University of Sussex

Investigating the effective population size of animals

Jennifer E. James

Submitted for the degree of Doctor of Philosophy

University of Sussex

September 2017

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree

Signature: ……..

University of Sussex

Jennifer E. James, Doctor of Philosophy

Investigating the effective population size of animals

Summary

In this thesis I have investigated variation in the effective population size (N_e) between species, and the impact that this population genetics parameter has on molecular evolution.

In Chapter 1 I review literature in order to outline our present understanding of variation in N_e, both between species and within a genome.

In Chapter 2 I determine whether island species have lower effective population sizes than their mainland counterparts. I found that island species did not differ substantially from mainland species in terms of molecular evolution, despite their considerably smaller ranges.

Chapter 3 examines the role of life history and demographic traits in shaping molecular evolution in mammals. Using mitochondrial DNA, I found significant correlations with species range for both genetic diversity (π_S) and the efficiency of selection (π_N/π_S). Both latitude and body mass are also predictive of π_S. However, these relationships are surprisingly weak. Additionally, no trait was predictive of nuclear molecular evolution.

In Chapter 4 I determine whether there is adaptive evolution in animal mitochondrial DNA using McDonald-Kreitman style tests. While mitochondrial evolution is dominated by deleterious mutations, mitochondria also experience adaptive evolution, such that 26% of all nonsynonymous mutations are fixed by adaptive evolution. I also found evidence to suggest that the rate of adaptive evolution is correlated to N_e.

In Chapter 5 I explore the relationship between π_N/π_S and π_S, two variables that are expected to depend on N_e. I quantified the relationship between π_S and π_N/π_S, after controlling for the statistical non-independence between the two, to show that as π_S doubles, π_N/π_S is reduced by 34%. I also investigated whether the slope of the regression between these variables is predicted by the shape parameter of the distribution of fitness effects.

In Chapter 6 I give a general overview of my research, and bring together the key findings of this thesis.

Acknowledgements

My first and most important thanks are due to Adam Eyre-Walker, whose excellent supervision and guidance has made all of the stages of my PhD enjoyable and enriching. I also thank the University of Sussex and NERC, who supported this research. Thanks are also due to three collaborators: Robert Lanfear, who gave me my introduction to the infamous PAML, and whose humour and enthusiasm spurred me on at the start of my PhD; Gwenael Piganeau, whose help and generosity allowed me to get to grips with big DNA datasets; and David Castellano, who is my comrade in running forward simulations and whose curiosity and insight has been a real help over my last projects.

In working at the University of Sussex, I have enjoyed the company of many lovely people who have generally made my time here great: there are far too many to name. However, I would particularly like to thank my office-mates past and present: Dan, Tom, Alex, Beth, Juraj and Viv. My last thanks go to my fantastic family and wonderful friends, who have been a great support and who have tolerated many alternating bouts of radio silence and complaining from me. I am very lucky to have you- you're all smashing.

Preface

The research presented here was conducted at the University of Sussex. Parts of this thesis have been accepted for scientific publication; details are as follows:

Chapter 2

James, J. E., Lanfear, R., Eyre-Walker, A. 2016. Molecular evolutionary consequences of island colonization. *GBE*, 8 (6), 1876-1888.

Chapter 4

James, J. E., Piganeau, G., Eyre-Walker, A. 2016. The rate of adaptive evolution in animal mitochondria. *Molecular Ecology*, 25 (1), 67-78

Chapter 5

James, J. E., Castellano, D. & Eyre-Walker, A. 2017. DNA sequence diversity and the efficiency of natural selection in animal mitochondrial DNA. *Heredity*, 118 (1), 88-95.

Table of Contents

List of Tables **ix**

List of Figures **x**

Chapter 1 **1**

1.1 General introduction 1

1.1.1 Introducing effective population size (N_e) 1

1.1.2 Methods of estimating N_e 1

1.1.3 Factors affecting effective population size 2

1.2 Variation in N_e between species 3

1.2.1 Patterns of genetic diversity 6

1.2.2 N_e and Lewontin's paradox 7

1.2.3 Patterns in measures of the efficiency of selection 8

1.3 Variation in N_e across the genome 10

1.3.1 Patterns of genetic diversity- variation in mutation rate or N_e? 10

1.3.2 Patterns of efficiency of selection 11

1.4 Objectives of this thesis 12

Chapter 2 **14**

Molecular Evolutionary Consequences of Island Colonisation

2.1 Abstract 14

2.2 Introduction 14

2.3 Methods 17

2.3.1 Dataset 17

2.3.2 Statistical tests 18

2.3.3 Polymorphism data 18

2.3.4 Substitution data 18

2.3.5 Adaptive evolution tests 19

2.4 Results 19

2.4.1 Dataset overview 19

2.4.2 Geography 20

2.4.3 Synonymous diversity (π_S) 21

2.4.4 Effective population sizes 26

2.4.5 Efficiency of selection 26

2.4.6 Adaptive evolution 32

2.4.7 Mutation rate 32

2.5 Discussion 33

Chapter 3 **38**

Investigating the life history and demographic traits that are predictive of Ne

3.1 Abstract 38

3.2 Introduction 38

3.3 Methods 40

3.4 Results 41

3.4.1 Relationships between-traits and mitochondrial molecular evolution 42

3.4.2 Relationships between traits and nuclear molecular evolution 46

3.4.3 Relationship between mitochondrial and nuclear molecular evolution 48

3.5 Discussion **49**

Chapter 4 **53**

The rate of adaptive evolution in animal mitochondria

4.1 Abstract 53

4.2 Introduction 53

4.3 Methods 56

4.3.1 Combining data 56

4.3.2 Estimating the rate of adaptive evolution 59

4.3.3 Independence 59

4.4 Results 60

4.5 Discussion **68**

Chapter 5 **72**

DNA sequence diversity and the efficiency of natural selection in animal mitochondrial DNA

5.1 Abstract 72

5.2 Introduction 72

5.3 Methods 73

5.3.1 Dataset 73

5.3.2 Relationship between π_N and π_N/π_S 74

5.3.3 Correcting for phylogenetic non-independence 74

5.3.4 Simulations 75

5.3.5 Calculating the DFE 76

5.4 Results 76

5.4.1 Simulating the method 77

5.4.2 Overall relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ 81

5.4.3 Correcting for phylogenetic non-independence 82

5.4.4 Taxonomic groups 83

5.4.5 Life history and demographic traits 83

5.4.6 Comparison of the slope to the shape of the DFE 86

5.5 Discussion **87**

Chapter 6 **92**

6.1 General Discussion 92

6.1.1 The effect of island colonisation on molecular evolution 92

6.1.2 Molecular evolution and life history 93

6.1.3 Adaptive evolution in animal mitochondria 93

6.1.4 Quantifying the relationship between neutral diversity and the efficiency of selection 94

6.2 Overview and Perspectives 95

Bibliography **98**

Appendices **113**

List of Tables

1.1) List of estimates of intermediate term N_e 5

2.1a and b) An overview of the sequences used in the analysis 20
2.2) Differences in π_S between island and mainland species 22
2.3) Differences in $\pi_N/(\pi_N+\pi_S)$ between island and mainland species 28
2.4a) 30
2.4b) Differences in ω between island and mainland species 31
2.5) Differences in DoS between island and mainland species 32
2.6) Differences in d_S between island and mainland species 33

3.1) Table of values of Pagel's λ 42
3.2) Results of correlation analyses in mitochondrial DNA 43
3.3) Results of multiple linear regression of traits on π_S(mt) and $\pi_N/(\pi_N+\pi_S)$(mt) 46
3.4) Results of correlation analyses in nuclear DNA 47
3.5) Results of multiple linear regression of traits on π_S(n) and $\pi_N/(\pi_N+\pi_S)$(n) 48

4.1) A summary of the DoS results 60
4. 2) The species for which DoS is significantly positive 61
4.3) Estimates of α calculated using a variant of the Messer-Petrov method 63
4.4) Predicted estimates α using the Messer-Petrov method under different DFEs 65
4.5) Results table showing estimates of α calculated using the parametric method 66
4.6) The strength and statistical significance of the correlation between ω_α and π_S. 68

5.1) Mean slope values from simulated datasets under various parameter combinations 80
5.2) Comparisons of the regression slope and intercept for different life history traits. 84
5.3a) Estimates of the DFE. 87
5.3b) Regression slope and intercept estimates. 87
5.4) Estimates of the DFE for different taxonomic and life history groups 89

List of Figures

1.1) Distribution of Ne estimates 6

2.1) The frequency distribution of the ratios of island:mainland species range areas 21
2.2) Relative island diversity plotted against total island-mainland divergence 25
2.3) Frequency distribution of the ratio of island π_S : island and mainland d_S 35

3.1a) The relationship between the global range of a species and π_S 44
3.1b) The relationship between the global range of a species and $\pi_N/(\pi_N+\pi_S)$ 44

4. 1) Histogram showing the frequency distribution of DoS values for the dataset. 61
4.2) Our estimate of α plotted against the frequency category of polymorphism 63
4.3) Bar chart of observed and expected values of p_n/p_s 67

5.1) The relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ in simulated data 78
5.2) The relationship between π_N/π_S and π_S in mammalian mitochondrial DNA 82
5.3) Comparison of the relationship between π_N/π_S and π_S of bats and rodents 83
5.4) The influence of species' traits on the relationship between π_N/π_S and π_S. 85

Chapter 1

1.1 General introduction

In this thesis, I explore the role of effective population size (N_e) on shaping patterns of molecular evolution. In this general introduction, I will begin by briefly outlining the population genetics concept N_e, how it is measured and the effect it has on the molecular evolution of a population. I will then discuss the numerous factors that are expected to influence N_e, before going on to review our current understanding of variation in N_e, focusing on the question: 'Does N_e vary across species, and within a genome?'.

1.1.1 Introducing effective population size (N_e)

The effective population size (N_e) is a central concept in population genetics, first introduced by Wright (1931). The N_e of a population is the size of an idealised Wright-Fisher population that would experience the same level of genetic drift as the population in question. There are a variety of ways to define Ne, depending on the quantity of interest: the two originally identified by Wright were the inbreeding effective population size and the variance effective population size. One of the most commonly used definitions today is the coalescent N_e (Charlesworth 2009). Generally, N_e is a measure of the extent to which a population is impacted by genetic drift, the stochastic fluctuation of allele frequencies in finite populations. These occur due to random differences in the survival and fecundity of individuals. The larger the population, the lower the proportional impact that random events will have on changes in allele frequency, and thus the larger the N_e the lower the impact of genetic drift. N_e has two major impacts on the molecular evolution of a population. Firstly, N_e determines the level of neutral and weakly selected genetic diversity in a population, such that the level of neutral diversity, $\pi_{neutral}$, is proportional to the product of N_e and μ, the mutation rate (Kimura 1984). Secondly it determines the efficiency of selection, i.e.the ability of natural selection to act upon deleterious and advantageous alleles, such that populations with a large Ne have more efficient selection. Mutations that occur in a population are effectively neutral when they have a selection coefficient, s, that is equal to or less than the inverse of the N_e of that population (Kimura 1962; Ohta 1992), and as such the ability of a population to select for or against advantageous and deleterious alleles is determined by the product of N_e and the intensity of selection. Therefore, N_e is of great biological importance in a number of fields, including evolutionary biology, ecology, and conservation.

1.1.2 Methods of estimating N_e

Unfortunately, N_e is difficult to estimate, and there are few estimates of N_e in wild populations. A number of methods for estimating N_e from genetic data have been developed: for a review see (Wang 2005; Wang

et al. 2016). These methods differ in the time-scale for which they estimate N_e: from measures that estimate contemporary N_e, which relate to present-day population size, to intermediate-term measures, to long term measures, which are determined over a period of time equivalent to N_e generations (Wang 2005). Estimates of the contemporary N_e are particularly important for species conservation (Luikart et al. 2010), however, recent work suggests that these methods are dominated by sampling error, and as such can have large biases and variances (Waples 2016). Estimates of medium and long-term N_e are the focus of this thesis/review: these are generally considered more relevant to explaining observed patterns of molecular evolution in populations, due to their longer evolutionary timescale (Nicolaisen & Desai 2012; Nicolaisen & Desai 2013; Wang et al. 2016). Medium-term N_e can be inferred from measures of the neutral genetic diversity using the formula $N_e = \pi_{neutral}/(4\mu)$, assuming that μ, the mutation rate per site per generation, is known (Charlesworth 2009; Wang 2005). However, if μ is not known, inferences about the N_e of a population can still be made from a number of proxies, including estimates of $\pi_{neutral}$ and census population sizes (Lanfear et al. 2014). In addition, as many life history traits are themselves correlated to census population sizes, they too can be used as proxies for N_e (Bromham 2011). In addition, the N_e of a population can be estimated from estimates of the efficiency of selection. We can also potentially obtain longer-term estimates of N_e using molecular traits such as d_N/d_S, the ratio of nonsynonymous to synonymous substitutions, and measures of codon usage bias, because these traits are shaped by the long term efficiency of selection acting on a population.

1.1.3 Factors affecting effective population size

Increasing our understanding and ability to estimate N_e is crucial. Without good estimates of N_e, many studies focus on proxies based on the census population size (N_c), and although N_e is expected to be related to N_c, only in a Wright-Fisher population will the N_e be equal to the N_c (Wright 1931). A Wright-Fisher population is an idealised population that consists of a finite number of diploid, hermaphroditic individuals that mate randomly, with discrete generations and no change in population size over time. Obviously, most real populations differ from a Wright-Fisher population in a number of ways. For example, many eukaryote species have two sexes, many have overlapping generations, many populations have some degree of spatial structure, and most will experience changes in size over time. These differences are important to our understanding of molecular evolution, as any factor that increases the variance in individual reproductive success will tend to decrease the N_e relative to N_c (Charlesworth 2009), such that the ratio of N_e/N_c tends to be less than 1, and is often in the range of 0.1-0.5 for wild populations (Frankham 1995; Palstra & Ruzzante 2008; Palstra & Fraser 2012; Waples et al. 2013). Therefore, in terms of molecular evolution, most populations will 'behave' as if they are smaller than their census numbers would indicate. However, it is worth noting that we do not expect the ratio of N_e/N_c to become infinitely small, as very extreme circumstances are required for the N_e/N_c of a wild population to be less than 0.01 (Waples 2016). Therefore, although we expect N_e to be correlated to N_c the relationship is clearly complicated; this was highlighted in a recent review by Palstra and Fraser (2012) who found better support for a loglinear relationship between the two parameters than the expected simple, linear relationship.

In addition, unlike in a Wright-Fisher population, real populations experience two additional evolutionary forces: selection and recombination. The effect of selection on N_e depends on the type of selection in question. Unidirectional selection on an allele will reduce N_e and genetic diversity for sites that are linked to the allele. This is the case both if selection is acting to remove an allele from a population, as any other mutations that are linked to that allele will also be removed (Charlesworth et al. 1993; Charlesworth 1994), and if selection is acting to fix an allele in a population, as any existing mutations around the selected locus will be lost if they are not in linkage with the selected allele (Maynard Smith & Haigh 1974). These processes are known as background selection and selective sweeps respectively, but can be collectively termed linked selection. Recombination serves to break up linkage between sites and so can mitigate these effects, which might result in average higher effective population sizes in species and genomic regions with high recombination rates (Betancourt & Presgraves 2002; Haddrill et al. 2007; Presgraves 2005). Balancing selection on the other hand maintains diversity in the population for longer than would be expected under neutrality: this form of selection is expected to create localised peaks in N_e at the site under selection (Delph & Kelly 2014). Although there are a few clear examples of balancing selection, for example, the MHC locus in mammals (Hughes & Nei 1988; Hughes & Nei 1989), it is generally considered to be rare (Asthana et al. 2005; Leffler et al. 2013). Finally, in real populations different genomic elements are inherited differently, which will influence their N_e. For example, in eukaryotes with XY sex-determination systems, there are three-quarters as many X-chromosomes as autosomes in the population, and so purely due to chromosome numbers we expect the N_e of the X-chromosome to be lower than that of the autosomes. Organelle DNA may show an even more drastic reduction in N_e relative to autosomes; not only is it inherited solely from the mother, it is almost entirely non-recombining (Charlesworth 2009). Additionally, sexual selection and variation in reproductive success between males and females can have a dramatic impact on the N_e of sex chromosomes and organelle genomes relative to the autosomes, however, whether the ratio will increase or decrease depends on the sex chromosome system, the direction of reproductive skew and the type of selection in question. (Johnson & Lachance 2012).

Due to the action of these diverse evolutionary factors, we expect variation in N_e to exist on a number of levels, both at the between species level, and within a genome. Here we explore recent advances in our understanding of this variation; we will also discuss hypotheses that reconcile observations from real data with theoretical expectations.

1.2 Variation in N_e between species

It is perhaps intuitive for there to be variation in N_e between species, as species clearly do vary in terms of N_c. It is possible to address this question directly for those species for which we have a direct estimate of μ, either from mutation-accumulation experiments or pedigree studies, and an estimate of neutral diversity, using the formula $\pi_{neutral}/(4\mu)$, as described previously. Table 1.1 provides a list of estimates for multicellular eukaryotes, compiled from the literature. We have added 8 additional N_e estimates to the 11 previously discussed in a study by Lynch et al. (2016). These 19 estimates of N_e vary by over two orders

of magnitude, an order of magnitude greater than the variation we observe in either μ or neutral genetic diversity; however, it is likely that we might observe greater levels of variation in a larger dataset (Leffler et al. 2012; Romiguier et al. 2014). The estimates of N_e are largely consistent with what we would expect considering species' census population sizes (see figure 1.1): insects have the largest effective population sizes in the dataset, followed by widespread and extremely abundant species such as herring and *Arabidopsis*. Species with lower effective population sizes tend to be larger and less abundant. While limited, this data illustrates that there is variation in N_e across species which appears to be somewhat related to N_c, and that variation in N_e may exceed variation in neutral genetic diversity across species. However, as μ is not known for many species, the dataset for which we can estimate N_e using this method is currently limited.

Species name	Common name	μ	Genetic diversity/ theta	N_e	References
Apis mellifera	Honey bee	6.80E-09	7.50E-03	276000	(Lynch et al. 2016; Wallberg et al. 2014)
Arabidopsis thaliana	Thale cress	6.95E-09	8.02E-03	289000	(Lynch et al. 2016)
Astatotilapia calliptera, Aulonocara stuartgranti, Lethrinops lethrinus	Cichlid fish	3.50E-09	1.20E-03	85700	(Malinsky et al. 2017)
Bombus terrestris	Bumblebee	3.60E-09	6.10E-03	1069444	(Lattorff et al. 2016; Lecocq et al. 2013; Liu et al. 2017)
Caenorhabditis briggsae	Rhabditidae nematode	1.33E-09	1.42E-03	267000	(Lynch et al. 2016)
Caenorhabditis elegans	Rhabditidae nematode	1.45E-09	3.14E-03	541000	(Lynch et al. 2016)
Chironomus riparius	Harlequin fly	4.20E-09	2.85E-02	1700000	(Oppold & Pfenninger 2017)
Clupea harengus	Atlantic herring	2.00E-09	3.20E-03	400000	(Feng et al. 2017; Barrio et al. 2016)
Daphnia pulex	Common water flea	5.69E-09	1.88E-02	826000	(Lynch et al. 2016)
Drosophila melanogaster	Fruit fly	5.17E-09	1.78E-02	863000	(Lynch et al. 2016)
Ficedula albicollis	Collared flycatcher	4.60E-09	3.95E−03	217391	(Burri et al. 2015; Smeds et al. 2016)
Heliconius melpomene	Postman butterfly	2.90E-09	2.40E-02	2070000	(Lynch et al. 2016)
Homo sapiens	Human	1.35E-08	1.14E-03	21100	(Lynch et al. 2016)
Mus musculus	House mouse	5.40E-09	3.83E-03	177000	(Lynch et al. 2016)
Oryza sativa	Asian rice	7.10E-09	1.50E-03	52800	(Lynch et al. 2016)
Pan troglodytes	Common chimpanzee	1.20E-08	1.38E-03	28800	(Lynch et al. 2016)
Poecilia reticulata	Guppy	4.89E−08	2.50E-03	12800	(Künstner et al. 2016)
Pristionchus pacificus	Diplogastridae nematode	2.00E-09	1.40E-02	1750000	(Lynch et al. 2016)
Prunus persica	Peach tree	8.16E-09	2.62E-03	80400	(Cao et al. 2014; Xie et al. 2016)

Table 1.1) List of estimates of intermediate term N_e, compiled from literature. Estimates are rounded to three significant figures. This dataset is an extension to that of Lynch et al (2016), see this manuscript for further references.

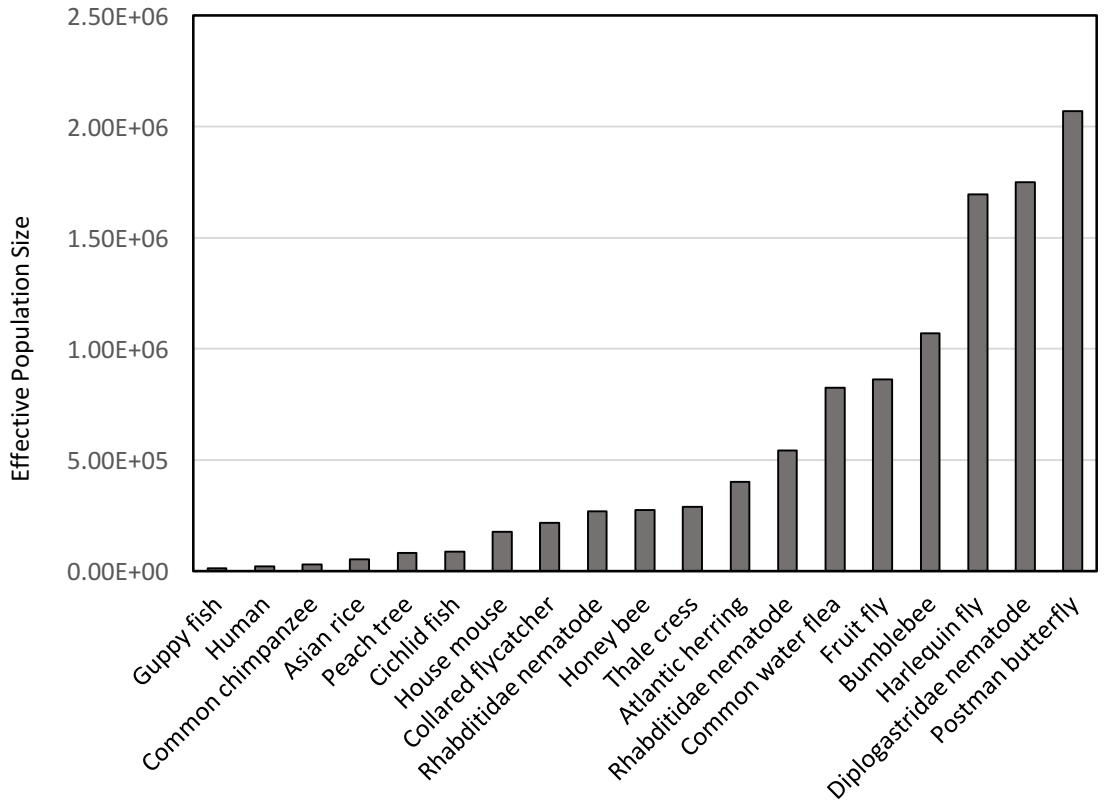

Figure 1.1) Distribution of N_e estimates

1.2.1 Patterns of genetic diversity

In order to overcome the problem that for many species μ is not known, a number of authors have addressed the question of whether there is variation in N_e by comparing levels of neutral genetic diversity between species, using genetic diversity as a proxy for N_e. This method has uncovered some broad-scale patterns that are consistent with the hypothesis that species with large census population sizes have larger effective population sizes. For example, arthropods are more diverse than vertebrates (Frankham 1996), marine species are more diverse than terrestrial species (Bazin et al. 2006) and outcrossing plants are more diverse than selfers (Chen et al. 2017). However, it is important to note that many studies of broad scale patterns across taxonomic groups do not control for the effects of phylogeny. This is important firstly for statistical reasons (Felsenstein 1985), and secondly because animal groups differ in a number of ways that could impact their genetic diversity and N_e, making results difficult to interpret (Bromham et al. 1996; Freckleton et al. 2002; Johnson & Seger 2001).

Animal groups often have different life histories, many of which are thought to be predictive of levels of genetic diversity. This is because many life history traits are themselves indicators of species abundance, and are therefore related to N_e. For example, body mass is known to be negatively correlated to N_e across most species (White et al. 2007). Additionally, body mass is associated with a suite of other life history traits, such that animals with large body masses are likely to have a greater longevity, low mass-specific metabolic rates, low numbers of offspring and long lifespans when compared to small animals (Bromham

2011; Jeschke & Kokko 2009; Lanfear et al. 2010; Oli 2004), thus potentially resulting in a relationship between N_e and many life history traits. Romiguier et al (2014) found relationships between traits such as adult body mass, fecundity and longevity and levels of neutral diversity in a wide range of animal species: their dataset included examples from groups as diverse as cnidarians, chordates, echinoderms, arthropods and nematodes. Surprisingly however, the authors did not find any relationship between genetic diversity and geographic variables related to species abundance.

This unexpected finding suggests that life history factors may in fact be better predictors of neutral genetic diversity/N_e than geographic factors (Romiguier et al. 2014). The reasons for this are unclear: it is possible that many geographic factors are poor proxies of N_c because while they may be indicative of the global range of a species, variables such as average species density are not known and are hard to estimate. Alternatively, life history traits may have additional explanatory power over molecular evolutionary traits as compared to more direct estimates of N_c, because life history traits can affect N_e, and the ratio of N_e to N_c, if they increase the variance in individual reproductive success. Waples (2016) found that across 63 species with varied life histories, a large proportion of the variance in N_e:N_c could be explained by a function of adult mortality and age at maturity. Even a very simple ratio of adult lifespan to age at maturity was strongly predictive of the ratio of N_e to N_c, such that as the ratio of adult lifespan to age at maturity increases, the ratio of N_e to N_c decreases (Waples et al. 2013). Research into the role of geographic factors on shaping variation in genetic diversity and N_e while controlling for the effects of life history may be needed in order to improve our understanding of the relative impacts of these factors on the N_e of a species.

1.2.2 N_e and Lewontin's paradox

Early studies of allozyme diversity revealed a surprising lack of variation in diversity between species which vary greatly in terms of N_c; an observation that came to be known as Lewontin's paradox (Lewontin 1974). This pattern has now been confirmed at the DNA level: a number of studies have found that the level of variation in neutral genetic diversity across species is far smaller than we would expect from the variation in their census population sizes (Leffler et al. 2012; Romiguier et al. 2014). This appears to be particularly true of mitochondrial DNA (mtDNA) diversity; for example, Bazin et al. (2006) reported that levels of mtDNA diversity are very similar across a wide range of taxa, to the extent that there are no significant differences in mtDNA diversity between invertebrates and vertebrates. Similarly, others have failed to find relationships between levels of mitochondrial diversity and measures of N_c in birds (Nabholz et al. 2009) or in mammals (Nabholz, Mauffrey, et al. 2008). These results are somewhat controversial, because Mulligan et al. (2006) showed that mtDNA diversity and nuclear diversity, measured as allozyme heterozygosity, are correlated in mammals (see also Nabholz et al. (2008)), suggesting that mitochondrial diversity does vary across animal groups. This result has been confirmed across many other animal groups (Piganeau & Eyre-Walker 2009). However, it is clear that levels of genetic variation are also limited in animal mitochondria.

It is possible that a lack of variation in genetic diversity simply reflects a lack of variation in N_e across species. However, two resolutions to Lewontin's paradox have been suggested, both of which depend on there being variation in N_e. Firstly, it is possible that levels of genetic diversity are heavily impacted by linked selection. Species with larger effective population sizes may be more affected by linked selection due to selection being more efficient in these populations, leading to greater reductions in neutral genetic diversity in these species (Gillespie 2000; Maynard Smith & Haigh 1974). There is some support for this hypothesis: reduction in diversity due to linked selection appears to be positively correlated with range size, and negatively correlated to body mass (Corbett-Detig et al. 2015). However, it seems unlikely that this effect will reduce genetic diversity sufficiently across species to explain Lewontin's paradox, as the authors estimate that at most, linked selection reduces genetic diversity by approximately 70%. We still expect genetic drift to be the dominant force determining levels of neutral diversity (Coop 2016; Ellegren & Galtier 2016). On the other hand, it is likely that the effects of linked selection are more dramatic in genomic regions which do not experience recombination, and so might be relevant to the patterns observed in mtDNA (Ballard & Whitlock 2004).

Alternatively, it has been suggested that the evolution of the mutation rate is responsible for the low range of values of genetic diversity (Lynch 2007; Lynch 2010). In this theory, known as the drift-barrier hypothesis, selection acts to reduce the mutation rate as far as possible. Because selection is only able to act on alleles where the s is greater than $1/N_e$, selection is better able to reduce μ in populations with a large N_e. As predicted by this theory, there does appear to be a negative scaling relationship with mutation rate per generation and N_e across a broad range of animal groups (Lynch 2010; Lynch 2011; Sung et al. 2012). This is particularly true when μ considered is corrected for the proportion of protein-coding DNA in the genome (Lynch et al. 2016). However, the relationship has yet to be tested on more restricted taxonomic scales. In addition, mutation rates in some species are far higher than would be predicted by this theory (Martincorena & Luscombe 2013). Therefore, it is not known to what extent this theory holds across species.

1.2.3 Patterns in measures of the efficiency of selection

The action of these confounding factors may make interpreting patterns of genetic diversity difficult, and could obscure the relationship between genetic diversity and N_e. A complimentary approach to investigate variation in N_e is to use measures of the efficiency of selection, such as the ratio of nonsynonymous to synonymous polymorphisms, π_N/π_S, or the ratio of nonsynonymous to synonymous substitutions, d_N/d_S, sometimes referred to as ω. These measures are normalised by synonymous rates, and therefore should not be affected by variation in μ. However, because of the different timescales of these measures, our predictions of how they should change with N_e are slightly different. π_N/π_S is a comparatively straightforward measure of the efficiency of selection: we can assume that nonsynonymous polymorphisms segregating in the population are deleterious (although some can be so weakly selected as to be selectively neutral). Advantageous mutations subject to directional selection will not significantly contribute to levels of polymorphism because of their rarity and their rapid fixation rates. As N_e increases, selection is more efficient and so able to remove more slightly deleterious polymorphisms segregating in

the population, resulting in a reduction in π_N/π_S. This is not the case for d_N/d_S: although non-synonymous substitutions may be deleterious or slightly deleterious, and have become fixed between species by genetic drift, nonsynonymous substitutions might also include advantageous mutations that have become fixed due to the action of positive selection. As N_e increases, the proportion of slightly deleterious mutations that become fixed in the population will decrease, however, the proportion of advantageous mutations that become fixed might increase, and therefore we cannot make clear predictions about how d_N/d_S will change with N_e (Ho et al. 2011; Kimura 1984; Kryazhimskiy & Plotkin 2008). One important caveat to the use of d_N/d_S and π_N/π_S is that both of these measures are based on the assumption that synonymous mutations are neutral. This is important to bear in mind when studying species with large effective population sizes, for which there is good evidence for selection on synonymous codon usage (Duret & Mouchiroud 1999; Hershberg & Petrov 2008; Kanaya et al. 2011). However, there is little evidence for selection at synonymous sites in animals with smaller effective population sizes, such as mammals (Duret 2002; Jia & Higgs 2008; Kanaya et al. 2011).

Recent results suggest that there is substantial variation in the efficiency of selection across species, as would be expected if there is variation in N_e between species. For example, Chen et al (2017) found evidence to suggest that nuclear π_0/π_4 (π_0/π_4 is a similar measure to π_N/π_S) is positively correlated with longevity in animals, while in plants π_0/π_4 is related to both longevity and mating system, such that outcrossers, which are expected to have higher effective population sizes, have more efficient selection and thus lower values of π_0/π_4 than selfers. In a study of amniotes, Figuet et al. (2016) found a relationship between d_N/d_S and body mass, longevity, and sexual maturity in nonavian amniotes, however, not in birds. The authors did find that estimates of π_N/π_S correlated to life history in birds, which suggests that N_e is also linked to life history traits in birds. A lack of relationship between d_N/d_S and N_e might occur if substitutions in birds are highly affected by adaptive evolution relative to other amniotes.

There is also evidence for variation in N_e in mitochondrial DNA. Popadin et al. (2007) found that the efficiency of selection, measured using substitution data, acting on mitochondrial DNA was significantly greater in small as opposed to large mammals, and that the efficiency of purifying selection in mitochondria is negatively correlated to generation time, i.e. positively correlated to N_e (Popadin et al. 2013). In addition, Piganeau and Eyre-Walker (2009) showed that *Pn*/*Ps* was correlated to *Ps* in mtDNA for many animal groups, again suggesting that there is variation in N_e between species. However, there are some results that are not predicted by a simple neutral model, reflecting complexities in patterns of the efficiency of selection. Firstly, while purifying selection in nuclear and mitochondrial DNA appears to be concordant, it is always higher in mitochondrial DNA (Havird & Sloan 2016; Popadin et al. 2013): this is not in agreement with the theoretical lower effective population sizes of mitochondria, suggesting that mitochondria might experience stronger selection. In addition, Eory et al. (2010) found that while the efficiency of selection appeared to be greater in murids (rodents) than in hominids (apes) overall, in accordance with their expected differences in N_e, the constraint on some types of site was greater in hominid genomes than in murid genomes. This may reflect strong selection in hominids to maintain elements of genomic structure such as splice sites.

1.3 Variation in N_e across the genome

We now explore evidence for variation in N_e within a species' genome. While linked selection may only play a minor role in determining patterns of variation between species, linked selection, both in the form of selective sweeps and background selection, is expected to be the most important force determining within-genome variation in N_e (Charlesworth et al. 1993; Maynard Smith & Haigh 1974). As such, there are two major predictors of local variation in N_e across the genome: firstly, the density of selected sites, which could result in more linked selection in some regions of the genome than others, and secondly, local recombination rates. Rates of recombination are known to differ across the genome, and in some cases quite substantially: the greater the recombination rate, the smaller the genomic window affected by linked selection and so the greater the variation in N_e (Ellegren & Galtier 2016).

1.3.1 Patterns of genetic diversity- variation in mutation rate or N_e?

Evidence for variation in N_e existing across the genome comes from a number of sources. Firstly, as is the case between species, genetic diversity is known to vary across the genome. This is particularly obvious from comparisons of regions such as sex-chromosomes, which are known to have low recombination rates, and the rest of the genome: these regions have relatively low levels of genetic diversity (Ellegren & Galtier 2016). More generally, we do observe substantial differences in genetic diversity across the genomes of species (Bachtrog & Andolfatto 2006; Begun & Aquadro 1992; Campos et al. 2012). However, this variation could be the result of either variation in N_e across the genome, or variation in the mutation rate across the genome (Hodgkinson & Eyre-Walker 2011). It is possible to distinguish between these possibilities, even though direct estimates of the local mutation rate across the genome are rarely known. The rate of neutral evolution is equal to the neutral mutation rate (Kimura 1984; Ohta 1992) and therefore the rate of neutral (i.e. synonymous) divergence, d_S, can be used as a proxy for μ. If there is a relationship between genetic diversity and the rate of recombination and/or the density of selected sites, and no relationship between these variables and d_S, this suggests the relationship is being driven by variation in N_e, not by variation in μ.

Interestingly, which of these factors drives variation in genomic diversity appears to vary between taxonomic groups. For example, in *Drosophila* there is no relationship between levels of divergence and recombination rates across the genome, while there is a strong relationship between levels of neutral diversity and recombination rates (Begun & Aquadro 1992; Charlesworth 1996). Therefore, there is variation in N_e across the *Drosophila* genome, highlighting the importance of selective sweeps and background selection in shaping the molecular evolution of this species (Andolfatto & Przeworski 2001; Betancourt & Presgraves 2002; Castellano et al. 2015; Presgraves 2005). By contrast, in humans, both diversity and divergence are positively correlated to recombination rate (Bullaughey et al. 2008; Hellmann et al. 2003; Lercher & Hurst 2002). A positive relationship between mutation rate and recombination rate is expected if recombination is itself mutagenic, or if recombination is associated with a mutagenic nucleotide context, for which there is some evidence (Arbeithuber et al. 2015; Francioli et al. 2015; Pratto et al. 2014; Strathern et al. 1995), and it is possible that this could drive a relationship

between the recombination rate and neutral diversity. Hellmann et al. (2003) found no evidence for a correlation between diversity and the recombination rate in humans once they had corrected for interspecies divergence, which supports this hypothesis. However, a recent review of patterns across multiple species suggests that humans may be unusual in there being a relationship between neutral divergence and recombination rate, while a relationship between neutral diversity and recombination is fairly ubiquitous across species (Cutter & Payseur 2013), although with a few exceptions such as *Caenorhabditis remanei* (Cutter 2008), rice (Flowers et al. 2012), or *Arabidopsis* species (Slotte et al. 2011). This suggests that variation in N_e across the genome is likely to be a general phenomenon across species.

1.3.2 Patterns of efficiency of selection

As in studies of inter-species differences in N_e, it is also possible to study differences in within-genome N_e by investigating differences in the efficiency of selection across the genome. A number of authors have found that the efficiency of selection on non-recombining sex chromosomes is greatly reduced: Y chromosome evolution in *Drosophila* species appears to be characterised by rapid degeneration, leading to high values of d_N/d_S for the Y relative to the X chromosome (Bachtrog & Charlesworth 2002; Bachtrog 2003), while bird W chromosomes (these are the heterogametic chromosomes in this group) also accumulate considerably more slightly deleterious nonsynonymous substitutions than Z chromosomes (Berlin & Ellegren 2006). The relationship between recombination rate and the efficiency of selection, and thus N_e, also holds across other genomic regions. In a study of two passerine bird species, Gossmann et al. (2014) found that d_N/d_S was significantly lower in high recombination regions, indicating more efficient purifying selection. This pattern appears to hold over a wide variety of bird species (Künstner et al. 2010). In addition, high recombination regions were significantly more likely to fix advantageous mutations than low recombination regions (Gossmann et al. 2014). As previously noted, similar results have been found in *Drosophila* species (Betancourt & Presgraves 2002; Haddrill et al. 2007). However, in keeping with findings from studies into patterns of genetic diversity, this relationship does differ between taxonomic groups as these patterns are not observed in primates. Bullaughey et al. (2008) found no significant differences in d_N/d_S values between high and low recombination regions across humans, chimps and macaques. This result was not driven by a lack of targets of selection in regions of low recombination. The authors also failed to find evidence for a relationship between d_N/d_S and recombination rate on fine or broad genomic scales.

There are a number of possible explanation for the observed differences between taxonomic groups. It may be that for there to be an observable pattern between recombination rates and proxies of N_e such as the efficiency of selection, there needs to be sufficient variation in recombination rates across the genome. *Drosophila* and birds are both notable in having regions of very low or no recombination (the dot chromosome in *Drosophila*) and very high recombination (microchromosomes in birds) respectively, which may be why we observe a pattern in these species (Gossmann et al. 2014; Haddrill et al. 2007). Species also differ in how tightly spaced genes are , and how much selection has been detected in non-coding regions (Lynch & Conery 2003), both of which will affect the density of selected sites, an

important predictor of variation in N_e. Therefore, although there is clearly intragenomic variation in the N_e, many general patterns between traits may not be obvious across species due to their different genomic structures. For example, Gossmann et al. (2011) found variation in N_e across the genome for a range of species, including humans, yeast, fruit flies and plants, but this variation is not consistently correlated to either recombination or the density of selected sites. Although overall levels of variation were modest, the authors did find sufficient variation to result in differences in the efficiency of selection. Using a modelling approach, the authors suggest that variation in the local mutation rate and variation in local N_e might play equal roles in determining variation in diversity.

1.4 Objectives of this thesis

Despite practical difficulties in obtaining direct estimates of N_e, a number of lines of evidence suggest that variation in N_e does exist, both between species and across the genome, highlighting the importance of stochastic events in shaping patterns of molecular evolution. However, we still do not fully understand the determinants of this variation. Perhaps most surprisingly, many past analyses have not found the predicted relationship between range size, a proxy of N_c, and N_e, as estimated from measures such as neutral genetic diversity and the efficiency of selection, as detailed above. This may be because past studies have tended to focus on uncovering patterns across a broad range of taxonomic groups: the distantly related species used in such studies are likely to exhibit variation in many life history traits, which can also affect N_e and therefore could obscure a relationship between N_e and range size. This issue can be overcome by the use of study systems in which species are known to vary in their global range, but not life history. In chapter 2, we consider this problem by comparing island endemic species to their mainland relatives. Island and mainland species offer a promising system in which to investigate the relationship between the global range of a species and its N_e, without the confounding factor of variation in life history. We explore whether it is possible to detect a difference in island species and mainland species in either π_S, π_N/π_S, or d_N/d_S in a manner which controls for phylogenetic non-independence. While this system has been investigated before, past work has focused solely on substitution data, and while some authors have found significant or marginally significant differences between island and mainland species in terms of d_N/d_S (Johnson & Seger 2001; Woolfit & Bromham 2005), others have found no significant differences between the two (Wright et al. 2009). By using polymorphism data to address the question of whether island species have smaller effective population sizes than mainland species, and by using a large and relatively taxonomically broad dataset, we were able to more directly test whether island species tend to have low effective population sizes. In chapter 3 we investigate the relationship between global species range and N_e more generally, using a dataset of mammalian mitochondrial and nuclear polymorphisms, and controlling for phylogenetic non-independence. We also consider the effect of a number of other life history traits, including mass, age at sexual maturity and longevity. This research fills two gaps in the literature: firstly, while a number of relationships between life history traits and π_S have been reported on a broad taxonomic scale (Romiguier et al. 2014) this results are yet to have been tested on a more restricted taxonomic group, and secondly, the relationship between molecular evolution and life history traits is currently not well understood in mitochondrial DNA. In chapter 4 we go on to

explore mitochondrial evolution further, by considering to what extent mitochondrial DNA undergoes adaptive evolution. In the last data chapter, chapter 5, we attempt to determine the relationship between neutral genetic diversity and the efficiency of selection. A negative relationship between these two variables is expected if the nearly neutral theory holds across species. By quantifying this relationship across mammalian mitochondrial DNA, we will assess the extent of variation in N_e between species, and demonstrate the effect of N_e on molecular evolution.

Chapter 2

Molecular Evolutionary Consequences of Island Colonisation

2.1 Abstract

Island endemics are expected to have low effective population sizes (N_e), firstly because some may experience population bottlenecks when they are founded, and secondly because they have restricted ranges. Therefore we expect island species to have reduced genetic diversity, inefficient selection and reduced adaptive potential compared to their mainland counterparts. We used both polymorphism and substitution data to address these predictions, improving on the approach of recent studies that only used substitution data. This allowed us to directly test the assumption that island species have small values of N_e. We found that island species had significantly less genetic diversity than mainland species; however, this pattern could be attributed to a subset of island species that appeared to have undergone a recent population bottleneck. When these species were excluded from the analysis, island and mainland species had similar levels of genetic diversity, despite island species occupying considerably smaller areas than their mainland counterparts. We also found no overall difference between island and mainland species in terms of the effectiveness of selection or the mutation rate. Our evidence suggests that island colonisation has no lasting impact on molecular evolution. This surprising result highlights gaps in our knowledge of the relationship between census and effective population size.

2.2 Introduction

Island species have long been considered to be under greater threat of extinction than their mainland counterparts (Frankham 1997; Johnson & Stattersfield 1990; Jones et al. 2003; Mckinney 1997; Purvis et al. 2000). Although extinction itself is caused by a number of stochastic factors, not least human activity (Burgess et al. 2013; Pimm et al. 1988), the susceptibility of island populations may also be a consequence of population genetics. Island species are likely to have experienced population bottlenecks at some point in their evolutionary history due to founder events during the initial island colonisation. As only a fraction of individuals from the original population found an island population, only a fraction of the original genetic diversity of the population will be maintained, and effective population sizes (N_e) will be small (Nei et al. 1975). In addition, island species are restricted to relatively small areas, which could impose long-term restrictions on census population sizes, and in turn on long-term N_e. Therefore it may be that island species are genetically vulnerable.

Low diversity and low N_e could theoretically reduce the adaptive potential of a species, as standing levels of genetic variation determine the alleles that are immediately available for evolution to act upon (Barrett & Schluter 2007; Hermisson & Pennings 2005; Messer & Petrov 2013). In addition, populations founded by a small number of individuals will experience increased inbreeding. Inbreeding results in an increasingly homozygous population, and therefore there is a greater risk that deleterious recessive alleles will be exposed (Charlesworth & Charlesworth 1987), which could have significant fitness costs. There is some evidence that bottlenecked species do experience a loss of fitness: for example, Frankham et al. (1999) demonstrated that laboratory populations of *Drosophila* showed reduced evolvability (in terms of ability to tolerate increasing concentrations of an environmental pollutant) after a bottleneck; while Briskie and Mackintosh (2003) uncovered a link between the severity of population bottlenecks and loss of fitness in birds.

In addition, species with low effective population sizes are expected to have inefficient selection, resulting in high levels of deleterious mutations segregating and a tendency to fix deleterious mutations. However, past studies investigating the differences in the efficiency of selection between island and mainland species have provided only limited support for this prediction. Johnson and Seger (2001) found some evidence that island species had less efficient selection, but this was for a small and taxonomically restricted dataset. Woolfit and Bromham (2005) used a much larger and more varied dataset; however, they reported a difference between island and mainland species that was only significant at the one-tailed level, while Wright et al. (2009) found no significant difference between island and mainland species. This may be because previous studies have focused on substitution rates as measures of the efficiency of selection, in particular the ratio of the rate of non-synonymous substitution to the rate of synonymous substitution (ω). The problem with considering substitution data is that a reduction in N_e is expected to increase the rate at which slightly deleterious mutations are fixed, but reduce the rate at which advantageous mutations are fixed, particularly if the rate of adaptation is limited by the supply of mutations. We therefore cannot make a clear prediction about the effect of N_e on ω. This issue can be addressed by using polymorphism data instead of substitution data, using the ratio of nonsynonymous to synonymous polymorphisms, because advantageous mutations, subject to directional selection, are not expected to significantly contribute to polymorphism (Ho et al. 2011; Kimura 1984; Kryazhimskiy & Plotkin 2008).

It seems likely that adaptive evolution might occur for at least some island species, despite their predicted low effective population sizes, due to the fact that the species is encountering a novel habitat. Although populations with large effective population sizes may have more efficient selection, we might also expect positive selection to play a significant role after colonisation events as species adapt to new environmental requirements and ecological niches. However, in making predictions regarding adaptive evolution it is important to consider the direction of colonisation. Although island species most commonly colonise an island from a nearby mainland, occasionally lineages that originated on islands re-colonise a mainland, providing an interesting contrast in terms of molecular evolution. Species colonising the mainland from islands are likely to experience population size increases, and therefore increases in N_e.

This could result in a spate of rapid molecular evolution in the new mainland population as advantageous mutations that were previously effectively neutral become fixed (Charlesworth & Eyre-Walker 2007; Takano-Shimizu 1999).

However, predictions about the molecular evolution of island species are predicated on the crucial assumption that island species do in fact have lower N_e and levels of genetic diversity than mainland species. Whether this is in fact the case is not certain, because census population size can sometimes be a poor indicator of genetic diversity (Bazin et al. 2006; Leffler et al. 2012; Lewontin 1974; Romiguier et al. 2014). Although some studies uncover a link between the two (for overview, see (Frankham 2012)), other authors have not found a relationship; for example, Bazin et al. (2006) and Nabholz et al. (2008) failed to find any strong relationship between mitochondrial diversity and traits associated with N_e (such as body mass), or between diversity and IUCN category, an index partly based on assessments of census population size. More generally, there is surprisingly little variation in levels of diversity between species; one recent paper reported a range of nucleotide diversities of 800-fold across a range of taxa, with most species falling within a range of 50-fold, many orders of magnitude smaller than their estimated census population size differences (Leffler et al. 2012). The determinants of genetic diversity remain poorly understood.

One possible complicating factor is the mutation rate. Both Nabholz et al. (2008) and Romiguier et al. (2014) found evidence suggesting that there are lineage-specific differences in the mutation rate, in mitochondrial and nuclear data respectively. How the mutation rate evolves is contentious: if selection is responsible for determining the mutation rate, populations with high effective population sizes should have the lowest mutation rates, because selection will be more effective at reducing the rate (Lynch 2010). This is because whether a mutation can be selected depends on the strength of selection being greater than $1/N_e$. However, support for this prediction remains mixed. For example, in previous studies of island-mainland systems (all of which controlled for phylogenetic non-independence), two found no difference in substitution rate between island and mainland lineages (Johnson & Seger 2001; Woolfit & Bromham 2005), while another found that it was mainland species that had higher rates of substitution (Wright et al. 2009), the opposite of what we might expect if the mutation rate depends on the population size. Another factor that may contribute to unexpected patterns of diversity is selection at linked sites: this reduces genetic diversity, particularly in genomic regions with low rates of recombination (Frankham 2012; Gillespie 2000; Maynard Smith & Haigh 1974). Linked selection may occur more frequently in populations with high values of N_e, reducing diversity more rapidly than in populations with a low N_e (Corbett-Detig et al. 2015). On the other hand, it could be that selective sweeps occur more commonly in species adapting to a new environment e.g. (Montgomery et al. 2010).

In summary, we expect island species to have low effective population sizes and because of this we expect them to have low genetic diversities. We also expect selection to be less efficient in island species, leading to higher ratios of nonsynonymous to synonymous polymorphism, and potentially to increases in the mutation rate (the mutation rate might increase to such an extent that island and mainland species

have similar diversities, but this is expected to take some time to occur). Whether we expect island species to have higher ratios of nonsynonymous to synonymous substitution depends on how much adaptive evolution there is, and how this is affected by N_e and the act of colonisation. If there is no adaptive evolution then island species are expected to have higher values of ω; however, adaptive evolution could potentially be either reduced in island species because of their low N_e or increased because of adaptation to a new environment, given that in most cases the island is the new environment that is colonised. Here we perform the first analysis of polymorphism data from a dataset of phylogenetically independent pairs of island and mainland species, and combine this with substitution data. The paired study design is crucial: there are a large number of life history traits that are known to influence molecular evolution (e.g. body size, fecundity, generation times) and could therefore act as confounding factors (Bromham 2011; Lanfear et al. 2013). Closely related island and mainland species have similar life-history traits, and even if there is variation it is not expected to be systematic, and so should not bias our results. Therefore, island colonisation itself should be the primary reason for any differences in molecular evolution between island and mainland species (Johnson & Seger 2001; Woolfit & Bromham 2005).

2.3 Methods

2.3.1 Dataset

The dataset was compiled by combining all of the independent island-mainland species comparisons used in two previous studies: 33 from Wright et al. (2009) and 34 from Woolfit and Bromham (2005). This dataset was then expanded using a keyword search ('endemic') of the Arkive species database (http://www.arkive.org/). One or more mainland relatives and outgroup species were then identified for each island species. This added 45 species comparisons to the dataset. Some comparisons contained a single island and mainland species, while some consisted of multiple island and/or mainland species. Each island-mainland comparison in our dataset can be considered as a phylogenetically independent contrast, i.e. they are statistically independent, in that the species in each comparison share a common ancestor to the exclusion of all other species in the dataset. Phylogenies were subsequently estimated for each comparison, as detailed in section 2.3.4. All phylogenies were checked for agreement with the literature, and apparent direction of colonisation was noted. In addition, the recorded range area of the species used was calculated from IUCN records (IUCN 2014) using ArcGIS (ESRI 2011). Endemic species of islands with very large areas (such as Madagascar) were excluded on the grounds that these species are unlikely to experience restricted ranges. The endemic species with the largest ranges in this analysis are found on Cuba. Protein coding sequences were collected from NCBI (www.ncbi.nlm.nih.gov/genbank/). Sequences were collected if there was an orthologous gene available for each of the island, mainland, and outgroup species in a comparison, or if there were multiple sequences of the same loci available for both the island and the mainland species in a comparison. A note was made of whether the sequences were nuclear, mitochondrial or chloroplast. All alignment files are available online from FigShare at: http://dx.doi.org/10.6084/m9.figshare.1296151.

2.3.2 Statistical tests

This study has a paired design, in that each island species/clade is compared to a closely related mainland species/clade, with each comparison occurring only once in each analysis. If a choice had to be made between comparisons (for example, if statistics from both the mitochondrial and nuclear genomes were available for a single comparison) the statistics that corresponded to the longest sequence alignment were used. This decision should reduce sampling error, because longer sequences are more representative than short sequences. We also calculated a relative value for each comparison, since values can differ considerably between different island-mainland comparisons. To do this we divided each island statistic by the sum of the island and mainland statistics; e.g. if the statistic being considered is X (for example, the nucleotide diversity), we calculate the relative value as X'(island) = X(island) / (X(island) + X(mainland)). Using this method, if the island and mainland values are the same then the relative island value will be 0.5. Therefore, to quantify the difference between island and mainland values we used a Wilcoxon signed-ranks tests to assess whether the median of the distribution of relative island values was significantly different from 0.5. In order to assign confidence intervals to our results we bootstrapped the data, using 1000 bootstrap datasets. For each bootstrap, the relative island values were randomly resampled (with replacement), and the mean of the relative values was calculated.

2.3.3 Polymorphism data

Sequences (of the same loci from the same species) were aligned by eye using Geneious; the alignment was then analysed using our own scripts. A number of statistics were recorded, including nucleotide diversity and number of polymorphisms. If a comparison included multiple island and/or multiple mainland species, average values of each statistic were taken across the species. Similarly, if multiple sequences from the same genome were available for a particular island/mainland comparison, the average value of the sequences was used. Therefore, each comparison is represented by a single island, mainland, and outgroup value of each polymorphism statistic for a particular genome.

The data was used to calculate $\pi_N/(\pi_N+\pi_S)$, where π_N is nonsynonymous diversity and π_S is synonymous diversity. This ratio is used because, unlike polymorphism counts, nucleotide diversity is unbiased by the number of chromosomes sampled. In addition, using total diversity as the denominator reduces the number of undefined values to those comparisons in which both the island and mainland species had no diversity and were therefore uninformative. Any comparisons with undefined values were excluded from the analysis.

2.3.4 Substitution data

Substitution data was calculated by aligning orthologs of island, mainland and outgroup species. If multiple sequences at different loci were available for all of the species in a comparison, sequences were concatenated prior to alignment; however, sequences from different genomes of the same organism were treated separately. The alignments were pruned so that they included equal numbers of island and mainland species to control for the node-density effect (Hugall & Lee 2007), and then used to generate phylogenetic trees with RaxMl (Stamatakis 2014), in combination with PartitionFinder (Lanfear et al.

2012). The trees were subsequently used to run the codeml progamme of PAML version 4.7 (Yang 2007), which calculated ω (d_N/d_S) for island, mainland, and outgroup branches of each tree, as well as separate d_N and d_S values for each branch.

2.3.5 Adaptive evolution tests

Polymorphism and substitution data were combined to test for differences in levels of adaptive evolution between island and mainland species. A variant of the direction of selection (DoS) statistic was used, calculated as: DoS = $d_N/(d_N+d_S) - \pi_N/(\pi_N+\pi_S)$ (Stoletzki & Eyre-Walker 2011). This statistic has the advantage over using the neutrality index in that it is defined for all datasets in which there is at least one substitution and one polymorphism, so fewer species comparisons had to be excluded; it is also expected to be unbiased (Stoletzki & Eyre-Walker 2011). Positive values indicate that the dynamics of evolution are dominated by positive selection and negative values that slightly deleterious mutations predominate.

2.4 Results

2.4.1 Dataset overview

To investigate the consequences of island colonisation on molecular evolution we compiled data for 112 island-mainland comparisons. In approximately 90% of cases, the inferred direction of colonisation is from mainland-to-island,. The data is dominated by mitochondrial sequences from birds, which comprise 40% of the species comparisons (Table 2.1a), but we have a reasonable number of mitochondrial sequence comparisons available for invertebrates (11%) and (non-avian) reptiles (13%), and a moderate number of nuclear sequence comparisons (approximately 20% of all available comparisons are nuclear DNA). The sequences used in this analysis are on average 750 nucleotide bases long. For 70 of our comparisons, multiple sequences from the same species were available, allowing us to conduct polymorphism analyses. The mean number of sequences available per species was 7. Again, this dataset is dominated by mitochondrial data from birds (Table 2.1b). For a full list of species used in this analysis, please see the archived data at: http://dx.doi.org/10.6084/m9.figshare.1296151.

Divergence	Mitochondrial	Nuclear	Chloroplast	Combined dataset
Amphibian	1	2	-	2
Bird	60	9	-	60
Invertebrate	15	3	-	15
Mammal	2	2	-	2
Plant	-	2	10	12
Reptile (non-avian)	18	14	-	21
Total	96	32	10	112

Table 2.1a)

Polymorphism	Mitochondrial	Nuclear	Chloroplast	Combined dataset
Amphibians	-	1	-	1
Bird	37	2	-	37
Invertebrate	11	1	-	11
Mammal	1	-	-	1
Plant	-	1	4	4
Reptile (non-avian)	11	9	-	16
Total	60	14	4	70

Table 2.1b)

Table 2.1a and 2.1b) An overview of the sequences gathered in this analysis, split by DNA type and taxonomic group. For analyses that combined data across DNA types, each species comparison appeared only once: the numbers of sequences available in these cases are given in the 'combined dataset' column. When choosing between sequences from different genomes for a particular comparison, we always used the longest sequence.

2.4.2 Geography

Island species are studied from a molecular evolutionary perspective because they are expected to have smaller populations than mainland species due to their small ranges. However, this assumption is rarely tested. In this study, the ranges of the species used were confirmed where possible using the IUCN database (IUCN 2014). The mean range of island species was 5,780 km^2, while for mainland species this mean range was over 4,080,000 km^2. The ratio of island to mainland range sizes did not exceed 0.25 for any of the comparisons used, and in the majority of cases island species had ranges which were less than 1% of the area of those of their mainland counterparts (Figure 2.1). Therefore we have evidence that the island species used in this study inhabit substantially smaller geographic regions than their mainland relatives, although we have no information on population density.

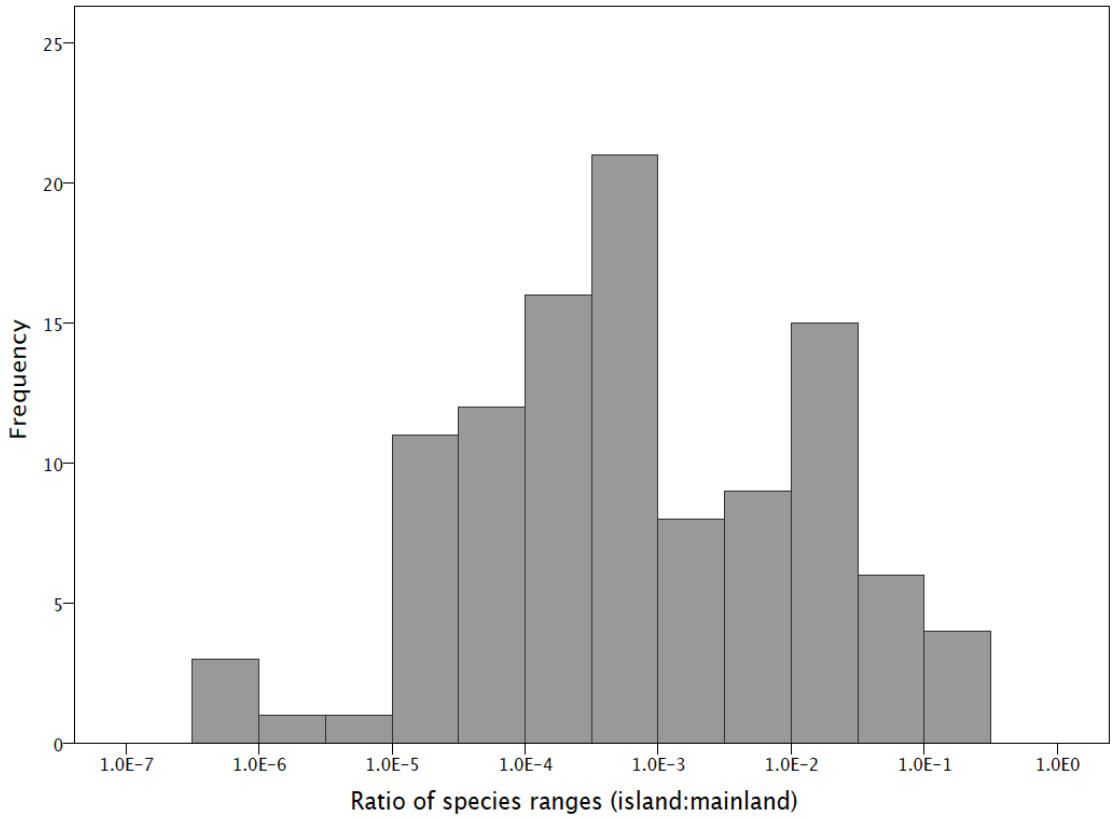

Figure 2.1) The frequency distribution of the ratios of island:mainland species range areas

2.4.3 Synonymous diversity (π_S)

We might expect island species to have lower diversity than their mainland counterparts for two reasons. Firstly, island species inhabit substantially smaller areas than their mainland relatives, resulting in a smaller census population size and hence potentially a smaller long-term N_e. Secondly, island populations are likely to be founded by few individuals, which again is expected to result in a small N_e. Since diversities can differ quite substantially between phylogenetic groups, we calculated relative values of island diversity from each comparison by dividing each by the sum of the island and mainland diversities. Therefore, if island π_S is significantly smaller than mainland π_S, the relative island values will be significantly lower than 0.5.

As expected, we find that island species have significantly lower π_S for both our combined dataset, and when we consider mitochondrial and nuclear DNA separately (Table 2.2). Chloroplast sequences show the opposite pattern, but as there are only 2 comparisons this is likely to be due to sampling error. When different taxonomic groups were considered separately, island birds and island reptiles both had significantly lower π_S than their mainland counterparts, while there was no significant difference between island and mainland invertebrates (Table 2.2) (for other groups we do not have enough data to make a valid comparison). Although we find that island species have lower diversity than mainland species, there is no significant correlation between relative island diversity and the ratio of the island and mainland ranges, either overall or for any subset of the data (see Table 2.2). However, despite being statistically

significant, the differences between mainland and island species are relatively modest. Island species have a mean π_S that is only 31% smaller than that of mainland species, and in about one third of cases, island species have higher π_S than their mainland relatives.

Dataset	*n*	Mean Island π_S	Mean Mainland π_S	Mean relative island π_S	Lower CI	Upper CI	Wilcoxon P-value	Spearman's rho of the correlation between the ratio of ranges and relative island π_S
Combined	70	0.027	0.039	0.36	0.29	0.45	**0.0013**	0.14
Chloroplast	2	0.0023	0.00058	0.85	0.69	1	1	-
Mitochondrial	60	0.032	0.052	0.38	0.29	0.46	**0.0041**	0.15
Nuclear	14	0.0015	0.0069	0.22	0.039	0.45	**0.039**	0.039
Bird	37	0.011	0.028	0.34	0.24	0.45	**0.0035**	0.041
Invertebrate	11	0.078	0.058	0.53	0.33	0.73	0.69	0.3
Reptile (non-avian)	16	0.037	0.052	0.27	0.095	0.44	**0.018**	0.034

Table 2.2) Differences in synonymous nucleotide diversities (π_S) between island and mainland species. The number of comparisons used in each analysis is given in the second column (*n*). The mean relative value of island π_S is given in the fifth column, with relative values calculated as: (island π_S)/(island π_S + mainland π_S). Any undefined values were excluded from the analysis. CIs for the relative island values of π_S are given in the sixth and seventh columns. A one- tailed Wilcoxon signed-ranks test on the relative island values was conducted, with the alternative hypothesis that the true island value is less than 0.5. The p-value of this test is given in the eighth column, with any statistically significant results highlighted in bold. Spearman's coefficient of rank correlation between the ratio of island to mainland species ranges and the relative island π_S is given in the last column. None of these correlations are statistically significant.

It is potentially possible to differentiate between the two possible causes of the lower diversity in island species by considering the ratio of island to mainland nucleotide diversity as a function of the time of divergence between the island and mainland species. In this analysis we use the level of synonymous divergence between island and mainland species/clades (d_S) as an estimator of the time at which species diverged since we lack information on colonisation times. However it should be noted that this is a crude estimator of the divergence time since d_S is dependent on both the time of divergence and the mutation rate.

If most of the reduction in diversity is due to a bottleneck during colonisation, then we expect the difference in island to mainland diversity to be greatest when the evolutionary divergence is shortest. In contrast, if diversity is largely determined by population sizes after colonisation then we might expect the ratio of island to mainland diversity to decline with evolutionary divergence. Consistent with the bottleneck hypothesis, we find that relative island synonymous diversity, π_S(island)/(π_S(island)+π_S(mainland)), a measure of the island diversity relative to mainland diversity, which is defined for all informative comparisons, is positively correlated to the synonymous divergence between island and mainland species across our combined dataset (Pearson's correlation r= 0.318, p= 0.012) (Figure 2.2). The correlation increases in strength if we restrict the analysis to mainland-to-island colonisation events (r = 0.384, p=0.004), and is negative, though non-significant, if we consider colonisations that occurred in the opposite direction (r = -0.129, p = 0.74). However, as there are few comparisons available in which the direction of colonisation is inferred to be island to mainland, we probably lack power to detect any significant trends in this group (see figure 2.2). The positive correlation that we have found appears to be driven by a group of island species/clades that are closely related to their mainland relatives, and are therefore likely to be recent colonists, and have no synonymous diversity (Figure 2.2), because the positive correlation disappears when species with no synonymous diversity are removed from the analysis (r = 0.214, p = 0.150). Although the low levels of diversity we have recorded could be a result of low levels of mutation and/or short sequences, this explanation is unlikely because we would expect equal numbers of island and mainland species to have low diversity (i.e. in figure 2.2 we would expect an equal number of points clustering at 1 on the y-axis as at 0), which is not what we observe.

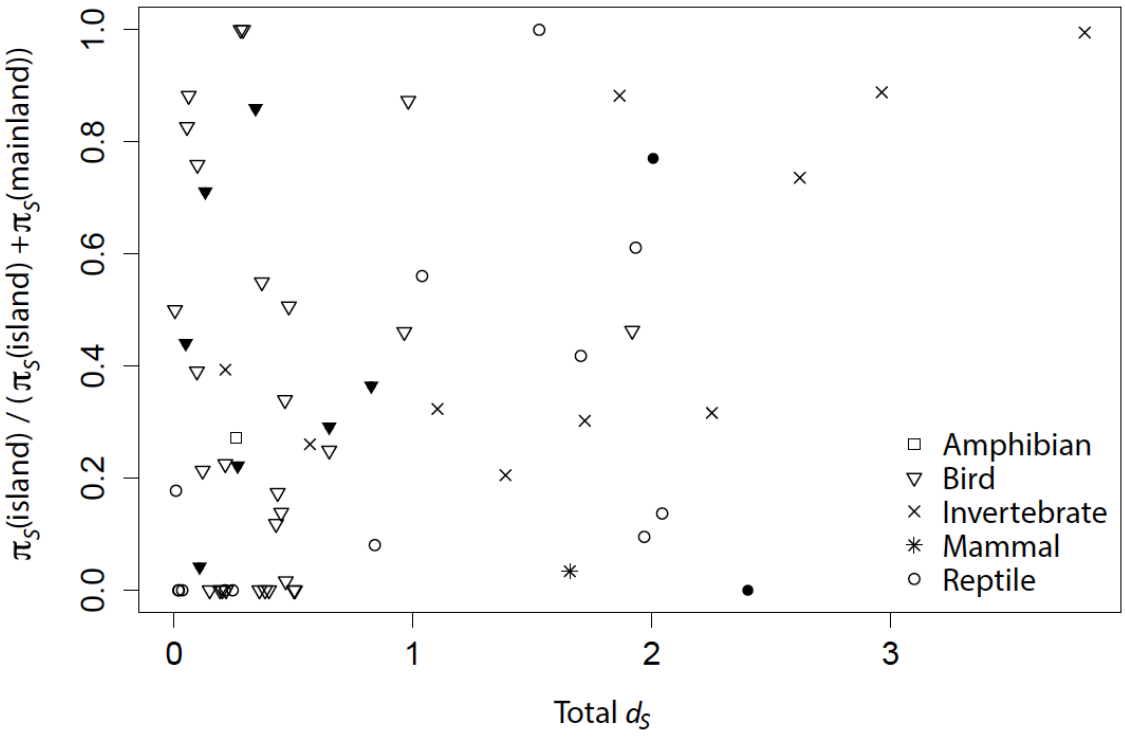

Figure 2.2) The ratio of island diversity to the combined island and mainland diversity, π_S(island)/(π_S(island)+ π_S(mainland)), where π_S is synonymous diversity, plotted against total divergence (d_S) between island and mainland species. Filled shapes indicate comparisons in which the inferred direction of colonisation is island to mainland.

Reptiles are disproportionately represented amongst the species with no genetic diversity in the island species/clades (6 out of 14 reptiles compared to 9 out of 35 birds and 0 out of 10 invertebrates). If each phylogenetic group is considered individually we find a significant positive correlation between relative island diversity and d_S for invertebrates (r = 0.752, p = 0.012) and positive but non-significant correlations for birds and reptiles (Figure 2.2) (we do not have enough data to study the other groups individually). As a group, birds appear to retain the highest levels of diversity, with some species seemingly not undergoing a population bottleneck during the colonisation event, perhaps because there are more individuals initially founding the island population and/or because there is continued migration from the mainland. This is compatible with the greater dispersal ability of birds compared to other animal groups. Reptiles on the other hand appear to experience a quite severe loss of diversity during founder events.

Although our results are consistent with the idea that the genetic diversity of island species is able to recover over time, either through continued immigration or the accumulation of new genetic diversity *in situ*, an alternative interpretation is that island species that are not diverse simply go extinct. This may be why only young species have low levels of diversity (out of 62 comparisons, only the chameleon *Archaius tigris* was moderately divergent without any synonymous diversity at all). These explanations are not necessarily mutually exclusive. Nevertheless it is surprising that aside from those species with no

synonymous diversity, in most cases island species have similar and in some cases more genetic diversity than their mainland counterparts. If we remove the comparisons in which island diversity is zero and re-analyse the data we find that the remaining island species do not have lower synonymous diversity than mainland species (Wilcoxon signed rank test, n = 48, p = 0.32). This suggests that island species/clades only have lower levels of diversity if they have recently (in terms of generations) undergone a population bottleneck.

However, it is possible that we do not observe a correlation between synonymous divergence and neutral diversity because for some of our comparisons the total level of synonymous divergence is very high. This is most likely due to the high mutation rates of mitochondria. As such, it is possible that some populations may have reached equilibrium, and therefore we would not expect to see any remaining effects of a bottleneck. This might obscure a relationship between neutral divergence and neutral diversity over shorter time scales. To account for this possibility, we reanlaysed the data, only including those comparisons for which values of total d_S were below a certain cut-off value (we used the values 1, 0.5 and 0.3). This does not increase the strength of the correlation, and in fact none of the correlations that employed d_S cut-off values were significant. Again, this supports the idea that colonisation events have surprisingly little impact on the genetic diversity of island species.

2.4.4 Effective population sizes

The fact that the genetic diversity of island species is generally not lower than that of mainland species suggests that they do not have lower effective population sizes. To investigate this, we estimated N_e by dividing π_S by d_S (using d_S, synonymous divergence, to approximate the mutation rate) and compared island species to their mainland counterparts. Note that these effective population size estimates can only be compared against each other (i.e. within each island-mainland comparison), since in effect we are dividing the diversity by the product of the mutation rate per generation and the number of generations since the mainland and island species diverged. Mainland species had significantly greater effective population sizes than island species overall (Wilcoxon signed-ranks test, n = 66, p = 0.030); however, the differences are small; on average we estimate island species to have an effective population size that is 69% that of mainland species (95% CIs: 51%, 89%). If we exclude those comparisons in which the island species had no synonymous diversity, the difference between island species and mainland species is no longer significant (n = 48, p = 0.566).

2.4.5 Efficiency of selection

Selection is expected to be less efficient in species with small N_e. However, we have found little evidence to suggest that island species have lower long-term effective population sizes than mainland species. It is therefore perhaps not surprising that we find little evidence for selection being less efficient in island species. Using polymorphism data we compared $\pi_N/(\pi_N+\pi_S)$ between island and mainland species and found that island species did not have significantly larger values of $\pi_N/(\pi_N+\pi_S)$ (Wilcoxon signed-ranks test, n = 48, p = 0.54). We also found no difference when considering different DNA types separately, or when considering different taxonomic groups separately (Table 2.3). We also find no correlation between

the relative island value of $\pi_N/(\pi_N+\pi_S)$ and the ratio of island and mainland range sizes ($r = -0.16$, $p = 0.38$). It should be noted however that most of the island species that have no synonymous polymorphisms also have no non-synonymous polymorphisms and hence are excluded from the analysis because $\pi_N/(\pi_N+\pi_S)$ is undefined.

Dataset	*n*	Mean Island $\pi_N/(\pi_N+\pi_S)$	Mean Mainland $\pi_N/(\pi_N+\pi_S)$	Mean relative island $\pi_N/(\pi_N+\pi_S)$	Lower CI	Upper CI	Wilcoxon p-value
Combined	48	0.18	0.093	0.50	0.40	0.60	0.54
Chloroplast	1	0.26	0.22	0.54	-	-	-
Mitochondrial	44	0.17	0.092	0.50	0.40	0.60	0.51
Nuclear	3	0.18	0.13	0.39	0	0.68	0.75
Bird	28	0.27	0.10	0.54	0.40	0.67	0.32
Invertebrate	10	0.035	0.055	0.54	0.33	0.73	0.36
Reptile (non-avian)	7	0.027	0.095	0.32	0.10	0.59	0.88

Table 2.3) Differences in $\pi_N/(\pi_N+\pi_S)$ between island and mainland species. The number of comparisons used in each analysis is given in the second column (*n*). The mean relative values of island $\pi_N/(\pi_N+\pi_S)$ is given in the fifth column, with relative values calculated as: (island $\pi_N/(\pi_N+\pi_S)$)/ (island $\pi_N/(\pi_N+\pi_S)$ + mainland $\pi_N/(\pi_N+\pi_S)$). Any undefined values were excluded from the analysis. CIs for the relative island value of $\pi_N/(\pi_N+\pi_S)$ are given in the sixth and seventh columns. A one- tailed Wilcoxon signed-ranks test on the relative island values was conducted, with the alternative hypothesis that the true island value is greater than 0.5. Statistically significant results are highlighted in bold.

We also find no significant differences between island and mainland species for ω (nonsynonymous divided by synonymous divergence) overall, or if we split the data by phylogenetic group or genome type (Table 2.4). However, there is an expectation that ω will increase during a population size expansion (Charlesworth & Eyre-Walker 2007; Takano-Shimizu 1999) and so we might expect island-to-mainland colonisations to show different patterns to mainland-to-island colonisations. If we restrict our analysis to mainland-to-island colonisations we still do not observe a significant difference between island and mainland ω overall, or for each genome, although if we split by phylogenetic group the result for birds is close to being statistically significant (Table 2.4). We also do not observe any significant difference in ω(mainland) / ω(island) between species that have colonised the island from the mainland, and the mainland from the island (independent samples t-test, $p = 0.315$), contrary to the results of Charlesworth and Eyre-Walker (2007). We find no correlation between ω(mainland) / ω(island) and the ratio of island and mainland range sizes ($r = -0.031$, $p = 0.77$).

Dataset	*n*	Mean Island ω	Mean Mainland ω	Mean relative island ω	Lower CI	Upper CI	Wilcoxon p-value
Combined	112	0.10	0.087	0.53	0.50	0.57	0.20
Chloroplast	10	0.34	0.16	0.70	0.57	0.83	0.11
Mitochondrial	96	0.042	0.051	0.52	0.49	0.57	0.38
Nuclear	32	0.37	0.24	0.52	0.43	0.62	0.68
Bird	60	0.083	0.062	0.54	0.50	0.59	0.17
Invertebrate	15	0.059	0.028	0.54	0.41	0.66	0.85
Plant	12	0.31	0.17	0.66	0.53	0.76	0.18
Reptile (non-avian)	21	0.092	0.11	0.50	0.41	0.59	0.76

Table 2.4a)

Dataset	*n*	Mean Island ω	Mean Mainland ω	Mean relative island ω	Lower CI	Upper CI	Wilcoxon p-value
I→M	14	0.16	0.19	0.45	0.34	0.57	0.50
M→I	98	0.095	0.071	0.54	0.49	0.60	0.11
M→I Chloroplast	9	0.26	0.15	0.69	0.49	0.87	0.20
M→I Mitochondrial	84	0.040	0.035	0.54	0.47	0.60	0.20
M→I Nuclear	29	0.39	0.23	0.53	0.39	0.66	0.62
M→I Bird	51	0.088	0.044	0.56	0.50	0.62	0.058
M→I Invertebrate	15	0.059	0.028	0.54	0.36	0.71	0.85
M→I Plant	11	0.24	0.16	0.64	0.48	0.80	0.32
M→I Reptile (non-avian)	17	0.069	0.073	0.51	0.37	0.64	0.85

Table 2.4b)

Table 2.4a and 2.4b) Differences in ω between island and mainland species. The number of comparisons used in each analysis is given in the second column (*n*). The mean relative value of island ω is given in the fifth column, with relative values calculated as: (island ω)/(island ω + mainland ω). CIs for the relative island values of ω are given in the sixth and seventh columns. A two- tailed Wilcoxon signed-ranks test on the relative island values was conducted, to test whether the distribution of island values was significantly different from symmetrical about 0.5. Statistically significant results are highlighted in bold. In a), the total dataset is analysed and then divided by DNA type and taxonomic group, while in b), the comparisons are split by colonisation direction; I→M refers to comparisons in which the colonisation direction was island-to-mainland, while M→I is mainland.

2.4.6 Adaptive evolution

Given that there seems to be little difference in N_e between island and mainland species we might expect colonisation of an island to lead to a burst of adaptive evolution, since the colonisers are experiencing a new environment that might have empty niches into which the species can adaptively evolve (this effect might have been reduced or eliminated if island species had lower N_e and rates of adaptation were mutation limited). To investigate whether colonisation leads to higher rates of adaptive evolution we estimated the rate of adaptive amino acid substitution along the island and mainland lineages using two approaches. First we calculated the direction of selection (DoS) statistic for each lineage. We find that on average DoS is negative in both island and mainland species (Table 2.5), indicating that slightly deleterious mutations are prevalent in our data. We find no significant difference in values of DoS between island and mainland species, either when considering the dataset as a whole, or when the results are analysed separately depending on the direction of colonisation. However, DoS is sensitive to slightly deleterious mutations segregating in the population, and therefore any changes in the relative frequencies of deleterious mutations between island and mainland species will influence DoS, potentially masking a signal of adaptive evolution (Nielson 2005). Unfortunately, we did not have sufficient polymorphism data to correct for slightly deleterious mutations by removing low frequency polymorphisms (Charlesworth & Eyre-Walker 2008; Fay et al. 2001) or applying more sophisticated methods that use the site frequency spectrum to estimate the distribution of fitness effects.

Dataset	*n*	Mean Island DoS	Mean Mainland DoS	p-value
Combined	50	-0.090	-0.056	0.619
I → M	8	-0.053	-0.020	0.401
M → I	42	-0.106	-0.063	0.827

Table 2.5) Differences in DoS between island and mainland species, for the combined dataset, and for the dataset split by the direction of colonisation. The number of comparisons used in each analysis is given in the second column (*n*), with the significance level of the Wilcoxon signed-ranks test given in the last column. I→M refers to comparisons in which the colonisation direction was island-to-mainland, while M→I is mainland-to-island.

2.4.7 Mutation rate

We also investigated potential differences in the mutation rates of island and mainland species. In this study we inferred the mutation rate from d_S, the number of synonymous substitutions, along the lineages leading to the mainland and island species (and where there were multiple island and mainland species, from their averages). N_e is predicted to influence mutation rate, and as we found no consistent differences in N_e between island and mainland species we do not expect mutation rate to differ between the two groups. This is in fact the case: comparing d_S values between island and mainland species revealed no significant difference (Table 2.6) (n = 111, p = 0.45). However, when different genomes were considered

separately, there was one statistically significant difference between island and mainland species for nuclear DNA (n = 30, p = 0.01). The trend in this instance was for mainland species to have higher values of d_S than island species.

Dataset	*n*	Mean Island d_S	Mean Mainland d_S	Mean relative island d_S	Lower CI	Upper CI	Wilcoxon p-value
Combined	111	0.35	1.15	0.49	0.45	0.53	0.45
Chloroplast	10	0.016	0.013	0.56	0.37	0.76	0.72
Mitochondrial	96	0.56	1.42	0.49	0.45	0.53	0.75
Nuclear	30	0.058	0.16	0.40	0.31	0.50	**0.010**

Table 2.6) Differences in d_S between island and mainland species. The number of comparisons used in each analysis is given in the second column (*n*). The mean relative value of island d_S is given in the fifth column, with relative values calculated as: (island d_S)/(island d_S + mainland d_S). CIs for the relative island values of d_S are given in the sixth and seventh columns. A two- tailed Wilcoxon signed-ranks test on the relative island values was conducted, to test whether the distribution of island values was significantly different from symmetrical about 0.5. Statistically significant results are highlighted in bold.

2.5 Discussion

It is generally assumed that island species will have smaller effective population sizes than mainland species. Island species are expected to have low effective population sizes initially because they are likely to have been founded by a small number of individuals (one pregnant female is sufficient) and hence experience a bottleneck. We find some evidence for this: some island species, which are very closely related to their mainland counterparts, have little or no diversity, consistent with these species experiencing extreme bottlenecks during colonisation. However, besides these species, island species have similar levels of diversity to mainland species. There is no evidence to suggest that island species have low long-term effective populations sizes, despite the fact that island species occupy considerably smaller ranges than mainland species; in this analysis, island species had ranges of on average 0.14% of the area of their mainland counterparts. Consistent with island and mainland species having similar effective population sizes, we find no evidence that natural selection is less efficient in island species.

Although our dataset consists of a large number of island and mainland comparisons, for most of our comparisons we have a single gene, and hence little data. This may result in noisy estimates of π_S and π_N and a lack of power in our analysis: this is particularly true of our estimates of π_N and the efficiency of selection, simply because nonsynonymous polymorphisms are rarer. It is therefore important to consider whether we are likely to be able to detect differences between island and mainland species if they exist.

The fact that we observe a significance difference in diversity between island and mainland species suggests that we do have the ability in this analysis (Table 2.2). However, the lower 95% confidence interval indicates that island species have at least 41% of the diversity of mainland species, far larger than the ratio of the ranges (0.14%). Even accounting for possible noise in our estimates of π_S, our data suggests that while there may be a difference in island and mainland species diversity, it is not of the same magnitude as we would expect from the difference in their census population sizes. Furthermore the difference in mainland and island diversity seems to be due to a few young island species with no diversity; if these are excluded island species have on average 92% of the diversity of mainland species (CIs: 65%, 128%). For our measures of the effectiveness of selection the upper 95% CIs suggest that $\pi_N/(\pi_N+\pi_S)$ could be up to 50% larger in island than mainland species, and that ω could be up to 33% larger. To put these numbers into context, ω is approximately 90% larger in primates than rodents (Eyre-Walker et al. 2002) and the ratio of nonsynonymous to synonymous polymorphisms, P_N/P_S, varies by almost 8-fold in plants (Gossmann et al. 2010) so the differences between island and mainland species are modest. The differences are also consistent with very moderate differences in N_e. For example, if we assume that all mutations are deleterious (though some can be effectively neutral) and the distribution of fitness effects is a gamma distribution, then the ratio of the ω values from two species with effective population sizes of N_1 and N_2 is expected to be $\omega_1/\omega_2 = \left(N_1/N_2\right)^{-\beta}$ (Welch et al. 2008), where β is the shape parameter of the gamma distribution. Analyses of both nuclear (Boyko et al. 2008; Eyre-Walker et al. 2006; Keightley & Eyre-Walker 2007, Gossmann et al. (unpublished results)) and mitochondrial data (James, Piganeau, et al. 2016), suggest that $\beta < 0.5$ in most species. If we conservatively assume that $\beta = 0.5$, a 1.33 ratio of island to mainland ω translates into a ratio of N_e values of 0.57. In other words, if island species had values of N_e that were below half that of their mainland relatives we ought to be able to detect a difference between them from the measures of the effectiveness of selection that we have used, and given the amount of data that we have.

Our results are perhaps not surprising. It is well established that the relationship between population size and genetic diversity is not straightforward, with levels of genetic diversity remaining remarkably constant across groups of organisms which are incredibly disparate in terms of census population size (Gillespie & Ohta 1996; Leffler et al. 2012; Lewontin 1974; Bazin et al. 2007). What is unique about the current data is that only closely related species are compared to each other- many of the island and mainland species pairs are in the same genus. They are therefore likely to share life history traits, many of which influence molecular evolution. In addition, our paired study design allows us to correct for phylogenetic effects (Lanfear et al. 2010). This is crucial, as it has been well demonstrated that molecular evolution is influenced by taxonomy. For example, Romiguier et al. (2014) demonstrated that levels of diversity differ between families but are similar within a family. Correcting for phylogenetic effects has allowed us to study the effects of island colonisation on molecular evolution across a wide range of taxa.

There are a number of possible explanations for our results. It is possible that island species do not have lower effective population sizes than their mainland counterparts: if island species are commonly founded by multiple individuals, and if gene flow is maintained throughout speciation, island species might inherit

much of the variation of the mainland species. We have evidence that this is true of some species: birds in particular appear to experience relatively few bottlenecks as a taxonomic group, which is probably due to their increased dispersal ability relative to other animals. However, after the initial colonisation event we might expect a reduction in the genetic diversity of island species over time, considering their restricted ranges. It is surprising that we see no evidence of this: even if we exclude those young island species with no diversity, the correlation between synonymous nucleotide diversity and synonymous divergence remains positive, but not significant (r = 0.214, p = 0.150). In addition, introgression is an unlikely explanation for the comparable neutral diversity of island and mainland species, because in the event of high levels of introgression we would expect the amount of synonymous diversity to be similar to that of synonymous divergence, assuming introgression is between the species being considered, rather than some other species we have not considered. In our analysis the majority of species have considerably higher levels of d_S than π_S (Figure 2.3) with island π_S being on average just 6% of the d_S between island and mainland species. This indicates that most of the island and mainland species pairs are diverging: losing shared polymorphisms and accumulating substitutions. This pattern is not expected if there is extensive gene flow. However, it might be that there is introgression into the island from another species we have not surveyed. This is difficult to rule out.

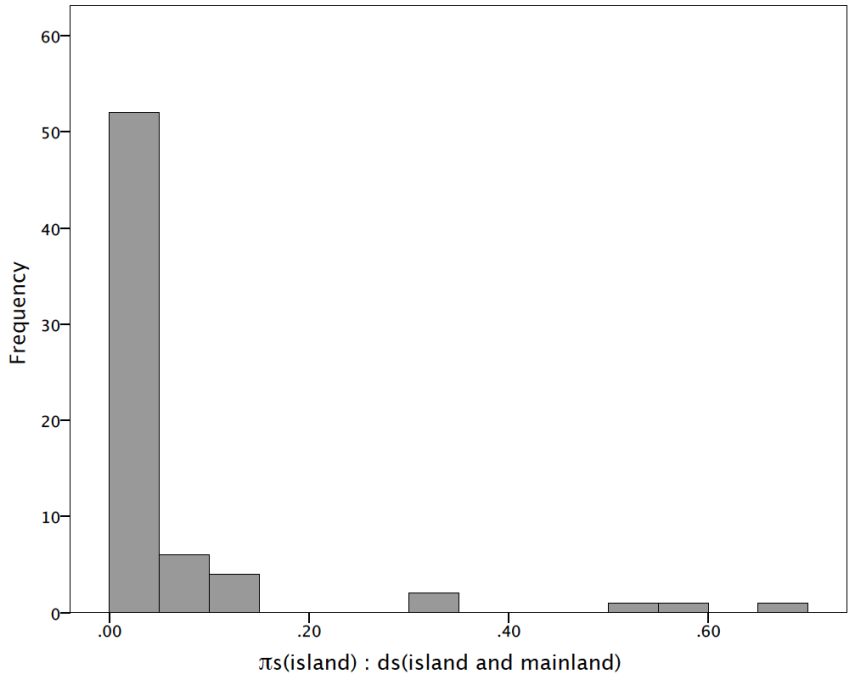

Figure 2.3) Frequency distribution of the ratio of island π_S : island and mainland d_S

There are also a number of factors that might obscure a relationship between effective population size and genetic diversity, which could explain our results. Firstly, it has been suggested that levels of diversity are relatively constant across species because of an inverse relationship between population size and the mutation rate per generation (Lynch 2007; Piganeau & Eyre-Walker 2009), a relationship for which there is some evidence (Lynch 2010; Sung et al. 2012). This is hypothesised to occur because populations with large effective population sizes can more effectively select for modifiers of the mutation rate. Therefore,

selection to reduce the mutation rate will be more effective in larger populations, resulting in lower mutation rates and hence levels of genetic diversity similar to those found in small populations. There is no evidence that this is the case in this analysis. When we analysed the levels of synonymous divergence, an indicator of the neutral mutation rate, we did not find a difference between island and mainland species, indicating that island species do not have higher mutation rates. In addition, there is no evidence, from considering the efficiency of selection, that island species have lower effective population sizes. This is perhaps not surprising, since the mutation rate is expected to increase when the effective population size is reduced, but only slowly. Finally, upon excluding those species with no diversity we do not find that diversity increases with divergence, which we might expect if higher mutation rates evolve over time in island species.

Secondly, it is also possible that there is selection on synonymous mutations, which might also obscure a relationship between genetic diversity and effective population size. If selection on synonymous codon use varied between sites and was directional we would find that as N_e increases, the proportion of effectively neutral mutations would decrease as selection becomes more efficient. If the distribution of fitness effects of synonymous mutations was exponential one would have a situation in which the increase in N_e was perfectly matched by a decrease in the proportion of mutations that were effectively neutral (Ohta 1992). However, there is no evidence that there is selection on synonymous codon usage in animal mitochondria (Jia & Higgs 2008). Furthermore, it has been suggested that selection on synonymous codon use is stabilising in nature, at least where the synonyms match different tRNAs (Qian et al. 2012), and under such a model we might expect the strength of selection, in terms of N_es to remain relatively constant (Charlesworth 2013)

Finally, it is also possible that the relationship between genetic diversity and the efficiency of selection is not straightforward due to selection at linked sites (Gillespie 2000; Maynard Smith & Haigh 1974). Gillespie has argued that if the rate of adaptive evolution is mutation limited then as population sizes increase so does the rate of adaptive evolution and hence the level of genetic hitch-hiking – a phenomenon that he has termed genetic draft. Some authors have found evidence to suggest that draft has an important role in reducing genetic diversity (Bazin et al. 2006; Corbett-Detig et al. 2015). However, studies generally report that draft has relatively weak effects which may not be powerful enough to reduce genetic diversity to observed levels, particularly in nuclear DNA (Andolfatto 2007; Corbett-Detig et al. 2015; Gossmann et al. 2011; Weissman & Barton 2012); the most extensive analysis of draft in the nuclear genome has shown that draft at most reduces diversity by 73% in a survey 40 eukaryotic species (Corbett-Detig et al 2015). Furthermore, there is no evidence in our data that draft is important. Firstly, if genetic draft was prevalent in our dataset we might expect different patterns for the organellar genomes, which have little or no recombination, and the nuclear genome (Campos et al. 2014). However, they behave in a qualitatively similar fashion between island and mainland species (for example, see Table 2.2 and 2.3). Secondly, we do not find a significant difference between island and mainland species in terms of their DoS. If selective sweeps were responsible for the low diversity of mainland species, we might expect mainland species to have greater values of DoS than their island counterparts. In addition, our

results indicate that it is deleterious mutations that are dominating evolutionary dynamics, rather than advantageous mutations. However, it is worth noting that the signal of adaptive evolution could be obscured by a shift in the distribution of fitness effects for island species. Correcting for this with the current dataset is difficult due to a lack of sufficient polymorphism data, although the results from our limited sample indicate that it is island species that undergo a greater degree of adaptive evolution, rather than species with large population sizes.

Romiguier et al. (2014) recently showed that geographic factors likely to influence population size are poor correlates of genetic diversity when diversity is considered across the full breadth of the animal kingdom. Surprisingly, they find that propagule size is the single best predictor of diversity. Those species with few large propagules had low genetic diversity, and those with a large number of small propagules had high genetic diversity, and were termed K and r strategists respectively. They suggest that K strategists might be able to maintain smaller population sizes because they invest substantially in their offspring, whereas r-strategists have to maintain large population sizes on average because they are more prone to population crashes. An alternative hypothesis is that propagule size is related to population density, and that the variance in population density is far greater than the variance in population range size, so that the degree to which species differ in effective and census population sizes is largely determined by density and not range size. However, this would only explain our results if population density was on average much higher on the islands than the mainland.

Alternatively, it may be that the mutation rate itself is an important determinant of diversity, particularly in organellar genomes (Bazin et al. 2006; Lynch et al. 2006; Nabholz, Mauffrey, et al. 2008). Although the issue is controversial, Nabholz et al. (2008) showed that the mutation rate is a major determinant of mitochondrial diversity, and as our dataset is dominated by mitochondrial sequences this could explain why we did not find a difference between island and mainland species, considering that we also did not find a difference in mutation rate between them. We found a positive correlation between the mutation rate, as measured by the rate of synonymous divergence, and levels of synonymous diversity, both for our entire dataset ($n = 138$, $r = 0.337$, $p < 0.001$), and considering mitochondrial sequences separately ($n = 112$, $r = 0.269$, $p = 0.004$), which lends some support to this theory, however, we are unable to recover this correlation if we correct for phylogenetic independence by comparing island and mainland species (i.e. π_S(island)/(π_S(island)+π_S(mainland)) is not significantly correlated to d_S(island) / (d_S(island) +d_S(mainland)).

In conclusion, our analysis demonstrates that island colonisation has had little impact on the molecular evolution of species in this dataset. For some species the initial colonisation event results in a period of low diversity, but this effect appears to be short-lived with no discernible lasting effects. Our results confirm that census population size is a poor correlate of effective population size**.**

Chapter 3

Investigating the life history and demographic traits that are predictive of N_e

3.1 Abstract

The role of life history and demographic traits in determining molecular evolution has received a lot of attention in recent years, however, past studies have tended to focus on very broad taxonomic scales. We focus on factors that are predictive of neutral genetic diversity, π_S, and the efficiency of selection, $\pi_N/(\pi_N+\pi_S)$ in a single animal class, the mammals, using a phylogenetically controlled dataset. For the first time, we find evidence to suggest that the range of a species is significantly related to both π_S and $\pi_N/(\pi_N+\pi_S)$ in mammalian mitochondria. We also find that both the average latitude at which a species is found and its body mass are predictive of mitochondrial neutral diversity. These results support the nearly neutral theory of molecular evolution. However, we do not find any evidence for a relationship between life history traits and molecular evolution in nuclear DNA, which may be due to lack of power in our dataset.

3.2 Introduction

The factors that govern molecular evolution are complex, and despite considerable interest from evolutionary geneticists, remain somewhat unclear. One of the most important variables determining the molecular evolution of a species is N_e, the effective population size, which governs the extent to which populations are affected by genetic drift: populations with a large N_e are only weakly affected by genetic drift, and are expected to have high levels of neutral genetic diversity and more efficient selection compared to populations with a small N_e (Charlesworth 2009; Lanfear et al. 2014; Ohta 1992). Unfortunately, N_e has been estimated in relatively few species. There are three principle approaches that have been used to estimate this parameter: dividing nucleotide diversity at putatively neutral sites by an estimate of the mutation rate per generation; estimating N_e from the fluctuation in allele frequencies using temporally sampled data; and estimating N_e from average levels of linkage disequilibrium. It is important to appreciate that these methods estimate N_e over very different time scales; when using nucleotide diversity, one is estimating N_e over a timescale of N_e generations, whereas the other two methods estimate N_e over a matter of just a few generations. Waples (2016) has shown that the temporal and linkage disequilibrium methods are subject to both large biases and variances; they tend to either grossly underestimate the effective population size, or estimate it to be infinite. However, estimating N_e using

nucleotide diversity is also problematic because it requires an estimate of the mutation rate per generation, which is known for relatively few species (Lynch et al. 2016). We are therefore left with estimating N_e from a variety of surrogate measures. Two that have been widely used in the molecular evolution literature are 1) nucleotide diversity at putatively neutral sites, and 2) measures of the efficiency of selection – either the rate of non-synonymous over the rate of synonymous substitution, or the ratio of nonsynonymous to synonymous nucleotide site diversity. Considering factors that correlate with these proxies should help us to uncover variables that are predictive of N_e.

Measures of neutral nucleotide diversity (henceforth π_S) are commonly used proxies for N_e, and a number of studies have been conducted in order to uncover factors that predict π_S. For example, recently Romiguier et al. (2014) found significant relationships between neutral nuclear diversity and a number of life history traits, including body mass, maximum longevity, fecundity, adult dispersal ability and propagule size. The relationships were as expected if π_S functions as a proxy for N_e. For example, large species, which are known to have small census population sizes relative to small species (White et al. 2007) and thus low effective population sizes, have lower values of π_S. However, the authors did not find any relationship between π_S and geographic variables that could indicate species abundance, such as distance between GPS records (Romiguier et al. 2014). More generally, the relationship between π_S and geographic traits tends to be weak. This may reflect the difficulty in choosing geographic traits that accurately reflect species abundance; while there are a number of species for which we have an estimate of global range, important parameters such as population density are very hard to measure and as such are rarely known. Therefore, geographic traits may be poor indicators of both census population sizes, and thus also effective population sizes.

However, it is important to note that any relationship between π_S and life history could in principle be driven by differences in the mutation rate, rather than N_e. This may explain some surprising results in mitochondrial DNA. Unlike in nuclear DNA, there are no obvious relationships between mitochondrial neutral diversity and life history traits in either mammals (Nabholz et al. 2008), or birds (Nabholz et al. 2009). It is possible that variation in the mitochondrial mutation rate in these groups could be the primary influence on variation in mitochondrial diversity. On the other hand, there is evidence that mitochondrial π_S is influenced by N_e: for example, it has been found that in mammals (Mulligan et al. 2006) and other animals (Piganeau & Eyre-Walker 2009) that mtDNA diversity and nuclear diversity, measured as allozyme heterozygosity, are correlated, suggesting that mitochondrial diversity does vary across animal groups. Although we would generally expect the forces that govern the molecular evolution of mitochondrial and nuclear DNA to be similar, because mitochondrial mutation rates are higher than nuclear rates, perhaps mutation rate variation has a greater impact on mitochondrial evolution than it does on nuclear evolution.

Measures of the efficiency of selection, such as variants of π_N/π_S or d_N/d_S, are normalised by synonymous rates and so should not be confounded by variation in the mutation rate. Using the measure d_N/d_S, Popadin et al. (2007) found that the efficiency of purifying selection on mitochondrial DNA is greater in small

mammals relative to large mammals. This indicates that N_e and life history traits do have an effect on mitochondrial molecular evolution. There are also significant positive relationships between nuclear d_N/d_S and mass, longevity and sexual maturity in non-avian amniotes (Figuet et al. 2016), in agreement with the results of studies that have measured π_S. However, interpreting d_N/d_S results is not straightforward, as a large d_N/d_S value could indicate either weak purifying selection, such that there are lots of deleterious nonsynonymous substitutions fixed in the population, or the role of adaptive evolution, if the population has fixed lots of advantageous nonsynonymous substitutions. Measures of the efficiency of selection that use polymorphism data, such as π_N/π_S, are easier to interpret: polymorphisms are not expected to include significant numbers of beneficial mutations, and so are more straightforward as indicators of the efficiency of selection. Currently there are relatively few analyses that consider the effects of life history on π_N/π_S: Chen et al. (2017) find a positive correlation between π_0/π_4 (a similar, albeit more conservative, measure to π_N/π_S) and longevity in both plants and animals, while Figuet et al. (2016) find a positive relationships between π_N/π_S and mass, longevity and age at sexual maturity in birds. Overall, large, long-lived, less fecund animals appear to have less efficient selection, just as we predict from their presumably smaller effective population sizes. However, this has yet to be shown in mitochondrial DNA.

In this study, we explore how life history and demographic traits affect the molecular evolution of both mitochondrial and nuclear DNA, using a large dataset of mammalian species, in order to fill this gap in the current literature. We are also interested in considering the joint effects of traits on molecular evolution: many life history traits are known to be related to each other (Bromham 2011), and so quantifying the relative impact of traits on variation in π_S and the efficiency of selection is important to our understanding of any patterns. Importantly, in all of our analyses we use methods that adequately control for phylogenetic non-independence. Our dataset is relatively taxonomically restricted compared to a number of studies, allowing us to investigate whether a number of recent findings hold at a finer taxonomic scale.

3.3 Methods

Mitochondrial and nuclear coding DNA sequences were downloaded from Mampol, a database of mammalian polymorphisms (Egea et al. 2007). Only those species for which there was a minimum of four sequenced individuals were included in this study. Sequences were aligned by eye, using Geneious version 7.0.6 (Kearse et al. 2012). Where multiple genes were sequenced for a single species, sequences were concatenated to produce longer alignments, however, nuclear and mitochondrial DNA were treated separately. Alignments were then analysed using our own scripts, in order to calculate our two summary statistics, synonymous nucleotide site diversity, π_S, and the efficiency of selection, $\pi_N/(\pi_N+\pi_S)$. We used average nucleotide site diversity, rather than raw polymorphism counts, because our alignments were of different lengths and contained different numbers of individual sequences. Unlike raw polymorphism counts, average nucleotide site diversity measures should not be biased by these differences. In addition, we use the statistic $\pi_N/(\pi_N+\pi_S)$ rather than the more common π_N/π_S in this study because it is less biased,

and results in fewer species for which the value of the ratio is undefined (Stoletzki & Eyre-Walker 2011). We added life history and demographic information to the species in our dataset by using the Pantheria database (Jones et al. 2009). In total, we have mitochondrial polymorphism data for 639 species, and nuclear polymorphism data for 159 species for which we also have at least some life history and/or demographic information. We do not have all trait information for all species, and as such the number of species available for each analysis differs depending on the trait in question. The number of species comparisons available for each analysis, *n*, is given in our results tables. In terms of sequence data, in our mitochondrial dataset an average of 4.5 individuals was available per species, and the mean length of the alignments used was 1100 nucleotides. The majority of the mitochondrial sequences used in this analysis are sections of the cytochrome b gene, although the dataset also includes other proteins, such as cytochrome c oxidase (cox1, cox2 and cox3) and subunits of NADH dehydrogenase. All mitochondrial genes are involved in ATP production, and are highly conserved across mammals. In our nuclear DNA dataset, there are on average 4 sequenced individuals available per species, and the mean length of the alignments used in the analysis is 860 nucleotides. Unfortunately, the genes encoded by the nuclear sequences are quite variable across species in this analysis. All polymorphism estimates, life-history data and sequence alignments used in this analysis can be found at: https://dx.doi.org/10.6084/m9.figshare.3084205.v1

Species cannot be considered as statistically independent datapoints, due to shared ancestry. In order to quantify this effect in our dataset, we first calculated the degree of phylogenetic signal in the dataset using Pagel's λ (reviewed in (Freckleton et al. 2002; Kamilar & Cooper 2013)), using the R package phylosignal (Keck et al. 2016). In order to remove the effects of phylogenetic non-independence from our dataset, we used a method of paired independent contrasts (Felsenstein 1985). Life history and molecular evolution traits were first log transformed, and then phylogenetic contrasts were calculated using the ape package in R (Paradis et al. 2004). The phylogenetic trees used in this study were created using TimeTree (Hedges et al. 2006). All analyses were conducted in R. Three datasets were constructed for this study: one including species for which there was mitochondrial sequence data available, one including species for which there was nuclear sequence data available, and one that only includes species for which there is both mitochondrial and nuclear sequence data available.

3.4 Results

We have collected a dataset of mitochondrial and nuclear polymorphism data for a wide variety of mammal species for which we also have life history and demographic information. We have mitochondrial polymorphism data for 639 species, and nuclear polymorphism data for 159 species, however, our life history data is not complete for all of the species in the dataset (i.e. we do not have information on all of the traits for every species).

In keeping with the higher mutation rate of mitochondria , we found that synonymous mitochondrial diversity, π_S(mt), was significantly greater than synonymous nuclear diversity, π_S(n): mean values of π_S

were 0.044 and 0.0086 for mitochondrial and nuclear sequences respectively, Wilcoxon rank-sum test p << 0.001. However, we find that mitochondrial nonsynonymous diversity π_N is significantly lower than that of nuclear DNA (means are 0.0020 and 0.0071 respectively, Wilcoxon rank-sum test p << 0.001). Additionally, mitochondrial sequences have lower values of $\pi_N/(\pi_N+\pi_S)$ than nuclear DNA (mean values were 0.074 and 0.42, Wilcoxon rank sum p-values << 0.001), which suggests that mitochondrial DNA has more efficient purifying selection. Similar results have been obtained using substitution data (Havird & Sloan 2016; Popadin et al. 2013). These results also hold if we only consider those species for which we have both mitochondrial and nuclear DNA (n=113), and if the tests are conducted in a paired manner (Wilcoxon signed rank tests: p << 0.001 for both π_S and $\pi_N/(\pi_N+\pi_S)$.

3.4.1 Relationships between traits and mitochondrial molecular evolution

Our aim in this study is to identify trends between molecular evolutionary traits and life history/demographic traits. However, in mammals many traits show some degree of phylogenetic signal, such that species that are more closely related are more similar to each other. This is true of our dataset: all life history, demographic and molecular traits we consider are significantly different between Orders, and show significant phylogenetic signal, as measured using Pagel's λ (Table 3.1). A number of traits had λ values that were very close to 1, which is indicative of strong phylogenetic signal, and a model of evolution that is similar to a Brownian motion model. In order to control for these phylogenetic effects and to remove the statistical non-independence between species, we conducted our analyses using a method of phylogenetic comparisons (Felsenstein 1985).

Dataset	Mass	Sexmat	Longevity	Range	Latitude	π_S	$\pi_N/(\pi_N+\pi_S)$
Mitochondria	1.0 (<0.01)	0.95 (<0.01)	0.92 (<0.01)	0.64 (<0.01)	0.84 (<0.01)	0.43 (<0.01)	0.32 (<0.01)
Nuclear	1.0 (<0.01)	0.98 (<0.01)	0.99 (<0.01)	0.59 (<0.01)	0.63 (<0.01)	<<0.001 (1)	0.32 (0.29)

Table 3.1) Table of values of Pagel's λ for life history, demographic and molecular traits. p-values are shown in parentheses- these indicate whether Pagel's λ was significantly different from 0.

We find a number of significant relationships between life history and demographic traits and mitochondrial molecular evolution (Table 3.2). We present here results using both Pearson's and Spearman's tests, however, the results from both are in agreement, and as such our conclusions are not dependent on the test used. π_S(mt) is significantly positively correlated to the global range of species, while $\pi_N/(\pi_N+\pi_S)$(mt) is significantly negatively correlated to global range (Figure 3.1a and 3.1 b). We also found a significant relationship between π_S(mt) and the average latitude of a species global range, such that π_S(mt) decreases with increasing distance from the equator, although we do not find a

significant relationship between $\pi_N/(\pi_N+\pi_S)$(mt) and latitude. In addition, we find that age at sexual maturity significantly negatively correlates to π_S(mt), in agreement with Nabholz et al. (2008), however, we do not find a relationship between longevity and π_S(mt). We find a negative correlation between π_S(mt) and mass with Spearman's correlation coefficient, but not Pearson's, and surprisingly there is no correlation between mass and $\pi_N/(\pi_N+\pi_S)$(mt).

Trait (Log values)	**(Log values)**	***n***	**Pearson's**		**Spearman's**	
			R	**p**	**rho**	**p**
Mass	π_S	537	-0.040	0.35	-0.093	**0.031**
	$\pi_N/(\pi_N+\pi_S)$	469	0.038	0.42	0.040	0.38
Longevity	π_S	225	-0.050	0.46	0.011	0.87
	$\pi_N/(\pi_N+\pi_S)$	204	0.042	0.55	0.084	0.23
Sexual Maturity	π_S	238	-0.17	**0.0091**	-0.10	0.11
	$\pi_N/(\pi_N+\pi_S)$	219	0.010	0.88	0.066	0.33
Range	π_S	556	0.32	**<0.001**	0.25	**<0.001**
	$\pi_N/(\pi_N+\pi_S)$	480	-0.22	**<0.001**	-0.18	**<0.001**
Distance from equator	π_S	556	-0.11	**0.013**	-0.090	**0.035**
	$\pi_N/(\pi_N+\pi_S)$	480	0.056	0.22	0.061	0.18

Table 3.2) Results of correlation analyses for two molecular evolutionary traits in mitochondrial DNA: π_S and $\pi_N/(\pi_N+\pi_S)$, with life history and demographic traits. Values are log-transformed before phylogenetic contrasts are calculated. The column *n* gives the number of contrasts available for each correlation. Significant results are in bold.

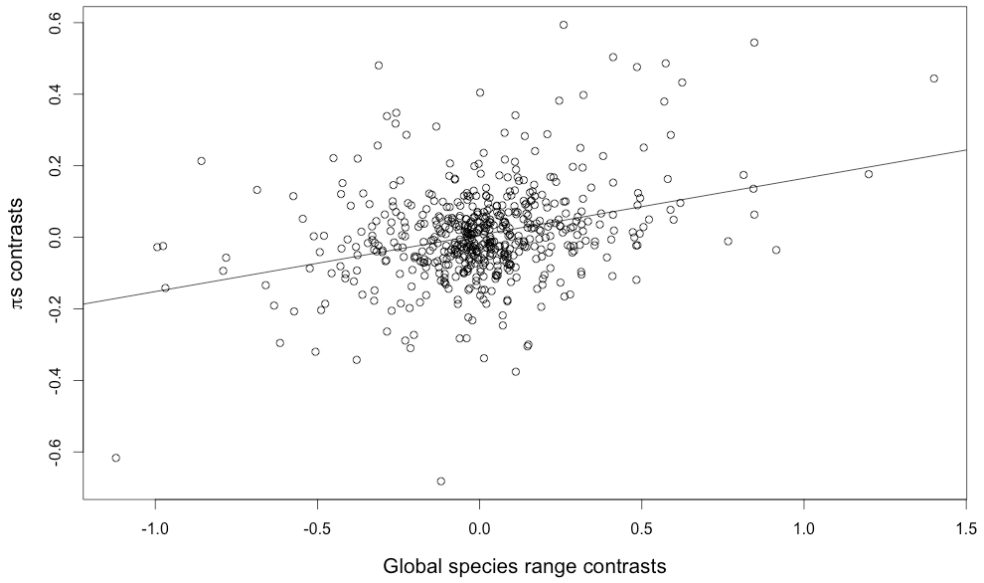

Figure 3.1a) The relationship between the global range of a species and π_S. Values plotted are log-transformed phylogenetic independent contrasts.

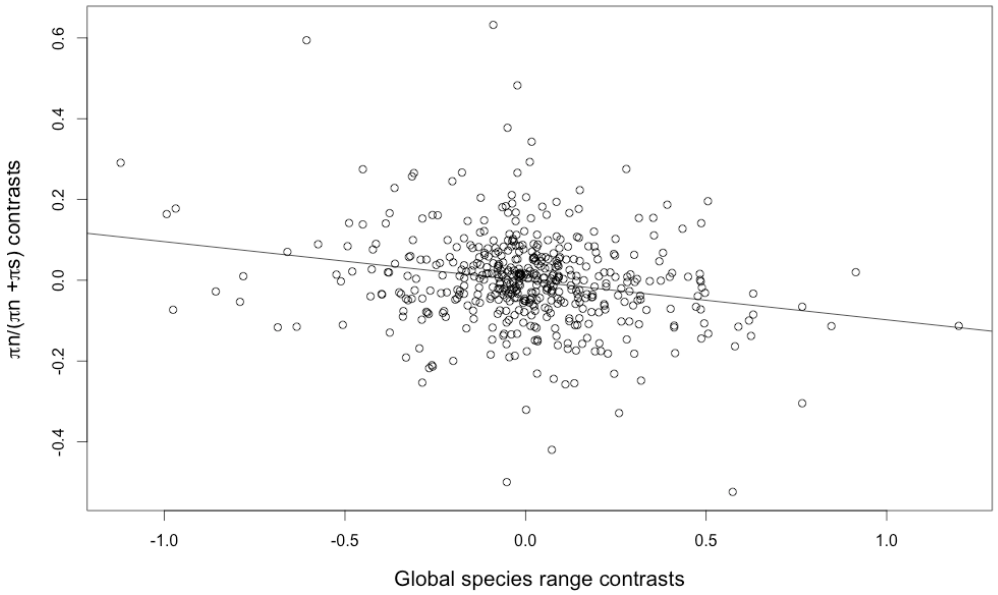

Fig 3.1b) The relationship between the global range of a species and $\pi_N/(\pi_N+\pi_S)$. Values plotted are log-transformed phylogenetic independent contrasts.

Our correlation results suggest that only range is a significant predictor of both π_S and $\pi_N/(\pi_N+\pi_S)$. However, as it is well established that many traits are related to each other, particularly in mammals, it is possible that our results could be driven by interactions with other variables, or that interactions between traits could obscure a relationship. To test this possibility, we consider the effects of multiple life history and demographic traits on molecular evolution jointly, using multiple linear regression models. We find that including interaction terms between life history and demographic traits does not increase the fit of our models for either π_S or $\pi_N/(\pi_N+\pi_S)$, and so we only present the results for models that do not include interaction terms.

We find that of our traits, only range and latitude are significantly predictive of π_S. Models that include sexual maturity and longevity were not a significantly better fit to the data than models which exclude these traits (number of contrasts = 154, Anova p = 0.55). In addition, excluding these traits resulted in an increase in the number of available contrasts for which we have the life history trait data of interest (n = 489). We found that while the effect of mass is not statistically significant at the $p < 0.05$ level, mass does appear to have a weak effect on π_S(mt), and a model including the effects of body mass, range and latitude on π_S was a marginally significantly better fit to the data than a model which excluded body mass (anova test p value = 0.086). Therefore for π_S the best fit model to our data, with an adjusted R-squared of 0.11 and a p-value of less than 0.001, includes the traits mass, range and latitude. The model results shown in Table 3.3. After standardising the regression coefficients, we find that range is the most important variable in determining π_S(mt). In contrast, only range has a significant effect on $\pi_N/(\pi_N+\pi_S)$ when we consider range, mass and latitude together (n = 424, adjusted R-squared value of 0.040 and a p-value 0.00015, Table 3.3), and is not a significantly better fit to the data than a model in which range is the only life history trait considered.

As has been previously noted, many traits are highly correlated in mammals, with implications for molecular evolution. In our dataset, longevity, sexual maturity, and even global species range are all positively correlated to mass, with only the average distance from the equator of a species' range not correlating to mass. However, despite these correlations, none of the variance inflation factors for the life history and demographic variables used in our models were above 1.1, suggesting that multicollinearity is not substantially affecting our model results.

Trait	π_S			$\pi_N/(\pi_N+\pi_S)$		
	Regression slope	**Sig.**	**Standardised regression coefficient**	**Regression slope**	**Sig.**	**Standardised regression coefficient**
Mass	-0.12	0.086	-0.074	0.088	0.16	0.067
Range	0.18	**<0.001**	0.32	-0.096	**<0.001**	-0.20
Distance from equator	-0.10	**0.038**	-0.091	0.046	0.37	0.044

Table 3.3) Results of multiple linear regression of life history and demographic traits on π_S and $\pi_N/(\pi_N+\pi_S)$. Values are log-transformed before phylogenetic contrasts are calculated. Regression slopes were forced through the origin.

3.4.2 Relationships between traits and nuclear molecular evolution

Again, all life history and demographic traits show significant phylogenetic signal in our nuclear dataset (Table 3.1). However, unlike mitochondrial DNA, there is no significant phylogenetic signal in nuclear molecular traits, and therefore phylogeny may be less predictive of nuclear molecular evolution than mitochondrial molecular evolution.

We find a negative correlation of π_S(n) with range which is only significant using Pearson's, and a negative relationship between $\pi_N/(\pi_N+\pi_S)$(n) and latitude which is only significant using Spearman's (Table 3.4). However, we do not find any other significant relationships between life history traits and π_S(n), or $\pi_N/(\pi_N+\pi_S)$(n). This may be due to lack of power in the nuclear dataset- nuclear DNA is generally far less diverse than mitochondrial DNA and we have far fewer species. To address the affect that species with low numbers of polymorphisms may have on our dataset, we repeated our analyses on a restricted dataset, consisting of only those species in which the number of polymorphisms was greater than 8, the median number of polymorphisms for species in the nuclear dataset. However, this did not greatly affect our results (results not shown).

Trait	(Log values)	*n*	Pearson's		Spearman's	
			R	**p**	**rho**	**p**
Mass	π_S	109	-0.071	0.46	-0.090	0.35
	$\pi_N/(\pi_N+\pi_S)$	115	0.023	0.81	0.061	0.51
Longevity	π_S	78	-0.18	0.11	-0.21	0.061
	$\pi_N/(\pi_N+\pi_S)$	85	0.094	0.39	0.14	0.19
Sexual Maturity	π_S	76	0.040	0.73	0.084	0.47
	$\pi_N/(\pi_N+\pi_S)$	82	-0.059	0.60	0.068	0.54
Range	π_S	95	-0.25	**0.017**	-0.16	0.12
	$\pi_N/(\pi_N+\pi_S)$	98	0.050	0.62	-0.028	0.78
Distance from equator	π_S	95	0.050	0.63	0.062	0.55
	$\pi_N/(\pi_N+\pi_S)$	98	-0.19	0.068	-0.21	**0.041**

Table 3.4) Results of correlation analyses for two molecular evolutionary traits in nuclear DNA: π_S and $\pi_N/(\pi_N+\pi_S)$, with life history and demographic traits. Values are log-transformed before phylogenetic contrasts are calculated. The column *n* gives the number of contrasts available for each correlation. Significant results are in bold.

As previously noted, it is possible that interactions between any of life history or demographic traits might obscure a relationship between traits and nuclear molecular evolution. We examined the effect of traits together using multiple linear regression models. A model which includes mass, range and latitude has an adjusted R squared of 0.035, p-value = 0.10, number of species comparisons = 92 (Table 3.5). However, compared to this model, a model which only considers the effect of range on π_S is a better fit to the data and is statistically significant (Adjusted R-squared = 0.049, p = 0.019), but the models are not significantly different from each other (anova test p = 0.67). Including sexual maturity and longevity as traits also does not improve the fit of the model. Therefore for π_S(n), the model that fits the data best is one that only includes range.

However, no traits seem to be predictive of $\pi_N/(\pi_N+\pi_S)$(n). Latitude may have a weak effect on $\pi_N/(\pi_N+\pi_S)$, but the effects are not significant at the $p < 0.05$ level, and the model is not significant overall (n = 96, Adjusted R-squared = 0.010, p = 0.27). This may be a power issue, however, it is not due to multicollinearity, as again all variance inflation factors were less than 1.1.

Trait	π_S			$\pi_N/(\pi_N+\pi_S)$		
	Regression slope	**Sig.**	**Standardised regression coefficient**	**Regression slope**	**Sig.**	**Standardised regression coefficient**
Mass	-0.084	0.51	-0.067	-0.066	0.45	-0.069
Range	-0.13	**0.018**	-0.25	0.025	0.50	0.064
Distance from equator	0.053	0.58	0.054	-0.12	0.084	-0.19

Table 3.5) Results of multiple linear regression of life history and demographic traits on π_S and $\pi_N/(\pi_N+\pi_S)$. Values are log-transformed before phylogenetic contrasts are calculated. Regression slopes were forced through the origin.

3.4.3 Relationship between mitochondrial and nuclear molecular evolution

There are a number of species in our dataset for which we have both mitochondrial and nuclear polymorphism data: therefore, we considered whether we can detect any relationships in the molecular evolution of these two genomes. Again, all our analyses are conducted using log-transformed phylogenetically independent contrasts. Surprisingly, using the 72 species for which both mitochondrial and nuclear π_S is not 0, we find a negative correlation between π_S(mt) and π_S(n), which is significant with Pearson's (R= -0.39, p = 0.00090) but not Spearman's (rho = -0.0065, p = 0.96). The correlation between these two variables appears to be driven by a single point- the contrast between *Papio papio* and *Papio cyanocephalus*. This contrast appears to be unusual: these two species vary greatly in their mitochondrial and nuclear diversity, and while *Papio cyanocephalus* has high mitochondrial diversity, *Papio papio* has high nuclear diversity. It is possible that this reflects the introgressive hybridization that has occurred in the genus *Papio* (Boissinot et al. 2014; Zinner et al. 2009). When this point is removed from the analysis the relationship is no longer significant (Pearson's R = -0.052, p = 0.67, Spearman's rho = 0.037, p = 0.76). There are 66 species for which we have $\pi_N/(\pi_N+\pi_S)$ estimates for both mitochondrial and nuclear DNA. The relationship between the two is not significant with either Pearson's (R = -0.063, p = 0.62) or Spearman's (rho = 0.0078, p = 0.95).

3.5 Discussion

In our analysis, we have considered the relationship between various traits and two frequently used proxies for N_e: neutral genetic diversity, measured as π_S, and the ratio of nonsynonymous to synonymous polymorphisms, $\pi_N/(\pi_N+\pi_S)$. We use $\pi_N/(\pi_N+\pi_S)$ as our measure of the efficiency of selection, rather than π_N/π_S, for two reasons. Firstly, it increases the number of species for which the ratio is defined, and so fewer species had to be excluded from the analysis, reducing bias and increasing power. Secondly, there is a tendency for the average value of a ratio to be biased upwards, an effect which can be somewhat mitigated by using this slightly unusual statistic (Stoletzki & Eyre-Walker 2011). In all of the results presented, we correct for phylogenetic non-independence using paired independent contrasts (Felsenstein 1985). This is crucial, because all of the life history traits and all mitochondrial molecular traits that were considered in this study were found to be significantly affected by phylogenetic inertia. In addition, because many life history traits are expected to covary (Bromham 2011), we have examined the influence of traits both individually and in combination. We demonstrate for the first time that in mammalian mitochondrial DNA, after correcting for phylogeny, the best predictor of both π_S and $\pi_N/(\pi_N+\pi_S)$ is global species range, which suggests that global species range is an important factor determining N_e. This can be interpreted in terms of the nearly neutral theory: species with larger global ranges are likely to have larger census population sizes, and therefore larger effective population sizes.

We also find the average latitude at which a mammal species is found is predictive of π_S(mt), which is also the case in birds (Smith et al. 2017). Mass may also be negatively correlated to π_S(mt), however, in a multiple linear regression, mass is only significant at the $p < 0.1$ level. Our multiple linear regression model suggests that although neither mass nor latitude is as strongly predictive of π_S, and by proxy N_e, as range, both variables have an effect on π_S(mt) that is independent from that of range, although only that of latitude is significant at the $p < 0.05$ level. This might be because latitude, i.e. the average distance from the equator at which a species is found, is a proxy for population density, which could explain why it has additional explanatory power over π_S(mt) in our multiple linear regression model. This would indicate that species found further from the equator are less abundant than species that have more equatorial distributions, resulting in a reduced census population size, N_c, and therefore reduced N_e, even for species with similar overall range sizes. Similarly, mass might also be predictive of population density: large species are expected to be less abundant than small species (White et al. 2007) and so have smaller effective population sizes.

However, we do not find a relationship between either mass or latitude and $\pi_N/(\pi_N+\pi_S)$(mt). Therefore, an alternative possibility is that the relationships we observe between π_S(mt) and these two variables are driven by differences in mutation rate per generation (henceforth the mutation rate) between species, and not differences in N_e. To explain the pattern we observe, mitochondrial mutation rates would have to decrease with body mass and with distance from the equator. In general, we expect large, long-lived species to have higher mutation rates, due to the fact that they have a relatively high number of germ line

cell divisions per generation (Lynch 2010; Gao et al. 2016), however, we observe a relationship between π_S(mt) and mass in the opposite direction to that which we would predict from this theory. Alternatively, our results could be explained by a negative relationship between mutation rate and latitude. There is a latitudinal gradient in species diversity which holds across a wide range of species (Gaston 2000; Hillebrand 2008; Fischer 1960), and it has been hypothesised that this is due to a latitudinal gradient of molecular evolution, with species experiencing more rapid rates of molecular evolution in the tropics (reviewed by Dowle et al. 2013). Species that live at lower altitudes experience higher levels of solar energy, which leads to high productivity and metabolic rates in the tropics, all of which can be mutagenic (Rohde 1992). In agreement with this hypothesis, a relationship between latitude and molecular evolution has been found across a number of taxonomic groups, including plant nuclear DNA (Wright et al. 2006) and mammal mitochondrial DNA (Gillman et al. 2009), however, the generality of this pattern has been questioned, as in some groups, e.g. birds (Bromham & Cardillo 2003, Smith et al. 2017) and Squamata reptiles (Rolland et al. 2016), the relationship between latitude and molecular evolutionary rates is contentious. Nevertheless, it is possible that a latitudinal gradient in mutation rate is influencing mammalian π_S(mt).

While our mitochondrial DNA results have an intuitive explanation, our results differ from those of some past studies. A number of studies have failed to find any link between geographic factors, such as species range, and proxies of N_e. In a recent survey of nuclear diversity, Romiguier et al (2014) did not observe any significant relationships between geographic factors, such as average distance between GPS records, and maximum distance between GPS records, and neutral nuclear diversity. Instead, the authors found significant relationships between π_S and a number of biological traits, including body mass, longevity and propagule size. The disagreement between our results could be due to the fact that the two studies consider patterns at very different taxonomic scales: our dataset consisted only of mammal species, while Romiguier et al (2014) used a wide range of animals, including cnidarians, nematodes, arthropods and chordates. These species are likely to vary greatly in their population densities, which may confound a relationship between geographic factors (such as species range) and N_e. However, across species that are relatively similar to each other, population density is likely to vary less widely, and therefore geographic factors might be better proxies for census population sizes in taxonomically restricted datasets. This might be why we observe a relationship between global species range and proxies of N_e, while other authors have not. We also do not find a strong relationship between other life history traits and mitochondrial or nuclear molecular evolution in our dataset. Again, this might be due to our taxonomically restricted dataset. It seems likely that differences between taxonomic groups might be at least partly driving the relationship between life history traits and molecular evolution in taxonomically diverse datasets such as those reported by Chen et al (2017) and Romiguier et al (2014), while within a single taxonomic class, in which life history traits are generally more similar, this relationship does not hold.

It is worth noting that the relationships that we observe, while being statistically significant, are fairly weak, and explain quite a low proportion of the variance in π_S(mt) and $\pi_N/(\pi_N+\pi_S)$, and that both these

traits experience strong phylogenetic inertia. Therefore, it is possible that some past studies lacked power to detect a relationship. For example, Nabholz et al. (2009) considered the relationship between a direct estimate of population size and mitochondrial diversity in birds however their dataset was relatively small, while in Nabholz et al. (2008) the authors use categorical variables as predictors of species range, e.g. 'limited', 'regional' and 'worldwide', which are less informative than direct estimates of species ranges. In addition, using a dataset of 48 phylogenetically-independent island mainland comparisons, James et al (2016) found no evidence for variation between island and mainland species in terms of π_N/π_S, despite the fact that island species had range sizes that were on average less than 1% that of their mainland counterparts. However, they did find that island species had significantly lower π_S than mainland species, although the difference between island and mainland species was low, and the significant results appeared to be driven largely by young island species with low levels of divergence from their mainland counterparts. This somewhat supports our results: range size has an effect on N_e, but the relationship is noisy and the effect is not as large as we would predict from the expected variation in N_e. The relationship between N_e and N_c is not straightforward (Palstra & Fraser 2012).

This may be why we do not find any strong relationship between any life history trait and mammalian π_S(n), or $\pi_N/(\pi_N+\pi_S)$(n), despite our expectations that similar evolutionary forces will affect both nuclear and mitochondrial DNA. This is in contrast to the results of Figuet et al. (2016), who used a dataset of amniote nuclear DNA to investigate the relationship between body mass, longevity, sexual maturity and d_N/d_S. The authors found that, within mammals, all of these traits were positively correlated to d_N/d_S. It may be that we do not recover any similar patterns in nuclear DNA due to a lack of power in our nuclear dataset: nuclear DNA is considerably less polymorphic than mitochondrial DNA, consistent with the fact that mitochondrial DNA has a far higher mutation rate, and as such such our estimates of π_S(n) and $\pi_N/(\pi_N+\pi_S)$ (n) may be more noisy and less reliable. In addition, we have fewer species for which nuclear polymorphism data is available. The lower power in our nuclear dataset may also be why we do not observe a correlation between nuclear and mitochondrial π_S, or $\pi_N/(\pi_N+\pi_S)$. However, if we subset our mitochondrial dataset down to the same number of species as our nuclear dataset, we still observe relationships between mitochondrial molecular traits and life history and demographic traits: importantly, we still observe a relationship between $\pi_N/(\pi_N+\pi_S)$(mt) and range (p value = 0.064, R = -0.17), although our relationship between π_S(mt) and range is no longer significant (p value = 0.19, R = 0.11). This suggests that the patterns we observe are not solely due to differences in the sizes of the mitochondrial and nuclear datasets. In addition in mitochondria, molecular evolution is highly affected by phylogeny, with more similar species having more similar values of π_S and $\pi_N/(\pi_N+\pi_S)$. This is not the case for nuclear DNA, as we do not detect significant phylogenetic inertia in either π_S(nuclear) or $\pi_N/(\pi_N+\pi_S)$(nuclear). Therefore, overall our data suggests that nuclear and mitochondrial DNA evolve somewhat differently.

Our findings suggest that variation in N_e is widespread, however, its determinants may vary substantially depending on the taxonomic range of species used in an analysis. Life history traits may be more strongly

predictive of N_e over a wide range of species, and while global species range is predictive of N_e, its effect may only be detectable when the species under consideration are taxonomically restricted.

Chapter 4

The rate of adaptive evolution in animal mitochondria

4.1 Abstract

We have investigated whether there is adaptive evolution in mitochondrial DNA, using an extensive dataset containing over 500 animal species from a wide range of taxonomic groups. We apply a variety of McDonald-Kreitman style methods to the data. We find that the evolution of mitochondrial DNA is dominated by slightly deleterious mutations, a finding which is supported by a number of previous studies. However, when we control for the presence of deleterious mutations using a new method, we find that mitochondria undergo a significant amount of adaptive evolution, with an estimated 26% (95% confidence intervals: 5.7% to 45%) of non-synonymous substitutions fixed by adaptive evolution. We further find some weak evidence that the rate of adaptive evolution is correlated to synonymous diversity. We interpret this as evidence that at least some adaptive evolution is limited by the supply of mutations.

4.2 Introduction

Mitochondrial DNA (mtDNA) is widely used in evolutionary and ecological studies for a number of reasons. Firstly, a large number of copies of mtDNA are present in every cell, and this abundance made it relatively easy to use prior to the advent of PCR. Secondly, in many animal lineages mtDNA is highly variable due to its high mutation rate (for example, see Brown et al. 1979; Denver et al. 2000; Lynch 2010), allowing the analysis of evolutionary events over short timescales. Finally, mtDNA is inherited asexually, usually solely from the mother (Birky 1995), which means that it can be considered as a single locus, with all sites sharing a common genealogy. This can make some inferences easier to make because a single tree is appropriate for representing the evolutionary history of the molecule. This is particularly important in phylogeography, a field in which mtDNA is widely used to trace the geographical origins and movements of groups of individuals within species (for review of use in humans, see (Torroni et al. 2006)). Because the use of mtDNA is so ubiquitous, the factors that govern its evolution have received considerable attention. In particular, the role of natural selection on the diversity and divergence of mtDNA has been a focus of research (Ballard & Whitlock 2004).

Initial studies that applied McDonald-Kreitman (MK) tests to mtDNA found an excess of non-synonymous polymorphisms, suggesting that slightly deleterious mutations are common in mtDNA (Rand & Kann 1998; Nachman 1998). Slightly deleterious mutations contribute to polymorphism, but over time purifying selection is expected to remove them from populations, and so they are not expected

to contribute substantially to between-species variation in mtDNA. That purifying selection could be an important force on the evolution of mtDNA is intuitive, as mtDNA contains important genes, whose protein products are vital for the mitochondrial oxidative phosphorylation process (Ballard & Whitlock 2004). Multiple lines of evidence support the purifying selection hypothesis. For example, mitochondrial encoded genes experience base composition constraints due to the hydrophobic nature of mitochondrial-encoded proteins (Naylor et al. 1995), mitochondrial proteins are highly conserved among mammal species (da Fonseca et al. 2008) and experimental work suggests that nonsynonymous changes in mtDNA can be eliminated very rapidly between generations (Fan et al. 2008; Stewart et al. 2008).

There is also evidence to suggest that mtDNA might also undergo adaptive evolution. In a study by Bazin et al. (2006), the authors failed to find the positive relationship between mitochondrial genetic diversity and effective population size expected under neutrality. They reasoned this could be due to widespread adaptive evolution on mtDNA. Using an MK test, the authors found evidence for adaptive evolution in mtDNA but not nuclear DNA in both vertebrates and invertebrates, although the signal was only significant for invertebrates. However, the reasons for the discrepancy between this study and previous results (e.g. Nachman (1998) and Rand and Kann (1998)) are unclear. Bazin et al. (2006) used a considerably larger dataset than any previous analysis, which may have allowed them to identify the pattern of adaptive evolution. However, there are also some statistical concerns associated with this study. Bazin et al. (2006) used the Neutrality Index (NI) in order to quantify the strength and direction of selection from MK tests (Rand & Kann 1996) As a ratio of ratios, the neutrality index will tend to be biased and to have high variance (Stoletzki & Eyre-Walker 2011). Also, studies that use the Neutrality Index must exclude datasets for which the index is undefined, which can produce an estimate of adaptive evolution which is biased upwards (Stoletzki & Eyre-Walker 2011).

However, at least in particular systems there is good evidence to suggest that adaptive selection is an important influence on mitochondrial evolution. Mitochondrial genes have a key role in the oxidative phosphorylation pathway, which provides a large proportion of a cell's required ATP. A change in metabolic demand may produce selection pressure on mitochondrial-encoded genes to meet the metabolic demands of the host organism's lifestyle, possibly in concert with nuclear genes (Ballard & Rand 2005; Dowling et al. 2007). For example, evidence suggests that in bats, key genes in the oxidative phosphorylation pathway underwent adaptive evolution, likely as a consequence of the increased metabolic demand associated with the evolution of flight (Shen et al. 2010). There is also some evidence for a reduction in selective constraints on mtDNA in flightless bird species (Shen et al. 2009), which is in accordance with this hypothesis. There has also been adaptive evolution of mitochondrial genes in snakes, resulting in extensive modification of a number of core proteins, changes that may underlie the extreme metabolic regulation and efficiency exhibited by snakes (Castoe et al. 2008). Similarly, there is evidence of adaptive mitochondrial evolution in the simian ('higher') primates, which may be linked to the major phenotypic differences between simians and prosimians, which include a relatively large neocortex and a long life span (Doan et al. 2004; Grossman et al. 2004). Finally, da Fonesca et al. (2008) found evidence for functionally significant changes in mitochondrial-encoded proteins across a range of mammal species.

In addition to experiencing direct selection, mitochondria may also undergo selective sweeps through hitch-hiking with other maternally inherited genetic elements: these could include the sex chromosomes, if the female is the heterogametic sex, and cytoplasmically inherited symbionts (Ballard & Whitlock 2004). There is some evidence for reduced mt DNA diversity in birds (where females are heterogametic) compared to mammals, which could indicate mitochondrial hitchhiking with sex chromosomes (Berlin et al. 2007); however, the observed pattern could also be explained by other biological factors, such as differential mitochondrial mutation rates (Hickey 2008). Evidence for the effects of infection by cytoplasmically inherited symbionts on mtDNA is far less controversial, particularly for *Wolbachia*, a common symbiont of arthropods. *Wolbachia* infection is known to drive sweeps in mtDNA, as the infection is usually initially associated with only one mitochondrial haplotype as it spreads through a population. This has the effect of reducing mitochondrial diversity in the population and driving increased mitochondrial divergence between populations (for reviews, see (Hurst & Jiggins 2005; Galtier et al. 2009)).

In our analysis we reconsider the question of whether mtDNA undergoes adaptive evolution. We address statistical concerns associated with removing undefined values from the dataset by using methods that are always defined for any informative dataset. We also control for the presence of slightly deleterious mutations using two contrasting methods. In the first we follow the non-parametric approach suggested by Messer and Petrov (2013), based on the insight of Fay et al. (2001), who suggested removing polymorphisms below a particular frequency when estimating the rate of adaptive evolution using MK-type approaches in order to remove deleterious mutations segregating in the population. Although this is a useful *ad hoc* method, the cut-off point for excluding polymorphisms is essentially arbitrary and theoretical work suggests that this method can still produce estimates of adaptive evolution that are biased downwards (Charlesworth & Eyre-Walker 2008). Messer and Petrov (2013) suggested an extension of this method, in which the rate of adaptive evolution is estimated for each frequency class of polymorphism in the site frequency spectrum (SFS). As higher frequency variants are considered, the estimate of adaptive evolution should increase as more slightly deleterious mutations are excluded. At high polymorphism frequency classes, the rate of adaptive evolution is expected to reach an asymptote that is close to the true level of adaptive evolution. This method cannot be applied directly to mtDNA, because clonality makes the SFS highly erratic, and furthermore any one dataset typically has too few polymorphisms to make inferences reliable. We therefore propose a method by which we can combine the divergence and SFS from different species in an unweighted and unbiased manner. We then investigate how our estimate of the rate of evolution changes as a function of the frequency of polymorphisms on this combined dataset. We also apply a variant of the method suggested by Eyre-Walker and Keightley (2009)and Boyko et al. (2008) to estimate the rate of adaptive evolution. Their methods use the SFSs at selected and neutral sites to infer the distribution of fitness effects (DFE) at the selected sites and then use this to make inferences about the level of adaptive evolution. Here we estimate the DFE from the ratio of the SFSs at selected and neutral sites, combining data across species, and then use this to make inferences about rates of adaptive evolution.

4.3 Methods

We used the dataset originally compiled by Bazin et al. (2006), which was built using the Polymorphix database (Bazin et al. 2005), in order to conduct our analysis. Our dataset consists of coding, mitochondrial DNA sequences, with alignments being on average 645 nucleotides long, and contains 514 species alignments. Mitochondrial-encoded genes are highly conserved across animals, and involved in the essential process of ATP production. Unfortunately, whole mitochondrial genomes, or longer sequence fragments, are not available for the majority of the species used in this analysis. For every species entry in the dataset, we had two or more sequences from the ingroup species and between one and five outgroup species. We calculated levels of non-synonymous and synonymous divergence using the method of Goldman and Yang (1994), as implemented in the codeml package of PAML version 4.7 (Yang 2007) from pairwise alignments between each outgroup and a randomly chosen sequence from the ingroup. We also estimated the SFS and calculated polymorphism summary statistics for every species in the dataset using our own scripts. All of our summary statistic estimates, in addition to all alignment files used in this analysis, are available to download from Figshare at: http://dx.doi.org/10.6084/m9.figshare.140849.

4.3.1 Combining data

Our methods for estimating the rate of adaptive evolution required us to combine data across species. We did this by dividing the SFS at non-synonymous and synonymous sites by the total number of polymorphisms, and the numbers of non-synonymous and synonymous substitutions by the total number of substitutions to yield normalised SFSs and divergences estimates for each species and its outgroup; we then combined data across species. Hence if $\hat{P}_{ni}(x)$ and $\hat{P}_{si}(x)$ are the observed numbers of non-synonymous and synonymous polymorphisms at frequency x of the site frequency spectrum of polymorphisms in the ith species the normalised values are

$$\hat{P}^*_{ni}(x) = \frac{\hat{P}_{ni}(x)}{\sum_{\text{all x}} \hat{P}_{ni}(x) + \sum_{\text{all x}} \hat{P}_{si}(x)} \qquad \hat{P}^*_{si}(x) = \frac{\hat{P}_{si}(x)}{\sum_{\text{all x}} \hat{P}_{ni}(x) + \sum_{\text{all x}} \hat{P}_{si}(x)} \tag{4.1}$$

and if $\hat{D}_{ni}$ and $\hat{D}_{si}$ are the numbers of non-synonymous and synonymous substitutions between the ith species and its outgroup then the normalised values are

$$\hat{D}^*_{ni} = \frac{\hat{D}_{ni}}{\hat{D}_{ni} + \hat{D}_{si}} \qquad \hat{D}^*_{si} = \frac{\hat{D}_{si}}{\hat{D}_{ni} + \hat{D}_{si}} \tag{4.2}$$

The overall SFS and divergences were then calculated as

$$\hat{P}_n^*(x) = \sum_{\text{all i}} \hat{P}_{ni}^*(x) \qquad \hat{P}_s^*(x) = \sum_{\text{all i}} \hat{P}_{si}^*(x) \tag{4.3}$$

and

$$\hat{D}_n^* = \sum_{\text{all i}} \hat{D}_{ni}^* \qquad \hat{D}_s^* = \sum_{\text{all i}} \hat{D}_{si}^* \tag{4.4}$$

In this way we combine data across species weighting each species equally.

We estimate the rate of adaptive evolution using two statistics. The proportion of the non-synonymous substitutions that are adaptive:

$$\alpha = 1 - \frac{D_s P_n}{D_n P_s} \tag{4.5}$$

and the rate of adaptive evolution relative to the rate of mutation

$$\omega_a = \frac{D_n}{D_s} - \frac{P_n}{P_s} \tag{4.6}$$

Theoretical analysis

If we assume that the mutation rate is low (i.e. $2N_e u << 1$) and synonymous mutations are neutral, then the expected number of synonymous polymorphisms segregating in *i* of *n* sequences in a haploid is

$$P_s(i) = \frac{2\theta}{i} \tag{4.7}$$

where $\theta = N_e u$, *u* the rate of mutation per generation per site, and, in the case of mtDNA, N_e is the effective population size of females.

If we assume that all non-synonymous mutations are deleterious (some can be sufficiently weakly selected to be neutral) and drawn from some distribution $G(S,\mathbf{V})$, where $S=2N_e s$ and *s* is the strength of selection acting against a deleterious mutation, and $\mathbf{V}$ is a vector of parameters of the distribution then the expected number of non-synonymous polymorphisms segregating in *i* of *n* sequences is

$$P_n(i) = \theta \int_0^\infty \int_0^1 G(S;\mathbf{V}) H(S,x) Q(n,i,x)\, dx\, dS \tag{4.8}$$

where

$$H(S,x) = 2\frac{1-e^{S(1-x)}}{x(1-x)(1-e^{S})} \quad \text{and} \quad Q(n,i,x) = \frac{n!}{i!(n-i)!}x^{i}(1-x)^{n-i}$$

H(S,x) is the time a deleterious mutation subject to selection *S* spends at a frequency *x* and *Q(n,i,x)* is the probability of observing a mutation in *i* of *n* sequences for a mutation at frequency *x*. To fold the SFS we add $P_n(i)$ to $P_n(n-i)$, and $P_s(i)$ to $P_s(n-i)$. The ratio of the SFSs depends simply on the parameters governing the distribution of fitness effects.

The expected divergence at synonymous and non-synonymous sites due to neutral and deleterious mutations is:

$$D_s = 2ut \tag{4.9}$$

$$D_n = 2ut\int_0^{\infty} G(S;\mathbf{V})\frac{-S}{1-e^{S}}dS \tag{4.10}$$

respectively, where $-S/(1-e^{S})$ is approximately N_e times the probability that a deleterious mutation will become fixed. The ratio D_n / D_s depends solely on the parameters that govern the distribution of fitness effects.

We assume here that the distribution of fitness effects is a gamma distribution:

$$G(S;\lambda,\phi) = \frac{\phi^{\lambda}S^{\lambda-1}e^{-\phi S}}{\Gamma(\lambda)} \tag{4.11}$$

where λ is the shape parameter and ϕ the scale parameter of the gamma distribution; the mean of the distribution is λ/ϕ.

We use the equations above for two purposes. First we use the equations to study the consequences of folding the SFS on our ability to infer the rate of adaptive evolution. If the true value of α is α_{true} then to a good approximation the expected number of non-synonymous substitutions, D_n', can be written as

$$D_n' = \frac{D_n}{1-\alpha_{true}} \tag{4.12}$$

if we assume advantageous mutations are rare and strongly selected (i.e. they can contribute to divergence, but contribute little to polymorphism). The value of α we estimate using the polymorphisms from each frequency category is

$$\alpha_{est}(i) = 1 - \frac{D_s P_n(i)}{D'_n P_s(i)} \tag{4.13}$$

4.3.2 Estimating the rate of adaptive evolution

We can also use the equations above to estimate the rate of adaptive evolution. The ratio of the non-synonymous and synonymous SFS is expected to depend solely on the parameters of the DFE under our model. We find the parameters of the DFE that minimise the squared difference between the observed and expected values of $P_n(i)/P_s(i)$ using the Nelder-Mead algorithm as implemented in the Mathematica routine Minimize (i.e. we found the best fitting parameters using least squares). We take into account that in our data there are on average 3.52 as many non-synonymous as synonymous sites. Using the inferred DFE we estimate the expected value of $d_n/d_s = \omega_{exp}$ due to deleterious mutations, where d_n and d_s are the numbers of non-synonymous and synonymous substitutions per site. If the observed value of $d_n/d_s = \omega_{obs}$ is greater than this we infer that there has been adaptive evolution. We can estimate the rate of adaptive evolution relative to the rate mutation as

$$\omega_a = \omega_{obs} - \omega_{\exp} \tag{4.14}$$

and the proportion of substitutions that are adaptive as

$$\alpha = \frac{\omega_{obs} - \omega_{\exp}}{\omega_{obs}} \tag{4.15}$$

4.3.3 Independence

In our final analysis we consider the correlation between ω_a and synonymous diversity (π_S), variables that are both calculated using the number of synonymous polymorphisms. As a consequence, ω_a and π_S are non-independent and we would expect the variables to be positively correlated due to sampling error alone (see equation 6). To remove this statistical non-independence, we split the synonymous polymorphisms into two independent groups by sampling from the hypergeometric distribution – in effect we randomly divide the synonymous sites into two equally sized groups. One group of synonymous polymorphisms is then used to estimate ω_a, while the other is used to estimate π_S. (This is equivalent to dividing our sequences into odd and even codons, as in (Smith & Eyre-Walker 2003)). This method removes the statistical non-independence due to sampling error, and has the advantage that effective population size will be the same for both groups of synonymous polymorphisms: this is important as we are interested in determining the effect of N_e on rates of adaptive evolution.

4.4 Results

To investigate rates of adaptive evolution in animal mitochondrial DNA we applied a variety of McDonald-Kreitman type analyses to mtDNA data from a broad variety of animal species. We used the dataset originally compiled by Bazin et al. (2006). This dataset comprises multiple sequences from over 1000 animal species with at least one outgroup for each species. We excluded any species entry without at least one outgroup for which the synonymous divergence between the ingroup and outgroup was less than 0.1 or greater than 1.0. We avoided outgroups that were very closely related to the ingroup because of problems in estimating the rate of adaptive evolution using MK type methods when divergences are short (Keightley & Eyre-Walker 2012). We also avoided very divergent outgroups because of mutation saturation at non-synonymous and synonymous sites, which leads to a tendency to underestimate the rate of non-synonymous relative to the rate of synonymous substitution (Yang 2006; dos Reis & Yang 2013). If multiple outgroups were available after this filtering we chose the most closely related outgroup. This left us with a dataset of 514 animal species.

The Direction of Selection (DoS) statistic is a simple unbiased summary of the data in an MK table (DoS = $D_n/ (D_n+D_s) – P_n/(P_n+P_s)$) (Stoletzki & Eyre-Walker 2011) (note that throughout this paper, D_x or P_x refer to the total number of subsitutions and polymorphisms, at sites of type x, respectively, whereas d_x and p_x refer to the numbers per site). Assuming that synonymous mutations are neutral, positive values of DoS indicate a pattern of adaptive evolution at non-synonymous sites, whereas negative values indicate the presence of slightly deleterious mutations. We find that DoS is negative for 347 of the 514 (68%) of species, indicating that in our dataset slightly deleterious mutations are dominating mitochondrial evolution (Figure 4.1). Qualitatively similar patterns are found in both vertebrates and invertebrates, and amongst the two largest invertebrate groups the arthropods and molluscs (Table 4.1). The median DoS value does not differ significantly between either vertebrates and invertebrates (Mann-Whitney test p = 0.07), or between molluscs and arthropods (p = 0.632).

Dataset	*n*	Prop negative DoS	Median DoS
Complete dataset	514	0.68	-0.052
Vertebrates	404	0.70	-0.059
Invertebrates	110	0.59	-0.027
Arthopods	67	0.63	-0.030
Molluscs	25	0.60	-0.019

Table 4.1) A summary of the DoS results. The median DoS values and the proportion of DoS values that were negative (column titled 'Prop negative DoS') are given for each dataset. *n* = number of species included in each dataset.

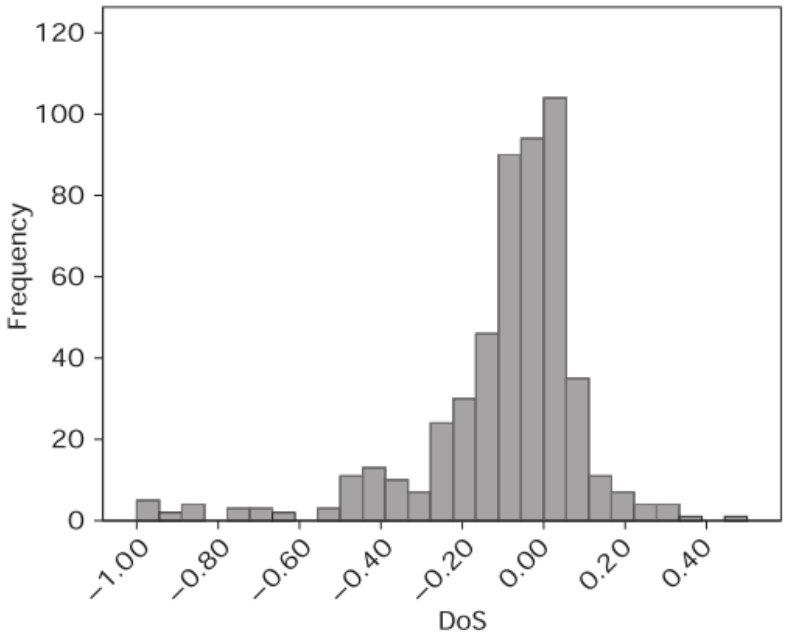

Figure 4.1) Histogram showing the frequency distribution of DoS values for the dataset.

Although the majority of DoS values are negative, a considerable minority are positive. Of these positive values, only one, *Anolis punctatus*, is significantly positive after correcting for multiple tests (results given in table 4.2); we estimate that ~80% of amino acid substitutions in this species have been fixed as a consequence of positive selection, but this is likely to be an overestimate due to the winner's curse (i.e. if evolution was rerun, we would expect to see a regression to the mean and a reduction in this high estimate).

Species	**DoS Value**	D_n	D_s	P_n	P_s	α	**Fisher's Exact Test (p-value)**
Anolis punctatus	0.36	21.47	14.63	29	97	0.80	>0.001

Table 4.2) The species for which DoS is significantly positive, after correcting for multiple tests. The DoS value, the number of nonsynonymous and synonymous substitutions, the number of nonsynonymous and synonymous polymorphisms, and an estimate of α for the species are given, followed by the p-value calculated using Fisher's exact test.

The above analyses indicate that slightly deleterious mutations dominate the evolutionary dynamics of mtDNA, and slightly deleterious mutations can obscure adaptive evolution in MK type analyses. Therefore we sought to estimate the rate of adaptive evolution while controlling for the presence of deleterious mutations.

Deleterious mutations are expected to segregate in the population at low frequencies. Therefore, we investigated how the estimate of adaptive evolution changes as a function of increasing frequency category of polymorphisms. We expect the estimate of adaptive evolution to increase and to eventually approach an asymptote as the frequency category of polymorphism being considered increases, because with increasing frequency category, more segregating deleterious mutations are removed from the population (Fay et al. 2001; Charlesworth & Eyre-Walker 2006; Messer & D. a Petrov 2013). Messer and Petrov (Messer & Petrov 2013) suggest estimating the rate of adaptive evolution from the asymptote. However, it should be noted that they advocate using the unfolded SFS, which were not able to do because of difficulties in inferring ancestral states in most of our species. Unfortunately, there are challenges in applying this method in the mtDNA of individual species for two reasons. Firstly, most species in the dataset have insufficient numbers of polymorphisms; and secondly, as mtDNA is largely clonal, all mtDNA sites share the same genealogy and hence the site frequency spectra tend to be highly erratic. Therefore we devised a method to combine data across multiple species.

The method requires that all datasets have the same number of sampled chromosomes and so we first randomly subsampled the number of chromosomes for each species down to 12 chromosomes, excluding species without sufficient numbers of sequenced chromosomes. We chose to reduce the data to 12 chromosomes because theoretical analyses suggest that this is the minimum sample size that is likely to yield a reasonable asymptote (results not shown). This subsampling reduced our total dataset to 372 species. We then combined the folded SFS and divergence data across all species from the reduced dataset, in a manner that weights the data for all species equally (see materials and methods). Finally we calculated our measure of adaptive evolution for each frequency category of polymorphism using the combined SFSs and divergence data. Following Messer and Petrov (2013) we fit a curve of the form $y = a + b\, e^{-cx}$ to the data by least squares, where a, b and c are parameters that are estimated; the estimate of adaptive evolution was taken as the value from this curve for the highest frequency class – i.e. by setting x in the equation above to 6. Confidence intervals were obtained by bootstrapping the data by species. We performed the estimation for both α, the proportion of non-synonymous substitutions inferred to have been fixed by positive adaptive evolution, and ω_a, the rate of adaptive non-synonymous substitution relative to the rate of mutation. The two statistics gave qualitatively similar patterns.

As expected, we find that our estimate of α is negative for low frequency classes and that it increases as the frequency of polymorphisms increases. However, our graph appears to asymptote at a value close to 0 when all the data are considered together, suggesting that there is little evidence of adaptive evolution in mtDNA (Figure 4.2, Table 4.3). The confidence interval on this estimate is surprisingly large given the number of datasets we have analysed. If we divide the dataset up into vertebrates and invertebrates, as Bazin et al. (2006) did, we find that the estimate of α is positive for invertebrates, and for the two biggest groups of invertebrates, arthropods and molluscs; however none of these estimates are significantly greater than zero, and there are no significant differences between any of the groups.

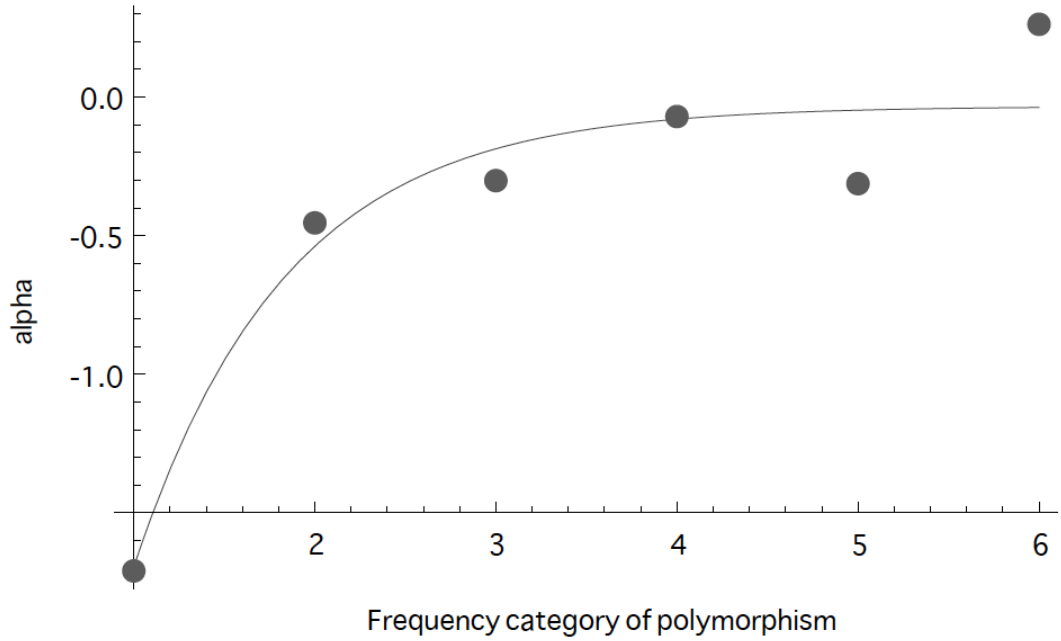

Figure 4.2) Graph of our estimate of α (using the Messer-Petrov method) plotted against the frequency category of polymorphism. An asymptotic curve of the form $y = a + b\ e^{(-c\ x)}$ was fitted to the data in order to obtain an estimate of α.

Dataset	α	Lower 95% CI	Upper 95% CI	Prop < 0
Complete dataset	-0.037	-0.4	0.28	0.58
Low divergence dataset	-0.21	-1.13	0.62	0.68
Vertebrates	-0.24	-0.62	0.21	0.83
Invertebrates	0.26	-0.16	0.57	0.08
Arthopods	0.23	-0.32	0.56	0.09
Molluscs	0.44	-6.36	0.87	0.30

Table 4.3) Results table showing estimates of α calculated using a variant of the Messer-Petrov method. α was estimated by fitting an exponential function to the graph of α plotted against polymorphism frequency category. A bootstrap was performed 100 times in order to calculate the 95% confidence intervals (CIs). The ' Prop < 0' column gives the proportion of bootstrap datasets in which the estimate of α was less than zero, providing one-tailed p-values for our results. Results are shown for the complete dataset, a 'low divergence dataset' that included species for which $0.1 < d_s < 0.5$, and for vertebrates, invertebrates, arthropods and molluscs separately.

There may be two reasons we may have underestimated the rate of adaptive evolution using the asymptotic method. First, some of our species pairs are quite divergent and d_n/d_s tends to be lower in highly divergent species (dos Reis & Yang 2013). To investigate this possibility, we repeated our analysis on a dataset that was restricted to those species pairs for which $0.1 < d_s < 0.5$. Unfortunately this reduced our dataset to just 83 species. In this reduced dataset we again found no evidence of adaptive evolution (Table 4.3)**.** Therefore it does not appear that the lower value of α is due to the underestimation of d_n/d_s on long branches.

Second, the low values of α could be due to the fact we have used the folded SFS. We folded the SFS because most of our outgroup species are too divergent to allow us to infer ancestral states. Unfortunately, folding the SFS is expected to yield a greater underestimation of α than not folding the SFS (Charlesworth & Eyre-Walker 2008). The potential severity of the underestimation can be estimated using the theory set out in Charlesworth and Eyre-Walker (2008)(see materials and methods). In the model we assume that synonymous mutations are neutral and that non-synonymous mutations are either deleterious or strongly beneficial. The fitness effects of deleterious non-synonymous polymorphisms are drawn from a gamma distribution and can be sufficiently weakly selected that they are effectively neutral. We parameterise the model such that it yields the observed value of the number of non-synonymous polymorphisms per site relative to the number of synonymous polymorphisms per site from the combined 12 chromosome dataset – i.e. we selected a shape parameter for the gamma distribution, and find the mean strength of selection acting on the deleterious mutations that will yield the correct value of p_n/p_s.

In Table 4.4 we give the estimate of α obtained from the last frequency class (this is a good approximation to the asymptotic value obtained by fitting a curve of the type used above), when 12 chromosomes have been sampled for various DFEs that are consistent with the data. It is evident from this analysis that estimates of α using the folded SFS are much more downwardly biased than estimates using the unfolded SFS and that the underestimate can be very substantial, particularly if there is little adaptive evolution and the DFE is platykurtic (i.e. high shape parameter values); for example, the estimate of α from the whole dataset (-0.037) is consistent with a true α value of 20% if the DFE has a shape parameter of 0.5 and true α value of 0.39 if the DFE has a shape parameter of 0.75.

Shape	Mean S	α True	α est (folded)	α est (unfolded)
0.25	810,000	0	-0.18	-0.037
0.5	3600	0	-0.39	-0.071
0.75	770	0	-0.64	-0.011
0.25	810,000	0.25	0.11	0.22
0.5	3600	0.25	-0.043	0.20
0.75	770	0.25	-0.23	0.17
0.25	810,000	0.5	0.41	0.48
0.5	3600	0.5	0.30	0.46
0.75	770	0.5	0.18	0.45

Table 4.4) The predicted estimated values of α (α est) using the Messer-Petrov method under different distributions of fitness effects that are consistent with the data (Shape = shape parameter of the DFE, Mean S = mean strength of selection), calculated for three different true values of α.

The tendency for the asymptotic method to underestimate the true value of α when the SFS is folded motivated the development of an alternative method. Several parametric methods have already been developed to estimate rates of adaptive evolution in which the distribution of fitness effects is estimated from the SFS, and this is then used to make inferences about the rate of adaptive evolution (Boyko et al. 2008; Eyre-Walker & Keightley 2009). Unfortunately, none of these methods can be applied to our data because of the way in which we have constructed our unweighted average SFSs and divergences. We therefore developed a new method in which we estimated the DFE from the ratio of the SFS at selected sites and the SFS at neutral sites (i.e. the p_n/p_s values for each frequency category) using least squares. We then used the DFE to estimate the expected value of d_n/d_s due to neutral and slightly deleterious mutations, inferring adaptive evolution if the observed value of d_n/d_s exceeded the observed value. The value of α and ω_a can easily be estimated from considering the difference between the observed and expected values of d_n/d_s. This method is similar to the second method presented by Eyre-Walker and Keightley (2009) which uses the method of Eyre-Walker et al. (2006) to estimate the DFE. This and the current method assume that any demographic or sampling process affects the non-synonymous and synonymous sites similarly – e.g. a demographic process that reduces synonymous singletons by 20% also reduces non-synonymous singletons by the same amount. In reality the demography does affect the ratio of the selected and neutral SFSs and hence the estimate of the DFE (Otto & Whitlock 1997), but simulations suggest the method is fairly robust (Eyre-Walker et al. 2006).

Using this new parametric method substantially increases our estimates of α over those obtained by the asymptotic method, and many are now significantly greater than zero (Table 4.5). For the entire dataset

we estimate that 26% of all non-synonymous substitutions have been fixed by positive selection ($p<0.001$). The estimate of α is higher in invertebrates ($\alpha = 0.45$, $p<0.001$) than vertebrates ($\alpha = 0.14$, $p = 0.16$) but the difference is not significant. Interestingly DFEs estimated from most of the datasets are remarkably similar. These DFEs fit the observed values of P_n/P_s well – the fit of the DFE to the whole dataset is shown in figure 4.3.

Dataset	α	**Lower 95% CI**	**Upper 95% CI**	**prop < 0**	**Shape**	**Mean S**
All	0.26	0.057	0.45	0	0.44	4600
Low divergence dataset	0.058	-0.91	0.57	0.49	0.27	140,000
Vertebrates	0.14	-0.18	0.42	0.16	0.45	3200
Invertebrates	0.45	0.12	0.61	0	0.44	7500
Arthropods	0.41	0.23	0.65	0	0.44	6900
Molluscs	0.61	-1.6	0.93	0.22	0.39	15000

Table 4.5) Results table showing estimates of α calculated using the parametric method. The data was bootstrapped 100 times to calculate 95% CIs; the proportion of bootstrap datasets in which the estimate of α was less than zero is given in the 'Prop < 0' column. The shape parameter of the DFE is given in the 'Shape' column, while the mean strength of selection is given in the 'Mean S' column. Results are shown for the complete dataset, a 'low divergence dataset' that included species for which $0.1< d_s < 0.5$, and for vertebrates, invertebrates, arthropods and molluscs separately.

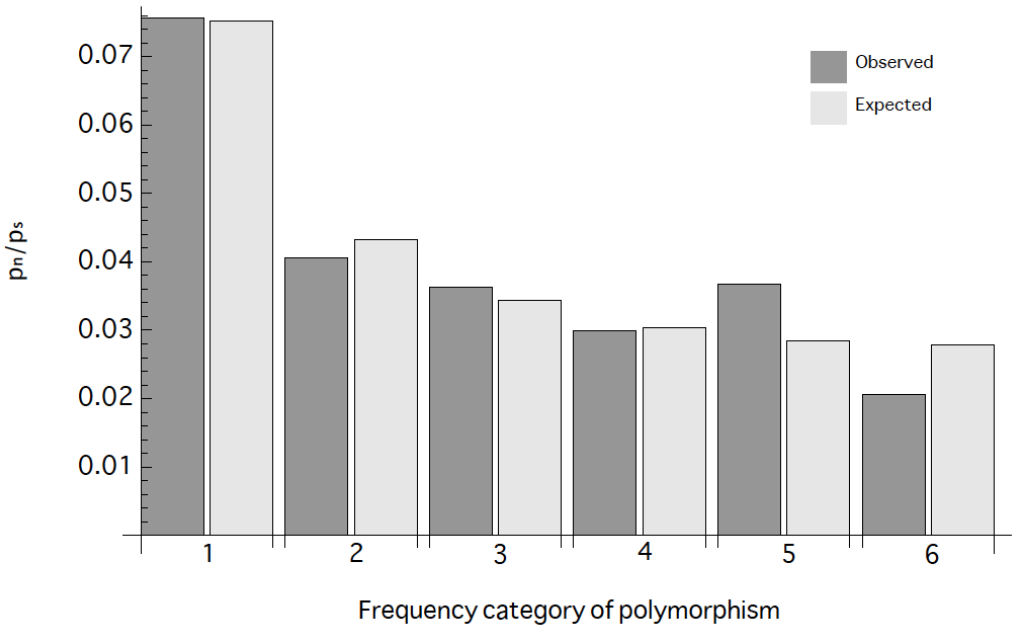

Figure 4.3) Bar chart of observed and expected values of p_n/p_s , as predicted from fitting a distribution to the DFE to the complete dataset, plotted against the frequency category of polymorphism.

We expect species with larger N_e to undergo more adaptive evolution than species with small N_e if the rate of adaptive evolution is limited by the supply of mutations. Such a correlation has been described for nuclear DNA across a small number of species (Gossmann et al. 2012). Unfortunately, we do not have estimates of N_e for most of our species as we do not have an estimate of the mutation rate per generation. However, previous analyses have suggested that synonymous diversity (π_S) in mtDNA is correlated to N_e (Piganeau & Eyre-Walker 2009). To investigate whether that was the case in this dataset we split P_s into two independent variables by sampling from a hypergeometric distribution. P_{s1} was used to estimate π_S in each species and P_{s2} was used to estimate a measure of the effectiveness of selection, $P_n/(P_n+P_s)$. We find that $P_n/(P_n+P_s)$ is significantly and negatively correlated to π_s (Spearman's rho = -0.47, $p < 0.0001$), suggesting that synonymous diversity is correlated to N_e.

To investigate whether the rate of adaptive evolution is correlated to N_e, we used the data from the above analysis, in which P_s was split into two halves: one half was used to estimate π_s and the other to estimate the rate of adaptive evolution using our parametric method. Up to this point in our analysis we have primarily concentrated on α, the proportion of substitutions that are advantageous, because it has a readily understood interpretation. However, differences in α between species can be due to differences in either the number of advantageous substitutions or the number of effectively neutral substitutions; this is not ideal for this analysis since both might be expected to be correlated to N_e. We therefore estimated ω_a, the rate of adaptive non-synonymous substitution relative to the rate of mutation (Gossmann et al. 2012). In order to use our method of calculating ω_a, it was necessary to divide the species used in the analysis into groups. We divided our species evenly into groups based on the synonymous diversity of each species: ω_a and average π_S could then be calculated for each group. We find that the correlation between ω_a and π_S is

always positive, however, the statistical significance of the correlation changes depending on the number of groups used; we only find a significant Spearman's correlation when we use 4 and 8 groups (Table 4.6). This is likely to be because the data is noisy. We conclude that there is some weak evidence that rates of adaptive evolution are correlated to levels of synonymous diversity in mitochondria.

Number of groups	Spearman's coefficient	p-value	Pearson's coefficient	p-value
4	1	0.042	0.94	0.065
6	0.54	0.27	0.4	0.43
8	0.86	0.007	0.48	0.23
10	0.26	0.47	0.52	0.13
12	0.45	0.15	0.53	0.077

Table 4.6) The strength and statistical significance of the correlation between ω_α and π_S. Results for both Spearman's and Pearsons tests are shown. In order to calculate ω_α using our method it was necessary to group the species used in this analysis: the number of groups is given in the first column of the table.

4.5 Discussion

We have investigated the evolutionary dynamics of mitochondrial evolution using variations of the McDonald-Kreitman test, in which the number of substitutions (i.e. differences between species) are contrasted to the number of polymorphisms (i.e. differences within species) at two categories of sites; those at which most mutations are neutral and those at which selection acts. Using this approach we find, as others have (Ballard & Kreitman 1994; Nachman 1998; Rand & Kann 1998; Nabholz, Mauffrey, et al. 2008) that the evolution of mtDNA is dominated by slightly deleterious mutations. However, when we control for these slightly deleterious mutations by estimating the distribution of fitness effects, we find evidence that mitochondria generally experience non-negligible level of adaptive evolution, with 26% of nonsynonymous substitutions fixed by positive selection. In this regard, our results broadly agree with the findings of Bazin et al. (2006) who found evidence of adaptive evolution in animal mitochondria, particularly in invertebrates, using the neutrality index.

We also find some weak evidence that the level of adaptive evolution in mitochondria is correlated to the effective population size, as measured by the level of synonymous genetic diversity. This is expected if the rate of adaptation is limited by the supply of mutations, because species with a high N_e are more likely to generate the advantageous mutations that allow them to adapt. It might also occur if adaptation occurs from standing genetic variation, if mutations that become advantageous were previously weakly selected or neutral. However, if advantageous mutations were previously strongly deleterious, we do not expect a

relationship between the rate of adaptation and N_e, because the diversity of strongly deleterious mutations is expected to be either independent of, or negatively correlated to, N_e. It should be noted here that we expect a correlation between ω_a and N_e but not necessarily between ω_a and π_s, because π_s is expected to be equal to $N_e u$, while ω_a is expected to be independent of the mutation rate. Hence variation in the mutation rate per generation will generate noise in the correlation between ω_a and π_s. The correlation between the rate of adaptive evolution and the level of neutral diversity is consistent with the results of Gossmann et al. (Gossmann et al. 2012) who observed a correlation between the rate of adaptive evolution and the effective population size in the nuclear genes of 13 independent pairs of animal, fungal and plant species.

Bazin et al. (2006) have suggested that there is little variation in the N_e of mtDNA between animal species because groups of animals with apparently very different census population sizes have similar synonymous diversities. However, Popadin et al. (2007) and Piganeau and Eyre-Walker (2009) have found evidence of variation in the effective population size of mtDNA between species, by showing that there is significant correlation between a measure of the effectiveness of selection and a correlate of N_e a result we have confirmed by showing that $P_n/(P_n+P_s)$ is significantly negatively correlated to synonymous diversity after accounting for the non-independence between these variables.

Our major result, that mitochondria undergo substantial levels of adaptive evolution, appears to be inconsistent with previous work indicating that nuclear genes in regions of low recombination undergo little or no adaptive evolution, at least in *Drosophila* (Betancourt et al. 2009; Campos et al. 2014). This is thought to be due to Hill-Robertson interference, whereby selection at one site interferes with selection at linked sites. There might be several reasons why mitochondria undergo adaptive evolution while non-recombining nuclear loci do not. First, although we have estimated that a substantial proportion of non-synonymous substitutions have been due to adaptive evolution in mtDNA, the absolute rate is low; the average value of d_n/d_S, estimated from the sum(normalised d_n)/sum(normalised d_s) = 0.036 so our overall estimate for ω_α = 0.009. This is considerably lower, for example, than the estimate in the nuclear genes of *Drosophila melanogaster* and *yakuba* (ω_a = 0.050 estimated from 6120 autosomal genes (Castellano Esteve et al. unpublished results)), and is perfectly consistent with the relationship between ω_a and the rate of recombination that is observed in *Drosophila* (Castellano et al. in press). Second, animal mitochondria typically have small genomes (Boore 1999) which will limit the total rate of deleterious and advantageous selection, and hence the level of Hill-Robertson interference.

Third, mitochondria might show relatively high levels of adaptive evolution because the genes they contain are essential for cell survival, and so we might predict mutations to be under intense selection, which will reduce the effects of Hill-Robertson interference. The reason is as follows. The more strongly selected an advantageous mutation is, the more likely it is to escape HRi even if the deleterious mutations are also more strongly selected, because, assuming there is no epistasis, the mutation load exerted by deleterious mutations is independent of the fitness effect of deleterious mutations (Haldane 1937). Mitochondria are also inherited in an unusual manner, experiencing what has been termed the mitochondrial bottleneck during germ cell development (Stewart & Larsson 2014), which could act to

expose variants to selection. This possibility is supported by experimental work which demonstrates that purifying selection can have a rapid and drastic impact on mtDNA, removing deleterious mutations in very few generations (Fan et al. 2008; Stewart et al. 2008).

It should be noted that our analysis assumes that synonymous mutations are effectively neutral. This may not be the case if mitochondria experience selection on synonymous codon usage for translational accuracy or efficiency. However, there is little evidence for this phenomenon in mitochondria (Pesole et al. 1999; Sun et al. 2009; Castellana et al. 2011). Our analysis is also limited in that we only consider individual protein coding sequences for each species as opposed to complete mitochondrial genomes, which were not available for the majority of species used in this analysis. We were also not able to take into account possible differences between the rates of adaptive evolution of different mitochondrial proteins, for which there is some evidence (Meiklejohn et al. 2007; da Fonseca et al. 2008). However, our approach should be robust to sampling error associated with alignment size, as this is expected to affect both synonymous and nonsynonymous sites equally. Importantly, this study has a major advantage over previous work in that our estimates of α should be unbiased, whereas past methods gave estimates of α that were biased downwards, even if a correction was made to remove deleterious polymorphisms from the analysis.

Our analysis does not account for demographic changes across species, i.e. we make the assumption that demography will affect nonsynonymous and synonymous sites equally, and thus will not influence our estimate of the DFE. This is a simplifying assumption, and it is known that demographic changes do in fact impact the nonsynonymous and synonymous site frequency spectra differently (Otto & Whitlock 1997). Methods that account for the impact of demography on the SFS tend to infer the demography and the DFE simultaneously, e.g. (Eyre-Walker 2006; Keightley & Eyre-Walker 2007), this is inappropriate for our analysis, because we use pooled data from across multiple species, all of which are likely to have different demographic histories. However, simulations suggest that we are able to estimate the shape parameter of the DFE relatively accurately even with demographic change, although demography does impact our ability to infer the mean of the distribution (Eyre-Walker et al. 2006). Therefore, although we do not account for the variation in demography between species, our estimates of the shape parameter of the DFE should be relatively accurate. It is also important to note that in our method we do not consider the variation in DFEs that exists across species, and instead calculate a DFE for the complete dataset. Although variation in the DFE is likely to exist across species, our results suggest that DFEs across the different animal groups included in this analysis are similar (for example, see table 4.5), and so applying one DFE to the complete dataset to gain an approximation of the rate of adaptive evolution is not unreasonable.

In summary, we have found evidence that mtDNA undergoes substantial amounts of adaptive evolution and that the rate of adaptive evolution is correlated to the diversity of the species being considered. These results have important implications for molecular ecology. MtDNA is widely used as a neutral genetic marker, however our results indicate that up to 45% of nonsynonymous substitutions could be fixed by

positive selection, a figure that rises to over 60% if we restrict our results to invertebrates. Therefore, adaptive evolution is likely to have a non-trivial impact on mitochondrial diversity. MtDNA diversity will, at least in part, reflect the amount of time since the last selective sweep, rather than demographic processes affecting the population.

Chapter 5

DNA sequence diversity and the efficiency of natural selection in animal mitochondrial DNA

5.1 Abstract

Selection is expected to be more efficient in species that are more diverse because both the efficiency of natural selection and DNA sequence diversity are expected to depend upon the effective population size. We explore this relationship across a dataset of 751 mammal species for which we have mitochondrial polymorphism data. We introduce a method by which we can examine the relationship between our measure of the efficiency of natural selection, the non-synonymous relative to the synonymous nucleotide site diversity (π_N/π_S), and synonymous nucleotide diversity (π_S), avoiding the statistical non-independence between the two quantities. We show that these two variables are strongly negatively and linearly correlated on a log scale. The slope is such that as π_S doubles π_N/π_S is reduced by 34%. We show that the slope of this relationship differs between the two phylogenetic groups for which we have the most data, rodents and bats, and that it also differs between species with high and low body mass, and between those with high and low mass-specific metabolic rate.

5.2 Introduction

Variation in effective population size between species is expected to have two important effects on molecular evolution. Firstly, the higher the effective population size (N_e), the greater the efficiency of natural selection in that population (Kimura 1984). This is because with increasing N_e, stochastic changes in allele frequencies have a proportionally lower impact, and therefore deleterious mutations are more likely to be removed (Corbett-Detig et al. 2015; Popadin et al. 2007). Secondly, the greater the N_e, the higher the level of neutral genetic diversity, with the level of neutral genetic diversity in a population determined by the product of N_e and the neutral mutation rate (Charlesworth 2009; Kimura 1984). We therefore expect neutral nucleotide diversity and the efficiency of selection to be correlated, as both are influenced by N_e. This prediction is well supported by a number of recent studies, both in nuclear (Galtier 2015) and mitochondrial DNA (Piganeau & Eyre-Walker 2009).

We can also make a specific prediction about the relationship between neutral genetic diversity and the efficiency of selection. If all synonymous mutations are neutral, we expect the nucleotide diversity at synonymous sites, π_S, to be equal to $4N_e\mu$. If we assume that all nonsynonymous mutations are

deleterious, although some may be sufficiently weakly selected that they are effectively neutral, we expect π_N, the nonsynonymous nucleotide site diversity, to be influenced by the mutation rate, the effective population size and the distribution of fitness effects (DFE). Assuming the DFE is a gamma distribution, we expect π_N, to be equal to $4N_e\mu\, k\, N_e^{-\beta}$, where β is the shape parameter of the gamma distribution of fitness effects and k is a constant that depends upon the mean strength of selection (Welch et al. 2008). Hence $\pi_N/\pi_S = kN_e^{-\beta}$, and $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ are expected to be linearly correlated to each other with a slope of $-\beta$ if N_e and μ are uncorrelated and if the DFE, and hence k, remains constant with changing N_e.

However, the relationship between π_S and π_N/π_S has not yet been quantified, and as such these prediction are yet to be tested using real data. The influence of the DFE on levels of standing genetic diversity is not well understood: although our past work suggests that the DFE is in fact be relatively constant across taxonomic groups, at least for animal mitochondrial data (James, Piganeau, et al. 2016), it is possible that the DFE will not remain constant with changing N_e across taxonomic groups (Lourenço et al. 2011; Martin & Lenormand 2006). However, N_e is still expected to be a major determinant of the efficiency of selection for a species. In addition, a number of factors influence π_S, and may confound a relationship between Ne and genetic diversity. This is clear from past studies that have failed to find a strong relationship between proxies of census and thus effective population size, and genetic diversity. For example, James et al. (see Chapter 3) found only a weak relationship between genetic diversity and species range, and between the efficiency of selection and range. The authors also found that other life history and demographic traits were predictive of genetic diversity but not of the efficiency of selection, which might be indicative of an influence of variation in the mutation rate (per generation) across species. This could result in variation in the relationship between π_S and π_N/π_S between different taxonomic groups.

Here we test whether the theoretical prediction of Welch et al. (2008) is upheld in mitochondrial DNA using polymorphism data from 751 mammals. We explore and quantify the relationship between neutral genetic diversity (synonymous site diversity, π_S) and the efficiency of selection (the ratio of nonsynonymous to synonymous site diversity, π_N/π_S) in mammalian mitochondria, using a new method. We compare the slope of the log-transformed relationship to the shape parameter estimated from the site frequency spectra under the assumption that the DFE is a gamma distribution. We also investigate whether the relationship between π_N/π_S and π_S differs between phylogenetic groups and according to demographic and life history parameters.

5.3 Methods

5.3.1 Dataset

Our dataset was constructed by downloading sequences from MamPol, a database of mammalian polymorphisms (Egea et al. 2007). Only protein-coding, mitochondrial DNA was used in this study, with

the majority of DNA sequences in this study being parts of the cytochrome b gene, although the dataset also contains sequences from other mitochondrial genes. As a group, mitochondrial genes are highly conserved, and are all involved in the oxidative phosphorylation pathway. To be included in our dataset, species had to be represented by a minimum of four sequenced individuals. Sequences for each species were concatenated where possible, to produce longer alignments, and then aligned using Geneious. The mean length of the alignments used in this study was 1290 base pairs, and on average our polymorphism estimates for each species were calculated over 14 individuals. We analysed the alignments using our own software to produce polymorphism estimates, and where available, we added life history and demographic data to the species in our dataset, using information from the PanTHERIA database (Jones et al. 2009). Our complete dataset contains 751 mammal species for which we have polymorphism data. All polymorphism estimates, life-history data and sequence alignments used in this analysis can be found on FigShare at: https://dx.doi.org/10.6084/m9.figshare.3084205.v1

5.3.2 Relationship between π_N and π_N/π_S

We use π_S, synonymous nucleotide site diversity, as a measure of neutral genetic diversity, and π_N/π_S, the ratio of nonsynonymous to synonymous nucleotide site diversity, as a measure of the efficiency of natural selection. These summary statistics are used, rather than raw counts of numbers of polymorphisms, to correct for the fact that the species in the dataset had variable numbers of sequenced loci, and that the sequences used were of different lengths.

Synonymous polymorphisms are used to calculate both π_S and π_N/π_S, and so we expect there to be a negative correlation between these variables just through sampling error. Therefore we removed the statistical non-independence between the variables by first dividing synonymous polymorphisms into three groups by randomly sampling from a hypergeometric distribution. We then used each group to calculate π_{S1}, π_{S2} and π_N/π_{S3} respectively. This is analogous to dividing each sequence into thirds (Piganeau & Eyre-Walker 2009; Smith & Eyre-Walker 2002). Since we were interested in the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$, we needed to ensure there were no cases in which either π_N or π_S were zero, however, removing species in which π_S or π_N/π_S was 0 would result in biased results. In addition, individual measurements of π_S and π_N/π_S will be subject to a large degree of sampling error. Therefore, to overcome these problems we first ranked species by π_{S1}, and based on this ranking the species were divided into evenly-sized groups. Average π_{S2} and π_N/π_{S3} values were calculated for each group. We then ran ordinary least squares regression between the log-transformed values of these averages.

5.3.3 Correcting for phylogenetic non-independence

To ensure that our results are not due to phylogenetic non-independence, we used paired- independent comparisons (PIC) (Harvey & Pagel 1991). Using DNA-based phylogenetic literature, we identified 186 sister pairs of species in our dataset, where sister pairs are defined as sharing a common ancestor to the exclusion of all other species in the dataset. We then repeated our method as before, using the ratio of π_{S1}

between species in a pair to rank and group pairs, and calculating the average ratio of π_{S2} and π_N/π_{S3} between species in a pair over each group. We then considered the relationship between the log of the average ratios of π_{S2} and π_N/π_{S3}.

5.3.4 Simulations

To investigate the performance of our method we ran two sets of simulations. In the first we used SFS_code (Hernandez 2008) to simulate loci with no intra-locus recombination, but with free recombination between loci. We ran simulations of loci of three different lengths: 1, 1000 and 10,000 codons. The number of loci simulated was reduced as the length of the loci was increased, such that we simulated 100 loci 1 codon long, 10 loci 1000 codons long and 3 loci 10,000 codons long. The synonymous sites in these codons were assumed to be neutral and the non-synonymous mutations to be deleterious and drawn from a gamma distribution. In each simulation the population size was set to 5000 but the arithmetic mean value of Ns and the value of $N\mu$ were changed to reflect changes in N (i.e. we take advantage of the fact that increasing s and u x-fold is equivalent to increasing the population size by x-fold because population genetic behaviour depends on the product of the effective population size and the other parameters). We set the mean Ns and $N\mu$ to be 100 and 0.001, 500 and 0.005, 1000 and 0.01, 2000 and 0.02, 4000 and 0.04, 8000 and 0.08, respectively (note that because of background selection the effective population size was not equal to the census population size). We simulated data under three different shape parameters for the gamma distribution: 01, 0.3 and 0.5. Each simulation was run for $15N$ generations for the population to equilibrate before the population was sampled, and for each combination of parameters a number of iterations were run, such that for each parameter combination at least 1000 synonymous polymorphisms were sampled (the exception was the set of simulations with 10,000 codons run with $N\mu = 0.08$, which were only run once since they ran so slowly; in these cases at least 100 synonymous polymorphisms were sampled).

In the second set of simulations we investigated the statistical properties of the method, and in particular our scheme for combining data from different species and/or genes. In each simulation we had 500 species, each of which had 1000 synonymous sites. We sampled effective population sizes from a gamma distribution with a shape parameter β_N, arbitrarily and without loss of generality, setting the expected value of N_e to one (the absolute value of N_e is not important in this context, since what matters is how the proportion of effectively neutral mutations changes with N_e, and this is independent of the absolute value). The expected number of synonymous polymorphisms, $E(P_s)$, and non-synonymous polymorphisms, $E(P_n)$, for species i were calculated as $E\left(P_{si}\right)=E\left(P_s\right)N_{ei}$ and $E\left(P_{ni}\right)=kN_{ei}^{-\beta_s}$ where β_s is the shape parameter of the distribution of fitness effects, and k is a constant that normalises the expected values of P_n such that $E\left(P_n\right)=0.2\,E\left(P_s\right)$, approximately the pattern that is observed in our data. The simulated values of P_n and P_s were generated by sampling from a Poisson distribution with the expected values as given above. The method then proceeded as detailed previously. We investigated the effects of altering the size of the groups, the average number of synonymous polymorphisms, variation in N_e and the shape parameter of the DFE. For each combination of parameters, we ran the simulation 100

times. Throughout our analyses we use ordinary least squares regression, however we also investigated the use of standard major axis regression in our simulations.

5.3.5 Calculating the DFE

In order to calculate the DFE of mitochondrial mutations, we combined the synonymous and nonsynonymous site frequency spectra (SFS) across species. We cannot calculate the DFE for individual species, firstly because mitochondria are inherited in a clonal manner, which can make the SFS highly erratic; and secondly because the majority of species have too few polymorphisms to allow us to make a reliable estimate of the DFE. We therefore combined SFS data across species in the dataset using the method of James et al. (2016); this method weights the data for each species equally, to produce an overall nonsynonymous and synonymous SFS for the dataset. We then inferred the DFE by fitting a gamma distribution to the ratio of nonsynonymous to synonymous polymorphism at each frequency category of the SFS using least squares. Full details of the method are given in James et al. (2016). This method required each species in the dataset to have a common number of sampled individuals: we therefore produced datasets in which the number of individuals (n) for each species was resampled down to a common number. We produced two resampled datasets, one in which $n = 5$ and one in which $n = 11$, however the sequence data are otherwise identical to that used in the previously described methods. Any species that did not have a minimum of n sampled individuals was excluded from the datasets, therefore our datasets sub-sampled to 5 and 11 individuals contained 564 and 256 species respectively. We again fitted regression models to the sub-sampled datasets, randomly splitting synonymous polymorphisms into three groups and calculating average values of π_S and π_N/π_S over groups of species as before.

To test whether the shape parameter of the DFE, as inferred from the SFS, was different from the slope of the regression between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ we bootstrapped the data by species 100 times, in each re-estimating the shape parameter of the DFE using the resultant SFS, and the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$.

5.4 Results

We have investigated the relationship between a measure of the efficiency of natural selection, $\log(\pi_N/\pi_S)$, and diversity, $\log(\pi_S)$, in mammalian mitochondria, using a polymorphism dataset of 751 species. If there is free recombination, the DFE is gamma distributed and the effective population size, N_e, is uncorrelated to mutation rate, u, then this relationship is expected to be linear with a slope equal to the shape parameter of the gamma distribution (Welch et al. 2008). However, it is not straightforward to investigate this relationship for three reasons. First, for many of our species either π_N or π_S is zero and hence one of our two statistics is undefined, however, to exclude these species will bias our results. Second, there will be a degree of variable error and 'noise' in our individual measurements of π_N and π_S. And third, $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ are not statistically independent; we therefore expect there to be a negative correlation between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ simply because of sampling error in π_S, which arises because we have

sequences of finite length. To overcome these problems, we randomly split the synonymous polymorphisms into three independent groups, using each to calculate a separate value of π_S. We used the first estimate of π_S to rank and group species, the second value as our estimate of π_S for the group and the third value to calculate π_N/π_S. We ran ordinary least squares regression between the $\log\left(\overline{\pi}_n / \overline{\pi}_{s3}\right)$ and $\log\left(\overline{\pi}_{s2}\right)$ where the means are for each group of species. Using this method reduces the variance in our dataset, such that the smaller the number of groups the greater the reduction in variance.

5.4.1 Simulating the method

To investigate the properties of the method we ran two sets of simulations. In the first we investigated the population genetics of the method; in particular, we were interested in ascertaining whether linkage affected the predictions determined under the assumption of free recombination. We simulated loci with 1, 1000 and 10,000 codons. There was free recombination between loci but no recombination within a locus. Synonymous mutations were assumed to be neutral and non-synonymous mutations to be deleterious and drawn from a gamma distribution. We altered the population size over nearly two orders of magnitude from a mean Ns value of 100 and an $N\mu$ value of 0.001, to values of 8000 and 0.08 respectively (where s is the strength of selection and μ is the mutation rate). We simulated data under three different shape parameters: 0.1, 0.3 and 0.5.

The results are shown in figure 5.1, where $\log(\pi_N/\pi_S)$ is plotted against $\log(\pi_S)$. Despite the fact that there is considerable background selection in the some of the simulations, such that BGS reduces synonymous diversity by more than 10-fold in the simulations with 10,000 codons and high $N\mu$ values, the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ is close to that expected (Figure 5.1). However, there is a slight but significant tendency to underestimate the slope. This underestimation does not depend on linkage.

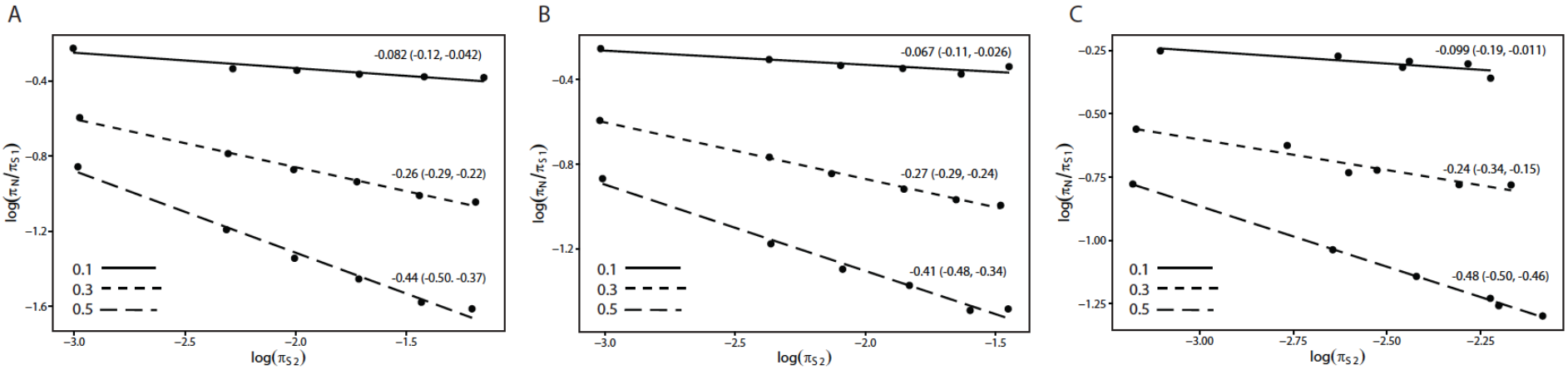

Figure 5.1) The relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ in simulated data when each non-recombining locus contains (A) 1 codon, (B) 1000 codons and (C) 10,000 codons . In each panel the lines from top to bottom show the results for different shape parameters of the DFE: 0.1, 0.3 and 0.5.

In the second set of simulations we sought to investigate the statistical properties of the method and in particular whether our method of combining data from different species and/or genes gave biased estimates. To do this we ran simulations in which we generated the number of non-synonymous (P_n) and synonymous (P_s) according to our model, analysing the resulting data according to the method detailed above – i.e. splitting the synonymous polymorphisms into three groups, and considering regressing $\log(P_n/P_s)$ against $\log(P_s)$ using ordinary least squares regression. We consider the relationship between $\log(P_n/P_s)$ and $\log(P_s)$ rather than $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ because theory predicts the relationship should be the same (Welch et al. 2008) and simulating the numbers of polymorphisms rather than the diversity is more straight-forward.

Our simulations suggest that the method is unbiased when the shape parameter of the gamma distribution is very small – i.e. when there is no expected relationship between $\log(P_n/P_s)$ and $\log(P_s)$ (see table 5.1). However, the method can either be upwardly or downwardly biased when the shape parameter of the gamma distribution is greater than zero (Table 5.1). When there is relatively little variation in N_e (higher values of β_{Ne}) and relatively few synonymous polymorphisms then the method tends to estimate the slope to be shallower than it should be. This bias can be ameliorated by using large groups of species/genes, but was not helped by using standard major axis regression; this led to a dramatic overestimation of the slope (data not shown). The bias in underestimating the slope is not surprising; the ability to estimate the relationship between $\log(P_n/P_s)$ and $\log(P_s)$ will depend upon the relative magnitudes of the variation in N_e and the sampling error in P_s; when the latter dominates the former then it is difficult for the method to determine which species/genes have high or low N_e. Surprisingly the method can also estimate the slope to be slightly steeper than it should be when there is substantial variation in N_e and few synonymous polymorphisms. However, so long as the mean number of polymorphisms per species/gene is reasonable (on average >8), and there is moderate variation in N_e then the method is largely unbiased if large groups of species/genes are used (see table 5.1). The bias is not likely to be very large in our dataset since the average number of synonymous polymorphisms is quite large (approximately 16) and we have substantial variation in P_s (we estimate β_{Ne} to be 1.5 assuming all the variation in P_s is due to variation in N_e).

β_s	E(P_s)	β_{Ne}	Group Size	Mean slope (SE)
0.001	2	1.5	10	-0.04 (0.02)
	4			0.03 (0.01)
	8			0.02 (0.01)
	16			0.00 (0.01)
	32			0.01 (0.01)
	2		20	0.03 (0.03)
	4			0.01 (0.02)
	8			0.02 (0.01)
	16			-0.00 (0.01)
	32			0.01 (0.00)
	2		50	-0.00 (0.04)
	4			0.02 (0.02)
	8			0.01 (0.01)
	16			0.00 (0.01)
	32			-0.01 (0.00)
	2	0.1	50	0.00 (0.02)
	4			0.01 (0.01)
	8			0.00 (0.01)
	16			0.00 (0.00)
	32			0.01 (0.00)
	2	10	50	0.08 (0.07)
	4			-0.05 (0.05)
	8			-0.01 (0.04)
	16			0.00 (0.00)
	32			0.01 (0.00)
0.5	2	1.5	10	-0.26 (0.03)
	4			-0.40 (0.01)
	8			-0.44 (0.01)

β_s	E(P_s)	β_{Ne}	Group Size	Mean slope (SE)
	16			-0.46 (0.01)
	32			-0.48 (0.00)
	2		20	-0.33 (0.03)
	4			-0.45 (0.02)
	8			-0.45 (0.01)
	16			-0.46 (0.01)
	32			-0.49 (0.00)
	2		50	-0.54 (0.05)
	4			-0.53 (0.02)
	8			-0.48 (0.01)
	16			-0.48 (0.01)
	32			-0.49 (0.00)
	2	0.1	50	-0.58 (0.01)
	4			-0.54 (0.01)
	8			-0.52 (0.01)
	16			-0.49 (0.00)
	32			-0.49 (0.00)
	2	10	50	-0.01 (0.07)
	4			-0.22 (0.06)
	8			-0.36 (0.04)
	16			-0.44 (0.02)
	32			-0.47 (0.01)

Table 5.1) The mean slope estimated from 100 simulated datasets under various parameter combinations. Symbols: $\beta_{s,}$ the shape parameter of the DFE; E(P_s), the average number of synonymous polymorphisms; β_{Ne}, shape parameter of the gamma distribution sampled to produce effective population sizes. The last column gives the mean and standard error of the slope from simulation run.

5.4.2 Overall relationship between log(π_N/π_S) and log(π_S)

Across the entire mammalian mitochondrial dataset, we find that log(π_N/π_S) is almost perfectly linearly related to log(π_S) after grouping species into 8 groups (Figure 5.2). The correlation is highly significant

(Pearson's R = -0.98, p <<0.001). The slope of the relationship is -0.60 (SE = 0.050) (Figure 5.2), which means that if diversity doubles the proportion of effectively neutral substitutions is reduced by $1- 2^{-0.60} =$ 34%. Similar results are obtained if we use 10, 20, 30 40 and 50 groups (see Appendices Table A5.1), with the correlation remaining highly significant (p << 0.005) for all numbers of groups. However, as the number of groups increases there is a trend for the slope of the line to become shallower. This is in accordance with the results of our simulations: as the number of groups increases, the lower the number of synonymous polymorphisms per group and the greater the bias in our method towards underestimating the slope.

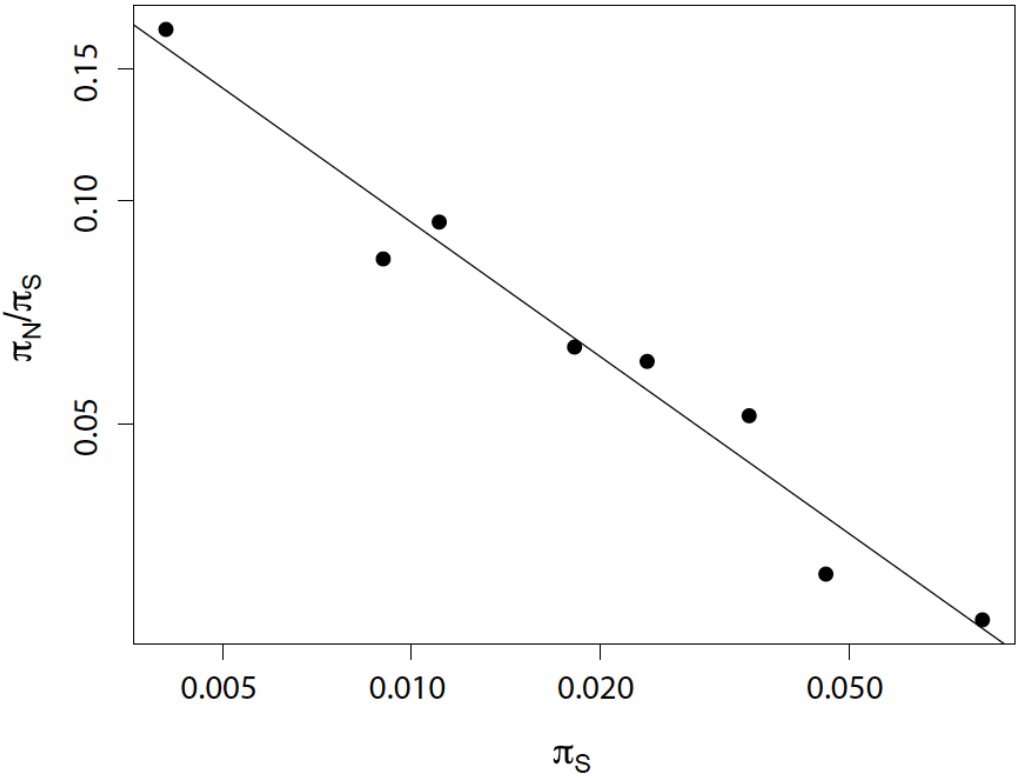

Figure 5.2) The relationship between π_N/π_S and π_S in mammalian mitochondrial DNA. Plotted on a log scale.

5.4.3 Correcting for phylogenetic non-independence

The correlation between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ is not due to phylogenetic non-independence between species. Using paired-independent contrasts, we repeated our analysis (exploring the correlation between the log of the ratio of π_N/π_S and the log of the ratio of π_S for each species pair, again dividing species into 8 groups) and found a linear correlation, of slope -0.60 (SE = 0.067), which is identical to that of our non-paired dataset. The correlation was also highly significant (Pearson's R = -0.96, p = 0.00011). Similar results are obtained when using 10, 20 and 30 groups (results not shown).

5.4.4 Taxonomic groups

There are a number of reasons why the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ might differ between species. To investigate this question we compared the two groups represented by the largest number of species in the dataset, bats (Chiroptera) and rodents (Rodentia) (178 and 226 species respectively). These are the two most species-rich groups of mammals. Results are shown in figure 5.3. While the correlation remains significant and negative in both bats and rodents, in rodents the slope of the line is far steeper (slope = -1.13, SE = 0.10; intercept = -2.97, SE = 0.16) than in bats (slope = -0.43, SE = 0.15; intercept = -2.08, SE = 0.26.) a difference that is statistically significant (p=0.0022). Again this result holds if we use 10 and 20 groups, although the difference is not significant with 30 groups (results not shown). So although π_N/π_S is substantially lower in bats than rodents, the efficiency of selection does not increase as rapidly with increasing π_S in bats as it does in rodents.

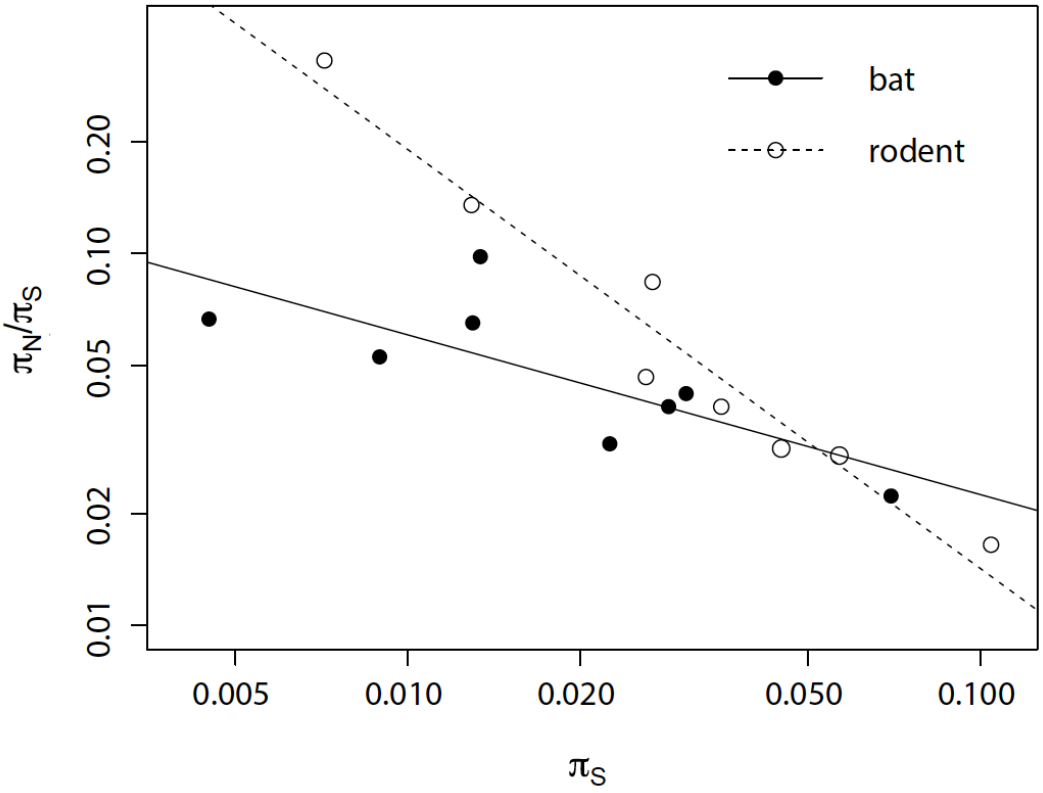

Figure 5.3) Comparison of the relationship between π_N/π_S and π_S of bats (Chiroptera) and rodents (Rodentia). Plotted on a log scale.

5.4.5 Life history and demographic traits

We also investigated whether we could detect any influence of life history or demographic traits on the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$. We looked at three traits: body mass, species range size and mass-specific metabolic rate (i.e. resting metabolic rate divided by body mass). We ranked the species by the trait in question, and split the species into two evenly-sized groups depending on the ranking. We then used the method as described previously (grouping the species into 8) to investigate the correlation of

$\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ separately for each group. Therefore we have an estimate of the relationship for those species in which the life history trait is low, and an estimate for those species in which the life history trait is high. Results are shown in table 5.2 and figure 5.4. Of the species in our dataset, we have 567 with body mass estimates, 588 with range area estimates and 157 species with mass-specific metabolic rate estimates.

We found that two of the life history traits we considered had a significant effect on the regression slope of $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$. The largest difference in regression slope was found to be between mammals with high and low mass-specific metabolic rates: the slope was steeper for mammals with high rates, such that π_N/π_S decreases more rapidly with increasing π_S in mammals with high as opposed to low mass-specific metabolic rates. We also found that the slope of the regression line was significantly different between mammals with low body mass and mammals with high body mass, with smaller mammals having a steeper regression line. Range size on the other hand did not appear to influence the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$.

	Low		**High**		
Trait	**Slope**	**Intercept**	**Slope**	**Intercept**	**Sig.**
Mass-specific metabolic rate	-0.49	-2.03	-1.32	-3.36	0.037
Mass	-0.91	-2.81	-0.36	-1.77	0.012
Range	-0.50	-2.09	-0.46	-2.08	0.74

Table 5.2) Comparisons of the regression slope and intercept for different life history traits. The species were always split into 8 groups in this analysis. The slope and the intercept of the regression line for each life history trait are given. The relationship between π_S and π_N/π_S was statistically significant for all the above subsets of the data, with $p < 0.05$. The linear models for species with high and low values of the life history trait were compared using an ANOVA test, the significance level of which is given in the last column.

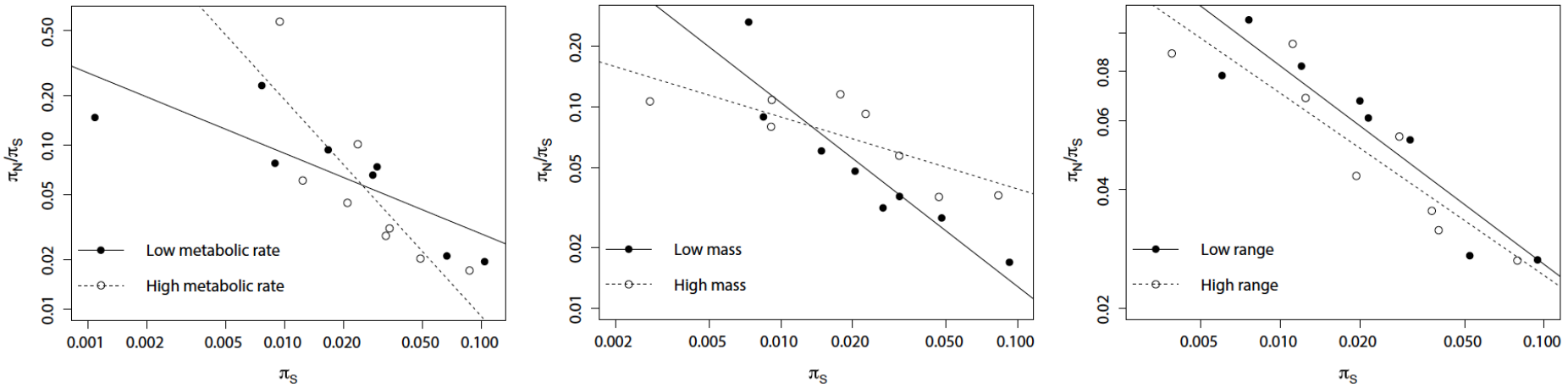

Figure 5.4) The influence of life history and demographic traits on the relationship between π_N/π_S and π_S. We consider whether the relationship is different for species with a higher value of a given trait as opposed to a lower value. From left to right, the traits considered are: mass-specific metabolic rate (referred to as 'metabolic rate' in the figure legend, $mL.O_2/hr/g$), body mass (g) and range size (km^2). Plotted on a log scale.

We also investigated the interaction of body mass and mass-specific metabolic rate on the regression slope between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$. In this analysis we increased the number of groups used in the regression to 20. We set up the model such that $\log(\pi_N/\pi_S)$ was the dependent variable, with $\log(\pi_S)$ as the covariate. We find that including a possible interaction between π_S and mass-specific metabolic rate does not significantly improve the fit of the model, and the interaction term is not significant (interaction term = -1.81, p = 0.21, such that if the value of $\log(\pi_S)$ increases by 1, the slope of the interaction between $\log(\pi_N/\pi_S)$ and log(mass-specific metabolic rate) decreases by the value of the term). While the interaction term between π_S and body mass was significant (interaction term = -0.13, p=0.022), if we repeat the analysis using higher numbers of groups (30 and 50) the interaction is no longer statistically significant.

5.4.6 Comparison of the slope to the shape of the DFE

If the DFE is a gamma distribution then the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ should be linear with a slope equal to the negative value of the shape parameter. Our analysis above suggests that the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ is linear, but is the slope equal to the shape parameter? To investigate this we used an independent method to estimate the shape parameter of the DFE. We sub-sampled the sequences for each species down to a common number of sequences (n=5 and n=11) and combined the site frequency spectra in a manner that weights every species equally, and then estimated the DFE by fitting a gamma distribution to the ratio of nonsynonymous to synonymous polymorphisms at each frequency category of the SFS, using the method of James et al. (2016). We then conducted our analysis of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ on the resampled datasets as before. To test whether the estimate of the shape parameter, as inferred from the SFSs, was significantly different to the negative value of the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ we bootstrapped the data by species 100 times. The results are shown in table 5.3. We found that our estimates for the shape parameter of the gamma distribution was considerably and significantly smaller than the negative value of the slope of the regression line between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ for both 5 and 11 sample datasets (t-test: $p << 0.001$ for both datasets, where the alternative hypothesis is that the true difference in means is less than 0). This is particularly striking considering that simulations indicate our regression method may estimate the slope to be shallower than it should be.

No. of chromo.	Shape	Lower CI	Upper CI	S	Lower CI	Upper CI
5	0.45	0.21	0.56	1120	289	34400
11	0.44	0.34	0.57	1890	574	12200

Table 5.3a) Estimates of the DFE. The first column gives the number of chromosomes sampled for each species in the dataset. The shape parameter of the gamma distribution of fitness effects is given in the 'shape' column, and the estimated strength of selection is given in the S column. Confidence intervals for each variable are labelled as 'Lower CI' and 'Upper CI'.

Dataset	Slope	Upper CI	Lower CI	Intercept	Upper CI	Lower CI
5	-0.63	-0.38	-0.79	-2.34	-1.96	-2.61
11	-0.74	-0.37	-0.91	-2.54	-2.01	-2.81

Table 5.3b) Regression slope and intercept estimates. The first column gives the number of individuals sampled for each species in the dataset. The slope and the intercept of the regression line are given. Confidence intervals for each variable are labelled as 'Upper CI' and 'Lower CI'.

5.5 Discussion

The relationship between π_N/π_S and π_S is of considerable biological interest. There are few estimates of N_e available, and so π_S is commonly used as a proxy for N_e. However, the extent to which π_S is related to measures of selective constraint was not previously known. In addition, Welch et al (2008), extending the work of Kimura (1962), were able to make specific predictions about how we expect π_S and π_N/π_S to be related under the nearly neutral theory: if the DFE is a gamma distribution, the relationship between these variables should be loglinear, and have the same slope as the shape parameter of the DFE. These theoretical predictions had not previously been tested with real data. Here, we show that $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ are strongly negatively and linearly correlated in mammalian mitochondria, so that as neutral genetic diversity doubles, the efficiency of selection also increases resulting in a 34% reduction in the number of effectively neutral polymorphisms.

Life history traits affected the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$: we find that the slope of the relationship is greater in species with high as opposed to low mass-specific metabolic rates;

therefore, with increasing genetic diversity, the increase in selective constraint is greater in species with high metabolic rates. We also find that the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ is shallower for species with high as opposed to low body mass. This may be due to a correlation between the two life history traits: species with low body mass are known to have higher mass-specific metabolic rates (Schmidt-Nielsen 1984; Suarez 1992). We also find that the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ is shallower in bats than in rodents. This does not appear to be driven by differences in life history traits: the bats and rodents in our dataset have very similar mass-specific metabolic rates (t-test p-value = 0.88, mean mass specific metabolic rate is 1.45 for rodents, 1.43 for bats), and whilst bats were found to be significantly smaller than rodents (t-test p-value < 0.0001, mean mass = 196.7 g for rodents, 30.7 g for bats), this would have been expected to generate the opposite pattern in terms of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$.

There are a number of reasons why the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ might differ between bats and rodents, and between species with high and low body mass and metabolic rate. It might be that the DFE differs between the two groups, or alternatively it might be that the relationship between π_S and N_e differs; for example, in some groups N_e and the mutation rate per generation might be negatively correlated (Lynch 2010). There might also be differences in how the strength of selection changes with N_e between different taxonomic groups and species with different life history traits, i.e., the parameter k in the equation $\pi_N/\pi_S = kN_e^{-\beta}$, which is related to the strength of selection, may also be a function of N_e. To differentiate between these possibilities, we estimated the DFE from the SFSs for each group. The results are presented in table 5.4. The shape parameter of the DFE estimated from the SFS mirrors the difference in slopes between high and low body mass and high and low metabolic rate, although only in the dataset in which $n = 5$ is the difference significant. It might therefore be that the DFE differs between groups with different body sizes and metabolic rates. In contrast, the estimate of the shape parameter of the DFE from the SFS is very similar in bats and rodents suggesting that the difference in the slope between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ is not due to differences in the shape of the DFE between these taxonomic groups. However, the SE associated with these estimates of the shape parameters are large: the estimates are likely to be less accurate because they are based on polymorphism data from far fewer species. This is also clear from our analysis of the slope of the regression between π_S and π_N/π_S, as shown in figure 5.3. Although the regressions are clearly significant, and the slopes are significantly different between bats and rodents, our estimates for these slopes have higher SEs and the relationship between π_S and π_N/π_S is not as clear, due to the fact that these analysis are based on data from fewer species.

It may be that the difference we observe in the slope between bats and rodents is due either to different relationships between π_S and N_e, or to differences in how the k parameter changes with N_e in the two groups. For example, if k were to increase with increasing N_e in bats but not rodents, the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ would be shallower for bats. This might occur because of the metabolic demands of flight, which could result in a more rapid increase in a higher mean strength of selection on bat mitochondria with increasing N_e. There is some evidence to support this possibility: for example, there are signs of adaptive mitochondrial evolution on the common ancestral lineage of bats, but

not rodents (Shen et al. 2010). In addition, Shen et al. (2009) have found a relationship between flight ability and the strength of selective constraint in bird mitochondrial DNA. However, there is also evidence to suggest that patterns of molecular evolution are different between these two groups, which may suggest a difference in the relationship between π_S and N_e. For example, Nabholz et al. (2008) have found that synonymous substitution rates are considerably lower in bats than in rodents, which could be due to lower mutation rates or longer generation times in bats than in rodents.

Group	***n***	**Shape**	**Sig.**
Rodents	5	0.51 (0.25)	n.s.
Bats	5	0.46 (0.13)	
Rodents	11	0.66 (0.11)	n.s.
Bats	11	0.59 (0.35)	
High body mass	5	0.13 (0.095)	<0.01
Low body mass	5	0.73 (0.16)	
High body mass	11	0.37 (0.11)	n.s.
Low body mass	11	0.59 (0.19)	
High metabolic rate	5	0.71 (0.30)	0.01
Low metabolic rate	5	0.059 (0.083)	
High metabolic rate	11	0.70 (0.28)	n.s.
Low metabolic rate	11	0.23 (0.076)	

Table 5.4) Estimates of the DFE for different taxonomic and life history groups. The group of species for which the DFE was calculated is given in the first column. The second column gives the number of individuals sampled for each species in the dataset The shape parameter of the DFE, with the standard error of the estimate in brackets, is given in the 'shape' column. The significance level of the difference in the DFE between the groups is given in the last column. n.s. = not significant.

Our results depart from the theoretical prediction that the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ is the same as the (negative) shape parameter of the DFE, as estimated from the SFS, with the shape parameter of the gamma distribution being significantly smaller than the slope of the regression

line. There are a number of possible reasons for this. Firstly, as Welch et al. (2008) note, if the DFE does not follow a gamma distribution, then the above predictions may not hold. Although the DFE is most commonly modelled as a gamma distribution (Boyko et al. 2008; Eyre-Walker et al. 2006; Eyre-Walker & Keightley 2007; Piganeau & Eyre-Walker 2003), some studies have found support for alternative distributions, such as the lognormal (Loewe & Charlesworth 2006), the normal (Nielsen & Yang 2003), and the bimodal beta distribution (Kousathanas & Keightley 2013). However, the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ is linear, which is consistent with the DFE being a gamma distribution; if it were log-normal, for example, the relationship should show curvature (Welch et al. 2008).

Secondly, it has been suggested that a negative relationship exists between N_e and μ, the mutation rate per site per generation. This is because in populations with larger effective population sizes natural selection is more efficient, and so should be more able to reduce the mutation rate μ (Lynch 2010). This introduces a negative interaction between genetic diversity and the efficiency of selection, which will tend to make the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ steeper. However, the importance of this effect remains unclear: although some studies have found evidence to suggest that there is a negative correlation between N_e and μ (Cutter et al. 2013; Lynch 2010; Piganeau & Eyre-Walker 2009), the theory predicts a relatively small effect of N_e on μ, which could be masked by the impact of other influences on μ, such as rate of sperm production and exposure to mutagens (Gao et al. 2016). In addition, the theory suggests that the mutation rate of a species is reduced as far as possible by selection, with genetic drift preventing selection from further reducing the rate (Lynch 2011). However, mutation rates vary widely between species, with some exceeding the upper limit predicted by this theory by orders of magnitude (Martincorena & Luscombe 2013).

Thirdly, selection on synonymous codons may influence the slope of $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$. Synonymous sites are commonly assumed to be neutrally evolving; however, codon usage bias in nuclear genes has been reported in a number of species, including *Saccharomyces pombe*, *Caenorhabditis elegans*, *Drosophila* and *Arabidopsis* (Duret & Mouchiroud 1999; Hershberg & Petrov 2008; Kanaya et al. 2011). Selection on synonymous sites is thought to occur in order to maximise translational efficiency by matching tRNA abundances (Kanaya et al. 2011) and to improve mRNA stability (Chamary & Hurst 2005). (Although not relevant to mtDNA, selection on synonymous sites may also maintain accurate splicing of mRNA (Carlini & Genut 2006)). If selection for optimal codons also reduces the number of nonsynonymous polymorphisms, this will tend to make the slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ steeper than that predicted by the shape of the distribution of fitness effects. However, if selection on synonymous codons only affects synonymous polymorphisms then it is expected to affect diversity and the efficiency of selection equally and therefore will not influence the slope of the relationship. In addition, the strength of selection in animals with low N_e may not be sufficient to select for optimal codon usage, and there is little evidence that selection acts on synonymous sites in mammals, or indeed vertebrates, which are thought to have relatively small effective population sizes (Duret 2002; Kanaya et al. 20011; although see Chamary et al. 2006). Jia and Higgs (2008) found that mitochondrial codon usage evolution is dominated by mutational effects.

Fourthly, as was previously mentioned, there could be a relationship between the k parameter and N_e across species. We make the assumption that the strength of selection, included in k, and the shape of the DFE remain constant as N_e changes; however, this is unlikely to be the case (although there is evidence to suggest that the strength of selection can be nearly independent of N_e in some evolutionary scenarios (Charlesworth 2013)). It may be that generally there is a negative relationship between k and N_e across the species in our dataset: this would result in making the slope of the relationship steeper than that predicted by the shape parameter of the gamma distribution.

Finally, hitchhiking may have an important influence on the relationship between genetic diversity and the efficiency of selection. It has been demonstrated across a broad range of species that hitchhiking and background selection remove more neutral diversity in species with larger census population sizes (and hence larger effective population sizes) (Corbett-Detig et al. 2015). Hitchhiking will result in a loss of genetic diversity and a reduction in N_e at linked sites. However, not all types of site will be affected equally by hitchhiking. Deleterious variants segregate at low frequencies in populations, and therefore are expected to reach their equilibrium frequencies relatively rapidly after a linked selection event. Neutral and advantageous variants on the other hand segregate at higher frequencies in populations, and so linked selection will result in a proportionally smaller loss of diversity for deleterious than for neutral or advantageous sites. Because selection on linked sites reduces both the efficiency of selection and the level of neutral genetic diversity, we expect the overall effect to be a steeper slope of the relationship between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ than that predicted by the shape parameter of the DFE, which is the average taken across all species. This is perhaps the most likely explanation of our results: mtDNA undergoes minimal recombination, and as such selection on linked sites will have a large impact on molecular evolution, increasing the slope between $\log(\pi_N/\pi_S)$ and $\log(\pi_S)$ (Birky & Walsh 1988; Castellano et al. 2015). Furthermore, it has been recently shown that mtDNA undergoes adaptive evolution in animals (James et al. 2016).

Chapter 6

6.1 General Discussion

In this thesis I have examined a number of issues relating to variation in the effective population size across animal species, in order to better understand factors that are predictive of N_e and to explore the effects of N_e on molecular evolution. Here I will briefly outline the avenues of enquiry pursued in this thesis, followed by a discussion of the results from each chapter. I will then go on to give a general overview of the work presented in this thesis, and discuss the implications of the results and a number of lines of enquiry for future work.

In chapters 2 and 3, I focused on uncovering the life history and demographic factors that are predictive of two molecular traits that are correlated to N_e, neutral genetic diversity and the efficiency of selection, using two different methodologies, both of which correct for phylogenetic non-independence. In chapter 2, I used an island-mainland system, i.e., I compared the molecular evolution of island species to that of their closely related mainland counterparts. In chapter 3 I investigated whether molecular evolutionary factors are correlated to a wide range of life history factors, using a much larger dataset of mammalian mitochondrial DNA. While I have found evidence to suggest that a species' range is correlated to its effective population size, in general the relationships that I detected are surprisingly weak. This is particularly evident in my analysis of island species; while island species are affected by bottlenecks, their molecular evolution is not strongly affected by their restricted range sizes. In chapter 4, I investigated the role of adaptive evolution in mammalian mitochondria, and developed a new method with which to estimate the proportion of adaptive substitutions that have occurred. I also considered whether rates of adaptive evolution are correlated to N_e in animal mitochondria, and whether a relationship exists between genetic diversity and the efficiency of selection in mitochondria. I found that mitochondrial diversity does indeed correlate to N_e. I greatly extended this analysis in chapter 5. I quantified the relationship between neutral sequence diversity and the efficiency of selection, two important proxies of N_e, in mammalian mitochondrial DNA. I have demonstrated that there is a strong correlation between these two variables, as predicted by the nearly neutral theory.

6.1.1 The effect of island colonisation on molecular evolution

In this chapter, we compared levels of genetic diversity and the efficiency of selection in island endemic species to those of their closely related mainland counterparts, assuming that island species, which we have found to have range sizes that are on average 1000 times smaller than those of mainland species, have small census population sizes and therefore small effective population sizes. This approach has two benefits: firstly, our paired analysis controls for phylogenetic non-independence, and secondly, because island and mainland species are closely related, they are likely to have similar life history traits.

Therefore, we can avoid the confounding effects of variation in life history in our analysis, and any variations we observe in molecular evolution between island and mainland species are likely to be due to differences in N_e. In order to conduct our analyses, we have compiled the first polymorphism dataset for island-mainland comparisons. We found that while island species do have significantly lower levels of genetic diversity than mainland species, the difference is relatively small. This result appears to be largely driven by young island species that have only recently diverged from their mainland counterparts. These species have little or no diversity, indicating the initial colonisation of an island does result in an extreme population bottleneck. However, our results indicate that older island species have similar levels of diversity to their mainland counterparts, and possibly also similar effective population sizes. Consistent with this, we find no significant differences between island and mainland species in terms of the efficiency of selection. This result suggests that an island colonisation event has little impact on the molecular evolution of island endemic species, despite their smaller ranges, and presumably smaller effective population sizes. It is surprising that N_c has so little effect on N_e.

6.1.2 Molecular evolution and life history

In chapter 3, we reinvestigated the question of whether the global range of a species is predictive of its N_e, using a considerably larger dataset of mammalian mitochondrial and nuclear DNA. In this analysis, we also consider whether other traits, including body mass, longevity and latitude, might also be predictive of N_e. In this dataset, we do find a positive correlation between range size and levels of neutral genetic diversity, and between range size and the efficiency of selection. This tallies with our expectations that species with larger range sizes have larger effective population sizes. In addition, we also detect an effect of latitude and mass on neutral genetic diversity, which could reflect a latitudinal gradient in mutation rate such that equatorial species tend to have higher mutation rates. Alternatively, this result could be interpreted as the influence of population density on N_e, assuming that larger, higher latitude species tend to have lower population densities than smaller, tropical species. While the effects we detect are significant, they are quite weak, as the regression slope is quite shallow and species range explains only a small proportion of the variance in genetic diversity and the efficiency of selection. This somewhat supports the results we describe in chapter 2. Our results do however differ from a number of past studies: this may be because our dataset is relatively taxonomically restricted and is also larger than datasets used in previous studies, and additionally we control for phylogeny in our analysis. We hypothesise that in a taxonomically restricted dataset, global species range is a fairly good predictor of relative census population sizes, whereas in a dataset which covers a very broad taxonomic range species are likely to vary substantially in factors such as population density, which could confound a relationship between range and census population size, and therefore also between range and N_e.

6.1.3 Adaptive evolution in animal mitochondria

In chapter 4 we investigate the evolutionary dynamics of animal mitochondria, using a variety of McDonald-Kreitman style tests. We find that mitochondrial evolution is dominated by slightly deleterious mutations, which can obscure the signal of adaptive evolution. We therefore developed a method for detecting adaptive mutations that controls for the presence of deleterious mutations. After applying this

method, we find that there has been significant adaptive evolution in mitochondrial DNA, with an estimated 26% of nonsynonymous mutations being fixed by positive selection. The level of adaptive evolution we observe in mitochondria, while significant, is quite low. This is consistent with the fact that mitochondria experience little or no recombination, which has previously been shown to limit the rate of adaptive evolution. This suggests that animal mitochondria experience strong selection compared to nuclear regions with similarly low rates of recombination, which is likely to be because these organelles encode a number of crucial genes in the ATP production pathway. We also detect a weak positive correlation between mitochondrial adaptive evolution and synonymous neutral diversity, a proxy for N_e, which indicates that adaptive evolution in mitochondria is affected by N_e. This highlights the role of N_e in determining the rate of adaptive evolution, and also suggests that adaptive evolution is limited by the supply of mutations. Our results have some important practical implications, because mitochondrial DNA is commonly used as a neutral molecular marker, under the assumption that mitochondrial diversity is solely affected by drift, and as such will reflect only the demographic processes experienced by the population. In fact, mitochondrial DNA experiences considerable adaptive evolution, and so mitochondrial diversity will be at least partly determined by the amount of time since the mitochondrial genome experienced a selective sweep.

6.1.4 Quantifying the relationship between neutral diversity and the efficiency of selection

Measures of genetic diversity are commonly used proxies of N_e, however, the extent to which they are related to measures of the efficiency of selection acting in a population has rarely been investigated. We address this question by exploring and quantifying the relationship between π_S, a measure of neutral genetic diversity, and π_N/π_S, a measure of the efficiency of selection, after first removing the non-independence between these two variables, using a dataset of mammalian mitochondrial DNA. We find that they are strongly, loglinearly correlated, a result which has subsequently been found to hold across a wide range of animal and plant species (Chen et al. 2017). We find that in animal mitochondria, as diversity doubles, the number of nonsynonymous variants segregating in the population is reduced by about a third. The slope of the relationship is affected by life history traits- we find that the slope is greater in species with high mass-specific metabolic rates, and in species with low body mass. This is likely to be because these traits are negatively correlated: species with low body mass have higher mass-species metabolic rates. We also observe variation in the slope of the relationship across taxa, with the slope of the line between π_S and π_N/π_S being steeper in rodents than in bats. The differences we observe might be due to a difference in the distribution of fitness effects (DFE) between these groups. This seems to be the case for species with different life history traits: while differences between the DFE between the groups are not always significant, they mirror the differences we observe in the slopes between these groups. However, there is no difference between the DFE of bats and rodents. The difference between the slopes that we observe between these taxonomic groups might instead suggest that the relationship between N_e and the strength of selection differs between taxonomic groups. While our results support the nearly neutral theory, they do not support the theoretical prediction that the slope of the relationship between $\log(\pi_S)$ and $\log(\pi_N/\pi_S)$ is the same as the shape parameter of the DFE: we estimate the line to be

significantly steeper than we would expect. Although it is not directly addressed by the authors, this pattern is also evident in the results of Chen et al. (2017), indicating that it holds across a range of eukaryote species. We also find that the slope of the relationship between $\log(\pi_S)$ and $\log(\pi_N/\pi_S)$ is steeper than we would expect from the shape parameter of the DFE within the *Drosophila* genome (Castellano et al., in prep).

6.2 Overview and Perspectives

Our research has highlighted a number of interesting patterns in molecular evolution, which are largely in line with the predictions of nearly neutral theory. N_e is a crucial population genetics parameter, important in determining both the level of standing genetic diversity and the efficiency of selection in a population, clearly demonstrated by the loglinear relationship between these variables discussed in chapter 5. However, the determinants of N_e are still unclear. We find, as others have (Lewontin 1974), that life history and demographic traits expected to be proxies of N_c and thus N_e are surprisingly poor predictors of molecular evolutionary traits. While we do find a relationship between range size and both genetic diversity and the efficiency of selection using a large dataset of mammalian mitochondrial sequences (see chapter 3), our analyses of variables that might influence molecular evolution are somewhat inconclusive. We found that the regression slopes for the relationships between molecular evolutionary traits and species ranges are shallow: life history traits generally explain a low proportion of the variance in both genetic diversity, π_S, (approximately 11%) and the efficiency of selection, $\pi_N/(\pi_N+\pi_S)$, (approximately 4%). Overall, the data is noisy: it may be that to detect a relationship between life history and demographic traits and N_e, a large dataset is needed.

The 'noisiness' of the data may be why there is some discrepancy between our results from chapter 2 and chapter 3. In chapter 2, we found little evidence to suggest that island species varied strongly in their molecular evolution from mainland species: crucially, mainland species did not have more efficient selection than island species, which suggests that the range of a species does not strongly impact its N_e. However, in chapter 3 we found that range did have a significant effect on molecular evolution, both in terms of neutral genetic diversity and the efficiency of selection. We may simply lack power in our island-mainland species dataset to detect relationships between demographic traits and molecular evolution. Alternatively, it is possible that there is a biological explanation for the discrepancy between our results for chapters 2 and 3. Island species do appear to differ from even closely related mainland species in their population densities, with island species having higher population densities, which is theorised to occur because island species have fewer competitors (Adler & Levins 1994; Buckley & Jetz 2007; Crowell 2017). However, the densities of island species would have to be in the order of thousands of times higher in order for island and mainland populations to have equal census population sizes, and so while this effect might reduce the strength of a relationship between N_e and range size, it is still somewhat surprising that we do not see a considerable difference in the molecular evolution of island and mainland species.

More generally, we were not able to determine why the relationship between N_c and N_e is weak, and by extension why life history and demographic traits are only poorly predictive of π_S, and $\pi_N/(\pi_N+\pi_S)$. This may be due to a number of other factors that are known to act on molecular evolutionary traits. One possible confounding factor is the mutation rate. Variation in mutation rate could exist between species and taxonomic groups: if mutation rates are higher in species with low effective population sizes, as suggested by Lynch (2010), it would obscure a relationship between N_e and π_S. However, we do not find any evidence to suggest variation in mutation rates between high N_e (mainland) and low N_e (island) species in chapter 2. Alternatively, other factors could affect the mutation rate; for example, our results suggest there could be a latitudinal gradient in mutation rate, with low latitude species having higher mutation rates (see chapter 3). The role of the mutation rate in determining levels of diversity is unclear, particularly as there are so few direct estimates of the mutation rate available (see Table 1.1 for all species for which there are direct estimates). As more estimates become available, the role of the mutation rate in shaping patterns of genetic diversity is likely to become clearer. A second confounding factor is the DFE: the molecular evolution of a species will be due to both its N_e and its DFE, and our results do suggest that there is some variation in the DFE across species. An interesting line of future research might be to develop models and methods that examine the joint effects of both of these factors on molecular evolution.

Another possible confounding factor is the action of linked selection. We expect linked selection to have a strong impact on molecular evolution in organelle genomes, which do not recombine. We explore the effects of linked selection on animal mitochondria in some detail in chapter 4: we find evidence for both purifying selection and substantial adaptive evolution, which could have the effect of obscuring a relationship between neutral genetic diversity and N_e. However, the effects of linked selection are not sufficient to completely remove any relationship between π_S and N_e, as we find a strong relationship between π_S and the efficiency of selection, and π_S and the rate of adaptive evolution. Linked selection might also be important to our understanding of the relationship between π_S and π_N/π_S. In chapter 5 we find that the regression slope for the relationship between neutral genetic diversity, π_S, and the efficiency of selection, π_N/π_S, is steeper than we would predict from the shape parameter of the gamma distribution of fitness effects. Further simulations could clarify whether this is a result of selective sweeps, or of other evolutionary and/or demographic factors. Initial simulation results suggest that hitch-hiking has a far greater impact on neutral diversity than on deleterious diversity, which can potentially explain why the slope of the relationship between π_N/π_S and π_S is steeper than expected.

It is also important to note that while we interpret variation in π_S and π_N/π_S as a result of variation in N_e across species, it is possible that nonequilibrium dynamics could also have a role in shaping this pattern. For example, if a population experiences a bottleneck, we expect both its neutral and its deleterious genetic diversity to be greatly reduced. As the population recovers in size, it will recover genetic diversity. However, we expect deleterious diversity to recover more rapidly than neutral diversity, simply because levels of deleterious diversity will have been lower initially, i.e., the population needs to accumulate fewer deleterious mutations than neutral mutations to recover its level of equilibrium

diversity. Therefore, if we sample individuals from a population that is not in equilibrium, which is still in the process of recovering diversity, we might observe a pattern of polymorphisms similar to the one we observe, in that π_N/π_S will be greater than we would expect for any given value of π_S, and in addition this pattern would not necessarily be a result of variation in the long term N_e of populations (Brandvain & Wright 2016). Identifying the role of nonequilibrium dynamics in shaping molecular evolution could perhaps be investigated by studying populations for which we have good demographic history data, or alternatively by using simulations.

Bibliography

Adler, G.H. & Levins, R., 1994. The Island Syndrome in Rodent Populations. *The Quarterly Review of Biology*, 69(4), pp.473–490.

Andolfatto, P., 2007. Hitchhiking effects of recurrent beneficial amino acid substitutions in the Drosophila melanogaster genome. *Genome Research*, 17(12), pp.1755–1762.

Andolfatto, P. & Przeworski, M., 2001. Regions of lower crossing over harbor more rare variants in African populations of Drosophila melanogaster. *Genetics*, 158(2), pp.657–665.

Arbeithuber, B. et al., 2015. Crossovers are associated with mutation and biased gene conversion at recombination hotspots. *Proceedings of the National Academy of Sciences of the United States of America*, 112(7), pp.2109–14.

Asthana, S., Schmidt, S. & Sunyaev, S., 2005. A limited role for balancing selection. *Trends in Genetics*, 21(1), pp.30–32.

Bachtrog, D., 2003. Adaptation shapes patterns of genome evolution on sexual and asexual chromosomes in Drosophila. *Nature Genetics*, 34(2), pp.215–219.

Bachtrog, D. & Andolfatto, P., 2006. Selection, recombination and demographic history in Drosophila miranda. *Genetics*, 174(4), pp.2045–2059.

Bachtrog, D. & Charlesworth, B., 2002. Reduced adaptation of a non-recombining neo-Y chromosome. *Nature*, 416(6878), pp.323–326.

Ballard, J.W.O. & Kreitman, M., 1994. Unraveling selection in the mitochondrial genome of Drosophila. *Genetics*, 138(3), pp.757–772.

Ballard, J.W.O. & Rand, D.M., 2005. The Population Biology of Mitochondrial DNA and Its Phylogenetic Implications. *Annual Review of Ecology, Evolution, and Systematics*, 36, pp.621–642.

Ballard, J.W.O. & Whitlock, M.C., 2004. The incomplete natural history of mitochondria. *Molecular Ecology*, 13(4), pp.729–744.

Barrett, R.D.H. & Schluter, D., 2007. Adaptation from standing genetic variation. *Trends in ecology & evolution*, 23(1), pp.38–44.

Barrio, A.M. et al., 2016. The genetic basis for ecological adaptation of the Atlantic herring revealed by genome sequencing. *eLife*, 5(e12081), pp.1–32.

Bazin, E. et al., 2005. Polymorphix: A sequence polymorphism database. *Nucleic Acids Research*, 33(Database Issue), pp.481–484.

Bazin, E., Glémin, S. & Galtier, N., 2006. Population size does not influence mitochondrial genetic diversity in animals. *Science*, 312(5773), pp.570–2.

Begun, D.J. & Aquadro, C.F., 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature*, 356(6369), pp.519–520.

Berlin, S. & Ellegren, H., 2006. Fast accumulation of nonsynonymous mutations on the female-specific W chromosome in birds. *Journal of Molecular Evolution*, 62(1), pp.66–72.

Berlin, S., Tomaras, D. & Charlesworth, B., 2007. Low mitochondrial variability in birds may indicate Hill-Robertson effects on the W chromosome. *Heredity*, 99(4), pp.389–396.

Betancourt, A.J. & Presgraves, D.C., 2002. Linkage limits the power of natural selection in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, 99(21), pp.13616–13620.

Betancourt, A.J., Welch, J.J. & Charlesworth, B., 2009. Reduced Effectiveness of Selection Caused by a Lack of Recombination. *Current Biology*, 19(8), pp.655–660.

Birky, C.W., 1995. Uniparental inheritance of mitochondrial and chloroplast genes: mechanisms and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 92(25), pp.11331–11338.

Birky, C.W. & Walsh, J.B., 1988. Effects of linkage on rates of molecular evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 85(17), pp.6414–6418.

Boissinot, S. et al., 2014. Neutral Nuclear Variation in Baboons (genus Papio) Provides Insights into their Evolutionary and Demographic Histories. *Am J Phys Anthropol.*, 155(4), pp.621–634.

Boore, J.L., 1999. Animal mitochondrial genomes. *Nucleic Acids Research*, 27(8), pp.209–210.

Boyko, A.R. et al., 2008. Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genetics*, 4(5), p.e1000083.

Brandvain, Y. & Wright, S.I., 2016. The Limits of Natural Selection in a Nonequilibrium World. *Trends in Genetics*, 32(4), pp.201–210.

Briskie, J. V & Mackintosh, M., 2003. Hatching failure increases with severity of population bottlenecks in birds. *PNAS*, 101(2), pp.558–561.

Bromham, L., 2011. The genome as a life-history character: why rate of molecular evolution varies between mammal species. *Phil. Trans. R. Soc. B*, 366, pp.2503–13.

Bromham, L. & Cardillo, M., 2003. Testing the link between the latitudinal gradient in species richness and rates of molecular evolution. *Journal of Evolutionary Biology*, 16(2), pp.200–207.

Bromham, L., Rambaut, a & Harvey, P., 1996. Determinants of rate variation in mammalian DNA sequence evolution. *Journal of Molecular Evolution*, 43(6), pp.610–621.

Brown, W.M., George, M. & Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 76(4), pp.1967–71.

Buckley, L.B. & Jetz, W., 2007. Insularity and the determinants of lizard population density. *Ecology Letters*, 10(6), pp.481–489.

Bullaughey, K., Przeworski, M. & Coop, G., 2008. No effect of recombination on the efficacy of natural selection in primates. *Genome Research*, 18(4), pp.544–554.

Burgess, S.C., Waples, R.S. & Baskett, M.L., 2013. Local adaptation when competition depends on phenotypic similarity. *Evolution*, 67(10), pp.3012–22.

Burri, R. et al., 2015. Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of Ficedula flycatchers. *Genome Research*, 25(11), pp.1656–1665.

Campos, J.L. et al., 2014. The relation between recombination rate and patterns of molecular evolution and variation in Drosophila melanogaster. *Molecular Biology and Evolution*, 31(4), pp.1010–1028.

Campos, J.L., Charlesworth, B. & Haddrill, P.R., 2012. Molecular evolution in nonrecombining regions of the Drosophila melanogaster genome. *Genome Biology and Evolution*, 4(3), pp.278–288.

Cao, K. et al., 2014. Comparative population genomics reveals the domestication history of the peach, Prunus persica, and human influences on perennial fruit crops. *Genome Biology*, 15(7), p.415.

Carlini, D.B. & Genut, J.E., 2006. Synonymous SNPs provide evidence for selective constraint on human exonic splicing enhancers. *Journal of Molecular Evolution*, 62(1), pp.89–98.

Castellana, S., Vicario, S. & Saccone, C., 2011. Evolutionary patterns of the mitochondrial genome in Metazoa: Exploring the role of mutation and selection in mitochondrial protein-coding genes. *Genome Biology and Evolution*, 3(1), pp.1067–1079.

Castellano, D. et al., 2015. Adaptive Evolution Is Substantially Impeded by Hill–Robertson Interference in Drosophila. *Molecular Biology and Evolution*, 33(2), pp.442–455.

Castoe, T. a. et al., 2008. Adaptive evolution and functional redesign of core metabolic proteins in snakes. *PLoS ONE*, 3(5), p.e2201.

Chamary, J. V. & Hurst, L.D., 2005. Evidence for selection on synonymous mutations affecting stability of mRNA secondary structure in mammals. *Genome Biology*, 6(9), p.R75.

Chamary, J. V, Parmley, J.L. & Hurst, L.D., 2006. Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nature reviews. Genetics*, 7(2), pp.98–108.

Charlesworth, B., 1996. Background selection and patterns of genetic diversity in Drosophila melanogaster. *Genetical Research*, 68(2), pp.131–150.

Charlesworth, B., 2009. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nature reviews. Genetics*, 10(3), pp.195–205.

Charlesworth, B., 2013. Stabilizing selection, purifying selection, and mutational bias in finite populations. *Genetics*, 194(4), pp.955–71.

Charlesworth, B., 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genetical Research*, 63(3), pp.213–227.

Charlesworth, B. & Charlesworth, D., 1987. Inbreeding Depression and its Evolutionary Consequences. *Ann. Rev. Ecol. Syst*, 18, pp.237–268.

Charlesworth, B., Morgan, M.T. & Charlesworth, D., 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics*, 134(4), pp.1289–1303.

Charlesworth, J. & Eyre-Walker, A., 2008. The McDonald-Kreitman test and slightly deleterious mutations. *Molecular biology and evolution*, 25(6), pp.1007–15.

Charlesworth, J. & Eyre-Walker, A., 2007. The other side of the nearly neutral theory, evidence of slightly advantageous back-mutations. *PNAS*, 104(43), pp.16992–7.

Charlesworth, J. & Eyre-Walker, A., 2006. The rate of adaptive evolution in enteric bacteria. *Molecular Biology and Evolution*, 23(7), pp.1348–1356.

Chen, J., Glemin, S. & Lascoux, M., 2017. Genetic diversity and the efficacy of purifying selection across plant and animal species. *Mol. Biol. Evol.*, 34(6), pp.1417–1428.

Coop, G., 2016. Does linked selection explain the narrow range of genetic diversity across species? *bioRxiv*, p.42598.

Corbett-Detig, R.B., Hartl, D.L. & Sackton, T.B., 2015. Natural Selection Constrains Neutral Diversity across A Wide Range of Species. *PLOS Biology*, 13(4), p.e1002112.

Crowell, K., 2017. Reduced Interspecific Competition among the Birds of Bermuda. *Ecology*, 43(1), pp.75–88.

Cutter, A.D., 2008. Multilocus patterns of polymorphism and selection across the X chromosome of Caenorhabditis remanei. *Genetics*, 178(3), pp.1661–1672.

Cutter, A.D., Jovelin, R. & Dey, A., 2013. Molecular hyperdiversity and evolution in very large populations. *Molecular Ecology*, 22(8), pp.2074–2095.

Cutter, A.D. & Payseur, B.A., 2013. Genomic signatures of selection at linked sites: unifying the disparity among species. *Nature Reviews Genetics*, 14(4), pp.262–74.

Delph, L.F. & Kelly, J.K., 2014. On the importance of balancing selection in plants. *New Phytol.*, 201(1), pp.45–56.

Denver, D.R. et al., 2000. High direct estimate of the mutation rate in the mitochondrial genome of Caenorhabditis elegans. *Science*, 289(5488), pp.2342–2344.

Doan, J.W. et al., 2004. Coadaptive evolution in cytochrome c oxidase: 9 of 13 subunits show accelerated rates of nonsynonymous substitution in anthropoid primates. *Molecular Phylogenetics and Evolution*, 33(3), pp.944–950.

Dowle, E.J., Morgan-Richards, M. & Trewick, S.A., 2013. Molecular evolution and the latitudinal biodiversity gradient. *Heredity*, 110(6), pp.501–510.

Dowling, D.K., Abiega, K.C. & Arnqvist, G., 2007. Temperature-specific outcomes of cytoplasmic-nuclear interactions on egg-to-adult development time in seed beetles. *Evolution*, 61(1), pp.194–201.

Duret, L., 2002. Evolution of synonymous codon usage in metazoans. *Current opinion in genetics & development*, 12(6), pp.640–649.

Duret, L. & Mouchiroud, D., 1999. Expression pattern and, surprisingly, gene length shape codon usage in Caenorhabditis, Drosophila, and Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 96(8), pp.4482–4487.

Egea, R. et al., 2007. MamPol: A database of nucleotide polymorphism in the Mammalia class. *Nucleic Acids Research*, 35(Database issue), pp.624–629.

Ellegren, H. & Galtier, N., 2016. Determinants of genetic diversity. *Nature Reviews Genetics*, 17(7), pp.422–433.

Eory, L., Halligan, D.L. & Keightley, P.D., 2010. Distributions of selectively constrained sites and deleterious mutation rates in the hominid and murid genomes. *Molecular Biology and Evolution*, 27(1), pp.177–192.

ESRI, 2011. ArcGIS Desktop.

Eyre-Walker, A. et al., 2002. Quantifying the slightly deleterious mutation model of molecular evolution. *Molecular biology and evolution*, 19(12), pp.2142–9.

Eyre-Walker, A., 2006. The genomic rate of adaptive evolution. *Trends in ecology & evolution*, 21(10), pp.569–75.

Eyre-Walker, A. & Keightley, P.D., 2009. Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. *Molecular Biology and Evolution*, 26(9), pp.2097–2108.

Eyre-Walker, A. & Keightley, P.D., 2007. The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, 8(8), pp.610–618.

Eyre-Walker, A., Woolfit, M. & Phelps, T., 2006. The distribution of fitness effects of new deleterious amino acid mutations in humans. *Genetics*, 173(2), pp.891–900.

Fan, W. et al., 2008. A Mouse Model of Mitochondrial Disease Reveals Germline Selection Against Severe mtDNA Mutations. *Science*, 319(5865), pp.958–962.

Fay, J.C., Wyckoff, G.J. & Wu, C., 2001. Positive and Negative Selection on the Human Genome. *Genetics*, 158(3), pp.1227–1234.

Felsenstein, J., 1985. Phylogenies and the Comparative Method. *The American Naturalist*, 125(1), pp.1–15.

Feng, C. et al., 2017. Moderate nucleotide diversity in the Atlantic herring is associated with a low mutation rate. *eLife*, 6, pp.1–14.

Figuet, E. et al., 2016. Life History Traits, Protein Evolution, and the Nearly Neutral Theory in Amniotes. *Molecular Biology and Evolution*, 33(6), pp.1517–1527.

Fischer, A.G., 1960. Latitudinal Variations in Organic Diversity. *Source: Evolution*, 14(1), pp.64–81.

Flowers, J.M. et al., 2012. Natural selection in gene-dense regions shapes the genomic pattern of polymorphism in wild and domesticated rice. *Molecular Biology and Evolution*, 29(2), pp.675–687.

da Fonseca, R.R. et al., 2008. The adaptive evolution of the mammalian mitochondrial genome. *BMC genomics*, 9(119), p.10.1186/1471-2164-9-119.

Francioli, L.C. et al., 2015. Genome-wide patterns and properties of de novo mutations in humans. *Nature Genetics*, 47(7), pp.822–826.

Frankham, R., 1997. Do island populations have less genetic variation than mainland populations? *Heredity*, 78, pp.311–27.

Frankham, R. et al., 1999. Do population size bottlenecks reduce evolutionary potential? *Animal Conservation*, 2, pp.255–260.

Frankham, R., 1995. Effective population size / adult population size ratios in wildlife : a review. *Genet. Res.*, 66, pp.95–107.

Frankham, R., 2012. How closely does genetic diversity in finite populations conform to predictions of neutral theory? Large deficits in regions of low recombination. *Heredity*, 108, pp.167–78.

Frankham, R., 1996. Relationship of Genetic Variation to Population Size in Wildlife. *Conservation biology*, 10(6), pp.1500–1508.

Freckleton, R.P., Harvey, P.H. & Pagel, M., 2002. Phylogenetic Analysis and Comparative Data : A Test and Review of Evidence. *The American Naturalist*, 160(6), pp.712–726.

Galtier, N., 2015. Adaptive Protein Evolution in Animals and the Effective Population Size Hypothesis. *PLoS genetics*, 12(1), p.e1005774.

Galtier, N. et al., 2009. Mitochondrial DNA as a marker of molecular diversity: A reappraisal. *Molecular Ecology*, 18(22), pp.4541–4550.

Gao, Z. et al., 2016. Interpreting the dependence of mutation rates on age and time. *PLoS Biology*, 14(1), p.e1002355.

Gaston, K.J., 2000. Global patterns in biodiversity. *Nature*, 405, pp.220–227.

Gillespie, J.H., 2000. Genetic Drift in an Infinite Population : The Pseudohitchhiking Model. *Genetics*, 155, pp.909–919.

Gillespie, J.H. & Ohta, T., 1996. Development of neutral and nearly neutral theories. *Theor. Pop. Biol.*, 49(2), pp.128–148.

Gillman, L.N. et al., 2009. Latitude, elevation and the tempo of molecular evolution in mammals. *Proceedings of the Royal Society B: Biological Sciences*, 276(1671), pp.3353–3359.

Goldman, N. & Yang, Z., 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Molecular biology and evolution*, 11(5), pp.725–736.

Gossmann, T.I. et al., 2010. Genome wide analyses reveal little evidence for adaptive evolution in many plant species. *Molecular Biology and Evolution*, 27(8), pp.1822–1832.

Gossmann, T.I. et al., 2014. Highly variable recombinational landscape modulates efficacy of natural selection in birds. *Genome biology and evolution*, 6(8), pp.2061–2075.

Gossmann, T.I., Keightley, P.D. & Eyre-Walker, A., 2012. The effect of variation in the effective population size on the rate of adaptive molecular evolution in eukaryotes. *Genome biology and evolution*, 4(5), pp.658–67.

Gossmann, T.I., Woolfit, M. & Eyre-Walker, A., 2011. Quantifying the variation in the effective population size within a genome. *Genetics*, 189, pp.1389–402.

Grossman, L.I. et al., 2004. Accelerated evolution of the electron transport chain in anthropoid primates. *Trends in Genetics*, 20(11), pp.578–585.

Haddrill, P.R. et al., 2007. Reduced efficacy of selection in regions of the Drosophila genome that lack crossing over. *Genome biology*, 8(2), p.R18.

Haldane, J.B.S., 1937. The effect of variation on fitness. *The American Naturalist*, 71(735), pp.337–349.

Harvey, P.H. & Pagel, M., 1991. *The comparative method in evolutionary biology.*, Oxford: Oxford University Press.

Havird, J.C. & Sloan, D.B., 2016. The roles of mutation, selection, and expression in determining relative rates of evolution in mitochondrial vs. nuclear genomes. *Molecular Biology and Evolution*, 33(12), pp.3042–3053.

Hedges, S.B., Dudley, J. & Kumar, S., 2006. TimeTree: A public knowledge-base of divergence times among organisms. *Bioinformatics*, 22(23), pp.2971–2972.

Hellmann, I. et al., 2003. A Neutral Explanation for the Correlation of Diversity with Recombination Rates in Humans. *The American Journal of Human Genetics*, 72(6), pp.1527–1535.

Hermisson, J. & Pennings, P.S., 2005. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics*, 169, pp.2335–52.

Hernandez, R.D., 2008. A flexible forward simulator for populations subject to selection and demography. *Bioinformatics*, 24(23), pp.2786–2787.

Hershberg, R. & Petrov, D.A., 2008. Selection on codon bias. *Annual review of genetics*, 42, pp.287–299.

Hickey, a J.R., 2008. An alternate explanation for low mtDNA diversity in birds: an age-old solution? *Heredity*, 100(5), p.443.

Hillebrand, H., 2008. On the Generality of the Latitudinal Diversity Gradient. *Evaluation*, 163(2), pp.2005–2008.

Ho, S.Y.W. et al., 2011. Time-dependent rates of molecular evolution. *Molecular ecology*, 20(15), pp.3087–101.

Hodgkinson, A. & Eyre-Walker, A., 2011. Variation in the mutation rate across mammalian genomes. *Nature Reviews Genetics*, 12(11), pp.756–766.

Hugall, A.F. & Lee, M.S.Y., 2007. The likelihood node density effect and consequences for evolutionary studies of molecular rates. *Evolution*, 61(10), pp.2293–307.

Hughes, A.L. & Nei, M., 1989. Nucleotide substitution at major histocompatibility complex class II loci: vidence for overdominant selection. *Proceedings of the National Academy of Sciences of the United States of merica*, 86(3), pp.958–962.

Hughes, A.L. & Nei, M., 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature*, 335(6186), pp.167–170.

Hurst, G.D.D. & Jiggins, F.M., 2005. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proceedings. Biological sciences / The Royal Society*, 272(1572), pp.1525–1534.

IUCN, 2014. IUCN Red List of Threatened Species. Available at: www.iucnredlist.org [Accessed September 1, 2014].

James, J.E., Lanfear, R. & Eyre-Walker, A., 2016. Molecular Evolutionary Consequences of Island Colonization. *Genome Biology and Evolution*, 8(6), pp.1876–1888.

James, J.E., Piganeau, G. & Eyre-Walker, A., 2016. The rate of adaptive evolution in animal mitochondria. *Molecular Ecology*, 25(1), pp.67–78.

Jeschke, J.M. & Kokko, H., 2009. The roles of body size and phylogeny in fast and slow life histories. *Evolutionary Ecology*, 23(6), pp.867–878.

Jia, W. & Higgs, P.G., 2008. Codon usage in mitochondrial genomes: Distinguishing context-dependent mutation from translational selection. *Molecular Biology and Evolution*, 25(2), pp.339–351.

Johnson, K.P. & Seger, J., 2001. Elevated Rates of Nonsynonymous Substitution in Island Birds. *Molecular biology and evolution*, 18(5), pp.874–881.

Johnson, N.A. & Lachance, J., 2012. The genetics of sex chromosomes: evolution and implications for hybrid incompatibility. *Ann N Y Acad Sci*, 1256, pp.E1–E22.

Johnson, T.H. & Stattersfield, A.J., 1990. A global review of island endemic birds. *Ibis*, 132(2), pp.167–180.

Jones, K.E. et al., 2003. Biological Correlates of Extinction Risk in Bats. *The American Naturalist*, 161(4), pp.601–614.

Jones, K.E. et al., 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals. *Ecology*, 90(9), p.2648.

Kamilar, J.M. & Cooper, N., 2013. Phylogenetic signal in primate behaviour, ecology and life history. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1618), pp.20120341–20120341.

Kanaya, S. et al., 2011. Codon Usage and tRNA Genes in Eukaryotes: Correlation of Codon Usage Diversity with Translation Efficiency and with CG-Dinucleotide Usage as Assessed by Multivariate Analysis. *Journal of Molecular Evolution*, 53(4), pp.290–298.

Kearse, M. et al., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28(12), pp.1647–1649.

Keck, F. et al., 2016. Phylosignal: An R package to measure, test, and explore the phylogenetic signal. *Ecology and Evolution*, 6(9), pp.2774–2780.

Keightley, P.D. & Eyre-Walker, A., 2012. Estimating the rate of adaptive molecular evolution when the evolutionary divergence between species is small. *Journal of Molecular Evolution*, 74(1–2), pp.61–68.

Keightley, P.D. & Eyre-Walker, A., 2007. Joint inference of the distribution of fitness effects of deleterious mutations and population demography based on nucleotide polymorphism frequencies. *Genetics*, 177(4), pp.2251–2261.

Kimura, M., 1962. On the probability of fixation of mutant genes in a population. *Genetics*, 47(391), pp.713–719.

Kimura, M., 1984. *The neutral theory of molecular evolution*, Cambridge: Cambridge University Press.

Kimura, M. & Maruyama, T., 1966. The mutational load with epistatic gene interactions in fitness. *Genetics*, 54(6), pp.1337–1351.

Kousathanas, A. & Keightley, P.D., 2013. A comparison of models to infer the distribution of fitness effects of new mutations. *Genetics*, 193(4), pp.1197–1208.

Kryazhimskiy, S. & Plotkin, J.B., 2008. The population genetics of dN/dS. *PLoS genetics*, 4(12), p.e1000304.

Künstner, A. et al., 2010. Comparative genomics based on massive parallel transcriptome sequencing reveals patterns of substitution and selection across 10 bird species. *Molecular Ecology*, 19(SUPPL. 1), pp.266–276.

Künstner, A. et al., 2016. The genome of the trinidadian guppy, Poecilia reticulata, and variation in the Guanapo population. *PLoS ONE*, 11(12), pp.1–25.

Lanfear, R. et al., 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular biology and evolution*, 29(6), pp.1695–1701.

Lanfear, R. et al., 2013. Taller plants have lower rates of molecular evolution. *Nature Communications*, 4(1879).

Lanfear, R., Kokko, H. & Eyre-Walker, A., 2014. Population size and the rate of evolution. *Trends in ecology & evolution*, 29(1), pp.33–41.

Lanfear, R., Welch, J.J. & Bromham, L., 2010. Watching the clock: studying variation in rates of molecular evolution between species. *Trends in ecology & evolution*, 25(9), pp.495–503.

Lattorff, H.M.G. et al., 2016. Effective population size as a driver for divergence of an antimicrobial peptide (Hymenoptaecin) in two common European bumblebee species. *Biological Journal of the Linnean Society*, 119(2), pp.299–310.

Lecocq, T. et al., 2013. Patterns of Genetic and Reproductive Traits Differentiation in Mainland vs. Corsican Populations of Bumblebees. *PLoS ONE*, 8(6), pp.16–22.

Leffler, E.M. et al., 2013. Multiple Instances of Ancient Balancing Selection Shared Between Humans and Chimpanzees. *Science*, 339(6127), pp.1578–1582.

Leffler, E.M. et al., 2012. Revisiting an old riddle: what determines genetic diversity levels within species? *PLoS biology*, 10(9), p.e1001388.

Lercher, M.J. & Hurst, L.D., 2002. Human SNP variability and mutation rate are higher in regions of high recombination. *Trends in Genetics*, 18(7), pp.337–340.

Lewontin, R., 1974. *The genetic basis of evolutionary change*, New York: Columbia University Press.

Liu, H. et al., 2017. Direct determination of the mutation rate in the bumblebee reveals evidence for weak recombination-associated mutation and an approximate rate constancy in insects. *Molecular Biology and Evolution*, 34(1), pp.119–130.

Loewe, L. & Charlesworth, B., 2006. Inferring the distribution of mutational effects on fitness in Drosophila. *Biology letters*, 2(3), pp.426–430.

Lourenço, J., Galtier, N. & Glémin, S., 2011. Complexity, pleiotropy, and the fitness effect of mutations. *Evolution*, 65(6), pp.1559–1571.

Luikart, G. et al., 2010. Estimation of census and effective population sizes: The increasing usefulness of DNA-based approaches. *Conservation Genetics*, 11(2), pp.355–373.

Lynch, M., 2010. Evolution of the mutation rate. *Trends in genetics*, 26(8), pp.345–52.

Lynch, M. et al., 2016. Genetic drift, selection and the evolution of the mutation rate. *Nature reviews. Genetics*, 17(11), pp.704–714.

Lynch, M., 2011. The lower bound to the evolution of mutation rates. *Genome Biology and Evolution*, 3(1), pp.1107–1118.

Lynch, M., 2007. *The Origins of Genome Architecture*, Sunderland: Sinauer Associates Inc.

Lynch, M. & Conery, J.S., 2003. The origins of genome complexity. *Science*, 302, pp.1401–4.

Lynch, M., Koskella, B. & Schaack, S., 2006. Mutation pressure and the evolution of organelle genomic architecture. *Science*, 311, pp.1727–30.

Malinsky, M. et al., 2017. Whole Genome Sequences Of Malawi Cichlids Reveal Multiple Radiations Interconnected By Gene Flow. *bioRxiv*, p.143859.

Martin, G. & Lenormand, T., 2006. A General Multivariate Extension of Fisher's Geometrical Model and the Distribution of Mutation Fitness Effects Across Species. *Evolution*, 60(5), pp.893–907.

Martincorena, I. & Luscombe, N.M., 2013. Non-random mutation: The evolution of targeted hypermutation and hypomutation. *BioEssays*, 35(2), pp.123–130.

Maynard Smith, J. & Haigh, J., 1974. The hitch-hiking effect of a favourable gene. *Genetical Research*, 23(1), pp.23–35.

Mckinney, M.L., 1997. Extinction Vulnerability and Selectivity : Combining Ecological and Paleontological Views. *Annu. Rev. Ecol. Syst.*, 28, pp.495–516.

Meiklejohn, C.D., Montooth, K.L. & Rand, D.M., 2007. Positive and negative selection on the mitochondrial genome. *Trends in genetics*, 23(6), pp.259–263.

Messer, P.W. & Petrov, D.A., 2013. Population genomics of rapid adaptation by soft selective sweeps. *Trends in ecology & evolution*, 28(11), pp.659–69.

Messer, P.W. & Petrov, D. a, 2013. Frequent adaptation and the McDonald-Kreitman test. *Proceedings of the National Academy of Sciences of the United States of America*, 110(21), pp.8615–20.

Montgomery, M.E. et al., 2010. Widespread selective sweeps affecting microsatellites in Drosophila populations adapting to captivity: Implications for captive breeding programs. *Biological Conservation*, 143(8), pp.1842–1849.

Mulligan, C.J., Kitchen, A. & Miyamoto, M.M., 2006. Comment on "Population Size Does Not Influence Mitochondrial Genetic Diversity in Animals." *Science*, 314, p.1390a.

Nabholz, B., Mauffrey, J.-F., et al., 2008. Determination of mitochondrial genetic diversity in mammals. *Genetics*, 178(1), pp.351–61.

Nabholz, B., Glémin, S. & Galtier, N., 2008. Strong variations of mitochondrial mutation rate across mammals- the longevity hypothesis. *Molecular biology and evolution*, 25(1), pp.120–130.

Nabholz, B., Glémin, S. & Galtier, N., 2009. The erratic mitochondrial clock: variations of mutation rate, not population size, affect mtDNA diversity across birds and mammals. *BMC evolutionary biology*, 9(54).

Nachman, M.W., 1998. Deleterious mutations in animal mitochondrial DNA. *Genetica*, 102(103), pp.61–69.

Naylor, G.J., Collins, T.M. & Brown, W.M., 1995. Hydrophobicity and phylogeny. *Nature*, 373(6515), pp.565–566.

Nei, M., Maruyama, T. & Chakraborty, R., 1975. The Bottleneck Effect and Genetic Variability in Populations. *Evolution*, 29(1), pp.1–10.

Nicolaisen, L.E. & Desai, M.M., 2013. Distortions in genealogies due to purifying selection and recombination. *Genetics*, 195(1), pp.221–230.

Nicolaisen, L.E. & Desai, M.M., 2012. Distortions in Genealogies Due to Purifying Selection Research article. *Mol. Biol. Evol.*, 29(11), pp.3589–3600.

Nielsen, R. & Yang, Z., 2003. Estimating the distribution of selection coefficients from phylogenetic data with applications to mitochondrial and viral DNA. *Molecular Biology and Evolution*, 20(8), pp.1231–1239.

Nielson, R., 2005. Molecular Signatures of Natural Selection. *Annu. Rev. Genet.*, 39, pp.197–218.

Ohta, T., 1992. The Nearly Neutral Theory of Molecular Evolution. *Annual Review of Ecology, Evolution, and Systematics*, 23(1992), pp.263–286.

Oli, M.K., 2004. The fast-slow continuum and mammalian life-history patterns: An empirical evaluation. *Basic and Applied Ecology*, 5(5), pp.449–463.

Oppold, A.M. & Pfenninger, M., 2017. Direct estimation of the spontaneous mutation rate by short-term mutation accumulation lines in Chironomus riparius. *Evolution Letters*, in press.

Otto, S.P. & Whitlock, M.C., 1997. The probability of fixation in populations of changing size. *Genetics*, 146(2), pp.723–733.

Palstra, F.P. & Fraser, D.J., 2012. Effective/census population size ratio estimation: A compendium and appraisal. *Ecology and Evolution*, 2(9), pp.2357–2365.

Palstra, F.P. & Ruzzante, D.E., 2008. Genetic estimates of contemporary effective population size: what can they tell us about the importance of genetic stochasticity for wild population persistence? *Molecular Ecology*, 17(15), pp.3428–3447.

Paradis, E., Claude, J. & Strimmer, K., 2004. APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, 20(2), pp.289–290.

Pesole, G. et al., 1999. Nucleotide substitution rate of mammalian mitochondrial genomes. *Journal of molecular evolution*, 48(4), pp.427–434.

Piganeau, G. & Eyre-Walker, A., 2003. Estimating the distribution of fitness effects from DNA sequence data: implications for the molecular clock. *Proceedings of the National Academy of Sciences*, 100(18), pp.10335–10340.

Piganeau, G. & Eyre-Walker, A., 2009. Evidence for variation in the effective population size of animal mitochondrial DNA. *PLoS ONE*, 4(2), p.e4396.

Pimm, S.L. et al., 1988. On the Risk of Extinction. *The American Naturalist*, 132(6), pp.757–785.

Popadin, K. et al., 2007. Accumulation of slightly deleterious mutations in mitochondrial protein-coding genes of large versus small mammals. *Proceedings of the National Academy of Sciences of the United States of America*, 104(33), pp.13390–13395.

Popadin, K.Y. et al., 2013. Purifying selection in mammalian mitochondrial protein-coding genes is highly effective and congruent with evolution of nuclear genes. *Molecular Biology and Evolution*, 30(2), pp.347–355.

Pratto, F. et al., 2014. Recombination initiation maps of individual human genomes. *Science*, 346(6211), pp.1256442–1256442.

Presgraves, D.C., 2005. Recombination enhances protein adaptation in Drosophila melanogaster. *Current Biology*, 15(18), pp.1651–1656.

Purvis, A. et al., 2000. Predicting extinction risk in declining species. *Proc. R. Soc.*, 267, pp.1947–52.

Qian, W. et al., 2012. Balanced codon usage optimizes eukaryotic translational efficiency. *PLoS Genetics*, 8(3).

Rand, D.M. & Kann, L.M., 1996. Excess amino acid polymorphism in mitochondrial DNA: contrasts among genes from Drosophila, mice, and humans. *Molecular biology and evolution*, 13(6), pp.735–748.

Rand, D.M. & Kann, L.M., 1998. Mutation and selection at silent and replacement sites in the evolution of animal mitochondrial DNA. *Genetica*, 102(103), pp.393–407.

dos Reis, M. & Yang, Z., 2013. Why do more divergent sequences produce smaller nonsynonymous/synonymous rate ratios in pairwise sequence comparisons? *Genetics*, 195(1), pp.195–204.

Rohde, K., 1992. Latitudinal Gradients in Species Diversity : The Search for the Primary Cause. *Nordic Society Oikos*, 65(3), pp.514–527.

Rolland, J. et al., 2016. Molecular evolutionary rates are not correlated with temperature and latitude in Squamata: an exception to the metabolic theory of ecology? *BMC Evolutionary Biology*, 16(95).

Romiguier, J. et al., 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, 515, pp.261–263.

Schmidt-Nielsen, K., 1984. *Scaling. Why is Animal Size So Important?*, Cambridge: Cambridge Univeristy Press.

Shen, Y.-Y. et al., 2010. Adaptive evolution of energy metabolism genes and the origin of flight in bats. *Proceedings of the National Academy of Sciences of the United States of America*, 107(19), pp.8666–8671.

Shen, Y.Y. et al., 2009. Relaxation of selective constraints on avian mitochondrial DNA following the degeneration of flight ability. *Genome Research*, 19(10), pp.1760–1765.

Slotte, T. et al., 2011. Genomic determinants of protein evolution and polymorphism in arabidopsis. *Genome Biology and Evolution*, 3(1), pp.1210–1219.

Smeds, L., Qvarnström, A. & Ellegren, H., 2016. Direct estimate of the rate of germline mutation in a bird. *Genome Research*, 26(9), pp.1211–1218.

Smith, B.T. et al., 2017. A latitudinal phylogeographic diversity gradient in birds. *PLoS Biology*, 15(4), pp.1–24.

Smith, N.G.C. & Eyre-Walker, A., 2002. Adaptive protein evolution in Drosophila. *Nature*, 415, pp.1022–1024.

Smith, N.G.C. & Eyre-Walker, A., 2003. Partitioning the variation in mammalian substitution rates. *Molecular biology and evolution*, 20(1), pp.10–17.

Stamatakis, A., 2014. RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics*, 30(9), pp.1312–3.

Stewart, J.B. et al., 2008. Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease. *Nat Rev Genet*, 9(9), pp.657–62.

Stewart, J.B. et al., 2008. Strong purifying selection in transmission of mammalian mitochondrial DNA. *PLoS Biology*, 6(1), p.e10.

Stewart, J.B. & Larsson, N.-G., 2014. Keeping mtDNA in Shape between Generations. *PLoS Genetics*, 10(10), p.e1004670.

Stoletzki, N. & Eyre-Walker, A., 2011. Estimation of the neutrality index. *Molecular biology and evolution*, 28(1), pp.63–70.

Strathern, J.N., Shafer, B.K. & McGill, C.B., 1995. DNA synthesis errors associated with double-strand-break repair. *Genetics*, 140(3), pp.965–972.

Suarez, R.K., 1992. Hummingbird flight: sustaining the highest mass-specific metabolic rates among vertebrates. *Experientia*, 46(6), pp.565–570.

Sun, Z. et al., 2009. Comparison of base composition and codon usage in insect mitochondrial genomes. *Genes & Genomics*, 31(1), pp.65–71.

Sung, W. et al., 2012. Drift-barrier hypothesis and mutation-rate evolution. *PNAS*, 109(45), pp.18488–18492.

Takano-Shimizu, T., 1999. Local Recombination and Mutation Effects on Molecular Evolution in Drosophila. *Genetics*, 153, pp.1285–1296.

Torroni, A. et al., 2006. Harvesting the fruit of the human mtDNA tree. *Trends in Genetics*, 22(6), pp.339–345.

Wallberg, A. et al., 2014. A worldwide survey of genome sequence variation provides insight into the evolutionary history of the honeybee Apis mellifera. *Nature Genetics*, 46(10), pp.1081–1088.

Wang, J., 2005. Estimation of effective population sizes from data on genetic markers. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 360(1459), pp.1395–409.

Wang, J., Santiago, E. & Caballero, A., 2016. Prediction and estimation of effective population size. *Heredity*, 117(4), pp.193–206.

Waples, R.S., 2016. Life-history traits and effective population size in species with overlapping generations revisited: the importance of adult mortality. *Heredity*, 117(4), pp.241–250.

Waples, R.S. et al., 2013. Simple life-history traits explain key effective population size ratios across diverse taxa. *Proceedings of the Royal Society B*, 280(1768), p.20131339.

Waples, R.S., 2016. Tiny estimates of the Ne/N ratio in marine fishes: Are they real? *Journal of Fish Biology*, 89(6), pp.2479–2504.

Weissman, D.B. & Barton, N.H., 2012. Limits to the rate of adaptive substitution in sexual populations. *PLoS genetics*, 8(6), p.e1002740.

Welch, J.J., Eyre-Walker, A. & Waxman, D., 2008. Divergence and polymorphism under the nearly neutral theory of molecular evolution. *Journal of Molecular Evolution*, 67(4), pp.418–426.

White, E.P. et al., 2007. Relationships between body size and abundance in ecology. *Trends in Ecology and Evolution*, 22(6), pp.323–330.

Woolfit, M. & Bromham, L., 2005. Population size and molecular evolution on islands. *Proc. R. Soc.*, 272(1578), pp.2277–2282.

Wright, S., 1931. Evolution in Mendelian Populations. *Genetics*, 16(2), pp.97–159.

Wright, S., Keeling, J. & Gillman, L., 2006. The road from Santa Rosalia: a faster tempo of evolution in tropical climates. *Proceedings of the National Academy of Sciences of the United States of America*, 103(20), pp.7718–22.

Wright, S.D. et al., 2009. Slower tempo of microevolution in island birds: implications for conservation biology. *Evolution*, 63(9), pp.2275–2287.

Xie, Z. et al., 2016. Mutation rate analysis via parent– progeny sequencing of the perennial peach. I. A low rate in woody perennials and a higher mutagenicity in hybrids. *Proc. R. Soc. B*, 283(1841), p.20161016.

Yang, Z., 2006. *Computational Molecular Evolution*, New York: Oxford University Press.

Yang, Z., 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Molecular biology and evolution*, 24(8), pp.1586–1591.

Zinner, D. et al., 2009. Mitochondrial phylogeography of baboons (Papio spp.) – Indication for introgressive hybridization? *BMC Evolutionary Biology*, 9(1), p.83.

Appendices

n	No. of species per group	Slope	S.E.	Intercept
8	93	-0.60	0.050	-2.23
10	75	-0.62	0.048	-2.28
20	37	-0.54	0.082	-2.16
30	25	-0.53	0.067	-2.15
40	18	-0.51	0.069	-2.11
50	15	-0.45	0.068	-2.03

Table A5.1) The relationship between $\log(\pi_S)$ and $\log(\pi_N/\pi_S)$ for a number of different group sizes. All relationships were highly statistically significant, with p values all << 0.005. *n* indicates the number of groups used in the analysis, while the approximate number of species per group is given in the second column. The slope, the standard error of the slope and the intercept of the regression line for each value of *n* are given.