasive samples. *Molecular Ecology Resources* **12**, 18-35.

Cullen C, Smiddy P (2012) Diet of short-eared owls *Asio Flammeus* over seven winters in County Cork, Ireland. *Biology and Environment: Proceedings of the Royal Irish Academy*, **7**, 1-7.

de Marinis AM, Marsseti M (1995) Feeding habits of the pine marten *Martes martes* L., 1758, in Europe: a review. *Hystrix*, **7**, 143-150.

Deagle BE, Jarman SN, Pemberton D, Gales NJ (2005) Genetic screening for prey in the gut contents from a giant squid (*Architeuth*is sp.). *Journal of Heredity*, **96**, 417-423.

Dozières A, Chapuis JL, Thibault S, Baudry E (2012) Genetic structure of the French red squirrel populations: implication for conservation. *PloS One*, **7**, e47607.

Durnin ME, Palsbøll PJ, Ryder OA, McCullough DR (2007) A reliable genetic technique for sex determination of giant panda (*Ailuropoda melanoleuca*) from non-invasively collected hair samples. *Conservation Genetics*, **8**, 715-720.

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

Emerson KJ, Merz CR, Catchen JM, Hohenlohe, PA, Cresko WA, Bradshaw WE, Holzapfel, CM (2010) Resolving post-glacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences USA*, **107**, 16196–16200.

Ermakov OA, Surin VL, Titov SV, Zborovsky SS, Formozov NA (2006) A search for Y-chromosomal species-specific markers and their use for hybridization analysis in ground squirrels (Spermophilus: Rodentia, Sciuridae). *Russian Journal of Genetics*, **42**, 429-438.

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

Excoffier L, Lischer HL (2010) Arlequin version 3.5: a new series of programs to perform population genetic analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567.

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.

Fernandes CA, Ginja C, Pereira I, Tenreiro R, Bruford MW, Santos-Reis M (2007) Species-specific mitochondrial DNA markers for identification of non-invasive

samples from sympatric carnivores in the Iberian Peninsula. *Conservation Genetics*, **9**, 681-690.

Fike JA, Rhodes OE (2009) Characterization of twenty-six polymorphic microsatellite markers for the fox squirrel (*Sciurus niger*) and their utility in gray squirrels (*Sciurus carolinensis*) and red squirrels (*Tamiasciurus hudsonicus*). *Conservation Genetics*, **10**, 1545-1548.

Finnegan L, Hamilton G, Perol J, Rochford J (2007) The use of hair tubes as an indirect method for monitoring red and grey squirrel populations. *Biology and Environment: Proceedings of the Royal Irish Academy*, **107b**, 55-60.

Finnegan LA, Edwards CJ and Rochford JM (2008) Origin of, and conservation units in, the Irish red squirrel (*Sciurus vulgaris*) population. *Conservation Genetics*, **9**, 1099-1109.

Frantz AC, Pope LC, Etherington TR, Wilson GJ, Burke T (2010) Using isolation-by-distance-based approaches to assess the barrier effect of linear landscape elements on badger (*Meles meles*) dispersal. *Molecular Ecology*, **19**, 1663-74.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhinking and background selection. *Genetics*, **147**, 915-925.

Goldberg CS, Woodruff K, Toldness R, Waits LP (2011) Robust molecular sex identification of beaver (*Castor canadensis*) from non-destructive samples. *Conservation Genetics Resources*, **3**, 729-731.

Gorrell JC, Boutin S, Raveh S, Neuhaus P, Côté SD, Coltman, DW (2012) Sexing the Sciuridae: a simple and accurate set of molecular methods to determine sex in tree squirrels, ground squirrels and marmots. *Molecular Ecology Resources*, **12**, 806-809.

Goudet J (1995) FSTAT: a computer program to calculate F statistics. *Journal of Heredity*, **86**, 485–486.

Grill A, Amori G, Aloise G, Lisi I, Tosi G, Wauters LA, Randi E (2009) Molecular phylogeography of European *Sciurus vulgaris*: refuge within refugia? *Molecular Ecology*, **18**, 2687-99.

Gunn MR, Dawson DA, Leviston A, Hartnup K, Davis CS, Strobeck C, Slate J, Coltman DW (2005) Isolation of 18 polymorphic microsatellite loci from the North

American red squirrel, *Tamiasciurus hudsonicus* (Sciuridae, Rodentia), and their cross-utility in other species. *Molecular Ecology Notes*, **5**, 650-653.

Gurnell J, Clark MJ, Lurz PWW, Shirley MDF, Rushton SP (2002) Conserving red squirrels (*Sciurus vulgaris*), mapping and forecasting habitat suitability using a Geographic Information System approach. *Biological Conservation*, **105**, 53-64.

Gurnell J, Lurz PPW, Pepper H (2001) *Practical Techniques for Surveying and Monitoring Squirrels*. Forestry Commission Practice Note 11, Forestry Commission, Edinburgh.

Gurnell J, Lurz PWW, Halliwell EC (2008) Red squirrel (*Sciurus vulgaris)*. In *The Mammals of the British Isles, Handbook* 4th Edition. Harris S, Yalden DW (Ed.)

Gurnell J, Lurz PWW, McDonald R, Pepper H (2009) Practical techniques for surveying and monitoring squirrels. *Forestry Commission Practice Note FCPN011*. Forestry Commission, Edinburgh.

Gurnell J, Lurz PWW, Shirely MDF, Cartmel S, Garson PJ, Magris L, Steele J (2004) Monitoring red squirrels (*Sciurus vulgaris*) and grey squirrels (*Sciurus carolinensis*) in Britain. *Mammal Review*, **34**, 51–74.

Hale ML, Bevan R, Wolff K (2001a) New polymorphic microsatellite markers for the red squirrel (*Sciurus vulgaris*) and their applicability to the grey squirrel (*S. carolinensis*). *Molecular Ecology Notes*, **1**, 47–49.

Hale ML, Lurz PW, Shirley MD, Rushton S, Fuller RM, Wolff K (2001b) Impact of landscape management on the genetic structure of red squirrel populations. *Science,* **293**, 2246–48.

Hale ML, Lurz PWW, Wolff K (2004) Patterns of genetic diversity in the red squirrel (*Sciurus vulgaris* L.): footprints of biogeographic history and artificial introductions. *Conservation Genetics*, **5**, 167–179.

Hales AL, Belant JL, Bird J (2007) Potential for sampling bias in diet studies of American Marten *Martes americana*. *Small Carnivore Conservation*, **37**, 18-20.

Halliwell EC (1997) *The ecology of red squirrels in Scotland in relation to pine marten predation*. PhD thesis, University of Aberdeen, Aberdeen, Scotland.

Hanski IJ, Pöyry M, Kuussaari, Pakkala K (1995) Multiple equilibria in metapopulation dynamics. *Nature*, **377**, 618-621.

Harpending RC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Harvie-Brown JA (1881) *The squirrel in Great Britain*. Proceedings of the Royal Physical Society of Edinburgh.

Hellborg L, Ellegren H (2003) Y chromosome conserved anchored tagged sequences (YCATS) for the analysis of mammalian male-specific DNA. *Molecular Ecology*, **12**, 283-91.

Helyar SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg MT, Cariani A, Maes GE, Diopere E, Carvalho GR, Nielsen EE (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. *Molecular Ecology Resources*, **11**, 123-36.

Henry P, Russello MA (2011) Obtaining high quality DNA from elusive small mammals using low-tech hair snares. *European Journal of Wildlife Research*, **57**, 429–435.

Henry P, Sim Z, Russello MA (2012) Genetic evidence for restricted dispersal along continuous altitudinal gradients in a climate change-sensitive mammal: the American Pika. *PloS One*, **7**, e39077.

Hewitt G (1999) Postglacial recolonization of European Biota. *Biological Journal of the Linnaean Society*, **68**, 87–112.

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

Hickey K (2011) *Wolves in Ireland*. Four Courts Press LTD, Dublin, Open Air.

Hitchings SP, Beebee TJC (1997). Genetic substructuring as a result of barriers to gene flow in urban *Rana temporaria* (common frog) populations: implications for biodiversity conservation. *Heredity*, **79**, 117–127.

Hoffman R (2010) Bugs beasts and business. Some everyday and long-term interactions between biology and economy in preindustrial Europe. In *Economic and biological interactions in pre-industrial Europe from the 13th to the 18th Centuries*. Firenze University Press.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322-1332.

Huxley L (2003).The Grey Squirrel Review. Profile of an invasive alien species. Grey squirrel (*Sciurus carolinensis*). ESI Dorset.

Karmacharya DB, Thapa K, Shrestha R, Dhakal M, and Janecka JE (2011) Noninvasive genetic population survey of snow leopards (*Panthera uncia*) in Kangchenjunga conservation area, Shey Phoksundo National Park and surrounding buffer zones of Nepal. *BMC Research Notes*, **4**, 516-524.

Kassa MT, Penmetsa RV, Carrasquilla-Garcia N, Sarma BK, Datta S, Upadhyaya HD, Varshney RK, von Wettberg EJB, Cook DR (2012). Genetic patterns of domestication in pigeonpea (*Cajanus cajan* (L.) Millsp.) and wild *Cajanus* relatives. *PLoS One*, **7**, 1-13.

Katsushima K, Nishida C, Yosida S, Kato M, Okanoya K, Matsuda Y (2010) A multiplex PCR assay for molecular sexing of the naked mole-rat (*Heterocephalus glaber*). *Molecular Ecology Resources*, **10**, 222-224.

King RA, Read DS, Traugott M, Symondson WOC (2008) Molecular analysis of predation: a review of best practice for DNA-based approaches. *Molecular Ecology*, **17**, 947-63.

Koprowski JL (2005) Management and conservation of tree squirrels, the importance of endemism, species richness, and forest condition. Pp. 245–250 in *Connecting mountain islands and desert seas, biodiversity and management of the Madrean Archipelago II*. United States Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fort Collins, Colorado.

Krauze-Gryz, D, Gryz J, Goszczyński J (2012) Predation by domestic cats in rural areas of central Poland: an assessment based on two methods. *Journal of Zoology*, **288**, 260-266.

Kurose N, Masuda R, Tatara M (2005) Fecal DNA analysis for identifying species and sex of sympatric carnivores: a noninvasive method for conservation on the Tsushima Islands, Japan. *The Journal of Heredity*, **96**, 688-697.

Kusahara M, Kamimura Y, Tamura N, Hayashi F (2006) A new pair of primers for molecular sexing of the Pallas squirrel, *Callosciurus erythraeus*, and variation in fetal sex ratio. *Mammal Study*, **31**, 87-92.

Le Fanu TV (1922) The squirrel in Ireland. The Irish Naturalists' Journal, **31**, 83-84.

Li T, O'Brien PCM, Biltueva L, Fu B, Wang J, Nie W, Ferguson-Smith MA, Graphodatsky AS and Yang F (2004) Evolution of genome organizations of squirrels (Sciuridae) revealed by cross-species chromosome painting. *Chromosome Research*, **12**, 317-335.

Librado P, Rozas J (2009) DnaSP v5: a software for com- prehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 451-1452.

Lurz PWW, Gurnell J, Magris L (2005) Mammalian Species, *Sciurus vulgaris*. *American Society of Mammalogists*, **769**, 1–10.

Lynch AB, McCann Y (2007) The diet of the pine marten (*Martes martes*) in Killarney National Park. *Biology and Environment: Proceedings of the Royal Irish Academy*, **107**, 67-76.

Lynch M, Ritland K (1999) Estimation of pairwise relatedness with molecular markers. *Genetics*, **152**, 1753-1766.

Mac Coitir N (2010) *Ireland's Mammals, Myths, legends and Folklore*. The Collins Press.

Martinoli AD, Bertolino SA, Preatoni, Balduzzi A, Marsan A, Genovesi P, Tosi G, Wauters L (2010) Headcount 2010: The multiplication of the grey squirrel introduced into Italy. *Hystrix*, **21**, 127-136.

Matejusová I, Doig F, Middlemas SJ, Mackay S, DouglasA, Armstrong JD, CunninghamCO, SnowM (2008) Using quantitative real-time PCR to detect salmonid prey in scats of grey *Halichoerus grypus* and harbour *Phoca vitulina* seals in Scotland: an experimental and field study. *Journal of Applied Ecology*, **45**, 632–640.

Matejusová I, Bland F, Hall AJ, Harris RN, Snow M, Douglas A, Middlemas SJ (2013) Real-time PCR assays for the identification of harbor and gray seal species and sex: A molecular tool for ecology and management. *Marine Mammal Science*, **1**, 186-194.

Mayle BA, Broome AC (2013) Changes in the impact and control of an invasive alien, the grey squirrel (Sciurus carolinensis), in Great Britain, as determined from regional surveys. *Pest Management Science*, **3**, 323-333.

McDevitt AD, Vega R, Rambau RV, Yannic G, Herman JS, Hayden TJ, Searle JB (2011) Colonization of Ireland: revisiting "the pygmy shrew syndrome" using mitochondrial, Y chromosomal and microsatellite markers. *Heredity*, **107,** 548-557.

McInnes CJ, Coulter L, Dagleish MP, Deane D, Gilray J, Percival A, Willoughby K, Scantlebury M, Marks N, Graham D, Everest DJ, McGoldrick M, Rochford J, McKay F, Sainsbury AW (2013) The emergence of squirrel pox virus in Ireland. *Animal Conservation*, **1**, 51-59.

Meganathan PR, Dubey B, Jogayya KN, Whitaker N, Haque I (2010) A novel multiplex PCR assay for the identification of Indian crocodiles. *Molecular Ecology Resources*, **10**, 744-747.

Meirmans PG, Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Middleton AD (1929) The red squirrel in Ireland*. The Irish Naturalists' Journal*, **2**, 149-150.

Mills L, Citta J, Lair K, Schwartz M, Tallmon D (2000) Estimating animal abundance using noninvasive DNA sampling: promise and pitfalls. *Ecological Applications*, **10**, 283–294.

Mondol S, Bruford MW, Ramakrishnan U (2013) Demographic loss, genetic structure and the conservation implications for Indian tigers. *Proceedings Royal Society B*, **280**: 20130496.

Montgomery WI, Lundy MG, Reid N (2012) "Invasional meltdown": evidence for unexpected consequences and cumulative impacts of multispecies invasions. *Biological Invasions*, **6**, 111-1125.

Moran S, Turner PD, O'Reilly C (2008) Non-invasive genetic identification of small mammal species using real-time polymerase chain reaction. *Molecular ecology resources*, **8**, 1267-1269.

Morin PA, Luikart G, Wayne RK and the SNP Working Group (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, **19**, 208–216

Mullins J, Statham MJ, Roche T, Turner PD, O'Reilly C (2010) Remotely plucked hair genotyping: a reliable and non-invasive method for censusing pine marten (*Martes martes*, L. 1758) populations. *European Journal of Wildlife Research*, **56**, 443-453.

Murray DC, Bunce M, Cannell BL, Oliver R, Houston J, White NE, Barrero RA, Bellgard MI, Haile J (2011) DNA-based faecal dietary analysis: a comparison of qPCR and high throughput sequencing approaches. *PloS One*, **6**, e25776.

Nei (1972) Genetic distance between populations. *American Naturalist*, **106**, 283-292.

Nei (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583-590.

Nei M and Kumar S (2000) Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Nichols RV, Königsson H, Danell K, Spong G (2012) Browsed twig environmental DNA: diagnostic PCR to identify ungulate species. *Molecular Ecology Resources*, **12**, 983-989.

O'Mahony D, Turner P, O'Reilly C (2012) Pine marten (*Martes martes*) distribution and abundance in Ireland: A cross-jurisdictional analysis using non-invasive genetic survey techniques. *Mammalian Biology*, **77,** 351-357.

O'Meara DB, Sheehy E, Turner PD, O'Mahony D, Harrington AP, Denman H, Lawton C, MacPherson J, O'Reilly C (2014) Non-invasive multi-species monitoring: real-time PCR detection of small mammal and squirrel prey DNA in pine marten (*Martes martes*) scats. *Acta Theriologica*, 59, 111-117.

O'Meara DB, Turner PD, Coffey L, O'Reilly C (2012) TaqMan® assays for species identification of the red squirrel (*Sciurus vulgaris*) and the grey squirrel (*Sciurus carolinensis*). *Conservation Genetics Resources*, **4**, 603-604.

O'Neill D, Turner PD, O'Meara DB, Chadwick EA, Coffey L and O'Reilly C (2013). Development of novel real-time TaqMan® PCR assays for the species and sex identification of otter (*Lutra lutra*) and their application to noninvasive genetic monitoring *Molecular Ecology Resources*, **13**, 877-883.

O'Reilly C, Statham M, Mullins J, Turner PD, O'Mahony D (2007) Efficient species identification of pine marten (*Martes martes*) and red fox (*Vulpes vulpes*)

scats using a 5′ nuclease real-time PCR assay. *Conservation Genetics*, **9**, 735-738.

O'Teangana D, Reilly S, Montgomery WI and Rochford J (2000) Distribution and status of the red squirrel (*Sciurus vulgaris*) and grey squirrel (*Sciurus carolinensis*) in Ireland. *Mammal Review*, **30**, 45-56.

Ogden R, Shuttleworth C, McEwing R, Cesarini S (2005) Genetic management of the red squirrel, *Sciurus vulgaris*: a practical approach to regional conservation. *Conservation Genetics*, **6**, 511–525.

Ogden R, McEwing R (2005) *Genetic Analysis of the Mid-Wales red squirrel (Sciurus vulgaris) population*. Wildlife DNA Services Ltd Pr oject Report No 05/004.

Paetkau D, Waits LP, Clarkson PL, Craighead L, Strobeck C (1997) An empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) populations. *Genetics*, **147**, 1943-1957.

Painter JN, Selonen V, Hanski IK (2004) Microsatellite loci for the Siberian flying squirrel (*Pteromys volans*). *Molecular Ecology Notes*, **4**, 119–121.

Palaiokostas C, Bekaert M, Davie A, Cowan ME, Oral M, Taggart JB, Gharbi K, McAndrew BJ, Penman, DJ, Migaud H (2013) Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. *BMC Genomics*, **14**, 566.

Palsbøll P (1999) Genetic tagging: contemporary molecular ecology. *Biological Journal of the Linnean Society*, **68**, 3-22.

Peakall R, Smouse PE (2006) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Pilgrim KL, Zielinski WJ, Schlexer FV, Schwartz MK (2012) Development of a reliable method for determining sex for a primitive rodent, the Point Arena mountain beaver (*Aplodontia rufa nigra*). *Conservation Genetics Resources*, **4**, 945-977.

Piry S, Luikart G, Cornuet JM (1999) Bottleneck: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, **90**, 502–503.

Pomp D, Good BA, Geisert RD, Corbin CJ, Conley AJ (1995) Sex identification in mammals with polymerase chain reaction and its use to examine sex effects

on diameter of day-10 or -11 pig embryos. *Journal of Animal Science*, **73**, 1408–1415.

Poole A, Lawton C (2009) The translocation and post release settlement of red squirrels Sciurus vulgaris to a previously uninhabited woodland. *Biodiversity and Conservation*, **18**, 3205-3218.

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

Queiros J, Godinho R, Alves PC (2013) The effect of microsatellite selection on population genetic inference: insights from an empirical study using cross-specific and specific markers. *International Mammal Congress*, Queen's University Belfast.

Raymond M. & Rousset F (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248-249.

Reis-Filho JS (2009) Next-generation sequencing. *Breast Cancer Research*, **11** (Supplement 3), S12.

Reports by the juries on the subjects in the thirty classes into which the exhibition was divided. Volume 2. (1851) *Royal Commission, London*. Spicer Brothers, Wholesale Stationers; W Clowes and Sons Printers.

Reumer JWF, van den Hoek Ostende LW (2003) Petauristidae and Sciurudae (Mammalia, Rodentia) from Tegelen, Zuurland, and the Maasvlakte (The Netherlands). *Distribution and Migration of Tertiary Mammals in Eurasia*, **10**, 455-467.

Rosel PE (2003) PCR-based sex determination in Odontocete cetaceans. *Conservation Genetics*, **4**, 647-649.

Rosellini S, Barja I, Piñeiro A (2008) The response of European pine marten (*Martes martes* L.) feeding to the changes of small mammal abundance. *Polish Journal of Ecology*, **56**, 497–504.

Rousset F (2009) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103-106.

Ruiz-González A, Rubines J, Berdión O, Gómez-Moliner BJ (2007) A non-invasive genetic method to identify the sympatric mustelids pine marten (*Martes martes*) and stone marten (*Martes foina*): preliminary distribution survey on

the northern Iberian Peninsula. *European Journal of Wildlife Research*, **54**, 253-261.

Sánchez A, Bullejos M, Burgos M, Hera C, Jiménez R, Diaz de la Guardia R (1996) High sequence identity between SRY HMG box from humans and insectivores. *Mammalian Genome*, **7**, 536-538.

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

Sheehy E (2013) *Investigating the role of the pine marten in Irish squirrel population dynamics.* PhD thesis, National University of Ireland Galway, Galway, Ireland.

Sheehy E, O'Meara DB, O'Reilly C, Smart A, Lawton C (2014) A non-invasive approach to determining pine marten abundance and predation. *European Journal of Wildlife Research*, 60, 223-236.

Shehzad W, Riaz T, Nawaz MA, Miquel C, Poillot C, Shah SA, Pompanon F, Coissac E, Taberlet P (2012) Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Molecular Ecology*, **21**, 1951-1965.

Shibata K, Bandoh K, Yaekashiwa N, Matsuzaki T, Tamate HB (2003) A simple method for isolation of microsatellites from the Japanese squirrel (*Sciurus lis*) without constructing a genomic library. *Molecular Ecology Notes*, **3**, 657–658.

Shorten M (1954) *Squirrels*. Collins, St. James's Place London.

Simpson S, Blackett T, Blampied N, Peniche G, Dozi A, Cornish N, Groombridge JJ (2013) Genetic structure of introduced populations : 120-year-old DNA footprint of historic introduction in an insular small mammal population. *Ecology and Evolution*, **3**, 614-628.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, **139**, 457-462.

Statham M (2005) *Development and use of DNA identification techniques for Irish mustelids*. PhD Thesis, Waterford Institute of Technology, Waterford, Ireland.

Statham M, Turner PD, O'Reilly C (2005) Use of PCR amplification and restriction enzyme digestion of mitochondrial D-loop for identification of mustelids in Ireland. The *Irish Naturalists' Journal*, **28**, 1-6.

Statham MJ, Turner PD, O'Reilly C (2007) Molecular Sex Identification of Five Mustelid Species. *Zoological Studies*, **46**, 600-608.

Steppan S J, Storz BL, Hoffmann RS (2004) Nuclear DNA phylogeny of the squirrels (Mammalia, Rodentia) and the evolution of arboreality from c-myc and RAG1. *Molecular Phylogenetics and Evolution*, **30**, 703–719.

Stewart JR, Lister AM (2001) Crypitc refugia and the origins of the modern biota. *Trends in Ecology and Evolution*, **16**, 608-613.

Strachan R (1998) The mammal detective, British Natural History Series, Whittet Books Ltd.

Sugimoto T, Nagata J, Aramilev VV, Belozor A, Higashi S, McCullough DR (2006). Species and sex identification from faecal samples of sympatric carnivores, Amur leopard and Siberian tiger, in the Russian Far East. *Conservation Genetics*, **7**, 799-802.

Taberlet P and Bouvet J (1992) Bear conservation genetics. *Nature*, **358**, 197.

Tajima F (1989) The effect of change in population size on DNA polymorphism. *Genetics*, **123**, 453-464.

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

Thorington RT Jr, Koprowski JL, SteeleMA, Whatton JF (2012) *Squirrels of the World*. John Hopkins University Press, Baltimore, Maryland.

Todd R (2000) Microsatellite loci in the Eurasian red squirrel, Sciurus vulgaris. *Molecular Ecology*, **9**, 2165-2166.

Tosh DG, Lusby J, Montgomery I, O'Halloran J (2008) First record of greater white-toothed shrew (*Crocidura russula*) in Ireland. *Mammal Review*, **38**, 321-326.

Trizio I, Crestanello B, Galbusera P, Wauters LA, Tosi G, Matthysen E, Hauffe HC (2005) (2005) Geographical distance and physical barriers shape the genetic structure of Eurasian red squirrels (*Sciurus vulgaris*) in the Italian Alps. *Molecular Ecology*, **14**, 469-81.

Valentini A, Miquel C, Ali Nawaz M, Bellemain E, CossiacE, Pompanon F, Gielly L, Cruaud C, Nascetti G, Wincker P, Swenson J, Taberlet P (2009) New perspectives in diet analysis

based on DNA barcoding and parallel pyrosequencing: the trnL approach. *Molecular Ecology Resources*, **9**, 51–60.

Valière N (2002) GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes*, **2**, 377-379.

Vallone PM, Butler JM (2004) AutoDimer: a screening tool for primer-dimer and hairpin structures. *BioTechniques*, **37**, 226-231.

van Oosterhout C, Hutchinson W, Willds D, Shipley P (2004) MICROCHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Vidal O, Pérez-Serra A, Pla C (2010) A sex determination protocol for the Iberian desman (Galemys pyrenaicus) based on a three primer amplification of DBX and DBY fragments with non-invasive samples. *Conservation Genetics*, **11**, 1185-1187.

Waits LP, Paetkau D (2005) Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. *Journal of Wildlife Management*, **69**, 1419-1433.

Warner P, O'Sullivan P (1982) The food of the pine marten (*Martes martes*) in Co. Clare. *Transactions of the International Congress on Game Biology*, **14**, 323-330.

Waters C (2012) *Post release monitoring of two translocated red squirrel (Sciurus vulgaris) populations*. PhD Thesis, National University of Ireland Galway, Galway, Ireland.

Weckworth BV, Musiani M, McDevitt AD, Hebblewhite M, Mariani S (2012) Reconstruction of caribou evolutionary history in Western North America and its implications for conservation. *Molecular Ecology*, **21**, 3610-3624.

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

Wright S (1943) Isolation by distance. *Genetics*, **28**, 114-138.

Zalewski A (2004) *Geographical and seasonal variation in food habits and prey size of the European pine martens*. [In: Martens and fishers (*Martes*) in human-altered environments: an international perspective DJ Harrison, AK Fuller, G Proulx, eds]. Springer, New York, 77-98.

Zhou YB, Newman C, Xu WT, Buesching CD, Zalewski A, Kaneko Y, Macdonald DW, Xie ZQ (2011). Biogeographical variation in the diet of Holarctic martens (genus Martes, mammalia: Carnivora: Mustelidae): adaptive foraging in generalists. *Journal of Biogeography*, **38**, 137-147.

Conservation Genet Resour (2012) 4:603–604
DOI 10.1007/s12686-012-9602-0

TECHNICAL NOTE

TaqMan assays for species identification of the red squirrel (*Sciurus vulgaris*) and the grey squirrel (*Sciurus carolinensis*)

Denise B. O'Meara · Peter D. Turner · Lee Coffey · Catherine O'Reilly

Received: 23 June 2011 / Accepted: 9 January 2012 / Published online: 15 January 2012

Abstract We have developed TaqMan based assays for species-specific identification of two species of squirrel found in the British Isles, the native red squirrel (*Sciurus vulgaris*) and the introduced north American grey squirrel (*Sciurus carolinensis*). These assays correctly identified tissue and hair samples of both species and there was no cross-species amplification. This is a useful method for non-invasive surveys to help conserve the red squirrel and manage the spread of the grey squirrel in the UK and Ireland.

Keywords Real-time PCR · Non-invasive · Hair-tubes · Mitochondrial DNA

Hair-tubes are inexpensive to construct and easy to install, and can potentially provide distribution information on squirrel populations (Gurnell et al. 2001). However, interpretation is limited by the possibility of more than one species depositing a hair sample and by the variations in hair colour found in both red and grey squirrels. Correct identification of samples is important as the species are found to overlap in range prior to the red squirrel being outcompeted. Over recent years, non-invasive genetic sampling has been used to census populations (Mullins et al. 2010). Real-time PCR has been successfully used to species-type fox (*Vulpes vulpes)*, pine marten (*Martes martes),* wood mouse *(Apodemus sylvaticus),* bank vole *(Clethrionomys glareolus),* common shrew *(Sorex araneus),* pygmy shrew *(Sorex minutus)* and water shrew *(Neomys fodiens)* (Moran et al. 2008; O'Reilly et al. 2008; Mullins et al. 2010). The red squirrel is of conservation concern in the UK and Ireland, due to the presence the North American grey squirrel, an introduced competitor (Finnegan et al. 2008). We have developed two TaqMan assays capable of identifying red and grey squirrel.

Red and grey squirrel tissue and hair samples from both Ireland and the UK were used to assess the geographic robustness of the assays (Table 1). Genomic DNA was extracted from tissue samples using the Zymo ZR Genomic DNA II kit (Zymo Research) used according to the manufacturer's instructions. DNA was extracted from hair as described in Mullins et al. (2010). Two species-specific real-time polymerase chain reaction (PCR) TaqMan™ MGB® probe assays were designed to target a short region of mitochondrial D-loop DNA using haplotypes AF1110001–AF111027 (Barratt et al. 1999) and AM412650–AM412675 (Finnegan et al. 2008). Sequences were aligned using MegAlign 5.05 (DNASTAR) and conserved species-specific regions were identified and used to design primers and probes using Primer Express 2 software (Applied Biosystems), targeting species-specific nucleotide polymorphisms towards the 3′ ends of the primers to enhance the specificity and sensitivity of the amplification. Primers and fluorescently labelled probes were purchased from MWG-Biotech AG (Table 2).

PCR with TaqMan MGB labelled probes consisted of 5 µl TaqMan Universal PCR Master Mix (Applied Biosystems), 100 nmol each primer, 100 nmol probe and 1 µl DNA (dilution determined empirically) in a total volume of 10 µl. Negative controls contained molecular grade water instead of DNA. All PCR reactions were carried out in an ABI 7300 real-time PCR system with MicroAmp Optical 96-well reaction plates (Applied Biosystems) and data was analysed using version 2.2.1 of the SDS software (Applied

D. B. O'Meara (✉) · P. D. Turner · L. Coffey · C. O'Reilly
Department of Chemical and Life Sciences, Waterford Institute of Technology, Cork Road, Waterford, Ireland
e-mail: domeara@wit.ie

Table 1 Samples of red and grey squirrels for TaqMan assay validation

Species	Sample type	Location	No. of samples
S. vulgaris	Tissue	Ireland	7
	Plucked hair	Ireland, Wales	30, 3
	Hair tube	Ireland, Scotland	30, 15
S. carolinensis	Tissue	Ireland	5
	Plucked hair	Ireland, England, Scotland	9, 25, 6
	Hair tube	Ireland, England	4, 2

Table 2 D-Loop primer and probe sequences used in TaqMan assay to identify red and grey squirrels

Species	Primer/Probe	Sequence (5′–3′)	Reporter
S. vulgaris	Red_F	TGTGAGTATTAATGTGCATGCTT	
	Red_R	CATAGAACATATCATGTTTAATCAACA	
	Red_P	AGCATGTGGTGGAGGTT	6-FAM
S. carolinensis	Grey_F	GGGGAATGTAAAATTGAAGGG	
	Grey_R	TTGGTCCAGTACAATAAATGTAAGAA	
	Grey_P	AGTATCTATGGACATGCTTATAT	VIC

Biosystems). The PCR program used was 2 min at 50°C, 10 min at 95°C, and 40 cycles of 15 s at 95°C and 1 min at 60°C. Samples were considered positive with a Ct value less than 35. The technique successfully distinguished between red and grey squirrels using tissue and hair DNA. Both assays were shown to be species-specific. The average Ct value for squirrel tissue was 20 and the Ct values for squirrel hair ranged from 19 to 30 depending on the DNA preparation. The primers and probes were found to be applicable to populations throughout Ireland and the UK. However, while primers were designed across a diversity of haplotypes, further analysis across the geographic range of these species would be needed to confirm the applicability outside of the UK and Ireland. The grey squirrel TaqMan assay may also be adapted for use in North America, both in its native range and the western USA where the grey squirrel is also invasive (Fimbel and Freed 2008). As squirrels have sometimes been found in pine marten scats (Lynch and McCann 2007), the TaqMan assays may be further developed to detect squirrels in faecal dietary analyses, similar to the approach taken by Matejusová et al. (2008). The advantage of this study is that information from hair-tube surveys may be improved with the aid of this molecular tool and lead to more effective and efficient management of squirrel populations. This method would also be useful in identifying the species prior to genotyping and subsequent genetic population analysis of non-invasively collected material.

Acknowledgments The authors would like to thank contributors of samples: Mark Ferryman, Juliet Robinson, Richard Wales, Ruth Carden, Colin Lawton, Catherine Waters, Paddy Sleeman, Huw Denman, Pat Whelan, Daniel Buckley, Jacinta Mullins, Siobhan Moran, David O'Neill, Eugene Burke, Ger O'Brien, Ann Fitzpatrick, Fintan Malone and anonymous senders. We would also like to thank property owners for access to their lands: Seán and Síle Murphy, Rosemary Higby and Mount Congreve Estate. This work was funded under the Embark Strand 1 initiative. We are grateful to two anonymous reviewers for valuable comments on the manuscript.

References

Barratt EM, Gurnell J, Malarky G, Deville R, Bruford MW (1999) Genetic structure of fragmented populations of red squirrel (*Sciurus vulgaris*) in the UK. Mol Ecol 8:55–63

Fimbel C, Freed S (2008) Monitoring western gray squirrels for landscape management in western Washington. Northwest Sci 82:299–308

Finnegan LA, Edwards CJ, Rochford JM (2008) Origin of, and conservation units in, the Irish red squirrel (*Sciurus vulgaris*) population. Conserv Genet 9:1099–1109

Gurnell J, Lurz P, Pepper H (2001) Practical techniques for surveying and monitoring squirrels. Forestry Commission

Lynch AB, McCann Y (2007) The diet of the pine marten (*Martes martes*) in Killarney National Park. Biol Environ Proc R Ir Acad 107:67–76

Matejusová I, Doig F, Middlemas SJ, Mackay S, Douglas A, Armstrong JD, Cunningham CO, Snow M (2008) Using quantitative real-time PCR to detect salmonid prey in scats of grey *Halichoerus grypus* and harbour *Phoca vitulina* seals in Scotland: an experimental and field study. J Appl Ecol 45:632–640

Moran S, Turner PD, O'Reilly C (2008) Non-invasive genetic identification of small mammal species using real-time polymerase chain reaction. Mol Ecol Resour 8:1267–1269

Mullins J, Statham MJ, Roche T, Turner PD, O'Reilly C (2010) Remotely plucked hair genotyping: a reliable and non-invasive method for censusing pine marten (*Martes martes*, L. 1758) populations. Eur J Wildl Res 56:443–453

O'Reilly C, Statham M, Mullins J, Turner PD, O'Mahony D (2008) Efficient species identification of pine marten (*Martes martes*) and red fox (*Vulpes vulpes*) scats using a 5′ nuclease real-time PCR assay. Conserv Genet 9:735–773

Acta Theriol
DOI 10.1007/s13364-013-0155-8

ORIGINAL PAPER

Non-invasive multi-species monitoring: real-time PCR detection of small mammal and squirrel prey DNA in pine marten (*Martes martes*) scats

Denise B. O'Meara · Emma Sheehy · Peter D. Turner · Declan O'Mahony · Andrew P. Harrington · Huw Denman · Colin Lawton · Jenny MacPherson · Catherine O'Reilly

Received: 30 April 2013 / Accepted: 27 May 2013

Abstract DNA identification of mammal species occurring in the diet of a predator is potentially a useful approach to remotely monitor the distribution of multiple species. This is important in Ireland, where it has been shown that the combined presence of the introduced bank vole and greater white-toothed shrew impact the distribution of the indigenous small mammals, the wood mouse and pygmy shrew. Direct monitoring of these species and their interactions requires trapping, a labour-intensive and costly approach. In this study, we applied an indirect method by genetically testing the presence of small mammals in pine marten scats collected during the National Pine Marten Survey (2005–2007) to map their distribution. We also included additional scats to investigate if less common prey items, the red squirrel and grey squirrel, could also be detected. This study demonstrates that all target species were genetically detected from pine marten scats. This strategy could be implemented as a monitoring programme for indigenous and introduced mammal species.

Keywords Molecular scatology · Mammals · Invasional meltdown · Ireland · *Apodemus sylvaticus* · *Sorex minutus* · *Myodes glareolus* · *Crocidura russula* · *Sciurus vulgaris* · *Sciurus carolinensis*

Communicated by: Cino Pertoldi

Electronic supplementary material The online version of this article (doi:10.1007/s13364-013-0155-8) contains supplementary material, which is available to authorized users.

D. B. O'Meara (✉) · P. D. Turner · C. O'Reilly
Molecular Ecology Research Group, Waterford Institute of Technology, Cork Road, Waterford, Ireland
e-mail: domeara@wit.ie

E. Sheehy · C. Lawton
Mammal Ecology Group, School of Natural Sciences, Ryan Institute, National University of Ireland, Galway, Ireland

D. O'Mahony
Ecological Management Group, Ormeau Business Park, Belfast BT7 2JA Northern Ireland, UK

A. P. Harrington
Mammals in a Sustainable Environment (MISE) Project, Waterford County Council, Tramore Civic Offices, Tramore, Co., Waterford, Ireland

H. Denman
Selectfor, Brechfa, Carmarthenshire, Wales, UK

J. MacPherson
MISE Project, The Vincent Wildlife Trust, Brechfa, Carmarthenshire, Wales, UK

Published online: 22 June 2013

Introduction

As an island situated on the westernmost periphery of Europe, Ireland has a limited assemblage of flora and fauna in comparison with its nearest neighbour Great Britain, and even more so compared with mainland Europe (Yalden 1999; Searle 2008). Irish mammals consist of a few natural colonisers and many naturalised species that are thought to have been introduced over thousands of years between the arrival of humans and AD 1500 (Searle 2008). A number of more recent introductions of mammal species have occurred over the last 100 years. In some cases, these recent arrivals have threatened the distribution of the current indigenous species' range, especially in the case of small mammals and squirrels (O'Teangana et al. 2000; Carey et al. 2007; Montgomery et al. 2012).

Indigenous small mammal species are limited to the wood mouse (*Apodemus sylvaticus*) and pygmy shrew (*Sorex minutus*). The rat (*Rattus norvegicus*) and house mouse (*Mus musculus*), both historically introduced, are also well

established (Yalden 1999). Recently introduced small mammal species include the bank vole (*Myodes glareolus*) and greater white-toothed shrew (*Crocidura russula*). The bank vole now occupies almost a third of the island in the southwest, and the greater white-toothed shrew is found in Counties Tipperary and Limerick in the south-central part of Ireland (Tosh et al. 2008; Montgomery et al. 2012). The bank vole is believed to have been accidentally introduced into Limerick (west of Ireland) during the 1920s (Stuart et al. 2007). The greater white-toothed shrew was first identified in 2007 in owl pellets collected from County Tipperary, although its origins are not clear (Tosh et al. 2008). Montgomery et al. (2012) found that the combined presence of the bank vole and greater white-toothed shrew had negative impacts on the pygmy shrew distribution, resulting in its replacement, and local extinction; a term coined for this is the 'invasional meltdown.' The indigenous Eurasian red squirrel (*Sciurus vulgaris*) has also suffered declines in recent decades due to the introduction of the North American grey squirrel (*Sciurus carolinensis*) just over 100-years ago (O'Teangana et al. 2000; Carey et al. 2007).

The current distribution maps of all species can be accessed through the National Biodiversity Data Centre's mapping system, 'Data from the Atlas of Mammals in Ireland 2010–2015' (www.biodiversityireland.ie). Small mammals have, in the past, been greatly under-recorded in Ireland, and consequently very few records for wood mouse exist prior to 2010, despite the general assumption that the species is widespread across Ireland. Similarly, the pygmy shrew was also under-recorded prior to a nationwide population genetic study by McDevitt et al. (2009). More visible species such as the red and grey squirrel have been better recorded through public participation surveys (Carey et al. 2007). Consequently, disparities exist in our current ability to monitor less visible species such as small mammals compared with larger species, and such inabilities are now of conservation concern given the recent discovery that introduced species may be displacing indigenous ones.

A possible solution to this may be the wide-scale dietary analysis of carnivore scats such as the pine marten, as its diet in Ireland has been described as generalist and opportunistic (Lynch and McCann 2007). Wood mouse is typically a favoured small mammal in pine marten diet, mostly in the Mediterranean region, whereas vole (bank and field vole) consumption tends to be higher in northern latitudes (de Marinis and Marsseti 1995; Zalewski 2004; Zhou et al. 2011). Greater white-toothed shrew occurred in the pine marten diet in Spain (de Marinis and Marsseti 1995; Rosellini et al. 2008). In Ireland, Lynch and McCann (2007) found wood mouse, pygmy shrew and bank vole in the pine marten diet, and Warner and O'Sullivan (1982) found wood mouse and pygmy shrew (bank vole was not present in the study site). Both studies found a low occurrence of red squirrel in the diet, but grey squirrel and the greater white-toothed shrew were not present in either study site. As pine marten are now distributed across a large proportion of Ireland (O'Mahony et al. 2012), they may well be a suitable predator to remotely study the distribution of indigenous, introduced and invasive species.

Although technically possible, hard part analysis of a large number of predator scats over a large sampling area to infer mammal distribution may not be feasible taking constraints (labour, time, specialist training and overall cost) into consideration. However, over the last number of years, molecular scatology has emerged as a reliable tool to study the diet of mammals (Deagle et al. 2005; Murray et al. 2011; Shehzad et al. 2012; Zarzoso-Lacoste et al. 2013).

When dietary background information is known from previous hard part or molecular dietary analysis, specific species can be targeted using primers designed only to amplify the intended species. This can be addressed using real-time PCR, a technique well established for diagnostic studies (Dooley et al. 2004). Such diagnostic tools have useful applications in dietary studies, as the primers are designed to target short regions of DNA, suitable for detection of low quantity or degraded DNA, an inherent problem in dietary studies due to the breakdown of prey DNA in the predator's gut and environmental degradation. Real-time PCR is also suitable for the detection of prey items that are less abundant in the predator's scat and the use of diagnostic primers enables the detection of low quantity target DNA from mixtures containing more abundant DNA from other items. Successful applications include studies by Matejusová et al. (2008); Bowles et al. (2011), and Murray et al. (2011), where real-time PCR was applied to mammal and bird dietary studies to target species of interest. All studies found that real-time PCR was an accurate and cost-effective approach to address specific dietary questions.

In this study, we used real-time PCR to detect the presence of small mammals and squirrels using DNA extracted from pine marten scats collected during the National Pine Marten Survey (NPMS) 2005–2007 (O'Mahony et al. 2012). As squirrels were expected to be a minor component of the diet, we also collected additional scats from sites where red squirrels were known to occur and tested scats from a captive pine marten where grey squirrel was a food item. The primary purpose of this study was to map the distribution of target prey species, wood mouse, pygmy shrew, bank vole, and greater white-toothed shrew, and to assess if the technique was suitable for the detection of species that were expected to occur at low frequency in the diet, i.e., red squirrel and grey squirrel.

Materials and methods

Sample collection

A set of 252 pine marten scats were used to test for the presence of small mammals and squirrels (Table 1). These

Table 1 Pine marten (PM) sample collection information and detection of red squirrel (RS), grey squirrel (GS), wood mouse (WM), pygmy shrew (PS), bank vole (BV) and greater white-toothed shrew (GWTS) from pine marten scats. Known presence of target species in sample collection (+); known absence of target species (−)

Small mammal and squirrel detection					Squirrel detection				
Target	NPMS		Waterford		Midlands		Captive marten		Total
PM	168		84		223		23		498
RS	1	+	2	+	7	+	0	−	10
GS	0	+	0	−	0	−	12	+	12
WM	30	+	7	+					37
PS	26	+	9	+					35
BV	8	+	0	+					8
GWTS	6	+	0	−					6

samples were collected for the NPMS during the summer months (June–September) 2005–2007 (N=168) and from County Waterford (N=84). These samples originated from 153 10-km^2 sites across Ireland. The NPMS sample locations are mapped in O'Mahony et al. (2012) and Fig. 1. Samples collected in Portlaw, County Waterford (southeast Ireland) were collected in August 2010. As the occurrence of red squirrel in the diet was expected to be low, additional samples were collected from Counties Laois and Offaly (Midlands) (N=223), where the presence of red squirrels and the absence of grey squirrels were confirmed by a squirrel trapping study that took place during the scat collection, from March 2010 to

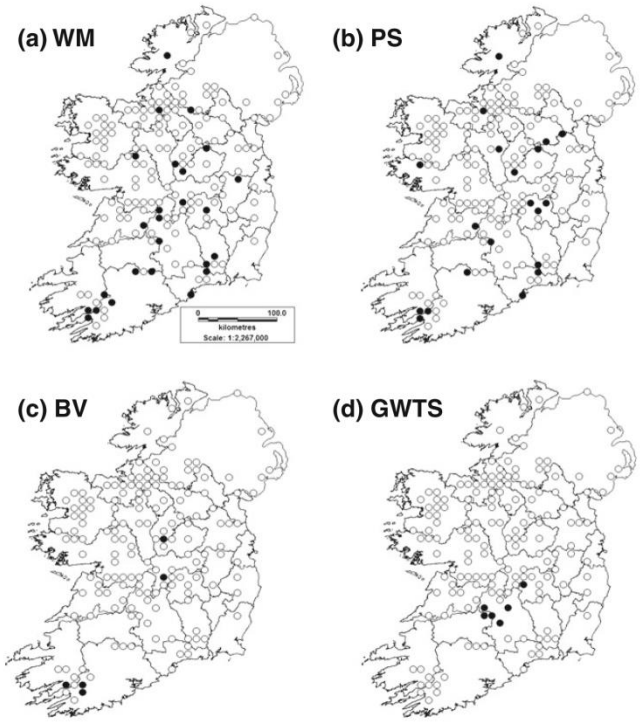

Fig. 1 Distribution of genetically identified small mammal species from DNA extracted from pine marten scats. Sample information can be found on Table 1. *Black circles*: positive detection, *white circles*: negative detection. Wood mouse (WM), pygmy shrew (PS), bank vole (BV), greater white-toothed shrew (GWTS)

September 2011 by E. Sheehy (Table 1). Grey squirrels have not been previously found to occur in the pine marten diet, and to establish a positive control, scats were collected from a captive pine marten that had been fed culled grey squirrels as part of its diet in Wales, UK (Table 1).

DNA analysis

Samples from NPMS were extracted and tested for pine marten DNA as part of the study of O'Mahony et al. (2012). For consistency, the same approach was taken with the remaining samples in this study. In brief, the DNA extraction was consistent with the study of O'Reilly et al. (2008), taking a small section (0.2 g) of faecal material from the outside of the scat. One DNA extract was isolated from each pine marten scat sample and genetically identified (Mullins et al. 2010). The samples from Table 1 (N=252) were tested for the presence of small mammal DNA (wood mouse, pygmy shrew, bank vole and greater white-toothed shrew) using the assays described by Moran et al. (2008), and the greater white toothed shrew assay was designed as part of this study. All scats (N=475) were tested for red and grey DNA using the assays described by O'Meara et al. (2012). Real-time PCR reactions were conducted as described by the associated studies, with the exception that DNA extracts were not diluted (1 μl of neat DNA). All real-time PCR reactions were performed singly, i.e., each target species was tested in isolation due to expected low quantities of target DNA that may be more accurately detected in singleplex real-time PCR reactions. Species were identified based on positive amplification with the species-specific assay and the corresponding Ct value was recorded (see Information box S1). Positive results were replicated at least twice for verification.

Greater white-toothed *shrew assay* design and sequencing

Conventional PCR primers (CrocF-482: 5′-CGCTTCTTCG CATTTCACTTT-3′, CrocR: 5′CATGTTAATGTAAAGAG GTCCGCTAC-3′) for the greater white-toothed shrew were designed to amplify and sequence a 482-bp region of mitochondrial cytb gene using haplotypes published by Brändli et al. (2005) (GenBank accession numbers: AY918341–AY918400). An internal primer (CrocF: 5′-ATAAGCCAAT GCATATTCTGAATTTTAG-3′) was subsequently designed to work in conjunction with CrocR to target a short 52-bp product using real-time RCR (designed to amplify degraded DNA). Primers were designed using Primer Express 2 software (Applied Biosystems) and ordered from Eurofins MWG Operon.

Conventional PCR consisted of 5 μl of GoTaq Hot Start Green Master Mix (Promega), 2 μM of each primer and 1 μl of DNA extract in a total volume of 10 μl. Negative controls contained water instead of DNA. The PCR programme consisted of 95 °C for 5 min followed by 40 cycles of 94 °C for 60 s, 56 °C for 60 s and 72 °C for 60 s followed by 5 min elongation at 72 °C. PCR products were separated and visualised on 2 % agarose gel stained with ethidium bromide. The PCR products were purified using DNA Clean and Concentrator-5 Kit (Zymo Research) and sequenced in both directions with BigDye Terminator Cycle Sequencing Kit 3.1 (Applied Biosystems). Sequences were obtained by running the products on an ABI 310 DNA sequencer (Applied Biosystems) and analysed using BLAST software at http://www.ncbi. nlm.nih.gov/BLAST/ (Altschul et al. 1997). The real-time PCR reaction followed the protocol described in Moran et al. (2008) using 1 μl of undiluted genomic DNA extract.

All genetically identified small mammal species were mapped onto Ordnance Survey Ireland (OSI) maps using MapInfo 11.0 GIS software using the same 10-km^2 grid as O'Mahony et al. (2012). Negative sites were also included for each small mammal species. We used the sample collection from the NPMS (N=168) to assess the occurrence of multiple small mammals in the diet, as this collection covered the largest geographical area and contained the highest number of species.

Results

In the field study, scat samples collected at all sites were firstly tested for the presence of pine marten DNA using real-time PCR and samples with a Ct value ≤33 were deemed positive and suitable for subsequent dietary analysis. This value was selected as a threshold above which it would be difficult to detect prey DNA. Prey DNA extracted from predator scats is known to be of lower quantity and quality than the species that deposited the scat (Matejusová et al. 2008). Genetically testing the pine marten scats also confirmed that the scats were of pine marten origin. The range of positive Ct values for the small mammal and squirrel assays were similar across species 22–37 (Table S2). Grey squirrel DNA was not detected in any of the field collected samples but occurred in 12 of the 23 scats collected from the captive pine marten. The Ct values varied both within and between species, indicating that scats contained varying quantities of target prey DNA ranging from abundant to moderate (see Information box S1).

The greater white-toothed shrew assay that was designed for this study was found to be species-specific. The presence of greater white-toothed shrew was further confirmed by PCR amplification with species-specific primers and DNA sequencing two positive samples. Samples with the lowest Ct values were selected for sequencing, as they contained the highest quantity of target DNA and were more likely to

produce the most accurate sequencing results. The 457-bp consensus sequence obtained from both samples was identical and was a new haplotype to those in the current database and showed maximum homology (99.8 %) to haplotypes from France and Switzerland (Brändli et al. 2005). The haplotype has been deposited in GenBank (accession number: JX424288).

Wood mouse was found to be the most common small mammal encountered in pine marten scats, followed by pygmy shrew, bank vole, greater white-toothed shrew, and red squirrel (Table 1). When mapped, wood mouse was detected in 25 of the 10-km^2 sites, pygmy shrew was found in 26 sites, bank vole in five sites, and greater white-toothed shrew in six sites (Fig. 1).

Based on the NPMS scats, prey DNA was detected from at least one small mammal species (i.e., wood mouse, pygmy shrew, bank vole or greater white-toothed shrew) in 32.74 % of scats. One species of mammal was detected in 25 % of scats, 6.55 % contained two species and 1.79 % contained three species. Wood mouse was most commonly found in single scats, with both wood mouse and pygmy shrew most often found to co-occur in the same scat (N=10), although geographically there was further overlap as some 10-km^2 sites contained more than one scat. The scats that contained DNA from three species contained wood mouse, pygmy shrew and greater white-toothed shrew in one scat, and the second scat contained wood mouse, pygmy shrew and bank vole. The greater white-toothed shrew was found to occur singly in the remaining samples (N=5). Bank vole occurred in the same scat as wood mouse on a single occasion, and bank vole was found not to co-occur with other species (pygmy shrew and wood mouse) in the remaining samples (N=6). Small mammals were also found to overlap geographically in different scats from the same site in several instances, especially in County Waterford.

The distribution of the bank vole found in pine marten scats in this study (Fig. 1) was within the known distribution of the species from 2010 (Montgomery et al. 2012), and despite the widespread distribution of the species, it did not feature prominently across its range in the pine marten diet. However, the distribution of the greater white-toothed shrew was larger than previously described for the species range during 2008 (Tosh et al. 2008; Montgomery et al. 2012). The results show that the greater white-toothed shrew occurred in a sample collected in County Laois, further north than previously described. Although the sample size is small, in the area where greater white-toothed shrew was detected, its occurrence appears not to cluster with other small mammal species more commonly detected elsewhere, and wood mouse and pygmy shrew did not prominently feature.

The red squirrel was detected as a low frequency item but was detected more often in the sample collection from the Irish midlands than in the samples from NPMS and County Waterford. Samples collected from the captive pine marten in Wales showed that grey squirrel could be detected in the pine marten diet, if it was present, but only between one and two days postfeeding. Grey squirrel did not occur in the field-collected scats in this study. Grey squirrels were not known to occur in the scat collection sites in the Irish Midlands or County Waterford, although there was a possibility of overlap between grey squirrel and pine marten ranges in the NPMS sample collection.

Discussion

This study aimed to develop a protocol for the detection of six species in the diet of pine marten, small mammals (wood mouse, pygmy shrew, bank vole and greater white-toothed shrew) and squirrels (red squirrel and grey squirrel), by exploiting a set of previously developed non-invasive species-specific assays; this study also designed two new assays for the detection and sequencing of the recently discovered greater white-toothed shrew. This offers two new approaches for the non-invasive identification of the greater white-toothed shrew, using real-time PCR and conventional PCR. The latter is useful for wildlife labs that do not have real-time PCR facilities, but the sensitivity of that assay is lower due to the larger size DNA fragment. The objective of this work was to investigate a non-invasive method that could be used to detect target species DNA in pine marten scats. This study reports the first successful application of real-time PCR to detect small mammals and squirrels in the diet of genetically identified pine marten scats. It is also the first study to use predator dietary analysis to indirectly map the distribution of indigenous and introduced small mammal prey species.

Small mammal detection

The detection of target small mammal DNA in scats was quite low; 67.23 % of samples from NPMS did not contain target small mammal DNA. However, previous dietary studies of the pine marten in Ireland have shown that other food items such as fruit, birds, reptiles and invertebrates form a significant part of the diet (Warner and O′Sullivan 1982; Lynch and McCann 2007). The occurrence of small mammals in the diet of pine marten in this study was found to be broadly comparable with other Irish studies (Warner and O′Sullivan 1982; Lynch and McCann 2007), with the exception of the pygmy shrew and red squirrel (detected more often in this study) that may be due to the increased sensitivity of real-time PCR (Matejusová et al. (2008).

The wood mouse and pygmy shrew were the most commonly detected species in this study (Table 1, Fig. 1). The pygmy shrew may be perceived to be an unpalatable prey item due to the presence of scent glands, but the species was detected relatively frequently in this study. Other carnivore

dietary studies have also found shrews to be an important item (e.g., Baltrūnaitė 2002). The bank vole presented at a low level in the pine marten diet, despite a greater possibility of it appearing in the diet than the greater white-toothed shrew, as that featured prominently in the pine marten diet within its known range.

It is likely that the greater white-toothed shrew has been longer established in Ireland than the time of its initial discovery in Counties Tipperary and Limerick in 2007 (Tosh et al. 2008). We have shown that the greater white-toothed shrew was present in those counties as well as in County Laois during 2005–2007, Fig. 1). Although the number of samples in this study that contained greater white-toothed shrew were low, it is worth noting that that pine marten may be preferentially predating on the greater white-toothed shrew, or alternatively, other potential prey items such as the pygmy shrew were already displaced by the presence of greater white-toothed shrew and bank vole, as described by Montgomery et al. (2012). The results in this study possibly support the latter. If that is the case, such displacement had occurred as early as 2005–2007. Although this theory needs further investigation, the tools developed in this study can be used to test this.

Squirrel detection

There was no evidence of grey squirrel DNA in any of the field collected scats, but as already mentioned, the majority of scats were collected from sites that were either outside the known range of the grey squirrel or where grey squirrel had been confirmed as absent during sample collection. However, we have shown that grey squirrel can only be detected one to two days after consumption, leaving a short opportunity to detect a species that rarely occurs in the diet. There was also no evidence of grey squirrel in the samples from the NPMS despite the potential for some overlap between the species (Carey et al. 2007; O'Mahony et al. 2012). Red squirrel was detected more often in the samples from the Midlands, but these samples were collected throughout the year and previous studies have shown that the consumption of red squirrel by pine marten is more likely to occur in the winter months in Ireland and Scandinavia (Warner and O'Sullivan 1982; Helldin 1999; Helldin 2000). Samples from NPMS and County Waterford were collected during the summer months and that may explain the slightly lower detection of red squirrel in that sample set. More intensive studies at a finer spatial scale in areas where both pine marten and grey squirrels are confirmed to be present are required to investigate any relationship between the species, in terms of predation or impacts on species range.

Advantages of a molecular approach

The increased throughput of samples with molecular analysis enables efficient large-scale sampling of small mammals and to infer mammal distribution. The detection of mammalian prey items in this study were comparable with previous hard part analysis studies, implying that the DNA extraction protocol and subsequent molecular detection techniques used in this study were found to extract both the host and mammalian prey DNA efficiently.

The primer selection criteria in molecular scatology studies have been described as a critical component for detecting low quantity DNA targets (Zarzoso-Lacoste et al. 2013), but the real-time PCR primers used in this study were developed to be species-specific and to amplify a short amplicon, thus making their application to molecular scatology studies very useful. This study demonstrates an efficient approach to facilitate high-throughput dietary analysis and could be used to screen large numbers of scats prior to additional quantitative hard part analysis, making this technique particularly useful for species that occur in low frequencies such as squirrels.

Warner and O'Sullivan (1982) reported that almost 3 % of the mammal prey items could not be identified to species. Traditional techniques used to identify prey remains from domestic cats in Poland note that intact skulls from small mammals are rarely found in carnivore scats, making identification of small bone remains more difficult (Krauze-Gryz et al. 2012). The visual identification of prey remains can be difficult due to the fragmentation process that takes place during digestion, making molecular detection a practical (and possibly necessary) alternative (Zarzoso-Lacoste et al. 2013).

Both Lynch and McCann (2007) and Warner and O'Sullivan (1982) identified scats by relying on the smell only, a method even expert field pine marten surveyors agree can be misleading (Ruiz-González et al. 2007; Balestrieri et al. 2011; Caryl et al. 2012). Subsequently, the overall dietary content could contain errors if the species had originally been misidentified. In this study, the scats were all identified to species using DNA identification, and the same DNA extract was used for dietary analysis, a cost-effective approach to increase the information gained from a single DNA extraction.

Conclusion

This study demonstrated the use of a predator diet to monitor small mammals in Ireland and has the potential to be used as a technique to monitor the spread and decline of small mammals. This technique can potentially be applied to the dietary analysis of other predators in Ireland including fox (*Vulpes vulpes*), American mink (*Mustela vison*) and even bird of prey pellet analysis. The protocol promises to greatly improve our understanding of multi-species systems and will be valuable for researchers, wildlife conservationists and the overall management of indigenous and introduced species.

Acknowledgments The MISE project is partly funded by the European Development Fund funded by INTERREG IVA (2007–2013). DOM and ES were supported by the EMBARK Strand 1 Initiative and the Irish Research Council for Science, Engineering and Technology (IRCSET), respectively. We would like to thank the associate editor, Allan McDevitt, anonymous reviewers and Jacinta Mullins for their constructive comments to help improve the manuscript.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Balestrieri A, Remonti L, Ruiz-González A, Vergara M, Capelli E, Gómez-Moliner B, Prigioni C (2011) Food habits of genetically identified pine marten (*Martes martes*) expanding in agricultural lowlands (NW Italy). Acta Theriol 56:199–207

Baltrūnaitė L (2002) Diet composition of the red fox (*Vulpes vulpes* L.), pine marten (*Martes martes* L.) and raccoon dog (*Nyctereutes procyonoides* Gray) in clay plain landscape, Lithuania. Acta Zool. Lituanica 4:362–368

Bowles E, Schulte PM, Tollit DJ, Deagle BE, Trites AW (2011) Proportion of prey consumed can be determined from faecal DNA using real-time PCR. Mol Ecol Resour 11:530–40

Brändli L, Lawson Handley L, Vogel P, Perrin N (2005) Evolutionary history of the greater white-toothed shrew (*Crocidura russula*) inferred from analysis of mtDNA, Y, and X chromosome markers. Mol Phylo and Evol 37:832–44

Carey M, Hamilton G, Poole A, Lawton C (2007) The Irish squirrel survey 2007. COFORD, Dublin

Caryl FM, Raynor R, Quine CP, Park KJ (2012) The seasonal diet of British pine marten determined from genetically identified scats. J Zool 288:52–259

de Marinis AM, Marsseti M (1995) Feeding habits of the pine marten *Martes martes* L., 1758, in Europe: a review. Hystrix 7:143–150

Deagle BE, Tollit DJ, Jarman SN, Hindell MA, Trites AW, Gales NJ (2005) Molecular scatology as a tool to study diet: analysis of prey DNA in scats from captive Steller sea lions. Mol Ecol 14:1831–42

Dooley JJ, Paine KE, Garrett SD, Brown HM (2004) Detection of meat species using TaqMan real-time PCR assays. Meat Sci 68:431–438

Helldin JO (1999) Diet, body condition, and reproduction of Eurasian pine martens *Martes martes* during cycles in microtine density. Ecography 22:324–336

Helldin JO (2000) Seasonal diet of pine marten *Martes martes* in southern boreal Sweden. Acta Theriol 45:409–420

Krauze-Gryz D, Gryz J, Goszczyński J (2012) Predation by domestic cats in rural areas of central Poland: an assessment based on two methods. J Zool 288:260–266

Lynch AB, McCann Y (2007) The diet of the pine marten (*Martes martes*) in Killarney National Park. Biol Environ Proc R Ir Acad 107:67–76

Matejusová I, Doig F, Middlemas SJ, Mackay S, Douglas A, Armstrong JD, Cunningham CO, Snow M (2008) Using quantitative real-time PCR to detect salmonid prey in scats of grey (*Halichoerus grypus*) and harbour (*Phoca vituline*) seals in Scotland: an experimental and field study. J Appl Ecol 45:632–640

McDevitt AD, Rambau RV, O'Brien J, McDevitt CD, Hayden T, Searle JB (2009) Genetic variation in Irish pygmy shrews *Sorex minutus* (Soricomorpha: Soricidae): implications for colonization history. Biol J Linn Soc 97:918–927

Montgomery WI, Lundy MG, Reid N (2012) "Invasional meltdown": evidence for unexpected consequences and cumulative impacts of multispecies invasions. Biol Invasions 6:111–1125

Moran S, Turner PD, O'Reilly C (2008) Non-invasive genetic identification of small mammal species using real-time polymerase chain reaction. Mol Ecol Res 8:1267–1269

Mullins J, Statham MJ, Roche T, Turner PD, O'Reilly C (2010) Remotely plucked hair genotyping: a reliable and non-invasive method for censusing pine marten (*Martes martes*, L. 1758) populations. Eur J Wildl Res 56:443–453

Murray DC, Bunce M, Cannell BL, Oliver R, Houston J, White NE, Barrero RA, Bellgard MI, Haile J (2011) DNA-based faecal dietary analysis: a comparison of qPCR and high throughput sequencing approaches. PLoS One 6:e25776

O'Mahony D, Turner P, O'Reilly C (2012) Pine marten (*Martes martes*) distribution and abundance in Ireland: a cross-jurisdictional analysis using non-invasive genetic survey techniques. Mamm Biol 77:351–357

O'Meara DB, Turner PD, Coffey L, O'Reilly C (2012) TaqMan assays for species identification of the red squirrel (*Sciurus vulgaris*) and the grey squirrel (*Sciurus carolinensis*). Conserv Genet Res 4:603–604

O'Reilly C, Statham M, Mullins J, Turner PD, O'Mahony D (2008) Efficient species identification of pine marten (*Martes martes*) and red fox (*Vulpes vulpes*) scats using a 5′ nuclease real-time PCR assay. Conserv Genet 9:735–73

O'Teangana D, Reilly S, Montgomery WI, Rochford J (2000) Distribution and status of the red squirrel (*Sciurus vulgaris*) and grey squirrel (*Sciurus carolinensis*) in Ireland. Mammal Rev 30:45–56

Rosellini S, Barj PA (2008) The response of European pine marten (*Martes martes* L.) feeding to the changes of small mammal abundance. Pol J Ecol 56:497–504

Ruiz-González A, Rubines J, Berdión O, Gómez-Moliner BJ (2007) A non-invasive genetic method to identify the sympatric mustelids pine marten (*Martes martes*) and stone marten (*Martes foina*): preliminary distribution survey on the northern Iberian Peninsula. Eur J Wildl Res 54:253–261

Searle JB (2008) The colonization of Ireland by mammals. In: Davenport JL, Sleeman DP, Woodman PC (eds) Mind the gap: postglacial colonization of Ireland. Irish Nat J Special Supplement: 109–115.

Shehzad W, Riaz T, Nawaz MA, Miquel C, Poillot C, Shah SA, Pompanon F, Coissac E, Taberlet P (2012) Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. Mol Ecol 8:1951–1965

Stuart P, Mirimin L, Cross TF, Sleeman DP, Buckley NJ, Telfer S, Birtles RJ, Kotlik P, Searle JB (2007) The origin of Irish bank voles (Clethrionomys glareolus) assessed by mitochondrial DNA analysis. Irish Nat J 28:440–446

Tosh DG, Lusby J, Montgomery I, O'Halloran J (2008) First record of greater white-toothed shrew *Crocidura russula* in Ireland. Mamm Rev 38:321–326

Warner P, O'Sullivan P (1982) The food of the pine marten (*Martes martes*) in Co. Clare. Trans Intern Congr Game Biol 14:323–330

Yalden DW (1999) The history of British mammals. Poyser, London

Zalewski A (2004) Geographical and seasonal variation in food habits and prey size of the European pine martens. In: Harrison DJ, Fuller AK (eds) Martens and fishers (*Martes*) in human-altered environments: an international perspective. Springer, New York, pp 77–98

Zarzoso-Lacoste D, Corse E, Vidal E (2013) Improving PCR detection of prey in molecular diet studies: importance of group specific primer set selection and extraction protocol performances. Mole Ecol Res 13:117–127

Zhou YB, Newman C, Xu WT, Buesching CD, Zalewski A, Kaneko Y, Macdonald DW, Xie ZQ (2011) Biogeographical variation in the diet of holarctic martens (genus *Martes*, mammalia: Carnivora: Mustelidae): adaptive foraging in generalists. J Biogeogr 38:137–147

Eur J Wildl Res
DOI 10.1007/s10344-013-0771-2

ORIGINAL PAPER

A non-invasive approach to determining pine marten abundance and predation

Emma Sheehy · Denise B. O'Meara · Catherine O'Reilly · Anthony Smart · Colin Lawton

Received: 1 June 2013 / Revised: 2 October 2013 / Accepted: 4 October 2013

Abstract A non-invasive approach was used to investigate variation in pine marten (*Martes martes*) abundance between the midlands and east of Ireland, and to determine the frequency of occurrence of squirrels and other small mammals in the diet. Remotely plucked hair samples were genotyped to differentiate between individual animals, and real-time polymerase chain reaction (PCR) was used to identify predator and prey DNA in scats. Macro analysis of prey remains was carried out on a sub sample of scats and the results from both methods are compared. Non-invasive techniques were successful in determining the presence and relative abundance of the pine marten at woodland level. As expected, abundance was found to be higher in the core population of the midlands than in the east. Pine martens were found to reach higher numbers per km^2 of forested habitat in Ireland than their British or European counterparts. Both traditional hard part analysis and molecular dietary analysis of mammalian prey yielded similar results. We provide the first evidence of the European pine marten predating upon the North American grey squirrel (*Sciurus carolinensis*) in its invasive range. While the grey squirrel was not available as a prey item in any of the midlands sites, it was available in the east, where it featured significantly more frequently in the diet than the native red squirrel. In both the midlands and the east the woodmouse is the most frequently occurring mammal in the diet.

Communicated by C. Gortázar

E. Sheehy (✉) · A. Smart · C. Lawton
Mammal Ecology Group, School of Natural Sciences, Ryan Institute, National University of Ireland, Galway, Ireland
e-mail: emmasheehy@gmail.com

D. B. O'Meara · C. O'Reilly
Molecular Ecology Research Group, Waterford Institute of Technology, Cork Road, Waterford, Ireland

Published online: 23 October 2013

Keywords Pine marten · Squirrel · Hair sampling · Genotyping · Macro faecal analysis · Prey DNA

Introduction

In the nineteenth and twentieth centuries, the European pine marten (*Martes martes*) population in Ireland experienced widespread decline as a result of habitat loss (large-scale deforestation) and heavy persecution (O'Sullivan 1983). Population censuses in the 1980s (O'Sullivan 1983) and again in 2005 (O'Mahony D, O'Reilly C, Turner P 2006 National pine marten survey of Ireland 2005) revealed that pine marten distribution in Ireland is still mainly concentrated around core populations in the west and midlands, along with several smaller populations in the south west and south east of the country. However, the pine marten population in the west and midlands of Ireland has undergone a range expansion in recent decades, as a result of increased habitat availability and connectivity through afforestation, and importantly protection by law (O'Mahony D, O'Reilly C, Turner P 2006 National pine marten survey of Ireland 2005). The most recent population estimate for the island of Ireland is 3060 individuals (O'Mahony et al. 2012), although there is still relatively little known about Irish pine marten population densities in the westernmost part of their European range. The European pine marten has traditionally been considered a forest specialist. Zalewski and Jedrzejewski (2006) estimated that 2 km^2 is the minimum area of forested habitat necessary to support an adult pine marten in the temperate forest zone. Despite an extremely fragmented forest landscape (with <11 % forested area, Ireland represents the lowest forested land cover in their range), previous population studies in Ireland (Lynch et al. 2006; Mullins et al. 2010) have found pine marten density can be higher in Ireland than is typical throughout their British and European ranges (Birks 2002; Zalewski and Jedrzejewski

2006; Mergey et al. 2011; Caryl et al. 2012a). The Irish studies were conducted on the smaller population pockets in the south west (Killarney, Co. Kerry) and the south east (Portlaw, Co. Waterford). A population density estimate for the core population, and also the part of their range where they are believed to be much less common (i.e., the east of the country), has yet to be determined.

In Ireland and Scotland, it has been anecdotally suggested that the recovering pine marten population may be inhibiting the spread of the invasive North American grey squirrel (*Sciurus carolinensis*), and indeed that the grey squirrel population has retracted in the presence of pine martens in both countries (Carey et al. 2007; Paterson and Skipper 2008). Published information on red (*Sciurus vulgaris*) and grey squirrel distribution in Ireland (Carey et al. 2007) and more recent studies on squirrel distribution (unpublished data from Sheehy and Lawton) have found the grey squirrel to be rare in the midlands of Ireland, but their range potentially overlaps with that of the pine marten in the east. It has been suggested that the European pine marten preys preferentially on the larger, less arboreal of the two squirrel species; however, there has been no evidence to date of the European pine marten predating upon the North American grey squirrel anywhere in its invasive range, which only overlaps to a small extent with that of the pine marten (see maps in Carey et al. 2007 and O'Mahony et al. 2012).

Non-invasive genetic studies to identify species distribution and population size have become important tools to aid the study of wild, and particularly elusive, carnivore populations such as martens (Mowat and Paetkau 2002; Williams et al. 2009). Hair sampling is commonly used to non-invasively survey for the presence of mammal species (e.g., Scotts and Craig 1988; Lindenmayer et al. 1999). Lynch et al. (2006) found hair traps (fur-snagging devices) both quick and reliable in detecting pine marten presence in broadleaved woodlands. Subsequently, Mullins et al. (2010) optimised a panel of microsatellite loci to identify unique genotypes within the Irish pine marten population, thus enabling distribution and abundance to be established reliably through non-invasive field studies such as hair trapping. Faecal analysis of scats is used to determine species distribution (e.g., Palomares et al. 2002) and individual identity (e.g., Ruiz-González et al. 2013) of carnivores. The use of molecular techniques in the analysis of carnivore diet has also become popular in recent years (Deagle et al. 2005; Dunshea 2009; Shehzad et al. 2012), as prey DNA found in scats can be identified to taxon and species level and is not dependent on hard parts surviving digestion. Molecular techniques have recently been optimised to specifically detect the presence of mammalian prey in the diet of the Irish pine marten (O'Meara et al. 2013).

Using these recently developed non-invasive techniques, this study firstly aims to quantify pine marten abundance in the fragmented forest habitat in their core range in the midlands of Ireland and in the east of the country where they are considered to be less common (O'Mahony D, O'Reilly C, Turner P 2006 National pine marten survey of Ireland 2005; O'Mahony et al. 2012). Secondly, the study aims to quantify the frequency of occurrence of small mammals in the diet with emphasis on red and grey squirrels. In the process, we aim to quantify scat density in the midlands and eastern regions and to compare the findings of both molecular and macro dietary analysis techniques.

Materials and methods

Field methods

Study area

The primary study area consisted of counties Laois and Offaly, in the midlands of Ireland, and the secondary study area was county Wicklow, in the east of the country, where the pine marten population is considered to be less abundant (O'Mahony et al. 2012) (Fig. 1). Hair trapping sites (abundance study, $n=5$) and scat based survey sites (dietary analysis, $n=23$) are described below.

Abundance study

Five sites were selected, a broadleaved and a predominantly coniferous woodland from each study area (with two broadleaved woods examined sequentially in Co. Wicklow) (Fig. 1). Site 1, Charleville Forest, Co. Offaly, is a mature broadleaved wood (ca. 113 ha) in which oak (*Quercus robur*) is the dominant tree species. Site 2, Clonad (ca. 143 ha), Co. Offaly, is a mixed, mainly coniferous woodland situated 1.5 km from Charleville, where Norway (*Picea abies*) and sitka spruce (*Picea sitchensis*) are the dominant tree species. Site 3, Cloragh (ca. 160 ha) is located in Ashford, Co. Wicklow. The dominant tree species present are sitka spruce and Douglas fir (*Pseudotsuga* spp.). Site 4, Knocksink nature reserve (ca. 60 ha), is located in Enniskerry, Co. Wicklow. It consists of mature oak (*Quercus petraea*) and mixed woodland. Site 5, Tomnafinnoge (ca. 80 ha) in south Wicklow, is a broadleaved woodland consisting mainly of mature oak (*Quercus petraea*) (Fig. 1). Sites 2 and 4 are considered discrete woodlands as there is no forested habitat within 1 km of these sites. Site 1 is separated from adjacent forest habitat by a new primary road. Site 3 is separated from adjacent forestry by the Vartry river. Site 5 is not separated from adjacent forestry by either natural or man-made boundaries, and as such represents the only site which is indiscrete in this study (Fig. 1). Spring based hair traps as described by Messenger and Birks (2000) were installed in sites 1 ($n=5$), 2 ($n=7$), 3 ($n=8$), and 4 ($n=4$) in March 2011 for a period of

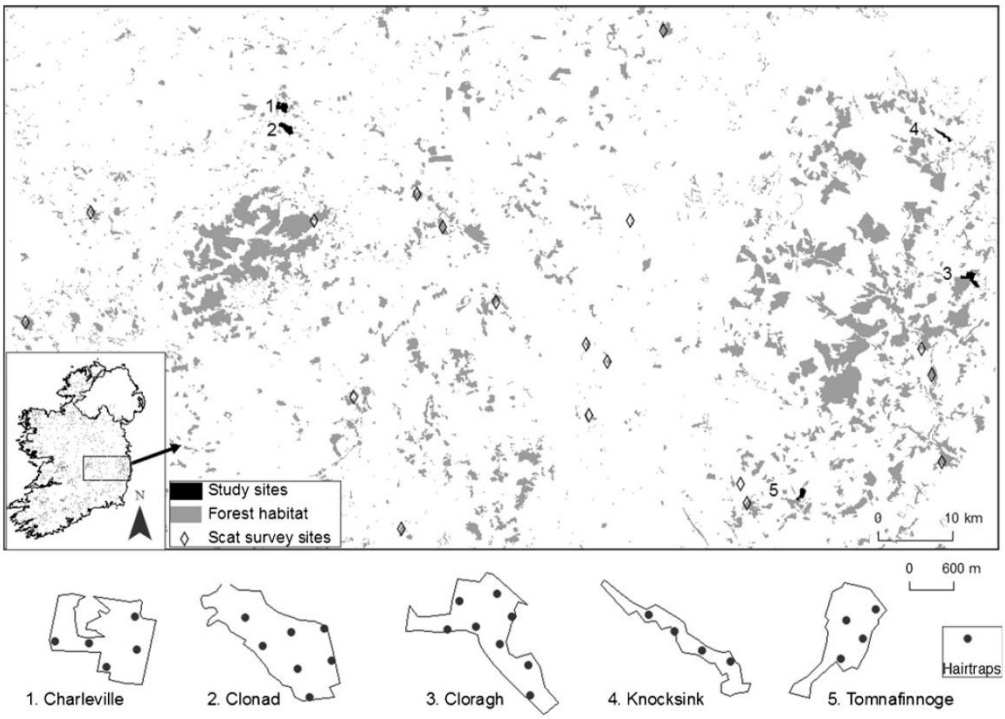

Fig. 1 Location of the study area in Ireland (*inset*) and the locations of the hair trap sites within the fragmented forested landscape. Non-forested land cover is *white*, forested habitat is *grey* and hair trap sites are represented in *black*. Scat survey sites are indicated by *open diamonds*. The position of each hair trap within the study sites are also shown

14 months, with the exception of site 4, where the hair traps were moved to site 5 in October 2011. Each trap was checked for samples and rebaited once per month with chicken and the tree trunk was smeared with marmalade. Traps were positioned 450 m apart at a density of one trap per 20 ha throughout the sites as previous home range sizes for the pine marten in Ireland have been reported as ≥0.2 km^2 (Birks 2002). Animals that were genetically identified in 3 or more months (including at least 1 month between November 2011 and April 2012) were assumed to be resident adults. Abundance values were then obtained by applying the number of residents adults identified at each site to the corresponding forested sampling area.

Dietary study

Scats were collected between March 2010 and August 2012 from 23 sites throughout the midlands and the east of Ireland (Fig. 1) and stored at −20 °C. In order to ensure reliability and validity of both scat density and dietary analysis, all scats collected as part of this study were subjected to DNA analysis in order to confirm they were of pine marten origin.

In order to determine whether a potential prey species was being preyed upon, it was necessary to establish first of all that it was available as a prey item. Of the 23 scat collection sites, 17 were classified as being either red squirrel or grey squirrel positive sites in accordance with the findings of concurrent squirrel distribution studies (Table 1) (unpublished data from Sheehy and Lawton). A sample of scats from the west of Ireland that were collected during the course of a red squirrel population study (Waters and Lawton 2011) were also included. Woodlands where both grey squirrels and pine marten were confirmed as being present together were identified as key sites and revisited where possible to increase the sample size of scats collected from these zones. In August 2012, a scent detection dog, specially trained to detect pine marten scat, carried out searches in three woodlands (including abundance study site 5, Tomnafinnoge and two further sites in the east). With the exception of the searches made by the scent detection dog, and a few scats that were collected opportunistically, the distance covered

Table 1 Scat surveys in sites where squirrel distribution studies had taken place (unpublished data from Sheehy and Lawton; Waters and Lawton 2011) resulted in a total of 361 scats for analysis of squirrel in the diet

Site	Location	RS	GS	RS site scats	GS site scats	RS prey	GS prey
Charleville	M	Y	N	162		4	
Clonad	M	Y	N	109		3	
Abbeyleix Demesne	M	Y	N	7		1	
Birr Castle	M	Y	N	1			
Ballykilcavan	M	Y	N	16			
Cappard	M	Y	N	6			
Emo Park	M	Y	N	9			
Ballyteige	M	Y	N	0			
Garryhinch	M	Y	N	11			
Derryclare	W	Y	N	6			
Croneybyrne	E	Y	N	1			
Clara	E	Y	N	1			
Tomnafinnoge[a]	E	N	Y		20		5
Dollardstown[a]	E	N	Y		8		
Ballygannon	E	N	Y		4		
Mullaghreelan[a]	E	N	Y		0		
Oakpark	E	N	Y		0		
Jenkinstown	E	N	Y		0		
Total				329	32	8	5
						2.4 % FO	15.6 % FO

RS red squirrel, *GS* grey squirrel, *M* midlands, *W* west, *E* east, *Y* squirrel species detected during field survey, *N* squirrel species not detected during field survey

[a] Site re-visited with scent detection dog

during the course of scat collection was used to estimate scat density, where scat density = (no. of scats collected)/(distance walked). All scats were collected by the lead author and transects were walked once per month in the abundance study sites, and in the rest of the sites either once or twice in total. A Fisher exact test was used to test for an overall difference in accuracy in scat identification in the field between the midlands and eastern sites, and to test for a significant difference in the frequency of occurrence of red and grey squirrels in the diet.

Laboratory methods

Abundance study

Molecular analysis

Genomic DNA was isolated from ($n = 158$) hair samples using The ZR Genomic DNA II Kit™ (ZYMO Research, California, USA) using the protocol for hair extraction (ZYMO RESEARCH Cat no. D3040). The DNA was eluted with 100 μl of deionised water. Real-time polymerase chain reaction (PCR) was used for species (targeting mitochondrial DNA) and sex determination (targeting ZFX and ZFY sequences on the X and Y chromosome) of the hair samples as described by Mullins et al. (2010). The C_t value for the ZFX gene was used to screen the samples for genotyping suitability. Samples with a C_t value of less than 36 were deemed to contain adequate quantities of nuclear DNA for genotyping.

Genotyping

Samples that were deemed suitable for genotyping were screened in duplicate at seven loci (Ma2-mini, Mel1, Gg7-mini, Mvi1341, Mvi1354, Mvis075, Ggu234) (Table 2). As the samples used for genotyping came from a non-invasive source (remotely plucked hairs), each sample was independently genotyped twice. Scores were only recorded if they were observed twice and exactly matched. Samples that were not replicated after the first two PCRs were repeated. Details of primers and multiplex setup are provided in Table 2. Fragment Analysis was conducted on an ABI PRISM 310® Genetic Analyser (Applied Biosystems) according to the manufacturer's instructions with the standard run module. Alleles were scored with GS500 LIZ™ size standard using GeneMapper software v3.7 (Applied Biosystems). Two authors independently called alleles.

Data analysis

The two genotyping replicates were compared to assess the data for genotyping errors including the presence of allelic drop out and false alleles using GIMLET version 1.3.4 (Valière 2002). PCR success rates were also calculated using GIMLET version 1.3.4. The occurrence of repeated genotypes was identified using GENALEX version 6.4.1 (Peakall and Smouse 2006) and the number of replicates or individual recaptures was recorded. GENALEX was also used to estimate the probability of identity (PID). A final dataset was created with duplicated data removed and MICROCHECKER version 2.2.3 (Van Oosterhout et al. 2004) was used to further identify possible genotyping errors, including the presence of null alleles, large allele dropout, and scoring errors as a result of stutter peak (using default settings).

Gametic phase linkage disequilibria by Fisher's method (1,000 dememorizations and 5,000 iterations) and deviations from Hardy–Weinberg equilibrium were assessed (default settings, exact tests) using GENEPOP version 4.0.10 (Rousset 2009). Observed (H_O) and expected (H_E) heterozygosities and the number of alleles (a), were calculated using GENALEX version 6.4.1 (Peakall and Smouse 2006), and allelic richness (R_s) was estimated using FSTAT 2.94 (Goudet 1995).

Table 2 Microsatellite primers used in pine marten genotyping

Locus	Primer sequence 5′–3′	Size range	Reference
Ma2-mini	F: YAK-CCATGTACTTTTCCTATCTTTTAGGA R: ATCTTGCATCAACTAAAAAT	131–141	O'Reilly (This study) Davis and Strobeck (1998)
Mel1	F: FAM-CTGGGGAAAATGGCTAAACC R: GCTCTTATAAATCTGAAAATTAGGAATTC	106–116	Bijlsma et al. (2000) Mullins et al. (2010)
Gg7-mini	F: FAM-GTTTTCAATTTTAGCCGTTCTG R: GCTCTTCACTCTGTGTGGCATCTAC	132–140	Davis and Strobeck (1998) O'Reilly (This study)
Mvi1341*	F: PET-GTGGGAGACTGAGATAGGTCA R: GTTTCTTGGCAACTTGAATGGACTAAGA	164–178	Vincent et al. (2003)
Mvi1354*	F: FAM-CCAACTGGAGCAAGTAAAT R: GTTTCTTCATCTTTGGGAAAGTATGTTT	200–212	Vincent et al. (2003)
Mvis075*	F: FAM-GAAATTTGGGGAATGCACTC R: GTTTCTTGGCAGGATAGGATGTGAGCT	145–155	Fleming et al. (1999)
Ggu234	F: PET-TTACTTAGAGGATGATAACTTG R: GAACTCATAGGACTGATAGC	84–90	Duffy et al. (1998)

Reverse primers marked with an asterisk (*) were modified with a 5′ sequence of GTTTCTT to promote non-templated nucleotide addition (Brownstein et al. 1996). The Ma2-F and Gg7-R primers were redesigned to produce a smaller product. The primers were used in two multiplex mixes. Mix A contained Gg7-mini and Mvi1354 and Mix B contained all the other primers. Each primer pair was at a final concentration of 0.5 μM. Microsatellite amplifications were performed in a total volume of 10 μl with 4 μl DNA extract, 1 μl primer mix and 5 μl GoTaq® Hot Start Green Master Mix (Promega). The PCR conditions were 95 °C for 5 min followed by 30 cycles of 95 °C for 30 s, 60 °C for 90 s and 72 °C for 30 s, followed by 72 °C for 30 min

Dietary study

Molecular analysis

Approximately 0.2 g of scat was used for DNA extraction as described by O'Reilly et al. (2008), and using the ZR Genomic DNA II Kit™(ZYMO Research). Pine marten DNA was verified as described above. All samples with a C_t value lower than 32 were classified as pine marten and those with a greater C_t value were classified as non pine marten and excluded from further analysis. To test for prey DNA in the confirmed pine marten scats, species-specific Taqman assays designed to detect red and grey squirrel DNA were used. All PCR reactions and probes were as described by O'Meara et al. (2012). A sub-sample of 160 scats (80 each from sites 1 and 2) were also tested for small mammal prey DNA; woodmouse (*Apodemus sylvaticus*), bank vole (*Myodes glareolus*), pygmy shrew (*Sorex minutus*), and greater white toothed shrew (*Crocidura russula*) (O'Meara et al. 2013). Samples with C_t values of 36 or higher were discounted and positive results were replicated for verification. Percentage frequency of occurrence (%FO) in the diet for each prey species was calculated as the number of scats in which the species' DNA was amplified/total no. of scats tested × 100.

Macro analysis

A sub-sample of 110 scats was subjected to traditional hard part analysis to identify mammalian prey using keys to identify mammal bones (Yalden and Morris 1990) and hairs (Teerink 1991). The subsample of 110 scats comprised 40 scats from both the Charleville and Clonad subsamples, respectively (which had been tested for squirrel and other small mammalian prey DNA), and a further 30 scats from the grey squirrel positive sites (which had been tested for squirrel DNA only). The results from molecular and macro analysis were compared and then combined to determine an overall frequency of occurrence in the diet for each prey species. A chi square test was used to investigate significant differences in the frequency of occurrence of each species according to molecular, macro and combined results. Regression analysis was used to investigate whether a relationship exists between the frequency of occurrence of prey items using molecular and macro techniques. %FO for mammalian prey species was calculated as in molecular analysis and percentage relative biomass of prey ingested (%BPI) was calculated as: weight of dried remains for each species/total weight of dried remains. Previous studies investigating the contribution the main food groups make in terms of biomass to the diet have used pre-established correction factors in such estimations (Lynch and McCann 2007; Caryl et al. 2012b); these correction factors were derived from feeding trials in which the weight of the food item eaten was divided by the dry weight of undigested matter later identified in scats (Lockie 1961; Balharry 1993; Jedrzejewska and Jedrzejewski 1998; Lanszki et al. 2007). Individual correction factors for the mammalian prey species investigated in the current study were not available as they are usually simply grouped together in feeding trials as 'small mammals'.

Results

Density study

A total of 157 hair samples were collected out of 273 baited hair traps. Sites 1 and 2 in the midlands yielded the highest

success rates with 91 % and 78 %, respectively. In site 3, 37 % of potential trapping events yielded hair samples. In site 4, one hair sample was obtained out of a possible 24, and this was the only hair sample in the study to test as negative for pine marten DNA. Site 5 yielded nine hair samples, a success rate of 37.5 %. A further hair sample was collected from a roadkill animal in July 2011, ca. 3 km from site 2, bringing the total number of hair samples to 158.

Of the 158 hair samples, 157 were successfully genetically identified as pine marten and 139 were successfully sex-typed. 109 samples had a ZFX C_t value ≤36 and of these, 104 were successfully genotyped (at six or more loci) (i.e., 95 % of samples that passed the screening process, or 66 % of all hair samples collected). This success rate varied between individual sites (Fig. 2).

A total of 25 individual genotypes were obtained from the 104 samples; one from the roadkill animal near site 2, and 24 from the abundance study sites (site 1, $n=6$ pine marten detected with five hair traps; site 2, $n=10$ pine marten detected with four to seven hair traps; site 3, $n=6$ pine marten detected with eight hair traps; site 5, $n=2$ pine marten detected with four hair traps). Pine marten abundance values ranged from 0 to 4.42 per km^2 including adult residents only (Table 3). Mean abundance values for the midlands and east were 3.13 and 1.01, respectively, with an overall abundance value of 1.99 pine marten/km^2 (Table 3). The number of hair samples genotyped from each individual ranged from 1 to 14 with a mean value of 4.29 (±1.6, 95 % confidence interval [CI]) and the number of months each animal was captured ranged from 1 to 7, with a mean value of 3.29 (±0.96, 95 % CI) (Table 4). One individual (a male) was detected in both site 1 and site 2 (January 2012 and May 2011, respectively). These sites are located 1.5 km apart (Fig. 1) and together they comprise less than 3 km^2 of forested habitat. In total, eight adult residents were detected in the two sites and a further eight that are assumed to be either sub-adult or non-resident individuals. With the exception of the one animal there was no further crossover of individuals detected between these two woodlands, despite their close proximity to one another and the lack of surrounding forested habitat.

Assessing genotyping errors

The proportion of positive PCRs ranged from 90 % to 100 % across loci and from 86 % to 100 % across samples. Analysis of genotyping error revealed the presence of allelic dropout rates of 0.08 at locus Ma2-mini, 0.10 at locus Ggu234 and 0.46 at locus Mvis1354, with no false alleles detected. The overall allelic dropout error rate across all loci was 0.09, 0.32 across all samples, and 0.08 across all PCRs. We found that there was no systemic evidence of scoring errors and the data was not shown to be affected by the systemic presence of null alleles or large allelic dropout. The cumulative PID was PI=1.9×10^{-3}, which is sufficient for the estimation of population size (0.01) (Mills et al. 2000).

Genetic variability

The number of alleles was low ranging from 2 at Ggu234 and Mvis075 to 4 at Gg7-mini (Table 5). Low levels of allelic richness per locus and per sample were also observed from 2 at Ggu234 and Mvis075 to 3.69 at Gg7-mini. Expected levels of heterozygosity averaged 0.386 and ranged from 0.106 at Mvis1354 to 0.570 at Ma2-mini. Observed levels of heterozygosity averaged 0.462 and ranged from 0.111 at Mvis1354 to 0.708 at Gg7-mini and Ma2-mini (Table 5). There were no significant deviations from Hardy–Weinberg expectations at any loci.

Dietary study

A total of 517 scats were collected between March 2010 and August 2011. Four hundred of these were collected in the midlands ($n=9$ sites) and 117 were collected in the eastern sites ($n=14$ sites) (Fig. 1). Overall, 86 % of scats collected in the midlands tested positive for pine marten DNA compared to 39 % in the east. As such, accuracy in the field was found to be significantly lower in the eastern region ($p<0.001$, Fisher exact). The majority of scats were included in scat density calculations as distance walked was known (333 out of 344 and 38 out of 46 in midlands and eastern sites, respectively). Thus scat density was estimated to be 1.745 scats/km in the midlands and 0.221 scats/km in the east. Although scat density was found to be higher in areas of higher pine marten occupancy, no statistically significant relationship was found between scat density and pine marten abundance. The scent detection dog succeeded in collecting seven pine marten scats from two of three woods visited over a 2-day period. When scats were categorised into red squirrel and grey squirrel sites (as per unpublished data from Sheehy and Lawton) a total of 329 were classified as coming from red squirrel positive sites, and 32 from grey squirrel positive sites (Table 1). Squirrels appeared in the diet as prey items at low frequencies during 9 months of the year, spring and early summer being most common (Table 6).

Molecular and macro analysis

Regression analysis found a linear relationship to exist between the %FO of mammalian prey items as detected by molecular and macro analyses ($y=-1.694+1.969x$, $R^2=0.895$, $p<0.05$) (Fig. 3). The woodmouse featured more frequently in the diet than any other mammal species in both the molecular ($\chi^2=41.17$) and the macro ($\chi^2=67.58$) analyses ($df=4$, $p<0.01$) (Fig. 4). Grey squirrels featured more frequently than red squirrels as prey items in both analyses,

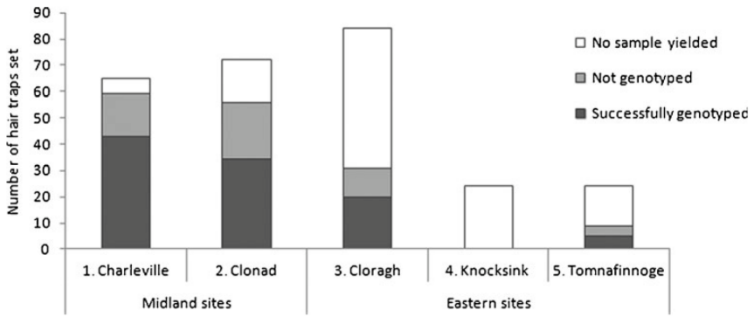

Fig. 2 The number of successful hair trapping events and the portion of hair samples successfully genotyped from each site

significantly so in the molecular analysis, despite the considerable difference in sample size ($p < 0.05$, Fisher exact).

When results from the reduced sample ($n = 110$) that was the subject of both molecular and macro analyses were combined, the %FO increased for all species; however, there was very little effect on the low frequency prey items (red squirrel, bank vole and pygmy shrew). The %FO for woodmouse and grey squirrel increased more considerably when results from both methods were combined, but not significantly so (Fig. 5). %BPI also found the woodmouse to be the most important prey species in the diet of the pine marten, followed by the grey squirrel where it was available, although ca. 30 % of mammalian remains could not be identified to species level (Fig. 6).

Discussion

Pine marten abundance

This study has provided an index of abundance for the pine marten population in both their core range and a considerably less populated part of their range in the east of Ireland. Abundance values are not directly comparable to European studies where density values were obtained from radio-tracking or snow tracking (e.g., Zalewski and Jedrzejewski 2006; Mergey et al. 2011), but are comparable to Irish studies that have used a combination of hair trapping and live-trapping to determine population density estimates. The abundance values obtained in this study suggest that the population in the midlands of Ireland (mean value of 3.13 adult residents/km^2) is currently living at a higher density than previously reported for the species in Europe (0.01–1.75 per km^2) (Zalewski and Jedrzejewski 2006) or Ireland (0.5–2 per km^2) (Lynch et al. 2006; Mullins et al. 2010) and thus quite possibly represents the highest density in their natural range. It is unclear why the European pine marten reaches these relatively high numbers in Ireland, particularly when their favoured habitat is so sparse and fragmented. In a review of available literature on European pine marten densities, Zalewski and Jedrzejewski (2006) found that between 41° and 68°N densities declined exponentially with decreasing winter temperature and increasing seasonality, and suggested that both winter severity and availability of rodents are

Table 3 The total number of pine marten identified by unique genotypes at each site, in total (All), in sites 1 and 2 combined (Midlands) and in sites 3, 4 and 5 combined (East)

	Site name	Site area	Total PM identified	Mean captures	No. adult residents	No. adult residents per km^2
Midlands	1. Charleville	1.13 km^2	6	4.67 (0.98)	5	4.42
	2. Clonad	1.43 km^2	10	2.8 (0.79)	3	2.10
East	3. Cloragh	1.58 km^2	6	2.83 (0.75)	2	1.27
	4. Knocksink	0.6 km^2	0	0.00	0	0.00
	5. Tomnafinnoge	0.8 km^2	2	2.5 (1.5)	1	1.25
Total	All	5.54 km^2	24	3.29 (0.46)	11	1.99
	Midlands	2.56 km^2	16	3.56 (0.63)	8	3.13
	East	2.98 km^2	8	2.75 (0.62)	3	1.01

Mean Captures = mean number of times each animal was captured per site (Std Er). No. adult residents = number of animals that were detected in ≥3 months including at least 1 month during November 2011 and April 2012

Table 4 The site each pine marten was recorded at and the sex assigned through DNA analysis

Animal	Site	Sex	GT	Months
CVF01	1. Charleville	Female	1	1
CVF02			3	3
CVF03			10	7
CVF04			6	4
CVM01		Male	9	7
CVM02			14	6
CDM05[a]			1	1
CDF01	2. Clonad	Female	4	4
CDF02			1	1
CDF03			3	3
CDF04			1	1
CDM01		Male	1	1
CDM02			3	2
CDM03			1	1
CDM04			12	8
CDM05[a]			1	1
CDM06			7	6
WWF01	3. Cloragh	Female	2	2
WWF02			3	3
WWM01		Male	7	5
WWM02			1	1
WWM03			6	5
WWM04			1	1
TFF01	5. Tomnafinnoge	Female	4	4
TFF02		Male	1	1
Roadkill			1	
Total			104	

GT the number of hair samples successfully genotyped for each animal, *Months* the number of months each animal was identified

[a] Animal was captured at two sites

limiting factors on populations. Thus it is possible that Ireland's relative lack of seasonality and mild winters (the moderating influence of the Atlantic gulf stream results in mean minimum winter temperatures of 2–6 °C) (MetEireann 2012) contribute to the observed high pine marten abundance. Other contributory factors may include lack of competition and lack of predators, with the red fox (*Vulpes vulpes*) representing the pine marten's only real competitor or predator in Ireland.

Zalewski and Jedrzejewski (2006) estimated that 2 km^2 is the minimum area of forested habitat necessary to support an adult pine marten in the temperate forest zone. However, this does not appear to apply to the core Irish pine marten population and both midlands sites, Charleville (1.13 km^2) and Clonad (1.43 km^2), sustain relatively high pine marten numbers in comparison to other extensively forested parts of Europe (Zalewski and Jedrzejewski 2006). The European pine marten has traditionally been considered a forest specialist; however, recent studies have found the species to be less

Table 6 Squirrels as detected as prey items in molecular and macro analysis of pine marten scats, including the date and site at which the scat was collected

Species	Date	Site	Molecular	Macro
RS	2010-03-22	Abbeyleix	Y	
RS	2010-11-16	Clonad	Y	N
RS	2011-03-31	Charleville	Y	Y
RS	2011-04-01	Charleville	Y	
RS	2011-05-11	Charleville	Y	
RS	2011-06-07	Clonad	Y	Y
RS	2011-06-07	Clonad	N	Y
RS	2011-09-30	Charleville	Y	
GS	2011-10-01	Tomnafinnoge	Y	
GS	2012-02-03	Tomnafinnoge	N	Y
GS	2012-03-13	Tomnafinnoge	Y	Y
GS	2012-04-18	Tomnafinnoge	N	Y
GS	2012-05-18	Tomnafinnoge	Y	N

RS red squirrel, *GS* grey squirrel, *Y* positive, *N* negative (blank = not tested)

Table 5 Descriptive statistics for microsatellite analysis of pine martens in four study sites in Ireland

	Ggu234	Mel1	Gg7-mini	Ma2-mini	Mvis075	Mvi1341	Mvi1354	Mean
N	24	24	24	24	24	24	18	23.14
a	2	3	4	3	2	3	3	2.86
R_s	2.0	2.99	3.69	3.0	2.0	2.75	3.0	2.78
a_s	87–93	108–116	132–142	131–137	151–153	170–180	200–208	
H_E	0.353	0.379	0.548	0.570	0.305	0.442	0.106	0.386
H_O	0.458	0.417	0.708	0.708	0.292	0.542	0.111	0.462
HW	0.145	0.861	0.775	0.889	0.834	0.632	0.996	

N denotes the number of individuals that successfully amplified at each locus, a is the number of alleles per locus, R_s is the allele size range, A_s is the allele size, H_E is the expected heterozygosity, H_O is the observed heterozygosity. There were no significant deviations from Hardy–Weinberg equilibrium

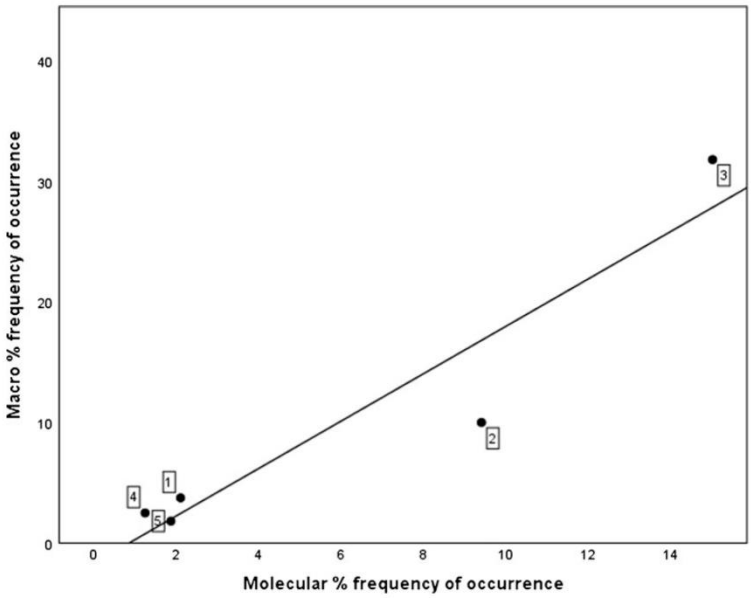

Fig. 3 A linear relationship was found to exist between the frequency of occurrence of mammalian prey items as determined using molecular and macro analyses ($y=-1.694+1.969x$, $R^2=0.895$, $p<0.05$)

restricted to large forests than previously believed and highlighted the importance of the surrounding landscape not only in providing habitat corridors but also in providing essential food resources and den sites throughout fragmented landscapes (Clevenger 1994; Pereboom et al. 2008; Mergey et al. 2011; Caryl et al. 2012a). In the current study, hair traps were only placed within the forested habitat and thus no data was obtained on the use of the surrounding landscape, or in the case of site 5, the adjacent forestry. Abundance values per km^2 are thus only applicable to forested area as use of surrounding landscape is not accounted for with this sampling technique. The fact that only one animal was detected both in sites 1 and 2 despite their close proximity to one another, supports the theory that these small woodlands can be

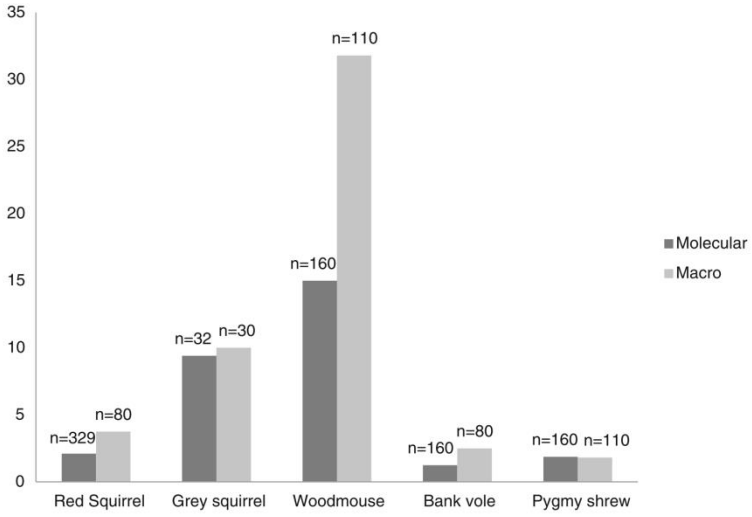

Fig. 4 Overall results for molecular and macro analysis of mammalian prey species in pine marten diet (n=number of pine marten scats tested). Those subject to macro analysis are a random sub-sample of the molecular samples except the sample from the grey squirrel positive sites. The woodmouse features significantly more frequently than any other prey species (molecular and macro analysis) and macro analysis detected significantly more woodmouse in the diet than molecular analysis ($p<0.01$, Fisher exact). Grey squirrels were more frequently detected than red squirrels as prey items (molecular analysis: $p<0.05$, Fisher exact)

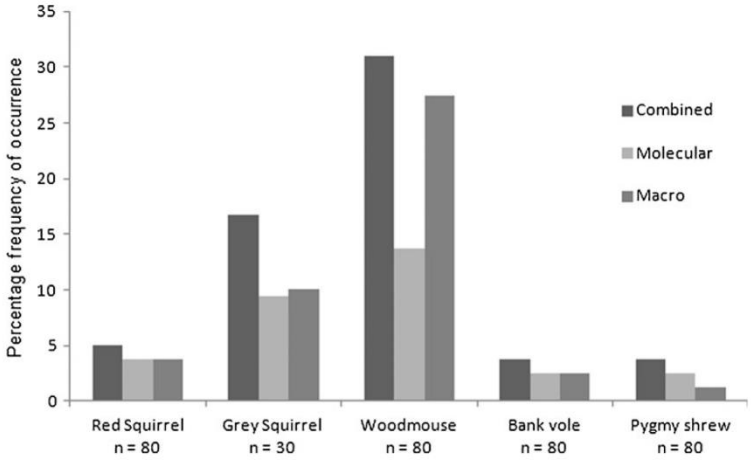

Fig. 5 Comparison of results from both molecular and macro analysis of pine marten scats, where all samples were subjected to both techniques (n = sample size). Frequency of occurrence was higher for all species detected when results were combined, and significantly so for the woodmouse (χ^2=6.04, df=1, p<0.05)

considered as relatively discrete in terms of pine marten occupancy; however, it is likely that surrounding non-forested landscape is also used to some extent.

Non-invasive techniques

When compared to other methods of determining pine marten density, such as snow-tracking (e.g., Zalewski 1999) (not feasible in Ireland), camera-trapping (e.g., Manzo et al. 2012) and radio-tracking, there are certain benefits and limitations to genetic tagging through hair samples. Genetic tagging does not provide specific biological information such as weight, reproductive activity and condition, nor does it provide detailed spatial information on home range size or territoriality. It is however the only non-invasive method that confidently differentiates between individual animals. The data collection process was non intrusive to the animal and time-efficient in terms of data collection. The hair traps themselves are inexpensive to construct and maintain. However it is clear from our data that hair trap density should be increased to optimise information gathered on the population. In the midlands sites there were more individuals detected per site over the study period than there were hair traps available per

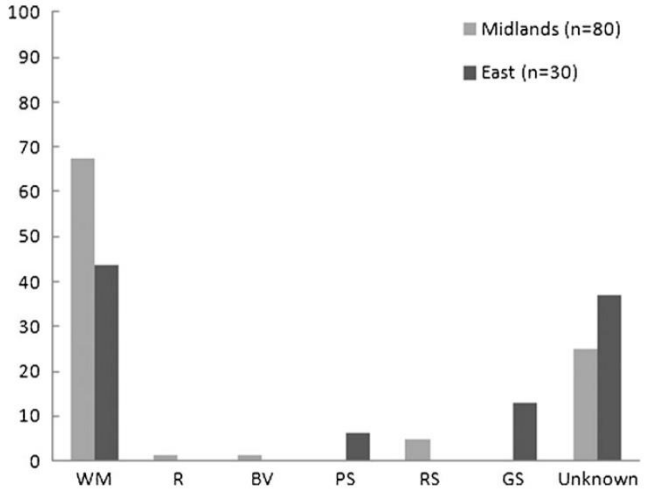

Fig. 6 The percentage relative biomass (%BPI) of mammalian prey items ingested as determined by macro analysis of pine marten scats. *WM* woodmouse, *R* rat, *BV* bank vole, *PS* pygmy shrew, *RS* red squirrel, *GS* grey squirrel

month; in site 1, six pine marten were detected using five hair traps and in site 2, ten pine marten were detected with as few as four hair traps (over the course of the sampling period, hair traps were persistently stolen from this site). It is possible that this resulted in an underestimation of the abundance value at this site as the number of hair traps available was not sufficient to give each individual present the opportunity to use a hair trap each month, in particular during the winter months when residency was being determined.

The proportion of hair samples collected that were successfully genotyped (66 %) could also be improved in future studies. A relationship has been found to exist between the number of hairs in a sample, and the probability it will be successfully genotyped (Mowat and Paetkau 2002). In the majority of hair trapping events in the current study more than ten hairs were captured, thus providing a relatively high amount of DNA. However, the samples were left in situ for a period of up to 1 month, which may have caused the DNA to degrade due to relatively high ambient temperatures and humidity. Lynch et al. (2006) suggested a survey period of 6 days is sufficient to detect pine marten presence (in lowland broadleaf woods). Screening the quantity of nuclear DNA in the samples using the sex typing assay prior to genotyping helped to increase the genotyping success (95 %) as the samples that were deemed to have insufficient high quality DNA did not proceed to the genotyping stage. This also helped reduce the overall cost and this technique combined with shorter sampling periods could substantially help improve the overall success rate in future studies.

The overall number of alleles and levels of heterozygosity in this study were very low. Mullins et al. (2010) also recorded low levels of genetic diversity (average H_E=0.35 and H_O= 0.34) in the Irish pine marten population, using a larger microsatellite panel than the current study. However, Mullins et al. (2010) used samples from a wider geographic range in Ireland than the current study. The low levels of genetic variability in both studies could be due to the low number of individuals that the current pine marten population have re-established themselves from. Furthermore, there has only been one mitochondrial DNA haplotype found in the contemporary Irish population (Jordan et al. 2012). The long-term effects of such low levels of genetic diversity in an expanding population are not known. However, the low diversity found in this study may also be partially due to the microsatellites used, as they were originally developed for use with other mustelids. This was also discussed as a reason for lower levels of genetic variability in the Iberian pine marten population by Ruiz-González et al. (2013).

An alternative form of quality screening to the method used in the current study involves preliminary analysis with a sub-group of microsatellites, as was undertaken by Ruiz-González et al. (2013). Samples that amplified well (>50 % positive PCRs) with the sub-group were then taken to the next stage of analysis. The approach taken in the current study may be more useful as data not used for genotyping at least provides further information on species and sex. The pre selection of DNA samples for genotyping removes samples that are unlikely to replicate or may cause a higher occurrence of genotyping error (Zhan et al. 2010) and thus may be more efficient.

Low genotyping error rates were reported in this study, and were at the lower end of the level of error when compared to other non invasive genetic studies using DNA extracted from hair (Broquet et al. 2007). Genotyping errors are easier to control and account for in small studies with fewer samples (Zhan et al. 2010). Mullins et al. (2010), also working with a relatively small dataset, similarly reported low genotyping errors. The high number of recaptures reported in the current study further supports the low occurrence of genotyping errors, and helps to validate the genotyping results. If an erroneous individual had been detected within the dataset, this individual would not affect the overall abundance estimates, as only animals identified within a site on at least three separate months were included in the abundance estimates.

Scat density

There are inherent problems with surveying for pine marten scats in areas of low population density (Birks et al. 2005); most notably misidentification of scats in the field, even by experienced surveyors. This problem is addressed in modern surveys by the use of genetic tests to confirm pine marten origin (O'Reilly et al. 2008; Balestrieri et al. 2011; Caryl et al. 2012b). The current study found that in the east of Ireland, where pine marten abundance is lower, a significantly lower portion of the scats collected were confirmed as being of pine marten origin than those collected in the midlands, where marten abundance is higher. A factor that may have contributed to this result is the likelihood that the surveyor was less discriminate about which scats were collected in the lower scat density sites. In areas where scats are more abundant, key features such as smell and shape are more easily taken into account, and the surveyor is likely to be more critical regarding the quality of the scat collected for the survey.

Furthermore, in areas of low pine marten population density, territorial scent marking behaviour may be greatly reduced (Macdonald et al. 1998). Lockie (1964) was the first to suggest that a relationship exists between the number of scats and pine marten abundance; however, in a review of nine previous scat surveys in the UK and Spain, Birks et al. (2005) found that the field relationship between scat abundance on transects and marten numbers was yet to be established. Indeed, whilst the current study found scat density to be higher in areas of higher pine marten abundance, regression analysis failed to define this possible relationship.

Scent detection dogs are increasingly being used in the study of elusive carnivores (Smith et al. 2003; Long et al.

2007; Reed et al. 2011), and have been found to have a superior detection rate to that of humans. In the current study, the scent detection dog was used over a 2-day period, and succeeded in collecting a total of 11 scats, seven of which were confirmed as pine marten through molecular analysis. Those that tested negative for pine marten DNA also tested negative for fox DNA (the species that pine marten scat is most likely to be misidentified as in Ireland), which suggests that the quality of the DNA in those samples was too degraded for genetic species identification. As such, it is not possible to determine whether the scats detected by the dog that tested negative for pine marten DNA were true or false negatives. The lead author only detected one pine marten scat during the 2-day survey without the aid of the dog, suggesting that the use of scent detection dogs in areas of low pine marten and low scat density can greatly improve sampling efficiency.

Dietary analysis techniques

In this study, both molecular and macro analyses detected prey species in similar proportions; therefore, molecular techniques can be accepted as a reliable method to detect mammals as prey items in pine marten diet. This is a useful tool in determining the small mammal composition of carnivore diet and also the spread (and possible decline) of both invasive and native mammal species in Ireland. However the macro analysis was significantly more sensitive in the detection of the woodmouse, which was the most frequently consumed mammal in the diet. It is recommended that any study aiming to determine exact frequencies of a species in the diet (as distinct from determining prey species presence or absence) be validated with traditional hard part (macro) analysis. In this study, a standard DNA extraction for both the species and dietary analysis was used (a cost effective strategy). However, to improve the molecular dietary detection of prey DNA, future molecular studies might increase the detection rate by sampling a larger amount of scat, extracting multiple samples from the same scat, or homogenising the scat prior to DNA extraction (see King et al. 2008). The woodmouse was found to occur in 31.8 % of scats tested, which is similar to the frequencies found in Northern Spain and Tuscany (De Marinis and Masetti 1995). Previous studies in Ireland have found the woodmouse to occur at around 13 % frequency (Lynch and McCann 2007) and 14.7 % (O'Meara et al. 2013) in pine marten scats.

Biomass or %BPI values could be better estimated for both macro and molecular analyses if feeding trials were conducted with captive pine marten to determine the appropriate correction factors for (a) the detection rates of the various mammalian prey species DNA after known amounts have been consumed and (b) the relationship between weight of dried remains and fresh weight ingested for red and grey squirrels as distinct from each other and from the 'small mammal' grouping.

Pine marten predation on squirrels

The absence of grey squirrel in the diet of the pine marten in the midlands most likely reflects their lack of availability as a prey item (Carey et al. 2007; unpublished data from Sheehy and Lawton). In Ireland and Scotland it has been speculated that the pine marten population has inhibited the grey squirrel population from spreading, and has even caused the grey squirrel population to crash in areas where they were once established (Carey et al. 2007; Caryl 2008; Paterson and Skipper 2008). No grey squirrel control measures had been carried out in any of the Irish midlands sites surveyed since the 1990s; therefore, human management of the alien squirrel population is not an explanatory factor in their rarity. Habitat is not a factor either, as red squirrel populations are found in the woodland, as until relatively recently were high numbers of grey squirrels. Whether predation was a factor in the retraction of the grey squirrel range historically is not possible to determine in retrospect, but evidence of predation on the alien squirrel species in the east confirms that the pine marten will indeed prey upon the grey squirrel, where it is available.

Molecular and macro analysis produced an overall frequency of occurrence for grey squirrel of 9.4 % and 10 %, respectively, in sites where grey squirrels are known to be present, that increased to 15.6 % when results from molecular and macro analysis were combined. The relative biomass of grey squirrels gave a similar estimate (13 % BPI). These figures are based upon a relatively small sample size however, and must be interpreted with caution as small sample sizes can cause a prey item to be either under or over represented in dietary analysis (Trites and Joy 2005). However, they do confirm that the North American grey squirrel forms part of the European pine marten diet when the two species' ranges overlap. Throughout the course of the current study, the grey squirrel was only confirmed as an available prey item in areas of low scat density, which made scat collection for dietary analysis in these areas very challenging. A larger sample size of scats will allow for a more robust dietary analysis in areas of low density. In contrast, the sample size of scats collected where red squirrels were confirmed as present was adequate to detect with confidence the frequency at which the red squirrel occurs in the diet. Red squirrels and pine marten have co-existed in Ireland and many other parts of Europe over many millennia, and the red squirrel has also appeared only as a very low frequency prey item in previous Irish pine marten dietary studies (Warner and O'Sullivan 1982; Lynch and McCann 2007; O'Meara et al. 2013). However, red squirrels have been recorded at higher frequencies in Russia and Sweden where other small mammal prey are less abundant (De Marinis and Masetti 1995). It is possible that the red squirrel's lower

frequency of occurrence in pine marten diet than that of the alien grey squirrel is a result of differences between red and grey squirrel ecology. Red squirrels live at lower densities (0.3–1.5 per ha) than grey squirrels (2–16 per ha) (Gurnell 1987) and would therefore be numerically less available as prey items. They are also lighter than the grey squirrel, capable of reaching the outermost branches of trees, and spend the vast majority of their foraging time in the canopy, whereas grey squirrels spend a larger proportion of their foraging time on the ground (Kenward and Tonkin 1986). This study introduces the possibility that there could be some form of density dependent effect on the grey squirrel population, where areas of high predator abundance might be discouraging the normally invasive grey squirrel from remaining in, or establishing in a woodland in the first place. The relationship between the native squirrel predator and the alien squirrel species is yet to be defined however, and what effect an increase in pine marten numbers in the east of Ireland will have on grey squirrel distribution and abundance merits further investigation.

Acknowledgements Emma Sheehy is funded by the Irish Research Council and Denise B. O'Meara by the Embark Strand 1 Initiative. Molecular work was supported by a grant from the European Squirrel Initiative. We thank Louise Wilson and Luna from Conservation Dogs and all of the woodland owners and managers especially David Hutton Bury, Donal Whelan and the Irish Forestry Unit Trust, Geoffrey Tottenham, Eamon Doran, National Parks and Wildlife Service (Department of Arts, Heritage and the Gaeltacht) and Michael Carey. Thanks to David O'Neill for guidance in the laboratory, the volunteers who helped in the field throughout and two anonymous reviewers for helpful advice on an earlier version of this manuscript. The authors acknowledge support of the HEA under PRTLI4 for licensing OSI Digital Imagery through the Ryan Institute.

References

Balestrieri A, Remonti L, Ruiz-González A, Vergara M, Capelli E, Gómez-Moliner B, Prigioni C (2011) Food habits of genetically identified pine marten (*Martes martes*) expanding in agricultural lowlands (NW Italy). Acta Theriol 56:199–207

Balharry D (1993) Factors affecting the distribution and population density of pine martens (*Martes martes* L.) in Scotland. PhD thesis, University of Aberdeen, Aberdeen, Scotland

Bijlsma R, Van De Vliet M, Pertoldi C, Van Apeldoorn R, Van de Zande L (2000) Microsatellite primers from the Eurasian badger, *Meles meles*. Mol Ecol 9:2215–2216

Birks J (2002) The pine marten. SP Press, London, The Mammal Society

Birks J, Messenger J, Braithwaite T, Davison A, Brookes R, Strachan C (2005) Are scat surveys a reliable method for assessing distribution and population status of pine martens? Martens and Fishers (*Martes*) in Human-Altered Environments:235–252

Broquet T, Ménard N, Petit E (2007) Noninvasive population genetics: a review of sample source, diet, fragment length and microsatellite motif effects on amplification success and genotyping error rates. Conserv Genet 8:249–260

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20:1004–1006, 1008–1010

Carey M, Hamilton G, Poole A, Lawton C (2007) Irish squirrel survey. COFORD, Dublin

Caryl F (2008) Pine marten diet and habitat use within a managed coniferous forest. PhD dissertation University of Stirling, Scotland

Caryl FM, Quine CP, Park KJ (2012a) Martens in the matrix: the importance of nonforested habitats for forest carnivores in fragmented landscapes. J Mammal 93:464–474. doi:10.1644/11-mamm-a-149.1

Caryl FM, Raynor R, Quine CP, Park KJ (2012b) The seasonal diet of British pine marten determined from genetically identified scats. J Zool 288:252–259

Clevenger AP (1994) Habitat characteristics of Eurasian pine martens *Martes martes* in an insular Mediterranean environment. Ecography 17:257–263

Davis C, Strobeck C (1998) Isolation, variability, and cross-species amplification of polymorphic microsatellite loci in the family Mustelidae. Mol Ecol 7:1776–1778

De Marinis AM, Masetti M (1995) Feeding habits of the pine marten (*Martes martes* Lineaus 1758) in Europe: a review. Hystrix 7:143–150

Deagle BE, Tollit DJ, Jarman SN, Hindell MA, Trites AW, Gales NJ (2005) Molecular scatology as a tool to study diet: analysis of prey DNA in scats from captive Steller sea lions. Mol Ecol 14:1831–1842

Duffy A, Landa A, O'Connell M, Stratton C, Wright J (1998) Four polymorphic microsatellites in wolverine, *Gulo gulo*. Anim Genet 29:63

Dunshea G (2009) DNA-based diet analysis for any predator. PLoS ONE 4:e5252

Fleming MA, Ostrander EA, Cook JA (1999) Microsatellite markers for American mink (*Mustela vison*) and ermine (*Mustela erminea*). Mol Ecol 8:1352–1355

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate *F*-statistics. J Hered 86:485–486

Gurnell J (1987) The natural history of squirrels. Christopher Helm, London

Jedrzejewska B, Jedrzejewski W (1998) Predation in vertebrate communities: the Bialowieza Primeval Forest as a case study, vol 135. Ecological Studies, Springer, Berlin

Jordan N, Messenger J, Turner P, Croose E, Birks J, O'Reilly C (2012) Molecular comparison of historical and contemporary pine marten (*Martes martes*) populations in the British Isles: evidence of differing origins and fates, and implications for conservation management. Conserv Genet 13:1195–1212

Kenward R, Tonkin J (1986) Red and grey squirrels: some behavioural and biometric differences. J Zool 209:279–304

King R, Read D, Traugott M, Symondson W (2008) Molecular analysis of predation: a review of best practice for DNA–based approaches. Mol Ecol 17:947–963

Lanszki J, Zalewski A, Horváth G (2007) Comparison of red fox *Vulpes vulpes* and pine marten *Martes martes* food habits in a deciduous forest in Hungary. Wildl Biol 13:258–271

Lindenmayer DB, Incoll RD, Cunningham RB, Pope ML, Donnelly CF, MacGregor CI, Tribolet C, Triggs BE (1999) Comparison of hairtube types for the detection of mammals. Wildl Res 26:745–753. doi:10.1071/WR99009

Lockie J (1961) The food of the pine marten *Martes martes* in west Ross-shire, Scotland. J Zool 136:187–195

Lockie JD (1964) Distribution and fluctuations of the pine marten, *Martes martes* (L.), in Scotland. J Anim Ecol 33:349–356

Long RA, Donovan TM, Mackay P, Zielinski WJ, Buzas JS (2007) Comparing scat detection dogs, cameras, and hair snares for surveying carnivores. J Wildl Manage 71:2018–2025

Lynch Á, Brown M, Rochford J (2006) Fur snagging as a method of evaluating the presence and abundance of a small carnivore, the pine marten (*Martes martes*). J Zool 270:330–339

Lynch Á, McCann Y (2007) The diet of the pine marten (*Martes martes*) in Killarney National Park. Biology & Environment: Proceedings of the Royal Irish Academy 107:67–76

Macdonald DW, Mace G, Rushton S (1998) Proposals for future monitoring of British mammals. Department of the Environment, Transport and the Regions, London, UK

Manzo E, Bartolommei P, Rowcliffe J, Cozzolino R (2012) Estimation of population density of European pine marten in central Italy using camera trapping. Acta Theriol 57:165–172

Mergey M, Helder R, Roeder J-J (2011) Effect of forest fragmentation on space-use patterns in the European pine marten (*Martes martes*). J Mammal 92:328–335

Messenger JE, Birks JDS (2000) Monitoring the very rare pine marten populations in England and Wales. Management and conservation aspects of small carnivores: human interactions. Backhuys, Leiden, the Netherlands

MetEireann (2012) Temperature in Ireland. http://www.met.ie/climate/temperature.asp?prn=1. Accessed 23 May 2013

Mills L, Citta J, Lair K, Schwartz M, Tallmon D (2000) Estimating animal abundance using noninvasive DNA sampling: promise and pitfalls. Ecol Appl 10:283–294

Mowat G, Paetkau D (2002) Estimating marten *Martes americana* population size using hair capture and genetic tagging. Wildl Biol 8: 201–109

Mullins J, Statham M, Roche T, Turner P, O'Reilly C (2010) Remotely plucked hair genotyping: a reliable and non-invasive method for censusing pine marten (*Martes martes*, L. 1758) populations. Eur J Wildl Res 56:443–453

O'Mahony D, O'Reilly C, Turner P (2006) National pine marten survey of Ireland (2005) Coford Connects. Environ Sci Technol 7:1–8

O'Mahony D, O'Reilly C, Turner P (2012) Pine marten (*Martes martes*) distribution and abundance in Ireland: a cross-jurisdictional analysis using non-invasive genetic survey techniques. Mamm Biol 77:351–357

O'Meara D, Sheehy E, Turner P, O'Mahony D, Harrington A, Denman H, Lawton C, MacPherson J, O'Reilly C (2013) Non-invasive multi-species monitoring: real-time PCR detection of small mammal and squirrel prey DNA in pine marten (*Martes martes*) scats. Acta Theriol:1–7. doi:10.1007/s13364-013-0155-8

O'Meara D, Turner P, Coffey L, O'Reilly C (2012) TaqMan assays for species identification of the red squirrel (*Sciurus vulgaris*) and the grey squirrel (*Sciurus carolinensis*). Conserv Genet Resour 4:603–604

O'Reilly C, Statham M, Mullins J, Turner P, O'Mahony D (2008) Efficient species identification of pine marten (*Martes martes*) and red fox (*Vulpes vulpes*) scats using a 5′ nuclease real-time PCR assay. Conserv Genet 9:735–738

O'Sullivan PJ (1983) The distribution of the pine marten (*Martes martes*) in the Republic of Ireland. Mammal Rev 13:39–44

Palomares F, Godoy JA, Piriz A, O'Brien SJ (2002) Faecal genetic analysis to determine the presence and distribution of elusive carnivores: design and feasibility for the Iberian lynx. Mol Ecol 11:2171–2182. doi:10.1046/j.1365-294X.2002.01608.x

Paterson WD, Skipper G (2008) The diet of pine martens (*Martes martes*) with reference to squirrel predation in Loch lomond and the Trossachs National Park. University of Glasgow, Glasgow, Scotland. BSc dissertation

Peakall ROD, Smouse PE (2006) genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Pereboom V, Mergey M, Villerette N, Helder R, Jean-Francois G, Lod T (2008) Movement patterns, habitat selection, and corridor use of a typical woodland-dweller species, the European pine marten (*Martes martes*), in fragmented landscape. Can J Zool 86: 983–991

Reed SE, Bidlack AL, Hurt A, Getz WM (2011) Detection distance and environmental factors in conservation detection dog surveys. J Wildl Manage 75:243–251

Rousset F (2009) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour 8: 103–106

Ruiz-González A, Madeira MJ, Randi E, Urra F, Gómez-Moliner BJ (2013) Non-invasive genetic sampling of sympatric marten species (*Martes martes* and *Martes foina*): assessing species and individual identification success rates on faecal DNA genotyping. Eur J Wildl Res 59:371–386

Scotts D, Craig S (1988) Improved hair-sampling tube for the detection of rare mammals. Wildl Res 15:469–472

Shehzad W, Riaz T, Nawaz MA, Miquel C, Poillot C, Shah SA, Pompanon F, Coissac E, Taberlet P (2012) Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. Mol Ecol 21: 1951–1965

Smith DA, Ralls K, Hurt A, Adams B, Parker M, Davenport B, Smith MC, Maldonado JE (2003) Detection and accuracy rates of dogs trained to find scats of San Joaquin kit foxes (*Vulpes macrotis mutica*). Anim Conserv 6:339–346

Teerink BJ (1991) Hair of west-European mammals, 1st edn. Cambridge University Press, Cambridge

Trites AW, Joy R (2005) Dietary analysis from fecal samples: how many scats are enough? J Mammal 86:704–712

Valière N (2002) GIMLET: a computer program for analysing genetic individual identification data. Mol Ecol Notes 2:377–379

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

Vincent I, Farid A, Otieno C (2003) Variability of thirteen microsatellite markers in American mink (*Mustela vison*). Can J Anim Sci 83: 597–599

Warner P, O'Sullivan P (1982) The food of the pine marten *Martes martes* in Co. Clare Transactions of the International Union of Game Biology 14:323–330

Waters C, Lawton C (2011) Red squirrel translocation in Ireland. Irish Wildlife Manuals, vol 51. National Parks and Wildlife Service, Department of the Environment, Heritage and Local Government, Dublin, Ireland

Williams BW, Etter DR, Linden DW, Millenbah KF, Winterstein SR, Scribner KT (2009) Noninvasive hair sampling and genetic tagging of co-distributed fishers and American martens. J Wildl Manage 73:26–34

Yalden DW, Morris PA (1990) Analysis of owl pellets. The Mammal Society, London

Zalewski A (1999) Identifying sex and individuals of pine marten using snow track measurements. Wildl Soc Bull 27:28–31

Zalewski A, Jedrzejewski W (2006) Spatial organisation and dynamics of the pine marten *Martes martes* population in Bialowieza Forest (E Poland) compared with other European woodlands. Ecography 29: 31–43

Zhan X, Zheng X, Bruford MW, Wei F, Tao Y (2010) A new method for quantifying genotyping errors for noninvasive genetic studies. Conserv Genet 11:1567–1571