Investigating the Population Structure of Daubenton's Bat in Europe: A Multiple Marker Approach

Katie Elizabeth Frith

Submitted in accordance with the requirements for the degree of PhD

The University of Sheffield
Department of Animal & Plant Sciences

July 2010

Acknowledgements

Firstly I would like express a big thank you to my supervisors Roger Butlin and John Altringham. I started this PhD with a limited experience of molecular ecology and Roger has been endlessly patient in explaining things to me and has helped to inspire in me a genuine interest in genetics. Technical aspects of this project have not always run smoothly, but Roger has provided support and encouragement throughout. I appreciate that his door has always been open whenever I have needed advice. John gave me the opportunity to work on this project, which is something I am extremely grateful for. But more than this I am grateful for the chance John gave me to work with bats. The fieldwork I experienced was without doubt the highlight of this project and I feel privileged to have been able to do this. Finally thank you to both of you for your suggestions to improve the thesis, especially at such short notice!

It has been a pleasure to work in TAB lab over the past few years. There are many people I wish to thank who, despite having big workloads of their own have helped me with patience and enthusiasm. In particular Andy Krupa, Gavin Horsburgh and Maria Mannarelli for help with lab-work, and Alain Frantz who provided invaluable advice on statistical analyses.

This project would not have been possible were it not for the years of fieldwork conducted by members of the Leeds bat team past and present. Thank you to everyone involved and to others who provided me with samples for this project: FERA, Veronika Laine, Stuart Piertney and Liselotte Andersen.

This project was funded by the Food & Environment Research Agency (FERA) and I would like to thank my supervisors at FERA, Alan MacNicoll and Helen Atterby for obtaining the funding that made this project possible.

Camille Jan helped me out massively during this project allowing me to use her primers and genotype data for my own analyses. I cannot express how indebted I am to her for this. She also proved an entertaining companion during our fieldwork trip to France, which is an experience I will never forget – especially that wedding!

The people I met at Sheffield made my time here particularly enjoyable and I feel lucky to have made some wonderful friends here. In particular I would like to mention original TAB team members Susan Johnston and Ian Cleasby with whom I have shared highs, lows, laughter and beetles.

I am grateful for the love and support I have received from my family during my PhD. Mum, Dad, Lucy , Marcus and my Grandparents; I realise that I have not always been easy to deal with but you ignored my pessimism and gave me the encouragement I needed to finish this thesis. Thank you.

I feel that I should mention the many hamsters past and present who have accompanied me through my PhD: Gladys, Pamplemousse, Cynthia, Moonpig, Nice, Jemba and especially Noon.

Finally I would like to mention David Green (Bean). You arrived halfway through this project and I cannot help but think that it is no coincidence that I started getting positive results at this time! You have always been there to listen to me ramble on about my thesis not just with kindness but with genuine enthusiasm – I cannot tell you how helpful this was. I hope that I can be as encouraging in all you do as you have been for me. Thank you Bean, this thesis is for you.

Abstract

Zoonotic diseases are of increasing concern on a global scale. In order to understand how such diseases spread it is vital to elucidate dispersal patterns of their hosts. Daubenton's bat (*Myotis daubentonii*) is one of the major reservoirs of European bat lyssavirus type-II which is transmissible to humans as rabies.

The spatial pattern of variation in X-linked and autosomal microsatellites plus mitochondrial sequence was investigated across *M. daubentonii* populations in Britain and Europe. Using a multi-marker approach allowed the exploration of phylogeographic processes, contemporary gene flow, barriers to gene flow, patterns of sex-biased differentiation and an insight into variable mating success between males and females.

Two mitochondrial clades were found in Europe that appeared to be consistent with the expansion of *M. daubentonii* from at least two glacial refugia. This was corroborated by similar broad-scale patterns of variation using microsatellites.

Differentiation was low for microsatellites over wide geographic areas. This implies that levels of geneflow among populations are high, but a pattern of isolation by distance suggests that *M. daubentonii* populations are not panmictic at a Europe-wide scale. Contrasting high levels of mitochondrial differentiation indicate that female dispersal is more restricted than males, probably as a consequence of female philopatry and male dispersal to swarming sites.

No significant features acting as barriers to dispersal or colonisation were detected, including the English Channel.

The ratio of X:autosomal effective population size was 0.84; greater than neutral expectations would predict. Although a number of factors could explain this, higher variability in mating success resulting in a reduced male effective population size seems the most likely explanation.

These results suggest that the potential for EBLV-2 to be transferred among Daubenton's colonies is high and that, as dispersal is male-biased, males may be more instrumental in EBLV-2 spread.

Table of Contents

Chapter 1: General introduction

Chapter 2: Methodological attempts to characterise Y-specific markers in Daubenton's bat

Chapter 3: Investigating the phylogeography of Daubenton's bat across Europe using the mitochondrial *ND1* gene

Chapter 4: The development and application of X-linked microsatellite markers to investigate population structure of Daubenton's bat in Europe

Chapter 5: Investigating the population structure of Daubenton's bat across Europe using a combination of autosomal microsatellites, X-linked microsatellites and mtDNA: the influence of sex-biased forces

Chapter 6: Final discussion

Appendix

List of Tables

List of Figures

Chapter 1: General Introduction

1.1 Overview

The spread of infectious diseases to humans facilitated by zoonotic pathogens is of increasing concern following the outbreak of several high profile epidemics on a global scale, notably H5N1 (bird flu) and the recently declared H1N1 swine flu pandemic (Neumann, Noda & Kawaoka 2009). In many of these cases wildlife serves as the primary reservoir, carrying pathogens that are transmissible to both humans and domestic livestock (Kruse, Kirekemo & Handeland 2004). An understanding of the spatial distribution of the principal reservoir is vital to predict the incidence of spill over infection to the human population and to mitigate for such risks (Ostfeld *et al.* 2005). As the frequency and level of contact between the host species and recipients will influence disease dynamics, it is imperative that dispersal rates and patterns for the host species are elucidated to allow predictions of future disease spread (Deal *et al.* 2000, Jin & Wang 2005).

Sex-biased dispersal is prevalent in many host organisms, with the dispersing sex likely to have a greater impact on disease transmission, both to other species and to conspecifics (Long *et al.* 2005). Therefore, it is important to consider this bias in epidemiological studies, as grouping both sexes together may result in an underestimate of disease spread if longer distance dispersal events by one sex are not included. It is notable that in the majority of mammal species females are philopatric, whereas males disperse away from the natal range (Greenwood 1980). The evolutionary causes for this disparity are debated, but for a review see Lawson-Handley & Perrin (2007). Consequently it is most often the case that, in mammals, males will be instrumental in spreading a zoonosis from one population to another, especially over a large geographic scale.

Traditionally sex-biased dispersal was measured using direct methods such as ringing or observation, however in practice this is technically difficult, especially so when the focal organism is highly vagile (Koenig *et al.* 1996). The past decade has seen an increase in the use of molecular markers to investigate population structure and dispersal. Whilst this has largely focussed on comparing nuclear markers between the sexes or comparing nuclear markers (Goudet *et al.* 2002, Vitalis 2002) and maternally inherited mitochondrial DNA (Broquet & Petit 2009, Prugnolle & de Meuus 2002),

there is increasing use of paternally inherited Y-chromosome markers (Greminger *et al.* 2010, Jobling & Tyler-Smith 2003) and X-chromosome markers (Schaffner 2004). Combining markers with different modes of inheritance and mutational properties affords us a better understanding of the role of sex-biased demographic processes in determining population structure than ever before.

Myotis daubentonii (Daubenton's bat) is known to be one of two host species for the rabies-like European bat lyssavirus type 2 (*Myotis dasycneme* being the other host) (Racey & Fooks 2005). Little is known about the epidemiology of EBLV-2 but it has been recorded in a number of European countries (UK, the Netherlands, Denmark, Germany, Switzerland) (Harris *et al.* 2006b) and there have been two confirmed instances of infection from *M. daubentonii* in humans (one in Scotland, the other in Finland), both of which resulted in death (Brooks *et al.* 2005, Fooks *et al.* 2003, Lumio *et al.* 1986). Passive and active surveillance in the UK suggests that when there is no *a priori* knowledge of lyssavirus infection the proportion of *M. daubentonii* individuals carrying antibodies for EBLV-2 in England is 1-5% (Harris *et al.* 2006a, 2009). When there is prior knowledge of EBLV-2 occurrence this figure rises to as high as 16% (Harris *et al.* 2009). Currently, we have a limited understanding of dispersal dynamics in *M. daubentonii*, although past work has suggested a male bias in dispersal (Angell 2008, Ngamprasertwong *et al.* 2008) and it appears that individuals do have the capacity to migrate tens to hundreds of kilometres (Hutterer *et al.* 2005, Urbancyzk 1990). The aim of this project was to examine population structure in *M. daubentonii*, with a view to elucidating past and present forces responsible for observed genetic structure. The expectation was that determining current rates of gene flow (effective dispersal) in both sexes and identifying past and present barriers to gene flow would afford us a better understanding of how EBLV-2 may spread between *M. daubentonii* colonies. This will make an important contribution to our knowledge of EBLV-2 epidemiology in the UK.

1.2 Bats

Bats make up the second most diverse mammalian taxon after the rodents, with over 1100 species currently recorded (Altringham 1996). The taxon has a widespread distribution and species are found occupying a diverse range of habitats in every continent except for Antarctica (Altringham 1996). Bats are both ecologically and economically important as they play a vital role in pollination (Knudsen & Tollsten

2008) and are good indicator species (Wickramasinghe *et al.* 2004), but they are also notable for their role as vectors of disease (Calisher *et al.* 2006). All bats are characterised by their ability to undertake powered flight which has a strong influence on their ecology and life history patterns in a manner which is very dissimilar to similarly sized ground dwelling mammals (Holmes & Austad 1994, Jones & Maclarnon 2001). In fact, for their size bats have a life history strategy which is more congruent with that of larger mammals (Wilkinson & South 2002). For example they are long lived, have low fecundity, take longer to reach maturity and have the ability to travel large geographic distances. They also display a diverse array of social systems, mating patterns and dispersal patterns (Altringham 1996, Burland & Worthington-Wilmer 2001, McCracken & Wilkinson 2000).

Traditionally the major taxonomic division within bats was that of the two suborders micro (echolocating bats) and megachiropterans (non-echolocating bats) (Simmons & Geisler 1998), with a monophyletic origin from an ancestor capable of powered flight. However advancing molecular techniques have allowed more sophisticated phylogenetic trees to be reconstructed and although the monophyletic origin of extant bat lineages is still supported, the grouping of the two original suborders is not. It is now widely accepted that the megachiropetra (Pteropodidae) form a sister taxon with the Rhinolophoidea that is termed the Yinpterochiroptera and the remaining bat lineages form a suborder termed the Yangochiroptera (Teeling *et al.* 2005). It also appears that echolocation arose independently in these two lineages, with the Pteropodidae never aquiring the ability to echolocate (Jones & Teeling 2006, Teeling 2009).

Of the nineteen accepted bat families five are recorded in Europe (Vespertilionidae, Rhinolophidae, Molossidae, Emballonuridae, Pteropodidae) representing over 45 species. In the UK, bats comprise over a quarter of species in the mammalian fauna with 17 species currently recorded and new cryptic species continually being added to the list as a consequence of increasingly sophisticated genetic and acoustic methods of species identification (Jan *et al.* 2010 in press, Mayer *et al.* 2007).

1.2.1 *Myotis daubentonii*

Myotis daubentonii (Daubenton's bat) (Figure 1.1) is a temperate species of insectivorous bat which has a widespread distribution from Ireland in the west to Japan in the east and Finland in the north to Turkey and Spain in the south (Figure 1.2).

Figure 1.1 *Myotis daubentonii* in hand

Although data are sparse, current records suggest that the population is large with approximately 150,000 individuals in Britain (Eurobats 2004) and it has even been suggested that in many parts of the range the population is expanding (Bogdanowicz 1994). Furthermore, genetic data on levels of genetic diversity and heterozygosity suggest that in Britain *M. daubentonii* has the characteristics of a large, outbreeding population (Angell 2008). The maximum age recorded in a Daubenton's individual is 22 years, although it is thought that a more realistic lifespan in the wild is 4-5 years (Schober & Grimmberger 1987).

Figure 1.2 Global distribution of *M. daubentonii*. Presence indicated by red coloration.

Remains of *M. daubentonii* have been found in lower and upper Pleistocene deposits in Eastern Europe (Bogdanowicz 1994) and Holocene deposits across Europe and the Far East (Ochman 2003) suggesting that it is has persisted in refugia in Europe during the last glacial period, with expansion during the current interglacial. A Europe-wide phylogenetic study placed *M. daubentonii* as being most closely related to *Myotis bechsteinii*, and within *M. daubentonii* there was evidence of four distinct mitochondrial clades which were thought to represent origins from different glacial refugia (Mayer & von Helversen 2001).

M. daubentonii is a trawling species and in terms of wing morphology has a medium to high aspect ratio ($span^2$/area) and a low wing loading (area of wings in relation to the weight of the bat), which allows efficient flight over open water and manoeuvrability around bank-side vegetation (Altringham 1996). *M. daubentonii* spends a large amount of foraging time over water where the main prey type is Diptera and Trichoptera (Flavin *et al.* 2001, Vaughan 1997). Consequently its distribution is tightly linked with riparian habitats, with habitat use studies suggesting that the smoothness of the water surface may also play a role in the fine scale structuring of individual distribution (Warren *et al.* 2000). As with many temperate bat species *M. daubentonii* undergoes a seasonal migration. In the summer, individuals are found in small roosts at lower elevations, where there is a roughly equal sex ratio and evidence for spatial segregation between the sexes (Angell 2008, Senior *et al.* 2005). In the autumn a migration event to swarming sites occurs where the bats mate prior to hibernation and the majority of recorded individuals are males (Kerth *et al.* 2003, Parsons & Jones 2003). Genetic and ringing studies have shown that bats of both sexes display a degree of philopatry to their natal area and that bats are also faithful to swarming sites. Swarming sites have a large catchment area (Angell 2008, Parsons & Jones 2003) and contain bats from many summer roosts, but bats from the same summer roosts do not necessarily go to the same swarming site (Angell 2008, Rivers *et al.* 2005, 2006).

In Europe *Myotis daubentonii* and *Myotis dasycneme* are notable as the vectors of the rabies related EBLV-2 virus (Harris *et al.* 2006b). However in Britain *Myotis dasycneme* is absent and *M. daubentonii* is the only known reservoir of EBLV-2.

1.3 Bats and disease

Wild animal populations may serve as reservoirs of zoonotic diseases, which can spill over to infect domesticated livestock or the human population (Daszak *et al.* 2000, Kruse *et al.* 2004). In many cases the diseases do not pose a severe threat to human health, but there have been notable instances of recent pandemics such as bird flu (Li *et al.* 2004) and swine flu (Neumann *et al.* 2009) which have the potential to cause widespread mortality in the human population. Bats are prominent vectors of disease transmission across the world (Calisher *et al.* 2006). Over 60 different viruses, from 10 virus families have been isolated from bats so far (Harris *et al.* 2006b). Perhaps the most notorious virus family carried by bats are the lyssaviruses. Lyssaviruses can be transmitted from bats to other mammals, including humans where they present themselves as rabies (Fooks *et al.* 2003, Warrell & Warrell 2004). Rabies is a fatal disease of the central nervous system which is of global importance, with approximately 55,000 cases reported worldwide on an annual basis, although only a few such cases are attributable to bats (World Health Organisation, 2008). Of all reported instances of rabies infection in humans only a handful of individuals who did not receive the post-infection vaccine have survived (Messenger *et al.* 2003). Those who have survived do so with chronic neurological impairments. Consequently understanding how lyssaviruses may spread between bat populations is of vital importance as this will inform our currently limited understanding of rabies epidemiology, which will have implications when it comes to the management of the disease.

European bat lyssaviruses (EBLVs) are related to the rabies virus and were first recorded in Europe in 1954 (Amengual 1997). There are two genotypes of EBLV in Europe: EBLV-1 has been isolated from a handful of different bat species, but the main carrier is thought to be *Eptesicus serotinus* (Freuling *et al.* 2009). EBLV-1 has been recorded in a number of different European countries (Denmark, Germany, the Netherlands, Poland, Russia, Slovakia, France, Spain) but thus far appears to be absent from Britain (Harris *et al.* 2006b). EBLV-2 has only been isolated from *Myotis daubentonii* and *M. dasycneme* (Van der Poel *et al.* 2005). It appears to have a less widespread geographical distribution than EBLV-1 (UK, Germany, Switzerland, Netherlands, Finland) but it is present in Britain (Harris *et al.* 2006b), where *M. daubentonii* is the only known reservoir, due to the absence of *M. dasycneme*.

Currently the prevalence of EBLV infection in European bat populations is unknown, as in most countries surveillance for EBLVs in bats is passive (i.e. testing dead bats for presence of the virus) (Harris *et al.* 2009). In addition, in a number of European countries rabies is still endemic in larger mammals such as the fox (Holmala & Kauhala 2006), which causes greater economic concern than EBLVs in bats. Surveillance has been more extensive in the UK where passive testing of dead bats has been conducted by the Veterinary Laboratories Agency (VLA) since 1987 (Harris *et al.* 2006a). Following the finding of a bat positive for EBLV-2 in 1996 a program of active surveillance in Britain has been carried out in conjunction with governmental organisations, NGOs and Universities (Harris *et al.* 2009). Active surveillance tests are for both previous exposure to the virus (in terms of antibodies) and for current infection (viral excretion in saliva). Results from active surveillance in the UK as a whole has suggested that where there is an *a priori* expectation of infection in bat populations antibody presence is 3-16% (Harris *et al.* 2009), and where there is no reason to expect the virus it is 1-5% (Harris *et al.* 2006a). Active surveillance results from Scotland only found antibodies in the blood of 32.9% of bats sampled, suggesting that previous exposure to the virus was high (Scottish Natural Heritage 2005).

Currently our understanding of how EBLVs are transmitted between bats and from bats to other organisms is limited (Freuling *et al.* 2009). Spill over infection of EBLV-1 to sheep, a pine marten, a domestic cat and one confirmed human case have been reported (Harris *et al.* 2006b). There have been two confirmed instances of EBLV-2 infection in humans but no recorded cases of spill over to other animals (Harris *et al.* 2006b). In humans EBLV infection thus far has been through a bite but it may also occur through contact with mucosal membranes, though there is less evidence for this (Johnson *et al.* 2008). Although it is confirmed that EBLV is secreted in saliva (Kuzmin *et al.* 2008) and thus the obvious infection route is via a bite (Fooks *et al.* 2003), laboratory experiments where *M. daubentonii* individuals were artificially infected with EBLV-2 via various routes indicated that *M. daubentonii* was surprisingly resilient to infection (Johnson *et al.* 2008). In fact Johnson (2008) *et al.* were only able to infect bats via direct intracranial inoculation. This suggests that bats may be able to tolerate EBLV, resulting in a silent infection, but that virus reactivation could occur due to specific stress factors (O'Shea *et al.* 2003). If bats are able to carry the infection asymptomatically, it is of concern because those coming into contact with an infected bat may be unaware of the infection risk. However it has also been shown that bats with

the active infection may behave strangely, becoming more aggressive and less co-ordinated than healthy animals (Fooks *et al.* 2003, Johnson *et al.* 2002).

Current management options of EBLVs in Europe largely consist of public awareness measures and habitat modifications, to exclude potentially infected bats from buildings where they may come into contact with people (Frantz & Trimarchi 1983). Because bats are highly mobile, containing an infection to a local area is impractical and because bats are protected species eradication of individuals is also not an option (Agreement on the Conservation of Bats in Europe (1992)). The major groups at risk are those who work directly with bats such as researchers and environmental consultants. Vaccination is a requirement for such people. Due to the low reporting of spill over instances the risk of infection to the general public is thought to be minimal, but the serious consequences of lyssavirus infection mean it must still remain a concern.

Elucidating the dispersal ecology of the vector species is of vital importance in improving our understanding of disease epidemiology and how EBLV-2 may spread. For example migration events in birds have been demonstrated as accounting for the spread of diseases into non-endemic areas such as Lyme disease (Smith *et al.* 1996) and West Nile Virus in the United States (Reed *et al.* 2003). As bats have high dispersal capabilities, doing this via direct observational methods is technically difficult, however genetic methods provide an alternative to investigate how dispersal patterns may have shaped population structure. Thus we can determine the magnitude of dispersal in *M. daubentonii*, whether this dispersal has a sex-bias and identify any barriers (geographic or otherwise) which may impede movement in *M. daubentonii*. All aforementioned factors will affect the spread on EBLV-2 between *M. daubentonii* colonies.

1.4 Dispersal

1.4.1 Background

Dispersal is a well documented phenomenon in natural populations which will have an impact on many important ecological and evolutionary factors including persistence of a species, community dynamics and population genetic structure to name a few. The mechanisms dictating dispersal pressures are debated and dispersal likely represents a trade off between the benefits/costs associated with dispersing to a new area and the

benefits/costs associated with philopatry (Bowler & Benton 2004). Costs associated with dispersal may include increased mortality whilst crossing unsuitable habitat and unfamiliarity with new areas (Van Vuren & Armitage 1994, Zollner & Lima 2001). Costs associated with philopatry include increased kin competition (Perrin & Lehman 2001), local resource competition and inbreeding (Perrin & Mazalov 2000). On the other hand, beneficial factors that favour dispersal include acquiring new resources, avoiding competition and inbreeding avoidance (Perrin & Mazalov 2000). The benefits associated with philopatry include familiarity with the natal area and kin co-operation. Thus the main evolutionary (or ultimate) mechanisms invoked to explain dispersal are kin competition (whether for resources or mates), inbreeding avoidance and environmental stochasticity which affects the distribution of resources in space and time and may make dispersal beneficial when resources are depleted in the natal area (Lawson-Handley & Perrin 2007).

The aforementioned forces can have differential influences depending on the sex under consideration and consequently the evolutionary stable balance between these factors often varies between the sexes (Lawson-Handley & Perrin 2007). When this occurs a sex-bias in dispersal may present itself (Greenwood 1980). Sex-biased dispersal (SBD) is supported by both theoretical and empirical work (Johnson & Gaines 1990) and in many cases is thought to be tightly linked to the mating system under consideration (Greenwood 1980). This is because the mating system will place differential pressures on the sexes in terms of levels of inbreeding and resource requirement. Mammalian mating systems can be broadly classified into four categories (Clutton-Brock & Harvey 1978, Clutton-Brock 1989): monogamy (successive matings with the same partner), promiscuity (both sexes mate with multiple partners, with no subsequent associations), polygyny (a male repeatedly mates with the same group of females) and polyandry (a female mates with the same group of males). For example, Greenwood (1980) conducted a review on dispersal strategies in birds and mammals and found that in birds (where many species are monogamous) females were more likely to disperse, whereas in mammals (where many species are polygynous) the reverse was true.

1.4.2 Evolutionary (Ultimate) causes of sex-biased dispersal

There are three main ultimate causes thought to promote one sex to disperse more than the other: Inbreeding avoidance, local mate/resource competition and kin interactions.

Although there is no general agreement as to which is overriding and it is likely that they will not be mutually exclusive. An excellent discussion of these forces can be found in Lawson-Handley and Perrin (2007) and will be briefly discussed here.

Inbreeding has a negative impact on individual and population fitness (Charlesworth & Charlesworth 1987), whether through decreased heterozygosity which can reduce adaptive potential (Day *et al.* 2007) or inbreeding depression which has more direct fitness consequences for the viability of offspring (Charlesworth & Charlesworth 1987, Rossiter 2001). Consequently if one sex disperses upon reaching maturity inbreeding is avoided for both sexes. If inbreeding was the sole force governing a sex bias in dispersal it would be expected that one sex would remain 100% philopatric but rarely is this true (there are exceptions including *Myotis bechsteinii* (Kerth *et al.* 2002)). In addition, if inbreeding was the only force driving SBD, it would be expected that in a polygynous system females would disperse because they stand to incur the highest inbreeding load, when often the reverse is true and males disperse (Waser *et al.* 1986). In fact empirical studies have shown that surprisingly high levels of inbreeding are tolerated in natural populations (Guillame & Perrin 2006). Thus, although avoidance of inbreeding is likely to play a role in mediating SBD, it probably acts in conjunction with other driving factors (Gandon 1999).

Resource competition has a powerful role to play in driving dispersal in one sex (Clarke 1978, Greenwood 1980). Resources can include territories, nesting sites, food and mates (Dobson 1982) and depending on the mating system will be limiting to the two sexes in asymmetric ways. In a polygynous or promiscuous mating system where there is variability in male mating success and females contribute the majority of juvenile care, availability of females (local mate competition) will be the limiting resource for males and nest sites/energy sources (local resource competition) will be the limiting resource for females (Perrin & Mazalov 2000, Pusey 1987), due to the higher energetic demands of raising young. Therefore it would be expected that males may disperse to find new mates, whereas it would favour females to stay in familiar/higher quality territory if resources are abundant.

Kin-cooperation is a further ultimate factor expected to dictate a sex-bias in dispersal when one sex relies on the building of social bonds but the other does not. For example co-operation among males to find new mates/territories will favour males to disperse

(Watson *et al.* 1994) and conversely co-operation amongst females to rear young will favour female philopatry (Lambin 1994, Packer & Pusey 1987). In addition, kin recognition could promote philopatry in both sexes as it would allow individuals to avoid mating events with close kin (Lehmann & Perrin 2002).

Ultimate factors will also be linked to the sociality of the sexes, with complex social interactions having the potential to increase the magnitude of a sex bias in dispersal (Greenwood 1980, Perrin & Goudet 2001). Furthermore the longevity of a species has been demonstrated to have a positive interaction with dispersal distance in the non-philopatric sex (Pusey 1987).

1.4.3 Proximate causes of sex-biased dispersal

There are likely to be many proximate explanations to invoke SBD and they will not be discussed in detail here but the most important factors are thought to be eviction, attraction to other groups and group dispersal (Lawson-Handley & Perrin 2007). Investigating the proximate (more immediate) factors that govern SBD will enhance our understanding of ultimate mechanisms. For example, if eviction is mediated by members of the same sex then it is likely that local resource competition is driving SBD, whereas if eviction is intersex then inbreeding avoidance is more likely.

1.4.4 Sex-biased dispersal patterns

As previously mentioned, whilst sex-biased dispersal is a common phenomena in many bird and mammal species rarely is it strict, with most species lying on a continuum. For example some members of the philopatric sex may disperse a short distance away from the natal territory even if the magnitude of this dispersal is not great and some members of the dispersing sex may remain in the natal area rather than dispersing to a new territory (Winters & Waser 2003). When both sexes disperse one sex may disperse farther than the other as has been reported in a number of mammalian species (Blundell *et al.* 2002, Fontanillas *et al.* 2004, Ji *et al.* 2001). Indeed it is important to make distinctions with regards to dispersal distance as short dispersal distances may relate to inbreeding avoidance whereas longer dispersal distances may occur when an individual is colonising a new territory (Perrin & Goudet 2001, Ronce *et al.* 2001).

It is also important to consider when in an organism's life dispersal occurs. For example, dispersal is usually discussed in the light of movement of juveniles. But breeding dispersal may also occur where individuals move to new breeding grounds or undergo seasonal migrations to areas in which they breed. This will also have important consequences on the ecology of a species, especially if gene flow results (Dobson 1982).

1.4.5 Philopatry and sex-biased dispersal in bats

Within bats there is a diverse array of social organisation, mating patterns and consequently patterns of sex-biased philopatry/dispersal (Altringham 1996, McCracken & Wilkinson 2000). The majority of bat species are classed as polygynous, but promiscuity is also commonplace and there are also recorded instances of monogamy (McCracken & Wilkinson 2000). In many cases it is better to think of bat mating systems as lying somewhere on a continuum between these extremes. However when the mating system leads to a situation where reproductive success is more variable in males than females (as is the case under polygyny and often under promiscuity) this is likely to lead to female philopatry and a male bias in dispersal because males will be competing for mates (Lawson-Handley & Perrin 2007). Further, because pregnancy places large energetic pressures on female bats (Encarnação & Dietz 2006), they may benefit by remaining faithful to natal summer roosts where there is high prey availability (Angell 2008). Direct and indirect (genetic) methods to study sex-biased dispersal in bats have shown that in many instances when mating systems are polygynous/promiscuous dispersal is weighted towards males, for example *Myotis bechsteinii* (Kerth *et al.* 2000, 2002), *Myotis myotis* (Castella *et al.* 2001), *Nyctalus noctula* (Petit & Mayer 1999), *Rhinolophus monoceros* (Chen *et al.* 2008). However there are exceptions which include philopatry or dispersal in both sexes and this is seen in *Myotis natterreri* (Rivers *et al.* 2006), *Plecotus auritus* (Burland *et al.* 1999) and *Miniopterus schreibersii* (Pereira *et al.* 2009), or when it is females which disperse as in *Saccopteryx bilineata* (Nagy *et al.* 2007).

1.4.6 Philopatry, dispersal and mating systems in *Myotis daubentonii*

Past work based upon using ringing, observation and genetics strongly suggests that in *M. daubentonii* both sexes display a degree of philopatry, although there is a bias

towards male dispersal (Angell 2008, Ngamprasertwong *et al.* 2008, Senior *et al.* 2005). Observation and parentage analysis also suggests that the mating system of *M. daubentonii* is promiscuous, although males roosting with females during the summer stand a higher chance of fathering offspring than males in single sex summer roosts (Senior *et al.* 2005). However, a large portion of paternity could not be attributed to resident males, suggesting that mating events at swarming sites play a vital role in gene flow in this species (Angell 2008). Male mating success at swarming sites is also likely to be variable, as observations of a highly skewed male sex ratio and chasing behaviour (Glover & Altringham 2008, Senior *et al.* 2005) suggest that mating at swarming sites occurs in a lek-like system, with males competing for access to females.

As with a number of temperate bat species, *M. daubentonii* form summer roosts where young are born and reared and then undergo a seasonal migration to swarming/hibernation sites. Studies of *M. daubentonii* populations in the Yorkshire Dales have demonstrated that in the summer there is substantial segregation by sex with most roosts predominantly single sex (discounting juveniles), although in some larger roosts a few males reside with females (Senior *et al.* 2005). It is thought that this segregation is largely to do with the energetic requirements of females during pregnancy and lactation (Barclay 1991). The all female roosts are generally found at lower elevations where insects are more abundant. Contrastingly, the all male roosts are found at higher elevations where insects are found in lower numbers, or more sporadically, and where it has been found that males have to forage for longer to meet their energetic demands (Senior *et al.* 2005).

Currently little is known about long distance dispersal events in *M. daubentonii* but maximum recorded distances on the continent have measured up to 260 km (Urbancyzk 1990). Further the wing morphology of *M. daubentonii* suggests that individuals of this species are capable of dispersing long distances (Altringham 1996). Radio tracking studies of *M. daubentonii* in the UK at swarming sites have shown that swarming sites have a large catchment area which can measure between 254 km^2-4118 km^2 (Parsons & Jones 2003). Ringing and genetic studies show that bats are faithful to swarming sites (Glover and Altringham 2008) and that swarming sites contain bats from many different summer roosts but bats from the same summer roosts do not necessarily go to the same swarming site (Angell 2008, Rivers *et al.* 2005). Thus there is the potential for gene flow and outbreeding over a wide geographic scale.

1.5 Measuring dispersal

Direct (or traditional) methods used to investigate dispersal require observation or capture of the organisms involved. Observational methods only really have the power to classify dispersal (i.e. does it occur and is it more prevalent in one sex than the other) and, whilst they may be useful for relatively sessile organisms, they are rendered extremely inaccurate for species with medium-high levels of vagility (Koenig *et al.* 1996, Nathan *et al.* 2003). If an organism can be caught and then monitored in some way this allows more accurate determination of dispersal magnitude. For example, traditional mark-recapture studies which mark an organism at a particular point of capture may be able to determine if dispersal is occurring by virtue of recaptures (Turchin & Theony 1993). However this relies on being able to recapture an animal, which is not always practical. Radio tracking provides a more accurate technique to follow the movement of individual organisms. However, it is expensive and can be time consuming (Parsons & Jones 2003). Further, although it may be useful to track individual animals (perhaps more relevant when the population is small and individuals have large home ranges, as with big carnivores (Schmidt 1998)), it cannot provide us with a quantifiable estimate of dispersal in the whole population. All direct methods are extremely limited when an organism is highly vagile, small and is found in large numbers (Wilson & Delahay 2001). This is especially pertinent for bats which meet all aforementioned criteria and have the additional caveat of nocturnal activity. In addition, although direct methods may be able to follow dispersal over an individual life-time, they are unable to measure levels of gene flow.

Indirect methods of dispersal measurement largely refer to genetic methods (Slatkin 1987) (though note that stable isotopes have an increasing role to play (Rubenstein & Hobson 2004)) and can be further split into direct and indirect genetic methods, depending on whether they measure dispersal in one generation (instantaneous dispersal) or whether they address the consequences of dispersal on observed patterns of genetic variation in a manner which may quantify gene flow (Bossart & Powell 1998, Broquet & Petit 2009).

For example, direct methods may be valid when one parent is known and thus levels of relatedness can be determined (Manel *et al.* 2005). It would be expected in such a case that levels of relatedness in relation to geographic distance away from the parent would

present a measure of natal dispersal. Assignment methods that determine the probability of an individual genotype originating from the population in which it was collected, or from some other population which has been sampled provide a measure of instantaneous dispersal (Manel *et al.* 2005). However sampling intensity required is high as it requires that at least some migrants have been sampled, which may be problematic if migration between populations is either very low or very high (Goudet *et al.* 2002, Paetkau *et al.* 2004).

Rather than looking at individual patterns of dispersal indirect genetic methods assess the consequences of dispersal for population structure after other factors such a mutation/drift equilibrium and selection have been taken into account (Slatkin 1987). Traditionally, indirect methods have been based on Wright's island model and F_{ST} statistic or some derivative of this method (Wright 1951). F_{ST} measures the level of differentiation between subpopulations (as a product of differences in allele frequencies) which is reached as a balance between mutation, genetic drift and gene flow (Wright 1931). The dispersal of individuals (if it results in gene flow) will affect this balance as allele frequencies are altered, and consequently if genes are freely exchanged between populations F_{ST} will be low (but note we are unable to know whether this is due to contemporary gene flow or recent common ancestry). Conversely if populations are isolated, F_{ST} will be high (once equilibrium is reached). Using Wright's island model, F_{ST} can be used to estimate the effective number of migrants (N_em_e) between subpopulations as $N_em_e \sim ¼ ((1/ F_{ST})-1)$. However the assumptions associated with Wright's island model (an infinite number of populations, with equally sized populations exchanging equal numbers of migrants between sub populations, in mutation/drift equilibrium and with no selection) are unrealistic in most natural systems which could lead to erroneous conclusions relating to migration (Whitlock & McCauley 1999).

Because the geographical distribution of a species is normally greater than the dispersal capability of any single individual, populations can become increasingly genetically differentiated with geographic distance (Balloux and Lugon-Moulin, 2002). Analysing the relationship between genetic differentiation and geographic distance (Isolation by distance (Wright 1943)) among individuals can give an indication of the magnitude of dispersal distance (Rousset 1997). The slope of this relationship in a two-dimensional space is expected to be equal to $1/4\pi D\sigma^2$ where D is equal to the effective population

density and σ is equal to the dispersal distance. Thus the magnitude of the slope tells us about both the density of the population at a given geographic distance and the relative dispersal distance of individuals. A shallow slope implies a low density or large dispersal distance, whilst a steep slope implies the opposite.

Recently, coalescent models based on sophisticated Bayesian or likelihood algorithms have offered an alternative approach to F_{ST} to explore population structure and other demographic parameters such as migration and population growth (Pearse & Crandall 2004). Coalescent models describe the properties of samples of genes by looking at their mutational and genealogical history, without assuming mutation-drift equilibrium and are thus able to use more of the information in genetic markers than traditional F_{ST} based methods (Beerli & Felsenstein 1999, Song *et al.* 2006). The use of the coalescent enables the likelihood of various demographic scenarios to be explored given the data and thus allows the history of population demography to be incorporated (Drummond *et al.* 2005). However, there are disadvantages to a coalescent approach, the main one being that sophisticated algorithms require high computational power and may be time consuming (Pearse & Crandall 2004). Further, in order to estimate parameters such as N_e or migration rate it is often required that some other parameter, such as mutation rate is known. For a review of coalescent methodologies and related software packages see Pearse & Crandall (2004) & Excoffier & Heckel (2006).

1.5.1 Measuring sex-biased dispersal

Behavioural differences between the sexes in terms of mating success, philopatry, dispersal distances and kin associations will affect the distribution of genetic variation within and among populations (Lawson-Handley & Perrin 2007, Prugnolle & De Meeus 2002).

Genetic markers located on different chromosomal systems will vary in the signatures of demographic history they record. For example, if markers have biparental inheritance then the contribution of both sexes will be averaged after one generation (Prugnolle & De Meeus 2002). However, if genetic markers have a uniparental mode of inheritance (e.g. maternally inherited mtDNA and the paternally inherited Y-chromosome in mammals), we can follow matrilines and patrilines to investigate male and female population histories more directly (Avise *et al.* 1987, Hedrick 2007, Jobling & Tyler-

Smith 2003). In addition heterogametic sex determination systems, in which one sex has two copies of a particular sex chromosome but the other has one, will record asymmetric male and female population histories (Ellegren 2009, Schaffner 2004).

Traditionally sex-biased population structure has been investigated either by comparing the same marker type (usually autosomal microsatellites) in both sexes separately (Goudet *et al.* 2002, Mossman & Waser 1999, Vitalis 2002) or by grouping the sexes and comparing marker types with different modes of inheritance (Avise 2004, Moritz *et al.* 1987). Although comparing the same marker type is advantageous in that differences are likely to reflect sex-biased demography (rather than differences in mutational properties, effective population size, etc.) it can only provide us with a measure of instantaneous sex-biased dispersal (Vitalis 2002). Because allele frequencies are randomised between males and females in the next generation, any signature of SBD is then lost. Further it has been shown that methods which seek to determine SBD based on biparental markers can only detect a bias if dispersal intensity is of an intermediate value (Goudet *et al.* 2002). In order to detect a bias sampling of individuals must be intensive.

Comparing biparentally with uniparentally inherited markers works on the logic that because one marker records the history of one sex but the other averages over both sexes any discrepancy in terms of population structure between the marker types may be due to behavioural differences between the sexes (Mossman & Waser 1999). In birds and mammals this comparison has traditionally been between maternally inherited mitochondrial DNA and autosomal markers, but the use of paternally inherited Y-chromosome markers is increasing in non-model organisms and has been used to describe sex differences in human population structure for quite some time (Jobling & Tyler Smith 2003).

Comparing different marker types has been the principal method for investigating SBD in mammals and has been applied in many different species including whales (Baker *et al.* 1998), primates (Melnick & Hoelzer 1992), humans (Mesa *et al.* 2000, Perez-Lezaun *et al.* 1999, Seielstad *et al.* 1998) and bats (Burland & Worthington-Wilmer 2001). For example if population structure is more pronounced in mtDNA than at autosomal markers this implies that females are philopatric and that males must be dispersing larger distances to homogenise genetic variation (e.g. Avise 2004, Castella *et al.* 2001,

Chen *et al.* 2008, Kerth *et al.* 2000). However, there are problems with this approach as differences between markers may not reflect sex-biased behaviour but could instead be a product of different mutation rates and/or different effective population sizes between marker types (Birky *et al.* 1989, Chesser & Baker 1996). For example the effective population size (N_e) of mtDNA and the Y-chromosome is 25% that of autosomes. Genetic drift has a more profound influence at smaller effective population sizes, which can act to increase differentiation. Consequently higher levels of differentiation for uniparental markers may just reflect the smaller N_e in comparison with autosomes. Although this problem could be overcome by correcting for the smaller N_e, in natural populations N_e of marker types may not fit assumptions due to non-random mating (Chesser & Baker 1996). As a result, the N_e of uniparental markers relative to autosomes may be greater or smaller than expected under neutral assumptions. In addition, differential mutation rates between markers will influence observed levels of genetic variation and may only be overcome if markers with similar mutation rates are used (Balloux *et al.* 2000).

A more direct method of investigating SBD in mammals would compare levels of variation between mtDNA and Y-chromosomes (Scozzari *et al.* 1999, Wilkins & Marlowe 2006, Wood *et al.* 2005). However, whilst mtDNA markers are widely available, Y-chromosome markers are less so (Petit *et al.* 2002) and have only been extensively used to investigate population structure in humans and other primates (Douadi *et al.* 2007, Erler *et al.* 2004, Jobling & Tyler-Smith 2003, Wilkins & Marlowe 2006). Further, when Y-chromosome markers have been applied to non-model organisms variability is sometimes so low that comparisons become uninformative on a population genetic level (Hellborg & Ellegren 2004). Thus the lack of Y-chromosome markers to investigate population structure is due to two main reasons: difficulty in isolating markers and low variability once markers are isolated. These two factors have recently been reviewed by Greminger *et al.* (2010) and will be discussed in more detail in chapter two.

Recently the X-chromosome has emerged as a useful marker to investigate sex-biased behaviour when used in combination with autosomal markers (Bustamante & Ramachandran 2009, Schaffner 2004). Because the X-chromosome has an N_e which is 0.75 that of autosomes (due to one male and two female copies) any deviation from this ratio that cannot be accounted for by selection or differences in mutation rate may be

attributed to unequal reproductive success and/or SBD (Hammer *et al.* 2008, Keinan *et al.* 2009) . Moreover, because the X-chromosome undergoes recombination when in females, loci on the X-chromosome have different population histories. This gives the potential to increase the power to determine sex-biased dispersal/mating patterns with an X:autosome comparison relative to a mitochondrial:Y comparison (Wilkins & Marlowe 2006). So far the utilisation of the X-chromosome in genetic studies has been largely restricted to humans (but see (Brandlii *et al.* 2005, Hagenblad *et al.* 2009, Melo-Ferreira *et al.* 2009)) but, due to the relative ease of characterising X-linked markers relative to Y-linked markers, it is likely that the use of the X-chromosome in a population genetic context will increase (Schaffner 2004).

It must be noted that although genetic methods present an incredibly useful tool to investigate SBD all have disadvantages and there could be other potential explanations to explain sex related genetic differences that do not relate to SBD. Therefore genetic methods are most useful when combined with field data on mating patterns and dispersal in the sexes which can be used to support inferences made from studies on population genetic structure (Lawson-Handley & Perrin 2007).

1.6 Population structure and SBD in bats: methods and results

Because bats are nocturnal, highly mobile and are found in large numbers, they are difficult to follow using traditional methods and consequently there have been many population genetic studies conducted on bats (Burland & Worthington-Wilmer 2001). The majority of published studies have used autosomal microsatellites to investigate contemporary population structure, mtDNA to investigate phylogeographic and phylogenetic relationships and a comparison of mtDNA/autosomal markers to infer SBD. There has been a recent increase in alternative and more sophisticated methods to study historic and contemporary population structure in bats, for example the use of autosomal microsatellite information to determine phylogeography (Flanders *et al.* 2008) and the use of multi-locus DNA sequence data to investigate male mediated introgression (Mao *et al.* 2010). However, to my knowledge, there are no published studies which have used the Y-chromosome or the X-chromosome to investigate population structure in bats.

The majority of published studies on European species on a small-medium geographic scale tend to report low but significant population structure when using autosomal markers, for example *M. bechsteinii* (Kerth *et al.* 2000), *N. noctula* (Petit & Mayer 1999), *M. myotis* (Castella *et al.* 2001). Because bats are highly capable dispersers gene flow is less restricted by distance or geographic barriers than in terrestrial small mammals, consequently many populations may be considered panmictic (Burland & Worthington-Wilmer 2001). However this may be dependent on species behaviour, some species can be considered migratory as they undergo seasonal migrations to swarming sites which facilitates gene flow between individuals from a large geographic area. For example *M. natterreri* (Rivers *et al.* 2005), *M. daubentonii* (Angell 2008, Ngampraseterwong *et al.* 2008), *M. schreibersii* (Pereira *et al.* 2009). But other species are far more sedentary in their behaviour and in such cases much greater levels of structure have been reported at autosomal markers. For example *R. ferruquinum* (Rossiter *et al.* 2000), *P. auritus* (Burland *et al.* 1999). Phylogeographic studies of bats in Europe tend to support post-glacial colonisation of Europe from one or more glacial refugia (Mayer & von Helversen 2001). Major glacial refugia that tend to be reported are Iberia (Ibanez *et al.* 2006) and/or the Balkans (Kerth *et al.* 2008). However Italy has also been invoked as a refugium for some species (Ruedi *et al.* 2008) and there is a limited understanding of the role other eastern areas might have played, due to lack of sampling from these regions (Flanders *et al.* 2008). The majority of studies investigating sex-biased dispersal in bats have used an mtDNA/autosomal comparison, although a few studies have used a biparental marker comparison between the sexes (Burland & Worthington-Wilmer 2001). In both cases female philopatry and male dispersal is the most often reported interpretation, although the degree of such a bias appears to vary between species and there have also been instances where female dispersal has been reported (Nagy *et al.* 2007) or when both sexes appear philopatric (Burland *et al.* 1999).

1.6.1 Population structure in *M. daubentonii*

There have been both population genetic and phylogenetic studies of *M. daubentonii* on a Europe-wide scale (Atterby *et al.* 2009, Mayer & von Helversen 2001), and smaller scale studies in the Yorkshire Dales (Angell 2008, Senior *et al.* 2005) and Scotland (Ngampraseterwong *et al.* 2008) which have aimed to address SBD. However, this is

the first study which has investigated both the population genetic and phylogeographic structure of *M. daubentonii* at a Europe-wide scale, with a view to addressing SBD.

Mayer & von Helversen (2001) conducted a study across Europe in which they investigated phylogenetic relationships within and among a number of bat species including *M. daubentonii*. They found four different mitochondrial clades in Europe (Germany, Byelorussia, Greece and Spain) suggesting postglacial colonisation of Europe by *M. daubentonii* from several refugia. A recent population genetic study on *M. daubentonii* using autosomal microsatellites (Atterby *et al.* 2009) focussed on samples collected in Britain and Western Europe. They found low population structure across the study area which suggested *M. daubentonii* populations were panmictic across a large scale. Britain appeared to be slightly differentiated from mainland Europe although there was no evidence that the English Channel provided a barrier to dispersal. Isolation by distance was demonstrated in Scotland/northern England suggesting that panmixia does not occur in these regions but there still appeared to be no significant geographic barriers to gene flow. On a smaller scale Ngamprasterwong *et al.* (2008) used mtDNA and microsatellites to study population genetic structure in Scotland. They found two mtDNA clades, which suggested that Scotland had been colonised by *M. daubentonii* from two different refugia. However they found little autosomal structure either within or among mtDNA clades suggesting that dispersal was male biased and/or significant mixing of individuals from different summer colonies occurred during mating at swarming sites. Another small-scale study conducted by Angell (2008) in the Yorkshire Dales used autosomal microsatellites to investigate population structure in male and female Daubenton's bats at both summer roosts and swarming sites. She found stronger population structure in females than males at summer roosts suggesting strong natal philopatry in females and dispersal of at least some males. Overall levels of differentiation across the study area were low suggesting high levels of gene flow, but when populations over 60 km apart were included there was evidence of isolation by distance, suggesting that gene flow is more restricted over a large geographic scale.

Levels and patterns of genetic diversity in a species are a product of both historic and contemporary forces (Hewitt & Butlin 1997). By using markers with different mutational properties at appropriate spatial scales the relative contribution of these forces in shaping population genetic structure can be determined. In summary, previous studies of *M. daubentonii* have found large scale structure in mitochondrial markers

which is a consequence of historical postglacial colonisation, but much weaker structure at autosomal markers on a small-medium scale as a product of widespread contemporary gene flow, which is probably mediated by mating of bats at swarming sites from a wide catchment area and longer male dispersal distances.

This thesis represents the largest scale genetic study on *M. daubentonii* to date, with samples from across Britain and Europe, but it also focuses on smaller scale patterns of genetic structure with extensive sampling in Yorkshire. By combining data from rapidly evolving autosomal microsatellites and mitochondrial markers with lower mutation rates, at both spatial scales, the aim is to gain a comprehensive picture of the role historic and contemporary forces have played in shaping current levels of genetic diversity and to identify past and present barriers to gene flow in *M. daubentonii*. The addition of X-linked microsatellites adds to our understanding of sex differences in dispersal and gene flow because an X/autosome comparison overcomes some of the problems inherent in inferring SBD from autosomal/mt markers. Understanding dispersal in *M. daubentonii* will enhance understanding of EBLV-2 epidemiology in Britain and Europe and will allow a better understanding of the potential for EBLV-2 spread.

1.7 Notes on samples

Wing biopsies were collected by members of Professor John Altringham's research group at the University of Leeds, with the following exceptions: DNA samples from Scotland were provided by Professor Stuart Piertney (University of Aberdeen). DNA samples from Finland were provided by Veronika Laine (University of Turkuu). DNA samples from Denmark were provided by Liselotte Andersen (Danmarks Miljoundersogelser Universitet). DNA samples from Belgium and Shropshire (Stokesay Castle) were provided by the Food & Environment Research Agency (FERA).

1.8 References

Altringham JD (1996) *Bats: Biology & Behaviour.* Oxford University Press, Oxford.

Amengual B, Whitby JE, King A, Cobbo J, Bourhy H (1997) Evolution of European bat lyssaviruses. *Journal of General Virology* 78 2319-2328.

Angell RL (2008) Population ecology and mating system of Daubenton's bat (*Myotis daubentonii*) PhD Thesis, University of Leeds.

Atterby H, Aegerter JN, Smith GC *et al.* (2009) Population genetic structure of Daubenton's bat (*Myotis daubentonii*) in western Europe and the associated occurrence of rabies. *European Journal of Wildlife Research* **56** 67-81.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology & Systematics* **18** 489-522.

Avise JC (2004) *Molecular Markers Natural History and Evolution.* Sinauer Associates, Sunderland, Massachusetts.

Baker CS, Medrana-Gonzalez L, Calambokidis J *et al.* (1998) Population structure of nuclear and mitochondrial DNA variation among humpback whales in the North Pacific. *Molecular Ecology* 7 695-707.

Balloux F, Brünner H, Lugon-Moulin N, Hausser J, Goudet J (2000) Microsatellites can be misleading: an empirical and simulation study. *Evolution* **54** 1414–1422.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11** 155-165.

Barclay RMR (1991) Population structure of temperate zone insectivorous bats in relation to foraging behaviour and energy demand. *Journal of Animal Ecology* **60** 165-178.

Beerli P, Felsenstein J (1999) Maximum likelihood estimation of migration rates and population numbers in two populations using a coalescent approach. *Genetics* **152** 763-773.

Birky CW, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration, mutation and drift: Equlibrium expectations, approach to equilibrium, effects of heteroplasmic cells and comparison to nuclear genes. *Genetics* **121** 613-627.

Blundell GM, Ben-David M, Groves P, Bowyer RT, Geffen E (2002) Characteristics of sex-biased dispersal and gene flow in coastal river otters: implications for natural recolonization of extirpated populations. *Molecular Ecology* **11** 289–303.

Bogdanowicz W (1994) *Myotis daubentonii. Mammalian Species* **475** 1-9.

Bossart JL, Powell P (1998) Genetic estimates of population structure and gene flow: limitations, lessons & new directions. *TREE* **13** 202-206.

Bowler DE, Benton TG (2005) Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biology Reviews* **80** 205–225.

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

Brooks SM, Aegerter JN, Smith GC *et al.* (2005) European bat lyssavirus in Scottish bats. *Emerging Infectious Diseases* **11** 572-578.

Broquet T, Petit EJ (2009) Molecular estimation of dispersal for ecology and population genetics. *Annual Review of Ecology, Evolution & Systematics* **40** 193-216.

Burland TM, Barratt EM, Nichols RA, Racey PA (1999) Mating patterns, relatedness and the basis of natal philopatry in the brown long-eared bat, *Plecotus auritus*. *Molecular Ecology* **10** 1309-1321.

Burland TM, Worthington-Wilmer J (2001) Seeing in the dark: molecular approaches to the study of bat populations. *Biological Reviews* **76** 389-409.

Bustamante CD, Ramachandran S (2009) Evaluating signatures of sex specific processes in the human genome. *Nature Genetics* **41** 8-10.

Calisher CH, Childs JE, Field HE, Holmes KV, Schountz T (2006) Bats: Important reservoir hosts of emerging viruses. *Clinical Microbiology Reviews* **19** 531-545.

Castella V, Ruedi M, Excoffier L (2001) Contrasted patterns of mitochondrial and nuclear structure among nursery colonies of the bat *Myotis myotis*. *Journal of Evolutionary Biology* **14** 708-720.

Charlesworth D, Charlesworth B (1987) Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology & Systematics* **18** 237-268.

Chen S, Jones G, Rossiter SJ (2008) Sex-biased gene flow and colonisation in the Formosan lesser horseshoe bat: inference from nuclear and mitochondrial markers. *Journal of Zoology* **274** 207-215.

Chesser RK, Baker RJ (1996) Effective sizes and dynamics of uniparentally and biparentally inherited genes. *Genetics* **144** 1225–1235.

Clarke A (1978) Sex ratio and local resource competition in a prosimian primate. *Science* **201** 163–165.

Clutton-Brock TH (1989) Female transfer and inbreeding avoidance in mammals. *Nature* **337** 70–71.

Clutton-Brock TH, Harvey PH (1978) Mammals, resources and reproductive strategies. *Nature* **273** 191-195.

Daszak P, Cunningham AA, Hyatt AD (2000) Emerging infectious diseases of wildlife – threats to biodiversity and human health. *Science* **287** 443-449.

Day SB, Bryant EH, Meffert LM (2007) The influence of variable rates of inbreeding on fitness, environmental responsiveness & evolutionary potential. *Evolution* **57** 1314-1324.

Deal B, Farello C, Lancaster M, Kompare T, Hannon B (2000) A dynamic model of the spatial spread of an infectious disease: the case of fox rabies in Illinois. *Environmental Modelling & Assessment* **5** 47-62.

Dobson FS (1982) Competition for mates and predominant juvenile male dispersal in mammals. *Animal Behaviour* **30** 1183–1192.

Douadi MI, Gatti S, Levrero F *et al.* (2007) Sex biased dispersal in western lowland gorillas (*Gorilla gorilla gorilla*). *Molecular Ecology* **16** 2247-2259.

Drummond AJ, Rambaut A, Shapiro B, Pybis OG (2005) Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology & Evolution* **22** 1185-1192.

Ellegren, H (2009) The different levels of genetic diversity in sex chromosomes and autosomes. *Trends in Genetics* **25** 278-284.

Erler A, Stoneking M, Kayser M (2004) Development of Y-chromosomal microsatellite markers for nonhuman primates. *Molecular Ecology* **13** 2921-2930.

Encarnação JA, Dietz M (2006) Estimation of food intake and ingested energy in Daubenton's bats (*Myotis daubentonii*) during pregnancy and spermatogenesis. *European Journal of Wildlife Research* **52** 221-227.

Excoffier L, Heckel G (2006) Computer programs for population genetics data analysis: a survival guide. *Nature Reviews Genetics* **7** 745-758.

Flanders J, Jones G, Benda P *et al.* (2009) Phylogeography of the greater horseshoe bat, *Rhinolophus ferrunquinum*: Contrasting results from mitochondrial and microsatellite data. *Molecular Ecology* **18** 206-318.

Flavin DA, Biggane SS, Shiel CB, Smiddy P, Fairley JS (2001) Analysis of the diet of Daubenton's bat *Myotis daubentonii* in Ireland. *Acta Theriologica* **46** 43-52.

Fooks AR, Brookes SM, Johnson N, McElhinney LM, Hutson AM (2003) European bat lyssavirus: an emerging zoonosis. *Epidemiology and Infection* **131** 1029–1039.

Fontanillas P, Petit E, Perrin N (2004) Estimating sex-specific dispersal rates with autosomal markers in hierarchically structured populations. *Evolution* **58** 886–894.

Frantz SC, Trimarchi CV(1983) *Bats in human dwellings:health concerns and management*. In: Decker, D.F. (Ed.), Proceedings of the First Eastern Wildlife Damage Control Conference. Cornell University, Ithaca, NY, pp. 299–308.

Freuling C, Vos A, Johnson N *et al.* (2009) Experimental infection of serotine bats (*Eptesicus serotinus*) with European bat lyssavirus type 1a. *Journal of General Virology* **90** 2943-2502.

Gandon S (1999) Kin competition, the cost of inbreeding and the evolution of dispersal. *Journal of Theoretical Biology* **200** 345–364.

Glover AM, Altringham JD (2008) Cave selection and use by swarming bat species. *Biological Conservation* **141** 1493-1504.

Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology* **11** 1103-1114.

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour* **28** 1140-1162.

Greminger M, Krutzen M, Schelling C, Pienkowska A, Wandeler P (2010) The quest for Y-chromosomal markers – methodological strategies for mammalian non-model organisms. *Molecular Ecology Resources* **10** 409-420.

Guillaume F, Perrin N (2006) Joint evolution of dispersal and inbreeding load. *Genetics* **173** 479–509.

Hagenblad J, Olsson M, Parker H, Ostrander EA, Ellegren H (2009) Population genomics of the inbred Scandinavian wolf. *Molecular Ecology* **18** 1341-1351.

Hammer MF, Mendex FL, Cox MP, Woerner AE, Wall JD (2008) Sex biased evolutionary forces shape patterns of human genomic diversity. *PLoS Genetics* **4** 1-8.

Harris SL, Brookes SM, Jones G *et al.* (2006a) Passive surveillance (1987 to 2004) of United Kingdom bats for European bat lyssaviruses. *The Veterinary Record* **159** 439-446.

Harris SL, Brookes SM, Jones G *et al.* (2006b) European bat lyssaviruses: Distribution, prevalence and implications for conservation. *Biological Conservation* **131** 193-210.

Harris SL, Aegerter JN, Brooks SM *et al.* (2009) Targeted surveillance for European bat lyssavirus in English bats (2003-06). *Journal of Wildlife Diseases* **45** 1030-1041.

Hedrick PW (2007) Sex: Difference in mutation, recombination, selection, gene flow & genetic drift. *Evolution* **61** 2750-2771.

Hellborg L, Ellegren H (2004) Low levels of nucleotide diversity in mammalian Y chromosomes. *Molecular Biology and Evolution* **21** 158-163.

Hewitt GM, Butlin RK(1997) Causes and consequences of population structure. In: *Behavioural Ecology: An Evolutionary Approach*, edited by J. R. Krebs and N. B. Davies. pp 203. Blackwell Science Ltd, Oxford.

Holmala K, Kauhala K (2006) Ecology of wildlife rabies in Europe. *Mammal Review* **36** 17-36.

Holmes DJ, Austad SN (1994) Fly now die later: Life history correlates of gliding and flying mammals. *Journal of Mammology* **75** 224-226.

Hutterer R, Ivanova T, Meyer-Cords C, Rodrigues L (2005) *Bat Migrations in Europe.* Naturschutz und Biologische Vielfalt.

Ibanez C, Garcia-Mudarra J, Ruedi M *et al.* (2006) The Iberian contribution to cryptic diversity in European bats. *Acta chiropterologica* **8** 277-297.

Jan C, Frith KE, Glover AM, Butlin R, Greenaway F, Ruedi M & Altringham JD (2010) *Myotis alcathoe* in the UK. *Acta Chiropterologica. In Press.*

Ji W, Sarre SD, Aitken N, Hankin RKS, Clout MN (2001) Sex-biased dispersal and a density-independent mating system in the Australian brushtail possum, as revealed by minisatellite DNA profiling. *Molecular Ecology* **10** 1527–1537.

Jin Y, Wang WD (2005) The effect of population dispersal on the spread of a disease. *Journal of Mathematical Analysis and Applications* **308** 343–364.

Jobling MA, Tyler-Smith C (2003) The human Y-chromosome: an evolutionary marker comes of age. *Nature Review Genetics* **4** 508-612.

Johnson ML, Gaines MS (1990) Evolution of dispersal: theoretical models and empirical tests using birds and mammals. *Annual Review of Ecology and Systematics* **21** 449–480.

Johnson N, Seldon D, Parsons G, Fooks AR(2002). European bat lyssavirus type 2 in a bat found in Lancashire. *Veterinary Record* **151** 455–456.

Johnson N, Vos A, Neubert L *et al.* (2008) Experimental study of European bat lyssavirus type 2 infection in Daubenton's bat (*Myotis daubentonii*). *Journal of General Virology* **89** 2662-2672.

Jones KE, MacLarnon (2001) Bat life histories: Testing models of mammalian life history evolution. *Evolutionary Ecology Research* **3** 465-476.

Jones G, Teeling EC (2006) The evolution of echolocation in bats. *TREE* **21** 149-156.

Keinan A, Mullikin JC, Patterson N, Reich D (2009) Accelerated genetic drift on the X chromosome during human migration out of Africa. *Nature Reviews Genetics* **41** 66–70.

Kerth G, Mayer F, Petit E (2002) Extreme sex-biased dispersal in the communally breeding, nonmigratory Bechstein's bat (*Myotis bechsteinii*). *Molecular Ecology* **11** 1491-1498.

Kerth G, Kiefer A, Trappmann C, Weishaar M (2003) High gene diversity at swarming sites suggest hot spots for gene flow in the endangered Bechstein's bat. *Conservation Genetics* **4** 491-499.

Kerth G, Petrov B, Conti A *et al.* (2008) Communally breeding Bechstein's bats have a social system that is independent from the postglacial history and location of the populations. *Molecular Ecology* **17** 2368-2381.

Knudsen JT, Tollsten L (2008) Floral scent in bat-pollinated plants: a case of convergent evolution. *Botanical Journal of the Linneaen Society* **119** 45-57.

Koenig WD, Van Vuren D, Hooge PN (1996) Detectability, philopatry and the distribution of dispersal distances in vertebrates. *Trends in Ecology & Evolution* **11** 514-517.

Kruse H, Kirkemo A-M, Handeland K (2004) Wildlife as source of zoonotic infections. *Emerging Infectious Diseases* **10** 2067-2072.

Kuzmin IV, Franka RF, Rupprecht CE (2008). Experimental infection of big brown bats (*Eptesicus fuscus*) with West Caucasian bat virus (WCBV). *Developmental Biology* **131** 327–337.

Lambin X (1994) Sex-ratio variation in relation to female philopatry in Townsend voles. *Journal of Animal Ecology* **63** 945–953.

Lawson-Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex biased dispersal. *Molecular Ecology* **16** 1559-1578.

Lehmann L, Perrin N (2002) Altruism, dispersal, and phenotype matching kin recognition. *American Naturalist* **159** 451–468.

Li KS, Guan Y, Wang J *et al.* (2004) Genesis of a highly pathogenic and potentially pandemic H5N1 influenza virus in eastern Asia. *Nature* **430** 209-213.

Long ES, Diefenbach, DR, Rosenberry CS, Wallingford BD, Grund MD (2005) Forest cover influences dispersal distance of white tailed deer. *Journal of Mammology* **86** 623-629.

Lumio J, Hillborn M, Roine R *et al.* (1986). Human rabies of bat origin in Europe. *Lancet* **1** 378.

Manel S, Gaggiotti O, Waples RS (2005) Assignment methods: matching biological questions with appropriate techniques. *TREE* **20** 136-142.

Mao X, Zhang J, Zhang S, Rossiter SJ (2010) Historical male-mediated introgression in horseshoe bats revealed by multi-locus DNA sequence data. *Molecular Ecology* **19** 1352-1366.

Mayer F, Dietz C, Kiefer A (2007) Molecular species identification boosts bat diversity. *Frontiers in Zoology* **4** 1-5.

Mayer F, von Helverson O (2001) Cryptic diversity in European bats. *Proceedings of the Royal Society of London, Series B* **268** 1825-1832.

McCracken GF, Wilkinson GS (2000) Bat mating systems. In: *Reproductive Biology of Bats*, edited by E. G. Crichton and P. H. Krutzsch. pp 321-362. Academic Press, London and San Diego.

Melnick DJ, Hoelzer GA (1992) Difference in male and female macaque dispersal lead to contrasting distributions of mitochondrial and nuclear DNA variation. *International Journal of Primatology* **13** 379-393.

Melo-Ferreira J, Alves PC, Freitas H, Ferrand N, Boursot P (2009) The genomic legacy from the extinct *Lepus timidus* to the three hare species of Iberia: contrast between mtDNA, sex chromosomes and autosomes. *Molecular Ecology* **18** 2643-2658.

Mesa NR, Mondragon MC, Soto ID *et al.* (2000) Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: pre- and post-Columbian patterns of gene flow in South America. *American Journal of Human Genetics* **67** 1277–1286.

Messenger SL, Smith JS, Orciari LA, Yager PA, Rupprecht CE (2003) Emerging pattern of rabies deaths and increased viral infectivity. *Emerging Infectious Diseases* **9** 151–154.

Mossman CA, Waser PM (1999) Genetic detection of sex-biased dispersal. *Molecular Ecology* **8** 1063-1067.

Nagy M, Heckel G, Voigt CC (2007) Female-biased dispersal and patrilocal kin groups in a mammal with resource-defence polygyny. *Proceedings of the Royal Society of London, Series B* **274** 3019-3025.

Nathan R, Perry G, Cronin J, Strand A, Cain M (2003) Methods for estimating long distance dispersal. *OIKOS* **103** 261-173.

Neumann G, Noda T, Kawaoka Y (2009) Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* **459** 931-939.

Ngamprasertwong T, Mackie IJ, Racey PA, Piertney SB (2008) Spatial distribution of mitochondrial and microsatellite DNA variation in Daubenton's bat within Scotland. *Molecular Ecology* **17** 3243-3258.

Ochman K (2003) Late Pleistocene and Holocene bats (Chiroptera) from the Komarowa Cave (Cracow-Czêstochowa Upland, Poland) – preliminary results. *Acta zoologica cracoviensia* **46** 73-84

O'Shea TJ, Shankar V, Bowen RA, Rupprecht CE, Wimsatt JH (2003) Do bats acquire immunity to rabies? Evidence from the field. *Bat Research News* **44** 161.

Ostfeld RS, Glass GE, Keesing F (2005) Spatial epidemiology: an emerging (or re-emerging) discipline. *Trends in Ecology and Evolution* **20** 328-336.

Packer C, Pusey AE (1987) Intrasexual cooperation and the sex-ratio in African lions. *American Naturalist* **130** 636–642.

Paetkau D, Slade R, Burden M, Estoup A (2004) Direct, real time estimation of migration using assignment methods: a simulation based exploration of accuracy and power. *Molecular Ecology* **13** 55-65.

Parsons KN, Jones G (2003) Dispersion and habitat use by *Myotis daubentonii* and *Myotis nattereri* during the swarming season: implications for conservation. *Animal Conservation* **6** 283-290.

Pearse DE, KA Crandall (2004) Beyond FST: analysis of population genetic data for conservation. *Conservation Genetics* **5** 585–602.

Pereira MJR, Salgueiro P, Rodrigues L, Coelho MM, Palmeirim JM (2009) Population structure of a cave-dwelling bat, *Miniopterus schreibersii:* Does it reflect history and social organisation? *Journal of Heredity* **100** 533-544.

Perez-Lezaun A, Calafell F, Comas D *et al.* (1999) Sex-specific migration patterns in central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA. *American Journal of Human Genetics* **65** 208–219.

Perrin N, Goudet J (2001) *Inbreeding, kinship and the evolution of natal dispersal.* In: Dispersal (eds Clobert J *et al.*), pp. 123–142. Oxford University Press, Oxford.

Perrin N, Lehmann L (2001) Is sociality driven by the costs of dispersal or the benefits of philopatry? A role for kin-discrimination mechanisms. *American Naturalist* **158** 471–483.

Perrin N, Mazalov V (2000) Local competition, inbreeding and the evolution of sex-biased dispersal. *American Naturalist* **55** 116–127.

Petit E, Mayer F(1999) Male dispersal in the noctule bat (*Nyctalus noctula*): where are the limits? *Proceedings of the Royal Society of London, Series B* **266** 1717-1722.

Petit E, Balloux F, Excoffier L (2002) Mammalian population genetics: why not Y? *Trends in Ecology & Evolution* **17** 28-33.

Prugnolle F, de Meuss T (2002) Inferring sex-biased dispersal from population genetic tools: a review. *Heredity* **88** 161-165.

Pusey AE (1987) Sex-biased dispersal and inbreeding avoidance in birds and mammals. *Trends in Ecology & Evolution* **2** 295–299.

Racey PA, Fooks A (2005) Rabies in bats in Britain. *Biologist* **52** 34-38.

Reed KD, Meece JK, Henkel JS, Shukla SK (2003) Birds, Migration and Emerging Zoonoses: West Nile Virus, Lyme Disease, Influenza A and Enteropathogens. *Clinical Medicine & Research* **1** 5-12.

Rivers NM, Butlin RK, Altringham JD (2005) Genetic population structure of Natterer's bats explained by mating at swarming sites and philopatry. *Molecular Ecology* **14** 4299-4312.

Rivers NM, Butlin RK, Altringham JD (2006) Autumn swarming behaviour of Natterer's bats in the UK: population size, catchment area and dispersal. *Biological Conservation* **127** 215-226.

Ronce O, Olivieri I, Clobert J, Danchin E (2001) *Perspectives on the study of dispersal evolution.* In: *Dispersal* (ed. Clobert J), pp. 314–357. Oxford University Press, Oxford.

Roussett FR (1997) Genetic differentiation and estimates of gene flow from F-statistics under isolation by distance. *Genetics* **145** 1219-1228.

Rossiter SJ, Jones G, Ransome RD, Barratt EM (2000) Genetic variation and population structure in the endangered greater horseshoe bat (*Rhinolophus ferrumequinum*). *Molecular Ecology* **9** 1131-1135.

Rossiter SJ (2001) Outbreeding increases offspring survival in wild greater horseshoe bats (*Rhinolophus ferrumequinum*). *Proceedings of the Royal Society B: Biological Sciences* **268** 1055-1061.

Rubenstein DR, Hobson KA (2004) From birds to butterflies: animal movement patterns and stable isotopes. *TREE* **19** 256-263.

Ruedi M, Walter S, Fischer MC, Scaravelli D, Excoffier L, Heckel D (2008) Italy as a major ice age refuge area for the bat *Myotis myotis* in Europe. *Molecular Ecology* **17** 1801-1814.

Schaffner SF (2004) The X-chromosome in population genetics. *Nature Reviews Genetics* **5** 43-51.

Schmidt K (1998) Variation in daily activity of free ranging Eurasian Lynx (*Lynx lynx*) in Poland. *Journal of Zoology* **249** 417-425.

Schober W, Grimmberger E (1987) *Die Fledermäuse Europas: Kennen . Bestimmen .* Schützen. Franckh'sche Verlagshandlung, Stuttgart.

Scozzari R, Cruciani F, Santolomazza P *et al.* (1999) Combined use of biallelic and microsatellite Y polymorphisms to infer affinities among African populations. *American Journal of Human Genetics* **65** 829-846.

Seielstad MT, Minch E, Cavalli-Sforza LL (1998) Genetic evidence for a higher female migration rate in humans. *Nature Genetics* **20** 278–280.

Senior P, Butlin RK, Altringham JD (2005) Sex and segregation in temperate bats. *Proceedings of the Royal Society of London, Series B* **272** 2467-2473.

Simmons NB, Geisler JH (1998) Phylogenetic relationships of *Icaronycteris, Archaeonycteris, Palaeochiropteryx* to extant bat lineages, with comments on the evolution of echolocation and foraging strategies in Microchiroptera. *Bulletin of the American Museum of Natural History* **23** 1–182.

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

Smith RP, Rand PW, Lacombe EH, Morris SR, Holmes DW, Caporale DA (1996) Role of bird migration in the long distance dispersal of *Ixodes dammini* the vector of Lyme disease. *The Journal of Infectious Diseases* **174** 221-224.

Song S, Dey DK, Holsinger KE (2006) Differentiation among populations with migration, mutation, and drift: Implications for genetic inference. *Evolution* **60** 1-12.

Teeling EC, Springer MS, Madsen O *et al.* (2005) A molecular phylogeny for bats illuminates biogeography and the fossil record. *Science* **307** 580-584.

Teeling EC (2009) Hear, hear: the convergent evolution of echolocation in bats? *TREE* **24** 351-354.

Turchin P, Thoeny WT (1993) Quantifying dispersal of southern pine beetles with mark –recapture experiments and a diffusion model. *Ecological Applications* **3** 187-198.

Urbancyzk Z (1990) Northern Europe's most important hibernation site. *Oryx* **24** 30-34.

Van der Poel WHM, Van der Heide R, Verstraten ER *et al.* (2005) European Bat Lyssaviruses, the Netherlands. Emerging Infectious Diseases 11 [serial on the Internet].

[01/12/05]. Available from: <http://www.cdc.gov/ncidod/EID/vol11no12/04-1200.htm>.

Vaughan N (1997) A review of the diets of British bats (Chiroptera). *Mammal Review* **27** 77-94.

Van Vuren D & Armitage KB (1994) Survival of dispersing and philopatric yellow bellied marmots: What is the cost of dispersal? *OIKOS* **69** 179-181.

Vitalis R (2002) Sex-specific differentiation and coalescence times: estimating sex-biased dispersal rates. *Molecular Ecology* **11** 125-138.

Warrell MJ, Warrell DA (2004) Rabies and other lyssavirus diseases. *The Lancet* **363** 959–969.

Warren RD, Waters DA, Altringham JD, Bullock DJ (2000) The distribution of Daubenton's bats (*Myotis daubentonii*) and pipistrelle bats (*Pipistrellus pipistrellus*) (Vespertilionidae) in relation to small-scale variation in riverine habitat. *Biological Conservation* **92** 85-91.

Waser PM, Austad SN, Keane B (1986) When should animals tolerate inbreeding? *American Naturalist* **128** 529–537.

Watson A, Moss R, Parr R, Mountford MD, Rothery P (1994) Kin landownership, differential aggression between kin and nonkin, and population fluctuations in red grouse. *Journal of Animal Ecology* **63** 39–50.

Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: F_{ST}=1/(4*Nm*+1). *Heredity* **82** 117-125.

Wickramasinghe L, Harris S, Jones G, Vaughan Jennings N (2004) Abundance and species richness of nocturnal insects on organic and conventional farms: effects of agricultural intensification on bat foraging. *Conservation Biology* **18** 1283-1292.

Wilkins JF, Marlowe FW (2006) Sex-biased migration in humans: what should we expect from genetic data? *Bioessays* **28** 290–300.

Wilkinson JS, South JM (2002) Life history, ecology & longevity in bats. *Aging Cell* **1** 124-131.

Wilson GJ, Delahay RJ (2001) A review of methods to estimate the abundance of terrestrial carnivores using field signs and observation. *Wildlife Research* **28** 151-164.

Winters JB, Waser PM (2003) Gene dispersal and outbreeding in a philopatric mammal. *Molecular Ecology* **12** 2251–2259.

Wood ET, Stover DA, Ehret C *et al.* (2005) Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. *European Journal of Human Genetics* **13** 867–876.

Wright S (1931) Evolution in Mendelian populations. *Genetics* **16** 97-159.

Wright S (1943) Isolation by distance. *Genetics* **28** 139-156.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15** 323-354.

Zollner PA, Lima SL (2001) Search strategies for landscape level interpatch movements. *Ecology* **80** 1019-1030.

Chapter 2: Methodological attempts to characterise Y-specific markers in Daubenton's bat

2.1 Introduction

The use of autosomal and mitochondrial DNA markers to study population history, population structure and sex-biased dispersal is widespread in ecology (Hedrick 2005, Moritz 1994). However, up until recently the Y-chromosome has, by comparison, remained neglected in ecological genetic studies (Greminger *et al.* 2010, Lawson-Handley & Perrin 2007, Petit *et al.* 2002). This is surprising as Y-chromosome markers would allow male-specific population parameters to be investigated; used in conjunction with autosomal and mtDNA, Y markers would provide a particularly informative picture of the contribution each sex makes to demographic parameters.

In mammals, males are the heterogametic sex and carry both an X and a Y chromosome (in birds the reverse is true; females are ZW and males ZZ). The vast majority of the Y-chromosome is non-recombining (excluding a small pseudo-autosomal region) and as such can be used to define patrilines and hence the evolutionary past of males (Stumpf & Goldstein 2001). Because it is thought that the majority of the Y-chromosome contains non-coding or 'junk' DNA (Skaletsky *et al.* 2003), it is not expected to be subject to the selective pressures of coding regions and should therefore be free to accumulate mutations. However, because the Y-chromosome behaves as a single linked locus if selection acts on one part of the chromosome this can have an influence on the rest of the chromosome, which means selective forces may be influencing patterns of variation in addition to demographic processes (Balloux 2010).

The potential of studying the Y-chromosome to make inferences about male-mediated population structure has long been recognised in humans, with the first polymorphic markers being discovered in 1985 (Casanova *et al.* 1985). Initially progress in identifying new markers was slow (Jobling & Tyler-Smith 1995) but since the human genome sequence was published (provisional sequence in 2000, final sequence in 2003 (Ensembl)) Y-markers have been far easier to identify and, as a result, numerous new markers have been discovered over the last decade (e.g. Jobling & Tyler-Smith 2003, Kayser *et al.* 2004, Repping *et al.* 2006) which have illuminated various aspects of human population history including mating systems (Dupanloup *et al.* 2003), sex-biased

dispersal and 'out of Africa' models of colonisation (Hammer *et al.* 1998, Wood *et al.* 2005, Underhill & Kivislid 2007). They have also have also had other applications including paternity testing and forensics (Gill *et al.* 2001, Jobling *et al.* 1997).

However despite the potential of Y-chromosome markers to look at sex-biased processes, the use of Y-chromosome markers in other species (especially non-model organisms and non-primates) has remained limited (Greminger *et al.* 2010, Petit *et al.* 2002). In order to be informative about populations a Y-chromosome marker should be specific to males, have single locus amplification and be sufficiently polymorphic. It is the difficulty in meeting all three criteria in most natural populations which has so far restricted the characterisation and use of Y-linked markers (Hellborg & Ellegren 2004, Lawson-Handley *et al.* 2006, Luo *et al.* 2007).

The unique architecture of the Y-chromosome is largely responsible for both the difficulty in characterising Y-specific single-copy markers and for the low levels of polymorphism observed. An X-Y sex determination system appears to have arisen independently in a number of lineages and both the X- and Y-chromosomes are thought to have originated from a pair of autosomes which subsequently developed sex determination functions that resulted in selection favouring suppression of recombination between the heterogametic chromosomes (Charlesworth 1996, Lahn & Page 1999, Rice 1996). An exception is the pseudo-autosomal region (PAR) on the Y-chromosome which recombines with the X during meiosis, though it is thought that the PAR is very small in most species; for example in humans it only represents 5% of the Y-chromosome (Graves *et al.* 1998, Rozen *et al.* 2003). The rest of the Y-chromosome is termed the Male-specific Region (MSY) and in many species is largely composed of several blocks of palindromic heterochromatic sequences and a smaller euchromatic region with extensive homology to the X-chromosome (Skaletsky *et al.* 2003).

Due to the highly repetitive nature of the heterogametic region it has largely remained ignored in most studies, as its utility for most biological applications is thought to be limited (Greminger *et al.* 2010). The euchromatic region shows more promise and contains three different sequence classes: a) ampliconic segments which are highly repetitive and show great similarity, over hundreds of kilobases throughout the MSY, b) X-transposed sequences which were transposed from the X- to the Y-chromosome millions of years ago and share high sequence similarity with regions on the X-

chromosome and c) X-degenerate sequences which were present on the ancient proto-sex chromosomes and comprise single-copy homologues to X-linked genes (Skaletsky *et al.* 2003). Consequently even within the euchromatic region it becomes apparent that characterisation of male-specific Y-linked markers is difficult due to high levels of sequence repetition, transposition and homology to the X-chromosome.

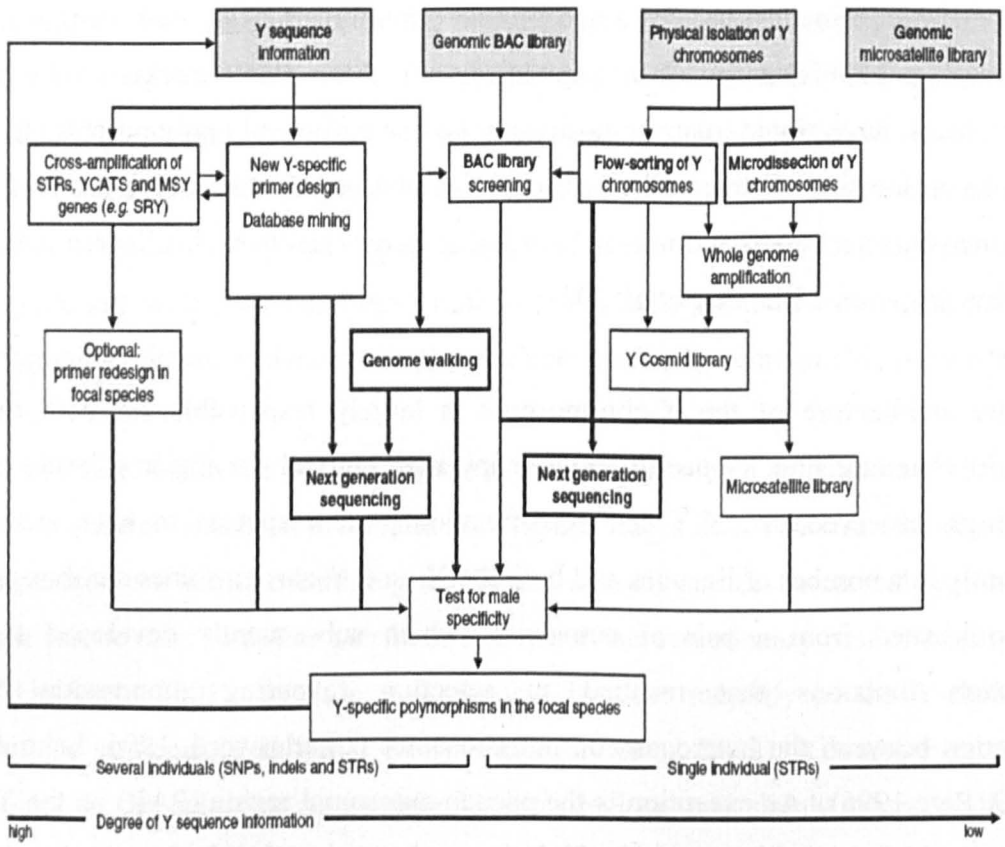


Figure 2.1 Methodological approaches to characterising single-copy Y-chromosome markers. Taken from Greminger *et al.* 2010.

There are an increasing number of methodologies (summarised in Figure 2.1 taken from Greminger *et al.* 2010) which have allowed the isolation of male-specific Y-chromosome markers in a number of species (Andres *et al.* 2008, Hailer & Leonard 2008, Lawson-Handley & Perrin 2006, Luo *et al.* 2007, Macdonald *et al.* 2006, Wallner *et al.* 2004). Yet even when Y-chromosome markers are characterised it is often the case that they are monomorphic (Hellborg & Ellegren 2004) or they are only polymorphic enough to be informative at a phylogenetic scale. The unique structure of the Y-chromosome and the lack of recombination offer explanations as to why this is so. After acquiring sex determining functions recombination between the X- and the Y-chromosome was gradually suppressed and the majority of functional variation on the

Y-chromosome was lost (Ohno 1967). Once the Y-chromosome no longer recombined the MSY essentially behaved like a single linked locus (Rice 1996). Linkage with no recombination is thought to have paved the way for degeneration of the Y-chromosome via a series of possible mechanisms (Charlesworth & Charlesworth 2000). These include the accumulation of deleterious mutations via background selection (Charlesworth 1994), Muller's ratchet (Gordo & Charlesworth 2001, Muller 1964), hitchhiking of beneficial mutations (Rice 1987) and lack of adaptation due to the interference of positive mutations with linked deleterious alleles (Peck 1994). The study of proto-Y chromosomes in a number of organisms indicates that these mechanisms may differ according to the species in question (Bachtrog 2005, Filatov 2004, Peichel *et al.* 2004).

This chapter aimed to characterise Y-specific, single-copy markers in *Myotis daubentonii* via a number of methods including the use of degenerate primers (Hellborg & Ellegren 2003), characterisation of Y-linked regions using AFLP (Griffiths & Orr 1999) and sequencing of a flow sorted *Myotis* Y-chromosome. All methods were unsuccessful in the discovery of Y-specific markers and the discussion section attempts to explain why and offer alternative methodologies.

2.2 Methods and results

2.2.1 DNA extraction

Genomic DNA was extracted from 3 mm wing biopsies from males and females using a standard ammonium acetate precipitation method adapted to a 96 well plate extraction (http://www.sheffield.ac.uk/molecol/protocols.html). Each DNA extract was diluted in 30 µl of Low TE buffer.

2.2.2 Y-chromosome anchor tagged sequences (YCATS)

Hellborg and Ellegren (2003) reported a new set of conserved Y-chromosome primers termed Y-chromosome anchor tagged sequences (YCATS), designed to amplify across mammalian orders. The primers are designed to bind to a conserved exonic sequence associated with a single-copy Y-linked gene and to amplify the adjacent intronic region. The primers were designed by aligning homologous Y-chromosome sequences from

humans and at least one other species (usually mice) and identifying regions that were conserved between these species. The corresponding gametologue from the X-chromosome of the two species was then aligned with the Y-sequences to determine Y specificity. Because humans and mice are distantly related, if there is sequence homology at particular regions it is likely that the sequences will be conserved in other species too. Hellborg & Ellegren (2003) were able to obtain a single-copy male-specific PCR product in 21 different species for at least four of their YCAT primers. Eleven primers were reported to produce a single male-specific PCR product in *M. daubentonii* DBY7, DBY8, DBY13, DBY14, SMCY1, SMCY3, SMCY11, SMCY16, SMCY17, UTY11 and ZFY2 (see Hellborg & Ellegren 2003 for details of primers).

In order to test for positive amplification, male-specificity and to determine optimal PCR conditions for each primer pair in *M. daubentonii* a gradient PCR was performed using the following conditions: 95 °C for 15 minutes, followed by 35 cycles of 94 °C for 30 seconds, gradient of 55 °C-60 °C for 90 seconds, 72 °C for 60 seconds and a final extension step of 60 °C for 30 minutes. Each PCR constituted a 12 µl reaction containing 2 µl of genomic DNA and 10 µl of Qiagen multiplex kit, which contained 0.2 µM of both the forward and reverse primer. Optimal conditions were determined when a single, clear PCR product of the expected size (determined using a 100 bp ladder) was obtained in at least three individuals for a particular temperature. Optimal annealing temperatures for each primer pair are reported in Table 2.1. The annealing step of further PCR reactions was conducted at the optimal temperature for the primer pair in question but all other conditions remained as above.

Table 2.1 Optimal T_m for each Y-CAT primer pair.

Locus	Optimal T_m (°C)
DBY7	58
DBY8	58
DBY13	56
DBY14	58
SMCY1	56
SMCY3	56
SMCY11	58
SMCY16	56
SMCY17	56
UTY11	56
ZFY2	56

To determine whether the primers produced a male-specific PCR product each primer pair was tested on at least five male and five female bats. Primers were deemed male-specific if they produced a consistent single PCR product in males but either no PCR product or a PCR product of a different size in females.

2.2.2.1 Sequencing

In order to confirm that the PCR products obtained in males and females were the same sequence, the PCR product of each primer pair from three males and three females was sequenced using the following method: Prior to sequencing each PCR product was cleaned using the ExoSap method with a ratio of 5 µl of PCR product to 1 µl of ExoSapIT. Big dye kit version 1.1 (Applied BioSystems) was chosen for the cycle sequencing reaction as it is optimised for products less than 500 bp in length. Cycle sequencing was performed in the forward and reverse direction. Standard cycle sequencing and ethanol precipitation methods were implemented prior to running samples on an ABI 3730 DNA sequencer. A negative control (ddH_2O) was added at the beginning of the PCR process and was also sequenced for each primer pair to investigate potential contamination of reagents.

2.2.2.2 Cleaning up sequences

ABI trace files were viewed in CodonCode Aligner version 3.0.2 (CodonCode Corp.). The raw trace files were assembled. Any trace files that were of poor quality or unreadable were then discarded. The beginning and end of each sequence within the assembly was viewed by eye for quality. Any portion of this sequence that was unreadable or had a low quality score (below 10) was removed. The sequence was then reassembled. Any position on a sequence that had been denoted with the degenerate base code was viewed by eye and the uncertainty was resolved. This resulted in sequence data from DBY7, DBY8, DBY13, DBY14, and SMCY17. We failed to obtain usable sequence data from the remaining 3 PCR products (SMCY11, SMCY16, UTY11) in males or females and subsequently focussed on the 5 loci for which there were sequences.

2.2.2.3 Sequence alignment

Multiple sequence alignments of the sequences from each locus were made using the Clustal-W algorithm (Thompson *et al.* 1994) in MEGA version 4.1 (Tamura *et al.* 2007). This was done to see if the sequences obtained from males and females were homologous, and if so to identify any differences between them.

2.2.2.4 BLAST

As the sequences obtained from the different individuals at each locus were virtually identical, one sequence from each locus was BLAST searched against the NCBI nucleotide reference collection using the BLASTn algorithm. This was done to identify the species and gene/chromosome location of the closest matches and so to determine if the sequences we obtained were located on the Y-chromosome or if they were located on the X-homologue or some other genomic region. Each sequence was also BLAST searched against the *Myotis lucifugus* genome assembly on Ensembl using the BLASTn algorithm. As the individual sequenced was a female we would expect that if the sequences we obtained were Y-specific they should not match strongly to any region in the *M. lucifugus* assembly.

2.2.2.5 Y-CAT results

Of the eleven primers tested, three failed to amplify in either males or females (SMCY1, SMCY3, ZFY2). The remaining eight primers produced PCR products in both male and female samples, which appeared to be of the expected size in both sexes (Figure 2.2). Sequence information was successfully obtained for males and females for five of the remaining eight primers (Accession numbers: HM596898-HM596902).

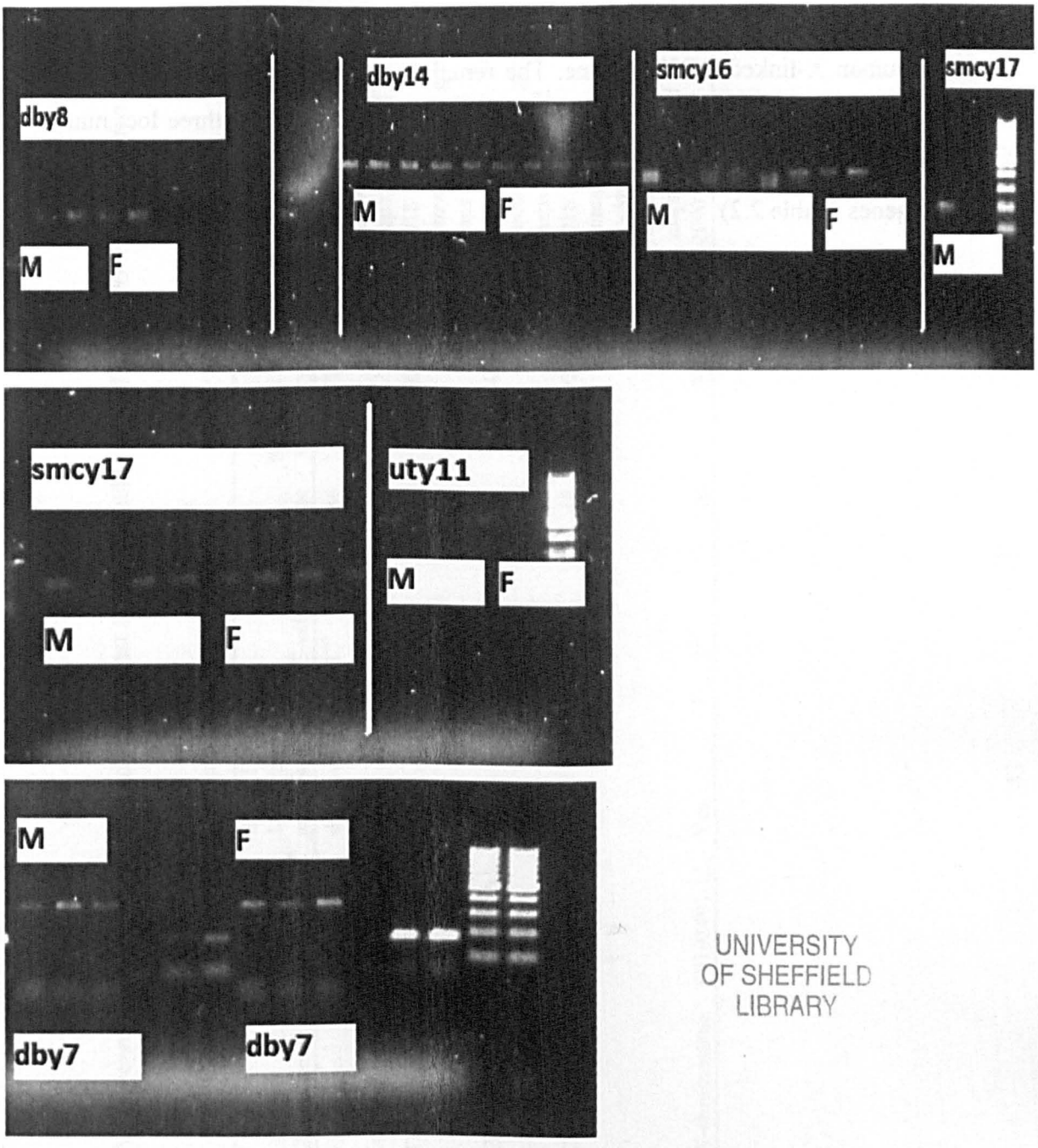


Figure 2.2 Agarose gel of YCAT PCR products. M and F denote male and female samples. Ladder size = 100 bp. Expected fragment sizes: DBY8 = 200 bp, DBY14 = 500 bp, SMCY16 = 450 bp, SMCY17 = 200 bp, UTY11= 550 bp, DBY7 = 400 bp.

Sequence alignments of PCR product from each locus revealed no polymorphism in homologous stretches of sequence either between or within sexes suggesting that the product amplified in males and females was the same.

PCR products from two loci returned the strongest BLAST searches from genes on the Y-chromosome, in the case of DBY 7 from another bat species and in the case of DBY 8 from a phocid seal. DBY 7 did not BLAST strongly to the *M. lucifugus* genome,

however DBY 8 produced a strong match to gene scaffold 2018, which partially maps to the human X-linked DDX3X gene. The remaining three loci returned the strongest BLAST searches from the X-chromosome of humans and pigs. All three loci matched strongly to regions on the *M. lucifugus* genome which partially mapped to human X-linked genes (Table 2.2).

Table 2.2 Closest matches from BLASTn searches of YCAT PCR products against a) NCBI nucleotide reference collection b) *Myotis lucifugus* genome

Locus	Sequence length	BLASTn closest species match	Gene/ Chromosome	BLASTn Accession	BLASTn hit to *M. lucifugus* genome	E-value	% coverage	% similarity	Gene scaffold	Additional Information
DBY 7	315 bp	Frosted sac-winged bat	DBY 7	EF584325.1	short, weak hits	8^{-35}	86	71	NA	NA
DBY 8	118 bp	Weddell seal	DBY	FJ813491.1	Yes	3^{-14}	87	80	2018	Human DDX3X gene maps to scaffold 2018
DBY 13	153 bp	Human	DBX 3	NG 012830.1	Yes	5^{-20}	81	80	2018	Human DDX3X gene maps to scaffold 2018
DBY 14	436 bp	Human	DBX 3	NG 012830.1	Yes	7^{-25}	18	92	2018	Human DDX3X gene maps to scaffold 2018
SMCY 17	151 bp	Pig DNA sequence	X-chromosome	FP102491.12	Yes	2^{-51}	100	90	2090	Human IQSEC2 gene (located on X-chromosome) maps to scaffold 2090

2.2.3 AFLP approach to isolating Y-specific genetic regions

Amplified fragment length polymorphism (AFLP) is a PCR based technique that uses restriction enzymes to selectively amplify restriction fragments from a total digest of genomic DNA (Vos *et al.* 1995). AFLP has the advantage that no prior sequence information is needed and it can therefore be used to identify markers in non-model organisms. In addition, because there are many possible combinations of selective primers that can be used, AFLP has the potential to produce many bands from the species of interest. The presence of SNPs and INDELs throughout the genome of an organism will alter the location of restriction sites, which means that different species and individuals will produce different banding patterns given the same primer combination. Because female bats do not have a Y-chromosome any restriction sites on the Y-chromosome that produce a band of a unique size should only be visible in male samples. We used this logic to identify bands only found in males (and therefore presumed to be located on the Y-chromosome) with a view to isolating the bands. If the bands can be isolated, sequenced and proved to be polymorphic this would provide the basis for Y-specific sequence marker identification in *M. daubentonii*.

2.2.3.1 Samples

We used six male and six female bats for the AFLP reactions.

2.2.3.2 AFLP procedure

We followed the AFLP procedure as outlined in Whitlock *et al.* (2008). Each DNA sample was digested using *Eco*RI and *Mse*RI enzymes and ligated to the *Eco* and *Mse* specific adaptors using T4 ligase. *Eco*RI and *Mse*RI pre-selective primers were used and a number of selective primer combinations (with fluorescently labelled *Eco*RI selective primer) were trialled. AFLP fragments were separated on a 3730 DNA sequencer and banding patterns were viewed using Genemapper version 3.9.

2.2.3.3 Identification of male-specific bands

For each primer combination male and female banding patterns were viewed alongside one another. If a band of a given size was consistently present in male but absent in

female samples it was presumed to be located on the Y-chromosome. Of the bands generated by this method, five bands from four different primer combinations were present in males only.

2.2.3.4 Isolation and characterisation of male-specific bands

In order to physically isolate the bands of interest the AFLP procedure was repeated as outlined above using a larger reaction volume of 50 µl. The PCR products from males and females were run on an ethidium bromide stained agarose gel from which the bands of interest could be excised. In order to obtain optimal separation of fragments the gel was 2.5% agarose concentration and was run for a minimum of 3 hours at 75 V. The gels were visualised on a UV box. Due to the number of bands present in each primer combination it was not possible to pinpoint the target band so a small cluster of bands with a 20 bp margin either side of the target product size was isolated. The cluster of bands was excised using a scalpel and each gel slice subsequently purified using the Promega gel purification kit. The purified products were run on a 1.5% syber safe gel to confirm presence and size. As each purified product contained bands other than the target sequence it was necessary to clone the products in order to isolate the band of interest. The pGEM-T easy kit was used for cloning. Following transformation each positive colony (white coloration indicating the uptake of PCR product by the plasmid) was picked into a well containing 100 µl of LB broth + 20% glycerol. The plate was then incubated overnight at 37 °C to allow colony growth. M13 PCR primers were used to amplify each plasmid using the following reaction conditions: 95 °C for 15 minutes, followed by 35 cycles of 94 °C for 30 seconds, 58 °C for 90 seconds, 72 °C for 60 seconds and a final extension step of 60 °C for 30 minutes. Each PCR constituted a 2 µl reaction containing 1 µl of genomic DNA and 1 µl of Qiagen multiplex kit, which contained 0.2 µM of both the forward and reverse primer. The resultant PCR products were run out on a 1.5% syber safe gel to confirm successful amplification and to quantify the concentration of positive PCR products.

2.2.3.5 AFLP results

Four selective primer combinations produced 102 bands, 5 of which were consistently found in males only (See table 2.3 for details). The 5 bands represent 4.9% of the total number of bands and this was roughly what we expected given the size of *Myotis* Y-

chromosomes relative to the rest of the genome (Bickham *et al.* 1986). However, we were unable to isolate any of the bands and determine the sequence composition.

Table 2.3 AFLP bands found only in male *M. daubentonii* samples

Pre selective primers	Selective Primers	Band Size
Eco-Mse	ATC CGG	253 bp
Eco-Mse	ATC CGG	126 bp
Eco-Mse	TCT CCG	282 bp
Eco-Mse	TCT CGC	300 bp
Eco-Mse	TCT CGT	200 bp

2.2.4 *De Novo* sequencing of a flow-sorted Y-chromosome

Due to the difficulty in characterising Y-specific markers in non-model organisms an alternative approach is to physically isolate the Y-chromosome from the species of interest and then sequence the isolated chromosome (Greminger *et al.* 2010). Next generation sequencing technologies increase the feasibility of this option as they can employ degenerate primers for *de novo* sequencing to produce hundreds of reads from the chromosome of interest (Schendure 2008). The reads may then be assembled into contigs and searched for SNPs (if several individuals have been sequenced) or individual reads may be searched for microsatellite motifs.

We employed this approach to sequence a flow sorted Y-chromosome from the bat *Myotis myotis*. As we only had chromosome PCR product from one individual we were unable to use the SNP detection approach and instead searched the sequence reads for microsatellite repeats.

2.2.4.1 Chromosome preparation and Degenerate Oligo-nucleotide (DOP) PCR

DOP PCR product from the amplification of a flow sorted Y-chromosome from the bat species *Myotis myotis* was obtained from the Sanger Institute, Cambridge. The metaphase chromosome was obtained from *M. myotis* fibroblast cell cultures and sorted using a fluorescence activated cell sorter as outlined in Volleth *et al.* (2002). Because *M. myotis* is presumed to have a karyotype that is representative of the ancestral type in

the Vespertilionidae it was hoped that the chromosomal architecture would be similar to that of *M. daubentonii.*

The flow sorted Y-chromosome was amplified using degenerate oligo-nucleotide (DOP) PCR. DOP PCR is a species-independent method that employs nucleotides of a partially degenerate sequence in combination with a low primer annealing temperature to ensure priming from evenly dispersed sites across a genome or chromosome (Telenius *et al.* 1992). The primer used was the 6MW primer (5'-CCG ACT CGA GNN NNN NAT GTG G- 3') and conditions were as follows: master mix containing 23 µl ddH_2O, 5 µl buffer, 5 µl 6MW (20 µM) primer, 4 µl dNTP, 2.5 µl 50 mM $MgCl_2$, 1 µl Bioline Taq, 5 µl DOP PCR product. PCR using tetrad thermal cycler and the following program: 94 °C for 3 minutes followed by 25 cycles of 1. 94 °C 1 minute 2. 62 °C 1.5 minutes 3. 72 °C 2 minutes. Then 72 °C for 10 minutes and a final extension step of 4 °C.

Following amplification, 3 µl of the PCR product was run on a 1% syber safe stained agarose gel at 100 V for 30 minutes with the control and 100 bp ladder. If the PCR is successful a clearly visible smear between 2 kb and 0.2 kb should be observed. In order to obtain sufficiently pure and concentrated product for 454 sequencing the remaining 45 µl of PCR product was purified using the QIAGEN PCR purification kit and eluted in a final volume of 20 µl. Purity of the product was confirmed using the Nanodrop method.

Sequencing of the PCR product using the Roche 454 method was carried out by the Genepool unit at the University of Edinburgh. For an overview of the methodology used see (http://genepool.bio.ed.ac.uk/roche/index.html). Sequencing was restricted to $1/8^{th}$ of a 454 plate as it was thought this would generate sufficient sequence information for the design of primers. Sequence data were returned in the form of short unassembled reads and contigs assembled *de novo* using the 454 assembler Newbler.

2.2.4.2 Screening for microsatellites

Two separate FASTA files were created; one containing the raw assembly data and one containing assembled contigs (http://xyala.cap.ed.ac.uk/Gene_Pool/Butlin_2008044/). A search for repeat regions among the sequences was conducted using SPUTNIK (www.cbib.u-bordeaux2.fr/pise/sputnik.html). A minimum repeat length of 20 bp was

included in the search criteria and only perfect microsatellite repeats were included to avoid PCR amplification problems.

As the 454 sequence reads contain a certain level of repetition microsatellite regions found more than once were discarded. Microsatellites with less than 30 bp of flanking DNA either side of the repeat were also discarded.

The remaining microsatellite repeats were then BLAST searched against the NCBI nucleotide reference collection using the BLASTn algorithm to identify if any of them were a match for Y-chromosome loci/genes. A BLASTn search of the repeat sequences against the *M. lucifugus* genome was also performed. If any strong hits were obtained it was assumed that the sequences were either not Y-specific or had homologues else where in the genome. Such repeats were than discarded.

2.2.4.3 Primer design

Primer3 (Rozen & Skaletsky 2000) was used to design primers from the remaining sequences. In each case the repeat region was highlighted and the following criteria used to select potential primer pairs: Primer size between 18-27 bp, T_m between 55-65 ^{o}C, Max T_m difference between pair 0.5 ^{o}C, CG clamp.

If the primer binding site selected by Primer3 was located on a region containing repeats this region was excluded so that a more suitable binding site could be found. If Primer3 failed to find any primer binding sites in the flanking sequence the CG clamp was removed and the maximum T_m difference was increased to 1.5 ^{o}C. If Primer3 still failed to find any binding sites the repeat region was dropped from further investigation. This resulted in 48 primer pairs designed to amplify the potential microsatellite regions (For details see Table 2.4).

Table 2.4. Details of primers designed to amplify microsatellites detected from sequenced *Myotis* Y-chromosome data

Locus Name	Forward	Reverse	Repeat Type	T m	Product size (bp)
C06	CCTGCAAGCAAACTTAGGAAG	TGAATTGAGACTGCTTTGTGG	AATC	57	102
C448	AAATGTGTCCTGCTGAAATGG	GGACCTAACTCCCTCATTTAACG	TG	58	103
C865	ATAATGAGCTCACTGGCATTTG	CGAGTAGCGTATGTGGTAGGG	AAAT	57	104
C1046	GTGGTTTGGCTGAAACAATG	AGGAAAGTACATGCCCCATC	TTGG	57	113
C680	GATACCCTGATTCAGAAAAGCTC	AGATTGCACCACGGCACT	TTTTG	57	107
C689	TCGAGTCTGTCATGTGGAATG	AGTGAGTTGGAGCCATCCAT	GAAT	58	115
C271	TCGAGATATATATGTGGGGAGTAGTG	ACAAGCCCAGAGACCTAGGAG	GGGCA	58	108
C273	TGATGATATATACCCATAAGCACACAG	TGGCCATGTGCATCTTTATG	CA	58	103
C536	TGTGGAAGTTGAAAATGAAGG	GTGGAGGCAACCAGTCAATA	GA	56	151
C1074	GCTTTCTTCATGTCAGCTAAGGA	TGTCAGACCATGTGTCTCAGTTT	AAAC	58	183
FY1	GGCGGGTGTGAAAGAGAG	TTCTCCAGACACTCCACTGC	AC	57	121
FY2	CGACTCGAGGTTTAGATGTTGG	TTTGCTAAAAGCGTCCTCAATC	AC	59	150
FY3	GGGATGGAATAGAGGTCCAAG	ACCAGGCACATGGTAGGC	AAAC	58	164
FY4	AAGAAAATCATTGGCGAGAGG	GACTTGGCCAGGACAGTCAG	AAAG	59	113
FY7	CCGACTCGAGCACATCAT	GGCTGGATAGGAGTCTTCTGA	AC	56	134
FY8	GTATGATGTGCGGTGGGAAC	GACTCGAGCCTTGAATGTGG	AG	59	192
FY9	GCTCAGAGGGAAGGCAAAG	CATCATGACTAAGCTTTCAACAGG	CG	58	189
FY11	TGTATCAGGTAAGAGTATGTTGTTTCG	TCCTGTAGGGTGGTAGCTTTTC	AC	58	154
FY12	CTTCACATATGGGTGCATTG	ACCTAACTCCCTCATTTAACGAC	AC	57	155
FY13	AGTTTGGGGGCATATCTGG	CTTCTCCAGACACTCCACTGC	AC	58	106
FY14	AAAAGTCGCCAGTTCAACTCC	GCCTCGGCAGCATAACATAC	AG	59	156
FY15	TGATGCAATAAGGCAAGGAAG	CAATCGATGTGTTTCCCTCAC	AAAAC	58	151
FY16	TGCTGGACTATAATGGAACTAGAGG	GCCAAAATCTCATGCACTTTC	AG	58	121
FY17	AATGGTGGCTGCTACTACTATGAC	AGTGGTTGTGGAGGTGGTG	ACACT	57	138
FY18	ACTCGAGCGTCTGATGTGG	ATCGCTATGTGGTCCGACTC	AG	58	116
FY19	TGCCGAAACGGTCAGATG	TCATGGGATTCTTATTCCTTTCG	AAAAT	60	137
FY20	AAATGTGTCCTGCTGAAATGG	GGACCTAACTCCCTCATTTAACG	AC	58	100
FY21	ACTAGGAAACAGGCAGATGAGG	AGTCGGACCACATTGAAAGC	ACCCC	58	192
FY22	GCCATGTGCATCTTTATGTG	GTAGACAGATGTGCGTATTTGC	AC	57	140
FY23	TCGAGTCTGTCATGTGGAATG	GTCTGGGCATTGAAGCAAG	AATG	58	154
FY24	CAGCGAAAGACTGACACACC	AAGGGTACGGACGTAGTTGG	AC	57	187
FY25	CTCCCATGGAGAAGAAAGTCC	CGACTCGAGATGGGATGTG	AC	58	117
FY26	AGTACTGGGTGTTGGACGTG	TTCTCCCTATTTAACCAATGAGC	AGCC	57	175
FY27	TGGCAGATCTTTGATTGCTG	ACTCGAGACACGGATGTGG	AAAG	58	134
FY28	GTGGCTTGAGAATTTCATAGGTC	GAAGGAGATGGGGAGGTTG	AC	57	104
FY29	GATGTGGCAGAGGATGCAG	AACATCACTTCCCTGCCATAAG	AG	58	154
FY30	TGGCAAGTAAGGAGACTCTGG	AGCAGGCACAGAAAGATTCAC	AAAT	58	160
FY31	TCCCTGTTAATTCTGGTGTGG	CGGGAGGTTCAACAGGTG	AG	58	182
FY32	ACCTATACCACGTGCATCACAC	AAAGTAGGTCTGAGAGTGCATGG	AC	58	122
FY33	GTGGGTTCCCACAAGCTAAG	ATTAAAGCACCAGGCTGAGG	AAAAC	58	155
FY34	CATTTGCATGTCTCTAATTCTCAGTG	AGCACATCCCACCCCATC	AC	60	107
FY35	TCGAGATTGCAATGTGGTTG	GCTGGACTATCATGGAACTAGAG	AG	59	139

		G			
FY36	CAAAGCCTCAGAGGAGGAAAG	GAGCCCGATGTGGTATGTTC	AAAT	58	156
FY37	GCCAGCTCCAGGTTTAGTCTC	TCCTCCAGAGGAGCAGAATG	ACCCC	58	104
FY38	GCAAAGTATGTGGTCCGACTC	ATGCAGGGGGTACGAGAAG	AG	58	136
FY39	CGACTCGAGCAGTCAATGTG	ATGCCCACGTAAACCCCTAC	AAAT	59	152
FY40	TGAATTGAGACTGCTTTGTGG	CCTGCAAGCAAACTTAGGAAG	ATTTG	59	102
FY41	ACTCGAGGGAAGTATGTGGAC	CTACCAGCCCCTCGAGAC	GT	58	151

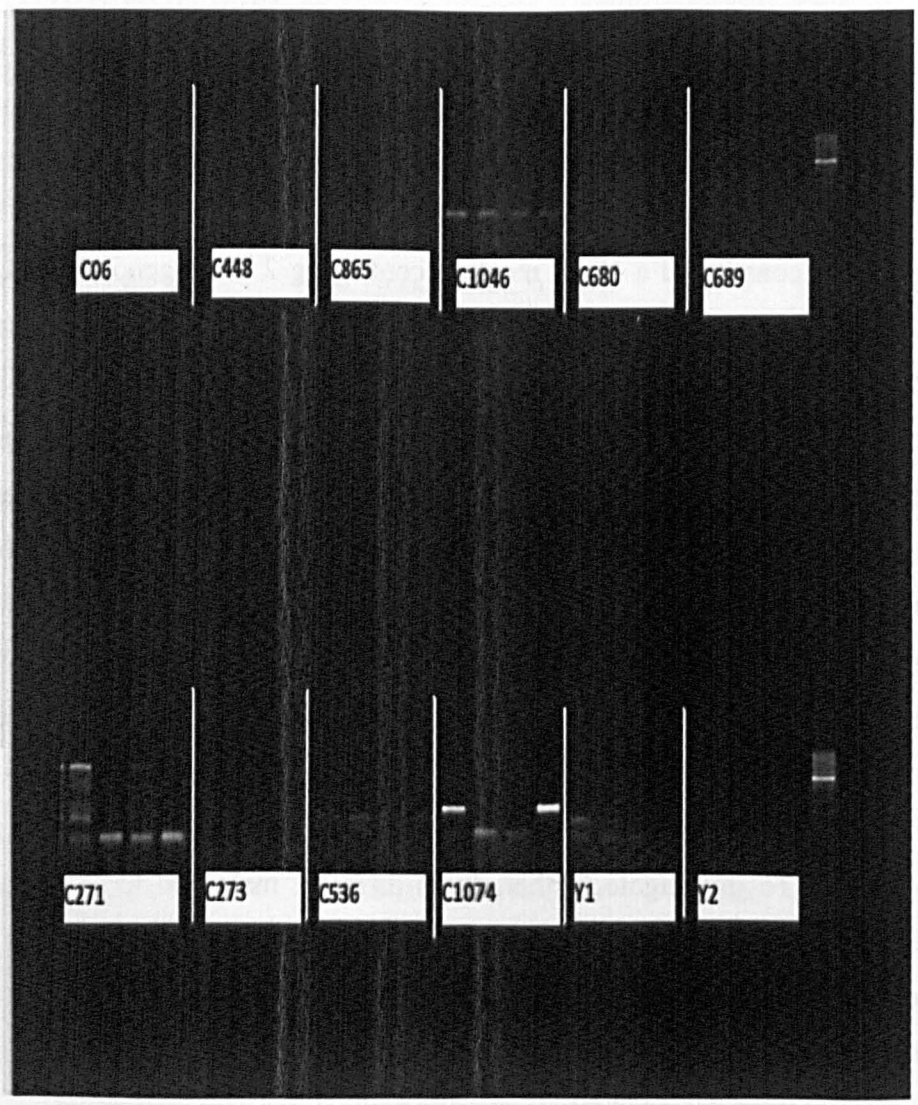
2.2.4.4. Primer testing

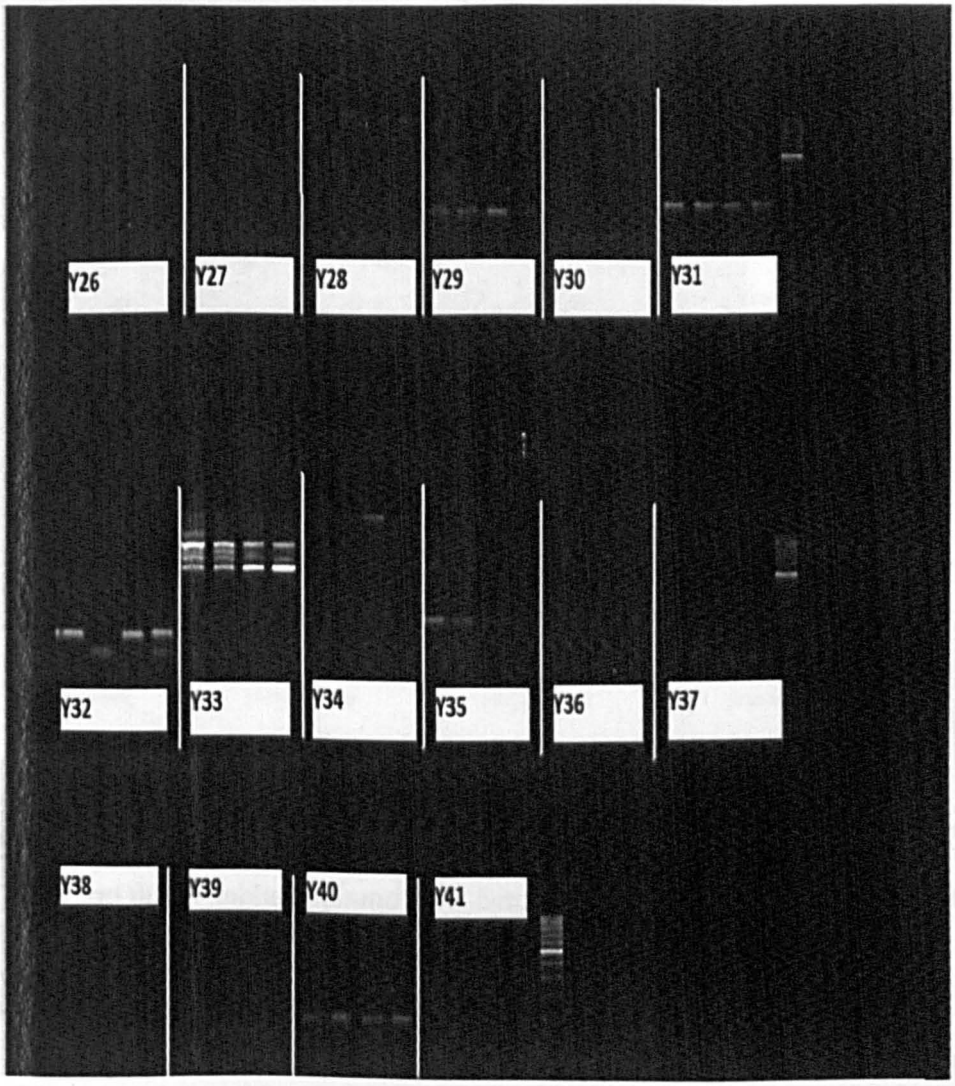
Initial primer testing involved PCR amplification of two male and two female *M. daubentonii* samples using a gradient programme with the following conditions: 95 °C for 15 minutes, followed by 35 cycles of 94 °C for 30 seconds, gradient of 55 °C-60 °C for 90 seconds, 72 °C for 60 seconds and a final extension step of 60 °C for 10 minutes. Each PCR constituted a 12 µl reaction containing 2 µl of genomic DNA and 10 µl of Qiagen multiplex kit, which contained 0.2 µM of both the forward and reverse primer. Optimal conditions were determined when a single, clear PCR product of the expected size (determined using a 100 bp ladder) was obtained in at least three individuals for a particular temperature. Products were visualised on a 1.5% syber safe gel. If a product of the expected size was present in males but absent in females it was taken as Y-specific. If a product of a similar size was also present in females further investigation was required. Out of the 48 primer combinations tested 15 produced no product in males or females, 11 produced a product of the same size in at least one male and one female, 14 produced a primer dimer and 8 produced a smear of product from 1 kb-100 bp in at least one male and one female (Table 2.5). Thus no product could be deemed Y-specific. To investigate further, three different male and three different female *M. daubentonii* samples were tested using those primers that produced a product. In addition the same primers were used to amplify ddH_2O (to rule out contamination of primers or reagents), one male and one female *M. myotis* sample (as this was the species the primers were designed from) and one male *Homo sapiens* sample.

As before PCR products were obtained in males and females, this was also the case in *M. myotis*. No product was obtained when water was used which ruled out contamination of reagents with male DNA. PCR products were obtained using human DNA, however they were not the same size as those obtained in bats.

2.2.4.5 PCR results

Many of the 48 microsatellite primers designed from the sequence data produced either no product or non-specific stutters (Figure 2.3). Eleven microsatellite primers did produce a clear band of approximately the expected size in males and females (Table 2.5). Of these primers, sequence data were obtained for locus C1074 only.




Y26
Y27
Y28
Y29
Y30
Y31
Y32
Y33
Y34
Y35
Y36
Y37
Y38
Y39
Y40
Y41

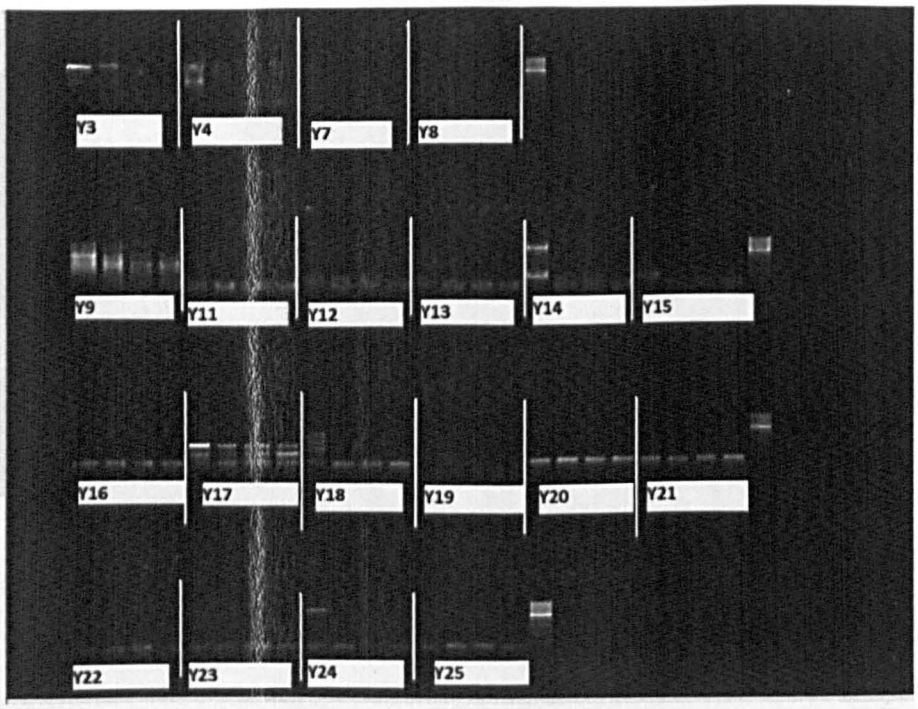


Figure 2.3 PCR products from putative 'Y' microsatellites. For each product the first two samples are male and the second two female. Ladder = 100 bp.

Table 2.5 PCR results from trials of putative 'Y' microsatellite primers in male and female *M. daubentonii* samples

Locus	Male1	Male2	Female1	Female2
C06	band	band	band	band
C448	band	band	band	band
C865	fail	fail	fail	fail
C1046	band	band	band	band
C680	fail	fail	fail	fail
C689	fail	fail	fail	fail
C271	stutter	stutter	stutter	stutter
C273	fail	fail	fail	fail
C536	stutter	stutter	stutter	stutter
C1074	2 bands	2 bands	2 bands	2 bands
FY1	2 bands	2 bands	2 bands	2 bands
FY2	stutter	band	band	fail

FY3	stutter	stutter	stutter	stutter
FY4	stutter	stutter	stutter	stutter
FY7	stutter	fail	fail	fail
FY8	fail	fail	fail	fail
FY9	stutter	stutter	stutter	stutter
FY11	primer dimer	primer dimer	primer dimer	primer dimer
FY12	band	band	band	primer dimer
FY13	stutter	primer dimer	primer dimer	primer dimer
FY14	stutter	band	band	primer dimer
FY15	band	primer dimer	primer dimer	primer dimer
FY16	band	band	band	band
FY17	stutter	stutter	stutter	stutter
FY18	stutter	band	band	band
FY19	fail	fail	fail	fail
FY20	primer dimer	primer dimer	primer dimer	primer dimer
FY21	primer dimer	primer dimer	primer dimer	primer dimer
FY22	primer dimer	primer dimer	primer dimer	primer dimer
FY23	stutter	primer dimer	primer dimer	primer dimer
FY24	stutter	primer dimer	primer dimer	primer dimer
FY25	band	primer dimer	primer dimer	primer dimer
FY26	fail	fail	fail	fail
FY27	fail	fail	fail	fail
FY28	fail	fail	fail	fail
FY29	band	band	band	band
FY30	fail	fail	fail	fail
FY31	band	band	band	band
FY32	band	fail	band	band
FY33	stutter	stutter	stutter	stutter
FY34	fail	stutter	stutter	stutter
FY35	stutter	stutter	stutter	fail
FY36	fail	fail	fail	fail
FY37	fail	primer dimer	primer dimer	primer dimer
FY38	primer dimer	primer dimer	primer dimer	primer dimer
FY39	primer dimer	primer dimer	primer dimer	primer dimer
FY40	band	band	band	band
FY41	fail	fail	fail	fail

2.2.4.6 Sequencing

In order to investigate whether the PCR products obtained in males and females were the same, whether the primers were amplifying the target sequence and whether they originated from *M. daubentonii*, or some other contaminant, PCR products from males and females for locus C1074 were purified and sequenced. Purification and sequencing were carried out as previously described in section 2.2.2.1.

The sequences were aligned against the expected sequence using the Clustal-W algorithm in MEGA version 4.1. Aligned sequences were then BLAST searched against the NCBI nucleotide reference collection using the BLASTn algorithm to identify closest matches in terms of species and chromosome.

2.2.4.7 C1074 sequencing results

The PCR product obtained from locus C1074 was sequenced in six male and six female *M. daubentonii* individuals. Aligning the sequences with the expected sequence at this locus showed that the 12 PCR products were very similar (although the length of clean sequence obtained was variable among samples) but that they were not the PCR product we expected. This suggested that the primers designed were binding to a region other than the target sequence. A BLAST search of the sequences against the NCBI nucleotide reference collection showed that the sequences did not match to the Y-chromosome of any other species (Table 2.6)

Table 2.6 BLAST search of C1074 from male and female *M. daubentonii* samples

Sample	Closest species match	Chromosome	Accession	E-value	% coverage	% similarity
MdM1	*H. sapiens*	13	NG007300.2	0.15	43	75
MdM2	*Mus* species	X	AL669976.11	0.14	35	80
MdM3	*Mus* species	7	AC090122.26	8^{-4}	64	74
MdM4	*Mus* species	2	BX284685.8	4^{-10}	43	82
MdM5	*Paa spinosa*	mt control region	EU597430.1	5^{-8}	61	75
MdM6	*Mus* species	7	AC124545.4	5^{-7}	68	76
MdF1	*Ricinus communis*	mRNA	XM002524655.1	0.16	26	81
MdF2	*Mus* species	4	AL807383.28	3^{-4}	31	85
MdF3	*H. sapiens*	6	AL590395.4	2^{-5}	49	80
MdF4	*Mus* species	10	AC137981.2	2^{-7}	42	84
MdF5	*Mus* species	7	AC124545.4	5^{-6}	71	76
MdF6	*Mus* species	7	AC124545.4	2^{-6}	62	75

2.2.4.8 Detection of non-microsatellite repeats

As the Y-chromosome is thought to contain a large number of repetitive elements which may also be found throughout the genome of the species of interest, the 48 targeted amplification sequences (see Table 2.4) were searched for repeat elements using the Repeat Masker web server (http://www.repeatmasker.org). Default settings were used with the following exceptions: DNA source changed to: Mammal (other than below) and satellites/simple repeats were excluded from masking.

2.2.4.9 Repeat search results

The repeat masker search identified 12 microsatellite loci that contained repeats other than the target microsatellites. Of these repeats five were SINE elements and three were LINE elements which may be found throughout the genome and possess the property of retro-transposition (Table 2.7).

Table 2.7 Repeat types located in the microsatellite loci other than the target microsatellites

Locus	Score	Matching repeat	Repeat Type
C1074	266	MER94B	DNA/hAT-Blackjack
C271	566	MST-int	LTR/ERVL-MaLR
C536	256	Ves2_ML	SINE/tRNA
C680	343	Alu	SINE/Alu
FY11	315	nhAT4a_ML	DNA/hAT-Charlie
FY11	362	MER112	DNA/hAT-Charlie
FY14	424	Ves2_ML	SINE/tRNA
FY15	350	Ves2_ML	SINE/tRNA
FY19	26	AT_rich	Low complexity
FY3	302	L2c	LINE/L2
FY30	321	Lx8b	LINE/L1
FY36	184	L2b	LINE/L2
FY37	400	B3A	SINE/B2

2.3 Discussion

This chapter describes three different approaches made to characterise and isolate Y-chromosome markers. No approach was successful in identifying a male-specific single-copy marker. The unusual genetic architecture of the Y-chromosome probably explains why we were unsuccessful, but for each method the specific reasons will vary and here they will be examined in turn.

2.3.1 YCATs

The YCAT primers were taken from Hellborg & Ellegren (2003) as they were species degenerate Y-chromosome primers, which had been demonstrated to produce a male-specific product in *Myotis daubentonii* previously (Hellborg & Ellegren 2003). The rationale behind the design of the YCATs is that homologous exonic Y-linked sequences are aligned from two distantly related species and if there is high sequence identity then primers can be designed to amplify the adjacent intronic region. But it should be noted that the Y-linked genes chosen all have X-linked homologues (in fact they may be termed X degenerate sequences) that have a common ancestor from the proto-sex chromosomes and share a high percentage of sequence identity in many species (Skaletsky *et al.* 2003). In some cases the Y-linked homologue is a pseudo-gene with similarity to the X-linked version. However, the gene families from which the YCAT primers were developed in this study (DBY, SMCY, UTY, ZFY) are transcribed into functional genes, which produce very similar protein isoforms whether they are X- or Y-linked and are therefore likely to share a very high degree of sequence similarity (Skaletsky *et al.* 2003). Consequently it is likely that the YCAT primers will bind not only to the Y-linked exons but also (and potentially preferentially) to their X-linked counterparts. This explanation was supported by BLAST searches of the sequences obtained from five YCAT PCR products against the NCBI nucleotide reference collection and the *Myotis lucifugus* genome. Three PCR products produced the strongest hit to X-linked genes from other species in the general BLAST and four products hit strongly to gene scaffolds on the *M. lucifugus* genome onto which human X-linked genes had previously been mapped. Only one sequence (DBY7) showed strong evidence of Y-linkage (strongest hit to DBY7 of another bat species and no strong hits to the *M. lucifugus* genome) suggesting it may warrant further investigation. However, in this study PCR products were also produced in females using this primer and the sequence identity between males and females was high. It is therefore possible that in *M. daubentonii* the DBY/DBX7 genes have not diverged sufficiently since the suppression of recombination, and that high sequence identity allows binding of DBY7 primers to the X-chromosome. It is also possible that, in *M. daubentonii* DBY/DBX are still recombining in both sexes, which would render these loci not male-specific. Indeed gene conversion between X and Y degenerate sequences in the MSY has been observed in humans and the *Felidae* (Pecon-Slattery *et al.* 2000, Rosser *et al.* 2009).

Given that Hellborg & Ellegren (2003) reported a single-copy, male-specific PCR product for 11 of their YCAT primers in *M. daubentonii* it is perhaps surprising that we could not repeat their results in the same species, despite testing numerous individuals. A possible explanation is contamination of either reagents or female samples with male DNA. Contamination of reagents can be ruled out as negative controls (water with all other reagents) failed to produce any PCR product. Although contamination of female samples cannot be entirely discounted it is unlikely as many females were tested. We do not know how many *M. daubentonii* individuals Hellborg & Ellegren (2003) tested their primers on, although their paper states that in all species at least one male and one female sample were tested. If only one individual of each sex was used then it is possible erroneous results could be obtained, potentially through mis-sexing an individual or contamination. Other studies have also reported negative amplification or non-male-specificity in YCAT primers (Luo *et al.* 2007, Tringali *et al.* 2008). Our lack of success may also relate to the evolutionary distance of *M. daubentonii* from the species from which the YCAT primers were designed (largely humans and mice). For example Erler *et al.* (2004) demonstrated that Y-specific amplification success in primates of primers designed from *Homo sapiens* was negatively correlated with divergence from the human lineage.

Given our results it would seem sensible, unless primers are unequivocally male-specific, to consider re-designing the YCAT primers based on male and female sequence data from the species in question. This would allow Y-specificity in primer binding and ensure a higher likelihood of obtaining a single-copy, male-specific product. Such an approach was adopted by Brandlii *et al.* (2005), who redesigned YCAT primers for use in the greater white-toothed shrew.

2.3.2 AFLP

Although we were able to identify potential Y-specific bands using an AFLP approach, technical difficulties prevented the physical isolation and hence sequence characterisation of these bands. AFLP has proved a successful method to isolate sex-specific bands in a number of species including plants (Hardenack *et al.* 2002, Peil *et al.* 2003), birds (Griffiths & Orr 1999), fish (Ezaz *et al.* 2004, Griffiths 2000) and reptiles (Quinn *et al.* 2007), but has sometimes been unsuccessful (Gao *et al.* 2010, Sriphairoj

et al. 2007, Wuertz *et al.* 2006). The technical difficulties associated with physically isolating the band of interest have been noted (Griffiths & Orr 1999). Potential mechanisms to overcome this include using acrylamide gels that allow better separation of fragments than the agarose gels used in this study, or using a cloning approach to isolate the band of interest from the selective PCR product (Quinn *et al.* 2009). However, these methods still require considerable technical expertise, expense and, in the case of acrylamide, exposure to hazardous substances.

2.3.3 454 Sequencing of a flow sorted Y-chromosome

When working on a non-model organism it is often the case that very little sequence information from the Y-chromosome is available. A possible strategy to overcome this is to physically isolate parts of, or a whole, Y-chromosome and obtain sequence information from this (Petit *et al.* 2002). We were able to obtain a WGA library from a flow sorted chromosome from a species in the same genus as *M. daubentonii* and use 454 sequencing technology to amplify short stretches of degenerately amplified sequence from this chromosome. Although we were able to characterise microsatellite repeats and design primers from the 454 data, when the primers were tested they did not produce a male-specific product and in many cases appeared to be amplifying a region other than the target.

Limited success in designing male-specific primers from physically isolated Y-chromosomes has also been reported in other species (Greminger 2007) and a number of unsuccessful unpublished attempts have also been made (see Greminger *et al.* 2010). Luo *et al.* (2007) were able to successfully characterise male-specific microsatellites from the Y-chromosome in the *Felidae* by partial sequencing of a cosmid library and the formation of a microsatellite enriched library from whole genome amplification of the flow sorted Y-chromosome. However despite identifying 77 microsatellites using this method only 24 were male-specific and none were polymorphic.

Our lack of success could be due to a number of reasons, including lack of Y-specificity of the flow sorted Y-chromosome and homology of isolated microsatellites to other regions on the Y-chromosome. During the flow cytometry process it is possible that contamination from autosomes could occur. If this is the case and we have autosomal material in the flow sorted Y-chromosome then DOP PCR could amplify large amounts

of autosomal material and thus microsatellites identified may not be located on the Y-chromosome (Jiang *et al.* 2005). Bacterial contamination is also possible during this process and presents similar problems.

Even if the 454 reads were from the *Myotis* Y-chromosome the unique architecture of the Y-chromosome is such that they may be located in highly repetitive heterochromatic regions that are not single-copy. The vast majority of mammalian Y-chromosomes are composed of heterochromatic sequences which are highly repetitive and even within the euchromatic region there are ampliconic sequences which exhibit marked similarity to other sequences on the MSY (Skaletsky *et al.* 2003). Consequently the primers designed may bind non-specifically to many regions on the Y-chromosome. In a number of cases we observed stutters of PCR product and this could reflect the non-specific binding of our Y microsatellite primers. In addition the Y-chromosome carries a high density of retro-elements which may be found throughout the genome. This has been observed in other species including the greater white-toothed shrew (Lawson-Handley & Perrin 2006), Cabrera's vole (Marchal *et al.* 2008) and a number of carnivores (Pecon-Slattery *et al.* 2000). It is also supported by a search of the sequences from which our primers were designed that detected a number of retro-elements. Finally the stretches of sequences we obtained could be located on the pseudo-autosomal region (PAR) of the Y-chromosome that is still undergoing recombination with the X-chromosome. The PAR of the human Y-chromosome is small, representing roughly 5% of the chromosome and this makes isolating sequences from this region unlikely. However, it is possible that the PAR in *M. daubentonii* is much larger and sequences isolated from this region are obviously not male-specific.

2.3.4 Future directions

Results from the methodologies used in this thesis suggest that a promising line of attack to identify single-copy, male-specific markers would be to redesign the YCAT primers to make them more specific to the Y-chromosome of the species in question. The sequences we obtained from males and females were near identical, which would prevent us doing this from the sequence data already available, but a technique such as genome walking that allows stretches of sequence away from the primer sites to be characterised may allow this approach to work (Rampias *et al.* 2009, Reddy *et al.*

2008). However, even if such an approach were successful in identifying male-specific, single-copy markers levels of polymorphism may still be low.

Although our 454 sequenced Y-chromosome was unable to produce anything useful NGS sequencing technologies are constantly improving, allowing longer reads to be obtained (Hudson 2008). Longer reads provide more sequence data to work with and may allow us to determine with greater ease the likelihood of any of the reads being Y-specific via a general nucleotide BLAST search. It would also be useful to sequence flow sorted Y-chromosomes from several individuals, as obtaining the same sequences from different individuals increases our confidence that they are from the Y-chromosome and not a product of contamination. This would also allow the alignment of contigs from several individuals and this could identify SNPs, or repeat length variation from which primers could then be designed.

A search of the sequence reads obtained from the flow sorted *M. myotis* Y chromosome revealed that a number of reads from which primers were designed contained repetitive elements (see Table 2.7). Early population genetic work on the human Y-chromosome determined the frequency of Alu repeats on the Y-chromosome to study the phylogenetic relationships of male lineages (Hammer 1994) and this is an approach which may also be applicable using repetitive elements on bat Y-chromosomes.

Finally, it is worth mentioning that although I was unable to successfully isolate sequence information from potentially male-specific AFLP bands, the success rate in characterising sex-specific markers from a wide range of organisms (see section 2.3.2) using this method suggests that it remains a promising approach.

2.4 References

Andres O, Ronn AC, Bonhomme M *et al.* (2008) A microarray system for Y-chromosomal and mitochondrial single nucleotide polymorphism analysis in chimpanzee populations. *Molecular Ecology Resources* **8** 529-539.

Bachtrog D (2005) Sex chromosome evolution: Molecular aspects of Y-chromosome degeneration in *Drosophila. Genome Research* **15** 1393-1401.

Balloux F (2010) The worm in the fruit of the mitochondrial DNA tree. *Heredity* **104** 419-420.

Bickham J, McBee K, Schlitter DA (1986) Chromosomal variation among seven species of *Myotis. Journal of Mammology* **67** 746-750.

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

Casanova M, Leroy P, Boucekkine C (1985) A human y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science* **230** 1403-1406.

Charlesworth B (1994) The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genetic Research* **63** 213-227.

Charlesworth B (1996) The evolution of chromosomal sex determination and dosage compensation. *Current Biology* **6** 149-162.

Charlesworth B, Charlesworth D (2000) The degeneration of the Y chromosome. *Philosophical transactions of the Royal Society of London Series B* **355** 1563-1572.

Dupanloup I, Pereira L, Bertorelle G *et al.* (2003) A recent shift from polygyny to monogamy in humans is suggested by the analysis of worldwide Y-chromosome diversity. *Molecular Biology & Evolution* **57** 85-97.

Erler A, Stoneking, Kayser M (2004) Development of Y-chromosomal microsatellite markers for nonhuman primates. *Molecular Ecology* **13** 2921-2930.

Ezaz MT Harvey SC, Boonphakdee C (2004) Isolation and Physical Mapping of Sex-Linked AFLP Markers in Nile Tilapia (*Oreochromis niloticus* L.). *Marine Biotechnology* **6** 435-445.

Filatov D (2004) Substitution rates in a new *Silene latifolia* sex-linked gene, SissX/Y. *Molecular Biology & Evolution* **22** 402-408.

Gao ZX, Wang HP, Yao H *et al.* (2010) No sex-specific markers detected in bluegill sunfish Lepomis macrochirus by AFLP. *Journal of Fish Biology* **76** 408-414.

Gill P, Brenner C, Brinkman B *et al.* (2001) DNA commission of the international society of forensic genetics: recommendations on forensic analysis using Y-chromosome STRs. *International Journal of Legal Medicine* **114** 305-309.

Gordo I, Charlesworth, B. (2001) Genetic linkage and molecular evolution. *Current Biology* **11** 1379-1387.

Graves JAM, Wakefield MJ, Toder R (1998) The origin and evolution of the pseudoautosomal regions of human sex chromosomes. *Human Molecular Genetics* **7** 1991-1996.

Greminger M (2007) The quest for the Y-development and application of male-specific markers in orangutans (*Pongo* sp) and bottlenose dolphins (*Tursiops* sp.). Master's thesis, University of Zurich, Switzerland.

Greminger M, Krutzen M, Schelling C, Pienkowska A, Wandeler P (2010) The quest for Y-chromosomal markers – methodological strategies for mammalian non-model organisms. *Molecular Ecology Resources* **10** 409-420.

Griffiths R (2000) Sex identification in birds. *Seminars in Avian & Exotic Pet Medicine* **9** 14-26.

Griffiths R, Orr K (1999) The use of amplified fragment length polymorphism (AFLP) in the isolation of sex-specific markers. *Molecular Ecology* **8** 671-674.

Hailer F, Leonard JA (2008) Hybridization among three native North American *Canid* species in a region of natural sympatry. *PLoS ONE* **3** e3333.

Hammer MF (1994) A recent insertion of an Alu element on the Y-chromosome is a useful marker for human population studies. *Molecular Biology & Evolution* **11** 749-761.

Hammer MF, Karafet T, Rasanayagam A *et al.* (1998) Out of Africa and back again: nested cladistic analysis of human Y-chromosome variation. *Molecular Biology & Evolution* **15** 427-441.

Hardenack S, Hauser E, Law T *et al.* (2002) Mapping of Sex Determination Loci on the White Campion (*Silene latifolia*) Y Chromosome Using Amplified Fragment Length Polymorphism. *Genetics* **160** 717-725.

Hedrick P (2005) *Genetics of Populations.* Jones & Bartlett.

Hellborg L, Ellegren H (2003) Y chromosome conserved anchor tagged sequences (Y-CATS) for the analysis of mammalian male-specific DNA. *Molecular Ecology* **12** 283-291.

Hellborg L, Ellegren H (2004) Low levels of nucleotide diversity in mammalian Y chromosomes. *Molecular Biology and Evolution* **21** 158-163.

Hudson ME (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular Ecology Notes* **8** 3–17.

Jiang ZW, Zhang XQ, Deka R, Jin L (2005) Genome amplification of single sperm using multiple displacement amplification. *Nucleic Acids Research* **33** 91.

Jobling MA, Pandya A, Tyler-Smith C (1997) The Y-chromosome in forensic analysis and paternity testing. *International Journal of Legal Medicine* **110** 118-124.

Jobling MA, Tyler-Smith C (1995) Fathers and sons: the Y-chromosome and evolution. *Trends in Genetics* **11** 449-456.

Jobling MA, Tyler-Smith C (2003) The human Y-chromosome: an evolutionary marker comes of age. *Nature Review Genetics* **4** 508-612.

Kayser M, Kittler R, Erler A *et al.* (2004) A comprehensive survey of human Y-chromosomal microsatellites. *American Journal of Human Genetics* **74** 1183-1197.

Lahn BT, Page DC (1999) Four evolutionary strata on the human X-chromosome. *Science* **286** 964-967.

Lawson-Handley LJ, Brandlii L, Perrin N (2006) Disentangling reasons for low Y-chromosome variation in the greater white-toothed shrew (*Crocidura russula*). *Genetics* **173** 935-942.

Lawson-Handley LJ, Perrin N (2005) Y-chromosome microsatellite isolation from BAC clones in the greater white-toothed shrew (*Crocidura russula*). *Molecular Ecology Notes* **6** 276-279.

Lawson-Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex biased dispersal. *Molecular Ecology* **16** 1559-1578.

Luo SJ, Johnson WE, David VA *et al.* (2007) Development of Y-chromosome intraspecific polymorphic markers in the *Felidae*. *Heredity* **98** 400-413.

MacDonald AJ, Sankovic N, Sarre SD *et al.* (2006) Y-chromosome microsatellites identified from the tammar wallaby (*Macropus eugenii*) and their amplification in three other macropod species. *Molecular Ecology Notes* **6** 1202-1204.

Marchal JA, Acosta MJ, Bullejos M *et al.* (2008) Origin and spread of the SRY gene on the X and Y chromosomes of the rodent *Microtus cabrerae*: Role of L1 elements. *Genomics* **91** 141-152.

Moritz C (1994) Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology* **3** 401-411.

Muller H (1964) The relation of recombination to mutational advance. *Mutation Research* **1** 2-9.

Ohno S (1967) *Sex chromosomes and sex-linked genes.* (Springer, Berlin, 1967).

Peccon-Slattery J, Murphy J, O'Brien SJ (2000) Patterns of diversity among SINE elements isolated from three Y-chromosome genes in carnivores. *Molecular Biology & Evolution* **17** 825-829.

Peccon-Slattery J, Sanner-Wachter L, O'Brien SJ (2000) Novel gene-conversion between X-Y homologues located in the non-recombining region of the Y-chromosome in *Felidae*. *PNAS* **97** 5307-5392.

Peck JR (1994) A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. *Genetics* **137** 597-606.

Peichel C, Ross J, Matson C *et al.* (2004) The master sex-determination locus in threespine sticklebacks is on a nascent Y-chromosome. *Current Biology* **14** 1416-1424.

Peil A, Flachowsky H, Schumann E, Weber WE (2003) Sex-linked AFLP markers indicate a pseudoautosomal region in hemp (*Cannabis sativa L). Theoretical & Applied Genetics* **107** 102-109.

Petit E, Balloux F, Excoffier L (2002) Mammalian population genetics: why not Y? *Trends in Ecology & Evolution* **17** 28-33.

Quinn AE, Geoges A, Sarre SD *et al.* (2007) Temperature sex reversal implies sex gene dosage in a reptile. *Science* **316** 411.

Quinn AE, Radder RS, Sarre SD *et al.* (2009) Isolation and development of a molecular sex marker for *Bassiana duperreyi*, a lizard with XX/XY sex chromosomes and temperature-induced sex reversal. *Molecular Genetics & Genomics* **281** 665-672.

Rampias TN, Fragoulis EG, Sideris DC (2009) A hybrid-specific, polymerase chain reaction-based amplification approach for chromosomal walking. *Analytical Biochemistry*, **388**, 342–344.

Reddy PS, Mahanty S, Kaul T *et al.* (2008) A high-throughput genome-walking method and its use for cloning unknown flanking sequences. *Analytical Biochemistry*, **381**, 248–253.

Repping S, van Daalen M, Brown LG *et al.* (2006) High mutation rates have driven extensive structural polymorphism among human Y-chromosomes. *Nature Genetics* **38** 463-467.

Rice WR (1987) Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* **116** 161-167.

Rice WR (1996) Evolution of the Y-chromosome in mammals. *Bioscience* **46** 331-343.

Rosser ZH, Balaresque P, Jobling MA (2009) Gene conversion between the X-chromosome and the male-specific region of the Y-chromosome at a translocation hotspot. *American Journal of Human Genetics* **85** 130-134.

Rozen S *et al.* (2003) Abundant gene conversion between arms of palindromes in human and ape Y-chromosomes. *Nature* **423** 873-876.

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology.* Humana Press, Totowa, NJ, pp 365-386.

Schedure J, Ji HL (2008) Next-generation sequencing. *Nature Biotechnology* **26** 1135-1145.

Skaltesky H, Kuroda-Kawaguchi T, Minx PJ *et al.* (2003) The male-specific region of the human Y-chromosome is a mosaic of discrete sequence classes. *Nature* **423** 825-837.

Smit AFA (1996) The origin of interspersed repeat elements in the human genome. *Current Opinion in Genetics & Development* **101** 77-92.

Sriphairoj J, Na-Nakorn U, Brunelli JP, Thorgaard GH (2007) No AFLP sex-specific markers detected in *Pangasianodon gigas* and *P. hypophthalmus. Aquaculture* **4** 739-743.

Stumpf MPH, Goldstein DB (2001) Genealogical and evolutionary inference with the human Y-chromosome. *Science* **291** 1738-1742.

Tamura KJ, Dudley M, Nei M, Kumar S (2007) mega4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24** 1596-1599.

Telenius H, Carter NP, Bebb CE *et al.* (1992) Degenerate oligo-nucleotide primer PCR: general amplification of target DNA by a single degenerate primer. *Genomics* **13** 718-725.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22** 4673-4680.

Tringali MD, Davis MC, Rodriguez-Lopez MA *et al.* (2008) Simultaneous use of the X- and Y-chromosome genes SMCX, SMCY, and DBY for sex determination in the Florida manatee (*Trichechus manatus latirostris*). *Marine Mammal Science* **24** 218-224.

Underhill PA, Kivisild T (2007) Use of Y-chromosome and mtDNA population structure in tracing human migrations. *Annual Review of Genetics* **41** 539-564.

Volleth M, Heller KG, Pfieffer RA, Hameister H (2002) A comparative ZOO-FISH analysis in bats elucidate the phylogenetic relationships between megachiropetra and five microchiropetran families. *Chromosome Research* **10** 477-497.

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

Wallner B, Piumi F, Brem G *et al.* (2004) Isolation of Y-chromosome specific microsatellites in the horse and cross species amplification in the genus *Equus*. *Heredity* **95** 158-164.

Whitlock R, Hipperson H, Mannarelli *et al.* (2008) An objective, rapid and reproducible method for scoring AFLP peak height data that minimizes genotyping error. *Molecular Ecology Resources* **8** 725-735.

Wood ET, Stover DA, Ehret C *et al.* (2005) Contrasting patterns of Y-chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. *European Journal of Human Genetics* **13** 867-876.

Wuertz S, Gaillard S, Barbisan F *et al.* (2006) Extensive screening of sturgeon genomes by random screening techniques revealed no sex-specific marker. *Aquaculture* **258** 665-668.

Chapter 3: Investigating the phylogeography of Daubenton's bat across Europe using the mitochondrial *ND1* gene

3.1 Introduction

Population structure comprises the distribution and level of genetic diversity among individuals of a species (Hewitt & Butlin 1997) and is a product of processes occurring in the historic past that determine how much overall variation is present, and contemporary forces which dictate how this variation is distributed at a local spatial scale. Historical processes include events that have acted on the population as a whole such as range expansion and bottlenecks (Goodman *et al.* 2001, Ibrahim *et al.* 1996). Contemporary forces include processes such as genetic drift and mutation which alter levels of variation at the individual level (Balloux & Lugon-Moulin 2002). How this variation is apportioned at a local scale is then dependent on levels of gene flow, which is affected by the distribution and density of individuals within a particular patch and also how well these patches are connected. Furthermore aspects of a species' life history such as mating structure and sex-biased dispersal patterns will influence gene flow and hence population structure (Chesser 1991, Greenwood 1980). By studying the genetic signature of populations of a species at different geographic scales we can attempt to understand the relative importance of contemporary versus historic forces in shaping observed patterns of variation (Templeton *et al.* 1995).

The field of phylogeography allows us to understand how historic forces have shaped observed patterns of variation by combining genetic data with information on the geographic distribution of individual samples (Avise 2000). Without the use of a phylogeographic approach it would be difficult to determine if patterns of genetic structure observed were a consequence of local processes or past events. Mitochondrial DNA is a good marker to study the phylogeography of a species owing to a number of properties (Avise *et al.* 1987). Firstly, it is a single-copy marker inherited without recombination so it can be used to define maternal lineages. Secondly, as it is inherited maternally it can be used to investigate female-specific population structure and therefore make inferences about sex-biased dispersal. Lastly mtDNA has a high mutation rate and so can be informative on an intra-species level. The high mutation rate of mtDNA also means that population structure at a local scale can be investigated by defining the degree of differentiation between populations. This information can then

be used to investigate levels of gene flow and how this may be affected by geographical features.

Mitochondrial data have been used extensively to investigate the phylogeography of mammals in Europe (e.g. Jaarola & Tegelstrom 1995, Santucci *et al.* 1998, Taberlet *et al.* 1998) highlighting that, for the period of climatic oscillations in the Pleistocene, species generally retreated to southern refugia in the Balkans, Italy and Iberia during cold periods, where populations became differentiated in isolation prior to expanding northwards when glaciers retreated (Hewitt 1999). The majority of studies using mtDNA to look at local population structure have highlighted the degree of female-mediated gene flow between populations, which is generally dependent upon the life history of the species in question (Melnick & Hoelzer 1992, Purdue *et al.* 2000).

However caution should be exercised when making inferences about population history using mtDNA. Because gene genealogies are influenced by chance any phylogeographic breaks in genetic structure may not coincide with geographic or other barriers. This is especially pertinent when the dispersal capabilities of the focal species are low or when population sizes are small and thus more easily affected by stochastic processes (Irwin 2002). Further, because all sites in mtDNA are linked any phylogenetic inferences made from mitochondrial data are only one realisation of the coalescent process which may not be congruent with population history (Moore 1992, Nichols 2001). An additional erroneous assumption made from mtDNA phylogeography is that the markers under consideration are neutral and therefore patterns of variation will reflect demographic processes alone. However, as the whole of mtDNA is linked selection acting on any base has the potential to alter patterns of variation across the entire mitochondrial genome, therefore the role of selection cannot be ignored (Balloux 2010). As a consequence care should be taken when making conclusions based on mtDNA markers and if possible findings should be corroborated by the additional use of markers from unlinked loci (Flanders *et al.* 2009).

Bats present an unusual example as, for their size, they have long generation times and low levels of fecundity. As such we might expect historic colonization by bats to be more similar to large mammals and birds rather than similarly sized small ground dwelling mammals (Austad & Fischer 1991, Wilkinson & South 2002). Because bats are dependent on mesic habitats to feed and reproduce (Altringham 1996) glacial cycles

will have had a large impact on their distribution. Phylogeographic studies of bats suggest that in Europe bats have been able to expand their range substantially from glacial refugia in southern Europe (Mayer & von Helversen 2001), traversing areas which act as significant barriers for similar sized terrestrial mammals, such as the Straits of Gibraltar (Ditchfield 2000, Juste *et al.* 2009). However other geographic features, notably mountain ranges such as the Alps appear to have acted as strong barriers to dispersal and colonisation, as evidenced by genetic discontinuities between populations on either side of mountain ranges (e.g. *Myotis myotis*) (Castella *et al.* 2001).

Because bats have high dispersal capabilities for mammals of their size it might be expected that genetic differentiation would be low and indeed this is often the case for nuclear markers (Burland *et al.* 1999, Kerth *et al.* 2003, Rivers *et al.* 2005). In contrast, mitochondrial studies on bats typically report high levels of genetic structuring, owing to female philopatry (Chen *et al.* 2006, Ngamprasertwong *et al.* 2008, Pereira *et al.* 2009). However the degree of structuring is variable between species and can range from strict female philopatry in some species such as *Myotis bechsteinii* (Kerth *et al.* 2002) to much lower levels in migratory species such as *Miniopteris schreibersii* (Pereira *et al.* 2009). There are exceptions where mtDNA studies have found either both sexes to be philopatric as with *Plecotus auritus* (Burland *et al.* 1999), or male philopatry as with *Macroderma gigas* (Worthington-Wilmer *et al.* 1994). As such it is not possible to make generalisations about sex-biased dispersal patterns in bats (Burland & Worthington-Wilmer 2001).

It is well documented that Daubenton's bats form summer colonies with significant segregation by sex (Senior *et al.* 2005), with adults undergoing migration to swarming sites during the autumn, where they mate (Angell 2008, Parsons & Jones 2003). Ringing and radio tracking studies have shown that *M. daubentonii* has the capacity to migrate tens of kilometres between summer and autumn sites, with the maximum recorded distance in the UK being 40 km (Angell 2008) and the maximum European distance 260 km (Urbancyzk 1990). Previous mtDNA phylogeographic studies of Daubenton's in mainland Europe highlighted four mtDNA clades, indicative of range expansion into western Europe following the last glacial maximum from at least two glacial refugia (Mayer and von Helversen 2001). A recent mtDNA study on *M. daubentonii* populations across Scotland revealed two distinct clades. The first one similar to the most basal clade in mainland Europe (Mayer & von Helversen 2001) and

the second similar to the most derived suggesting that, as with other mammalian species, the UK has been colonised by Daubenton's from multiple refugia (Ngamprasertwong *et al.* 2008). Population genetic studies on Daubenton's have demonstrated that there is restricted female gene flow at a local scale using either a comparison between mtDNA and autosomal microsatellites (Ngamprasertwong *et al.* 2008) or by comparing nuclear microsatellite structure between the sexes (Angell 2008). However information on the population structure of *M. daubentonii* at a regional scale outside the UK is scarce (but see Atterby *et al.* 2009).

In this chapter, I analyze sequences from the mitochondrial *ND1* gene at different spatial scales. At a European-wide scale the phylogeographic structure of populations spanning south west (Iberia) to north eastern Europe (Scandinavia) is investigated. This is the first study which has sampled *M. daubentonii* so extensively across Europe and should illuminate the patterns by which Europe has been colonised by *M. daubentonii* following the last ice age, and the extent to which geographic features such as the Pyrenees and the English Channel have acted as isolating barriers. Populations were also sampled at a local scale to assess how genetic structure and gene flow are affected by female movement. This should resolve patterns of phylogeographic and population genetic structure in *M. daubentonii* in the UK and mainland Europe and will be combined with other markers in later chapters to investigate sex-biased dispersal.

3.2 Methodology

3.2.1 Sampling

184 individuals were sampled from twenty localities from nine European countries and five *ND1* haplotypes from Scottish populations were also included (Ngamprasertwong *et al.* 2008, see Table 3.1). Within each country localities were further grouped into regions (Table 3.1) (Figure 3.1). The sites were chosen to allow comprehensive sampling of *M. daubentonii* across north west and south west Europe where the range of this species is continuous (Societas Europaea Mammalogica). However, as permission must be obtained from the relevant authority to collect wing punch samples it was not possible to sample continuously across the sampling area. Samples were taken from a mixture of summer roosts and autumn swarming sites. Only DNA samples from adult bats (16 females, 163 males) were used to ensure that they were post-dispersal individuals.

Figure 3.1 Location of sampling sites across Europe. The size of each circle is proportional to the number of samples.

Table 3.1 Sampling localities of *M. daubentonii* individuals sequenced at the *ND1* gene

Sample site	Region	Country	Site Type	N individuals	Latitude	Longitude
Ilkley (IK)	Yorkshire	England	Summer	12	53.931582	-1.814155
Dow Cave (DC)	Yorkshire	England	Swarming	11	54.1645	-2.026002
Windy Pits (WP)	Yorkshire	England	Swarming	2	54.24488	-1.120664
Westerham (KT)	Southern England	England	Swarming	7	51.25139	0.079189
Blakenham (SX)	Southern England	England	Swarming	5	50.92607	-0.760655
Withcall (WC)	Withcall	England	Swarming	8	53.366839	-0.0086263
Pluherlin (PL)	Northern France	France	Swarming	8	47.696167	-2.363167
Le cros (LC)	Southern France	France	Swarming	22	44.3206	1.777547

Roquefort (RQ)	Southern France	France	Swarming	20	44.036111	0.349167
EAG (EG)	North Spain	Spain	Swarming	7	42.0175	-2.48
Itxulegor (IT)	North Spain	Spain	Swarming	6	43.072043	-2.81354
Grotta Marelli (GM)	Grotta Marelli	Italy	Swarming	5	45.863028	45.863028
Tiberio (TI)	Tiberio	Italy	Swarming	7	43.371778	12.748139
Liezele (LB)	Belgium	Belgium	Unknown	10	51.0602	4.28006
Bornem (BB)	Belgium	Belgium	Unknown	10	51.09962	4.23491
Marble Arch Cave (IR)	Northern Ireland	Ireland	Swarming	12	54.245641	-7.815902
Gouffre de Cathy (CA)	Switzerland	Switzerland	Swarming	2	46.495556	6.137472
Gouffre de Pleine Lune (LU)	Switzerland	Switzerland	Swarming	3	46.535528	6.329111
Aasla (FI)	Finland	Finland	Summer	14	60.27996	21.96269
Mønsted Chalkmines (DE)	Denmark	Denmark	Unknown	13	56.45024	9.18892
EU447270 (SC1)	Scotland	Scotland	Na	Na	56.44372	-4.08769
EU447271 (SC2)	Scotland	Scotland	Na	Na	56.68231	-2.76243
EU447272 (SC3)	Scotland	Scotland	Na	Na	57.75094	-4.62248
EU447273 (SC4)	Scotland	Scotland	Na	Na	56.64179	-3.15159
EU447274 (SC5)	Scotland	Scotland	Na	Na	55.15225	-4.31492

3.2.2 DNA extraction

Genomic DNA was extracted from 3 mm wing biopsies using a standard ammonium acetate precipitation method adapted to a 96 well plate extraction (http://www.sheffield.ac.uk/molecol/protocols.html). Each DNA extract was diluted in 30 µl of Low TE buffer. The concentration of DNA in a random sample of 24 extracts was low (mean value 4.78 ng/µl) so extracts were not diluted further.

3.2.3 Primers

A 1330 bp portion of the mitochondrial *ND1* gene was PCR amplified. The *ND1* gene was chosen as it is variable at the intra-species level but has a low enough mutation rate to be informative at a phylogeographic scale. The primers used were L2985 (5'-CCTCGATGTTGGATCAGG-3') and H4419 (5'-GTATGGGCCCGATAGCTT-3') (Ngamprasertwong *et al.* 2008).

3.2.4 PCR reactions

Each PCR constituted a 12 µl reaction containing 2 µl of genomic DNA and 10 µl of Qiagen multiplex kit, which contained 0.2 µM of both the forward and reverse primer. PCR conditions were as follows; 95 °C for 15 minutes, followed by 35 cycles of 94 °C for 30 seconds, 58 °C for 90 seconds, 72 °C for 60 seconds and a final extension step of 60 °C for 30 minutes.

3.2.5 Sequencing

Prior to sequencing each PCR product reaction was cleaned using the ExoSap method, with a ratio of 5 µl of PCR product to 1 µl of ExoSapIT (USB Corp.). Big dye kit version 3.1 (Applied Biosystems) was chosen for the cycle sequencing reaction as it is optimised for larger product sizes. Cycle sequencing was performed in the reverse direction only. Standard cycle sequencing and ethanol precipitation methods were implemented prior to running samples on an ABI 3730 DNA sequencer.

3.2.6 Sequence data

3.2.6.1 Cleaning up sequences

ABI trace files were viewed in Codoncode Aligner version 3.0.2 (Codoncode Corp.). The raw trace files were assembled. Any trace files that were of poor quality or unreadable were then discarded. The beginning and end of each sequence within the assembly was viewed by eye for quality. Any portion of this sequence that was unreadable or had a low quality score was removed. The sequence was then reassembled. Any position on a sequence that had been denoted with the universal

ambiguity code was viewed by eye and the uncertainty was resolved. The sequences were then searched for SNPs. When a SNP was found the sequences were viewed at this position to confirm the validity of the SNP. Subsequently three sequences were found with an unusually high number of unique SNPs. A nucleotide BLAST search of these sequences against the NCBI database found that they belonged to species other than *M. daubentonii* due either to misidentification during sampling or data recording errors (misidentified species were found to be two *Myotis alcathoe* individuals and one *Myotis natterreri* individual) and they were subsequently removed.

3.2.6.2 Sequence alignment

Multiple sequence alignments of the 184 *ND1* sequences were made using the Clustal-W algorithm (Thompson *et al.* 1994) in MEGA version 4.1. (Tamura *et al.* 2007). This was in addition to five *ND1* haplotypes from Scottish populations of *M. daubentonii* deposited in Genbank by Ngamprasertwong *et al.* (2008) (see Table 3.1 for details). Because sequencing was performed in the reverse direction the sequence files were reversed to show the forward complement. As the sequence files varied in length they were trimmed to a homologous section of 617 bp to allow comparability between sequences whilst retaining an informative number of SNPs.

3.2.7 Analysis

3.2.7.1 Sequence diversity

Levels of polymorphism were characterised for each population using a number of descriptors: Number of haplotypes, haplotype diversity, segregating sites, nucleotide differences and nucleotide diversity, estimated with DnaSP version 5 (Librado *et al.* 2009).

3.2.7.2 Phylogeographic relationships

Phylogenetic relationships between *ND1* haplotypes were initially investigated by constructing a neighbour joining tree using the Kimura two-parameter distance in MEGA 4.1 (Tamura *et al.* 2007) rooted with the homologous stretch of the *ND1* gene from *Myotis bechsteinii* (*M. bechsteinii* was chosen to root the tree as a phylogeny of

European bats showed *M. bechsteinii* and *M. daubentonii* to be closely related (Mayer & von Helversen 2001)). The robustness of the tree was validated through bootstrapping using 1000 replicates.

Phylogenetic trees were also reconstructed using a Bayesian approach in BEAST v 1.4.8 (Drummond & Rambaut 2007) with the following parameters: HKY with invariable sites, with three codon partitions and all parameters unlinked. In each case a strict molecular clock was assumed but no clock rate was specified. Three independent searches were run for 10,000,000 generations each. The trees from all runs were combined using log combiner v 1.4.8, re-sampling every 1,000th tree which resulted in a total of 3,000 trees. TreeAnnotator v 1.4.8 was used to choose one tree out of the 3,000 remaining trees based on the likelihoods. Trees were viewed using FigTree v 1.2.

In order to look at the relationships between haplotypes, TCS version 1.21 (Clement *et al.* 2000) was used to create a statistical parsimony network of *ND1* haplotypes. Relative to phylogenetic tree construction TCS was advantageous for studying intraspecific gene genealogies as it allowed the possibility that the ancestral haplotype was still present. For the purpose of this analysis the connection limit was fixed at 20 steps.

3.2.7.3 NCPA

In order to examine the geographic structure underlying genetic variation of the *ND1* gene a nested clade phylogeographic analysis (NCPA) test was performed. NCPA tests for association between haplotypes and geographical distribution and has the potential to suggest causal processes using an inference key (Templeton 2004). As such it can suggest possible population histories from the sampled species range. NCPA uses the haplotype network produced by TCS and defines a clade hierarchy based on this network. Nesting begins at the tips of the network identifying clades separated by a one step mutation and proceeds step by step until a nesting level is reached that incorporates the whole cladogram. NCPA assumes a null hypothesis of random association between genetic and geographic structure within each clade nesting level. When the null hypothesis is rejected, NCPA outlines causal processes which may help explain the association. Causal processes can range from contemporary events such as restricted gene flow or historic factors such as population fragmentation. Inferences may differ

depending on clade and nesting level but NCPA will suggest the process which best explains the overall cladogram structure. There has been criticism of the reliability of the outcome produced by NCPA as it is suggested that local processes operating in a population can result in false inferences on a range-wide scale (Panchal & Beaumont 2007, Petit 2008). However NCPA remains one of the most frequently used tools in phylogeography and has been validated by its use in conjunction with supporting analyses and prior knowledge of the population in question (Garrick *et al.* 2008, Templeton 2008). Therefore whilst causal processes suggested by NCPA should not be accepted in isolation, they can be used to provide support for hypotheses when used in conjunction with other analyses.

3.2.7.3.1 NCPA input files

In order to incorporate geographic information into Aneca (Panchal & Beaumont 2007) it is necessary to create a distance file defining the geographic co-ordinates of each sampling site and the radius surrounding that site. The default radius given in this case was 3 km as it is assumed that bats found within this radius have a high probability of belonging to the sampled population (Altringham pers comm.). NCPA also requires that any areas where the organism occurs but has not been sampled are taken into account. In this case sampling was relatively comprehensive across the study area (Figure 3.1) bar central France. A mid point in central France was selected (lat: 47.372972 long: 6.243908) and given a radius of 100 km to account for the lack of sampling in this area.

3.2.7.4 Demographic expansion

Three coalescent models of demographic expansion were tested using BEAST: constant size, expansion growth and exponential growth. A model was applied as used for the reconstruction of the haplotype trees (see section 3.2.7.2) and Tracer v.1.4.8 (Rambaut & Drummond 2007) was used to view the combined log files of each algorithm. In order to test which prior was the most likely mode of population growth two factors were taken into consideration: Firstly the likelihoods of the three priors were compared by calculating the Bayes factor (in natural log units). Secondly the marginal posterior probability density distributions of the growth rates of the exponential and expansion growth models were viewed to see if they overlapped zero. If they did not have a

significant overlap with zero the hypothesis of a constant population in equilibrium cannot be accepted.

In order to further investigate historical demography of *M. daubentonii* in Europe Arlequin version 3.11 (Excoffier *et al.* 2005) was used to calculate three statistics. The first is Tajima's D (Tajima 1989) which tests whether θ_s is equal to θ_π, as expected in the absence of selection in a population of a constant size. The second is Fu's F_S (Fu 1997) statistic which calculates the probability of observing a random sample from a standard neutral population with no more alleles than the observed value given the mean number of pairwise differences. A negative D or F_S value signifies an excess of rare polymorphisms, which is suggestive of demographic expansion (or positive selection). Finally, pair-wise mismatch distributions (Slatkin & Hudson 1991) between individuals were calculated and compared against a model of sudden demographic expansion using Harpending's raggedness index *r* of the observed distribution and the sum of the squared deviation (SSD) between observed and expected distributions (Rogers & Harpending 1992). If the *r* value and SSD are non-significant the null hypothesis of sudden range expansion cannot be rejected. Significance levels were computed by bootstrapping using 1000 replicates. If an expansion was detected the time of expansion in generations (*t*) was estimated from $\tau = 2ut$. Where τ (tau) is calculated as the time to expansion in mutational units and *u* is the mutation rate per generation for the DNA sequence being studied. A substitution rate of 2% per base per million years was chosen for *ND1*, as this is typical of many mammalian mitochondrial sequences (Powell *et al.* 1986). The generation time for *M. daubentonii* was assumed to be 2 years.

3.2.7.5 Population structure

Arlequin version 3.11 (Excoffier *et al.* 2005) was used to calculate Φ_{ST} (a measure analogous to F_{ST} based on pair-wise distances between sequences). Φ_{ST} was used to define the degree of differentiation between localities using pair-wise comparisons. Restricted gene flow between localities can be inferred if the value of Φ_{ST} is significantly greater than zero. The higher the value of Φ_{ST} between any given pair of regions, the greater is the degree of genetic differentiation between them.

Analysis of molecular variance (AMOVA) with 10,000 permutations was used to investigate the distribution of genetic variability partitioned among the 14 regions

(Table 3.1), among sites within regions and within the 21 sites (note that the Scottish haplotypes were grouped into a single sample site).

In order to assess whether there was any correlation between genetic and geographic distance, an isolation-by-distance (IBD) test was applied to the whole data set. This involved a correlation of pair-wise comparisons between geographic distance between sites (km) (calculated in a distance matrix, using the haversine formula by Roger Butlin) and genetic distance between sites (Φ_{ST}). Note that as some Φ_{ST} measures were close to 1, Φ_{ST} was not linearised as this would produce values close to infinity for some pair-wise comparisons. A Mantel test with 10,000 permutations was performed in the isolation by distance web service version 3.14 (Jensen *et al.* 2005) to test the significance of an association between genetic and geographic distance. The process was repeated for samples from England only as sampling in England was more comprehensive than from any other country in this study.

3.3 Results

As the results of phylogenetic analyses detected the presence of two mitochondrial clades (denoted A and B, see Figure 3.2) all subsequent analyses were conducted on the two mitochondrial clades separately and the two clades combined. This was with the exception of the TCS network and nested clade analyses which were applied to the whole data set only. This was done to determine whether past fragmentation in glacial refugia, or a more simple explanation such as isolation by distance could account for observed patterns in the data (see Irwin 2002).

3.3.1 Sequence analysis and variation

In total, 37 of 617 nucleotide positions were found to be variable resulting in 29 unique *ND1* haplotypes. Three of the 37 (8.1%) variable positions involved transversion mutations and all were non-synonymous. Some haplotypes were widespread and common to a number of different localities (notably Hap 0 and Hap 6) whilst the vast majority were unique to a population or region (Table 3.2) (Figure 3.2). The two Italian locations were the only sites that did not contain any haplotypes common to other regions. Of the sampling localities that were accounted for by five or more individuals, all were polymorphic bar EAG in Spain.

Genetic diversity values showed some variability between localities (Table 3.3.). Unsurprisingly, the highest nucleotide diversity values were found in those localities which shared a mixture of haplotypes from clade A and B. Haplotype diversity was generally high for most localities but tended to be lower in areas where clades A and B met, such as the UK.

Table 3.2 *ND1* haplotype distribution among sampling locations. See Table 3.1 for location abbreviations

Haplotype	Locations	Countries
0	LU, CA, RQ, PL, LC, SX, KT, IK, DC, WC, WP, IR, BB, LB, EG, IT, SC,	Switzerland, France, England, Ireland, Scotland, Belgium, Spain
1	SX	England
2	RQ	France
3	RQ	France
4	RQ	France
5	KT, WP, IT, FI	England, Spain, Finland
6	KT, DC, IK, SC, DE, FI	England, Scotland, Denmark, Finland
7	BB, LB	Belgium
8	BB	Belgium
9	BB	Belgium
10	GM	Italy
11	GM, TI	Italy
12	GM, TI	Italy
13	PL	France
14	PL	France
15	PL	France
16	CA	Switzerland
17	IR	Ireland
18	IR	Ireland
19	WC	England
20	FI	Finland
21	FI	Finland
22	TI	Italy
23	LC	France
24	SC	Scotland
25	DE	Denmark
26	DE	Denmark
27	DE	Denmark
28	DE	Denmark

Table 3.3 Descriptive statistics of genetic diversity in a 617 bp fragment of the mtDNA *ND1* gene from 21 sample sites across Europe.

Sample site	Number of haplotypes	Haplotype diversity	Segregating Sites	Nucleotide Differences	Nucleotide diversity
Ilkley	2	0.303	15	4.56	0.007367
Dow Cave	2	0.545	15	8.18	0.013261
Windy Pits	2	0	16	16.00	0.015559
Withcall	2	0.25	2	0.50	0.00081
Westerham	3	0.524	16	7.43	0.01204
Blakenham	2	0.4	1	0.40	0.000648

Pluherlin	4	0.821	4	1.64	0.002663
Le cros	2	0.455	1	0.46	0.000737
Roquefort	4	0.5	3	0.59	0.000904
EAG	1	0	0	0.00	0
Itxulegor	2	0.6	0	9.60	0.01556
Grotta Marelli	3	0.7	2	0.80	0.001297
Tiberio	3	0.714	2	1.14	0.001852
Liezele	2	0.533	16	8.53	0.01383
Bornem	4	0.711	16	8.51	0.013794
Marble Arch Cave	3	0.591	3	1.23	0.001989
Cathy	2	na	na	1.00	na
Pleine Lune	1	0	0	0.00	0
Aasla	4	0.648	3	0.77	0.001247
Mønsted	5	0.803	5	1.29	0.002087
Chalkmines Scotland	3	0.7	16	9.44	0.015235
Total	29	0.786	38	7.48	0.012121

The neighbour joining tree (Figure 3.2) provided strong bootstrap support (> 90%) for a grouping of haplotypes into two clades. However, there was little support for further grouping within the two major clades. The Bayesian trees produced by BEAST were highly similar and no more informative than the NJ tree and so are not shown. Based on the statistical parsimony network (Figure 3.3), there is a clear distinction between *ND1* haplotypes which, as with the NJ tree, can be broadly split into two major clades (denoted A and B) separated by ten interior nodes with 2.4% sequence divergence between the most internal haplotypes (Figure 3.3). The two haplotypes with the highest ancestral probabilities were Hap 0 and Hap 6 respectively. Hap 0 was found in populations from south west Europe and Hap 6 was found in populations in north east Europe. However, where these two regions met (central and southern England, Scotland, Belgium) haplotypes from the two clades were found in sympatry. Haplotypes from Italy differed from the ancestral haplotype for clade B by more than two steps in all cases suggesting greater divergence than other haplotypes within clade B. Unusually, some samples from Itxulegor in northern Spain were represented by a haplotype (Hap 5) shared with Finnish samples, despite a large geographic distance between these two sites.

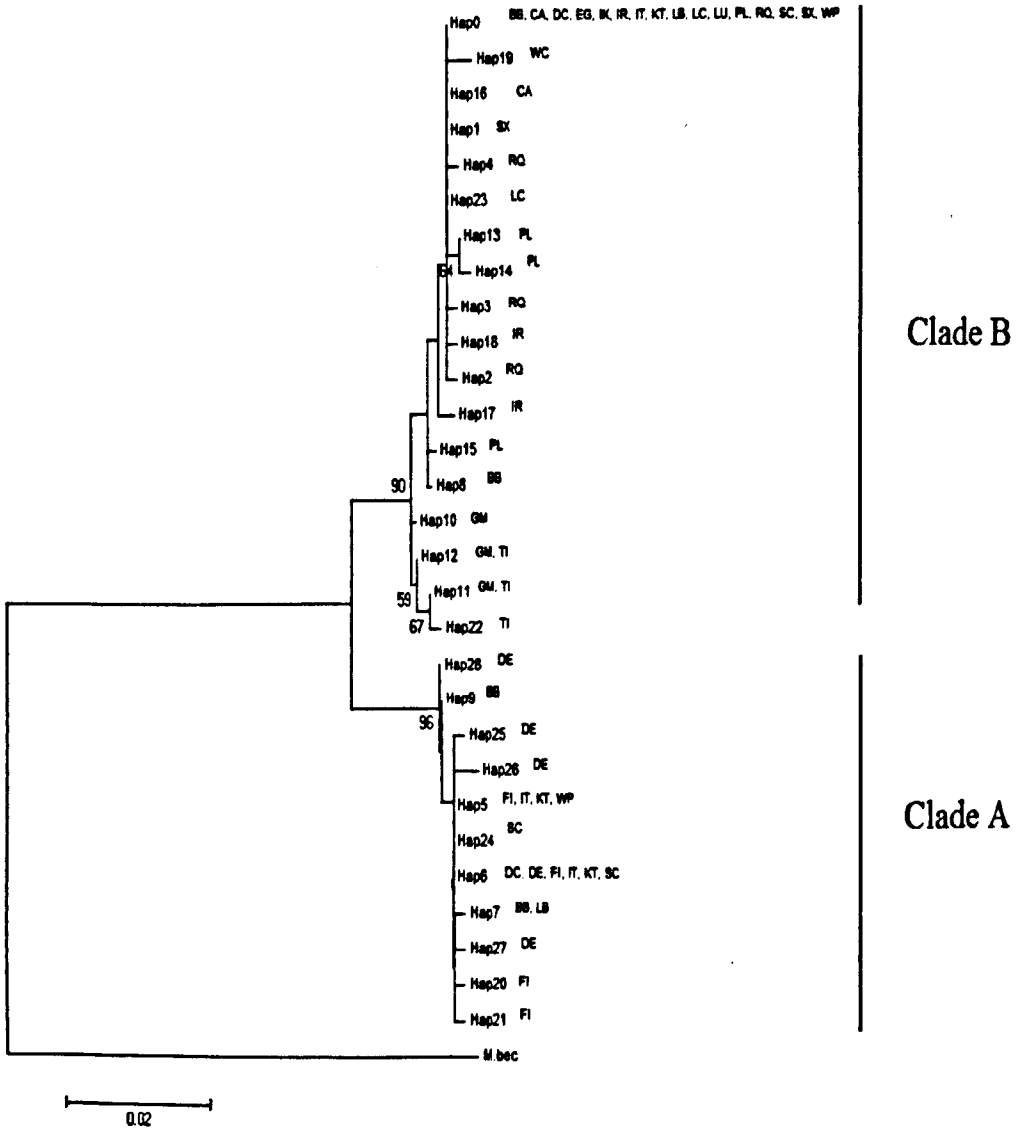


Figure 3.2 Phylogenetic relationships among haplotypes of *M. daubentonii.* Bootstrap consensus of a neighbour joining analysis on k2p distances with 1000 replicates. Values at main nodes are bootstrap support percentages (shown for values >50%). For details of site names corresponding to haplotypes see table 3.1.

3.3.2 NCPA

The nested cladogram constructed by NCPA showed 12 first level, 7 second level and 3 third level nested clades (Figure 3.3.). Haplotypes belonging to the clade found predominantly in north east Europe grouped separately from haplotypes belonging to the clade found predominantly in south west Europe. There were associations between haplotypes and their sampling localities at all nesting levels (Table 3.4.). Significant levels of geographic association were observed for clades 1-8, 2-2, 3-3 and the overall cladogram (Table 3.4.). Inferences made by the Templeton key as regards the mechanism underlying clade distribution are summarised in Table 3.5. For clades 1-5,

1-8 and 2-2 a pattern of isolation by distance is supported. Isolation by distance is also invoked to explain the total cladogram, suggesting restricted gene flow over a large geographic scale.

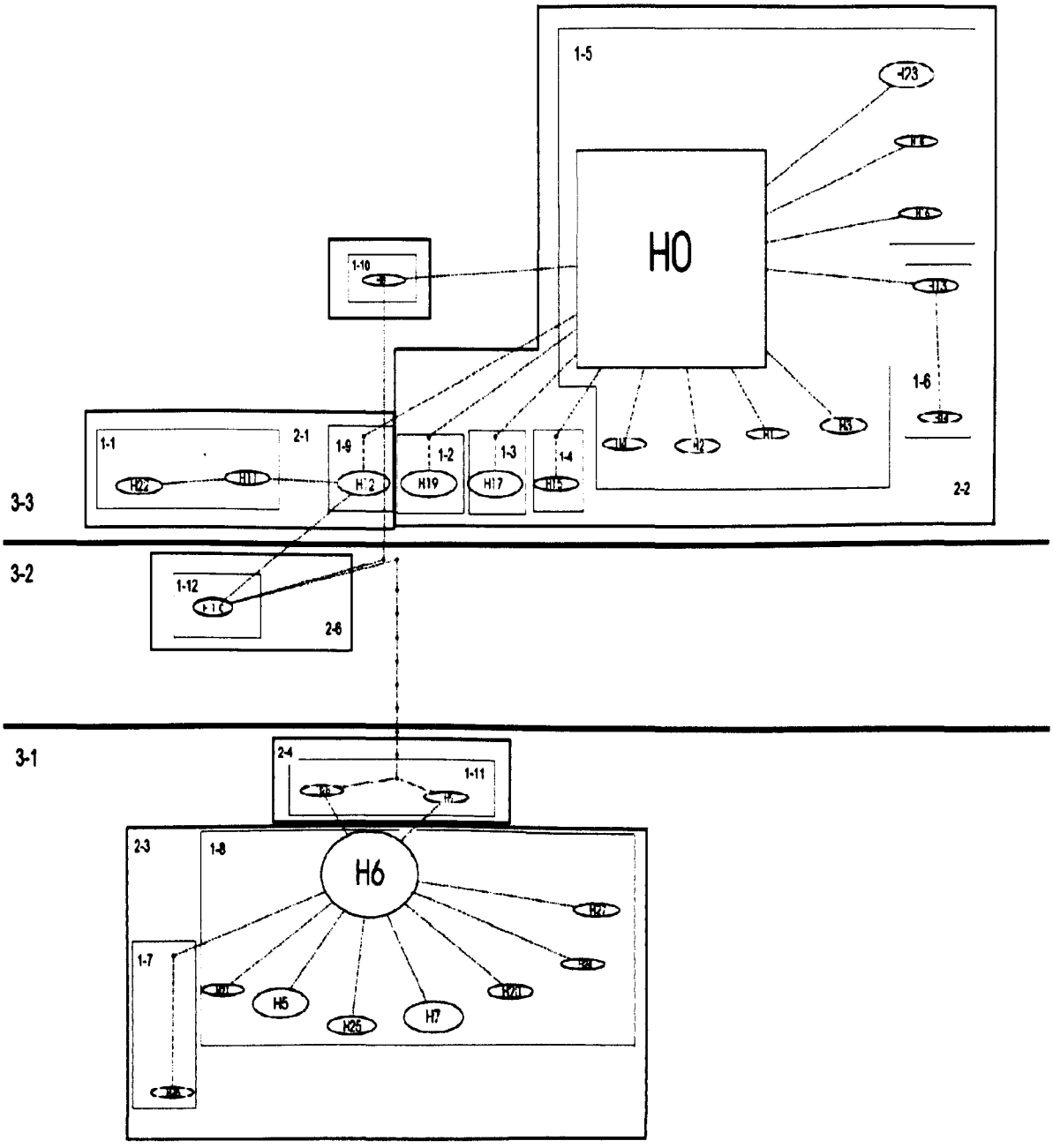


Figure 3.3 Minimum spanning network for *ND1* haplotypes. Nesting levels are denoted by boxes. Ellipse size is proportional to haplotype frequency and filled dots are hypothesised mutational steps. Nesting levels 3-2 and 3-3 correspond to clade B and nesting level 3-1 corresponds to clade A.

Table 3.4 Nested contingency analysis of geographical associations for *M. daubentonii ND1* data. Clades showing no genetic or geographic variation are excluded. Values followed by * are statistically significant $P < 0.05$

Clade	χ^2 statistic
1-1	1.88
1-5	132.46
1-8	185.03*
1-11	2.00
2-1	1.06
2-2	232.80*
2-3	4.35
3-1	5.69
3-3	157.92*
Total cladogram	166.18*

Table 3.5 Demographic inferences from NCPA analysis for the *ND1* gene in *M. daubentonii*

Clade	Inference chain	Inferred pattern
1-5	1-2-3-4 NO	Isolation by distance
1-8	1-2-3-4 NO	Isolation by distance
2-2	1-2-3-4 NO	Isolation by distance
Total cladogram	1-2-3-4 NO	Isolation by distance

3.3.3 Demographic expansion

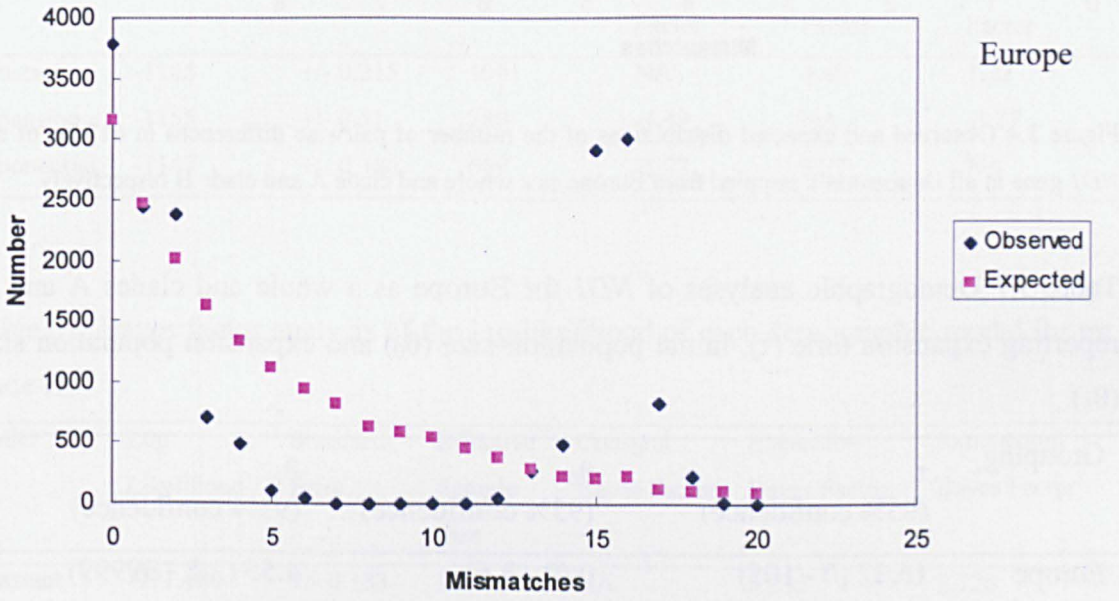
Tajima's D and Fu's F_S were negative and significant for both clade A and clade B suggesting demographic expansion of both clades (Table 3.6). Fu's F_S was also negative and significant for Europe as a whole (Table 3.6), but Tajima's D was positive and not significant.

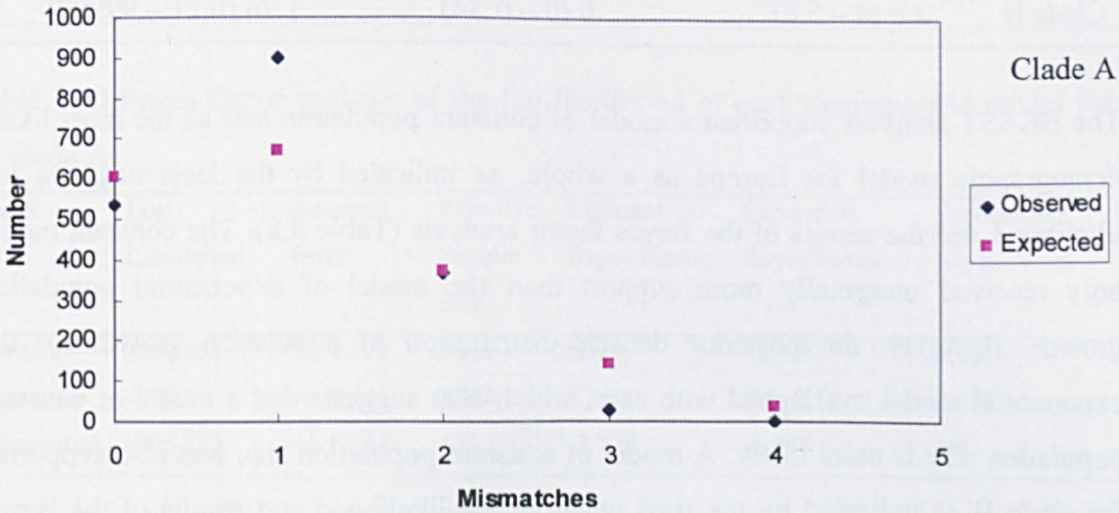
Table 3.6 Tajima's D and Fu's F_S statistics for Europe and mitochondrial clades A and B

	Tajima's D	P-value	Fu's F_S	P-value
Europe	0.44	0.77	-24.57	0.001
Clade A	-1.69	0.017	-3.438	0.001
Clade B	-1.83	0.012	-28.88	0.001

A non-significant P value was obtained for the mismatch distribution of Europe as a whole (ssd = 0.07 P = 0.16, r = 0.05, P = 0.46) and clade B individuals (ssd = 0.01 P = 0.6, r = 0.07, P = 0.73), which is surprising considering the apparent bimodal

distribution of the mismatch analyses (see Figure 3.4). However the confidence intervals for the upper and lower bound of the expected mismatch values were very wide in both cases (Europe; lower = 0 – 1027, upper = 1526 – 6955 clade B; lower = 0 – 517, upper = 1391 – 5796). The departure from expectation was significant for clade A individuals (ssd = 0.02 $P < 0.05$, r = 0.16, $P < 0.01$). Given an estimated mutation rate of 2% per million years for the *ND1* gene and a generation time in *M. daubentonii* of two years, timing of expansion for Europe as a whole is roughly 132,000 years bp (0 – 850,000) and clade B is roughly 21,000 years bp (0 - 47,000).





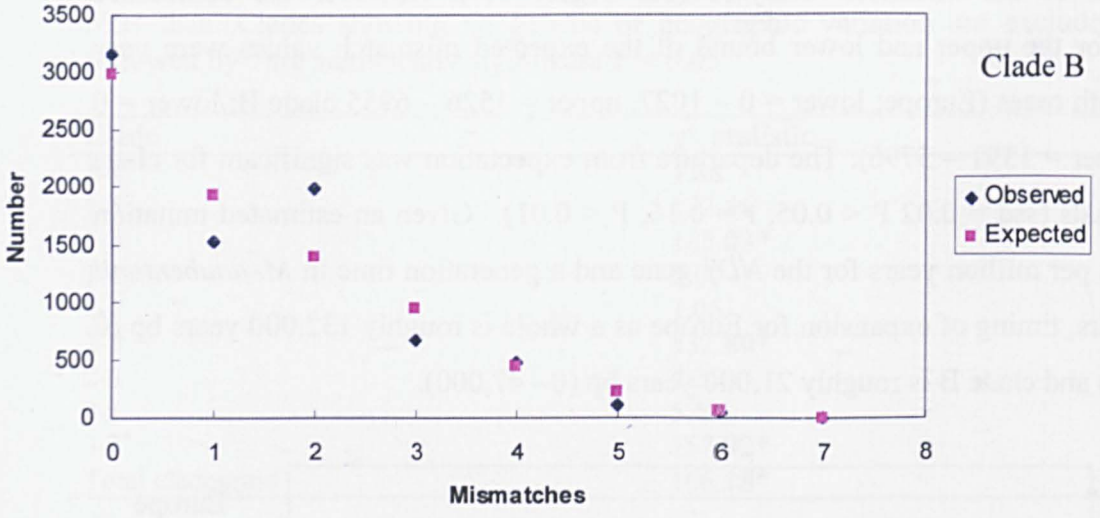


Figure 3.4 Observed and expected distributions of the number of pairwise differences in 617 bp of the *ND1* gene in all Daubenton's sampled from Europe as a whole and clade A and clade B respectively.

Table 3.7 Demographic analyses of *ND1* for Europe as a whole and clades A and B reporting expansion time (τ), initial population size (θ_0) and expansion population size (θ_1)

Grouping	τ (95% confidence)	θ_0 (95% confidence)	θ_1 (95% confidence)
Europe	16.32 (0 - 105)	0 (0 - 3.6)	4.55 (2.2 - 99999)
Clade A	1.1 (0.61 – 1.71)	0 (0 – 0.11)	99999 (3.83 – 99999)
Clade B	2.6 (0 – 5.8)	0 (0 – 0.34)	1.74 (0.21 – 99999)

The BEAST analyses supported a model of constant population size as the most likely demographic model for Europe as a whole, as indicated by the least negative log likelihood and the results of the Bayes factor analysis (Table 3.8). The constant model only received marginally more support than the model of exponential population growth. However, the posterior density distribution of population growth for the exponential model overlapped with zero, which also suggests that a model of constant population size is more likely. A model of constant population size was also supported for clade B as indicated by the least negative log likelihood and results of the Bayes factor analysis (Table 3.9). In this case the constant model only received marginally more support than the model of population expansion and the posterior density distribution of the growth rate for the expansion model did not overlap with zero, which suggests that the model of constant population size cannot be accepted. A model of

population expansion was supported as the most likely demographic model for clade A as indicated by the least negative log likelihood, results of the Bayes factor analysis (Table 3.10) and the fact that the posterior density distribution of the growth rate for the expansion model did not overlap with zero.

Table 3.8 Bayes factor analysis of the log-likelihood of each demographic model for Europe as a whole

Model	Log Likelihood	Standard Error	Effective Sample Size	Constant Bayes Factor	Expansion Bayes Factor	Exponential Bayes Factor
Constant	-1145	+/- 0.215	1091	NA	4.49	1.22
Expansion	-1155	+/- 0.51	250	-4.49	NA	-3.27
Exponential	-1147	+/- 0.386	669	-1.22	3.27	NA

Table 3.9 Bayes factor analysis of the log-likelihood of each demographic model for mt clade B

Model	Log Likelihood	Standard Error	Effective Sample Size	Constant Bayes Factor	Expansion Bayes Factor	Exponential Bayes Factor
Constant	-971.484	+/- 0.183	1290	NA	0.235	2.726
Expansion	-972.489	+/- 167	907	-0.235	NA	2.491
Exponential	-976.189	+/- 0.175	531	-2.727	-2.491	NA

Table 3.10 Bayes factor analysis of the log-likelihood of each demographic model for mt clade A

Model	Log Likelihood	Standard Error	Effective Sample Size	Constant Bayes Factor	Expansion Bayes Factor	Exponential Bayes Factor
Constant	-897.051	+/- 0.177	3048	NA	-0.209	3.788
Expansion	-896.57	+/- 0.131	3387	0.209	NA	3.997
Exponential	-905.773	+/- 0.738	153	-3.778	-3.997	NA

3.3.4 Population structure

The AMOVA results ($P < 0.001$) showed that for both Europe as a whole and for the clades separately most of the variation could be explained by either the 'Among groups' or the 'within populations' component with the 'within group' factor accounting for the smallest portion of the variance (see Table 3.11). This suggests that whilst there was a considerable level of variation among individuals within sites, there was still marked structure between regions. The overall Φ_{ST} values were high and significant ($P < 0.001$) at 0.55 (Europe), 0.42 (Clade A), 0.59 (Clade B) suggesting that female mediated gene flow was restricted.

Table 3.11 AMOVA results indicating the proportion of variation at the individual level, site level and regional level.

	Europe	Clade A	Clade B
% variation within populations	44.56%	57.23%	40.80%
% variation among populations	7.54%	2.83%	0.88%
% variation among regions	47.90%	39.94%	58.32%

Out of the 210 pair-wise comparisons between sample sites in Europe as a whole 123 (56%) had a Φ_{ST} value significantly different from zero (Table 3.13).

Out of the 45 pair-wise comparisons between sample sites for clade A individuals only 15 were significant (33%) (Table 3.12). Because of the small number of clade A individuals in some sample sites even though Φ_{ST} values were often high they were not significant.

Out of 153 pair-wise Φ_{ST} comparisons between sample sites for clade B individuals 57 were significant (37%) (Table 3.14). Note that due to the small sample size of KT (N = 2) and WP (N = 1) Φ_{ST} between these two sites was not calculated as it would result in a large negative value.

Table 3.12 Pair-wise Φ_{ST} values between sample sites for clade A individuals. Value in bold are significant ($P < 0.05$).

	KT	LB	BB	DE	IK	NYM	IT	FI	SC	DC
KT	0.0000									
LB	0.8222	0.0000								
BB	0.5032	-0.0811	0.0000							
DE	0.0349	**0.5037**	**0.4111**	0.0000						
IK	0.6875	**1.0000**	**0.7436**	0.0984	0.0000					
NYM	NA	1.0000	0.6667	0.2879	1.0000	0.0000				
IT	0.2500	**1.0000**	**0.7710**	**0.4722**	**1.0000**	0.0000	0.0000			
FI	-0.1756	**0.6084**	**0.4878**	**0.1212**	0.0667	0.2308	**0.4324**	0.0000		
SC	0.0000	0.8222	0.5032	0.0349	0.6875	0.3333	0.7692	0.1098	0.0000	
DC	0.5385	**1.0000**	**0.6667**	0.0368	0.0000	1.0000	**1.0000**	0.0094	0.5385	0.0000

Table 3.13 Pair-wise Φ_{ST} values between sample sites for all European samples. Values in bold are significant ($P < 0.05$).

	LU	CA	SX	RQ	LC	KT	LB	BB	GM	TI	EAG	IT	DE	PL	DC	IK	IR	NYM	SC	WC	FI
LU	0																				
CA	0.25	0																			
SX	0.132	0.172	0																		
RQ	0.139	0.178	0.016	0																	
LC	0.052	0.315	0.163	**0.182**	0																
KT	0.015	0.101	0.097	**0.326**	**0.366**	0															
LB	0.13	0.058	0.219	**0.431**	**0.46**	0.081	0														
BB	0.374	0.316	**0.444**	**0.637**	**0.66**	0.089	0.027	0													
GM	**0.784**	**0.708**	**0.769**	**0.773**	**0.806**	0.207	0.241	**0.403**	0												
TI	**0.729**	**0.681**	**0.74**	**0.78**	**0.808**	**0.318**	**0.334**	**0.478**	0.158	0											
EAG	0	0.588	0.073	0.014	0.161	0.161	0.275	**0.497**	**0.864**	**0.81**	0										
IT	0.25	0.158	0.355	**0.607**	**0.636**	0.061	0.076	0.062	0.333	**0.428**	0.432	0									
DE	**0.936**	**0.927**	**0.938**	**0.949**	**0.954**	**0.693**	**0.547**	**0.355**	**0.92**	**0.92**	**0.92**	**0.498**	0								
PL	0.009	0.027	0.097	**0.222**	**0.298**	0.133	0.248	**0.462**	0.616	**0.648**	0.155	**0.381**	**0.913**	0							
DC	0.307	**0.248**	0.379	**0.573**	**0.597**	0.014	-0.04	0.061	0.349	**0.428**	**0.43**	0.107	**0.392**	**0.396**	0						
IK	**0.714**	**0.689**	**0.743**	**0.836**	**0.848**	**0.411**	**0.274**	0.074	**0.701**	**0.734**	**0.774**	0.158	**0.102**	**0.736**	0.1	0					
IR	0.312	0.326	**0.368**	**0.454**	**0.508**	**0.262**	**0.344**	**0.538**	0.701	**0.719**	**0.429**	**0.485**	**0.923**	**0.372**	**0.476**	**0.774**	0				
NYM	0.25	0	0.444	0.734	0.765	0.294	0.315	0.295	0.398	**0.515**	0.588	0.5	**0.595**	0.441	0.358	0.054	0.579	0			
SC	0.109	0.002	0.234	**0.514**	**0.551**	0.163	0.146	0.049	0.254	**0.368**	0.323	0.159	**0.609**	0.263	0.116	**0.258**	**0.39**	0.447	0		
WC	**0.788**	**0.741**	**0.763**	**0.737**	**0.774**	**0.382**	**0.413**	**0.57**	**0.852**	**0.832**	**0.847**	**0.533**	**0.946**	**0.61**	**0.521**	**0.797**	**0.692**	**0.662**	**0.474**	0	
FI	**0.957**	**0.951**	**0.957**	**0.959**	**0.963**	**0.72**	**0.573**	**0.379**	**0.942**	**0.938**	**0.966**	**0.519**	**0.133**	**0.93**	**0.416**	**0.105**	**0.937**	0.642	**0.647**	**0.961**	0

Table 3.14 Pair-wise Φ_{ST} values between sample sites for clade B individuals. Values in bold are significant ($P < 0.05$).

	LU	SX	RQ	KT	LB	BB	GM	EG	PL	DC	CA	IK	IR	WC	IT	TI	LC	SC
LU	0.0000																	
SX	0.1321	0.0000																
RQ	0.1394	0.0161	0.0000															
KT	0.0000	0.0000	0.0516	0.0000														
LB	0.0000	0.0400	0.0300	0.0000	0.0000													
BB	0.0909	0.0064	0.0266	0.0625	0.1111	0.0000												
GM	**0.7842**	**0.7692**	**0.7727**	**0.8333**	**0.8500**	**0.7380**	0.0000											
EG	0.0000	0.0729	0.0141	0.0000	0.0000	0.1515	**0.8636**	0.0000										
PL	0.0091	0.0969	**0.2215**	0.0963	0.1287	0.0714	**0.6160**	0.1554	0.0000									
DC	0.0000	0.0000	0.0516	0.0000	0.0000	0.0625	**0.8333**	0.0000	0.0963	0.0000								
CA	0.2500	0.1720	0.1779	0.4737	0.5385	0.1111	**0.7079**	0.5882	0.0268	0.4737	0.0000							
IK	0.0000	0.2903	0.2642	0.0000	0.0000	0.2632	**0.7436**	0.0000	0.1323	0.0000	0.0000	0.0000						
IR	0.3121	**0.3675**	**0.4544**	**0.3836**	**0.4082**	0.3478	**0.7010**	**0.4294**	**0.3724**	**0.3836**	0.3262	0.2421	0.0000					
WC	**0.7876**	**0.7634**	**0.7369**	**0.8230**	**0.8362**	**0.7500**	**0.8522**	**0.8474**	**0.6104**	**0.8230**	**0.7405**	0.7612	**0.6918**	0.0000				
IT	0.0000	0.1321	0.1394	0.0000	0.0000	0.0909	**0.7842**	0.0000	0.0091	0.0000	0.2500	0.0000	0.3121	**0.7876**	0.0000			
TI	**0.7290**	**0.7399**	**0.7802**	**0.7773**	**0.7948**	**0.7202**	0.1584	**0.8095**	**0.6483**	**0.7773**	**0.6813**	**0.6915**	**0.7189**	**0.8316**	**0.7290**	0.0000		
LC	0.0517	0.1625	**0.1822**	0.1247	0.1451	0.1722	**0.8060**	0.1612	**0.2976**	0.1247	0.3151	0.0410	**0.5076**	**0.7737**	0.0517	**0.8079**	0.0000	
SC	0.0000	0.1321	0.1394	0.0000	0.0000	0.0909	**0.7842**	0.0000	0.0091	0.0000	0.2500	0.0000	0.3121	**0.7876**	0.0000	**0.7290**	0.0517	0.0000

The Mantel test demonstrated significant isolation by distance across all populations in Europe, supporting inferences made from the nested clade analysis ($r = 0.28$, $P = 0.01$, slope = 6.6×10^{-4}, Figure 3.5). However when analyses were performed separately on clades A and B the Mantel tests were non-significant (Clade A; $r = 0.20$, $P = 0.06$, Figure 3.6a. Clade B; $r = 0.14$, $P = 0.27$, Figure 3.6b). Isolation by distance was not supported when using data from English populations alone ($r = 0.42$, $P = 0.07$, slope = 1.94×10^{-3}, Figure 3.7).

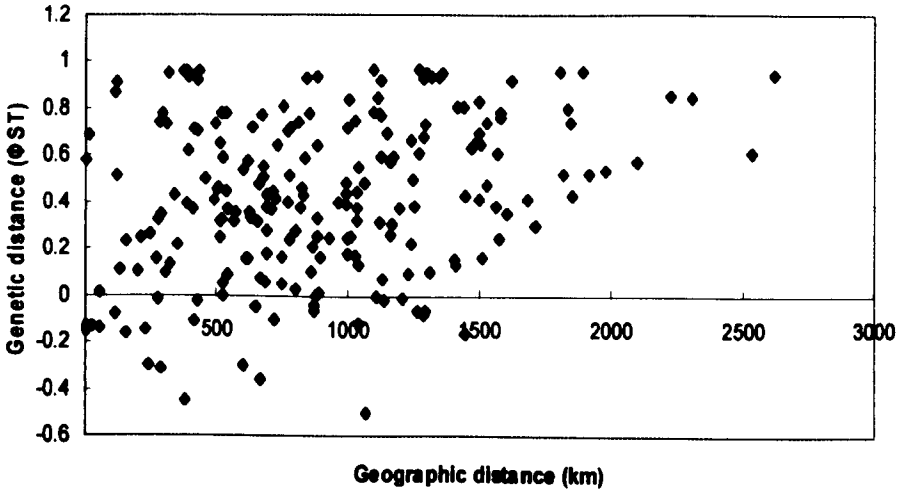
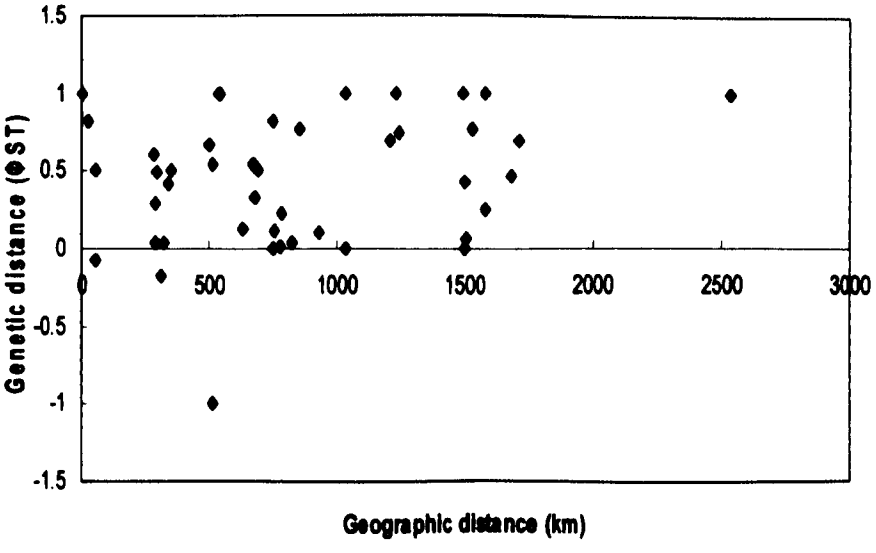


Figure 3.5 Relationship between the geographic distance (km) Vs genetic distance (Φ_{ST}) between Daubenton's populations across Europe.



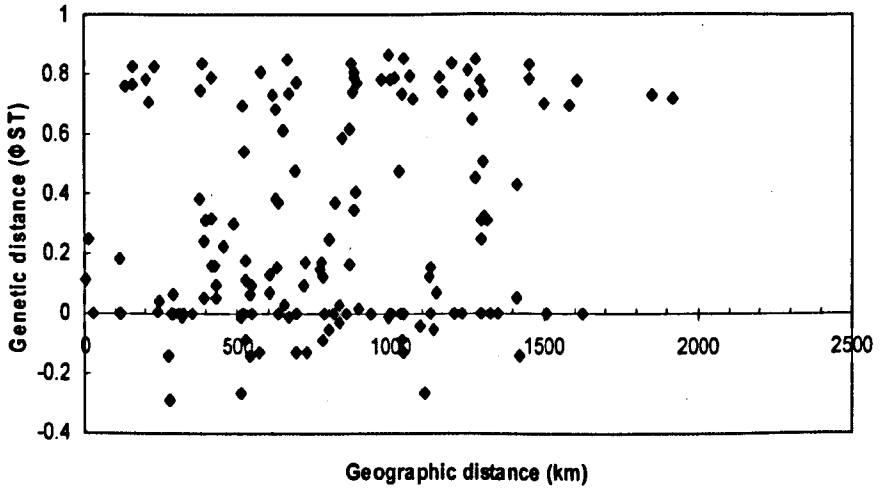


Figure 3.6 Plot of geographic distance (km) Vs genetic distance (Φ_{ST}) between Daubenton's populations across Europe for a) Clade A b) Clade B

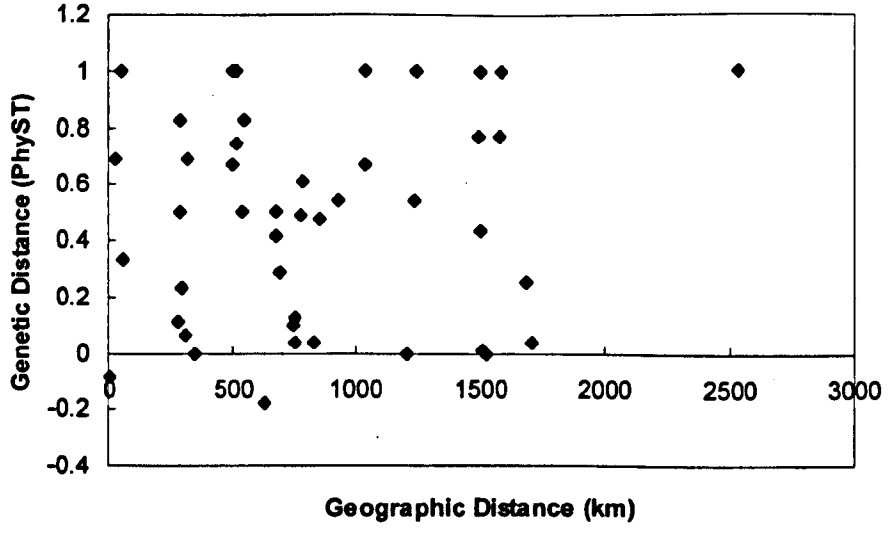


Figure 3.7 Plot of the geographic distance (km) Vs genetic distance (Φ_{ST}) between Daubenton's populations across England.

3.4 Discussion

3.4.1 Phylogenetic inferences

3.4.1.1 How divergent are populations of *Myotis daubentonii* in Europe?

This study has highlighted two intra-specific mitochondrial clades across western Europe. Individuals from populations in south west Europe contained haplotypes associated with clade B and individuals from north east Europe contained haplotypes

associated with clade A. Where these two regions met, in the UK and Belgium, populations consisted of individuals belonging to both clades. The minimum spanning network of *ND1* haplotypes provided additional support for this split with ten interior nodes between the two clades suggesting deep divergence. However, within each clade there was little divergence between haplotypes. The level of sequence divergence between the two clades was 2.4% which is within the intra-specific range for bats (1-6.8% for the *ND1* gene) (Kerth *et al.* 2008) and certainly far less than the sequence divergence between *M. daubentonii* and the closely related *M. bechsteinii* (11.6% divergence, this study).

A recent study on Daubenton's populations across Scotland also found two mitochondrial clades (Ngamprasertwong *et al.* 2008) which are consistent with the two clades in this study. Our south western clade B was the same as their clade B in southern and far northern Scotland and our north eastern clade A was the same as their clade A in central Scotland. A further phylogeographic study of *M. daubentonii* across Europe (Mayer and von Helversen 2001) found four mitochondrial clades with the most basal clade the same as clade B in the Scottish study (i.e. south west clade from this study) and the more derived the same as clade A in the Scottish study (the north east clade in this study) (Ngamprasertwong *et al.* 2008). Mayer & von Helversen (2001) also found that within their clades there was little divergence between *M. daubentonii* individuals from across widely dispersed European sites, which is in support of this study that found little divergence within clades A and B, irrespective of geographic location.

However, although the phylogenetic analyses highlighted two mitochondrial clades, the processes which resulted in the formation of this phylogeographic discontinuity should be considered carefully. Phylogeographic discontinuities of this kind are usually interpreted to represent a signature of past isolation between populations as a consequence of geographic barriers, behavioural differences or past fragmentation with subsequent contact. However theoretical and empirical work has demonstrated that because genealogies are influenced by chance, such patterns may also occur in the absence of any barriers to gene flow (Bridle *et al.* 2004, Irwin 2002). This may be especially relevant for single locus markers, inherited from one parent (as is the case for mtDNA) that evolve through bifurcating genealogies, which can produce two divergent clades irrespective of barriers to gene flow (Barton & Wilson 1995).

Because clade A was predominantly found in north eastern Europe and clade B in south western Europe it is tempting to suggest that there were at least two refugial populations of *M. daubentonii* in Europe which were separated in the past but came into contact when climatic conditions across Europe became suitable. However nested clade analysis (NCA) highlighted isolation by distance as the most likely process responsible for the observed pattern of genetic variation in this study. Notably NCA found no evidence for past fragmentation. This presents the possibility that all individuals sampled belong to one large European population, with individuals belonging to either clade A or clade B by virtue of the large geographic distance between them.

Each scenario will be considered in turn.

3.4.1.2 Scenario one: Demographic expansion of *M. daubentonii* from two glacial refugia in Europe

The sequence divergence level between mitochondrial clades A and B of 2.4% suggests that the split between clades preceeded the last glacial maximum 18,000 years ago, although assuming a mutation rate of 2% per million years this still places a split within the Pleistocene (Petit *et al.* 1999). Coalescent analyses, mismatch distributions and negative values for Tajima's D and Fu's F_S statistics are indicative of a demographic expansion of *M. daubentonii* populations in Europe during the current interglacial. The estimated timing of this expansion for Europe as a whole (using data obtained from mismatch distributions) is 132,000 years bp and for clade B it is 21,000 years bp. The finding of fossil remains of *M. daubentonii* in Holocene deposits across Europe suggests that the species was certainly widespread in Europe by the Holocene climatic optimum 5,500-6,000 years bp (Ochman 2003). In addition it should be noted that high levels of population structure (as demonstrated in this study by high Φ_{ST} values), which may occur if effective migration rates are low during an expansion can account for a bimodal mismatch distribution (Ray *et al.* 2003).

The minimum spanning network suggests Haplotype 0 was ancestral in clade B and indeed this haplotype is widespread and common throughout Daubenton's populations in south west Europe. A number of studies on both bats and other taxa have identified Iberia as an important refugium (Branco *et al.* 2008, Juste *et al.* 2009, Ruedi & Castella

2003, Veith *et al.* 2004). However samples from the two Spanish sites in this study were represented by Haplotype 0 only from clade B which was in contrast to all other populations in south west Europe which consisted of at least two haplotypes. Furthermore the two Spanish populations exhibited low levels of nucleotide diversity. This is in contrast to classic theory which predicts that ancestral populations should harbour greater genetic variability than more derived populations due to the loss of variation during northwards colonisation (Hewitt 1996, but see Petit *et al.* 2003). In the absence of samples from other potential southern refugia (southern Spain and the Balkans), it is impossible to say which of these regions acted as refugium for *M. daubentonii* clade B. Indeed it is possible that the sample sites in Spain were located too far north to pick up any signature of Iberian refugia or that there were multiple refugia in Iberia which remain unsampled.

The wide distribution of Haplotype 0 indicates that the Pyrenees were not a barrier to colonisation for *M. daubentonii* (note that this is still relevant whether past fragmentation or IBD accounts for the presence of two clades). Similarly the Pyrenees were not a barrier to dispersal for other vagile species such as the Greater mouse eared bat (Ruedi *et al.* 2008), the European hedgehog (Santucci *et al.* 1998) and the European Brown bear (Taberlet & Bouvet 1994), only acting as a barrier for species with restricted motility such as *Chorthippus* grasshoppers (Cooper *et al.* 1995), alder (King & Ferris 1998) and beech (Demesure *et al.* 1996). But there are exceptions, for example there is a hybrid zone in the Pyrenees for the chiffchaff, which is a species that does have a high dispersal capacity (Salomon 1989).

Although haplotypes from Italy grouped into clade B, they differed from the ancestral haplotype by at least two steps. Furthermore all haplotypes in Italy were unique to Italy but some were shared between the two Italian populations despite a geographic distance between the two sites of 418 km. This suggests that Italy may have been a minor refugium for Daubenton's, which still retains a degree of isolation from other *M. daubentonii* populations in Europe. A lack of sampling from Italy in previous phylogeographic studies prevents many comparisons from being drawn but the Italian peninsula has been invoked as a major refugium for *Myotis myotis* (Ruedi *et al.* 2008). The Alps have acted as a significant isolating barrier for a number of other mammalian species (Heckel *et al.* 2005, Taberlet *et al.* 1998) although not for birds where genetic discontinuities are not apparent (Bruderer & Jenni 1991). It is plausible to suggest that

populations in Italy originated from a refugium in the Balkans. In the past the level of the Adriatic sea was far lower than it is now (Shackleton 1986) and served as a colonization route for many other species including the brown bear (Taberlet & Bouvet 1994), rock partridge (Lucchini & Randi 1998), brown trout (Splendiani *et al.* 2006) and pond turtle (Fritz *et al.* 2005). However, we would require samples from the Balkan region to further explore this hypothesis.

The only two populations consisting exclusively of clade A individuals were found in Denmark and Finland. Both populations had high haplotype diversity but relatively low nucleotide diversity. Again in the absence of samples from regions potentially acting as refugia for north eastern Europe it is impossible to say where the ancestral individuals from clade A originated. *M. daubentonii* is found continuously from western Europe to Japan (Michell-Jones *et al.* 1999) so there are a number of possibilities for refugia including the Balkans, the Carpathians and central Asia (Deffontaine *et al.* 2005). There has been little consideration of Scandinavia in European phylogeographic work on bats making comparison with other species difficult. However a recent Eurasian study on *Rhinolophus ferrumequinum* (Flanders *et al.* 2009) found strong evidence that Asia Minor and eastern Europe acted as ancient glacial refugia and sources of colonisation into western Europe. Studies on rodents also suggest that populations in northern Europe are likely to have originated from Asian refugia rather than from classic southern European refugia (Brunhoff *et al.* 2003, Jaarola & Searle 2002).

3.4.1.3 Colonisation of the British Isles

Populations from Britain and Belgium shared haplotypes that were common to both clades A and B. As Britain and Belgium lie in a position between south west and north east Europe this pattern is likely attributed to post-glacial admixture between clade A and clade B populations, which had accumulated high levels of genetic divergence during isolation (Petit *et al.* 2003). Multiple patterns of recolonization in this manner have been found in other taxa in the UK including the common frog (Knopp & Merila 2008), water vole (Piertney *et al.* 2005), common shrew (Searle & Wilkinson 1987), bank vole, field vole, pygmy shrew (Searle *et al.* 2009) and oak (Cotrell *et al.* 2002). This sharing of haplotypes between the UK and mainland Europe suggests that the English Channel and the North Sea have not acted as a barrier to past dispersal events. Similarly the presence of the ancestral haplotype for clade B in Northern Ireland

suggests that the Irish Sea has also not proved a barrier to colonization. Large bodies of water do not appear to be a barrier to colonization in other bat species (*Pipistrellus pipistrellus* (Racey *et al.* 2007), *Plecotus auritus* (Juste *et al.* 2004), *Eptesicus serotinus* (Juste *et al.* 2009), thus it is unsurprising that *M. daubentonii* has colonized the British Isles from the continent.

The formation of the late glacial land bridge between the English Channel and the continent provides one possible route of colonization and has aided dispersal to the UK for organisms with far lower dispersal abilities than bats such as *Chorthippus* grasshoppers (Cooper *et al.* 1995). In addition, previous phylogeographic work on Daubenton's has alluded to a Lusitanian connection of Daubenton's populations during the last glacial maxima as evidenced by the presence of a haplotype found in the UK, Ireland, Sweden, Germany and Portugal (Ngamprasertwong *et al.* 2008). That haplotype is equivalent to Haplotypes from clade B in this chapter suggesting that this may be a plausible colonization route. Lusitanian influences have also been suggested for Leisler's bat (Boston *et al.* 2007) and the Natterjack toad (Rowe *et al.* 2005) and may be a consequence of human influences, such as transport along major trade routes (Yalden 1982). Again a lack of sampling farther south than north Spain prevents any firm conclusions from being drawn.

3.4.1.4 Scenario two: isolation by distance across Europe

Because the dispersal capabilities of a species are usually less than the geographic range over which it is found we might expect that, as populations become further away geographically, that they will also become more genetically distant (Balloux & Lugon-Moulin 2002). Such a process can produce phylogeographic breaks which are a consequence of the lack of gene flow between individuals from areas separated by a large geographic distance. This may be especially relevant if the species is continuously distributed but the sampling intensity has been low. Through a series of simulations, Irwin (2002) demonstrated that phylogeographic breaks in a continuously distributed population could occur in the absence of any real barrier to gene flow. This became increasingly likely the smaller the N_e and the lower the N_m of the species in question. Bridle *et al.* (2004) also conducted analyses on toad populations in Sulawesi to show that a simple pattern of isolation by distance could just as easily explain the observed patterns of genetic variation as could past fragmentation in areas of endemism. Whilst it

is tempting to find patterns in the data which are associated with *a priori* expectations, these studies demonstrate the pitfalls of doing so. A recent paper by Chen *et al.* (2010) investigating the population history of *Anopheles* mosquitoes in Asia recognised that multiple hypotheses (including IBD) could produce the patterns observed in their data. Through rigorous testing of the possible demographic scenarios and comparisons with other species from the same region they were able to demonstrate that past fragmentation followed by range expansion was the most likely scenario, but again this highlights the importance of considering multiple hypotheses.

Nested clade analysis invoked a pattern of isolation by distance to explain how the geographic distribution of Daubenton's populations had affected the genetic structure across Europe. This was also supported by a significant Mantel test across the entire study area, but not by Mantel tests within clades A and B that found no evidence for IBD at a smaller geographic scale. Isolation by distance over a large geographical area may be attributed to the colonisation history of a species, rather than contemporary gene flow, with females forming new breeding colonies in a stepwise fashion away from the source population (Bossart & Prowell 1998, Pogson *et al.* 2001). A pattern of isolation by distance across Europe has been found for a number of other species including the Eurasian badger (Pope et al 2006), the eider duck (Tiedeman *et al.* 2004), the European rabbit (Branco *et al.* 2008), the greater horseshoe bat (Rossiter *et al.* 2000) and the greater mouse eared bat (Castella *et al.* 2001). However IBD is not apparent in the colonization patterns of some European bat species. For example Bechstein's bat demonstrates no pattern of IBD for mitochondrial markers, which is probably a consequence of extreme female philopatry (Kerth *et al.* 2002). Conversely there is a lack of evidence for IBD in the serotine which is a product of its high dispersal capacity and preference for flying in open areas (Juste *et al.* 2009).

Given the evidence for IBD, the possibility that individuals in this study belong to a single population cannot be discounted. However phylogeographic breaks under IBD are more likely to occur when N_e and N_m are small. Although the N_e for mtDNA is one quarter of autosomal N_e past population surveys (Harris *et al.* 2006), genetic work (Angell 2008) suggests that *M. daubentonii* populations are large. Although N_m of females may be low due to female philopatry (see section 3.4.3.1 for further discussion) *M. daubentonii* individuals do have high dispersal capabilities, as they are capable of powered flight. Furthermore although a Mantel test across the entire study area was

significant, it was not significant within clades. A pattern of IBD over a large geographic scale that is not apparent on a smaller scale can occur due to mixing of individuals from divergent populations. This would lend more support to the idea of two populations separated by past fragmentation, rather than one population with IBD.

Despite the fact that NCPA reported an inference of IBD at both the overall cladogram and lower nesting levels these results should be interpreted with caution. NCPA has been criticised as local processes operating in a population can result in false inference at a range-wide scale. For example applying NCPA in an analysis of simulated data sets Panchal & Beaumont (2008) found a high false positive rate of inference, with the most common outcomes being restricted gene flow with isolation by distance and contiguous range expansion. Isolation by distance is also the most frequently reported inference made on published empirical data sets (Petit 2008). In addition the fact that some individuals from the north eastern clade A were found in Spain could partially explain why NCPA found IBD at a number of nesting levels. Work reported in later chapters that consider multiple markers with biparental inheritance should afford a greater insight into the likelihood of IBD versus past fragmentation.

This study found no significant isolation by distance for Daubenton's populations in England. An absence of isolation by distance at a regional scale is usually taken to signify gene flow between local populations (Balloux & Lugon-Moulin 2002). However looking at England as a whole it seems more likely that the lack of IBD was an artefact of a number of potential factors including insufficient sampling localities between the northern-most point in Yorkshire and the southern-most point in Sussex and insufficient time for UK populations to have reached equilibrium between gene flow and genetic drift (Slatkin 1993).

3.4.2 Notes of caution

Mitochondrial DNA markers have been used extensively over the last few decades to investigate intra and inter-specific phylogenetic relationships and to make inferences about the demographic history of a species (Avise 1987). However results from mtDNA markers alone should be interpreted with caution, and many have argued that inferences made from mtDNA should be corroborated with nuclear markers (Ballard & Whitlock 2004, Rubinoff & Holland 2005). Because mtDNA does not undergo

recombination, sampling is essentially from a single locus. Using a single locus to investigate population history is associated with a large sampling error due to the stochastic nature of the coalescent process. Therefore phylogenetic inferences made from mitochondrial data are only one realisation of the coalescent process which may not be congruent with population history (Edwards & Beerli 2000, Funk & Omland 2003, Nichols 2001). This is further complicated for mtDNA as it behaves as a single linked locus and so patterns observed may not be a consequence of purely demographic processes (Balloux 2010). This study used the *ND1* gene to investigate population history. The *ND1* gene is associated with metabolic function and it is therefore likely to be strongly influenced by selection. For both mitochondrial clades A and B, Tajima's D and Fu's F_S statistics were significantly negative, meaning that we could not accept the null hypothesis of neutral evolution. Although this could represent a signature of rapid demographic expansion it could also mean that patterns and processes observed were highly influenced by selective forces acting on the *ND1* gene, rather than demographic processes (Balloux 2010). The use of multiple, unlinked nuclear loci in later chapters should allow us to determine with greater certainty whether this is the case. Although some studies comparing nuclear and mitochondrial markers report conflicting interpretations (Johnson *et al.* 2003) many others appear to broadly corroborate findings based on mtDNA (Brito *et al.* 2007, Flanders *et al.* 2009, Toon *et al.* 2007).

3.4.3 Contemporary gene flow

3.4.3.1 What is the degree of differentiation between populations? What does this tell us about female gene flow between colonies?

The mean Φ_{ST} value in this study was 0.55, demonstrating high levels of differentiation between populations. Because mtDNA is maternally inherited this is indicative of strong female philopatry and is in concordance with Φ_{ST} values for other *Myotis* species (Castella *et al.* 2001, Kerth *et al.* 2002). However, the degree of genetic structure between individual populations was highly variable in this study and the overall high Φ_{ST} value may be an artefact of comparing populations containing haplotypes from divergent mitochondrial clades rather than a reflection of current female gene flow. However, high Φ_{ST} values within clade A (0.42) and clade B (0.59) also support restricted female gene flow.

Within clade A it is apparent that the only region significantly differentiated from every other region is Belgium. This is in contrast to clade B which suggests a lack of differentiation between Belgium and the UK. This suggests that Belgium received clade A colonists from a southern European route, whereas the UK was colonised from northern Europe via a North Sea route.

In general, populations within the same region showed low levels of genetic differentiation between each other, in comparison to the mean Φ_{ST} value. This is in agreement with the AMOVA results which only attributed 7.4% of genetic variation to among site differences. Other work investigating local population structure of *M. daubentonii* has suggested that, although females are philopatric, the level of philopatry is not strict (Angell 2008, Ngamprasertwong *et al.* 2008). This suggests that at a local scale female gene flow may be a significant force. A potential explanation for this is swarming behaviour whereby individuals from different summer roosts undergo a seasonal migration to hibernacula where they mate. Swarming behaviour is a common phenomena in temperate bat species (Altringham 1996, Glover & Altringham 2008) and genetic studies show that bats at swarming sites may come from many summer roosts and that bats from the same summer roost visit a number of different swarming sites (Angell 2008, Rivers *et al.* 2005) . Consequently, it is possible that individuals captured at different swarming sites within a local region originated from the same summer colony and thus have high levels of relatedness.

Across Europe as a whole and within clades AMOVA attributed a large portion of variation to differentiation among regions over a wide geographic scale. Strong differentiation between colonies in distinct regions has been reported for female bats of other species including *M. bechsteinii* (Kerth *et al.* 2002) and *M. myotis* (Castella *et al.* 2001). This is a consequence of female philopatry, as females are highly dependent on roosts with stable temperatures and quality foraging areas to meet the energetic demands of pregnancy and lactation (Racey & Swift 1981), they often return to the roost where they were born which constrains the distance they move away from their natal colony. This is supported not only by genetic analyses but also ringing studies (Angell 2008, Kapfer *et al.* 2008, Rieger *et al.* 1996).

It should be noted that because the majority of tissue sampling was conducted at swarming sites most of the samples included in mtDNA analyses were male. This can

have important consequences when making interpretations regarding contemporary sex-biased gene flow. Because mtDNA is maternally inherited it is assumed that mtDNA population structure is reflective of female-specific processes. However if males are sampled which have dispersed away from the natal range then this could have the effect of altering mtDNA population structure such that the underlying pattern of structure due to female genetic drift and gene flow is blurred as a consequence of male movement in the current generation. For example Chen *et al.* (2008) examined mtDNA population structure for both sexes separately in the Formosan horseshoe bat and found greater population structure in females than males which they interpreted as more extreme female philopatry, with males remaining broadly faithful to maternal roosts over a larger geographic scale. Therefore whilst strong mtDNA structure was found, it is possible that if samples were female, different levels of population structure would have been demonstrated that would have altered conclusions made about levels of sex-biased dispersal.

3.4.4 Summary

This chapter has highlighted the importance of investigating the signature of mtDNA variation across different geographic scales to understand how both contemporary and historic process may have acted to shape genetic variation. The finding of two divergent mitochondrial clades suggest that geographic features have played an important role in shaping patterns of variation in *M. daubentonii*, however it is not clear whether this is a consequence of isolation by distance or past fragmentation during glacial cycles. Because mtDNA is a single locus marker genealogies are highly stochastic and only present one realisation of the coalescent process. The use of multiple unlinked loci in later chapters should enable us to better discriminate between the processes responsible for the two mitochondrial clades.

On a contemporary time scale high levels of differentiation between colonies suggests that female philopatry is a significant force in maintaining population structure and high levels of variation between regions indicates that regional populations are relatively isolated in terms of female dispersal. Thus geographic features affect historical and contemporary processes differentially and as a consequence working at different scales allows us to understand genetic patterns that we see in the present. High levels of female population structure in *M. daubentonii* suggest that it may be male dispersal

between colonies that is the major force in current gene flow. However as mtDNA is maternally inherited information from markers with other modes of inheritance and mutational dynamics are needed to allow us to more fully explore both sex-biased dispersal patterns and current gene flow.

3.5 References

Altringham JD (1996) *Bats: Biology & Behaviour.* Oxford University Press, Oxford.

Angell RL (2008) Population ecology and mating system of Daubenton's bat (*Myotis daubentonii*) PhD Thesis, University of Leeds.

Austad SN, Fischer KE (1991) Mammalian aging, metabolism and ecology: Evidence from the bats and marsupials. *Journal of Getontology* **46** 47-53.

Avise JC (2000) *Phylogeography: the history and formation of species,* Harvard University Press.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology & Systematics* **18** 489-522.

Ballard JWO, Whitlock MC (2004) The incomplete history of the mitochondria. *Molecular Ecology* **13** 727-744.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11** 155-165.

Balloux F (2010) The worm in the fruit of the mitochondrial DNA tree. *Heredity* **104** 419-420.

Barton NH, Wilson I (1995) Genealogies and geography. *Philosophical Transactions of the Royal Society of London B* **349** 49-59.

Bosart JL, Prowell D (1998) Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *TREE* **13** 202-206.

Boston E, Prodohl P, Montgomery A (2007) Mitochondrial phylogeography of the Leisler's bat (*Nyctalus leisleri*) in Ireland and across Europe: new perspectives. *Bat Research News* **48** 184.

Branco M, Monnerot M, Ferrand, N Templeton AR (2008) Postglacial dispersal of the European rabbit (*Oryctolagus cuniculus*) on the Iberian peninsula reconstructed from nested clade and mismatch analyses of mitochondrial DNA variation. *Evolution* **56** 792-803.

Bridle JR, Pedro PM, Butlin RK (2004) Habitat fragmentation and biodiversity: testing for the evolutionary effects of refugia. *Evolution* **58** 1394-1396.

Brito PH (2007) Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (*Strix aluco*). *Molecular Ecology* 16 3423-3437.

Bruderer B, Jenni L (1991) Migration across the Alps. In: Gwinner E (ed) *Bird Migration.* Springer, Berlin Heidelberg New York, pp 60-77.

Brunhoff C, Galbreath KE, Feorov VB, Cook JA, Jaarola M (2003) Holarctic phylogeography of the root vole (*Microtus oeconomus*): implications for late Quaternary phylogeography of high latitudes. *Molecular Ecology* 12 957-968.

Burland TM, Barratt EM, Beaumont MA, Racey PA (1999) Population genetic structure and gene flow in a gleaning bat, *Plecotus auritus*. *Proceedings of the Royal Society of London, Series B* 266 975-980.

Burland TM, Worthington-Wilmer J (2001) Seeing in the dark: molecular approaches to the study of bat populations. *Biological Reviews* 76 389-409.

Castella V, Ruedi M, Excoffier L (2001) Contrasted patterns of mitochondrial and nuclear structure among nursery colonies of the bat *Myotis myotis*. *Journal of Evolutionary Biology* 14 708-720.

Chen B, Pedro PM, Harbach RE *et al.* (2010) Mitochondrial DNA variation in the malaria vector *Anopheles minimus* across China, Thailand and Vietnam: evolutionary hypothesis, population structure and population history. *Heredity* in press.

Chen SF, Rossiter SJ, Faulkes CG, Jones G (2006) Population structure and demographic history of the endemic Formosan lesser horseshoe bat (*Rhinolophus monoceros*) *Molecular Ecology* 15 1643-1656.

Chen S, Jones G, Rossiter SJ (2008) Sex-biased gene flow and colonisation in the Formosan lesser horseshoe bat: inference from nuclear and mitochondrial markers. *Journal of Zoology* 274 207-215.

Chesser RK (1991) Influence of gene flow and breeding tactics on gene diversity within populations. *The Genetics Society of America* 129 573-583.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* 9 1657-1660.

Cooper SJB, Ibrahim KM, Hewitt GM (1995) Postglacial expansion and genome subdivision in the European grasshopper *Chorthippus parallelus*. *Molecular Ecology* 4 49–60.

Cotrell JE, Munro RC, Tabener HE, Gillies ACM, Forrest JI, Deans J, Lowe AJ (2002) Distribution of chloroplast DNA variation in British oaks (*Quercus robur* and *Q. petraea*): the influence of postglacial colonisation and human management. *Forest Ecology & Management* 156 181-195.

Deffontaine V, Libois R, Kotlik P *et al.* (2005) Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*). *Molecular Ecology* 14 1727-1739.

Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution* **50** 2515-2520.

Ditchfield AD (2000) The comparative phylogeography of neotropical mammals: patterns of intraspecific mitochondrial DNA variation among bats contrasted to non-volant small mammals. *Molecular Ecology* **9** 1307-1318.

Drummond AJ, Rambaut A (2007) "BEAST: Bayesian evolutionary analysis by sampling trees." *BMC Evolutionary Biology* **7** 214.

Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* **54** 1839-1854.

Excoffier LG, Laval F, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1** 47-50.

Flanders J, Jones G, Benda P *et al.* (2009) Phylogeography of the greater horseshoe bat, *Rhinolophus ferrunquinum*: Contrasting results from mitochondrial and microsatellite data. *Molecular Ecology* **18** 206-318.

Fritz U, Fattizzo T, Guicking D *et al.* (2005) A new cryptic species of pond turtle from southern Italy, the hottest spot in the range of the genus *Emys* (Reptilia, Testudines, Emydidae). *Zoologica Scripta* **34** 351–371.

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

Funk DJ, Omland KE (2003) Species level paraphyly and polyphyly: Frequency, causes and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology Evolution & Systematics* **34** 397-423.

Garrick RC, Dyer RJ, Beheregaray LB, Sunnucks P (2008) Babies and bathwater: a comment on the premature obituary for nested clade phylogeographical analysis. *Molecular Ecology* **17** 1401-1403.

Glover AM, Altringham JD (2008) Cave selection and use by swarming bat species. *Biological Conservation* **141** 1493-1504.

Goodman SJ, Tamate HB, Wilson R *et al.* (2001) Bottlenecks, drift & differentiation: the population structure and demographic history of sika deer (*Cervus nippon*) in the Japanese archipelago. *Molecular Ecology* **10** 1357-1370.

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour* **28** 1140-1162.

Harris SL, Brookes SM, Jones G *et al.* (2006) European bat lyssaviruses: Distribution, prevalence and implications for conservation. *Biological Conservation* **131** 193-210.

Heckel G, Burri R, Fink S, Desmet JF, Excoffier L (2005) Genetic structure and colonization processes in European populations of the common vole, *Microtus arvalis*. *Evolution* **59** 2231–2242.

Hewitt GM (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnaean Society* **58** 247–276.

Hewitt GM (1999) Post-glacial recolonization of European biota. *Biological Journal of the Linnaean Society* **68** 87-112.

Hewitt GM, Butlin RK (1997) Causes and consequences of population structure. In: *Behavioural Ecology: An Evolutionary Approach*, edited by J. R. Krebs and N. B. Davies. pp 203. Blackwell Science Ltd, Oxford.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* **77** 282-291.

Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution* **56** 2383-2394.

Jaarola M, Tegelstrom H (1995) Colonisation history of North European field voles (*Microtus agrestis*) revealed by mitochondrial DNA. *Molecular Ecology* **4** 299-310.

Jaarola M, Searle JB (2002) Phylogeography of field voles (*Microtus agrestis*) in Eurasia inferred from mitochondrial DNA sequences. *Molecular Ecology* **11** 2613-2621.

Jensen JL, Bohonak AJ, Kelly ST (2005) Isolation by distance, web service. *BMC Genetics* **6** 13.

Johnson JA, Toepfer JE, Dunn PO (2003) Contrasting patterns of mitochondrial and microsatellite population structure in fragmented populations of greater prairie-chickens. *Molecular Ecology* **12** 3335-3347.

Juste J, Ibàñez C, Muñoz J *et al.* (2004) Mitochondrial phylogeography of the long-eared bats (*Plecotus*) in the Mediterranean Palaearctic and Atlantic Islands. *Molecular Phylogenetics and Evolution* **31** 1114–1126.

Juste J, Bilgin R, Munoz J, Ibanez C (2009) Mitochondrial DNA signatures at different spatial scales: from the effects of the Straits of Gibraltar to population structure in the meridinol serotine bat (*Eptesicus isabellinus*) *Heredity* **103** 178-187.

Kanuch P, Kristin A (2006) Altitudinal distribution of bats in the Pol'ana Mts area (Central Slovakia). *Biologia* **61** 605-610.

Kapfer G, Rigot T, Holsbeek L, Aron S (2008) Roost and hunting site fidelity of female and juvenile Daubenton's bat *Myotis daubentonii* (Kuhl, 1817) (Chiroptera:Vespertilionidae). *Mammalian Biology* **73** 267-275.

Kerth G, Mayer F, Petit E (2002) Extreme sex-biased dispersal in the communally breeding, nonmigratory Bechstein's bat (*Myotis bechsteinii*). *Molecular Ecology* **11** 1491-1498.

Kerth G, Kiefer A, Trappmann C, Weishaar M (2003) High gene diversity at swarming sites suggest hot spots for gene flow in the endangered Bechstein's bat. *Conservation Genetics* **4** 491-499.

Kerth G, Petrov B, Konti A *et al.* (2008) Communally breeding Bechstein's bats have a stable social system that is independent from the postglacial colonization history and location of the populations. *Molecular Ecology* 13 2358-2381.

King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology* 4 95–103.

Knopp T, Merila J (2008) The postglacial recolonization of Northern Europe by *Rana arvalis* as revealed by microsatellite and mitochondrial DNA analyses. *Heredity* 102 174-181.

Kunz TH (1982) Roosting ecology of bats. In: *Ecology of Bats*, edited by T. H. Kunz. Pp 1-55. Plenum Press, New York and London.

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

Lopez de Heredia U, Carrio´n S, Jime´nez P, Collada C, Gil L (2007). Molecular and palaeoecological evidence for multiple glacial refugia for evergreen oaks on the Iberian Peninsula. *Journal of Biogeography* **34** 1505–1517.

Lucchini V, Randi E (1998) Mitochondrial DNA sequence variation and phylogeographical structure of rock partridge (*Alectoris graeca*) populations. *Heredity* **81** 528–536.

Mayer F, von Helverson, O. (2001) Cryptic diversity in European bats. *Proceedings of the Royal Society of London, Series B* **268** 1825-1832.

McCracken GF, Wilkinson, G.S. (2000) Bat mating systems. In: *Reproductive Biology of Bats*, edited by E. G. Crichton and P. H. Krutzsch. pp 321-362. Academic Press, London and San Diego.

Melnick DJ, Hoelzer GA (1992) Differences in male and female macaque dispersal lead to contrasting distributions of nuclear and mitochondrial DNA variation. *International Journal of Primatology* 13 379-393.

Moore WS (1992) Inferring phylogenies from mtDNA variation: Mitochondrial gene-trees versus Nuclear-gene trees. *Evolution* **49** 718-726.

Nabholz B, Glemin S, Galtier N (2008) Strong variations of mitochondrial mutation rate across mammals – the longevity hypothesis. *Molecular Biology & Evolution* **25** 120-130.

Ngamprasertwong T, Mackie IJ, Racey PA, Piertney SB (2008) Spatial distribution of mitochondrial and microsatellite DNA variation in Daubenton's bat within Scotland. *Molecular Ecology* **17** 3243-3258.

Nichols R (2001) Gene trees and species trees are not the same. *TREE* **16** 358-364.

Ochman K (2003) Late Pleistocene and Holocene bats (Chiroptera) from the Komarowa Cave (Cracow-Częstochowa Upland, Poland) – preliminary results. *Acta zoologica cracoviensia* **46** 73-84.

Panchal M, Beaumont MA (2007) The automation and evaluation of nested clade phylogeographic analysis. *Evolution* **61** 1466–1480.

Peng PK, Sutton SL, Fletcher CR (1992) Spatial and temporal distribution patterns of flying Diptera. *Journal of Zoology* **228** 329-340.

Pereira MJR, Salgueir P, Rodrigues L, Coelho MM, Palmeirim JM (2009) Population structure of a cave-dwelling bat, *Miniopterus schreibersii:* Does it reflect history and social organisation? *Journal of Heredity* **100** 533-544.

Petit E, Excoffier L, Mayer F (1999) No evidence of bottleneck in the postglacial recolonization of Europe by the noctule bat (*Nyctalus noctula*). *Evolution* **53** 1247–1258.

Petit RJ, Aguinagalde I, de Beaulieu JL *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300** 1563-1565.

Petit RJ (2008) The coup de grâce for the nested clade phylogeographic analysis? *Molecular Ecology* **17** 516–518.

Pierpaoli M, Riga F, Trocchi V, Randi E (1999) Species distinction and evolutionary relationships of the Italian hare (*Lepus corsicanus*) as described by mitochondrial DNA sequencing. *Molecular Ecology* **11** 1805–1817.

Piertney SB, Stewart WA, Lambin X *et al.* (2005) Phylogeographic structure and postglacial evolutionary history of water voles (*Arvicola terrestris*) in the United Kingdom. *Molecular Ecology* **14** 1435-1444.

Pogson GH, Taggart C, Mesa KA, Boutilier RG (2001) Isolation by distance in the atlantic cod, *Gadus morhua,* at large and small geographic scales. *Evolution* **55** 131-146.

Pope LC, Domingo-Roura X, Erven K, Burke T (2006) Isolation by distance and gene flow in the Eurasian badger (*Meles meles*) at both a local and broad scale. *Molecular Ecology* **15** 371-386.

Powell JR, Caccone A, Amato JD, Yoon C (1986) Rates of nucleotide substitution in Drosophila mitochondrial DNA and nuclear DNA are very similar. *PNAS* **83** 9090-9093.

Purdue JR, Smith MH, Patton JC (2000) Female philopatry and extreme spatial genetic heterogeneity in white-tailed deer. *Journal of Mammology* **81** 179-185.

Racey PA, Swift SM (1985) Feeding ecology of *Pipistrellus pipistrellus* (Chiroptera: Vespertilionidae) during pregnancy and lactation, I, Foraging Behaviour. *Journal of Animal Ecology* **54** 205-215.

Racey PA, Barratt EM, Burland TM *et al.* (2007) Microsatellite DNA polymorphism confirms reproductive isolation and reveals differences in population genetic structure

of cryptic Pipistrelle bat species. *Biological Journal of the Linnaean Society* **90** 539–550.

Ray N, Currat M, Excoffier L (2003) Intra-deme molecular diversity in spatially expanding populations. *Molecular Biology & Evolution* **20** 76-86.

Rieger I (1996) How do Daubenton's bats, *Myotis daubentonii*, use their day roosts? *Mammalian Biology* **61** 202-214.

Rivers NM, Butlin RK, Altringham JD (2005) Genetic population structure of Natterer's bats explained by mating at swarming sites and philopatry. *Molecular Ecology* **14** 4299-4312.

Rogers AR, Harpending H (1992) Population-growth makes waves in the distribution of pairwise genetic-differences. *Molecular Biology and Evolution* **9** 552–569.

Rossiter SJ, Jones G, Ransome RD, Barratt EM (2000) Genetic variation and population structure in the endangered greater horseshoe bat, *Rhinolopus ferrumequinum*. *Molecular Ecology* **9** 1131-1135.

Rossiter RJ, Benda P, Dietz C, Zhang S, Jones G (2007) Rangewide phylogeography in the greater horseshoe bat inferred from microsatellites: implications for population history, taxonomy and conservation. *Molecular Ecology* **16** 4169-4174.

Rowe G, Harris DJ, Beebee TCJ (2005) Lusitania revisited: A phylogeographic analysis of the natterjack toad *Bufo calamita* across its entire biogeographical range. *Molecular Phylogenetics and Evolution* **39** 335-346.

Rubinoff D, Holland BH (2005) Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology* **54** 952–961.

Ruedi M, Castella V (2003). Genetic consequences of the ice ages on nurseries of the bat *Myotis myotis*: a mitochondrial and nuclear survey. *Molecular Ecology* **12** 1527–1540.

Ruedi M, Walter S, Fischer MC, Scaravelli D, Excoffier L, Heckel D (2008) Italy as a major ice age refuge area for the bat *Myotis myotis* in Europe. *Molecular Ecology* **17** 1801-1814.

Salomon M (1989) Song as a possible reproductive isolating mechanism between two parapatric forms. The case of the chiffchaffs *Phylloscopus C. collybita* and *P.C. brehmii* in the western Pyrenees. *Behaviour* **111** 720-790.

Santucci F, Emerson BC, Hewitt GM (1998) Mitochondrial DNA phylogeography of European hedgehogs. *Molecular Ecology* **7** 1163-1172.

Searle JB, Wilkinson PJ (1987) Karyotypic variation in the common shrew (*Sorex araneus*) in Britain – a 'Celtic fringe'. *Heredity* **59** 345-351.

Searle JB, Kotlik P, Rambau RV *et al.* (2009) The Celtic fringe of Britain: Insights from small mammal phylogeography. *Proceedings of the Royal Society B – Biological Sciences* **276** 4287-4294.

Senior P, Butlin RK, Altringham JD (2005) Sex and segregation in temperate bats. *Proceedings of the Royal Society of London, Series B* **272** 2467-2473.

Shackleton NJ (1986) Paleogene stable isotope events. *Palaeogeography, Palaeoclimatology, Palaeoecology* **57** 91-102.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47** 264-279.

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

Splendiani A, Giovannotti M, Cerioni PN, Caniglia ML, Caputo V(2006) Phylogeographic inferences on the native brown trout mtDNA variation in central Italy. *Italian Journal of Zoology* **73** 179–189.

Taberlet P, Bouvet J (1994) Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. *Proceedings of the Royal Society B: Biological Sciences* **255** 195–200.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-S (1998) Comparative phylogeography and postglacial colonisation routes in Europe. *Molecular Ecology* **7** 453-464.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123** 585-595.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24** 1596-1599.

Templeton AR, Routeman E, Phillips CA (1995) Separating population structure from population history: A cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum. Genetics* **140** 767-782.

Templeton AR (2004) Statistical phylogeography: methods for evaluating and minimizing inference errors. *Molecular Ecology* **13** 789–809.

Templeton AR (2008) Nested clade analysis: An extensively validated method for strong phylogeographic inference. *Molecular Ecology* **17** 1877-1880.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22 4673-4680.

Tiedeman R, Paulus KB, Scheer M, Von Kistowski KG, Skirnnison K, Bloch D, Dam M (2004) Mitochondrial DNA and microsatellite variation in the eider duck (*Somateira mollissima*) indicate stepwise postglacial colonisation of Europe and limited current long distance dispersal. *Molecular Ecology* **13** 1481-1494.

Toon AT, Mather PB, Baker AM *et al.* (2007) Pleistocene refugia in an arid landscape: analysis of a widely distributed Australian passerine. *Molecular Ecology* **16** 2525-2541.

Urbanczyk Z (1990) Northern Europe's most important hibernation site. *Oryx* **24** 30–34.

Veith M, Mayer C, Samraoui B, Donaire D, Bogaerts S (2004) From Europe to Africa and vice versa: evidence for multiple intercontinental dispersal in ribbed salamanders (genus *Pleurodeles*). *Journal of Biogeography* **31** 159–171.

Wilkinson JS, South GM (2002) Life history, ecology and longevity in bats. *Aging cell* 1 124-131.

Worthington Wilmer J, Moritz C, Hall L, Toop, J. (1994) Extreme population structuring in the threatened ghost bat, *Macroderma gigas*: Evidence from mitochondrial DNA *Proceedings of the Royal Society of London, Series B* **257** 193-198.

Yalden DW (1982) When did the mammal fauna of the British Isles arrive? *Mammal Review* **12** 1-57.

Chapter 4: The development and application of X-linked microsatellite markers to investigate population structure of Daubenton's bat in Europe

4.1 Introduction

The use of genetic markers to examine levels and patterns of genetic diversity within and among populations and to explore the demographic and evolutionary factors which contribute to these patterns is the basis for population genetics (Li 1948). Markers on autosomes (Selkoe & Toonen 2006), mitochondrial DNA (Avise *et al.* 1987) and, in humans, the Y-chromosome (Jobling & Tyler-Smith 2003) have been extensively explored to describe such patterns, make inferences about the history of a population and to explain contemporary patterns of gene flow. In addition to these markers, the X-chromosome contains a potentially large source of information which has remained under-utilised in population genetics (Schaffner 2004).

Yet the X-chromosome has a number of properties which render it an extremely useful source of information. Firstly, as the X-chromosome is present in a single-copy in males it is accessible to haplotyping in males (Schaffner 2004), something which is not yet practical using autosomal markers (but see Kidd *et al.* 2008, Pool *et al.* 2010). Although haplotyping has long been possible using mitochondrial markers (Cann *et al.* 1987), and in males using Y-chromosome markers (Lucotte *et al.* 1989), both marker types have a uniparental mode of inheritance. In contrast the X-chromosome is biparentally inherited and thus has the potential to be informative about both male and female population histories (Bustamante & Ramachandran 2009, Wilkins & Marlowe 2006). In females, the X-chromosome is present in two copies and in this state undergoes recombination. Consequently different regions on the X-chromosome can be analysed which may have different underlying histories (Schaffner 2004). This multi-locus approach should provide a more accurate picture of population history than mtDNA or the Y-chromosome which (Brumfield *et al.* 2003), because they are essentially single locus markers, only crudely record the history of a population and are more likely to be influenced by selection (Balloux 2010). The X-chromosome is thought to be older in evolutionary terms than the Y-chromosome and mutates more slowly than mtDNA (Hedrick 2007). This, combined with the accessibility to multiple loci, means that it is

possible to make inferences about history occurring farther back in time than either mtDNA or Y-chromosome markers allow (Harris & Hey 1999, Yotova *et al.* 2007).

Further, due to the mode of inheritance of the X-chromosome, X-linked markers may potentially record more marked signatures of population structure than autosomal markers (Schaffner 2004, Vicoso & Charlesworth 2006). The X-chromosome is present in two copies in females and a single-copy in males; this means that under neutral expectations the effective population size of the X-chromosome should be three quarters that of autosomes (Laporte & Charlesworth 2002). As genetic drift is more pronounced the smaller the N_e, this will act to reduce genetic diversity on the X-chromosome (Ellegren 2009, Hedrick 2007). As the X-chromosome spends proportionally more time in females than males the mutation rate will be reduced, as in most species females have a lower mutation rate than males (Bartosch-Harlid *et al.* 2003, Ellegren 2007, Li *et al.* 2002). In combination these factors act to reduce diversity on the X-chromosome, which may allow population history to be recorded more accurately due to the loss of shared ancestral polymorphisms and hence result in more pronounced genetic structure than when using autosomal markers (Pereira *et al.* 2007, Yu 2002). However, it should be noted that when behavioural differences are taken into account deviations from this expectation may result (Caballerro 1994, Charlesworth 2001). For example if female gene flow is greater than males, this increases the N_e of the X-chromosome relative to that of autosomes (Keinan *et al.* 2009).

Finally as the X-chromosome only recombines when it is in females, it has a lower recombination rate than autosomes and thus signatures of linkage disequilibrium, as a consequence of demographic history, may be more pronounced (Kong *et al.* 2002, Laan *et al.* 2005, Pool *et al.* 2006). However, it must be noted that a higher overall recombination rate on the X-chromosome can act to counter this trend (Nachman 1998).

Despite the aforementioned properties and the ease of haplotyping it is only recently that the X-chromosome has begun to be used more extensively as a marker to investigate human population history (Schaffner 2004). Studies on humans have tended to focus on the genetic history of modern human lineages, either globally (Jaruzelska *et al.* 1999, Yu 2002) or within regions in comparison to African populations, to investigate the traditional 'out of Africa' hypothesis of modern human origins (Lambert *et al.* 2010, Templeton 2002). The X-chromosome has sometimes proved a better

marker, tracing human history farther back in time than global studies using Y-chromosome or mitochondrial markers had previously allowed (Bourgeois *et al.* 2009). Patterns of diversity and linkage disequilibrium within ethnic groups have allowed population structure and levels of admixture to be determined (Laan *et al.* 2005). Studies comparing X-chromosome Vs autosomal patterns of genetic diversity are increasing, as any deviation from the expected N_e ratio of 0.75 may be a consequence of demographic processes (Ellegren 2009, Pool & Nielsen 2007), such as population expansion, or social processes, for example when there is variance in male and female reproductive success (Hammer *et al.* 2008, Kienan *et al.* 2009). As previously mentioned, because the X-chromosome and autosomes provide access to many unlinked markers they record multiple aspects of a population history, and therefore X:Autosomal comparisons have the power to be more informative than traditional MT:Y comparisons when comparing male- and female- specific processes (Wilkins & Marlowe 2006).

Outside humans (and model organisms such as *Drosophila* species (Singh *et al.* 2007, Visco & Charlesworth 2009)), work on the X-chromosome has mainly focussed on speciation and evolutionary history and is confined to a handful of studies. Brandlii *et al.* (2005) reconstructed the history of the greater white-toothed shrew using a combination of X-linked, Y-linked and mitochondrial markers. However, X-linked variation was too low to be informative in any capacity. Further work on mouse and hare species has used the X-chromosome to investigate admixture and introgression between closely related species within Europe (Geraldes *et al.* 2005, 2008, Melo-Ferreira *et al.* 2009). However, it appears the X-chromosome was not involved in introgression, probably as a consequence of X-linked hybrid male unfitness. Most recently a population genomics approach was used to investigate linkage disequilibrium patterns on the X-chromosome of an inbred population of Scandinavian wolves (Hagenblad *et al.* 2009).

For their body size, bats are long lived and share life history characteristics typically attributed to much larger mammals (Wilkinson & South 2002). The capacity for powered flight means that over large geographic distances genetic structure is less pronounced than that of similarly-sized ground-dwelling mammals (Altringham 1996). In addition, there is no typical pattern of sex-biased dispersal or social system among related species (McCracken & Wilkinson 2000). All population genetic studies of bats

to date have investigated population structure using autosomal or mitochondrial markers, or a combination of the two (Burland & Worthington-Wilmer 2001, Flanders *et al.* 2009, Kerth *et al.* 2008). There is no published work that I know of on the X-chromosome.

Myotis daubentonii is a medium sized species of bat, with a distribution from Ireland in the west to Japan in the east and a fairly continuous presence throughout. Past work on Daubenton's has shown little genetic structure for autosomal markers either at a local scale (Angell 2008, Ngamprasertwong *et al.* 2008) or at a continental scale (Atterby *et al.* 2009), suggesting that panmixia is common in areas where suitable habitat is present. However, mitochondrial markers show considerably more structuring, which has been attributed to female philopatry (Ngamprasertwong *et al.* 2008). Results from studies on the social and breeding system are more complex as Daubenton's undergo a behaviour termed swarming, where bats gather from a number of different summer colonies to mate (Parsons & Jones 2003), although it has been found that mating also occurs between male bats sharing summer roosts with females (Senior *et al.* 2005). Either way, male reproductive success is likely to be more variable than that of females. Because the X-chromosome is expected to show more pronounced genetic structure than autosomes, which is further increased when males are the major force in dispersal, we may be able to pick up signatures of population structure in both sexes which are not apparent using autosomal markers alone. However, because male gene flow can be more variable this could act to increase the N_e of the X-chromosome, which would decrease our ability to pick up signatures of population history and this should be taken into account.

In this chapter I use nine X-linked markers developed in bats to investigate population structure of *M. daubentonii*, at a continental and a local scale, and to investigate barriers to gene flow between populations at a wide geographic scale. In further chapters results from the X-linked markers will be considered in combination with autosomal and mitochondrial markers to provide an accurate picture of *M. daubentonii* population structure in Europe, and to investigate sex-biased differences in genetic structure, reflecting demographic and social processes.

4.2 Methods

4.2.1 Development of X-chromosome microsatellite markers

4.2.1.1 Isolation of X-linked microsatellite regions

The website Orthomam (http://www.orthomam.univ-montp2.fr/orthomam/html/) serves as a repository for single-copy orthologous sequences from autosomes and the X-chromosome of mammalian genomes. Sequences can be searched according to chromosome membership and the number of species carrying a copy of the orthologous sequence can be specified. The following criteria: sequence present in at least six different mammals, located on the X-chromosome, were used to search for sequences found on the X-chromosome of *Myotis lucifugus* and common to at least 5 other mammal species. This was for the reason that if the sequence was present in another *Myotis* species and conserved across mammalian orders then it had a reasonable chance of being conserved in *M. daubentonii*.

The sequences were converted to FASTA format and a search for repeat regions among the sequences was conducted using SPUTNIK (http://www.cbib.u-bordeaux2.fr/pise/sputnik.html). A minimum repeat length of 20 bp was included in the search criteria and only perfect microsatellite repeats were included to avoid PCR amplification problems.

The sequences containing repeat regions were BLAST searched using the BLASTn algorithm against the *M. lucifugus* genome on Ensembl to identify the contig each sequence was located on. This was done to identify sequences on the same contig. Sequences were subsequently named according to contig membership.

4.2.1.2 Primer design

Any sequences that did not have at least 30 base pairs flanking both sides of the repeat region were excluded as this was deemed insufficient sequence for primers to bind to. Primer3 version 4 (Rozen & Skaletsky 2000) was used to design primers from the remaining sequences. In each case the repeat region was highlighted and the following

criteria used to select potential primer pairs: Primer size 18-27 bp, T_m 55-65 °C, Max T_m difference between pair 0.5 °C, CG clamp. If the primer binding site selected by Primer3 was located on a region containing repeats this region was excluded so that a more suitable binding site could be found.

If Primer3 failed to find any primer binding sites in the flanking sequence the CG clamp was removed and the maximum T_m difference increased to 1.5 °C. If Primer3 still failed to find any binding sites the repeat region was dropped from further investigation. This resulted in 92 primer pairs designed to amplify the X-linked repeat regions.

4.2.1.3 Primer testing

4.2.1.3.1 Unlabelled testing

Primer pairs were named according to the contig the repeat-containing sequence was located on. Ninety seven unlabelled forward and reverse primers were ordered in each case. To test whether the primers produced a PCR product of the expected size a gradient PCR was performed on a MJ Research tetrad thermal cycler with the following conditions: 95 °C for 15 minutes, followed by 35 cycles of 94 °C for 30 seconds, gradient of 55 °C-60 °C for 90 seconds, 72 °C for 60 seconds and a final extension step of 60 °C for 30 minutes. Each PCR constituted a 12 µl reaction containing 2 µl of genomic DNA and 10 µl of Qiagen multiplex kit, which contained 0.2 µM of both the forward and reverse primer.

Each primer pair was tested on three different *M. daubentonii* individuals to ascertain optimal primer annealing temperature.

5 µl of each PCR product was run on a 1.5 % agarose gel, stained with syber safe for 45 minutes at 100 V.

4.2.1.3.2 Labelled testing

Of the unlabelled primer pairs tested, 37 produced a clear band of consistent size. In such instances the forward primer of each pair was fluorescently labelled using either FAM or HEX. The labelled primer pairs were PCR amplified in a HEX/FAM duplex

using 6 individuals (3 males and 3 females) and the following conditions: 95 °C for 15 minutes, followed by 35 cycles of 94 °C for 30 seconds, gradient of 55 °C-60 °C for 90 seconds, 72 °C for 60 seconds and a final extension step of 60 °C for 30 minutes. Each PCR constituted a 12 µl reaction containing 2 µl of genomic DNA and 10 µl of Qiagen multiplex kit, which contained 0.2 µM of both the forward and reverse primer.

The PCR product was diluted in 30 µl of ddH_20 and 0.5 µl of the dilution was mixed with 9.5 µl of formamide, containing ROX size standard. The PCR product dilution was then run on an ABI3730 capillary sequencer (Applied Biosystems).

4.2.1.4 Determination of X-specificity

Genemapper version 3.9 was used to visualise the amplified regions. Any primer pairs producing messy, unreadable banding patterns across all temperatures and individuals were discarded from further testing. The remaining primer pairs (those producing a clear peak of the expected size with little or no background noise) were subsequently tested at the optimum temperature using 11 male and 13 female samples. This was done to test for a) X-specificity and b) levels of polymorphism.

If a primer pair produced heterozygous peaks in males it was assumed that the amplified sequence was located in an autosomal region and not on the X-chromosome and was thus excluded from further testing. Any primer pairs producing monomorphic peaks in all individuals were also excluded. In addition primer pairs producing peaks that were difficult to score were excluded from further testing. This resulted in 8 remaining repeat regions which we were confident were X-linked and were variable in size between individuals (Table 4.1).

The 24 individuals used for testing were all from the same population (Ilkley cemetery). Cervus version 3 (Kalinowski *et al.* 2007) was used to test for the presence of null alleles. In order to gain a more accurate idea of allele size range for the 8 markers 10 individuals from southern France were then tested. This was in addition to 3 microsatellite primers designed by Jan *et al.* (in prep.) that were already proven to be X-linked in bats.

Allele size ranges from 11 X-linked microsatellites used in analyses are given in Table 4.1

Table 4.1 Multiplex conditions and locus information for X-linked microsatellites used in this thesis

Multiplex	Locus	Repeat motif	Allele size range (bp)	Volume in master mix (µl)	Tm	Dye	Reference
1	Efnb	(TG)n	127-132	4	58	Fam	This thesis
1	Mcart	(AC)n	135-163	4	58	Fam	This thesis
1	H4	(TC)n	152-129	2	58	Fam	Jan (2010)
1	Kcne	(AC)n	200-218	1	58	Hex	This thesis
1	Cxcr	(AG)n	211-242	1	58	Hex	This thesis
2	Shroom	(AT)n	126-134	2	58	Fam	This thesis
2	Pou	(AG)n	195-213	1	58	Fam	This thesis
2	Nap	(AGC)n	193-207	1	58	Hex	This thesis
3	Nlgn	(AC)n	184-215	2	58	Hex	This thesis
3	Pcam	(GT)n	221-243	2	58	Ned	Jan (2010)
3	P22	(TC)n	278-304	2	58	Fam	Jan (2010)

4.2.1.5 Multiplex design

The minimum and maximum size range for each locus was plotted using Excel. Pre-PCR multiplexes were designed so that any loci with overlapping size ranges would be labelled with a different dye (FAM, HEX or NED). This resulted in three multiplexes (Table 4.1). Each multiplex was tested on six individuals using a gradient PCR as above (see section 4.2.1.3.2.). Banding patterns were viewed on Genemapper v3.9. In cases where allelic drop out occurred the pre-PCR multiplex was altered so that the concentrations of primers designed to amplify weak or absent alleles were increased relative to the concentration of strong alleles. Multiplex conditions were as above for the gradient PCR, but with annealing temperatures appropriate to each multiplex. For final multiplex conditions, see Table 4.1.

4.2.1.6 Loci dropped from final data analysis

After genotyping all samples, two loci were dropped from the analysis: Cxcr and H4. This was due to a combination of difficulty in scoring these loci due to peak stutter and a high rate of amplification failure.

4.2.2 Investigating population structure using X-linked microsatellites

4.2.2.1 Samples

Sampling locations were selected to give reasonable coverage of Daubenton's samples across the study region (Figure 4.1). The study region comprised samples which were mainly from the UK, with additional sampling from sites in mainland Europe which contain likely source populations for colonisation or current gene flow. However, as permission must be obtained from the relevant authority to collect wing punch samples, it was not possible to sample continuously across the region. Thus sampling points comprised 46 sampling locations from 11 European countries. For purposes of analysis, the sampling locations were grouped into 17 regions, defined as within country populations located no more than 120 km apart. Wing biopsies from adult bats were collected and processed as described in Chapter 2. In addition DNA samples from Scotland, Denmark, Finland and Belgium were provided from outside sources (see Chapter 1 for details). This resulted in samples from 942 adult bats (Table 4.2).

Figure 4.1 Sampling localities of *M. daubentonii* individuals across Europe. Large circle for Yorkshire Dales indicates the region of all sample sites in this area.

4.2.2.2 DNA extraction

DNA was extracted from wing punches using an ammonium acetate precipitation method, adapted to a 96 well plate extraction as described in Chapter 2.

4.2.2.3 Microsatellite amplification

A PCR of each 96 well sample plate was performed on an MJ Research Tetrad thermal cycler for each multiplex. For PCR conditions, see multiplex development section 4.2.1.5. In addition to the samples to be tested each plate contained a negative control of water plus the PCR reagents to rule out contamination. Further, a random sample of 90

DNA extracts was chosen and amplified again for each multiplex to test repeatability of genotyping.

4.2.2.4 Analysis

Microsatellite fragments were sized using Genemapper version 3.9. The data were formatted in Microsoft Excel using Microsatellite toolkit (Park 2001), which tests for invalid alleles and formats data for input into population genetics software programs. There were a lot of non-canonical alleles present in the microsatellite loci used in this study, suggesting flanking region variation and/or imperfect repeats. Consequently, a stepwise model of mutation was deemed inappropriate and further analyses were restricted to using an infinite-alleles model.

4.2.2.5 Linkage Disequilibrium

Linkage Disequilibrium (LD) between pairs of loci was tested both within sample sites (grouped according to region) and across all sample sites in Genepop on the web (http://genepop.curtin.edu.au/) using an MCMC approach with 1000 batches and 10,000 iterations per batch. Tests for LD were restricted to males only, to avoid the problem of mixing haploid and diploid individuals. P-values were adjusted using a sequential Bonferroni correction.

4.2.2.6 Population structure

Due to a small sample number in some localities and close geographical proximity of others, the 46 populations were pooled into 17 groups according to region. Pair-wise F_{ST} values between the 17 regions were calculated using Arlequin version 3.11 (Excoffier *et al.* 2005). To allow the inclusion of female samples in Arlequin the two alleles from each locus in females were separated and treated as haploid.

In addition, an AMOVA test was performed on data from males only to assign the amount of variation due to individual Vs locality Vs regional structure. AMOVA was performed in Arlequin.

To visualise structure among regions a non-metric multi-dimensional scaling (MDS) plot was applied to the F_{ST} matrix.

4.2.2.7 Isolation by distance

Pair-wise F_{ST} and Rousset's $F_{ST}/(1- F_{ST})$ values for each of the original 46 sample populations were calculated using Arlequin. A corresponding matrix of the natural log of pair-wise geographic distance (km) was also produced (calculated from latitude and longitude, using the haversine formula, by Roger Butlin). A plot of Rousset's genetic distance ($F_{ST}/(1- F_{ST})$) Vs the natural log of Euclidean geographic distance (km) was produced and the significance of any correlation analysed using a Mantel test with 1000 permutations, performed using Isolation by distance Web service (http://ibdws.sdsu.edu/~ibdws/).

4.2.2.8 Clustering analyses

Although sampling locations were defined according to a 3 km radius, most sampling locations were from swarming sites meaning they were likely to represent bats from a much larger radius. To assess the number of genetically distinct populations in the study area the clustering programs STRUCTURE (Pritchard *et al.* 2000) and BAPS (Corander *et al.* 2008) were used. STRUCTURE employs a Bayesian clustering algorithm based on reducing deviations from linkage equilibrium and HWE to estimate the number of clusters (K) in the data set. The posterior probability for different values of K is estimated, assuming both HWE and linkage equilibrium within clusters. The latest version of STRUCTURE allows prior sampling location to be given, which assists in cluster assignment (Hubisz *et al.* 2009). This feature is useful when structure is weak, perhaps due to a small number of samples or markers. Furthermore, using prior sampling information does not tend to lead STRUCTURE to find false population structure (Hubisz *et al.* 2009). As we only had nine markers and small sample size in some areas, in addition to the high mobility of bats, it was expected *a priori* that population structure would be weak. Thus STRUCTURE was run both with and without sampling location as a prior and the results compared. The length of the burnin period was 1,000,000 iterations and the number of MCMC repetitions after the burnin was 1,000,000. The admixture ancestry model, with alpha set separately for each population and correlated allele frequencies were used with default settings. The program was run

first without any prior location information and then again using location as a prior. The optimum number of clusters was estimated by performing ten independent runs at K=1-10. Ten was assumed to be the upper limit of clusters within the study area. The true number of clusters present was assumed to be the value of K which maximised the posterior probability of the data, Pr($X|K$). In other words the K with the least negative log likelihood value was preferred. When K increased monotonically the value of K with least variance between runs was chosen as the likely maximal K within the study area. Further, the absolute rate of change of K (ΔK) was calculated following Evanno *et al.* (2005). However, it should be noted that although ΔK reveals the highest level of structure present (i.e. the smallest meaningful value of K), it does not reveal more detailed levels of structure which may be present in the population. Thus results from ΔK and Pr($X|K$) are both considered.

After finding the maximal appropriate K, percentage membership of each cluster within each region was calculated and represented in a pie chart to visualise how clustering is apportioned between regions.

In addition to using STRUCTURE, the program BAPS was used to estimate K. Frantz *et al.* (2009) have shown that it is sensible to use two different methods of estimating population structure in this way, as the underlying assumptions and algorithms used by different programs may differ. Thus if two programs produce different outputs it suggests problems with the data or unclear genetic structure. It may also allow resolution in cases where one program does not produce a definitive answer in terms of the most likely K. Further, BAPS allows the incorporation of geographical sampling information in terms of co-ordinates, whereas assignment of populations or individual locations in STRUCTURE is based on arbitrary categories. As with STRUCTURE, BAPS employs a Bayesian clustering algorithm to estimate K. However rather than estimating K based on reducing deviations from HWE and linkage equilibrium it does so by finding the optimal value of K which groups clusters according to similarity in allele frequencies. Rather than using an MCMC method it uses a greedy stochastic algorithm to search for the most likely K. Also as with STRUCTURE, it is possible to use location information as a prior. Again BAPS was run using no prior location information, using prior information in terms of pre-defined groups of individuals (set by sample identity), and using prior information in terms of pre-defined groups set by sampling location which was defined using latitude and longitude. In each case, BAPS

was run independently, five times each with the maximum K set to 5, 10 and 15. The value of K with the least negative logml value was deemed to be the optimal number of clusters. Congruence of both K and logml between runs and maximum K values was checked.

4.2.2.9 Detecting migrants

Geneclass version 2 (Piry *et al.* 2004) was used to detect any individuals which could be first generation migrants. L = L_home / L_max was selected as the parameter for likelihood computation. An MCMC approach with 10,000 repetitions and Paetkau (2004) re-sampling algorithm was used to detect the probability that each individual is a resident (i.e. not a first generation migrant) of the locality to which it was assigned. Any individuals with a P value below 0.01 were assumed not to be a resident and the probability of assignment to a different population was determined.

This was done for a particular group of localities. Individuals from England and individuals from France were grouped according to country to identify individuals that were not residents of their sampled region.

4.2.2.10 Barrier effects

To test the effect of the English Channel as a barrier to Daubenton's dispersal, whilst eliminating the effect of geographical distance, a distance-based redundancy analysis (dbra) was employed in DISTLM version 5 (Anderson 2004). DISTLM first performs marginal tests on a set of predictor variables to identify which variables are correlated with genetic distance. Each variable is considered independently from other variables. Next conditional tests are performed which allow one to examine the extent to which any of the predictor variables explains genetic diversification better than geographic distance alone. P values were obtained using 9999 permutations of the rows and columns of the multivariate residual matrix. For this analysis only regions in England (Dales, NYM, Withcall, southern England), Belgium and northern France were considered. For the input files genetic distance was defined as a matrix of pair-wise F_{ST} values between the 26 localities within the aforementioned regions (see Table 4.2 for details). A file containing the response variables was generated by assigning sample sites a value depending on whether they were from regions in England (1) or regions in

northern France and Belgium (2). The geographic co-ordinates (latitude/longitude) of each locality were included as a co-variate in column format.

4.3 Results

4.3.1 Samples

Of the 942 bats genotyped, only those with data for six or more loci were retained for further analyses. A PCA analysis (conducted by Camille Jan) of the remaining samples identified four individuals with dubious species identification (one sample from France and three from the Yorkshire Dales) and these samples were discarded. This left a total of 782 individuals, comprising 670 males and 112 females (Table 4.2, Figure 4.1). Data regarding allele size ranges and numbers of alleles per locus can be found in Table 4.3. HWE tests were not included as the numbers of females within sample sites in the dataset were too small to test for deviations from HWE reliably.

Table 4.2. Number of individuals genotyped by location

Sampling Location (clustering id in brackets)	Region	Country	Number of individuals with 6 or more loci genotyped
Bornem (1)	Belgium	Belgium	8
Liezele (2)	Belgium	Belgium	10
Demark (3)	Denmark	Denmark	18
Finland (4)	Finland	Finland	26
Roquefort (5)	South France	France	25
Fougeres (6)	North France	France	19
Pluherlin (7)	North France	France	23
Tremblay (8)	North France	France	15
Frayssinet (9)	South France	France	4
La Pyramide (10)	South France	France	11
Le cros (11)	South France	France	21
Senchet (12)	South France	France	12
Grotta Marelli (13)	Marelli	Italy	10
Tiberio (14)	Tiberio	Italy	13
Boho Cave (15)	Ireland	Ireland	6
Marble Arch Cave (16)	Ireland	Ireland	19
Poland (17)	Poland	Poland	13
Itxulegor (18)	North Spain	Spain	20

Sierra de Entezia (19)	North Spain	Spain	6
EIN Alta Garrotxa (20)	North Spain	Spain	8
Gouffre Cathy (21)	Cathy	Switzerland	23
Gouffre de la Pleine Lune (22)	Lune	Switzerland	18
Sussex (23)	South UK	England	22
Kent (24)	South UK	England	28
Windy Pits (25)	NYM	England	15
Scottish North clade a (26)	Scotland	Scotland	10
Scottish North clade b (27)	Scotland	Scotland	9
Sottish South clade b (28)	Scotland	Scotland	10
Stokesay Castle (29)	South UK	England	9
Withcall (30)	Withcall	England	7
Addingham (31)	Yorkshire Dales	England	23
Arncliffe (32)	Yorkshire Dales	England	6
Bar Pot (33)	Yorkshire Dales	England	13
Barden (34)	Yorkshire Dales	England	32
Browgill (35)	Yorkshire Dales	England	6
Buckden (36)	Yorkshire Dales	England	25
Buckden Gavel (37)	Yorkshire Dales	England	18
Coverbridge (38)	Yorkshire Dales	England	28
Dow cave (39)	Yorkshire Dales	England	31
Dowkabottom (40)	Yorkshire Dales	England	5
Fountains Abbey (41)	Yorkshire Dales	England	3
Grassington (42)	Yorkshire Dales	England	52
Ilkley (43)	Yorkshire Dales	England	47
Kettlewell Bridge (44)	Yorkshire Dales	England	28
Malham Tarn (45)	Yorkshire Dales	England	4
Wensley Bridge (46)	Yorkshire Dales	England	23

Table 4.3. Per Locus allele information

Locus	% missing data	Allele size range (bp)	Number of alleles
Efnb	7.8	126-133	8
Shroom	8.3	122-132	8
Pou	6.3	180-209	18
Nap	7.8	165-216	12
Kcne	7.3	199-225	16
Mcart	3.8	136-163	12
Nlgn	10.9	176-201	15
P22	21.4	227-324	18
Pcam	8.6	217-245	26

4.3.2 Linkage disequilibrium (LD)

There was significant LD between a number of locus pairs across all localities, however notable that the reported value of Chi^2 in each case was infinity. Chi^2 may take a value of infinity when the expected result is zero (or close to zero). This makes the test unreliable and so although significant linkage was reported these results should be interpreted with caution (Table 4.4). After Bonferroni corrections, within region LD was only demonstrated in the Yorkshire Dales (Table 4.5) suggesting that there may be some within region population structure in this area.

Table 4.4 Significant LD ($P < 0.05$ after Bonferroni correction) between pairs of loci across all localities

Locus Pair	Chi^2	Df	P value
Efnb + Pou	Infinity	32	Highly significant
Efnb + Nap	Infinity	32	Highly significant
Pou + Nap	Infinity	32	Highly significant
Kcne + Mcart	Infinity	34	Highly significant
Efnb + Nlgn	Infinity	32	Highly significant
Pou + Nlgn	Infinity	32	Highly significant
Mcart+Pcam	Infinity	32	Highly significant

Table 4.5 Significant ($P < 0.05$ after Bonferroni correction) LD between pairs of loci within localities.

Population	Locus 1	Locus 2	P-Value	S.E.
Yorkshire Dales	Efnb	Pou	0	0
Yorkshire Dales	Efnb	Nap	0	0
Yorkshire Dales	Pou	Nap	0	0
Yorkshire Dales	Efnb	Kcne	0.001053	0.000094
Yorkshire Dales	Pou	Kcne	0.000043	0.000033
Yorkshire Dales	Pou	Mcart	0.000244	0.000040
Yorkshire Dales	Kcne	Mcart	0	0
Yorkshire Dales	Efnb	Nlgn	0	0
Yorkshire Dales	Nap	Nlgn	0.000285	0.000044
Yorkshire Dales	Pou	Nlgn	0	0
Yorkshire Dales	Kcne	Nlgn	0.000157	0.000040
Yorkshire Dales	Mcart	Nlgn	0.000016	0.000008
Yorkshire Dales	Efnb	Pcam	0.001068	0.000118
Yorkshire Dales	Pou	Pcam	0.000480	0.000107
Yorkshire Dales	Kcne	Pcam	0	0
Yorkshire Dales	Mcart	Pcam	0	0
Yorkshire Dales	P22	Pcam	0.000760	0.000126

4.3.3 Population structure

F_{ST} across all loci and localities was low but significant at 0.034 ($P < 0.01$). Pair-wise comparisons of the 17 regions revealed very little differentiation between regions in Europe. Out of 136 pair-wise comparisons only 27 (20%) were significant (Table 4.6). Most regions were only significantly different from one or two other regions. Spain and southern France were differentiated from most populations in northern Europe but not those in southern Europe. There was evidence of significant differentiation of populations within the UK. For example between southern England and Scotland/NYM and at a smaller scale between the Yorkshire Dales and the North York Moors. It should also be noted that sample size differed between localities (see Table 4.2) and this may affect the distribution of significant values.

In order to analyse partitioning of genetic variation using AMOVA, localities were grouped into 17 regions. AMOVA was performed in Arlequin. The vast majority of the variance (96.56%) could be explained by the within locality factor (i.e. at the individual level), with very little, but significant, variation explained by either the among-groups

(1.46%) or among-localities (1.98%) factor. This suggests that there is very little genetic subdivision between colonies and that this lack of subdivision persists over a large geographic scale.

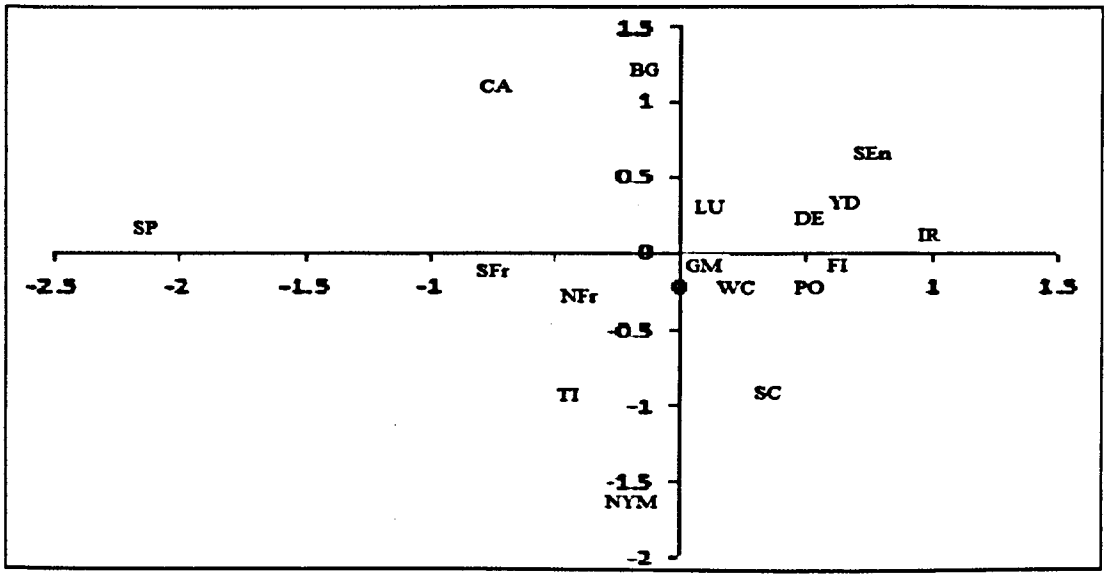


Figure 4.2 MDS plot based on F_{ST} values among European Daubenton's populations. BG = Belgium, CA = Gouffre Cathy, DE = Denmark, FI = Finland, GM = Grotta Marelli, IR = Ireland, LU = Gouffre de la Pleine Lune, NFr = Northern France, NYM = North York Moors, PO = Poland, SC = Scotland, SEn = Southern England, SFr = Southern France, SP = Spain, TI = Tiberio WC = Withcall, YD = Yorkshire Dales.

Population structure, as evidenced by the MDS plot (Figure 4.2), appeared to roughly mirror geographic distances. All regions in England and Ireland were more similar to regions in north eastern Europe, except the North York Moors which appeared closer to populations in Italy and France. Spain was the most divergent region.

Table 4.6 Pair-wise F_{ST} values between regions in Europe for X-linked microsatellite markers. Values in bold are significant ($P < 0.05$).

	Dales	Withcall	Moors	Belgium	Finland	Denmark	N France	S France	Ireland	Marelli	Tiberio	Poland	S England	Spain	Cathy	Lune	Scotland
Dales	0.000																
Withcall	0.007	0.000															
Moors	**0.090**	0.059	0.000														
Belgium	0.022	0.062	**0.172**	0.000													
Finland	0.004	-0.019	0.070	0.052	0.000												
Denmark	0.011	0.023	0.077	0.053	-0.003	0.000											
N France	**0.035**	0.010	0.057	0.040	0.039	0.035	0.000										
S France	**0.048**	0.024	0.056	0.044	**0.055**	0.059	0.000	0.000									
Ireland	0.017	-0.020	**0.126**	0.065	0.005	0.049	**0.055**	**0.074**	0.000								
Marelli	0.027	-0.005	-0.007	0.071	0.000	0.010	0.008	0.007	0.042	0.000							
Tiberio	**0.064**	0.009	0.024	0.090	0.048	0.086	0.024	0.014	0.060	-0.014	0.000						
Poland	0.014	-0.042	0.106	0.070	-0.014	0.026	0.030	0.059	-0.012	0.015	0.041	0.000					
S England	-0.001	0.004	**0.147**	0.018	0.001	0.025	**0.043**	**0.063**	0.004	0.048	0.083	0.009	0.000				
Spain	**0.099**	0.070	**0.128**	0.067	**0.114**	**0.111**	0.023	0.028	**0.125**	0.080	0.057	0.091	**0.106**	0.000			
Cathy	**0.038**	0.065	**0.160**	-0.011	0.066	0.053	0.021	0.032	**0.078**	0.068	0.087	0.067	0.036	0.032	0.000		
Lune	0.010	0.032	**0.111**	-0.002	0.024	0.023	0.014	0.022	0.033	0.015	0.047	0.034	0.008	0.064	-0.001	0.000	
Scotland	0.029	-0.048	**0.072**	**0.095**	-0.001	0.058	0.025	**0.041**	0.012	-0.004	-0.003	-0.013	0.045	**0.079**	**0.092**	0.056	0.000

The Mantel test demonstrated significant isolation by distance across the sampling area ($R^2 = 0.1036$, $r = 0.322$, slope = 9.9×10^{-5}, $P < 0.001$, Figure 4.3).

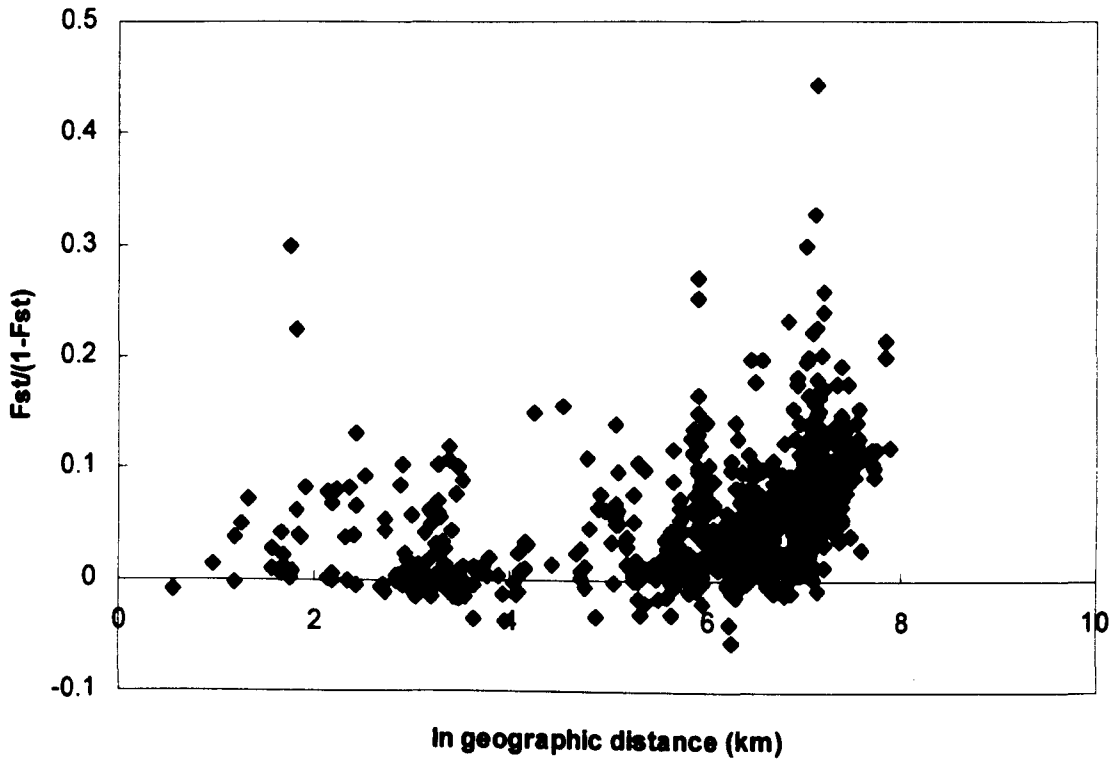


Figure 4.3 Plot of the correlation between Rousset's genetic distance ($(F_{ST}/(1-F_{ST}))$) and the natural log of Euclidean geographic distance (km) for X-linked microsatellites in Daubenton's populations across Europe.

4.3.4 Clustering analysis using data from males and females

4.3.4.1 Structure

Results from the STRUCTURE analysis for both sexes did not provide a simple answer. Using no prior location information there was no detectable structure and the best K was found to be one. When prior sampling location was included to aid clustering, likelihood values ($Pr(X|K)$) increased monotonically, although above K=4 (Figure 4.4a) there was substantial variance among runs. A plot of the absolute rate of change of K (Evanno *et al.* 1995) showed that the highest level of structure was at K=2 (Figure 4.4b). When using the locprior model, STRUCTURE also calculates a value 'r' for a given K which allows a measure of the informativeness of the pre-defined location. If r

is close to or below one it suggests prior locations are informative, but if it is much greater than one either there is no real population structure or the population structure is independent of the pre-defined populations. Values of r for K2, K3 and K4 were 14.718, 1.2080 and 2.0546 respectively. This suggested that there may have been a signature of population structure at K3 and K4 that the Evanno *et al.* (2005) method failed to pick up. K=3 highlighted France/Spain/Switzerland/Italy, Denmark/Poland/Finland/England and Scotland to be separate clusters with substantially mixed membership in Belgium, Ireland and some sampling sites in England. At K=4, Finland stood out as a separate cluster (Figure 4.5).

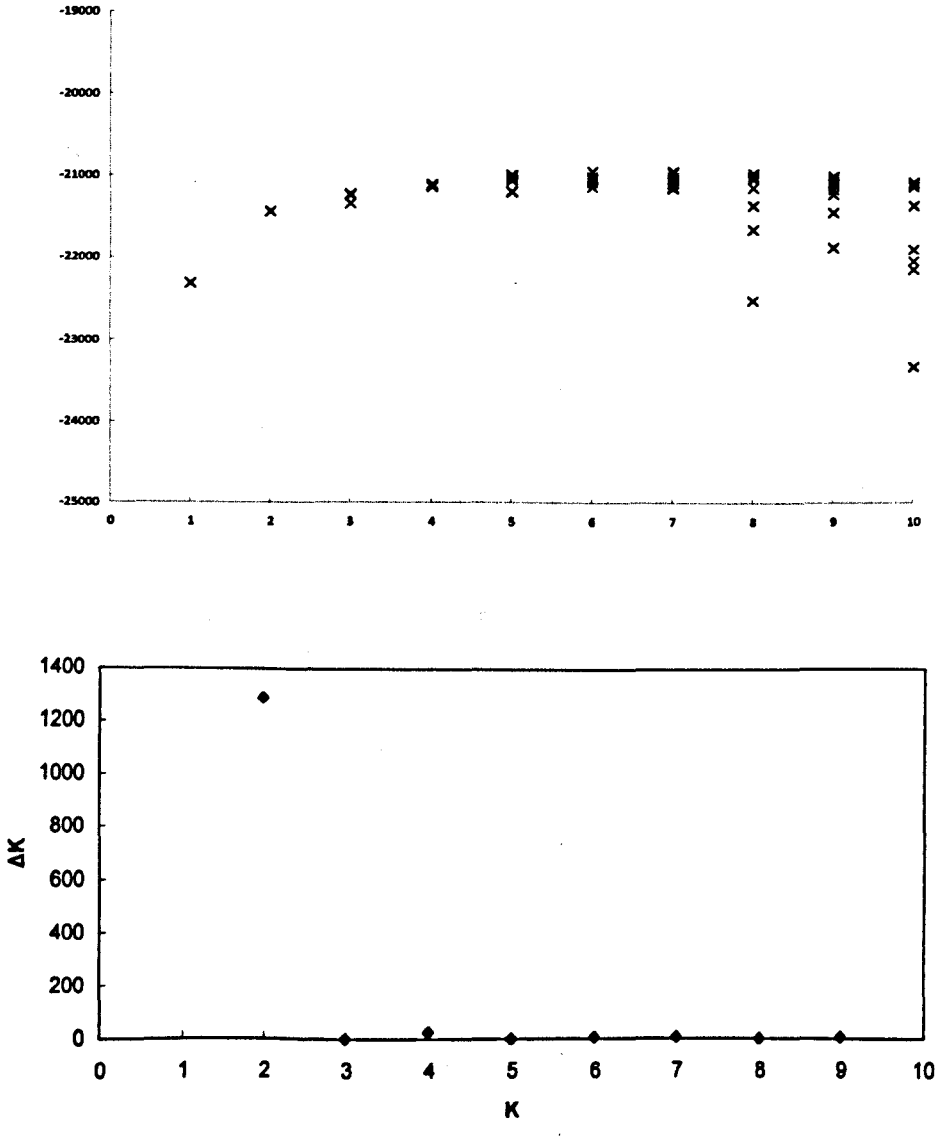


Figure 4.4. a) plot of log likelihood values for K 1-10 for ten independent runs. b) plot of absolute rate of change of log likelihood values for increasing values of K.

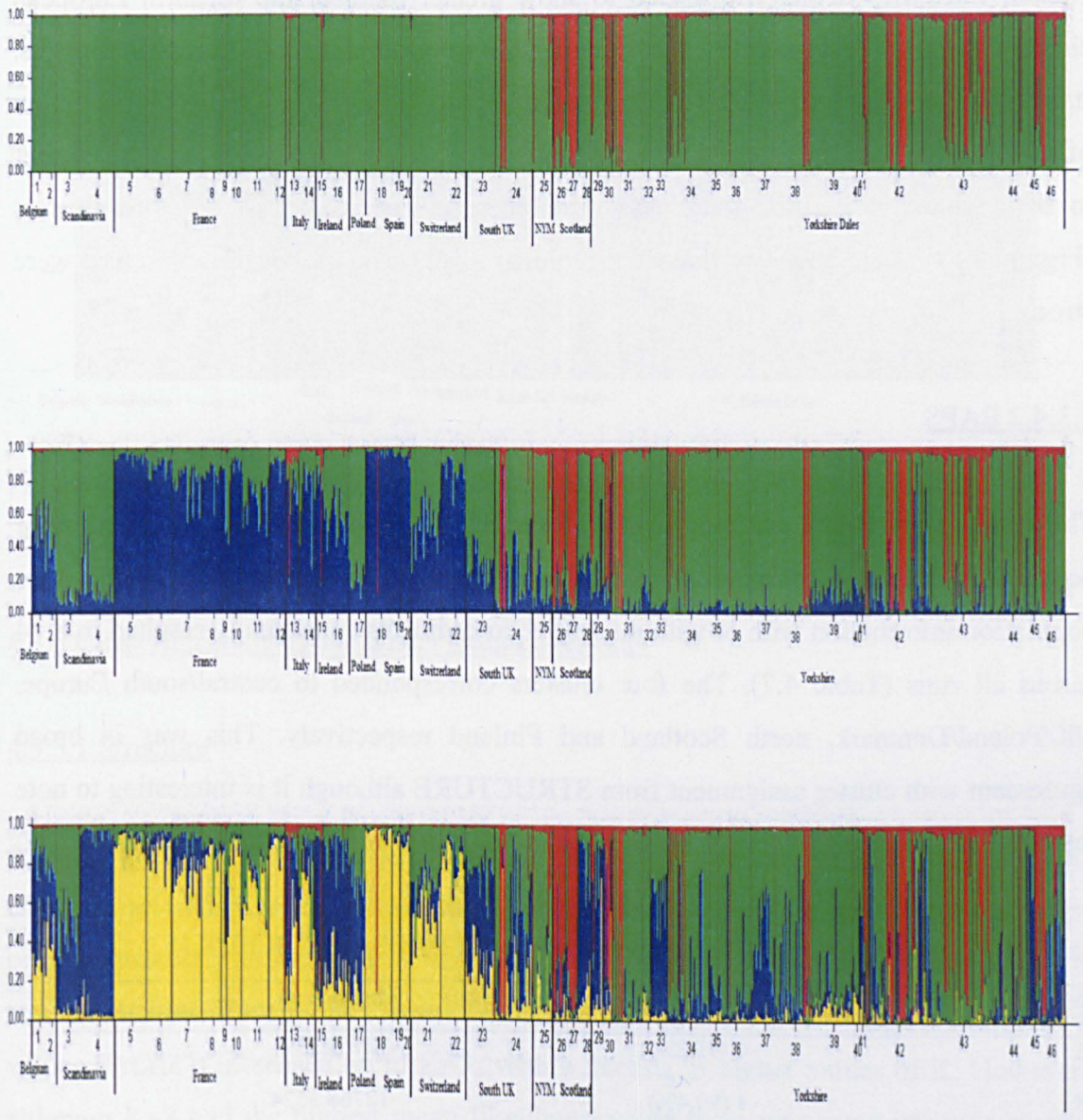


Figure 4.5 Bar plot of STRUCTURE output for a) K2 b) K3 and c) K4. Each line represents the proportion of assignment (Q) of each individual to a given cluster. Sample sites are separated by black lines. For details of sample site names corresponding to numbers see Table 4.1.

It is notable that out of the 46 localities sampled only 11 (Finland, Fougeres, Tremblay, La Pyramide, Senchet, Itxulegor, Sierra de Entezia, EIN Alta Garrotxa, Arncliffe, Dow Cave) assigned confidently (Q > 80%) to one cluster at K=4. The majority of localities contained individuals with membership to each of the four clusters (Figure 4.5c). Collating results from localities into regions highlighted only regions in Spain, northern France, southern France and Finland as containing over 80% membership to a particular

cluster. Localities in the UK tended to show greater membership to north European clusters. Belgium had roughly equal membership to north and south European clusters. It is also interesting to note that the two Italian regions assigned a significant proportion of membership to clusters more typically associated with north eastern regions. Further to this, within individuals there was often membership from two or more clusters (Figure 4.5). This suggests that at a regional scale gene flow and admixture were strong.

4.3.4.2 BAPS

The results from BAPS are congruent with the STRUCTURE analysis. Using no prior population or spatial information resulted in the most likely K being one. Using prior population information with or without spatial co-ordinates consistently resulted in K=4 across all runs (Table 4.7). The four clusters corresponded to central/south Europe, UK/Poland/Denmark, north Scotland and Finland respectively. This was in broad agreement with cluster assignment from STRUCTURE although it is interesting to note that the Irish and the southern Scottish populations clustered with mainland Europe.

Table 4.7 Logml values for five independent runs of BAPS at K5, K10, K15

Maximum K assumed	Most likely K (% confidence)	Logml
5	4 (91.2%)	-12769.237
5	4 (99.4%)	-12768.3524
5	4 (99.9%)	-12768.2555
5	4 (89.2%)	-12769.237
5	4 (100%)	-12769.237
10	4 (91%)	-12769.237
10	4 (91%)	-12769.237
10	4 (99%)	-12769.237
10	4 (91%)	-12769.237
10	4 (91%)	-12769.237
15	4 (100%)	-12768.2555
15	4 (100%)	-12768.2555
15	4 (100%)	-12768.2555
15	4 (100%)	-12768.2555
15	4 (100%)	-12768.2555

Figure 4.6 BAPS bar plot. Colours represent assignment to one of four clusters.

4.3.5 Clustering analysis using data from males only

4.3.5.1 Structure

As with the STRUCTURE output for both sexes, no structure can be detected when using no prior population information and the most likely K is one. However, when prior population information is considered there is more convincing evidence of structure within Europe than when females are included in the data set. Likelihood values ($\Pr(X|K)$) increased up to K=3, with a decline at higher values of K. However, although K=3 had the highest mean likelihood value there was some variance among runs (Figure 4.7). A plot of the absolute rate of change of K (Evanno *et al.* 1995) showed that the highest level of structure was at K=2 (Figure 4.7). There appeared to be significant structuring at both K=2 and K=3. However the substantial variance between runs for K =3 and the high absolute rate of change at K=2 made K=2 the most plausible level of structure within the study area. At K=2 the study area appeared to be split into the UK/north eastern Europe and central/southern Europe, with substantial levels of mixed membership in Belgium, Ireland and Gouffre Cathy (Switzerland). In contrast to the data set containing females, most populations comprised majority membership in one cluster and there were few cases where individuals in a population were from different clusters. This suggested that signatures of admixture were weaker when considering only males. There is little change in the outcome at K=3 although it is notable that Finland formed a separate cluster (Figure 4.8).

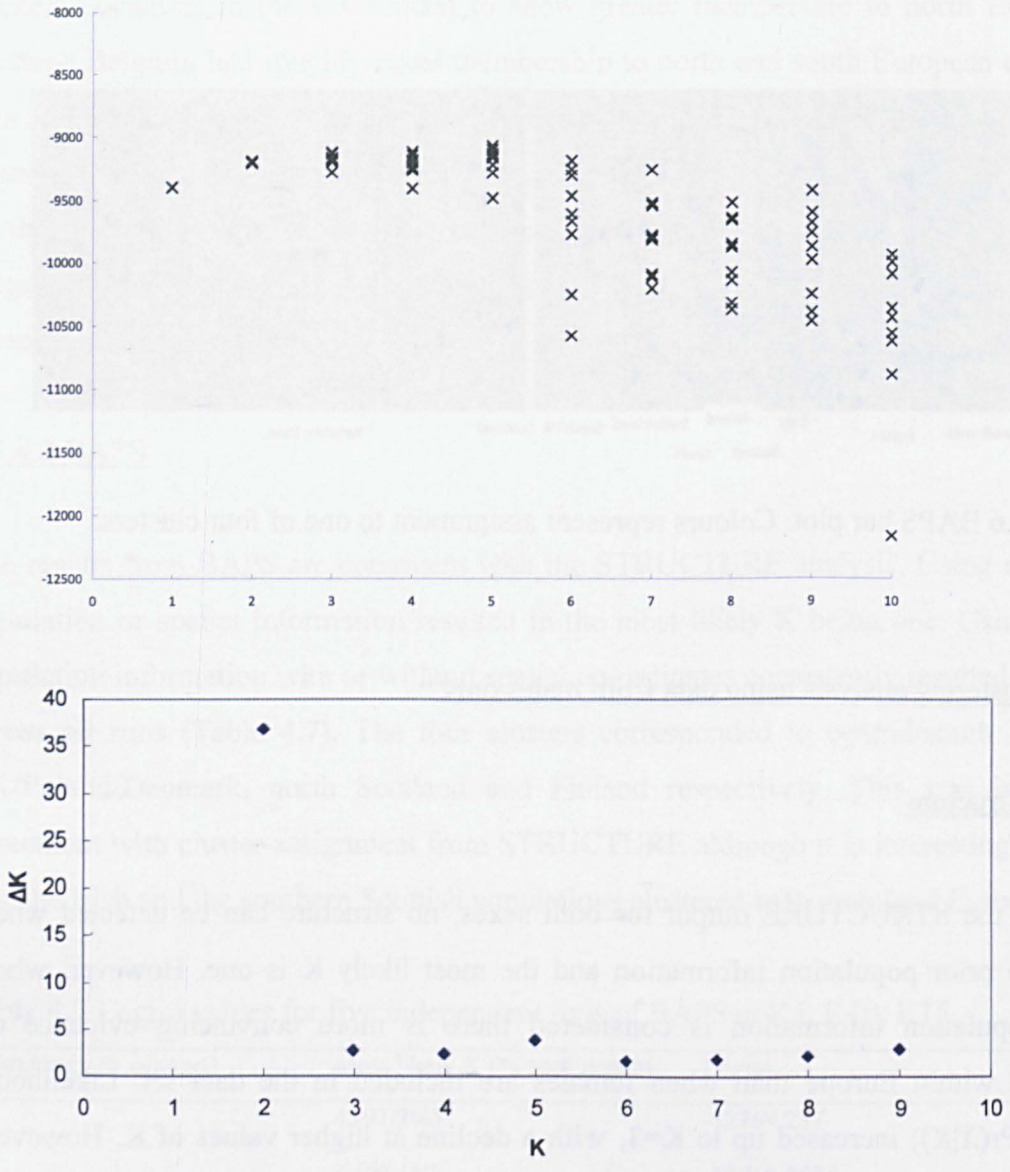
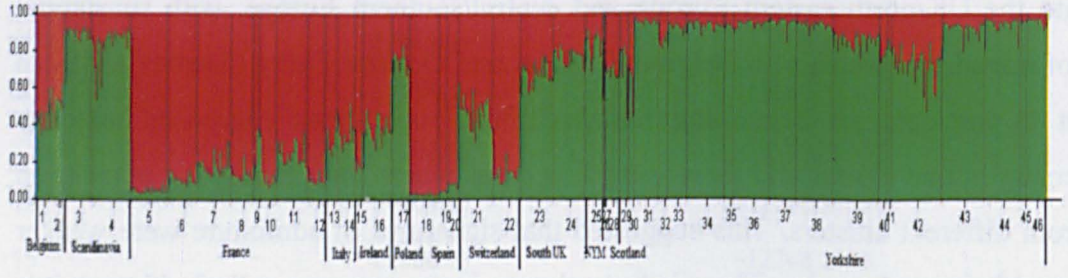


Figure 4.7 a) log likelihood values for K 1-10 for ten independent runs for male only data. b) absolute rate of change of log likelihood values for increasing values of K for male-only data.



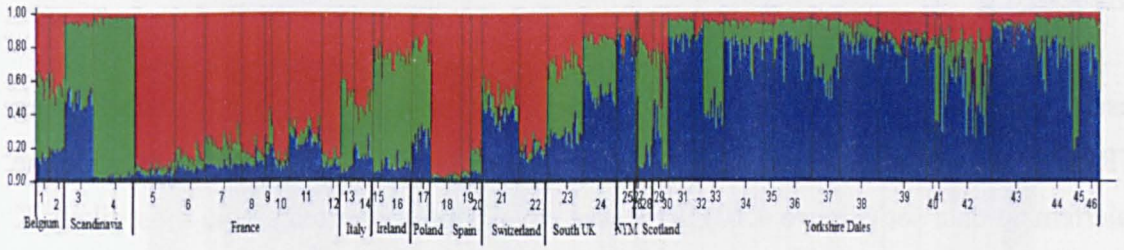


Figure 4.8 Bar plot of STRUCTURE output for male only data a) K2 b) K3. Each line represents the proportion of assignment (Q) of each individual to a given cluster. Sample sites are separated by black lines. For details of sample site names corresponding to numbers see Table 4.1.

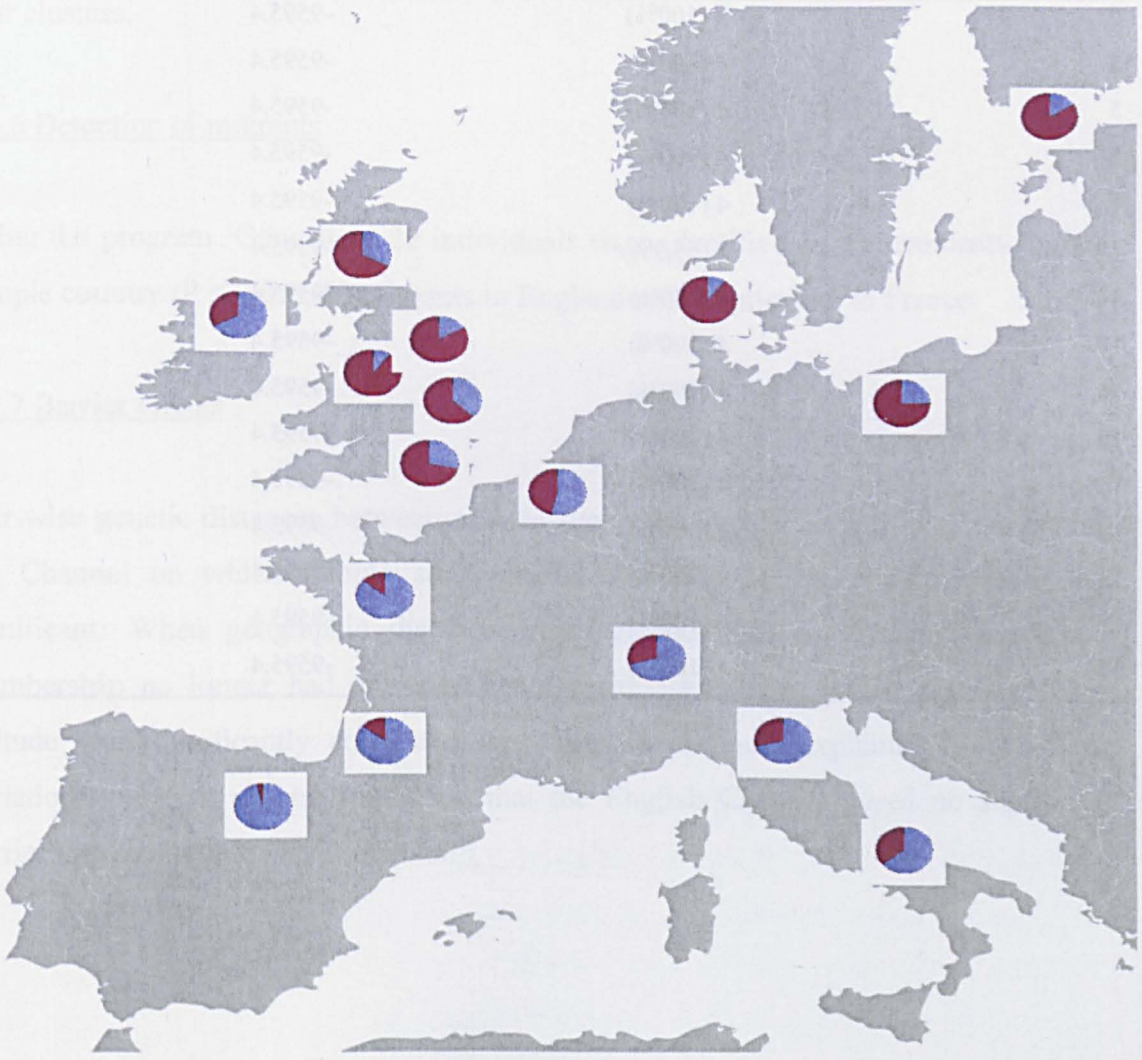


Figure 4.9 Regional percentage cluster assignment at K=2 for male *M. daubentonii* individuals according to geographic location.

4.3.5.2 BAPS

Results from the BAPS analysis suggested more clusters within Europe than when using STRUCTURE to analyse the data set and appeared to be more congruent with the male/female data set (Figure 4.10). K=4 was consistently reported as the most likely K across all runs (Table 4.8). The clusters corresponded to UK/north east Europe, Finland, Ireland, southern Europe. Interestingly Gouffre Cathy in Switzerland clustered with UK/north east Europe.

Table 4.8 Logml values for five independent runs of BAPS at K5, K10, K15 for males only

Maximum K assumed	Most likely K (% confidence)	Logml
5	4 (100%)	-9595.4
5	4 (100%)	-9595.4
5	4 (100%)	-9595.4
5	4 (100%)	-9595.4
5	4 (100%)	-9595.4
10	4 (65.9%)	-9595.4
10	4 (100%)	-9595.4
10	4 (100%)	-9595.4
10	4 (100%)	-9595.4
10	4 (100%)	-9595.4
15	4 (100%)	-9595.4
15	4 (100%)	-9595.4
15	4 (100%)	-9595.4
15	4 (100%)	-9595.4
15	4 (100%)	-9595.4

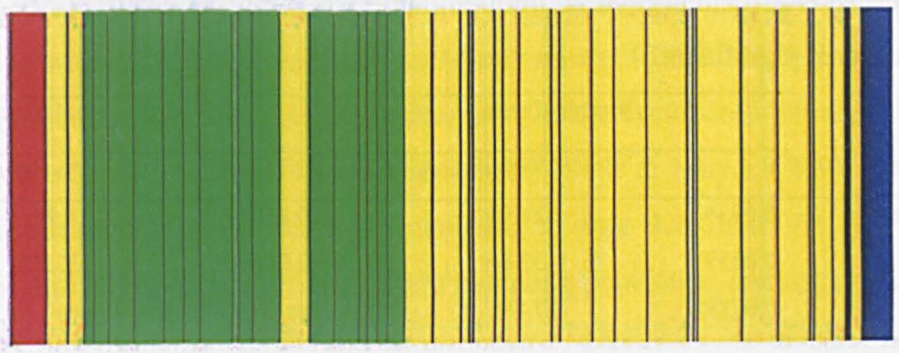


Figure 4.10 BAPS bar plot for male only data. Colours represent assignment to one of four clusters.

4.3.6 Detection of migrants

Using the program 'Geneclass' 22 individuals were identified as non-residents to their sample country ($P < 0.01$), 10 migrants to England and 12 migrants to France.

4.3.7 Barrier effects

Pair-wise genetic distances between sample sites were correlated both with the side of the Channel on which sample sites resided and latitude, but longitude was non significant. When geographic distance was included as a co-variable sample site membership no longer had any significant relationship with genetic distance. Only latitude was significantly correlated with genetic distance, explaining 34% of the variation. Thus it can be concluded that the English Channel posed no significant barrier to gene flow for *M. daubentonii.*

Table 4.9 Barrier effect of the English Channel on genetic distance in *M. daubentonii*. Results are shown for marginal and conditional tests. Only values of P less than 0.05 were considered significant.

	Marginal tests		Conditional tests	
Variable Set	P	% variation	P	% variation
Sample site	0.0339	26.99	0.4501	1.94
Latitude	0.0159	33.86	0.0159	33.86
Longitude	0.0789	7.89	na	na

4.4 Discussion

4.4.1 Patterns of differentiation

Over the entire study area a mean F_{ST} value of 0.034 suggested weak but significant population structure. However pair-wise comparisons varied widely, ranging from -0.069 to 0.172 (see Table 4.6), and few were significant. Moreover, AMOVA apportioned the vast majority of differentiation to variation at the individual level, suggesting overall levels of differentiation between both local populations and wider geographic regions to be very weak. In comparison to other studies of Daubenton's and temperate bat species our mean F_{ST} value was notably large and this was indicative of some overall population structure. For example Ngamprasertwong *et al.* (2008) reported a mean F_{ST} of 0.017 for Daubenton's across Scotland, Angell (2008) reported a value of 0.003 for summer sites across the Yorkshire Dales and 0.0043 for swarming sites in Yorkshire and Atterby *et al.* (2009) reported F_{ST} values which ranged from 0 to 0.05 for populations across western Europe. It could be that a relatively high mean F_{ST} in this study was reflective of the area over which populations were sampled. Ngamprasertwong *et al.* (2008) and Angell (2008) concentrated sampling on a comparatively small geographic area and indeed our F_{ST} values are more in concordance with those of Atterby *et al.* (2009) whose sampling was also at the continental scale. Differentiation between populations is expected to be higher purely as a product of large geographic distance between some populations, which automatically restricts gene flow between them. This is further supported by the finding of significant isolation by distance (IBD) in this study. The overall magnitude of our mean F_{ST} value was comparable to that for other bat species within Europe, however it was more similar to values found for non-migratory species such as *Rhinolophus ferrumequinum* (F_{ST} =

0.043, Europe wide study) (Rossiter *et al.* 2007) and *Myotis bechsteinii* (F_{ST} = 0.03, Europe wide study) (Kerth *et al.* 2008) than for bats with dispersal capabilities in concordance with *M. daubentonii* such as *Myotis myotis* (0.022, Study located around the Alps) (Castella *et al.* 2001), *Myotis natterreri* (0.017, Yorkshire) (Rivers *et al.* 2005) and *Nyctalus noctula* (0.006, Europe wide study) (Petit & Mayer 1999). Again this could be a question of scale but it is interesting to note that AMOVA still assigned small but significant levels of differentiation to among population and region factors.

All aforementioned work used autosomal microsatellites and this is the first study on bats using only X-linked markers to investigate population structure. As a consequence, we may expect to find more pronounced genetic structure due to the lower mutation rate of the X-chromosome and smaller effective population size in comparison to autosomes, under neutral expectations (Hedrick 2007). This discrepancy becomes even more pronounced under male-biased dispersal, due to reduced X-chromosome gene flow relative to autosomes (Ellegren 2009, Wilkins & Marlowe 2006). Population genetics work on humans has indeed found differences in the level of structure reported when using autosomal Vs X-linked markers, which has been attributed to skewed levels of gene flow between the sexes (Bustamante & Ramachandran 2009). Further, past work on Daubenton's strongly suggests that dispersal is male biased (Angell 2008, Ngamprasertwong *et al.* 2008) which would act to reduce the ratio of X:A effective population size and hence increase X-linked structure. If this is the case it may explain, in part, why our mean F_{ST} value appeared large in comparison to other studies. Work in later chapters on autosomal markers in the same individuals should help to illuminate the matter.

Results from this study and other work on temperate bat species were in stark contrast to population genetics studies on similarly size ground dwelling mammals. For example, Heckel *et al.* (2005) studied populations of the common vole across Europe and found limited dispersal with strong differentiation, which is the typical pattern for small rodents. Due to the high mobility of Daubenton's bat substantial gene flow over large geographic areas is not surprising and the weak population structure in *M. daubentonii* is more typical of large mammals. For example, grey wolves show significant but weak structure which tends to be dictated more by environmental factors than physical barriers to gene flow (Geffen *et al.* 2004). Studies on the Eurasian badger in western Europe also reveal low levels of differentiation among populations, with

broad scale genetic structure a reflection of IBD patterns (Pope *et al.* 2005). The capacity for powered flight and the wing morphology of Daubenton's, allows for long distance movement and the capability to traverse areas that would otherwise act as a barrier for ground dwelling rodents (Altringham 1996). Consequently past work has found that *M. daubentonii* dispersal is more restricted by climatic conditions due to the energetic demands of flight and lower abundance of insects at high altitudes than by physical or distance related barriers (Angell 2008, Ngamprasertwong *et al.* 2008, Russo 2002).

Low levels of structure and high levels of admixture between populations could be a consequence of swarming behaviour (Kerth *et al.* 2003, Rivers *et al.* 2005). As with a number of other temperate bat species *Myotis daubentonii* undergoes a seasonal migration from summer roosts to swarming sites, where mating occurs (Rivers *et al.* 2005, 2006). At any particular swarming site bats may come from a large catchment area and mating between bats from different summer roosts would act to homogenise differentiation between them (Angell 2008). This may be particularly pertinent in some areas of the continent where there are few suitable swarming sites meaning that bats migrate from a wider area to visit them. This may in part explain why there appeared to be greater structure in the UK (as evidenced by unique cluster grouping of Scotland and significant F_{ST} values between the North York Moors and Scotland/Ireland) than in continental Europe.

4.4.2 Population structure across Europe

Results from clustering analyses showed that there was some genetic structuring of Daubenton's populations across Europe. However, using both programs structure was only found when prior information on sampling location was given, suggesting that it was weak (Hubitz *et al.* 2009). Using the program STRUCTURE a stronger signal of structure was apparent when using only male data than when including data from both sexes. The two major clusters in the male-only data set geographically represented central/southern Europe and north eastern Europe/Britain. Although STRUCTURE also gave the strongest support to two clusters when using data from both sexes, there was little geographical coherence, with most localities comprising individuals with substantial membership from both clusters. Because males only have one X-chromosome this means that we know the phase of a particular allele at a locus

(Schaffner 2004). Consequently it is possible to determine linkage patterns between loci. This study found highly significant evidence for linkage disequilibrium (LD) between a number of locus pairs across the study area. LD across a population is suggestive of population structuring due to the maintenance of different allele frequencies among subpopulations (Slatkin 2008). Because STRUCTURE aims to minimise deviations from linkage disequilibrium within clusters this means we can detect population structure when the phase is known and LD is present which would not be possible when using data from a diploid locus. This suggests that the information gained from knowing the phase of the alleles outweighs information from adding additional individuals to the analyses. However, as the results from the BAPS analyses were very similar using male and female data and the male-only data set, this suggested that BAPS was unable to use the additional information from linkage patterns to aid population clustering.

The central/southern Europe and north eastern Europe/Britain clusters were broadly in agreement with the finding of two mitochondrial clades in this study (Chapter 3), which means that differentiation between these two areas could, in part, reflect historical colonization from two different sources (e.g. Brito 2007, Yannic *et al.* 2008). Populations in southern Britain, Belgium and Gouffre Cathy in Switzerland comprised individuals with a significant proportion of membership from both clusters, indicative of admixture between individuals from northern and southern Europe. Interestingly individuals from the Irish populations shared more membership with southern Europe than they did with the UK in the STRUCTURE analysis, and formed a separate cluster completely when BAPS was used. Using autosomal markers Atterby *et al.* (2009) found Irish populations of Daubenton's to be genetically distinct from both Britain and mainland Europe. Work on other species found in Ireland and continental Europe also suggests that Irish populations are genetically divergent from Britain (Boston *et al.* 2007, Teacher *et al.* 2009) and indeed some species are unique to Ireland, western France and Iberia (Corbett 1960). A Lusitanian connection between Iberia and Ireland has been invoked to explain these patterns (Boston *et al.* 2007) and it has been suggested that Ireland may have had climatic conditions suitable for a glacial refuge (Stewart & Lister 2001). For example the pine marten (Davison *et al.* 2001) and the pygmy shrew (Mascheretti *et al.* 2003) exhibit greater genetic similarity to populations in southern Europe than to Britain. There is also evidence that trade links between areas in Lusitania could have facilitated the transport of some species to Ireland (Yalden *et al.*

1982). Contrary to results from the clustering analyses, pair-wise F_{ST} values between Ireland and other regions showed that there was significant differentiation between Ireland and southern Europe but not Ireland and Britain, suggesting that contemporary gene flow among *M. daubentonii* populations from mainland Britain and Ireland may occur.

At K=3 Finland formed a separate cluster, which again could reflect historical processes (there is debate as to whether Daubenton's from eastern Europe belong to a distinct subspecies from western Europe based on cranial morphology (Bogdanowicz 1994)) but it could also be a product of the large geographic distance between the sample site in Finland and other sites in this study. As pair-wise F_{ST} comparisons between Finland and other populations were generally non-significant (south France and north Spain being exceptions) this favours the former explanation. Work on other vagile mammals has highlighted populations in Scandinavia as being genetically distinct (Fedorov *et al.* 2001). For example, Eurasian badgers in Finland form a discrete genetic cluster, despite little structure being found across the rest of mainland Europe (Pope *et al.* 2005).

In general, localities in England shared greater similarity, in terms of cluster membership, to north eastern Europe than central/southern Europe. Atterby *et al.* (2009) also found that the UK was distinct from Mainland Europe. However they did not sample eastern Europe so comparisons cannot be drawn. Again this may be a reflection of historic colonization of Europe by *M. daubentonii* from two different clades, perhaps with proportionally greater colonization of the UK from a north eastern clade than from the south. This suggests that although it may not be possible to identify migrant individuals to the UK from Scandinavia, there is the potential to detect migrants from France and other regions in southern Europe.

4.4.3 Isolation by distance

These results should be interpreted with caution, as it has been found that when using Bayesian clustering algorithms both STRUCTURE and BAPS perform well when F_{ST} values are at and above 3% but when F_{ST} values fall to 2% both programs have a tendency to find clusters with false certainty (Frantz *et al.* 2009, Latch *et al.* 2006) Although our mean F_{ST} was 3.4%, above the 3% threshold, for some pair-wise

comparisons F_{ST} was much lower than this. Clustering algorithms also have a tendency to find clusters, based not on actual barriers to gene flow, but on differences between populations due to isolation by distance (Frantz *et al.* 2009, Rosenberg *et al.* 2005). Indeed a Mantel test found significant isolation by distance across all localities sampled in this study. Further, IBD has been demonstrated in Daubenton's populations in Scotland, England and mainland Europe (Angell 2008, Atterby *et al.* 2009, Ngamprasertwong *et al.* 2008). If IBD is present across a large geographic area, it suggests that there are no significant geographical or behavioural barriers to gene flow and that differentiation between sampling sites is a product of the overall species range being greater than individual dispersal capabilities (Balloux & Lugon-Moulin 2002). Indeed the maximum recorded dispersal distance for a Daubenton's individual is 260 km (Urbancyzk 1990) and the maximum pair-wise distance between sample sites in this study was far greater at 2618 km. As a pattern of IBD was detected the idea of within-cluster panmixia may be inappropriate, as individuals may only be exchanging gene flow with neighbouring groups. IBD was found to be strong in this study, but the slope was shallow, which suggests that dispersal distances are large (see Chapter 5 for estimates of effective dispersal distance) and that gene flow could occur between bats from distant colonies (Slatkin 1993).

4.4.4 Testing for a potential barrier effect of the English Channel

Results from the dbra analysis suggested that the English Channel posed no significant barrier to X-chromosomal gene flow for Daubenton's. This appears slightly at odds with results from the STRUCTURE analysis which highlights England as considerably different in terms of cluster membership to regions in northern France and Belgium (Figure 4.9). However it should be noted that when STRUCTURE is run with no prior location information the most likely K is 1, this suggests that across the study area population structure is weak. When location information is provided England might appear differentiated from France/Belgium because a greater portion of migrants to Britain are from north eastern Europe. Thus whilst dispersal across the Channel probably occurs, it seems likely that fewer bats cross the Channel than if the same distance were land. Data from autosomal markers show that Pipistrelle species are also able to cross the English Channel (Racey *et al.* 2007) and indeed a number of other studies on bats suggest that large bodies of water pose little or no barrier to gene flow, even in species such as *Plecotus auritus* for which long distance flight is uneconomical

(Juste *et al.* 2004). Further, Atterby *et al.* (2009) found low genetic differentiation for autosomal markers among *M. daubentonii* populations on the continent and in Britain, which provides further supporting evidence that *M. daubentonii* individuals can and do cross the Channel. However in some species such as *Rhinolophus ferrumequinum*, which is non-migratory, the Channel does present a barrier to gene flow (Rossiter *et al.* 2000). This is also the case with *M. myotis* and the straits of Gibraltar (Castella *et al.* 2000). Although the dbra analysis provided no evidence of the Channel being a barrier to gene flow, geographic distance explained only 34% of differentiation between populations. This suggests there must be some other factor limiting gene flow which could come down to a number of effects such as behavioural differences, population size or differences in environmental clines.

4.4.5 Summary

In this chapter I have shown that there is very little population structure for *M. daubentonii* populations within Europe, with individuals from large geographic areas effectively in panmixia. However panmixia did not extend to the entire study area as evidenced by a pattern of isolation by distance. Isolation by distance appeared to be a factor limiting gene flow, with little evidence that the English Channel posed a barrier and no other obvious barrier to gene flow.

4.5 References

Altringham JD (1996) *Bats: Biology & Behaviour.* Oxford University Press, Oxford.

Anderson MJ (2004) DISTLM v.5: a FORTRAN computer program to calculate a distance-based multivariate analysis for a linear model. Department of Statistics, University of Auckland, New Zealand.

Angell RL (2008) Population ecology and mating system of Daubenton's bat (*Myotis daubentonii*) PhD Thesis, University of Leeds.

Atterby H, Aegerter JN, Smith GC *et al.* (2009) Population genetic structure of Daubenton's bat (*Myotis daubentonii*) in western Europe and the associated occurance of rabies. *European Journal of Wildlife Research* **56** 67-81.

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology & Systematics* **18** 489-522.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11** 155-165.

Balloux F (2010) Mitochondrial phylogeography: The worm in the fruit of the mitochondrial DNA tree. *Heredity* **104** 419-420.

Bourgeois S, Yotova V, Wang S *et al.* (2009) X-chromosome lineages and the settlement of the Americas. *American Journal of Physiological Anthropology* **140** 417-428.

Bartosch-Harlid A, Berlin S, Smith NGC, Møller AP, Ellegren H (2003). Life history and the male mutation bias. *Evolution* **57** 2398–2406.

Bogdanowicz W (1994) *Myotis daubentonii. Mammalian Species* **475** 1-9.

Boston E, Prodohl P, Montgomery A (2007) Mitochondrial phylogeography of the Leisler's bat (*Nyctalus leisleri*) in Ireland and across Europe: new perspectives. *Bat Research News* **48** 184.

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

Brito PH (2007) Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (*Strix aluco*). *Molecular Ecology* **16** 3423-3437.

Brumfield RT, Beerli P, Nickerson DA, Edwards SV (2003) The utility of single nucleotide polymorphisms in inferences of population history. *TREE* **18** 249-256.

Burland TM, Worthington-Wilmer J (2001) Seeing in the dark: molecular approaches to the study of bat populations. *Biological Reviews* **76** 389-409.

Bustamante CD, Ramachandran S (2009) Evaluating signatures of sex specific processes in the human genome. *Nature Genetics* **41** 8-10.

Caballero A (1994) Developments in the prediction of effective population size. *Heredity* **73** 657–679.

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. *Nature* **325** 31–36

Castella V, Ruedi M, Excoffier L, Ibanez C, Arlettaz R, Hausser J (2000) Is the Gibraltar strait a barrier to gene flow for the bat *Myotis myotis* (Chiroptera: Vespertilionidae). *Molecular Ecology* **9** 1761-1772.

Castella V, Ruedi M, Excoffier L (2001) Contrasted patterns of mitochondrial and nuclear structure among nursery colonies of the bat *Myotis myotis*. *Journal of Evolutionary Biology* **14** 708-539 720.

Charlesworth B (2001) The effect of life-history and mode of inheritance on neutral genetic variability. *Genetics Research* **77** 153–166.

Corander J, Marttinen P, Siren J, Tang J (2008) Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9** 539.

Corbett GB (1961) Origin of the British insular races of small mammals and of the 'Lusitanian' fauna. *Nature* **191** 1037–1040.

Davison A, Birks JDS, Brookes RC, Messenger JE, Griffiths HI (2001) Mitochondrial phylogeography and population history of pine martens *Martes martes* compared with polecats *Mustela putorius*. *Molecular Ecology* **10** 2479–2488.

Ellegren H (2007) Characteristics, causes and evolutionary consequences of male-biased mutation. *Proceedings of the Royal Society of London Series B.* **274** 1–10.

Ellegren H (2009) The different levels of genetic diversity in sex chromosomes and autosomes. *Trends in Genetics* **25** 278-284.

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

Excoffier LG, Laval T, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1** 47-50.

Fedorov VB, Stenseth NC (2001) Glacial survival of the Norwegian lemming (*Lemmus lemmus*) in Scandinavia: evidence from mitochondrial DNA variation. *Proceedings of the Royal Society of London Series B.* **268** 809-814.

Flanders J, Jones G, Benda P *et al.* (2009) Phylogeography of the greater horseshoe bat, *Rhinolophus ferrunquinum*: Contrasting results from mitochondrial and microsatellite data. *Molecular Ecology* **18** 206-318.

Frantz AC, Cellina S, Krier A, Schley L, Burke T (2009) Using spatial bayesian methods to determine the genetic structure of a continuously distributed population: structure or isolation by distance? *Journal of Applied Ecology* **46** 493-505.

Geffen E, Anderson MJ, Wayne RK (2004) Climate and habitat barriers to dispersal in the highly mobile grey wolf. *Molecular Ecology* **13** 2481-2490.

Geraldes A, Rogel-Gaillard C, Ferrand, N. (2005) High levels of nucleotide diversity in the European rabbit (*Orydolagus cuniculus*) SRY gene. *Animal Genetics* **36** 349-351.

Geraldes A, Basset P, Gibson B *et al.* (2008) Inferring the history of speciation in house mice from autosomal, X-linked, Y-linked and mitochondrial genes. *Molecular Ecology* **17** 5349-5363.

Hagenblad J, Olsson M, Parker H, Ostrander EA, Ellegren H (2009) Population genomics of the inbred Scandinavian wolf. *Molecular Ecology* **18** 1341-1351.

Hammer MF, Mendex FL, Cox MP, Woerner AE, Wall JD (2008) Sex biased evolutionary forces shape patterns of human genomic diversity. *PLOS* **4** 1-8.

Harris EE, Hey JX (1999) chromosome evidence for ancient human histories. *PNAS* **96** 3320–3324.

Heckel G, Burri R, Fink S, Desmet JF, Excoffier L (2005) Genetic structure and colonization processes in European populations of the common vole, *Microtus arvalis*. *Evolution* **59** 2231–2242.

Hedrick PW (2007) Sex: Difference in mutation, recombination, selection, gene flow & genetic drift. *Evolution* **61** 2750-2771.

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

Jan C (2010) PhD Thesis, University of Leeds.

Jaruzelska J, Zietkiewicz E, Batzer M *et al.* (1999) Spatial and Temporal Distribution of the Neutral Polymorphisms in the Last ZFX Intron: Analysis of the Haplotype Structure and Genealogy. *Genetics* **152** 1091-1101.

Jobling MA, Tyler-Smith C (2003) The human Y-chromosome: an evolutionary marker comes of age. *Nature Review Genetics* **4** 508-612.

Juste J, Ibàñez C, Muñoz J *et al.* (2004) Mitochondrial phylogeography of the long-eared bats (*Plecotus*) in the Mediterranean Palaearctic and Atlantic Islands. *Molecular Phylogenetics and Evolution* **31** 1114–1126.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16** 1006-1009.

Keinan A, Mullikin JC, Patterson N, Reich D (2009) Accelerated genetic drift on the X chromosome during human migration out of Africa. *Nature Reviews Genetics* **41** 66–70.

Kerth G, Mayer F, Koenig B (2002) Mitochondrial DNA (mtDNA) reveals that female Bechstein's bats live in closed societies. *Molecular Ecology* **9** 793-800.

Kerth G, Petrov B, Conti A *et al.* (2008) Communally breeding Bechstein's bats have a social system that is independent from the postglacial history and location of the populations. *Molecular Ecology* **17** 2368-2381.

Kidd JM, Cheng Z, Graves T, Fulton B, Wilson RK, Eichler, EE (2008) Haplotype sorting using human fosmid clone end-sequence pairs. *Genome Research* **18** 2016-2023.

Kong A, Gudbjartsson DF, Sainz J *et al.* (2002) A high-resolution recombination map of the human genome. *Nature Reviews Genetics* **31** 241–247.

Laan M, Wiebe V, Khusnutdinova, E (2005) X-chromosome as a marker for population history: linkage disequilibrium and haplotype study in Eurasian populations. *European Journal of Human Genetics* **13** 452–462.

Lambert CA, Connelly CF, Madeoy J, Qiu R, Olson MV, Akey JM (2010) Highly punctuated patterns of population structure on the X chromosome and implications for African evolutionary history. *American Journal of Human Genetics* **86** 34-44.

Laporte V, Charlesworth B (2002) Effective population size and population sub division in demographically structured populations. *Genetics* **162** 501-519.

Latch EK, Dharmarajan G, Glaubitz JC, Rhodes OE (2006) Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics* **7** 295-302.

Li CC (1948) *An introduction to population genetics.* University Press, Peking.

Li W-H, Yi S, Makova K (2002) Male-driven evolution. *Current Opinion in Genetics & Development* **12** 650–656.

Lucotte G, Guerin P, Halle L, Loirat F, Hazout S (1989)Y chromosome DNA polymorphisms in two African populations. *American Journal of Human Genetics* **45** 16–20.

Mascheretti S, Rogatcheva MB, Gunduz I, Fredga K, Searle JB (2003) How did pygmy shrews colonise Ireland? Clues from a phylogenetic analysis of mitochondrial cytochrome-B sequences. *Proceedings of the Royal Society of London Series B.* **270** 1593-1599.

McCracken GF, Wilkinson GS (2000) Bat mating systems. In: *Reproductive Biology of Bats*, edited by E. G. Crichton and P. H. Krutzsch. pp 321-362. Academic Press, London and San Diego.

Melo-Ferreira J, Alves PC, Freitas H, Ferrand N, Boursot P (2009) The genomic legacy from the extinct *Lepus timidus* to the three hare species of Iberia: contrast between mtDNA, sex chromosomes and autosomes. *Molecular Ecology* **18** 2643-2658.

Nachman MW (1998) Y chromosome variation of mice and men. *Molecular Biology & Evolution* **15** 1744–1750.

Ngamprasertwong T, Mackie IJ, Racey PA, Piertney SB (2008) Spatial distribution of mitochondrial and microsatellite DNA variation in Daubenton's bat within Scotland. *Molecular Ecology* **17**,3243-3258.

Paetkau D, Slade R, Burden M, Estoup A (2004) Direct, real-time estimation of migration rate using assignment methods: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13** 55-65.

Park SDE (2001), Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection (Ph.D. thesis) University of Dublin

Parsons KN, Jones G (2003) Dispersion and habitat use by *Myotis daubentonii* and *Myotis nattereri* during the swarming season: implications for conservation. *Animal Conservation* **6** 283-290.

Pereira R, Gomes I, Amorim A, Gusmao L (2007) Genetic diversity of 10 X chromosome STRs in Northern Portugal. *International Journal of Legal Medicine* **121** 192-197.

Petit E, Mayer F (1999) Male dispersal in the noctule bat (*Nyctalus noctula*): where are the limits? *Proceedings of the Royal Society of London Series B: Biological Sciences* **266** 1717-1722.

Piry S, Alapetite A, Cornuet J-M, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: A Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity* **95** 536-539.

Pool JE, DuMont VB, Mueller JL, Aquadro CF (2006) A scan of molecular variation leads to the narrow localization of a selective sweep affecting both afrotropical and cosmopolitan populations of *Drosophila melanogaster*. *Genetics Society of America* **172** 1093-1105.

Pool JE, Nielsen R (2007) Population size changes re-shape patterns of genomic diversity. *Evolution* **61** 3001-3006.

Pool JE, Hellmann I, Jensen JD, Nielsen R (2010) Population genetic inference from genomic sequence variation. *Genome Research* **20** 291-300.

Pope LC, Domingo-Rouara X, Burke, T. (2005) Isolation by distance and geneflow in the Eurasian badger (*Meles meles*) at both a local and broad scale. *Molecular Ecology* **15** 371-386.

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

Racey PA, Barratt EM, Burland TM *et al.* (2007) Microsatellite DNA polymorphism confirms reproductive isolation and reveals differences in population genetic structure of cryptic Pipistrelle bat species. *Biological Journal of the Linnean Society* **90** 539–550.

Rivers NM, Butlin RK, Altringham JD (2005) Genetic population structure of Natterer's bats explained by mating at swarming sites and philopatry. *Molecular Ecology* **14** 4299-4312.

Rivers NM, Butlin RK, Altringham JD (2006) Autumn swarming behaviour of Natterer's bats in the UK: population size, catchment area and dispersal. *Biological Conservation* **127** 215-226.

Rosenberg N, Saurabh S, Ramachandran S, Zhao C, Pritchard J, Feldman M (2005) Clines, clusters, and the effect of study design on the influence of human population structure. Public Library of Science, *Genetics* **1** 660–671.

Rossiter SJ, Jones G, Ransome RD, Barratt EM (2000) Genetic variation and population structure in the endangered greater horseshoe bat (*Rhinolophus ferrumequinum*). *Molecular Ecology* **9** 1131-1135.

Rossiter SJ, Benda P, Dietz C, Zhang S, Jones G (2007) Rangewide phylogeography in the greater horseshoe bat inferred from microsatellite: implications for population history, taxonomy. *Molecular Ecology* **16** 4699-4714.

Rozen S, Skaletsky H (2000) Primer 3 on the www for general users and for biologist promgrammers. *Methods in Molecular Biology* **132** 365-386.

Russo D (2002) Elevation affects the distribution of the two sexes in Daubenton's bats *Myotis daubentonii* (Chiroptera: Vespertilionidae) from Italy. *Mammalia* **66** 543-551.

Schaffner SF (2004) The X-chromosome in population genetics. *Nature Reviews Genetics* **5** 43-51.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9** 615-629.

Senior P, Butlin RK, Altringham JD (2005) Sex and segregation in temperate bats. *Proceedings of the Royal Society of London, Series B* **272** 2467-2473.

Singh N, Macphersen J, Jensen J, Petrov, D. (2007) Similar levels of X-linked and autosomal nucleotide variation in african and non-african populations of *Drosophila melanogaster. BMC Evolutionary Biology* **7** 202.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47** 264–279.

Slatkin M (2008) Linkage disequilibrium: understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics* **9** 477-485.

Stewart JR, Lister AM. (2001) Cryptic northern refugia and the origins of the modern biota. *TREE* **16** 608–613.

Teacher AGF, Garner TWJ, Nichols RA (2009) European phylogeography of the common frog (*Rana temporaria*): Routes of postglacial colonisation into the British Isles, and evidence for an Irish glacial refugium. *Heredity* **102** 490-496.

Templeton AR (2002) Out of Africa again and again. *Nature* **416** 45-51.

Urbancyzk Z (1990) Northern Europe's most important hibernation site. *Oryx* **24** 30-34.

Vicoso B, Charlesworth B (2006) Evolution on the X-chromosome: unusual patterns & processes. *Nature Reviews Genetics* **7** 645-653.

Vicoso B, Charlesworth B (2009) Recombination rates may affect the ratio of X to autosomal non-coding polymorphism in African populations of *Drosophila melanogaster. Genetics* **181** 1699-1701.

Wilkins JF, Marlowe FW (2006) Sex-biased migration in humans: what should we expect from genetic data? *BioEssays* **28** 280-300.

Wilkinson GS, South JM (2002) Life history, ecology and longevity in bats. *Aging Cell* **1** 124-131.

Yalden DW (1982) When did the mammal fauna of the British Isles arrive? *Mammal Review* **12** 1–57.

Yannic G, Bassett P, Hausser J (2008) Phylogeography and recolonisation of the Swiss Alps by the Valais shrew (*Sorex antinori*), inferred with autosomal and sex-specific markers. *Molecular Ecology* **17** 4103-4118.

Yotova V, Lefebvre JF, Kohany O *et al.* (2007) Tracing genetic history of modern humans using X-chromosome lineages. *Human Genetics* **122** 431-443.

Yu N (2002) DNA Polymorphism in a worldwide sample of human X chromosomes. *Molecular Biology & Evolution* **19** 2131–2141.

Chapter 5: Investigating the population structure of Daubenton's bat across Europe using a combination of autosomal microsatellites, X-linked microsatellites and mtDNA: the influence of sex-biased forces

5.1 Introduction

The genetic structure of a natural population is the product of historic forces dictating how much overall variation is present, and contemporary forces which apportion variation at a local scale (Hewitt & Butlin 1997). Historical processes affect the population as a whole and may include events such as range expansions or population bottlenecks (Goodman *et al.* 2001, Ibrahim *et al.* 1996). Within a population the balance between mutation and genetic drift determines the distribution of genetic variation (Balloux & Lugon-Moulin 2002), but behavioural processes such as dispersal, aggregation of individuals and mating patterns will act to alter the dynamics of this balance (Lawson-Handley & Perrin 2007). Further, differential rates of mutation, modes of inheritance, recombination and selection within the genome and between the sexes will result in different regions of the genome having different underlying genealogical histories (Hedrick 2007).

Despite this complexity, many phylogeographic and population genetic studies rely on one marker type or a single locus, which can lead to erroneous conclusions when inferring the history of a population (Balloux 2010, Selkoe & Toonen 2006). This is for a number of reasons including size homoplasy (especially relevant for rapidly evolving markers such as microsatellites (Estoup *et al.* 2002)), independent evolution (especially relevant for single locus markers (Moore 1992, Nichols 2001)), incomplete lineage sorting (Maddison & Knowles 2006), inheritance bias between the sexes and the occurrence of sex-biased dispersal (Ellegren 2009, Hedrick 2007). By combining markers with different modes of inheritance and rates of evolution, historic and contemporary processes responsible for population structure can be determined and any ambiguities arising from conclusions based on different marker types have greater potential to be addressed and resolved. For example, a number of studies which have combined mitochondrial and autosomal microsatellite markers have been able to gain further insight into phylogeographic processes operating before and after the last glacial maximum, which would have been obscured if only one marker type was considered

(greater horseshoe bat (Flanders *et al.* 2009), tawny owl (Brito 2007), field vole (Heckel *et al.* 2005)).

Recently the X-chromosome has emerged as a useful marker to study population structure, both in isolation and when used in combination with autosomal markers to investigate sex-biased behavioural differences (Schaffner 2004). In the majority of mammal species the X-chromosome is present in two copies in females and a single-copy in males. Consequently the biparental, but asymmetric mode of inheritance allows markers located on the X-chromosome to be informative about both male and female population histories. This can be illustrated by considering cases in which the ratio of the effective population size (N_e) of X-chromosomal markers relative to autosomal markers departs from 0.75, as expected under neutral conditions. Deviations from this ratio can imply mutational differences, selection, unequal reproductive success and differences in the magnitude of dispersal between the sexes (Bustamante & Ramachandran 2009, Keinan & Reich 2010, Labuda *et al.* 2010).

A handful of recent studies on human populations demonstrate how this approach has been used to evaluate the role of sex-specific processes in shaping patterns of variation in the human genome. For example, Keinan *et al.* (2009) compared patterns of sequence variation on the X-chromosome and autosomes of African and non-African populations and reported that the X-chromosome had a smaller N_e in the non-African populations. They concluded that a smaller number of effective mating females and long distance male migration, resulting in increased autosomal gene flow relative to X-chromosomal gene flow, was accounting for this pattern. However, when Hammer *et al.* (2008) compared N_e of X-chromosomes and autosomes in a number of populations globally, they found the opposite result; an X:autosomal N_e ratio above 0.75. They concluded that although there were a number of possible factors that could explain this, a larger number of effective mating females, as would be observed under polygyny was the most parsimonious explanation.

Although these studies came to opposite conclusions (see Bustamante & Ramachandran 2009 for a discussion of possible reasons), they illustrate that autosomal and X-chromosomal markers can be compared to investigate sex-biased processes in a manner similar to more traditional mtDNA/autosomal and mtDNA/Y-chromosomal comparisons (e.g. Lorde *et al.* 2000). In fact comparing X-chromosomes and autosomes

may be more informative than comparing markers with uniparental inheritance (mtDNA and the Y-chromosome). Because both autosomes and the X-chromosome recombine during meiosis (the X-chromosome undergoing recombination when in females), loci on the X-chromosome and autosomes have independent population histories and this has the potential to increase the power to determine sex-biased processes with an X:autosome comparison relative to a mitochondrial:Y comparison (Wilkins & Marlowe 2006). Because mtDNA and Y-chromosomes do not recombine, they are essentially single locus markers. This can render them misleading when investigating demography as the stochastic nature of the coalescent process means that they only represent a single version of population history (Harpending *et al.* 1998, Nichols 2001). In addition markers on the Y-chromosome and mtDNA may be linked to regions under selection, which can obscure purely demographic processes (Balloux 2010). Indeed a number of studies have advocated the use of multiple, unlinked loci to investigate population structure, rather than replying on single locus markers (Ballard & Whitlock 2004, Rubinoff & Holland 2005).

However, whilst different levels of variation on X-chromosomes and autosomes may be a consequence of sex-biased differences in behaviour, it should also be noted that other factors may act to alter the N_e ratio of 0.75. For example because of the mode of inheritance of the two marker types, the X-chromosome is generally more influenced by genetic drift (due to the relatively smaller N_e) and patterns of population structure tend to be more pronounced irrespective of sex-biased processes. For example, Pool & Nielsen (2007) demonstrated that, in a range of taxa, certain demographic processes affecting the entire population (i.e. both sexes together) tended to lead to an elevated X-chromosome diversity relative to autosomes.

Furthermore, it is argued that deviations from 0.75 can be influenced by differential selection on the X-chromosome and autosomes. Because the X-chromosome is hemizygous when in males, this increases the likelihood that local adaptation will lead to higher levels of X-chromosome differentiation (Hedrick 2007). Moreover background selection is expected to alter relative levels of X:autosomal variation depending on whether it is negative or positive. For example, background selection on the X-chromosome was invoked to explain a reduction in X effective population size in derived populations of the house mouse (Baines & Harr 2007).

An additional complicating factor can arise from the different mutation rate in males and females. It is well known that overall rates of mutation are unequal between males and females (Goeting-Minesky & Makova 2006, Haldane 1947, Li *et al.* 2002). Males are often reported as having a higher mutation rate, although there are exceptions (Crow 2006, Huang *et al.* 2002). Because the X-chromosome spends proportionally more time in females than males we may expect diversity to be lower due to the lower female mutation rate.

The potential effects of demography, selection and mutation are all factors that can be considered when comparing patterns of diversity and N_e between autosomal and X-linked markers. Indeed after accounting for all the aforementioned effects Hammer *et al.* (2008) found that sex-biased forces were still required to explain the elevated levels of diversity on the X-chromosome.

In this chapter, I use a combination of haplotypes from the mitochondrial *ND1* gene, autosomal microsatellites and X-linked microsatellites to investigate the population structure of *M. daubentonii* in Europe. This is with a view to elucidating population dynamics pre- and post-glacially (by comparing a mitochondrial gene with a relatively low mutation rate and microsatellites with much higher rates of mutation) and to investigate mating behaviour and effective dispersal (gene flow) differences between the sexes (both by comparing mtDNA with maternal inheritance and X-linked/autosomal markers with biparental inheritance and by comparing structure and N_e of X-linked/autosomal markers to explore reasons for a deviation from neutral expectations).

Previous work on Daubenton's bat has suggested that there is a male bias to dispersal, as determined by greater population structure of mt Vs autosomal markers in Scotland (Ngamprasertwong *et al.* 2008) and greater microsatellite population structure in female bats than in males in the Yorkshire Dales (Angell 2008). In addition, there may be higher variance in reproductive success in males than in females, as evidenced by variable male paternity at swarming sites and summer roosts (Angell 2008, Senior *et al.* 2005). This is the first study which has investigated such sex-biased behavioural differences in *M. daubentonii* at a continental scale and the first study applying X-linked markers to investigate population structure in bats.

5.2 Methods

5.2.1 Samples

Details of samples used for combined analyses with mitochondrial and X-linked markers can be found in chapters 3 and 4 respectively. Sample sites for autosomal markers were the same as those used for X-linked markers (see chapter 4 for details). This resulted in wing punch samples from 848 adult bats comprising 100 females and 748 males.

5.2.2 DNA extraction

DNA was extracted from wing punches using an ammonium acetate precipitation method, adapted to a 96 well plate extraction as described in chapter 2.

5.2.3 Autosomal microsatellite amplification

14 microsatellite markers designed for cross-species amplification in bats by Camille Jan were used to amplify autosomal microsatellites. See Jan *et al.* (2010 in prep.) for details of markers and PCR conditions.

5.2.4 Analysis

5.2.4.1 Autosomal data

Microsatellite fragments were sized by Camille Jan using Genemapper version 3.9 (Applied Biosystems). The data were formatted in Microsoft Excel using Microsatellite toolkit (Park 2001), which tests for invalid alleles and formats data for input into population genetics software programmes. Non-canonical alleles were present in six of the microsatellite loci used in this study, suggesting flanking region variation and/or imperfect repeats. Consequently, a stepwise model of mutation was deemed inappropriate and further analyses using all 14 loci were restricted to using an infinite alleles model. The 8 loci which mutated in a stepwise fashion were analysed separately in some cases using a step-wise mutation model.

5.2.4.2 Diversity statistics

Measures of mean number of alleles per locus and observed heterozygosity were calculated in Microsatellite Toolkit. F_{IS} values were calculated for each locus and population, and then deviations from Hardy-Weinberg expectations were calculated using FSTAT version 2.9.3.2 (Goudet 1995). FSTAT was also used to calculate per-population allelic richness.

5.2.4.3 Population structure

Pair-wise F_{ST} values between the 17 regions (see Table 4.1 for details) were calculated in Arlequin version 3.11 (Excoffier *et al.* 2005).

In addition, an AMOVA test was performed to assign the amount of variation due to individual Vs locality Vs regional structure. AMOVA was performed in Arlequin.

5.2.4.4 Isolation by distance

Pair-wise F_{ST} and Rousset's (1997) $F_{ST}/(1- F_{ST})$ values for each of the original 46 sample populations were calculated using Arlequin. A corresponding matrix of the natural log of pair-wise geographic distances (km) was also produced (calculated from latitude and longitude, using the haversine formula, by Roger Butlin). A plot of Rousset's genetic distance ($F_{ST} /(1- F_{ST})$) Vs the natural log of Euclidean geographic distance was produced and the significance of any correlation analysed using a Mantel test with 1000 permutations performed using the isolation by distance web service (ibdws.sdsu.edu/).

5.2.4.5 Bottleneck

The program 'bottleneck' version 1.2.0.2 (Cornuet & Luikart 1997) was used to test for departure from expected heterozygosity per locus. Bottleneck defines expected heterozygosity not in terms of HWE but in terms of the number of alleles observed for a locus given the sample size, assuming neutrality and mutation-drift equilibrium. The rationale behind the test is that if a population has experienced a recent bottleneck (between 0.2-4 N_e generations in the past) then we should see an excess of

heterozygosity because, following a bottleneck, populations lose alleles faster than they lose heterozygosity. Conversely, if there has been a recent population expansion, we should observe a heterozygote deficiency because the number of alleles will increase faster than gene diversity. The program 'bottleneck' (Cornuet & Luikart 1997) was run using the step-wise mutation model (for the 8 loci with SMM dynamics) on all samples grouped into one European population and on samples grouped according to the 17 regions. The Wilcoxon test was used to test the significance of departure from expected H_e as it is more reliable when using a small number of loci and allows a test of both H_e excess and deficiency.

5.2.4.6 Clustering analyses

Clustering analyses for autosomal data were performed as described in chapter 4 using the programs STRUCTURE (Pritchard 2000) and BAPS (Corander *et al.* 2004). Prior population information was given as in chapter 4, representing 46 sample sites.

5.2.4.7 Detecting migrants

Geneclass version 2 (Piry *et al.* 2004) was used to detect any individuals which could be first generation migrants. L = L_home / L_max was selected as the parameter for likelihood computation. An MCMC approach with 10,000 repetitions and Paetkau (2004) re-sampling algorithm was used to detect the probability that each individual is a resident (i.e. not a first generation migrant) of the population to which it was assigned. Any individual with a P value below 0.01 was assumed not to be a resident and the probability of assignment to a different population was determined.

Geneclass analysis was restricted to detecting migrants between England and France. Populations from England and populations from France were grouped according to country to identify individuals which were not residents of their sampled population.

5.2.5 Comparison of autosomal/mitochondrial/X-linked markers

5.2.5.1 F_{ST}

In order to compare pair-wise F_{ST} values obtained using the three marker types, matrices of pair-wise F_{ST} (or in the case of mt markers Φ_{ST}) values from X-linked markers Vs autosomal markers, X-linked Vs mitochondrial clade A, X-linked Vs mt clade B, Autosomal Vs mt clade A and autosomal Vs mt clade B were plotted. The significance of any correlation between the matrices was determined using a Mantel test performed in GeneAlEx version 6.1 (Peakall & Smouse 2006).

5.2.5.2 Effective number of migrants

In order to compare levels of differentiation between the marker types, whilst accounting for N_e differences the effective number of migrants per generation (N_m) was calculated for each marker type as follows: Autosomal $N_m = [(1/F_{ST})-1)]/4$, X-linked $N_m = [(1/F_{ST})-1)]/3$, mt $N_m = [(1/F_{ST})-1]$.

5.2.5.3 Isolation by distance

The slope of the regression between genetic and geographic distance is expected to be equal to $1/4\pi D\sigma^2$, where D is equal to the effective population density and σ is equal to the effective dispersal distance (standard deviation of parent-offspring distances; Slatkin 1993). Thus the magnitude of the slope has the potential to tell us about the density of the population at a given geographic distance or the relative dispersal distance of individuals, if either value is already known. The approximate density of *M. daubentonii* individuals (per km^2) across Europe was calculated using the following information: The number of *M. daubentonii* individuals in the UK is estimated to be 150,000 (Harris *et al.* 2006) and the area of the UK is 244,820 km^2, as the density of *M. daubentonii* is expected to be slightly higher on the continent (John Altringham, pers. comm.) 150,000 was increased to 200,000 to take this into account and divided by the total area of the UK. This resulted in an estimate of 0.82 *M. daubentonii* individuals per km^2 of the study area. Because we needed to know effective density, and it is assumed that effective density will be less than census density, effective density was set at a range of 0.2 bats per km^2 to 1 bat per km^2. As both the slope and the density of

individuals are known we can use this information to calculate a range of effective dispersal distances using the following formula: effective dispersal distance = $\sqrt{1/4\pi Ds}$, where s is equal to the slope of the regression. This was done for all marker types, although not for data from mt clades A and B separately as the relationship between genetic and geographic distance was not significant for these analyses (see chapter 3).

5.2.5.4 Correlation between latitude/longitude and diversity

Population genetics theory predicts that during range expansion populations in more recently colonised areas will have lower genetic diversity than ancestral populations due to the loss of alleles through genetic drift during an expansion (note that in areas of admixture diversity may increase due to the mixing of lineages). To test this hypothesis and to examine the relative likelihood of a northwards V westwards expansion in *M. daubentonii,* regressions of latitude Vs measures of genetic diversity and longitude Vs measure of genetic diversity were performed in Excel. For autosomal microsatellites, diversity was defined as either observed H_e or allelic richness. For X-linked microsatellites diversity was defined as either allelic richness or expected H_e and for mitochondrial DNA diversity was defined as nucleotide diversity. Note that for mitochondrial data, analyses were performed on all sample sites and then on groups of samples split according to whether they contained individuals from clade A or clade B.

5.2.5.5 Correlation between clustering analyses and mitochondrial clades

The optimal number of clusters reported for autosomal and X-linked microsatellite data was two. As this coincided with the presence of two mitochondrial clades in the study area, a correlation of percentage assignment/clade membership for each region between each marker type was performed, and the significance of the correlation coefficient was tested using Excel. Note that mitochondrial clade B was presumed to be analogous to the microsatellite cluster found in central/southern Europe and mitochondrial clade A presumed analogous to the microsatellite cluster in Britain/north east Europe. To visualise this information in a geographic context, a pie chart comprising percentage membership to each cluster/clade for each marker type was superimposed onto a map of Europe according to geographic locality.

5.2.6 Comparison of X-linked/autosomal markers

5.2.6.1 Detecting migrants

The results from the 'detect migrants' function of Geneclass for X-linked Vs autosomal microsatellites were compared to see if either marker type proved more informative on geographic origin of individuals.

5.2.6.2 Comparison of effective population size between markers

Two methods were applied to calculate N_e (θ) for autosomal and X-linked microsatellites. Firstly the program Migrate version 3.1.2 (Beerli & Felsenstein 2001) was used to calculate θ for the European population as a whole using either X-linked θ ($3N_e u$) or autosomal θ ($4N_e u$). As θ is proportional to the effective population size multiplied by the mutation rate (u), we can use this measure to indirectly compare effective population size using the two marker types. The maximum likelihood mode of Migrate was chosen over the Bayesian method as it produced more consistent results over trial runs using a subset of the data. For both marker analyses the infinite-alleles model was used. Default parameters were used for each run with the following exceptions: The number of short chains was increased to 20, number of recorded genealogies 5000, a burn-in of 1000 trees and an interval of 40. For long chains the number of initial chains was increased to 4, number of recorded genealogies 100000, a burn-in of 1000 trees and an interval of 40. To increase the efficiency of searching four heated chains were used with temperatures of 1, 1.5, 3 and 1000.

Secondly because expected heterozygosity (H_e) for each marker type was known, we were also able to use this information to estimate θ using the following equations: For autosomal microsatellites $H_e = 4N_e u/(4N_e u+1)$ and for X-linked microsatellites $H_e = 3N_e u/(3N_e u+1)$. Therefore for autosomal microsatellites $\theta = H_e/(4-4H_e)$ and for X-linked microsatellites $\theta = H_e/(3-3H_e)$.

5.3 Results

5.3.1 Autosomal markers

Of the bats genotyped, only those with information for at least 8 of the 14 autosomal loci were included in further analyses. This resulted in data from 848 adult bats comprising 100 females and 748 males (see Table 5.2 for details). Details of allele size range and mutation model for each locus can be found in Table 5.1. Observed heterozygosity for each population ranged from 0.63 to 0.82. Number of alleles and allelic richness, averaged over all loci ranged from 4.43-10.14 and 3.45-4.21, respectively (Table 5.2). Deviations from Hardy-Weinberg expectations are provided in Table 5.2 as Wright's F_{IS} statistic across all loci. Eleven populations exhibited significant ($P < 0.05$) deviations from Hardy-Weinberg expectations, although after applying Bonferroni corrections none of the populations showed significant departure for more than two loci, bar population 36 which deviated from HWE at 4 loci. The loci out of HWE differed between populations suggesting that null alleles and non-neutrality are not issues.

Table 5.1. Per-locus allele size range and mutational model information. Loci are classed as following either a step-wise mutational (SMM) or infinite alleles mutational (IAM) model, as determined using Microsatellite Toolkit.

Locus	Allele size range (bp)	Mutational Model
A13	352-278 bp	SMM
A18	138-170 bp	SMM
A2	139-143 bp	SMM
A24	338-358 bp	IAM
A45	134-175 bp	IAM
B8	392-444 bp	IAM
CloneA2b	391-425 bp	SMM
ES43	369-397 bp	SMM
F19	411-441 bp	IAM
G2	343-373 bp	SMM
G30	245-310 bp	IAM
G31	197-225 bp	SMM
G6	102-142 bp	SMM
MSchreib3	256-291 bp	IAM

Table 5.2. Levels of genetic diversity per sample site and Wright's F_{IS} estimated from autosomal microsatellite data and averaged over all loci. Significant F_{IS} estimates (P <0.05) are highlighted in red. See Table 4.2 for sample site abbreviations.

Sample site	Sample size	H_o	SD	No Alleles	SD	Allelic Richness	F_{IS}
BB	15	0.77	0.03	8.36	3.27	4.01	0.04
LB	14	0.72	0.04	6.86	2.48	3.85	0.085
DE	19	0.76	0.03	8.79	2.97	4.01	0.068
FI	31	0.76	0.02	10.00	3.59	3.95	0.051
FG	16	0.73	0.03	8.36	2.65	3.94	0.082
PL	16	0.77	0.03	8.86	2.80	3.95	0.026
TR	15	0.80	0.03	7.79	2.36	4.01	0.006
FN	5	0.76	0.05	5.14	1.61	3.92	0.06
RQ	28	0.79	0.02	9.57	3.03	3.95	0.008
LP	14	0.80	0.03	8.36	2.84	3.99	-0.006
LC	20	0.78	0.02	9.00	3.28	3.91	0.009
SE	18	0.79	0.03	9.07	3.45	4.00	0.011
GM	11	0.76	0.03	8.29	2.84	4.13	0.053
TI	16	0.77	0.03	9.50	3.08	4.21	0.071
BH	6	0.77	0.05	4.93	1.73	3.56	-0.038
MA	21	0.73	0.03	7.57	3.84	3.45	-0.018
PO	26	0.80	0.02	9.00	3.26	3.95	-0.004
IT	25	0.77	0.02	7.64	3.08	3.69	-0.013
SI	6	0.75	0.05	5.36	1.69	3.76	0.017
EG	9	0.74	0.04	5.93	1.98	3.59	0.02
CA	26	0.81	0.02	9.86	3.78	4.03	0.003
LU	19	0.76	0.03	9.50	3.28	4.10	0.062
SX	22	0.82	0.02	9.07	3.63	3.99	-0.008
KT	29	0.77	0.02	9.86	3.57	3.92	0.046
SC	14	0.78	0.03	7.79	2.33	4.10	0.062
WP	15	0.73	0.03	7.79	3.09	3.82	0.064
SCAN	12	0.71	0.04	6.00	1.92	3.49	0.022
SCBN	12	0.70	0.04	5.86	2.18	3.45	0.019
SCBS	12	0.77	0.04	6.14	2.14	3.79	0.032
WC	9	0.78	0.04	6.71	2.64	3.87	-0.001
AHM	10	0.77	0.04	6.57	2.82	3.77	0.015
AR	6	0.75	0.05	4.93	2.09	3.73	0.025
BP	6	0.77	0.05	5.21	1.97	3.75	0.003
BD	32	0.79	0.02	9.36	3.50	3.95	0.016
BR	4	0.70	0.06	4.43	1.55	3.80	0.075
BU	40	0.74	0.02	9.86	3.98	3.87	0.068
BG	19	0.78	0.03	8.71	3.22	3.96	0.035
CB	32	0.82	0.02	9.50	3.72	3.90	-0.023
DC	32	0.77	0.02	9.14	3.37	3.83	0.025
DK	6	0.80	0.05	5.29	2.16	4.02	0.025
FA	4	0.78	0.06	4.71	1.44	3.99	0.054
GR	49	0.80	0.02	10.14	3.55	3.95	0.01
IK	44	0.75	0.02	9.79	3.60	3.90	0.065
KB	29	0.79	0.02	9.64	3.54	3.97	0.012
MT	6	0.63	0.06	4.50	1.51	3.84	0.199
WB	27	0.78	0.03	8.79	3.02	3.86	0.011

5.3.1.1 Population differentiation

F_{ST} across all loci and populations was low but significant at 0.014 ($P < 0.01$). Pair-wise comparisons of the 17 regions revealed some differentiation between sample sites in Europe. Out of 153 pair-wise comparisons 46 (30%) were significant ($P<0.05$, without multiple-test correction) (Table 5.3). Most regions in the British Isles were not significantly differentiated from each other, bar Ireland which was differentiated from every UK population except Withcall. The Italian regions were significantly differentiated from French and Spanish populations, but not from most of the regions further north and east. Spain exhibited significant differentiation from most regions in the UK and north/east Europe and conversely the Scandinavian populations were differentiated from France/Spain but not from other mainland European populations.

In order to analyse partitioning of genetic variation using AMOVA, sample sites were grouped into 17 regions. The vast majority of the variance (98.63%) could be explained by the within population factor (i.e. at the individual level), with very little, but significant, variation explained by either the among-groups (1.1%) or among-populations (0.27%) factor. This suggests that there is very little genetic subdivision between colonies and that this lack of subdivision persists over a large geographic scale.

The Mantel test demonstrated significant Isolation by Distance across Europe ($r = 0.29$, slope = 0.64^{-5}, $R^2 = 0.084$, $P < 0.05$ Figure 5.1)

Table 5.3 Pair-wise F_{ST} values between *M. daubentonii* regional populations in Europe using autosomal microsatellites. Values in bold are significant ($P < 0.05$).

	BB	N	F	SFR	NFR	GM	TI	IR	P	SP	CA	LU	SUK	NYM	SCOT	WC	YD
BB	0.000																
N	0.000	0.000															
F	0.010	-0.004	0.000														
SFR	0.002	0.013	**0.019**	0.000													
NFR	0.009	**0.018**	**0.018**	0.003	0.000												
GM	0.013	-0.007	**0.030**	**0.028**	**0.029**	0.000											
TI	0.019	-0.014	0.000	**0.034**	**0.044**	-0.003	0.000										
IR	**0.043**	**0.029**	**0.059**	**0.036**	**0.024**	0.002	**0.048**	0.000									
P	-0.011	-0.002	0.001	0.003	-0.003	0.000	0.010	**0.034**	0.000								
SP	0.004	**0.025**	**0.035**	-0.003	0.004	**0.044**	**0.061**	**0.043**	0.015	0.000							
CA	-0.007	-0.006	0.009	-0.004	-0.002	0.005	0.018	0.012	-0.004	-0.004	0.000						
LU	0.002	0.001	0.012	-0.005	-0.005	0.011	0.023	0.014	-0.001	0.000	-0.015	0.000					
SUK	**0.016**	**0.022**	**0.028**	**0.016**	**0.022**	0.008	**0.033**	**0.051**	0.000	**0.031**	**0.019**	**0.016**	0.000				
NYM	0.013	0.022	**0.041**	0.006	0.014	0.038	**0.044**	**0.060**	0.016	0.020	0.015	0.013	0.019	0.000			
SCOT	**0.019**	0.014	**0.027**	**0.025**	**0.027**	-0.007	0.022	**0.029**	0.000	**0.040**	0.017	0.017	0.000	0.019	0.000		
WC	-0.006	-0.003	-0.006	-0.012	0.001	0.006	-0.002	0.046	-0.018	0.009	-0.005	-0.018	-0.017	-0.007	-0.006	0.000	
YD	0.005	**0.013**	**0.024**	0.003	**0.009**	0.014	**0.032**	**0.040**	-0.002	**0.016**	0.006	0.005	0.004	0.004	**0.009**	-0.013	0.000

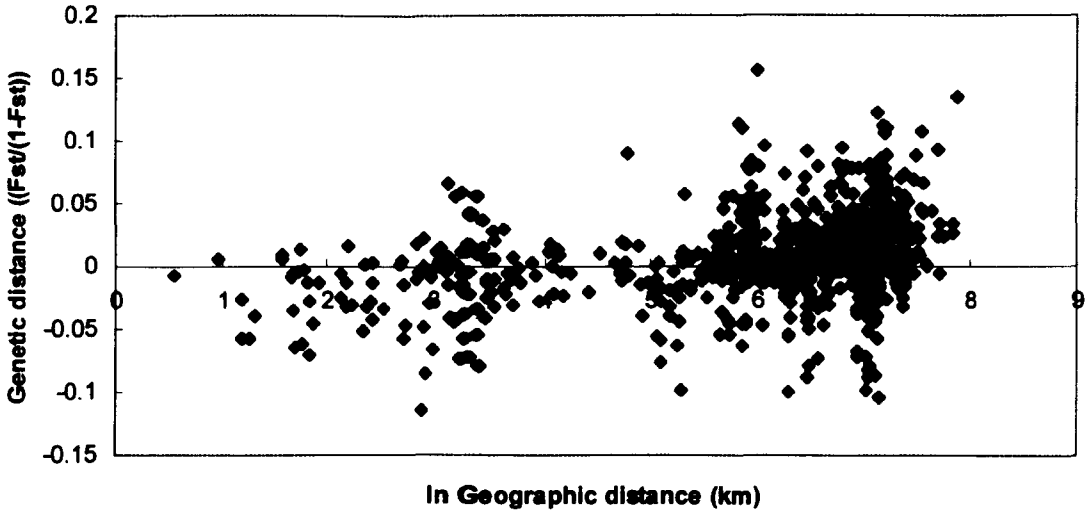


Figure 5.1. Plot of the correlation between Rousset's genetic distance ($F_{ST}/(1-F_{ST})$) and the natural log of Euclidean geographic distance (km) for autosomal microsatellites in *M. daubentonii* populations across Europe.

5.3.1.2 Bottleneck

When all populations were considered together significant heterozygote deficiency was observed (Wilcoxon test $P < 0.05$) suggesting recent (within the last 0.2-4 N_e generations) population expansion. When regions were considered individually Finland, France, Ireland, Italy, Spain, Switzerland, southern UK and the Yorkshire Dales all exhibited significant heterozygote deficiency ($P < 0.05$). Withcall showed significant heterozygote excess and a shifted mode ($P < 0.05$), suggesting a recent bottleneck. The remaining regions showed no deviations from expected heterozygosity ($P > 0.05$).

5.3.1.3 Clustering analyses

When no prior location information was included very little structure could be found using either STRUCTURE or BAPS and the most likely K was one. Using STRUCTURE when prior population information was used to aid clustering likelihood values, ($Pr(X|K)$) increased up until K=6, although above K=2 there was substantial variance among runs (Figure 5.2a.). A plot of the absolute rate of change of K (Evanno *et al.* 1995) showed that the highest level of structure was at K=2 (Figure 5.2b). Although there is some suggestion of significant clustering when K=3. Following this, a

plot of percentage Q values for each sample site is shown when K=2 and when K=3 in Figure 5.3.

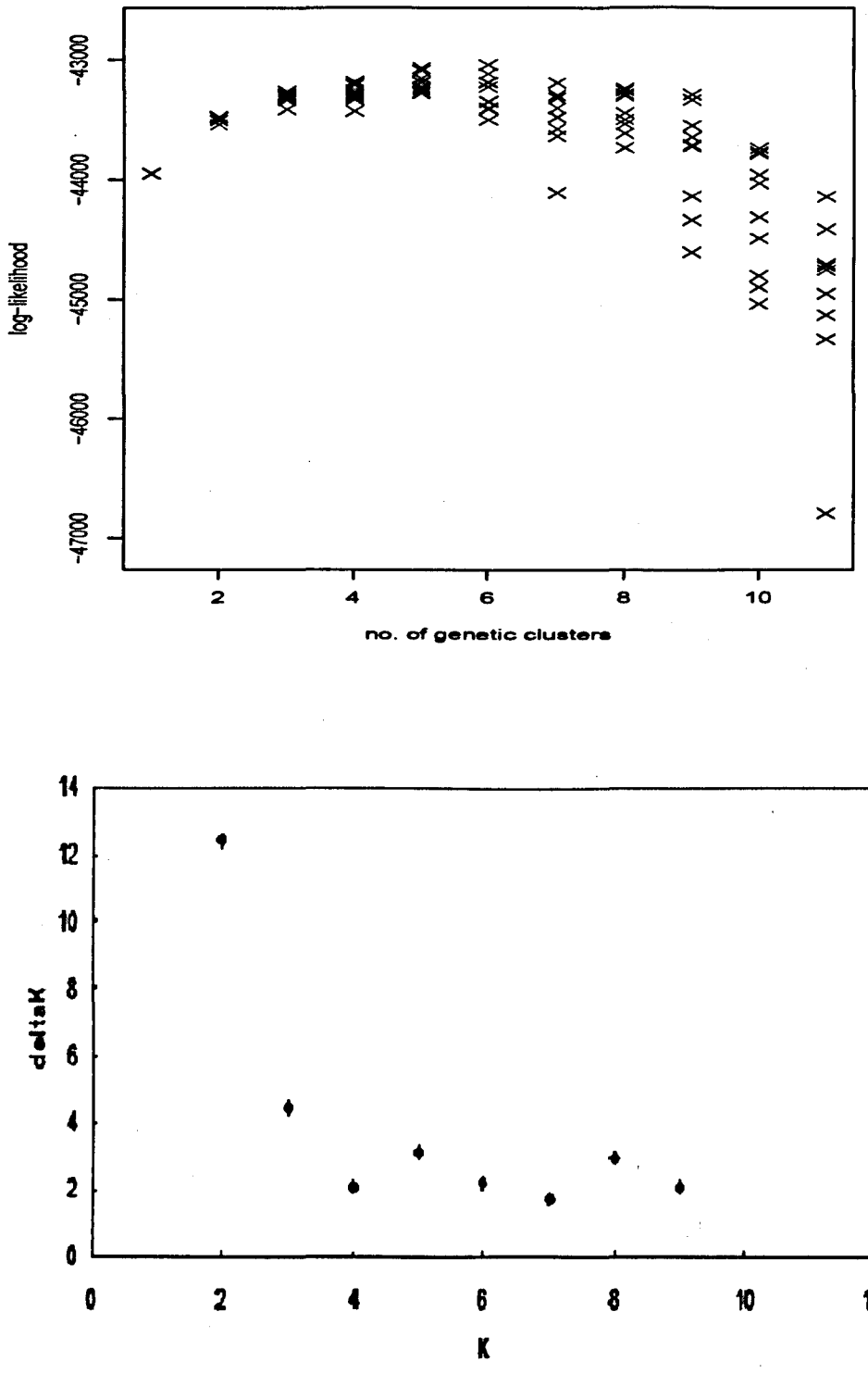


Figure 5.2 a) Log-likelihood values for K=1-10 for ten independent runs each. b) Absolute rate of change of log-likelihood values for increasing values of K.

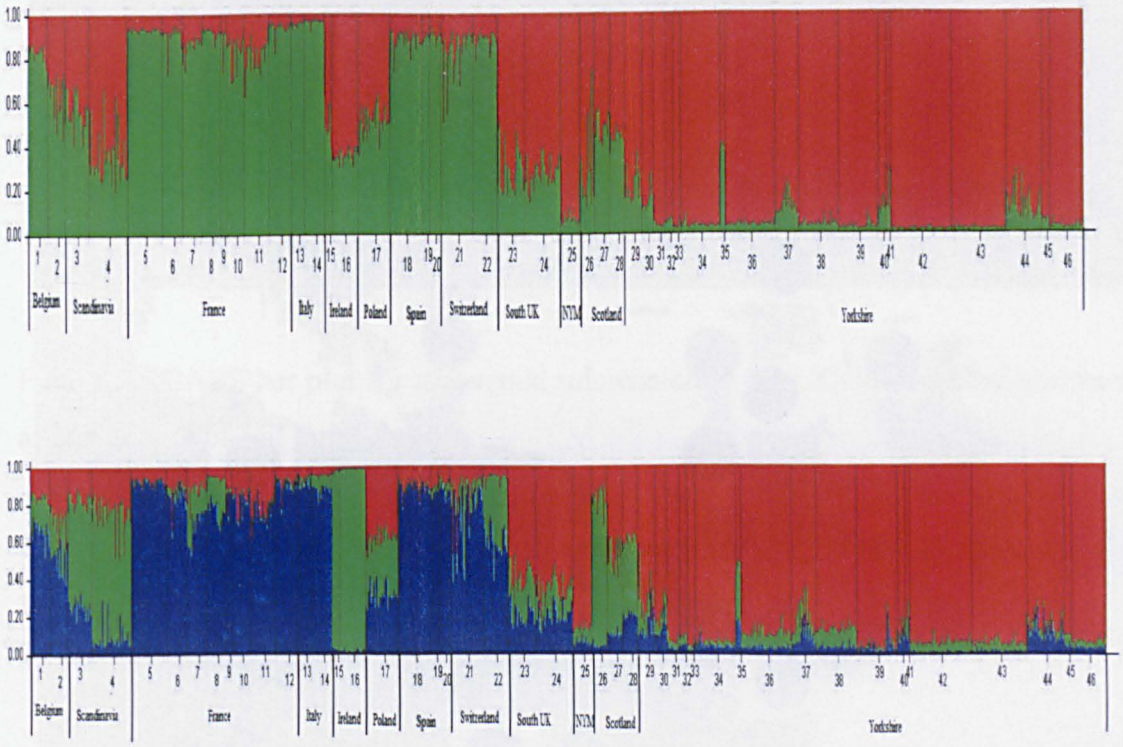


Figure 5.3. Bar plot of structure output for a) K2 b) K3. Each line represents the proportion of assignment (Q) of each individual to a given cluster. Populations are separated by black lines. For details of population names corresponding to numbers see chapter 4.

At K=2 there is a clear separation of samples from the Yorkshire Dales from Mainland Europe, with populations in north eastern Europe (Denmark, Finland, Poland), Belgium and Ireland exhibiting substantially mixed membership (Figure 5.3a). At K=3 (Figure 5.3b) Britain and mainland Europe are still clearly split, and there is further separation of Ireland and Scottish Northern Clade A populations. Denmark and Finland show majority membership to this Irish/north Scottish cluster.

Figure 5.4. Pie charts representing per region cluster assignment at K=2 superimposed according to geographic location. Cluster A is purple and cluster B is blue.

The most likely number of clusters reported when using BAPS at all prior values of K was 3 (Table 5.3, Figure 5.5). With England, mainland Europe/Scotland and Ireland forming three separate clusters. Note also that Browgill Cave in Yorkshire clustered with mainland Europe/Scotland rather than with England. There was no suggestion of a separate grouping of Scandinavia using BAPS.

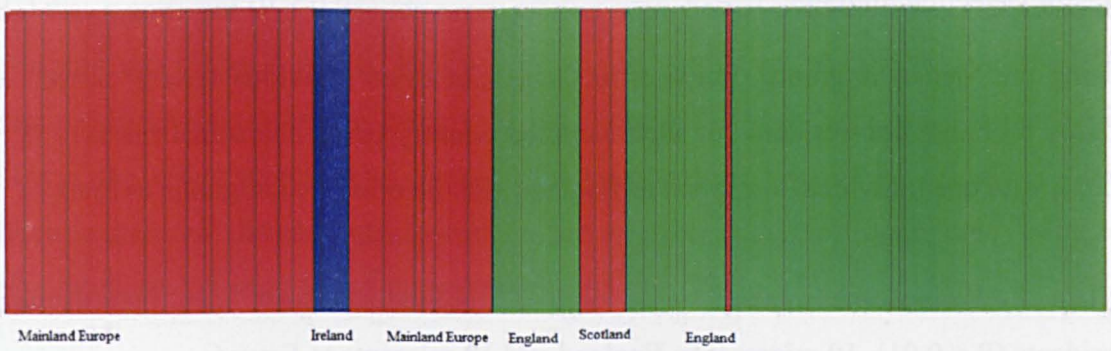


Figure 5.5 BAPS bar plot for autosomal microsatellite data. Coloured blocks represent assignment to one of three clusters.

Table 5.4. Logml values for most likely K determined using BAPS at maximum K=5, 10 and 15.

Maximum K assumed	Most likely K (% confidence)	Logml
5	3 (100%)	-44585.24
5	3 (100%)	-44585.24
5	3 (100%)	-44585.24
5	3 (100%)	-44585.24
5	3 (100%)	-44585.24
10	3 (100%)	-44585.24
10	3 (100%)	-44585.24
10	3 (100%)	-44585.24
10	3 (100%)	-44594.05
10	3 (100%)	-44585.24
15	3 (100%)	-44594.05
15	3 (100%)	-44585.24
15	3 (100%)	-44594.05
15	3 (100%)	-44585.24
15	3 (100%)	-44585.24

5.3.1.4 Detecting migrants

Using the 'detect migrants' function in Geneclass three individuals were classed as highly probable non-residents of their sampled country using autosomal markers ($P < 0.01$), with one individual a migrant to England and the other two migrants to France.

When X-chromosome markers were used 22 individuals were identified as non-residents ($P < 0.01$), 10 migrants to England and 12 migrants to France.

It should be noted that X-linked and autosomal markers did not identify the same individuals as potential migrants.

5.3.2 Comparison of autosomal/X-linked/mitochondrial analyses

5.3.2.1 Population differentiation and isolation by distance

Table 5.5. Comparison of effective number of migrants (N_m), AMOVA and range of effective dispersal distances (km) for the three marker types.

	Marker Type				
Statistic	*ND1*	*ND1* Clade B	*ND1* Clade A	Autosomal microsatellite	X-linked microsatellite
N_m	0.8	0.7	1.4	17.6	9.5
% variation within populations	44.56%	40.80%	57.23%	98.63%	96.56%
% variation among populations	7.54%	0.88%	2.83%	0.27%	1.98%
% variation among regions	47.90%	58.32%	39.94%	1.10%	1.46%
Effective dispersal distance (km)	11-25	NA	NA	36-81	28-63

Overall N_m was greatest for autosomal markers, followed by X-linked microsatellites and then mt markers (Table 5.5). For both X-linked and autosomal microsatellites the vast majority of variation was at the individual level, with very little differentiation explained by either the among locality or region factors. For the *ND1* gene, approximately equal levels of variation were attributable to the individual and among region factors, with much less variation between localities (Table 5.5). This suggests that there is significant mitochondrial gene flow at a local scale but that gene flow is far lower over large geographic distances.

IBD was significant ($P < 0.05$) for every marker type, however when mitochondrial data were split into clades A and B, IBD was no longer significant. As the slope of the IBD relationship and approximate density of bats across the study area were known, this allowed the effective dispersal distance for each marker type to be calculated. The effective dispersal distance was lowest for the mitochondrial *ND1* gene, then X-linked microsatellites followed by autosomal microsatellites (Table 5.5).

Of the Mantel tests of pair-wise genetic distance between the marker types the only significant correlations were between X-linked and Autosomal F_{ST} and X-linked and overall mitochondrial Φ_{ST} (Table 5.6). In general there appeared to be a trend towards higher F_{ST} values for X-linked microsatellites over autosomal microsatellites (Figure 5.6b) and unexpectedly high X-linked F_{ST} values were often linked with comparatively lower mitochondrial Φ_{ST} values (Figure 5.6a).

Table 5.6 Correlations between pair-wise F_{ST} values from different marker types. Significant correlations are highlighted in bold.

Comparison	R^2	P
Autosomal/X-linked	**0.08**	**0.03**
Autosomal/MT all	0.0008	0.41
Autosomal/MT Clade B	0.05	0.12
Autosomal/MT Clade A	0.01	0.27
X-linked/MT all	**0.08**	**0.02**
X-linked/MT Clade B	0.09	0.06
X-linked/MT Clade A	0.04	0.11

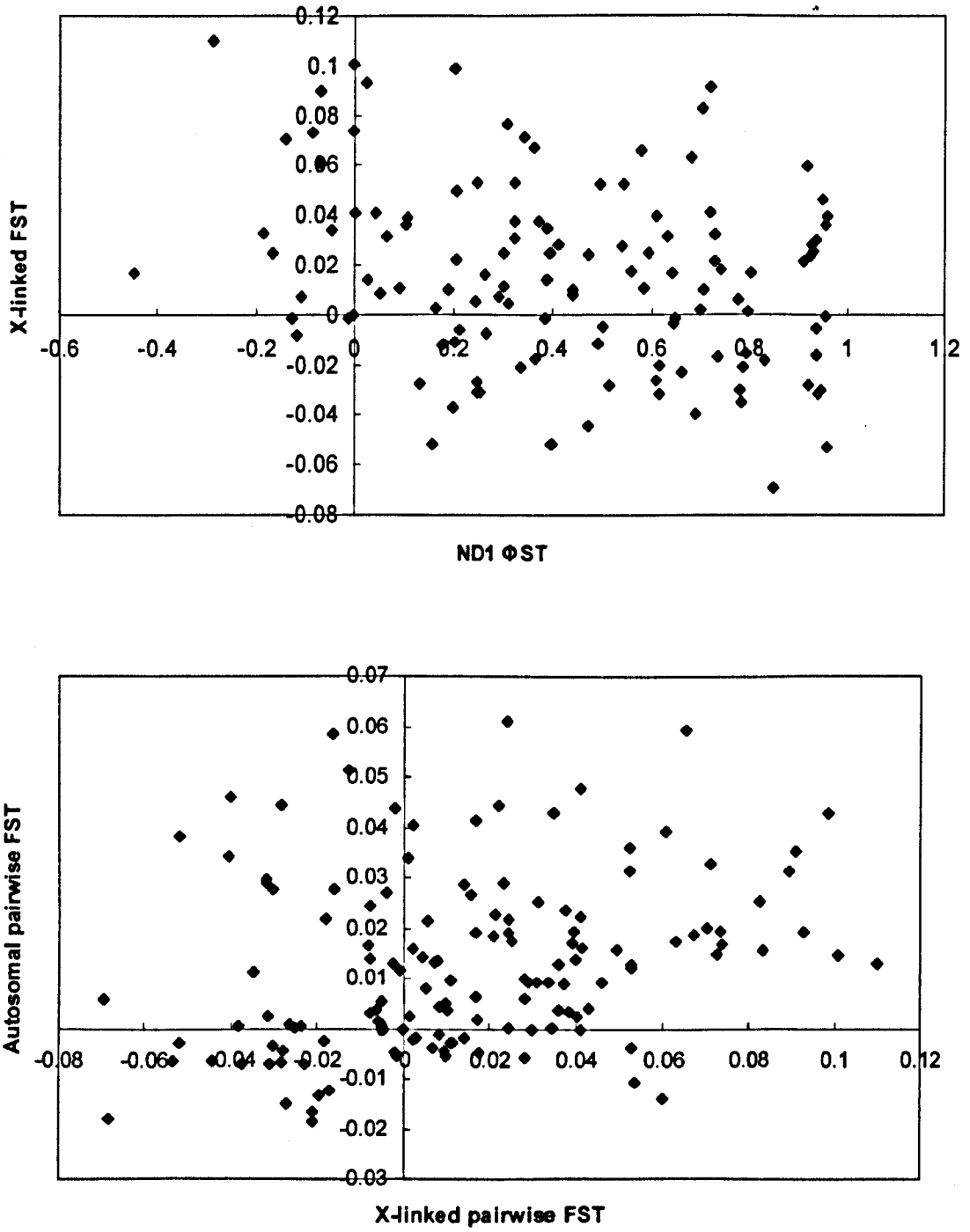


Figure 5.6. Correlation between a) mitochondrial Φ_{ST} and X-linked F_{ST} b) autosomal and X-linked F_{ST}

Analyses with all three marker types resulted in classification of genetic structure into two groups (two mitochondrial clades and two clusters for X-linked and autosomal microsatellites). Of the correlations performed between marker types, assignment of individuals to X-linked clusters was significantly correlated with assignment to mitochondrial clusters ($P < 0.01$) and autosomal clusters ($P < 0.01$). However,

assignment to autosomal clusters was not correlated with assignment to mitochondrial clades (Table 5.7).

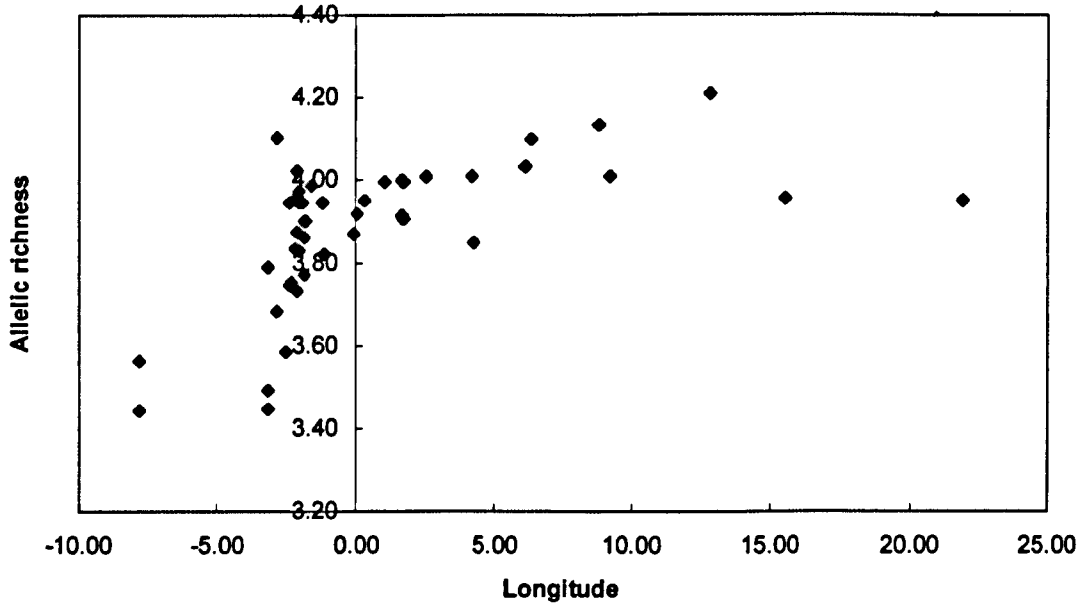
Table 5.7. Correlation between cluster assignment between the marker types. Significant correlations are highlighted in bold.

Comparison	R^2	P value
Mitochondrial/Autosomal	0.16	0.1617
Mitochondrial/X-linked	0.56	**0.0021**
Autosomal/X-linked	0.67	**0.0003**

Viewed in a geographic context, regions in France and Spain all show majority membership to the southern cluster/clade (B) for each marker type. For the northern cluster/clade (A) the Yorkshire Dales and Finland exhibit majority membership. Belgium has mixed membership in all three cases (Figure 5.8). However there are also some striking differences (Figure 5.8). Using mitochondrial markers most regions within the UK (bar the Yorkshire Dales and some Scottish localities) comprised individuals belonging to the southern mitochondrial clade, however when X-linked or autosomal microsatellites are used individuals from British regions have substantial membership to clusters more associated with the UK/north east Europe than southern Europe. In general there was congruence in cluster assignment between X-linked and autosomal markers, however autosomal markers appear to place England in a more distinct cluster than X-linked markers. This is reflected in the percentage of assignment to the two autosomal and X-linked clusters for north eastern European/UK populations. Using X-linked markers these regions are all strongly assigned to a "northern" cluster, however when autosomal markers are used the regions in north east Europe now have a substantial assignment to the cluster which appears more strongly associated with southern Mainland Europe.

5.3.2.2 Geographic correlates of diversity

Of the three marker types autosomal microsatellites showed a significant increase in allelic richness with increasing longitude ($P < 0.01$, $R^2 = 0.3$, Figure 5.7a) and X-linked microsatellites showed a significant increase in expected H_e with increasing longitude ($P < 0.01$, $R^2 = 0.28$, Figure 5.7b). All other comparisons were non-significant.



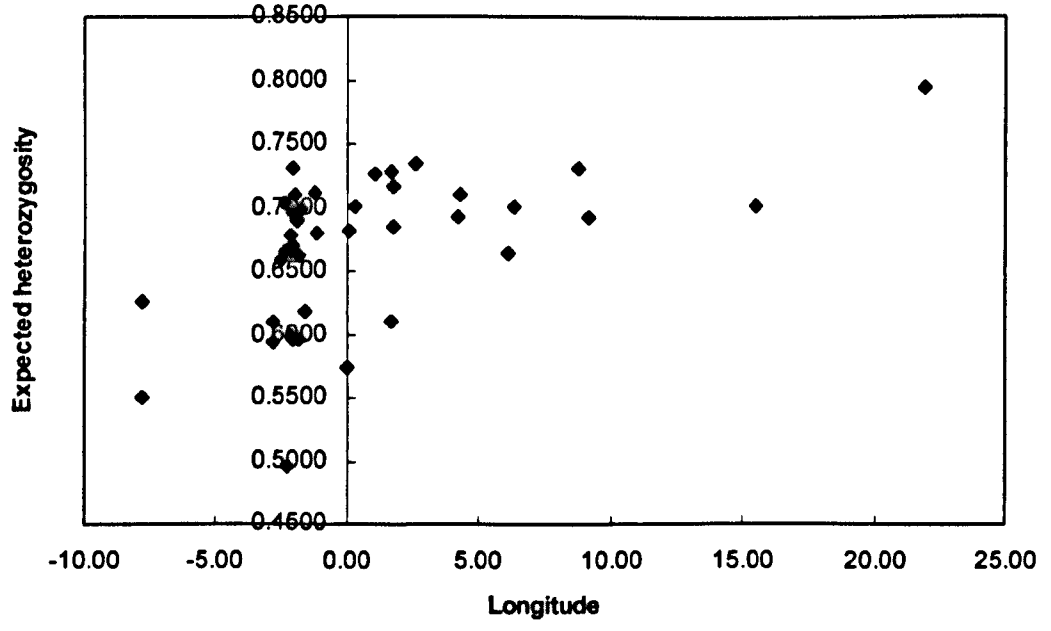


Figure 5.7 Correlation between longitude and a) allelic richness for autosomal microsatellites b) Expected H_e for X-linked microsatellites across Europe.

5.3.2.3 θ for X-linked and autosomal microsatellites

Estimates of θ (N_{eu}) using the Migrate method were as follows: Autosomal microsatellite θ = 1.26 (ln(L) = -1896.5, 95% confidence intervals; lower = 1.24, upper

= 1.29), X-linked microsatellite θ = 1.06 (ln(L) = - 113.9, 95% confidence intervals; lower = 1.03, upper = 1.10). This is an X:Autosomal ratio of 0.84, which is greater than the ratio we expect under neutral conditions (0.75).

Estimates of θ obtained from H_e were as follows: autosomal microsatellite θ = 1.50, X-linked θ = 1.21. This is an X:Autosomal ratio of 0.81, which is similar to the ratio obtained using Migrate.

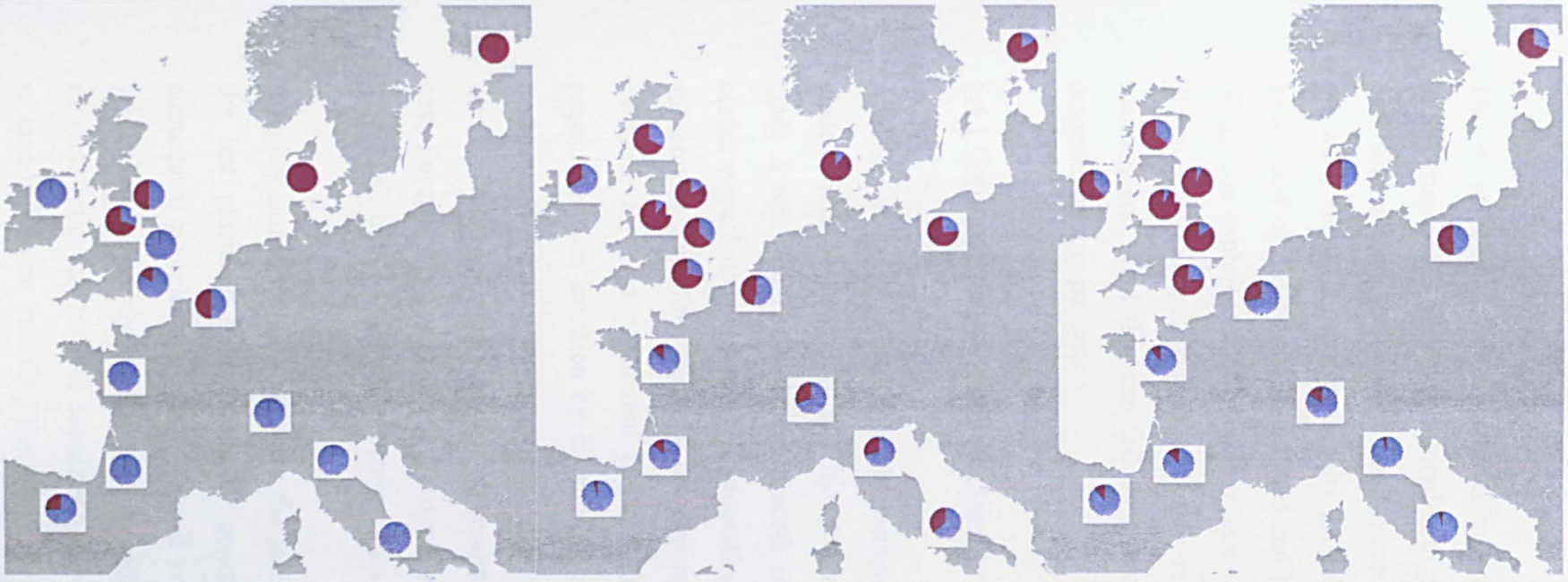


Figure 5.8 Percentage assignment of individuals to cluster A or B according to region for a) mtDNA b) X-linked microsatellites c) autosomal microsatellites. Cluster A is purple and cluster B is blue. Note that no Scottish populations were included in the mtDNA figure as genetic data was obtained from haplotypes deposited in Genbank, rather than from sequencing of individuals.

5.4 Discussion

The use of genetic markers with different modes of inheritance and mutational patterns in this chapter has allowed exploration of processes operating at different scales. On the one hand congruence between the markers increases confidence in the conclusions. On the other hand, contrasting results afford a better resolution of recent demographic events and of differential mutational and behavioural processes. The remainder of this discussion explores similar patterns between marker types and the implications from a phylogeographic perspective and then moves on to look at discrepancies between markers and what this can tell us about more recent events and about sex-biased dispersal/mating patterns.

5.4.1 Congruent patterns between mitochondrial and microsatellite markers

Analyses of genetic diversity at the mitochondrial *ND1* gene revealed two intraspecific clades within western Europe. Clade B was found in isolation in southern Europe, and clade A was found in isolation in north eastern Europe. In intermediate areas the two clades were found in sympatry, suggesting expansion from two different refugia with subsequent gene flow between the two populations. However the lack of any strong evidence for past population fragmentation meant that the idea of one European population with isolation by distance could not be ruled out. Clustering analyses for nuclear and X-linked microsatellites also supported a division of European populations into two clusters which roughly represented central/southern Europe and Britain/north east Europe. The congruence between all three marker type lends support to the hypothesis that the two clades represent a signal of past population fragmentation.

Roughly similar patterns between the three marker types suggest that isolation during the last glacial maximum shaped genetic diversity at all marker types, but that subsequent postglacial contact allowed contemporary gene flow between populations which had previously been isolated. Broadly congruent patterns between mt and nuclear markers reflecting historic population structure have also been found for a broad range of taxa in Europe including the tawny owl (Brito 2007), white-tailed eagle (Honnen *et al.* 2010), nine-spined stickleback (Shikano *et al.* 2010), grayling (Koskinen *et al.* 2002), common shrew (Yannic *et al.* 2008) and the European rabbit (Queney *et al.*

2001), and in other bat species, for example *Rhinolophus ferrumequinum* (Flanders *et al.* 2009) and *Miniopterus schreibersii* (Pereira *et al.* 2009). Because the *ND1* gene is under selection it evolves relatively slowly in comparison to neutral markers which are not protein coding. Therefore, it is potentially useful for resolving ancient demographic events. Conversely it may be less informative when investigating population structure present as a consequence of processes operating on a more recent time frame, because it may not evolve sufficiently rapidly to record current gene flow (Avise 2004). It should also be noted that because *ND1* is a coding gene, patterns of population structure at this marker may reflect selection as opposed to purely demographic processes (Balloux 2010). Microsatellite markers in this study are expected to be neutral and allow the corroboration of past events inferred from the *ND1* gene but also allow greater insight into processes which have occurred recently (Balloux & Lugon-Moulin 2002).

A comparison of percentage assignment of regional populations to either mt clade A/B or microsatellite clusters A/B revealed a strong correlation between mt and X-linked markers, but this comparison was not significant between mt and nuclear markers. This suggests that X-linked microsatellites may better record signatures of ancient events than autosomal microsatellites. There are a number of possible explanations for this. It could be a consequence of reduced admixture of the X-chromosome in comparison to autosomes, for example humans show increased X-chromosome admixture between regional populations due to more widespread female dispersal and gene flow (Bedoya *et al.* 2006). If the opposite is true in *M. daubentonii* and female dispersal is less prevalent, then this may act to reduce X admixture as males only transmit one copy of the X-chromosome. Between species, reduced X-chromosome introgression has been observed for other mammals such as mouse (Macholan *et al.* 2007) and hare species (Melo-Ferreira *et al.* 2009), probably as a consequence of X-linked hybrid male unfitness. Another explanation may involve the increased linkage disequilibrium recorded on the X-chromosome. Population history has been shown to have an impact on background levels of linkage disequilibrium within populations that is further enhanced within recombination-poor genomic regions in human populations in Europe (Pereira *et al.* 2007). Because the X-chromosome has a lower overall recombination rate than autosomes (as the X-chromosome only recombines when in females), X-linked markers could therefore record the footprint of a demographic event for longer than autosomal markers, which may account for the greater congruence in population structure between X-linked and mitochondrial markers than between autosomal and

mitochondrial markers. Alternatively the signal of a demographic event could persist for a similar time in both marker types, but because haplotypes can be observed directly in males using X-linked markers it may be more readily detectable on the X-chromosome (Schaffner 2004).

On a Europe-wide scale all marker types exhibited a pattern of isolation by distance. However, within mtDNA clades there was no significant IBD pattern. IBD is only expected when equilibrium between mutation and genetic drift has been reached (Slatkin 1993). Therefore it is possible that populations comprising individuals from the two clades are still expanding and have not had time to reach equilibrium (Crispo & Hendry 2005). The IBD slope was far greater for the mitochondrial *ND1* gene (as a whole) than for either of the microsatellite markers. As mt sequences have a smaller effective population size (1/4 of autosomes, 1/3 of the X-chromosome) it may be expected (under neutral expectations) that the slope would be steeper. However, even after considering differences in effective population size between the markers the slope was still steeper than expected. On a continental scale that is larger than the dispersal capability of an organism, patterns of IBD are usually interpreted as a reflection of colonisation in a stepping stone fashion (Ibrahim *et al.* 1996). This suggests that to a certain extent gene flow is limited by the maximum dispersal distance of *M. daubentonii*. Under mutation-drift equilibrium the slope of the relationship between genetic and geographic distance is a product of the effective dispersal distance and the effective density of a population (Rousset 2000). As slope and effective density were known, this allowed the effective dispersal distance to be calculated for each marker type. Effective dispersal distance for mitochondrial DNA was smaller (11-25 km) than for either microsatellite marker (autosomes = 36-81 km, X-chromosome = 28-63 km) and may reflect the shorter colonisation distances covered by females when forming a new breeding colony. As autosomes and the X-chromosome are transmitted by both sexes the shallower slope of IBD for both markers may reflect longer distances covered by males when moving to a new breeding colony. However, it should be noted that the pattern of IBD observed for mtDNA markers overall could also be an artefact of the mixing of populations from the two mitochondrial clades. A similar pattern of IBD differences between nuclear and mitochondrial markers has been found in *Rhinolophus monoceros* in Asia, which was also attributed to reduced female gene flow (Chen *et al.* 2008). Similarly IBD across Europe in the Eurasian badger was explained by philopatry (although this was not sex-biased) (Pope *et al.* 2005). Conversely *Myotis myotis*

populations in Europe demonstrated significant IBD at nuclear markers but no association between genetic and geographic distance for mitochondrial DNA. However, this was attributed to extreme female philopatry (Castella *et al.* 2001).

5.4.2 Contrasting patterns between mitochondrial and microsatellite markers

Although patterns of genetic structure were broadly similar across a large geographic scale, on a regional scale there were discrepancies. For example, populations composed exclusively of one mitochondrial clade contained individuals with substantial (over 20%) levels of admixture between the two main microsatellite clusters for both the X-chromosome and autosomes, for autosomal markers this was especially prevalent in north eastern European populations suggesting different contact zones for microsatellite V mitochondrial markers. A similar study comparing population history in the tawny owl in Europe also found that the position of contact between two clades differed between mitochondrial and nuclear markers (Brito 2007). Because zones in which admixture occurs are dependent on the balance between selection and dispersal, these two factors may produce different signals across loci and consequently congruence between markers may break down (Barton & Hewitt 1985). As behavioural processes such as sex-biased dispersal could also account for the discrepancies, they will be discussed later in this section in the light of measures of population differentiation.

Because microsatellites have a high mutation rate in comparison with coding regions they are traditionally more appropriate for investigating local and contemporary population structure (Selkoe & Toonen 2006). Microsatellites did not reveal substantially greater population structure than the *ND1* gene, and this is likely due to the high dispersal capacity of *M. daubentonii* which will act to homogenise gene frequencies over larger geographic scales (Burland & Worthington-Wilmer 2001). This is in contrast to similarly sized ground dwelling mammals, e.g. the common vole (Heckel *et al.* 2005), but is congruent for results from many other bat species which also show little nuclear structure at a wide spatial scale, for example *Myotis nattereri* (Rivers *et al.* 2005) and *Myotis myotis* (Castella *et al.* 2001). Indeed a study investigating mitochondrial and nuclear population structure of *M. daubentonii* in Scotland found two distinct mitochondrial clades with considerable levels of nuclear gene flow within and between clades (Ngamprasertwong *et al.* 2008). However in this study using microsatellites did highlight Britain and Ireland as more distinct than when using

mitochondrial markers. A similar pattern was found in a previous Europe-wide study of *M. daubentonii* population structure using nuclear microsatellites (Atterby *et al.* 2009), and a study of *Rhinolophus ferrumequinum* (Flanders *et al.* 2009), both of which placed Britain in separate clusters to mainland Europe. It is possible that since *M. daubentonii* colonised Britain post-glacially, a founder effect has occurred and that subsequent genetic drift in isolation resulted in differentiation at nuclear markers that was not apparent at mtDNA.

Overall F_{ST}/Φ_{ST} was far greater for the *ND1* gene than for either X-linked or autosomal microsatellites. F_{ST} values may be different by virtue of the differences in effective population size between markers which would automatically act to increase mitochondrial differentiation relative to nuclear markers, due to the enhanced effects of genetic drift (Hedrick 2007). In addition, microsatellite markers are highly variable and so less suited to recording differentiation between populations if equilibrium between mutation and genetic drift has not been reached (Avise 2004, Birky *et al.* 1989). A further issue is microsatellite allele size homoplasy which would again lead to an inappropriate conclusion in terms of levels of population differentiation (Estoup *et al.* 2002). Finally because estimates of population differentiation are dependant on within population homozygosity, if loci are highly variable within populations (as is more likely to be the case for microsatellites than mitochondrial sequences) then F_{ST} estimates may be low even if alleles are not shared between subpopulations (Hedrick 2005). Indeed there are a number of instances in natural populations where differences in F_{ST} have been attributed to effective population size differences between markers and insufficient time to reach mutation/drift equilibrium for nuclear markers as a consequence of a large expanding population (e.g. Olive sea snake (Lukoschek *et al.* 2008), Marlin (Buonaccorsi *et al.* 2001)). However, even after taking these considerations into account, mitochondrial Φ_{ST} is still far greater than expected. This is likely to be a product of behavioural differences between the sexes, notably female philopatry and male dispersal. *M. daubentonii* roosts in predominantly single sex nursery colonies (discounting juveniles) and past genetic and field work has indicated that females are philopatric and that males have a tendency to disperse over greater distances (Angell 2008, Atterby *et al.* 2009, Ngamprasertwong *et al.* 2008, Senior *et al.* 2005). If this is the case it would explain why maternally inherited mitochondrial DNA exhibits substantially greater differentiation than markers which are transmitted by both males and females. Similar studies comparing mitochondrial and microsatellite markers

in temperate bat species have also found increased mitochondrial differentiation relative to autosomal markers (*R. ferrumequinum* (Rossiter *et al.* 2000), *M. bechsteinii* (Kerth *et al.* 2002), *Nyctalus noctula* (Petit & Mayer 1999)).

The AMOVA results also support restricted maternal gene flow at a regional scale. For both autosomal and X-linked microsatellites the vast majority of variation could be explained at the individual level, suggesting few restrictions to gene flow at a local or wider scale. However, although AMOVA attributed a large amount of mitochondrial variation to individual differences, a roughly equal amount of variation could be explained by differences between regions. The majority of individuals sampled in this study were male and were caught at autumn swarming sites. Swarming is a mating behaviour that occurs at cave sites where bats are present from many different summer roosts from a wide catchment area (Angell 2008, Kerth *et al.* 2003, Parsons & Jones 2003, Rivers *et al.* 2005, 2006). Consequently within-region differentiation for mitochondrial markers may be low because the bats sampled at a swarming site had their origins in a number of different maternal colonies.

It should be noted that because the majority of tissue sampling was conducted at swarming sites most of the samples included in mtDNA analyses were male. This can have important consequences when making interpretations regarding contemporary sex-biased gene flow. Because mtDNA is maternally inherited it is assumed that mtDNA population structure is reflective of female-specific processes. However if males are sampled which have dispersed away from the natal range then this could have the effect of altering mtDNA population structure such that the underlying pattern of structure due to female genetic drift and gene flow is blurred as a consequence of male movement in the current generation. For example Chen *et al.* (2008) examined mtDNA population structure for both sexes separately in the Formosan horseshoe bat and found greater population structure in females than males which they interpreted as more extreme female philopatry, with males remaining broadly faithful to maternal roosts over a larger geographic scale. Therefore whilst strong mtDNA structure was found, it is possible that if samples were female, different levels of population structure would have been demonstrated that would have altered conclusions made about levels of sex-biased dispersal.

Pair-wise comparisons of F_{ST} values between populations using the different marker types revealed a lack of any congruence between microsatellite and mitochondrial pair-wise differentiation. In other words, just because two populations appear highly differentiated using one marker type does not mean that this pattern will hold using a different set of markers. This could be because mitochondrial differentiation reflects historic colonisation patterns, whereas among population differentiation using microsatellites reflects contemporary gene flow.

When a population undergoes expansion it is expected that the ancestral population will harbour more diversity due to the random loss of variation in populations at the leading edge of the expansion (Hewitt 1996). Note, however, that when admixture occurs between previously isolated populations the opposite may be true (Petit *et al.* 2003). To test this hypothesis and distinguish between an east to west Vs south to north Expansion, measures of genetic diversity for each marker type were correlated against both latitude and longitude. A significant relationship was found between autosomal allelic richness and longitude and between X-microsatellite expected H_e and longitude, with populations in eastern areas being more diverse than western populations. A decline in genetic diversity from east to west has also been reported in Europe for *Myotis mystacinus* (Daniel Buckley, pers comm.) which suggests that eastern areas (potentially the Balkans, although lack of sampling in this region prevents any firm conclusions from being drawn) may have acted as a major refugium for *M. daubentonii*. However, as the decline in allelic richness is not uniform across Europe this suggests that other areas, such as Iberia, may also have acted as minor refugia. The Balkans have been suggested as a major refuge for a number of other taxa in Europe, including bats for example *R. ferrumequinum* (Rossiter *et al.* 2000), *Myotis myotis* (Ruedi & Castella 2003), *N. noctula* (Petit *et al.* 1999), European brown bear (Taberlet & Bouvet 1994), *Chorthippus* grasshoppers (Cooper *et al.* 1995) and beech trees (Demesure *et al.* 1996). However as *M. daubentonii* is found from western Europe through to eastern Asia the possibility of other eastern refugia cannot be discounted (Deffontaine *et al.* 2005, Stewart & Lister 2001), unfortunately a lack of sampling from areas outside western Europe prevents an exploration of this hypothesis.

Results from the 'Bottleneck' analysis of autosomal data also support an expansion as evidenced by the detection of heterozygote deficiency, although it must be noted that

this suggests a recent demographic expansion, rather than providing evidence for a post-glacial range expansion. There is evidence that many bat populations in Europe are declining, potentially as a consequence of the negative results of agricultural intensification on insect numbers (Wickramasinghe *et al.* 2004). However *M. daubentonii* populations in Britain may be increasing, possibly due to an increase in the number of artificial water bodies, which provides more optimal foraging habitat for *M. daubentonii* (Bat Conservation Trust, August 2002). One population in Lincolnshire (Withcall) showed signs of a recent bottleneck due to the detection of significant heterozygote excess. This population may be isolated since there are few suitable swarming sites nearby, consequently this could account for a bottleneck in this population. However, a note of caution about the interpretation of these results should be made. 'Bottleneck' assumes no admixture between populations, however clustering analyses provided strong evidence for admixture and a departure from the assumptions of the program may therefore lead to erroneous results.

Given these results for autosomal microsatellites, it is surprising that there was no significant relationship between diversity and latitude/longitude for mitochondrial markers. In the case of mitochondrial markers low number of samples in some populations may account for the lack of a detectable relationship.

5.4.3 Comparison of results between autosomal and X-linked microsatellite analyses

5.4.3.1 Population structure of *M. daubentonii* in Europe using autosomal and X-linked microsatellites

Bayesian clustering analysis using both marker types highlighted two clusters as representing the highest level of population structuring of *M. daubentonii* in Europe. Overall the results were largely congruent as evidenced by a highly significant correlation between assignment to cluster A or B between the two markers. However X-linked markers appeared to record higher levels of admixture within populations than autosomal markers. This may be the opposite to what we expect considering the overall higher levels of differentiation on the X-chromosome in terms of F_{ST} and increased linkage disequilibrium. At K=3, further discrepancies arise. Both autosomal and X-linked markers highlight England and mainland Europe as two separate clusters, but autosomal markers place Ireland and north Scotland as highly distinct, with Scandinavia

showing majority membership to this cluster. X-linked markers place Finland as highly distinct with Ireland, southern Scotland and Poland showing majority membership and Denmark with roughly equal membership to the Finnish and English clusters. In a similar study on global human populations, Ramachandran *et al.* (2004) also used clustering analyses and found that X-linked microsatellites provided similar but lower resolution of population structure than autosomal microsatellites. They attributed this to a reduction in power due to a lower number of X-linked loci used to resolve population structure in comparison to autosomal loci. When the number of loci used was equivalent, population structure between the two marker types displayed far greater levels of similarity. However it should be noted that in a study on the population genetic structure of drug metabolising response enzymes in humans, Wilson *et al.* (2001) were able to rule out a lack of resolution on the X-chromosome due to data paucity and instead invoked female dispersal to explain the lower levels of structure detectable using X-linked markers. In our study, we used 9 X-linked microsatellites and 14 autosomal microsatellites; consequently the increased resolution afforded using autosomal markers could be due to the increased power to detect discrete clusters due to the information from the additional markers. Further, more alleles were genotyped for autosomal microsatellites, which also enhances the power to resolve population structure, although not to the same extent as additional loci (Goudet *et al.* 2002).

Overall levels of differentiation (F_{ST}) were higher for X-linked microsatellites than autosomal microsatellites. There are several possible, but not mutually exclusive explanations for this. Firstly, differentiation may be expected to be higher on the X-chromosome by virtue of the smaller effective population size which, under neutral expectations, is 75% that of autosomes. When the effective population size is smaller genetic drift has a greater impact, consequently populations and individuals may be more differentiated due to the accelerated reduction of genetic diversity on the X-chromosome (Vicoso & Charlesworth 2006). Indeed in a number of cases in human populations it has been found that differences in F_{ST} values between autosomes and the X-chromosome can be accounted for by the smaller N_e of the X-chromosome (Ramachandran *et al.* 2004). However, even after accounting for the smaller effective population size of the X-chromosome by calculating N_m, levels of differentiation are still greater for X-linked microsatellites (Table 5.4). This suggests that factors other than the mode of X inheritance must also be considered to explain the elevated levels of differentiation for the X-chromosome markers.

The second factor that could act to increase levels of differentiation on the X-chromosome is selection. Because the X-chromosome is found in a hemizygous state when in males, it is possible that local adaptation may produce higher levels of differentiation on the X-chromosome (e.g. Baines & Harr 2007, Begun & Whitley 2000). Although markers chosen for this study were assumed to be selectively neutral, linkage on the X-chromosome is greater than on autosomes, as a consequence of recombination only occurring when the X-chromosome is in females. As a result markers used in this study could be linked to regions under selection, but this is unlikely to affect all loci.

The third explanation involves demographic differences between males and females. Because females have two copies of the X-chromosome but males only have one, we might expect that if female gene flow exceeded that of males (above a certain threshold) then the effective population size of X would tend towards that of autosomes (Ellegren 2009, Hammer *et al.* 2008). Conversely if male gene flow was the driving force then the effective population size of the X-chromosome may be reduced relative to 0.75. If there are unequal effective male and female population sizes as a consequence of variable mating success in one sex or if sex-biased dispersal (provided it results in gene flow) is present this may act to increase or decrease levels of differentiation relative to neutral expectations (Keinan *et al.* 2009). As the populations are more differentiated on the X-chromosome, this would suggest that female mating success is more variable than males, which seems unlikely as the mating system of *M. daubentonii* is promiscuous and paternity assignment has shown that some males achieve a higher chance of fathering offspring than others (Angell 2008, Senior *et al.* 2005). A more parsimonious explanation is that males disperse greater distances than females, acting to reduce X gene flow relative to autosomes over a wide geographic scale. Apart from males which reside with females in summer roosts, it is thought that most *M. daubentonii* males achieve successful matings at swarming sites (Angell 2008, Ngamprasertwong *et al.* 2008). Bats may travel to swarming sites from a large catchment area, which would mean that males disperse greater distances before mating (Parsons & Jones 2003). A similar pattern of increased X differentiation in global populations of humans was found by Keinan *et al.* (2009) in which long range male migration was invoked to explain the higher levels of X Vs autosomal differentiation.

5.4.3.2 Detection of migrants using autosomal Vs X-linked microsatellites

Clustering analyses using both autosomal and X-linked microsatellites highlighted populations in Britain as being reasonably distinct from populations in France on the basis of allele frequencies/linkage patterns. Logically it follows that it should be possible to detect recent migrants (or first generation offspring) from France to the UK or vice versa. Both marker types revealed a number of individuals as likely migrants, but a greater number of migrants were revealed using X-lined markers despite there being fewer X-loci. Because the phase of the X-chromosome marker genotypes in males is known, this means we can determine patterns of linkage disequilibrium (LD) which population assignment software can use as additional information (Piry *et al.* 2004). Thus, recorded LD on X-chromosomes potentially makes X-linked markers more useful for detecting non-resident individuals in populations. However, it should also be noted that such individuals are only possible migrants with an assigned probability and the true number of migrants is unknown, therefore although autosomal markers identified fewer potential migrants they may identify such migrants with greater accuracy. Furthermore although both marker types identified potential migrants, the individuals identified were not the same.

5.4.3.3 Demography of X and autosomal microsatellites: what can contrasting results tell us?

As previously mentioned under 'Neutral expectations' (see section 5.1), we expect the effective population size of the X-chromosome to be 75% that of autosomes. Because genetic drift is accelerated when N_e is smaller, we would also expect a corresponding reduction in genetic diversity at neutral markers on the X-chromosome (Schaffner 2004, Yu *et al.* 2002).

Using expected heterozygosity to calculate θ the X:autosomal ratio of θ ($4N_{eu}$ for autosomal microsatellites and $3N_{eu}$ for X-linked microsatellites, where u = mutation rate) for *M. daubentonii* across Europe was 0.81 and using the Migrate method it was 0.84 (see section 5.2.3.2). Given the F_{ST} results we may have expected the ratio to be below 0.75, which could explain the elevated levels of X-differentiation, so this result is slightly surprising and is highly indicative that demographic factors other than the mode

of inheritance of the chromosomes may be responsible for the increased F_{ST} values using X-linked markers. It should also be noted that we have made the assumption that u is similar for X-linked and autosomal microsatellites (u was set to constant for all loci in the Migrate analysis), if this is not the case it may result in deviations from θ, and should remain a consideration.

This leaves the problem of why the X:autosomal θ ratio deviates from expectation and tends towards unity. Again a number of possible factors could explain this and include 1) Background selection acting to increase diversity on the X-chromosome 2) Demographic processes affecting the entire population 3) Sex-biased demographic forces.

Because recessive, deleterious mutations are maintained at a lower frequency and removed from populations more quickly, neutral alleles on the X-chromosome are less likely to be linked to deleterious mutations (Charlesworth *et al.* 1987) and consequently theoretical and empirical studies predict that linked negative selection will act to increase X diversity relative to autosomal diversity (Begun & Whitely 2000, Lu & Wu 2005, Thornton & Long 2002). This could partially explain the elevated θ observed for X-linked markers in this study but, given the data available, it is difficult to test. However, this is a very general mechanism that should also apply to studies on human populations and in such cases background selection does not appear to exert a significant influence (Hammer *et al.* 2008).

Using data from a number of different taxa, Pool & Nielsen (2007) demonstrated that, following a bottleneck with subsequent population expansion, the X-chromosome recovers diversity more quickly than autosomes and that this effect could be quite long-lasting. This was also supported by a study on humans by Hammer *et al.* (2008) who found that an ancient bottleneck coupled with population growth could produce an X:autosome ratio as high as 0.85. In the case of *M. daubentonii* it is likely that populations underwent an ancient bottleneck when glaciations made much of Europe uninhabitable and subsequently grew as suitable habitat became available. This could, at least in part, explain why the X:autosomal ratio is higher than expected.

Sex-biased forces

Many of the data in this study, and previous studies by other researchers (Angell 2008, Ngamprasertwong *et al.* 2008, Senior *et al.* 2005), are strongly suggestive of a male bias in dispersal. If male dispersal was responsible for influencing the N_e of the X-chromosome then we might expect X-chromosome N_e to be reduced relative to 'neutral expectations' rather than increased. However, Laporte & Charlesworth (2002) found that sex-biased migration only weakly skews X:autosomal diversity unless populations are strongly subdivided. Considering the lack of strong genetic structure across Europe, this does not appear to be the case for *M. daubentonii*. Regardless of sex-biased dispersal (SBD), greater variance in male mating success could explain why the θ ratio is higher than anticipated. If some males achieve greater paternity than others but females are roughly equally likely to achieve maternity this means that male N_e decreases relative to females and so the N_e of the X-chromosome will increase (Cabalerro 1995, Charlesworth 2001). Work on paternity in *M. daubentonii* by Senior *et al.* (2005) and Angell (2008) in the UK has suggested that some males, which reside with females during the summer, may stand a greater chance of fathering offspring than males which reside in bachelor roosts. However, it is thought that the majority of mating events in *M. daubentonii* occur at swarming sites. Work by Glover & Altringham (2008) suggests that mating at swarming sites is a competitive event (as evidenced by a highly skewed male sex ratio and the observation of chasing behaviour), with males competing for access to females in a lek-like system. Consequently mating success between males is also likely to be skewed at swarming sites. Therefore it does appear that individual male paternity is variable and this could explain the higher than expected X:autosomal θ ratio. Higher variance in male mating success has also been suggested as an explanation to invoke a higher than expected X:autosomal N_e ratio in humans (Hammer *et al.* 2008).

5.4.4 Conclusion

The use of markers with different mutation rates and modes of inheritance in this study has allowed us to investigate processes operating at different spatial and time scales which may be responsible for the observed genetic structure of *M. daubentonii* populations in Europe. Mitochondrial DNA has revealed broad-scale phylogeographic structure which is likely a consequence of genetic divergence during glacial isolation. The use of microsatellite markers corroborated these results but also revealed

substantial levels of contemporary gene flow at a local and medium geographic scale. Results from both marker types suggest that there are few significant barriers to dispersal in this species, including the English Channel. Using markers with different sex-based modes of inheritance has allowed us to elucidate aspects of sex-specific demography in this species both in terms of dispersal (male biased) and mating success (greater variability in males). X-linked microsatellites increased our confidence in these conclusions as markers from multiple loci reduce the stochasticity associated with comparing single locus mitochondrial markers.

5.5 References

Angell RL (2008) Population ecology and mating system of Daubenton's bat (*Myotis daubentonii*) PhD Thesis, University of Leeds.

Atterby H, Aegerter JN, Smith GC *et al.* (2009) Population genetic structure of Daubenton's bat (*Myotis daubentonii*) in western Europe and the associated occurrence of rabies. *European Journal of Wildlife Research* **56** 67-81.

Avise JC (2004) *Molecular Markers Natural History and Evolution.* Sinauer Associates, Sunderland, Massachusetts.

Baines JF, Harr B (2007) Reduced X-linked diversity in derived populations of house mice. *Genetics* **175** 1911-1921.

Ballard JWO, Whitlock MC (2004) The incomplete history of the mitochondria. *Molecular Ecology* **13** 727-744.

Balloux F (2010) Mitochondrial phylogeography: The worm in the fruit of the mitochondrial DNA tree. *Heredity* **104** 419-420.

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11** 155-165.

Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16** 113–148.

Bedoya G, Montoya P, Garcia J *et al.* (2006) Admixture dynamics in Hispanics: A shift in the nuclear genetic ancestry of a South American population isolate. *PNAS* **103** 7234-7239.

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *PNAS* **8** . 4563-4568.

Begun DJ, Whitley P (2000) Reduced X-linked nucleotide polymorphism in *Drosophila simulans. PNAS* **97** 5960–5965.

Birky CW, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration, mutation, and drift — equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics* **121** 613–627.

Brito PH (2007) Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (*Strix aluco*). *Molecular Ecology* **16** 3423-3437.

Buonaccorsi VP, Starkey E, Graves JE (2001) Mitochondrial and nuclear DNA analysis of population subdivision among young-of-the-year Spanish mackerel (*Scomberomorus maculatus*) from the western Atlantic and Gulf of Mexico. *Marine Biology* **138** 37-45.

Burland TM, Worthington-Wilmer J (2001) Seeing in the dark: molecular approaches to the study of bat populations. *Biological Reviews* **76** 389-409.

Bustamante CD, Ramachandran S (2009) Evaluating signatures of sex specific processes in the human genome. *Nature Genetics* **41** 8-10.

Caballero A (1995) On the effective size of populations with separate sexes, with particular reference to sex linked genes. *Genetics* **139** 1007-1011.

Castella V, Ruedi M, Excoffier L (2001) Contrasted patterns of mitochondrial and nuclear structure among nursery colonies of the bat *Myotis myotis*. *Journal of Evolutionary Biology* **14** 708-720.

Charlesworth B, Coyne JA, Barton NH (1987) The relative rates of evolution of sex chromosomes and autosomes. *American Naturalist* **130** 113-146.

Charlesworth B (2001) The effect of life history and mode of inheritance on neutral genetic variability. *Genetic Research* **63** 213-227.

Chen S, Jones G, Rossiter SJ (2008) Sex-biased gene flow and colonisation in the Formosan lesser horseshoe bat: inference from nuclear and mitochondrial markers. *Journal of Zoology* **274** 207-215.

Cooper SJB, Ibrahim KM, Hewitt GM (1995) Postglacial expansion and genome subdivision in the European grasshopper *Chorthippus parallelus*. *Molecular Ecolology* **4** 49–60.

Corander J, Waldmann P, Marttinen P, Sillanpää MJ (2004). BAPS 2: enhanced possibilities for the analysis of genetic population structure. *Bioinformatics* **20** 2363-2369.

Cornuet JM, Luikart G (1997) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144** 2001-2014.

Crispo E, Hendry AP (2005) Does time since colonization influence isolation by distance: a meta-analysis. *Conservation Genetics* **6** 665-682.

Crow JF (2006) Age and sex effects on human mutation rates: an old problem with new complexities. *Journal of Radiation Research* Supp. B75–B82.

Deffontaine V, Libois R, Kotlik P *et al.* (2005) Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*). *Molecular Ecology* **14** 1727-1739.

Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution* **50** 2515-2520.

Ellegren H (2009) The different levels of genetic diversity in sex chromosomes and autosomes. *Trends in Genetics* **25** 278-284.

Estoup A, Jarne P, Cournuet JN (2002) Homoplasy and mutation model at microsatellite loci and their consequence for population genetic analysis. *Molecular Ecology* **11** 1591-1604.

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

Excoffier LG, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1** 47-50.

Flanders J, Jones G, Benda P *et al.* (2009) Phylogeography of the greater horseshoe bat, *Rhinolophus ferrunquinum*: Contrasting results from mitochondrial and microsatellite data. *Molecular Ecology* **18** 206-318.

Glover AM, Altringham JD (2008) Cave selection and use by swarming bat species. *Biological Conservation* **141** 1493-1504.

Goetting-Minesky MP, Makova LD (2006) Mammalian male mutation bias: impacts of generation time and regional variation in substitution rates. *Journal of Molecular Evolution* **63** 537–544.

Goodman SJ, Tamate HB, Wilson R *et al.*(2001) Bottlenecks, drift & differentiation: the population structure and demographic history of sika deer (*Cervus nippon*) in the Japanese archapeligo. *Molecular Ecology* **10** 1357-1370.

Goudet J (1996) FSTAT Version 1.2: a computer program to calculate F-statistics. *Journal of Heredity* **86** 485-486.

Goudet J, Perrin N, Waser P (2002) Tests for sex-biased dispersal using bi-parentally inherited genetic markers. *Molecular Ecology* **11** 1103-1114.

Haldane JBS (1947) The mutation rate of the gene for hemophilia, and its segregation in males and females. *Annals of Eugenics* **13** 262–271.

Hammer MF, Mendex FL, Cox MP, Woerner AE, Wall JD (2008) Sex biased evolutionary forces shape patterns of human genomic diversity. *PLoS Genetics* **4** 1-8.

Harpending HC, Batzer MA, Gurven M *et al.* (1998) Genetic traces of ancient demography. *PNAS* **17** 1961-1967.

Harris SL, Brookes SM, Jones G *et al.* (2006) European bat lyssaviruses: Distribution, prevalence and implications for conservation. *Biological Conservation* **131** 193-210.

Heckel G, Burri R, Fink S, Desmet JF, Excoffier L (2005) Genetic structure and colonization processes in European populations of the common vole, *Microtus arvalis*. *Evolution*, **59**, 2231–2242.

Hedrick PW (2005) A standardized genetic differentiation measure. *Evolution* **59** 1633-1638.

Hedrick PW (2007) Sex: Differences in mutation, recombination, selection, gene flow & genetic drift. *Evolution* **61** 2750-2771.

Hewitt GM (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linneaen Society* **58** 247–276.

Hewitt GM, Butlin RK (1997) Causes and consequences of population structure. In: *Behavioural Ecology: An Evolutionary Approach*, edited by J. R. Krebs and N. B. Davies. pp 203. Blackwell Science Ltd, Oxford.

Honnen AC, Hailer F, Kenntner N *et al.* (2010) Mitochondrial DNA and nuclear microsatellites reveal high diversity and genetic structure in an avian top predator, the white-tailed sea eagle, in central Europe. *Biological Journal of the Linneaen Society* **99** 727-737.

Huang Q-Y, Xu F-H, Shen H *et al.* (2002) Mutation patterns at dinucleotide microsatellite loci in humans. *American Journal of Human Genetics* **70** 625–634.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity* **77** 282-291.

Jan C (2010) PhD Thesis, University of Leeds.

Keinan A, Mullikin JC, Patterson N, Reich D (2009) Accelerated genetic drift on the X chromosome during human migration out of Africa. *Nature Reviews Genetics* **41** 66–70.

Keinan A, Reich D (2010) Can a sex-biased human demography account for the reduced effective population size of chromosome X in non-Africans? *Molecular Biology & Evolution* Advance online publication, May 2010, *doi: 10.1093.*

Kerth G, Mayer F, Petit E (2002) Extreme sex-biased dispersal in the communally breeding, nonmigratory Bechstein's bat (*Myotis bechsteinii*). *Molecular Ecology* **11** 1491-1498.

Kerth G, Kiefer A, Trappmann C, Weishaar M (2003) High gene diversity at swarming sites suggest hot spots for gene flow in the endangered Bechstein's bat. *Conservation Genetics* **4** 491-499.

Koskinen MT, Nilsson J, Je Veselov A *et al.* (2002) Microsatellite data resolve phylogeographic patterns in European grayling, *Thymallus thymallus, Salmonidae. Heredity* **88** 391-401.

Labuda D, Lefebvre JF, Nadeau P, Roy-Gagnon MH (2010) Female to male breeding ratio in modern humans – An analysis based on historical recombinations. *American Journal of Human Genetics* **86** 353-363.

Laporte V, Charlesworth B (2002) Effective population size and population subdivision in demographically structured populations. *Genetics* **162** 501-519.

Lawson-Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex biased dispersal. *Molecular Ecology* **16** 1559-1578.

Li W-H, Yi S, Makova K (2002) Male-driven evolution. *Current Opinion in Genetics & Development* **12** 650–656.

Lorde LB, Watkins WS, Bamshad MJ *et al.* (2000) The Distribution of Human Genetic Diversity: A Comparison of Mitochondrial, Autosomal, and Y-Chromosome Data. *American Journal of Human Genetics* **66** 979-988.

Lu J, Wu C-I (2005) Weak selection revealed by the whole genome comparison of the X-chromosome and autosomes of human and chimpanzee. *PNAS* **102** 4063-4067.

Lukoscheck V, Waycott M, Keogh JS (2008) Relative information content of polymorphic microsatellites and mitochondrial DNA for inferring dispersal and population genetic structure in the olive sea snake, *Aipysurus laevis*. *Molecular Ecology* **17** 3062-3077.

Macholan M, Munclinger P, Sugerkova M *et al.* (2007) Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution* **61** 746-771.

Maddision WP, Knowles LL (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* **55** 21-30.

Melo-Ferreira J, Alves PC, Freitas H, Ferrand N, Boursot P (2009) The genomic legacy from the extinct *Lepus timidus* to the three hare species of Iberia: contrast between mtDNA, sex chromosomes and autosomes. *Molecular Ecology* **18** 2643-2658.

Moore WS (1992) Inferring phylogenies from mtDNA variation: Mitochondrial gene-trees versus Nuclear-gene trees. *Evolution* **49** 718-726.

Ngamprasertwong T, Mackie IJ, Racey PA, Piertney SB (2008) Spatial distribution of mitochondrial and microsatellite DNA variation in Daubenton's bat within Scotland. *Molecular Ecology* **17** 3243-3258.

Nichols R (2001) Gene trees and species trees are not the same. *TREE* **16** 358-364.

Paetkau D, Slade R, Burden M & Estoup A (2004) Direct, real time estimation of migration using assignment methods: a simulation based exploration of accuracy and power. *Molecular Ecology* **13** 55-65.

Park SDE (2001), Trypanotolerance in West African Cattle and the Population Genetic Effects of Selection. PhD thesis, University of Dublin.

Parsons KN, Jones G (2003) Dispersion and habitat use by *Myotis daubentonii* and *Myotis nattereri* during the swarming season: implications for conservation. *Animal Conservation* **6** 283-290.

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

Pereira MJR, Salgueiro P, Rodrigues L *et al.* (2009) Population structure of a cave-dwelling bat, *Miniopterus schreibersii:* Does it reflect history and social organisation? *Journal of Heredity* **100** 533-544.

Pereira R, Gomes I, Amorim A, Gusmao L (2007) Genetic diversity of 10 X-chromosome STRs in Northern Portugal. *International Journal of Legal Medicine* **121** 192-197.

Petit E, Mayer F (1999) Male dispersal in the noctule bat (*Nyctalus noctula*): where are the limits? *Proceedings of the Royal Society of London, Series B* **266** 1717-1722.

Petit E, Excoffier L, Mayer F (1999) No evidence of bottleneck in the postglacial recolonization of Europe by the noctule bat (*Nyctalus noctula*). *Evolution* **53** 1247–1258.

Petit RJ, Itziar A, de Beaulieu JL *et al.* (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. *Science* **300** 1563-1565.

Piry S, Alapetite A, Cornuet, JM, Paetkau D, Baudouin L, Estoup A (2004) GeneClass2: A Software for Genetic Assignment and First-Generation Migrant Detection. *Journal of Heredity* **95** 536-539.

Pool JE, Nielsen, R. (2007) Population size changes re-shape patterns of genomic diversity. *Evolution* **61** 3001-3006.

Pope LC, Domingo-Rouara X, Burke T (2005) Isolation by distance and geneflow in the Eurasian badger (*Meles meles*) at both a local and broad scale. *Molecular Ecology* **15** 371-386.

Pritchard JK (2000) Inference of population structure using multi-locus genotype data. *Genetics* **155** 945-959.

Queney G, Ferrand N, Wise S *et al.* (2001) Stationary Distributions of Microsatellite Loci Between Divergent Population Groups of the European Rabbit (*Oryctolagus cuniculus*). *Molecular Biology & Evolution* **18** 2169-2178.

Ramachandran S, Rosenberg NA, Zhivotovsky LA, Feldman MW (2004) Robustness of the inference of human population structure: A comparison of X-chromosomal and autosomal microsatellites. *Human Genomics* **1** 87-97.

Rivers NM, Butlin RK, Altringham JD (2005) Genetic population structure of Natterer's bats explained by mating at swarming sites and philopatry. *Molecular Ecology* **14** 4299-4312.

Rivers NM, Butlin RK, Altringham JD (2006) Autumn swarming behaviour of Natterer's bats in the UK: population size, catchment area and dispersal. *Biological Conservation* 127 215-226.

Rossiter SJ, Jones G, Ransome RD, Barratt EM (2000) Genetic variation and

population structure in the endangered greater horseshoe bat (*Rhinolophus ferrumequinum*). *Molecular Ecology* **9** 1131-1135.

Rousset FR (1997) Genetic differentiation and estimates of gene flow from F-statistics under isolation by distance. *Genetics* **145** 1219-1228.

Rousset FR (2000) Genetic differentiation between individuals. *Journal of Evolutionary Biology* **13** 58–62.

Rubinoff D, Holland BH (2005) Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology* **54** 952–961.

Ruedi M, Castella V (2003). Genetic consequences of the ice ages on nurseries of the bat *Myotis myotis*: a mitochondrial and nuclear survey. *Molecular Ecology* **12** 1527–1540.

Schaffner SF (2004) The X-chromosome in population genetics. *Nature Reviews Genetics* **5** 43-51.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9** 615-629.

Senior P, Butlin RK, Altringham JD (2005) Sex and segregation in temperate bats. *Proceedings of the Royal Society of London, Series B* **272** 2467-2473.

Shikano T, Shimada Y, Herczeg G, Merila J (2010) History vs. habitat type: explaining the genetic structure of European nine-spined stickleback (*Pungitius pungitius*) populations. *Molecular Ecology* **19** 1147-1161.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47** 264–279.

Stewart JR, Lister AM (2001) Cryptic Northern refugia and the origins of the modern biota. *TREE* **16** 608-613.

Taberlet P, Bouvet J (1994) Mitochondrial DNA polymorphism, phylogeography, and conservation genetics of the brown bear *Ursus arctos* in Europe. *Proceedings of the Royal Society B: Biological Sciences*, **255**, 195–200.

Thornton K, Long M (2002) Rapid divergence of gene duplicates on the *Drosophila melanogaster* X-chromosome. *Molecular Biology & Evolution* **19** 918-925.

Vicoso B, Charlesworth B (2006) Evolution on the X-chromosome: Unusual patterns and processes. *Nature Reviews Genetics* **7** 645-653.

Wickramasinghe L, Harris S, Jones G, Vaughan-Jennings N (2004) Abundance and species richness of nocturnal insects on organic and conventional farms: effects of agricultural intensification on bat foraging. *Conservation Biology* **18** 1283-1292.

Wilkins JF, Marlowe FW (2006) Sex-biased migration in humans: what should we expect from genetic data? *BioEssays* **28** 280-300.

Wilson JF, Weale ME, Smith AC *et al.* (2001) Population genetic structure of variable drug response. *Nature Genetics* **29** 265-269.

Yannic G, Basset P, Hausser J (2008) Phylogeography and recolonization of the Swiss Alps by the Valais shrew (*Sorex antinorii*), inferred with autosomal and sex-specific markers. *Molecular Ecology* **17** 4418-4433.

Yu N, Chen FC, Ota S *et al.* (2002) Larger genetic difference within Africans than between Africans and Eurasians. *Genetics* **161** 269-274.

Chapter 6: Final Discussion

6.1 Summary of findings

This thesis aimed to elucidate contemporary and historic processes responsible for observed levels and patterns of genetic variation in *Myotis daubentonii* populations across Europe. This was done using a variety of marker systems with different mutation rates, patterns of recombination and modes of inheritance. It is increasingly recognised that using a single marker type to investigate population genetic structure can lead to erroneous results (Brito *et al.* 2007, Flanders *et al.* 2009), whether it is due to the stochasticity associated with using a single locus (Avise 2004, Balloux 2010) or using markers with mutational properties that are inappropriate to the time and spatial scale of the question being addressed (Selkoe & Toonen 2006). Using a multi-marker approach allowed the exploration of phylogeographic processes, levels of contemporary gene flow, barriers to gene flow, patterns of sex-biased differentiation and an insight into variable mating success between males and females.

We found that, on a phylogeographic scale, there were two mitochondrial clades present which are likely to represent the expansion of *M. daubentonii* from at least two glacial refugia, with subsequent contact in intermediate areas. Microsatellite markers with higher mutation rates were able to corroborate these findings on a broad geographic scale. Post-glacial expansion into western Europe from the south and the east is a common pattern reported in bats and other taxa (Flanders *et al.* 2009, Hewitt 1999, Petit *et al.* 1999, Ruedi & Castella 2003, Ruedi *et al.* 2008). As bats are dependent on mesic habitats they would have been able to colonise Europe as climatic and habitat conditions became suitable (Altringham 1996).

In terms of contemporary gene flow, there appeared to be little genetic differentiation even over a wide geographic scale. A pattern of isolation by distance at all marker types suggested that *M. daubentonii* was limited by the maximal distance it can disperse rather than by any other barriers. In addition there seemed to be no barrier effect of the English Channel to either contemporary gene flow or historic colonisation. Because bats are capable of powered flight, lack of population structure over wide geographic areas is often reported (Burland & Worthington-Wilmer 2001). However, this does vary among species, with migratory bats exhibiting far lower levels of population structure

(Angell 2008, Petit & Mayer 1999, Rivers *et al.* 2005) than sedentary species (Kerth & Petit 2005, Rossiter *et al.* 2006). *M. daubentonii* is a swarming species and individuals found at swarming sites have been shown to arrive from a wide catchment area (Angell 2008, Parsons & Jones 2003). As mating occurs at swarming sites this acts to reduce genetic structure at a regional scale. Bodies of water appear to act as differential barriers to movement in bat species. For example strong genetic discontinuities have been reported between populations of some species on either side of the Straits of Gibraltar but not others (Garcia-Mudarra *et al.* 2009). Other bat species do appear capable of crossing the English Channel such as *Pipistrellus pipistrellus* (Racey *et al.* 2007), but it proves a barrier to species with lower dispersal capabilities such as *Rhinolophus ferrumequinum* (Rossiter *et al.* 2006).

Using markers with different sex-based modes of inheritance gave an insight into sex-biased processes. Far greater levels of differentiation in mitochondrial DNA than in microsatellites and a smaller N_m value suggested that females were not dispersing as far as males. Similar discrepancies in genetic structure between mitochondrial and nuclear markers have also been reported in *M. daubentonii* populations in Scotland (Ngamprasertwong *et al.* 2008) and in other temperate bat species (Castella *et al.* 2001, Kerth *et al.* 2002, Petit *et al.* 2001). However the magnitude of the discrepancy in some other bat species is far higher, suggesting stricter female philopatry (Kerth *et al.* 2002). Our results suggest that although dispersal was male biased in *M. daubentonii*, female philopatry was not strict, and whilst individuals of both sexes probably exhibit a degree of dispersal behaviour males disperse over longer distances.

The additional information provided by comparing X-linked and autosomal differentiation supported this finding and the use of X-linked and autosomal markers also allowed us to determine that male mating success was probably more variable than that of females in Daubenton's bat. This is the first study that I know of which has used X-chromosome markers to investigate population structure in a bat. However, similar work on humans that compared autosomal and X-linked markers has also been able to demonstrate differences in diversity and structure between the marker types that may be a consequence of sex-biased processes (Bustamante & Ramachandran 2009, Hammer *et al.* 2008, Keinan *et al.* 2009).

The main evolutionary mechanisms invoked to explain SBD are inbreeding avoidance, local resource competition, kin-interactions or a combination of these factors (Lawson-Handley & Perrin 2007). It has been suggested that the mating system of the species in question is central to dictating which of these factors is dominant in driving SBD (Greenwood 1980), as the mating system will affect which resources are limiting to each sex, the sex likely to incur the highest inbreeding load and whether either sex is likely to benefit from the formation of social bonds. Most bats appear to lie on a continuum between polygyny and promiscuity, where in either case mating success is probably more variable for males than it is for females (McCracken & Wilkinson 2000). Past work investigating paternity in *M. daubentonii* has suggested that males roosting with females during the summer stand a higher chance of fathering offspring than males in single sex roosts (Angell 2008, Senior *et al.* 2005). However this is probably a relatively unusual mating behaviour in bats and there were a significant number of offspring whose parentage could not be ascribed to summer males (Angell 2008). Instead it is likely that the majority of mating events in Daubenton's occur at swarming sites, however male mating success at swarming sites is also likely to variable due to male-male competition for mates (Glover & Altringham 2008, Senior *et al.* 2005). This study found that the θ (N_{eu}) ratio between X-linked and autosomal markers was higher than that expected under neutrality. Although there are a few potential reasons for this, one explanation is that male mating success was more variable than females. This would increase X-chromosome gene flow relative to autosomes, acting to increase the N_e of the X-chromosome (Ellegren 2009, Hedrick 2007). If males do not stand an equal chance of fathering offspring, such that the number of females becomes limiting, it may benefit them to disperse to swarming sites where they stand a higher chance of mating than if they stayed close to the summer roosts. Swarming would also play a role in inbreeding avoidance as bats from many summer colonies have been documented at a particular swarming site (Angell 2008, Glover & Altringham 2008, Parsons & Jones 2003, Veith *et al.* 2004) which would allow gene flow between bats from a wide catchment area. Thus it is probably a combination of inbreeding avoidance and male-male competition for resources (Russo 2002, Senior *et al.* 2005) that drives dispersal in *M. daubentonii*.

Understanding population structure and dispersal patterns in disease vectors is of vital importance in modelling disease epidemiology and predicting disease spread. Spatial epidemiological models recognise the importance of including spatially explicit information, and accurate measures of spatial parameters such as dispersal are paramount for the realistic construction of such models (Deal *et al.* 2000, Jin & Wang 2005).

M. daubentonii is one of two primary vectors of the rabies like EBLV-2 in Europe and the sole vector of the disease in Britain (Harris *et al.* 2006, Johnson *et al.* 2006). Current understanding of EBLV-2 epidemiology is limited. There is no clear picture of how the virus is transmitted between bats (Harris *et al.* 2006, Johnson *et al.* 2008), bats may carry the virus with no symptoms of ill health (Sims *et al.* 1963) and the distribution of outbreaks is patchy (Amengual *et al.* 1997, Harris *et al.* 2006).

This study found low levels of population structure among *M. daubentonii* populations in Britain and Europe, with an overall pattern of isolation by distance suggesting that *M. daubentonii* has the potential to disperse tens of kilometres. Furthermore our genetic data suggested that males may disperse greater distances than females, which may render males more instrumental in the spread of EBLV-2. Additional support for long distance male movement is provided by social network analyses and ringing work by Rigby in the Yorkshire Dales (thesis 2010), which demonstrated that male bats are more likely than females to be involved in movements between summer roosts and movements from summer roosts to swarming sites. Long distance dispersal (LDD) events are especially important in modelling disease spread as, whilst short dispersal movements affect the spread of a disease front, long distance movements have the potential to create outbreaks ahead of the disease front (Filipe & Maule 2004, White *et al.* 1995). For example, migration events in birds have been demonstrated as accounting for the spread of diseases into non-endemic areas, such as Lyme disease (Smith *et al.* 1996) and West Nile Virus in the United States (Reed *et al.* 2003). Long distance movements have also proved a critical factor in the spread of bovine tuberculosis both in terms of direct movement of cattle (Gilbert *et al.* 2005) and movements of the wildlife reservoir the European badger (Pope *et al.* 2007). A study on the spread of a prion disease among populations of mule deer also found that it was long

distance movements (migration) rather than shorter dispersal events which were the most likely mechanism for disease spread (Conner & Miller 2008).

There is no evidence for sex-specific host association of rabies related diseases, however if one sex is more likely to be involved in LDD events then this sex will be instrumental in spreading the disease. For example in ticks (the vectors of Lyme disease) males are more likely to associate with birds than females (De Meeus *et al.* 2002). As birds have the capacity to move long distances, male ticks attached to birds will be important propagators of Lyme disease over a wide geographic scale. A male bias in dispersal has also been demonstrated in the raccoon, which is a vector of rabies in North America, and this has had subsequent implications for management of rabies in racoon populations (Cullingham *et al.* 2008).

There is evidence that bats can survive rabies infection for extended periods and that apparently healthy bats can exist in an asymptomatic carrier state (Johnson *et al.* 2008, O' Shea *et al.* 2003). Combined with the low levels of population structure in *M. daubentonii* (indicating widespread dispersal) this suggests that there is strong potential for EBLV-2 spread among Daubenton's colonies. However it should also be noted that dispersal, especially over long distances, places energetic demands on an organism (Sulkin *et al.* 1959). Thus apparently healthy carriers may become immunosuppressed and hence more likely to succumb to the virus (Harris *et al.* 2006). This has been demonstrated in birds where long distance movements are hampered in individuals carrying bird flu (Weber & Stilianakis 2007). Consequently infected bats dispersing over longer distances may incur EBLV-2 related mortality before any potential spill-over infection to other bats can occur. Indeed it has been demonstrated that environmental stress factors may be responsible for inducing active rabies presence in previously dormant carriers in some bat species (Sulkin *et al.* 1959).

Despite the high potential for *M. daubentonii* to transmit rabies among colonies, cases of EBLV-2 infection in Europe have only been reported in isolated clusters, suggesting that the mechanisms of transmission are complex and/or virulence is low. However, the presence of antibodies in *M. daubentonii* colonies in Britain is estimated at 1-16% (Brookes *et al.* 2005, Harris *et al.* 2009) suggesting that previous exposure to the virus is high. Currently knowledge of how EBLV-2 is transmitted both among bats and to other organisms is limited (Harris *et al.* 2006). A study which aimed to artificially infect

bats with EBLV-2 was only able to do so via direct intracranial inoculation, which suggests that resistance to infection is high (Johnson *et al.* 2008). When the virus is active it is detectable in the saliva of bats and bats are more likely to become aggressive suggesting that spread may occur orally. The two cases of spill-over infection to humans were demonstrated to have occurred as the consequence of a bite from a Daubenton's bat infected with EBLV-2 (Racey & Fooks 2005).

Widespread post-glacial expansion and the lack of any demonstrable contemporary barriers to movement suggest that there are few physical barriers to dispersal/colonisation in *M. daubentonii*. The English Channel does not appear to pose a barrier to gene flow and this conclusion is to that of Atterby *et al.* (2009) who found very little differentiation between *M. daubentonii* populations in Britain and populations in mainland Europe. However, isolation by distance across Europe suggests that *M. daubentonii* populations are not panmictic over a large geographic area and that intrinsic dispersal capabilities limit gene flow beyond a certain distance. In the UK active EBLV-2 infection has been reported in a number of populations over a wide geographic range from Scotland to the South of England (Banyard *et al.* 2009, Harris *et al.* 2006, Ngamprasertwong *et al.* 2008). Furthermore, passive and active surveillance indicates that EBLV-2 infection is endemic at low levels in Britain, rather than being a sporadic occurrence due to the movement of infected individuals from the continent. There have been no reported instances of EBLV-2 infection in France or other countries which border the English Channel. It should be noted that surveillance for rabies in mainland Europe is often focussed on other organisms, such as foxes, that represent a greater public health risk (Cliquet *et al.* 2009), and so the possibility that surveillance for EBLV-2 infection in bats in some countries is not extensive enough to detect infected populations should not be discounted. Therefore the movement of infected bats between Britain and the continent, resulting in the spread of EBLV-2 to previously uninfected areas, remains a possibility.

In summary, widespread gene flow at a regional scale and lack of barriers to gene flow suggest that *M. daubentonii* has the capacity for widespread transmission of EBLV-2 among colonies. Because males appear to disperse over longer distances they may be more instrumental in the transmission of EBLV-2 than females. EBLV-2 is probably endemic in Britain, with individuals from the UK having the potential to spread the disease to other countries in mainland Europe.

In terms of improving management of EBLV-2 the additional information about dispersal and population structure can give greater realism to epidemiological models that aim to predict the spread of disease. For example, by using markers with different sex-based modes of inheritance I was able to demonstrate that female gene flow (effective dispersal) occurred over a distance roughly one third the effective dispersal of both sexes combined (11-25 km for mtDNA Vs 36-81 km for autosomal microsatellites). Current methods to control EBLV-2 in Europe rely on excluding bats from buildings (Frantz & Trimachi 1983), informing the public and vaccination of people likely to come into close contact with bats (Racey & Fooks 2005). Because *M. daubentonii* is found in relatively large numbers and has the potential to disperse tens of kilometres, containing infections to isolated areas when outbreaks occur is impractical. In the UK and the rest of Europe bats are protected species so culling of infected colonies is also not an option. However, recent work by Rigby (2010) demonstrated that in summer colonies there are high intra-roost associations between individuals and low inter-colony associations, with the exception of some 'cut-point' male bats that move between roosts. These cut-point male bats represent a likely source of disease transmission between colonies and if such individuals can be identified, testing these bats for EBLV-2 infection might be a more practical and cost-effective way of limiting disease transmission between roosts than the monitoring of whole colonies. Active surveillance is a positive move (Harris *et al.* 2009) and should be directed to areas in which there is no prior reason to believe EBLV-2 is present as well as areas in which there is. Although differentiation between populations in Britain and mainland Europe is low, it is sufficient to allow us to potentially determine the origin of a bat given genetic data. Therefore we could use genetic assignment tests to determine the likely origin of an infected individual, which would allow us to determine the directionality of infections and how EBLV-2 may spread to previously uninfected areas. For example, using the autosomal (Jan *et al.* in prep) and X-linked microsatellites described in this thesis I was able to indentify individuals on either side of the English Channel that potentially did not originate from their sampled population.

6.4 Future work

EBLV-2 infections occur in isolated clusters and are apparently absent from some European countries. It is clear that more research on the mechanisms of transmission is

needed to understand how EBLV-2 may spread among colonies, given the high levels of gene flow. This would allow us to assess the potential health risk that EBLV-2 presents to the public, and would also allow more informed management decisions to be made.

For this thesis bats were collected from a wide distribution in Europe, however additional samples from populations in eastern Europe would prove more illuminating on the origins and relationships of *M. daubentonii* populations. For example this work suggested an eastern and a southern European refugium for *M. daubentonii*, but due to lack of samples in Balkan regions, eastern Europe and Germany it was only possible to speculate on the exact location of such refugia. Sampling from England (especially Yorkshire) was more intensive than in any other country. Although there were practical reasons for this, more intense sampling of populations in continental Europe would allow us to determine patterns of geneflow and identify barriers to dispersal more accurately.

This is the first study that has used X-chromosome markers to investigate population structure in bats. Evidence from studies in humans has suggested that, provided large numbers of loci are used, an X:autosome comparison may be more powerful than a mitochondrial:Y-chromosome comparison to investigate sex-biased differences in population structure (Wilkins & Marlowe 2006). Although this study was able to demonstrate differences in X and autosomal structure that were attributed to sex-biased dispersal, we only used nine X-linked loci. Increasing the number of X-linked loci would allow a more powerful comparison to be made and would allow us to be more confident in our conclusions.

Although some female samples were used the majority of individuals in this study were males. It would be interesting to have more samples from females so that differences in genetic structure could be investigated on a sex by sex rather than marker by marker basis.

Although attempts were made to identify and utilise polymorphic Y-chromosome markers in *M. daubentonii* we were unable to do so. The use of Y-chromosome markers would allow far more insight into male population genetic processes and sex-biased dispersal (Petit *et al.* 2002). It would also allow us to look at phylogeographic

relationships of patrilines in *M. daubentonii*. The increasing use of genomics and next generation sequencing methods should make it easier to identify genetic regions which can be mapped to the Y-chromosome in *M. daubentonii*. Thus there is future potential for the development and application of Y-linked markers in this species (Greminger *et al.* 2010).

Finally, whilst genetic data gives us insight into dispersal that would be difficult to obtain using field methods, direct field measures of dispersal should also be used to corroborate such findings. For example the use of ringing in male and female *M. daubentonii* individuals by Angell (2008) and Rigby (2010) demonstrated that males have a tendency to move greater distance than females, providing further confirmation of findings using genetic markers in this thesis. The use of geographic-information systems (GIS) will also be informative in identifying landscape areas and features which either promote or inhibit dispersal in Daubenton's (Manel *et al.* 2003). A multi-faceted approach to investigating dispersal will be the most accurate way of determining movement patterns in *M. daubentonii* that can inform out understanding of EBLV-2 incidence and spread.

6.5 References

Altringham JD (1996) *Bats: Biology & Behaviour.* Oxford University Press, Oxford.

Amengual B, Whitby JE, King A, Cobbo J, Bourhy H (1997) Evolution of European bat lyssaviruses. *Journal of General Virology* **78** 2319-2328.

Angell RL (2008) Population ecology and mating system of Daubenton's bat (*Myotis daubentonii*) PhD Thesis, University of Leeds.

Atterby H, Aegerter JN, Smith GC *et al.* (2009) Population genetic structure of Daubenton's bat (*Myotis daubentonii*) in western Europe and the associated occurrence of rabies. *European Journal of Wildlife Research* **56** 67-81.

Avise JC (2004) *Molecular Markers Natural History and Evolution.* Sinauer Associates, Sunderland, Massachusetts.

Balloux F (2010) Mitochondrial phylogeography: The worm in the fruit of the mitochondrial DNA tree. *Heredity* **104** 419-420.

Banyard AC, Johnson N, Voller K *et al.* (2009) Repeated detection of European bat lyssavirus type 2 in dead bats found at a single roost site in the UK. *Archives of Virology* **154** 1847-1850.

Brito PH (2007) Contrasting patterns of mitochondrial and microsatellite genetic structure among Western European populations of tawny owls (*Strix aluco*). *Molecular Ecology* **16** 3423-3437.

Brookes SM, Aegerter JN, Smith GC *et al.* (2005) European bat lyssavirus in Scottish bats. *Emerging Infectious Diseases* **11** 572-578.

Burland TM, Worthington Wilmer J (2001) Seeing in the dark: molecular approaches to the study of bat populations. *Biological Reviews* **76** 389-409.

Bustamante CD, Ramachandran S (2009) Evaluating signatures of sex specific processes in the human genome. *Nature Genetics* **41** 8-10.

Castella V, Ruedi M, Excoffier L (2001) Contrasted patterns of mitochondrial and nuclear structure among nursery colonies of the bat *Myotis myotis*. *Journal of Evolutionary Biology* **14** 708-720.

Cliquet F, Picard-Meyer E, Barrat J *et al.* (2009) Experimental infection of foxes with European bat lyssaviruses type-1 and type-2. *BMC Veterinary Research* **5** 19.

Conner MM, Miller MW (2008) Movement patterns and spatial epidemiology of a prion disease in mule deer population units. *Ecological Applications* **14** 1870-1881.

Cullingham CI, Pond BA, Kyle CJ *et al.* (2008) Combining direct and indirect genetic methods to estimate dispersal for informing wildlife disease management decisions. *Molecular Ecology* **17** 4874-4886.

Deal B, Farello C, Lancaster M, Kompare T Hannon B (2000) A dynamic model of the spatial spread of an infectious disease: the case of fox rabies in Illinois. *Environmental Modelling & Assessment* **5** 47-62.

De Meeus T, Beati L, Delaye C *et al.* (2002) Sex biased genetic structure in the vector of lyme disease, *Ixodes ricinus*. *Evolution* **56** 1802-1807.

Ellegren H (2009) The different levels of genetic diversity in sex chromosomes and autosomes. *Trends in Genetics* **25** 278-284.

Filipe JAN, Maule MM (2004) Effects of dispersal mechanisms on spatio-temporal development of epidemics. *Journal of Theoretical Biology* **226** 125-141.

Flanders J, Jones G, Benda P *et al.* (2009) Phylogeography of the greater horseshoe bat, *Rhinolophus ferrunquinum*: Contrasting results from mitochondrial and microsatellite data. *Molecular Ecology* **18** 206-318.

Frantz SC, Trimarchi CV (1983) Bats in human dwellings: health concerns and management. In: Decker, D.F. (Ed.), *Proceedings of the First Eastern Wildlife Damage Control Conference.* Cornell University, Ithaca, NY, pp. 299–308.

Garcia-Mudarra JL, Ibanez C, Juste J (2009) The Straits of Gibraltar: barrier or bridge to Ibero-Moroccan bat diversity? *Biological Journal of the Linnean Society* **96** 434-450.

Gilbert M, Mitchell A, Bourn D *et al.* (2005) Cattle movements and bovine tuberculosis in Great Britain. *Nature* **435** 491-496.

Glover AM, Altringham JD (2008) Cave selection and use by swarming bat species. *Biological Conservation* **141** 1493-1504.

Greenwood PJ (1980) Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour* **28** 1140-1162.

Greminger M, Krutzen M, Schelling C, Pienkowsk A, Wandeler P (2010) The quest for Y-chromosomal markers – methodological strategies for mammalian non-model organisms. *Molecular Ecology Resources* **10** 409-420.

Hammer MF, Mendex FL, Cox MP, Woerner AE, Wall JD (2008) Sex biased evolutionary forces shape patterns of human genomic diversity. *PLoS Genetics* **4** 1-8.

Harris SL, Brookes SM, Jones G *et al.* (2006) European bat lyssaviruses: Distribution, prevalence and implications for conservation. *Biological Conservation* **131** 193-210.

Harris SL, Aegerter JN, Brooks SM, McElhinney LM, Jones G, Smith GC & Fooks AR (2009) Targeted surveillance for European bat lyssavirus in English bats (2003-06). *Journal of Wildlife Diseases* **45** 1030-1041.

Hedrick PW (2007) Sex: Difference in mutation, recombination, selection, gene flow & genetic drift. *Evolution* **61** 2750-2771.

Hewitt GM (1999) Post-glacial recolonization of European biota. *Biological Journal of the Linnean Society* **68** 87-112.

Jan C (2010) PhD Thesis, University of Leeds.

Jin Y, Wang WD (2005) The effect of population dispersal on the spread of a disease. *Journal of Mathematical Analysis and Applications* **308** 343–364.

Johnson N, Vos A, Neubert L *et al.* (2008) Experimental study of European bat lyssavirus type 2 infection in Daubenton's bat (*Myotis daubentonii*). *Journal of General Virology* **89** 2662-2672.

Keinan A, Mullikin JC, Patterson N, Reich D (2009) Accelerated genetic drift on the X chromosome during human migration out of Africa. *Nature Reviews Genetics* **41** 66–70.

Kerth G, Mayer F, Koenig B (2002) Mitochondrial DNA (mtDNA) reveals that female Bechstein's bats live in closed societies. *Molecular Ecology* **9** 793-800.

Kerth G, Petit E (2005) Colonization and dispersal in a social species, the Bechsteins bat (*Myotis bechsteinii*). *Molecular Ecology* **14** 3943-3950.

Lawson-Handley LJ, Perrin N (2007) Advances in our understanding of mammalian sex biased dispersal. *Molecular Ecology* 16 1559-1578.

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

McCracken GF, Wilkinson GS (2000) Bat mating systems. In: *Reproductive Biology of Bats*, edited by E. G. Crichton and P. H. Krutzsch. pp 321-362. Academic Press, London and San Diego.

Ngamprasertwong T, Mackie IJ, Racey PA, Piertney SB (2008) Spatial distribution of mitochondrial and microsatellite DNA variation in Daubenton's bat within Scotland. *Molecular Ecology* **17** 3243-3258.

O'Shea TJ, Shankar V, Bowen RA, Rupprecht CE, Wimsatt JH (2003) Do bats acquire immunity to rabies? Evidence from the field. *Bat Research News* **44** 161.

Parsons KN, Jones G (2003) Dispersion and habitat use by *Myotis daubentonii* and *Myotis nattereri* during the swarming season: implications for conservation. *Animal Conservation* **6** 283-290.

Petit E, Mayer F (1999) Male dispersal in the noctule bat (*Nyctalus noctula*): where are the limits? *Proceedings of the Royal Society of London, Series B* **266** 1717-1722.

Petit E, Excoffier L, Mayer F (1999) No evidence of bottleneck in the postglacial recolonization of Europe by the noctule bat (*Nyctalus noctula*). *Evolution* **53** 1247–1258.

Petit E, Balloux F, Goudet J (2001) Sex-biased dispersal in a migratory bat: A characterization using sex-specific demographic parameters. *Molecular Ecology* **55** 635-640.

Petit E, Balloux F, Excoffier L (2002) Mammalian population genetics: why not Y? *Trends in Ecology & Evolution* **17** 28-33.

Pope LC, Butlin RK, Wilson GJ *et al.* (2007) Genetic evidence that culling increases badger movement: implications for the spread of bovine tuberculosis. *Molecular Ecology* **16** 4919-4929.

Racey PA, Fooks A (2005) Rabies in bats in Britain. *Biologist* **52** 34-38.

Racey PA, Barratt EM, Burland TM *et al.* (2007) Microsatellite DNA polymorphism confirms reproductive isolation and reveals differences in population genetic structure of cryptic Pipistrelle bat species. *Biological Journal of the Linnean Society* **90** 539–550.

Reed KD, Meece JK, Henkel JS, Shukla SK (2003) Birds, Migration and Emerging Zoonoses: West Nile Virus, Lyme Disease, Influenza A and Enteropathogens. *Clinical Medicine & Research* **1** 5-12.

Rigby EL (2010) Population ecology of Daubenton's bat. *PhD thesis*, University of Leeds.

Rivers NM, Butlin RK, Altringham JD (2005) Genetic population structure of Natterer's bats explained by mating at swarming sites and philopatry. *Molecular Ecology* **14** 4299-4312.

Rossiter SJ, Jones G, Ransome RD, Barratt EM (2000) Genetic variation and population structure in the endangered greater horseshoe bat (*Rhinolophus ferrumequinum*). *Molecular Ecology* **9** 1131-1135.

Rossiter SJ, Jones G, Ransome RD, Barratt EM (2006) Causes and consequences of genetic structure in the greater horseshoe bat (*Rhinolophus ferrumequinum*). Pp. 213-226. In: *Functional and Evolutionary Ecology of Bats* (A. Zubaid, G. F. McCracken and T. H. Kunz, eds). Oxford University Press, New York

Ruedi M, Castella V (2003). Genetic consequences of the ice ages on nurseries of the bat *Myotis myotis*: a mitochondrial and nuclear survey. *Molecular Ecology* **12** 1527–1540.

Ruedi M, Walter S, Fischer MC, Scaravelli D, Excoffier L, Heckel D (2008) Italy as a major ice age refuge area for the bat *Myotis myotis* in Europe. *Molecular Ecology* **17** 1801-1814.

Russo D (2002) Elevation affects the distribution of the two sexes in Daubenton's bats *Myotis daubentonii* (Chiroptera: Vespertilionidae) from Italy. *Mammalia* **66** 543-551.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters* **9** 615-629.

Senior P, Butlin RK, Altringham JD (2005) Sex and segregation in temperate bats. *Proceedings of the Royal Society of London, Series B* **272** 2467-2473.

Sims RA, Allen R, Sulkin SE (1963) Studies on the pathogenesis of rabies in insectivorous bats. *Journal of Infectious Diseases* **112** 17-27.

Smith RP, Rand PW, Lacombe EH, Morris SR, Holmes DW, Caporale DA (1996) Role of bird migration in the long distance dispersal of *Ixodes dammini* the vector of Lyme disease. *The Journal of Infectious Diseases* **174** 221-224.

Sulkin SE, Krutzsch PH, Allen R, Wallis C (1959) Studies on the pathogenesis of rabies in insectivorous bats. *Journal of Infectious Diseases* **110** 369-388.

Veith M, Beer N, Kiefer A, Johannesen J, Seitz A (2004) The role of swarming sites for maintaining gene flow in the brown long-eared bat (*Plecotus auritus*). *Heredity* **93** 342-349.

White PCS, Harris S, Smith GC (1995) Fox contact behaviour and rabies spread : a model for the estimation of contact probabilities between urban foxes at different population densities and its implications for rabies control in Britain. *Journal of Applied Ecology* **32** 693-706.

Weber TP, Stilianakis NI (2007) Ecological immunology of avian influenze (H5N1) in migratory birds. *Emerging Infectious Diseases* **13** 1139-1143.

Wilkins JF, Marlowe FW (2006) Sex-biased migration in humans: what should we expect from genetic data? *BioEssays* **28** 280-300.

7. Appendix

Table 7.1 Sample site information for all sample sites included in this study

Site Name	Code	Region	Country	Latitude	Longitude
Addingham Mill	AHM	Yorkshire	England	53.939	-1.863
Arncliffe	AR	Yorkshire	England	54.144	-2.103
Bar Pot	BP	Yorkshire	England	54.147	-2.381
Barden	BD	Yorkshire	England	54.012	-1.922
Browgill Gave	BR	Yorkshire	England	54.195	-2.308
Buckden	BU	Yorkshire	England	54.191	-2.093
Buckden Gavel Lead Mine	BG	Yorkshire	England	54.200	-2.070
Kettlewell Bridge	KB	Yorkshire	England	54.147	-2.052
Grassington	GR	Yorkshire	England	54.070	-2.006
Coverbridge	CB	Yorkshire	England	54.278	-1.780
Dow Cave	DC	Yorkshire	England	54.165	-2.026
Dowkabottom Cave	DB	Yorkshire	England	54.116	-2.076
Fountains Abbey	FA	Yorkshire	England	54.108	-1.582
Ilkley	IK	Yorkshire	England	53.932	-1.814
Malham Tarn	MT	Yorkshire	England	54.100	-2.162
Wensley Bridge	WB	Yorkshire	England	54.300	-1.862
Windy Pits	WP	North York Moors	England	54.245	-1.121
Withcall	WC	Withcall	England	53.367	-0.009
Stokesay Castle	SC	Southern England	England	52.429	-2.831
Kent	KT	Southern England	England	51.251	0.079
Sussex	SX	Southern England	England	52.105	1.083
Scottish North Clade A	SCAN	Scotland	Scotland	NA	NA
Scottish North Clade B	SCBN	Scotland	Scotland	NA	NA
Scottish South Clade B	SCBS	Scotland	Scotland	NA	NA
Marble Arch Cave	MA	Ireland	Ireland	54.246	-7.816
Boho Cave	BH	Ireland	Ireland	54.348	-7.804
Bornem	BB	Belgium	Beligum	51.100	4.235
Liezele	LB	Belgium	Beligum	51.060	4.280
Finland	FI	Finland	Finland	60.280	21.963
Denmark	DE	Denmark	Denmark	56.450	9.189
Fougeres	FO	Northern France	France	48.350	-1.200
Frayssinet	FN	Northern France	France	44.338	1.704
La Pyramide	LP	Southern France	France	44.088	1.744
Le Cros	LC	Southern France	France	44.321	1.778
Pluherlin	PL	Northern France	France	47.696	-2.363
Senchet	SE	Southern France	France	44.096	1.685
Tremblay	TR	Southern France	France	48.950	2.573
Roquefort	RQ	Southern France	France	44.036	0.349
Grotta Marelli	GM	Grotta Marelli	Italy	45.863	8.780
Tiberio	TI	Tiberio	Italy	43.372	12.748
Poland	PO	Poland	Poland	52.401	15.536
Ein Alt Garroxta	EG	Northern Spain	Spain	42.018	-2.480
Itxulegor	IT	Northern Spain	Spain	43.072	-2.814
Sierra de Entezia	SI	Northern Spain	Spain	42.836	-2.284
Cathy	CA	Switzerland	Switzerland	46.496	6.137
Pleine Lune	LU	Switzerland	Switzerland	46.536	6.329

7.1 PCR conditions & Primers as described in Jan *et al.* 2010

A total of 23 primer pairs designed to amplify autosomal microsatellites in Vespertilionidae were designed by Camille Jan (2010). For details of primer sequences, dye and allele size range see table 7.2.

Two multiplex PCR reactions (see Table 7.2 for details) were performed on samples using the following conditions: 95 °C for 15 minutes, followed by 35 cycles of 94 °C for 30 seconds, 58.5 °C for 90 seconds, 72 °C for 60 seconds and a final extension step of 60 °C for 30 minutes. Each PCR constituted a 2 µl reaction containing 2 µl of dried genomic DNA and 2 µl of Qiagen multiplex kit, which contained 0.2 µM of both the forward and reverse primer for each primer pair.

The PCR product was diluted in 30 µl of ddH_20 and 0.5 µl of the dilution was mixed with 9.5 µl of formamide, containing ROX size standard. The PCR product dilution was then run on an ABI3730 capillary sequencer (Applied Biosystems).

Multiple x N°	Dye	Marker	Forward Primer	Reverse Primer	Colour	Size min	Size max
1	Fam	C112b	GGCAAATGCAAGTTGGTAAATAAAC	GGCATTGGCCATTTTGTAAAG	blue	380	500
1	Fam	Clone A2a	TGGCCCATGCTCATCATC	CTGGTCTCAACTGGGTGCTC	blue	97	135
1	Fam	G2	TGAAAAGAACTGGAGAGGCTTT	AGATTGATGAATGTGAAAGGTCAG	blue	341	376
1	Fam	G31	GATCACCAATCATGTAAGGTTCAC	AAGTCAAGGCCAAGCAAGTC	blue	186	281
1	Ned	A2	TTTGTTGAGTAAATGAGTGGATGAATG	GTCTCCCTCTCCCCTGGAAC	yellow	135	176
1	Ned	A24	GTGGTATGAAATAACCAGTTCACTTTG	CAGACTGCATTACTGAAGAAATTATGG	yellow	320	486
1	Pet	A13	GCCACACTCTGAGCCTAGAAGG	TTCCCACCAGCCTTGCTC	red	335	489
1	Pet	EF15	GATCGCAGTCCCTTCC	GCTTATGGGGAGAAATGAG	red	200	268
1	Pet	G6	GGCTTTTTGAAAAGACTGAGG	ACATCAGCCAGTTCCTGTTC	red	100	140
1	Vic	ES43	AAGGGGGAGAGGAGTGG	GCTGCGTGTCCAGAGG	green	366	410
1	Vic	Kpa22	TCAGAAGGGCAAGTATTGTTAGTATC	TGGAGAAAGCTTGTGACTGC	green	135	168
1	Vic	Kpa24	GGCCAGGATGCACAGGAG	CTCCAATCAGCACGTGGATTC	green	235	317
2	Fam	C112a	GCAGGTGTGTGCCTGTG	GTCTCCTAAGTCAGGACTCCATC	blue	50	180
2	Fam	Clone A2b	TCAGCTGTTTCCTGTCAACTTTC	TCCTGTCTTTGATGTAGGAATGG	blue	398	436
2	Fam	Kpa16	TGCCATCAGGATTCAGTTTC	ACACCCTGAAAGCCACATC	blue	211	319

2	**Ned**	**A18**	**GCAGCTGACCAATGATTC**	**TTCAAACTTCCCTGCTAGTC**	**yellow**	**112**	**256**
2	Ned	D15	AAATTCTTTCCCTCCAAAGTGG	GCACGCTCAGACTCCTTCC	yellow	331	379
2	**Pet**	**A45**	**GGTAAAAATCTCATCATCTGC**	**AGGCATAAAACAGCAAGC**	**red**	**120**	**201**
2	**Pet**	**F19**	**TGTAGCTAGCCATGGAGAAGG**	**AAATGGTTACATTACAGAAAATGCTC**	**red**	**403**	**448**
2	**Pet**	**G30**	**GGCATGAACATGGAGTGAGG**	**GCTAGAAGTTATGGTCAATGTTCCTG**	**red**	**243**	**335**
2	**Vic**	**B8**	**AAATACCTGAGTGAGAACATTTAGTGGAG**	**CTCATTAACTTCATTGGTAAGTGTTGTACC**	**green**	**368**	**440**
2	Vic	H23	TTGTCTACTAGCATTTGTCCAGTG	ATAGCTATGTTGCCTAACCTATTTACTC	green	134	208
2	**Vic**	**Mschreib3**	**AGCCAGGCACAGCTCAC**	**GTTTTCTTTGGCATCTGAAGG**	**green**	**256**	**304**

Following genotyping of the whole *M. daubentonii* sample set 9 primer pairs were dropped from further analysis due to scoring difficulties. Fourteen primer pairs (highlighted in bold) were retained.