The University Of Sheffield.

The genetic architecture of a reproductive life-history trait in a wild passerine

By:

Jenny Claire Armstrong

A thesis submitted in partial fulfilments of the requirements for the degree of

Doctor of Philosophy

The University of Sheffield

Department of Animal & Plant Sciences

September 2015

Acknowledgements

I would like to gratefully acknowledge the funding provided by the Natural Environment Research Council (NERC), enabling me to undertake this PhD journey.

I would like to express my sincerest gratitude to my supervisor, Professor Jon Slate, for giving me this opportunity; for his vast patience in answering endless (and often repetitive) questions; for constructive criticism that managed to be both honest and encouraging, spurring me to work harder and be a better scientist; and finally for giving me the space to just get on with it, but for always being quick (and willing) to respond when I asked for help.

Thanks to the post-docs who were part of my lab group; Dr Anna Santure, for all her help and for putting up with me; Dr Jocelyn Poissant, Dr Matthew Robinson and Dr Isabelle de Cauwer.

A great deal of thanks must go to the people and institutions whose dedication to long term monitoring and data collection make projects such as this possible. In particular, Professor Ben Sheldon & team at the Edward Grey Institute at Oxford University, for providing all the data for the great tits of Wytham Woods, and Professor Marcel Visser & team at the Netherlands Institute of Ecology for all resources relating to the great tits of Hoge Veluwe. Thanks also to Professor Martien Groenen, Dr Kees van Oers, Dr Nikkie van Bers and Dr Veronika Laine.

I would further like to acknowledge;

The wonderful community that is TAB LAB. Academic credentials aside, the supportive, friendly, social and cake-filled environment that exists on B1 floor in the Animal and Plant Sciences Department has made the PhD experience a genuine pleasure;

My wonderful parents, for always supporting me, and for their unwavering belief in me, giving me the confidence to believe in myself (also for the chocolate support packages which never failed to cheer me up and fuelled most of my PhD); my brother, whose constant and encouraging presence means a great deal; my gran, whose hospitality and house in the Highlands of Scotland provided much needed sanctuary during times when I needed a break.

Thanks to Stephen Samworth for taking me on as a bird ringing trainee and getting me into the field, for being a patient and kind tutor in bird ID and data collection and giving me the opportunity to view up close (and take pictures of) my beloved little study species among a multitude of others. Thanks also to others in the Sorby Breck ringing group who allowed me to accompany / assist them on various bird ringing expeditions.

Finally, I wouldn't have gotten very far without the support of some truly great friends. In particular Isabel Winney, Marjorie Lundgren & Samiya Selim, for cocktails, coffee, park walks and everything in between that took the stress out of the PhD process and made it memorable; Helen Southern, a kindred spirit, with whom many valuable discussions were had during our espresso martini tour of Sheffield; Sarah Aynsley, for horseback adventures that got me out of the office; Philippa Gullet & Cheryl Mills for guidance and inspiration; and Hannah Ives, for having to live with me throughout it all.

Abstract

Understanding the capacity for species to respond to changes in their environment and the rate at which they are able to do so is a key topic in evolutionary biology and of increasing importance in wildlife conservation and management. However the mechanisms involved in mediating these responses are poorly understood. Specifically, while reactive responses may be advantageous in the short term persistent directional changes in environmental conditions may require a more profound response in order for organisms to adapt and persist successfully. Here I use data from two long-term studies of the great tit *Parus major* and apply a range of statistical techniques to dissect the genetic architecture of laying date, a reproductive life-history trait, to discern the extent to which a genetic component of variation contributes to observed phenotypic variation. A heritable component of variation exists in both populations, but specific regions of the genome contributing to trait variation could not be detected by quantitative trait loci mapping (Chapters 2 & 3), genome-wide association (Chapters 2 & 3) or chromosome partitioning (Chapter 4) analyses. These findings are consistent with a highly polygenic basis for variation in laying date, variation maintained by many genes of small effect. Attempts to increase the statistical power by combining two phenotypic datasets to increase overall sample size (Chapter 3) and increasing marker density (Chapter 5) drew similar conclusions, with an absence of genome-wide significant QTL. Despite evidence of a strong association on chromosome 3 (Chapter 5), an overall lack of consistency between analyses and datasets on regions exhibiting the highest associations suggests that power to detect genomic regions, particularly when variation may be determined by many variants of small effect, is low. I conclude that while genetic variation exists, environmental factors and phenotypic plasticity likely account for much of the variation in laying date.

Contents

		page
List of Figures		7
List of Tables		10
List of Appendices		12
Chapter 1:	**Introduction**	14
	1.1 Climate change and trophic mismatches	**15**
	1.1.1 Breeding phenology	17
	1.1.2 Passerine birds and shifts in laying date	19
	1.2 Study species	**24**
	1.2.1 Great tits	24
	1.2.2 Predator and prey phenology	26
	1.2.3 Wytham Woods (UK)	26
	1.2.4 Hoge Veluwe (NL)	28
	1.3 Patterns of laying date in *Parus major*	**30**
	1.3.1 A genetic component to laying date	32
	1.4 Gene mapping and quantitative genetics	**34**
	1.4.1 Identifying genes responsible for heritable variation in wild animals	35
	1.4.2 Mapping QTL	38
	1.4.3 Partitioning variance	40
	1.4.4 SNP markers	41
	1.4.5 Great tit genetic resources	42
	1.5 Aims and objectives	**43**
Chapter 2:	**Quantitative genetic analyses for loci contributing to variation in laying date**	44
	2.1 Abstract	**45**
	2.2 Introduction	**46**

2.3 Methods **49**
2.3.1. Phenotype data 49
2.3.2. Pedigree 49
2.3.3. Marker data 49
2.3.4. Statistical analysis 50
2.4 Results **59**
2.4.1. Complex interactions 59
2.4.2. Variance components analysis 60
2.4.3. Genome-scan and genome-wide association analysis 60
2.5 Discussion **64**

Chapter 3: A comparative and combined genome mapping analysis of variation in laying date in two populations of *Parus major* 71

3.1 Abstract **72**
3.2 Introduction **73**
3.3 Methods **77**
3.3.1. Phenotype data 77
3.3.2. Pedigree data (A matrix) 78
3.3.3. Marker data 78
3.3.4. Statistical analysis 80
3.4 Results **86**
3.4.1. Variance components analysis 86
3.4.2. QTL scan and GWAS results (NL) 86
3.4.3. QTL scan and GWAS results (NLUK) 89
3.5 Discussion **92**

Chapter 4: Partitioning additive genetic variation for a reproductive trait in an ecological model organism 95

4.1. Abstract **96**
4.2. Introduction **97**
4.3. Methods **101**
4.3.1. Phenotype information 101
4.3.2. Pedigree & marker information 101
4.3.3. Chromosome partitioning model method & statistical testing 102
4.4. Results **105**
4.4.1. Likelihood ratio testing 105
4.4.2. Linear modelling 110
4.5. Discussion **114**

Chapter 5: The genetic architecture of laying date; results from a high density chip 116

5.1. Abstract **117**
5.2. Introduction **118**
5.3. Methods **123**
5.3.1. Processing the Axiom array 123
5.3.2. GWAS analysis 125
5.3.3. Chromosome partitioning analysis 127
5.4. Results **129**
5.4.1. GWAS 129
5.4.2. Chromosome partitioning 132
5.5. Discussion **135**

Chapter 6: Discussion 139

Appendices 151

Bibliography 159

List of Figures

Chapter 1

Figure 1.1. Great tit (*Parus major*) 25

Figure 1.2. Map of Wytham Woods, Oxford (UK) highlighting the nine research areas and position of nest boxes (black dots) 28

Figure 1.3. Location and map of Hoge Veluwe national park in the Netherlands 29

Chapter 2

Figure 2.1. A tree diagram illustrating the interaction between female age (years) and altitude (m) and subsequent laying dates associated with each group 59

Figure 2.2. A QTL (a) and GWAS (b) scan for variants contributing to variation in laying date. Nominal significance thresholds of 7.38 (a) 1.3 (b) and genome-wide significance thresholds of 14.11 (a) 5.05 (b) are indicated by dashed blue lines 63

Chapter 3

Figure 3.1. Genome scans using linkage (a) and association (b) mapping to detect genetic variants contributing to variation in laying date in the Hoge Veluwe (NL) great tit population. Dashed lines represent nominal (a) 7.38; b) 1.3) and genome-wide significant (a) 14.11; b) 5.05) thresholds 88

Figure 3.2. Genome scans using linkage (a) and association (b) mapping to detect genetic variants contributing to variation in 91

laying date for the combined Hoge Veluwe (NL) and Wytham Woods (UK) great tit population dataset. Dashed lines represent nominal (a) 7.38; b) 1.3) and genome-wide significant (a) 14.11; b) 5.05) thresholds

Chapter 4

Figure 4.1. Results of the chromosome partitioning analysis, showing the relationship between the variance (h^2) explained by chromosomes in the UK and NL great tit populations. For clarity chromosomes that explain no variance are not labelled 107

Figure 4.2. Linear modelling for variance attributed to chromosomes in relation to number of genes (a) and size (Mbp) (b) for the UK great tit population. For clarity chromosomes that explain no variance are not labelled 111

Figure 4.3. Linear modelling for variance attributed to chromosomes in relation to number of genes (a) and size (Mbp) (b) for the NL great tit population. For clarity chromosomes that explain no variance are not labelled 112

Figure 4.4. Linear modelling for variance attributed to chromosomes in relation to number of genes (a) and size (Mbp) (b) for the NLUK dataset. For clarity chromosomes that explain no variance are not labelled 113

Chapter 5

Figure 5.1. Schematic view of the Affymetrix genotyping 'AffyPipe' protocol for standard and 'best practice' (indicated by the dashed line) workflows (from (Nicolazzi et al. 2014)) 124

Figure 5.2. GWAS scan for regions associated with variation in lay date. The blue dashed lines represent significance thresholds of nominal (1.301) and genome-wide (7) significance 131

Figure 5.3. Results of the chromosome partitioning analysis where the amount of variance explained by each chromosome (labelled directly on the figure) is plotted relative to number of genes (a) and size (Mbp) (b) 134

List of Tables

Chapter 1

Table 1.1. Adapted from Visser et al. (2003). Patterns of laying date in multiple populations of blue (BT) and great (GT) tits, obtained by regressing the annual mean population laying date against year for all populations separately. 23

Chapter 2

Table 2.1. Description of the variables included as terms in the initial version of the animal model and whether they were designated as 'fixed' or 'random' effects 51

Table 2.2. Statistics associated with the random effects included in the animal model, the likelihood ratio test (LRT) score obtained by comparison of the logarithm of the odds scores of the models excluding (LOD1) and including (LOD2) the relevant term. 52

Table 2.3. The top 10 highest associations from GWAS analysis including a permutation test of 1000 iterations to test for significance (P1df). 62

Chapter 3

Table 3.1. Statistical testing of random effect terms included in the final versions of the polygenic animal models for the Netherlands (a) and combined (b) datasets, using logarithm of the odds (LOD) scores obtained from ASReml, and subsequent likelihood ratio test (LRT) of significance 82

Table 3.2. The top 10 associations from GWAS analyses of laying date in the Hoge Veluwe (NL) great tit population. A permutation test 87

(1000 iterations) was performed to test for significance (P1df)

Table 3.3. The top 10 associations from GWAS analyses of laying date in the combined NL / UK great tit population dataset. A permutation test (1000 iterations) was performed to test for significance (P1df) 90

Chapter 4

Table 4.1. Estimates of heritability for laying date in the UK, NL and combined NLUK great tit datasets, as obtained from animal modelling and chromosome partitioning (G_{matrix}), showing standard error 105

Table 4.2 Likelihood ratio test (LRT) results for chromosome partitioning analysis performed on the Wytham Woods (UK), Hoge Veluwe (NL) and combined population (NLUK) datasets. 108

Chapter 5

Table 5.1. Statistical testing of random effect terms included in the final versions of the mixed effects model for the UK great tit population using a high density SNP chip, using logarithm of the odds (LOD) scores obtained from ASReml, and subsequent likelihood ratio test (LRT) of significance 126

Table 5.2. The top 10 associations from GWAS analysis, showing the chi-square values results of a permutation test of 1000 iterations (P1df) 130

Table 5.3. Likelihood ratio test (LRT) results for chromosome partitioning analysis performed on the Wytham Woods (UK) population using genotype information from a 500K SNP chip. 133

List of appendices

Chapter 2

Appendix 2.1 Coefficients and standard errors for fixed effects included in the final animal model, for heritability of laying date in the Wytham Woods great tit population 152

Appendix 2.2 Quantile-Quantile (QQ) plot of an association scan for laying date in the Wytham Woods (UK) great tit population before ($\lambda = 1.08 \pm 4.9x10^{-4}$) (a) and after ($\lambda = 0.98 \pm 6.5x10^{-4}$) (b) correction for population stratification 153

Chapter 3

Appendix 3.1. Coefficients and standard errors for fixed effects included in the final animal model for Hoge Veluwe (NL) and combined (NLUK) dataset 154

Appendix 3.2. Quantile-Quantile (QQ) plot of an association scan for laying date in the Hoge Veluwe (NL) great tit population before ($\lambda = 1.01 \pm 4.9x10^{-4}$) (a) and after ($\lambda = 0.99 \pm 4.6x10^{-4}$) (b) correction for population stratification 155

Appendix 3.3. Quantile-Quantile (QQ) plot of an association scan for laying date in the combined Wytham Woods (UK) / Hoge Veluwe (NL) dataset before ($\lambda = 1.00 \pm 8.5x10^{-4}$) (a) and after ($\lambda = 0.99 \pm 5.6x10^{-4}$) (b) correction for population stratification 156

Chapter 5

Appendix 5.1 Coefficients and standard errors for fixed effects included in the final mixed effects model, for heritability of laying date in the Wytham Woods (UK) great tit population 157

Appendix 5.2 Quantile-Quantile (QQ) plot of an association scan for laying date in the Wytham Woods (UK) great tit population before (λ = 1.04 $\pm$ 4.74x10^{-6}) (a) and after (λ = 1.00 $\pm$ 5.65x10^{-6}) (b) correction for population stratification using a 500K SNP chip 158

Chapter 1

Introduction

Identifying the potential for and the rate at which species can adapt to changes in their environment has consequences not only for a greater understanding of evolutionary biology, but also practical implications for the management and conservation of wildlife (Seddon et al. 2005; Romanov et al. 2009; Narum et al. 2013; Charmantier et al. 2014). However, identifying the mechanisms associated with advantageous adaptations in wild populations is fraught with difficulty. By definition, un-manipulated wild populations are subject to a host of physical, natural and climatic forces beyond the control, measurement and potentially even detection of scientists.

Increasingly long-term studies of wild organisms are yielding substantial quantities of data that, together with genetic information, can be used to estimate the genetic component of variation, or heritability of traits, and their underlying genetic architecture (Réale et al. 2003; Charmantier et al. 2008; Kruuk & Hill 2008; Slate et al. 2010; Poissant et al. 2012; Bérénos et al. 2014). Classical theory suggests that the rate of response of any given organism to selection is based on the amount of additive genetic variation present in the organism's genome (Fisher 1930). Identifying the extent to which observed phenotypic variation is maintained by an underlying genetic component may give us an insight into the potential for a microevolutionary response to selection.

1.1 Climate change and trophic mismatches

A body of evidence has been amassed documenting the direct effect of climate change on organisms across taxa (Gordo & Sanz 2010; Pozsgai & Littlewood 2011; Visser & Both 2005), often manifested as shifts in species' distribution (Parmesan & Yohe 2003) and phenology (the timing of seasonal activities (Walther et al. 2002)). Food availability dictates many of the responses observed and predicted in organisms, as climate change has the potential to induce trophic mismatches where the timing of peak food abundance is no longer aligned with peak food requirement (Stenseth & Mysterud 2002). The extent of the mismatch, and

the capacity of affected organisms to adapt will ultimately determine the persistence of many organisms with the rate and form of response varying substantially between species (Donnelly et al. 2011; Visser & Both 2005).

While some studies have suggested asynchronies could be corrected with very little adverse impact to some organisms that are reliant on the changing distribution / availability of a food source (van Asch et al. 2007; Stireman et al. 2005), for others the consequences could potentially be more severe (Memmott et al. 2007; Saino et al. 2010; Thomas et al. 2001; Watanuki et al. 2009). The interplay of many factors ultimately determines the outcome of mismatches, particularly the life history strategies of organisms.

Specialist species are at much greater risk of fitness losses than generalists as they are likely to be slow or incapable of utilise alternative food sources in the event a preferred source becomes unavailable (van Asch & Visser 2007; Miller-Rushing et al. 2010; Miller-Struttmann et al. 2015; Kowalczyk et al. 2015). Migratory species display increased vulnerability to trophic mismatches and population declines (Both & Visser 2001; Both et al. 2006; Saino et al. 2010). Reproductive strategy may also determine species responses to mismatches (Gregory et al. 2009; Kerby & Post 2013; Vedder et al. 2013).

Short-lived species may also be more resilient to environmental changes than longer lived species as rapid turnover of generations facilitates rapid adaptation (Perry et al. 2005; Vedder et al. 2013). In contrast species with slow life-histories unable to produce rapid demographic responses in response to climate-induced trophic mismatches may experience population declines (Post & Forchhammer 2008; Vors & Boyce 2009; Gienapp, Lof, et al. 2013).

Population demography, also has the potential to counterbalance the negative impact of asynchronies on with food availability. Fitness losses encountered by individuals breeding out of sync with prey availability may bestow fitness gains on individuals breeding in synchrony with prey as a consequence of relaxed competition (Reed, Jenouvrier, et al. 2013; Reed, Grotan, et al. 2013). Understanding the fitness consequences of mismatches is essential to be able to identify those species and populations that are most likely to be negatively affected by climate change (Donnelly et al. 2011).

1.1.1 Breeding phenology

For many species the main strategy employed to alleviate the negative consequences of inadequate food availability, and to regain synchrony with the phenology of a food source, is a change to the timing of breeding (Visser 2008; Both et al. 2004). Adaptive breeding responses to shifts in plant phenology as a result of climate change have been observed in a number of mammals including red deer and red squirrels (Moyes et al. 2011; Réale et al. 2003; Burthe et al. 2011).

Birds are capable of adapting their breeding strategy by making adjustments to a complex suite of stages in their reproduction, namely the date they choose to lay their eggs (laying date), their clutch size and the timing of onset of the incubation period (Visser et al. 2004; Ardia et al. 2006; Visser et al. 1998; van Balen 1973; Stoleson & Beissinger 1995; Cresswell & McCleery 2003; Goodenough et al. 2009). It is well established that the principal environmental cue mediating these behaviours is photoperiod (Lyon et al. 2008; Silverin et al. 2008; Dawson 2008), but birds are able to vary these behaviours to cope with environmental variation by means of 'supplementary' cues (Lyon et al. 2008; Dawson 2008; Lattin et al. 2016).

The physical (neuroendocrine and endocrine) mechanisms underpinning breeding phenology can operate under a range of conditions, and this physiological plasticity (Dawson 2008; Lattin et al. 2016) enables birds to fine-tune their behaviour by incorporating information from supplementary cues such as food availability, temperature, rainfall, and the phenology of other organisms (Lyon et al. 2008; Dawson 2008).

For marine bird species such as the common guillemot and black-legged kittiwake food availability (mediated by pre-breeding sea surface temperature) has driven observable variation in patterns of laying date (Reed et al. 2009; Votier et al. 2009; Shultz et al. 2009). Many birds are also capable of altering their breeding behaviour to take advantage of any beneficial conditions they experience such as unexpected food abundance.

Increased food availability prior to laying shown to advance the onset of laying in starlings (Källander & Karlsson 1993; Meijer & Langer 1995), tree swallows (Nooker et al. 2005) and seabirds such as murres and kittiwakes (Shultz et al. 2009). At the extreme, opportunistic breeders such as crossbills have been observed to breed whenever food is abundant, within the limits of their physiological capacity (Deviche 1997; Hahn 1998; Deviche & Sharp 2001). The onset of breeding on birds has also been shown to occur in response to rainfall, in species inhabiting tropical or arid regions such as the spotted antbird (Wikelski et al. 2000), Rufous-winged sparrow (Deviche et al. 2006) and Darwin's finch (Hau et al. 2004).

Temperature is the cue attributed with most of the variation in laying date in birds (Cresswell & McCleery 2003; Silverin et al. 2008; Visser et al. 2009; Schaper et al. 2012; de Jong et al. 2015). However the precise pattern of temperature influencing laying date varies both within and between species. Eider ducks, tawny owls and starlings have all been shown to advance the onset of laying in response to milder winters (Fossøy et al. 2014; Solonen 2014; Williams et al. 2015). In contrast laying date in the majority of passerines has been attributed to spring (March and April) temperature (Forchhammer et al. 1998; Crick &

Sparks 1999; Dunn 2004; Visser et al. 2006; Porlier et al. 2012; Schaper et al. 2012; Gienapp, Lof, et al. 2013; Thorley & Lord 2015).

Using captive great tit populations, a model species for studies of avian phenology, a series of experiments over the course of three years representing some of the most in-depth ever conducted revealed the complex nature of the relationship between spring temperature and laying date (Schaper et al. 2012). Broad conclusions drawn from observations made over the course of these experiments suggested that the specific nature of temperature pattern experienced by birds, not just warm temperatures in general, influence the onset of laying (Visser et al. 2009; Schaper et al. 2012). Seasonal increases in temperature, as opposed to mean or daily temperature variation were found to have the most significant impact on female's decisions to start laying (Schaper et al. 2012).

Schaper et al. (2011) in a previous study were also able to demonstrate that visual cues in the form of vegetation phenology (associated with the subsequent abundance of invertebrate prey) did not affect timing of breeding. Attributing behavioural variation to specific supplementary cues is difficult given that they are inherently connected, with cues such as temperature and rainfall determining food availability and abundance in subsequent periods (Wikelski et al. 2000; Shultz et al. 2009; Schaper et al. 2011). Untangling these cues and identifying the proximate causes of behaviour is important for our understanding of how species are likely to respond to climate change.

1.1.2 Passerine birds and shifts in laying date

Laying date is one of the life history traits best documented in relation to phenological responses to climate change, with evidence having been amassed from numerous long-term ornithological studies. Trends for earlier egg-laying have been directly attributed to changes in temperature (Crick & Sparks 1999; Crick et al. 1997; McCleery & Perrins 1998)

with numerous studies documenting advances in laying date (Bergmann 1999; Koike & Higuchi 2002; Pearce-Higgins et al. 2005; Torti & Dunn 2005; Møller et al. 2006; D'Alba et al. 2010; Charmantier & Gienapp 2014).

Research on passerine birds accounts for some of the most concentrated efforts to investigate the effect of climate change on organismal phenology. The abundance of many species and their short life-histories make them ideal candidates for investigating systematic trends. Data available from long-term studies of numerous passerines have described varying trends in laying date between species (Both et al. 2004; Moussus et al. 2011; Visser & Both 2005).

Reviews have helped build a picture of patterns of avian breeding behaviour in response to climate change worldwide. Parmesan & Yohe (2003) identified 78 out of 168 species for which long-term trends for earlier laying have been recorded in the literature. The trend for earlier laying is particularly pronounced in passerines. For 45 (out of 57) species for which evidence was collated documenting trends for earlier egg-laying 36 were passerines (Dunn 2004), a pattern consistent with other reviews (Crick et al. 1997; Crick & Sparks 1999).

Earlier laying was also consistently related to temperature in passerines. Crick & Sparks (1999) noted that whilst patterns of laying date were found to be accounted for by changes in climate in the majority (86%) of species of all orders, earlier laying was strongly associated with pre-breeding (March / April) temperatures in 7 species (winter wren, dunnock, blackcap, willow warbler, spotted flycatcher, long-tailed tit & greenfinch) all passerines (Crick & Sparks 1999).

Variation is also evident between populations of the same species, some showing a clear response, with earlier egg-laying associated with increased temperatures (Dolenec & Dolenec 2011; Winkler et al. 2002; Vidyasagar 1999; Both & Visser 2001; Cresswell &

McCleery 2003; Crick et al. 1997), others showing no advancement in laying date at all (Visser et al. 2003; Gienapp et al. 2006; Ahola et al. 2012). Heterogeneity between populations in their breeding phenology and response to increasing temperatures has been most well documented in 'model' tit and flycatcher species.

Contrasting patterns of laying date have been observed between populations of great and blue tits from several long-term study sites (Table 1.1). Significant advances in laying date in great tits from the UK, Belgium and the Netherlands (Visser et al. 2003; Husby et al. 2010; Matthysen et al. 2011) were not exhibited by populations in France, Finland, Russia or other Dutch sites (Table 1.1) (Visser et al. 2003). Similarly in 12 populations of blue tits only 4 displayed significant advances in laying date (Table 1.1) (Visser et al. 2003; Potti 2009).

Patterns of laying date also vary between populations of tits and flycatchers breeding in the same location. Both blue tit and pied flycatcher populations in Spain have advanced their laying date but the flycatchers have done so a slower rate than the sympatric and resident tit species, with blue tits and flycatchers laying approximately 0.41 / 0.2 days per year earlier respectively (1984-2007) (Potti 2009). Whilst there was no difference in the slope of the response the blue tits responded much more strongly to temperatures earlier in the season (Potti 2009), consistent with other studies showing the relationship between laying date and temperature strongest for great tits during the period 2nd April – 7th May, and for flycatchers 5th-28th May (Ahola et al. 2007).

A migratory species the breeding phenology of pied flycatchers is constrained by their arrival date as they are not exposed to temperatures or vegetation / prey phenology at nesting locations before their arrival at nesting sites. Several populations of pied flycatcher have shown no trend for earlier laying (Sanz et al. 2003; Both et al. 2004; Ahola et al. 2007; Ahola et al. 2012) resulting in reduced reproductive success (Merilä, Kruuk, et al. 2001;

Sanz et al. 2003; Both et al. 2004). Out of 25 populations of pied / collared flycatcher 9 have exhibited a significant shift to earlier laying (Both et al. 2004) with the Hogue Veluwe pied flycatcher population advancing their laying date by 10 days over a 20 year period. Despite these advances the shift is insufficient to maintain synchrony with their prey (Both & Visser 2001; Both et al. 2004; Both et al. 2006), a situation mirrored by the Hogue Veluwe great tit population, which are also failing to track the changing phenology of the winter moth (Visser et al. 1998; Visser et al. 2004; Nussey et al. 2005).

Table 1.1. Adapted from Visser et al. (2003). Patterns of laying date in multiple populations of blue (BT) and great (GT) tits, obtained by regressing the annual mean population laying date against year for all populations separately. Countries are France (F), Belgium (B), the Netherlands (NL), United Kingdom (UK), Russia (RU), Finland (FIN) and Spain (SP). Results are from Visser et al. (2003) unless otherwise stated. Significant results are highlighted in bold. Years of data collection: Visser et al. (2003); 1979-1998, Matthysen et al. (2011); 1979-2007, Husby et al. (2010); 1973-2006, Ahola (2009); 1953-2008, Potti (2009); 1984-2007.

			Number of Years		**p-value laying date versus year**		**Slope laying date versus year (s.e.)**		**Source**
Study site	**Country**	**Coordinates**	**GT**	**BT**	**GT**	**BT**	**GT**	**BT**	
Pirio (Corsica)	F	42˚23’ N, 08˚45’ E	0	20	n.a.	0.963	n.a.	-0.008 (0.168)	
Ventoux	F	44˚35’ N, 05˚25’ E	20	20	0.175	0.158	-0.367 (0.260)	-0.326 (0.221)	
Hutsepot	B	51˚01’ N, 03˚70’ E	20	20	**0.025**	**0.015**	-0.509 (0.209)	-0.535 (0.200)	
Boswachter	B	51˚16’ N, 04˚29’ E	29	29	**0.001**	**<0.001**	-0.416 (0.111)	-0.412 (0.110)	Matthysen et al. (2011)
Liesbos	NL	51˚35’ N, 04˚40’ E	20	20	0.383	0.134	-0.177 (0.198)	-0.337 (0.215)	
Wytham Wood	UK	51˚47’ N, 01˚20’ W	33	0	**<0.001**	n.a.	-0.367 (0.090)	n.a.	Husby et al. (2010)
Treswell	UK	53˚18’ N, 00˚51’ W	20	20	**0.021**	0.093	-0.538 (0.213)	-0.417 (0.236)	
Oosterhout	NL	51˚55’ N, 05˚50’ E	20	19	0.626	0.112	-0.101 (0.204)	-0.409 (0.244)	
Warnsborn	NL	52˚05’ N, 05˚50’ E	20	20	**0.017**	0.134	-0.404 (0.153)	-0.290 (0.185)	
Hoge Veluwe	NL	52˚05’ N, 05˚50’ E	33	0	**0.007**	n.a.	-0.196 (0.079)	n.a.	Husby et al. (2010)
Vlieland	NL	53˚15’ N, 05˚00’ E	20	20	0.227	0.493	-0.200 (0.160)	-0.086 (0.123)	
Zvenigorod	RU	55˚44’ N, 36˚51’ E	19	0	0.544	n.a.	-0.118 (0.190)	n.a.	
Oulu	FIN	65˚03’ N, 25˚26’ E	13	0	0.200	n.a.	-0.534 (0.392)	n.a.	
Haukipudas	FIN	65˚05’ N, 25˚34’ E	16	0	0.395	n.a.	-0.236 (0.269)	n.a.	
Harjavalta & Askainen	FIN	61˚20’ N, 22˚10’ E / 60˚30’ N, 21˚45’ E	57	0	0.97	n.a.	0.0013 (0.039)	n.a.	Ahola (2009)
La Hiruela	SP	41˚04’ N, 03˚27’ E	0	21	n.a.	**0.022**	n.a.	-0.410 (0.170)	Potti (2009)

1.2 Study species

1.2.1 Great tits

Passeriformes are the largest order of birds containing roughly 60% of all species, many of which are prolific and widespread. Readily nesting in nest boxes (Perrins 1965; Kruuk 2004; Husby et al. 2012) and with short generation spans, the ease with which many species of passerine can be studied makes them ideal subjects for quantitative genetic analyses, with long-term studies capable of generating substantial quantities of data (Visser et al. 2003; Brommer et al. 2005; Wilkin et al. 2009; Tarka et al. 2010; Ahola et al. 2012; Husby et al. 2012; Vedder et al. 2013; Thorley & Lord 2015).

The great tit (*Parus major*) is a social, common and widespread passerine with a distribution that extends through Europe, North Africa and Asia (BirdLife International 2015). Prolific in Europe with a population size that appears to be increasing (BirdLife International 2015), the species is not currently of conservation concern. Great tits occupy a wide range of habitats including deciduous and mixed woodland (Gosler et al. 2013) and are instantly recognisable as a common garden bird in Europe, with a black cap and breast stripe, yellow chest, white cheeks, and blue / green colouration on the back and wings (Figure 1.1). Though superficially similar, males are typically distinguished from females by having a broader and more defined black chest stripe.

Figure 1.1. Great tit (*Parus major*) (Photo credit: authors own)

Great tits are primarily insectivorous during the spring and summer, feeding their chicks on caterpillars of the winter moth (*Operophtera brumata*) (Perrins 1965; van Balen 1973; van Noordwijk et al. 1995; Visser et al. 2006; van Asch et al. 2007) but will forage on seeds and berries when numbers of their insect prey are low (Snow & Perrins 1998). Great tits are a monogamous and territorial bird (Krebs 1971; Krebs 1982; Drent et al. 2003; Wilkin et al. 2006). Egg-laying in Europe generally begins in April. Like many passerines the great tit is a cavity nester and readily utilises nest-boxes (Gosler et al. 2013; Garroway et al. 2013).

Clutch sizes generally range from six to eleven eggs in Europe (Haftorn 2012). In the UK the British Trust for Ornithology's (BTO) Next Record scheme has documented an eight day advance in laying date in the species since the 1960's (Crick & Sparks 1999). Variable patterns of laying date in the species have been observed between populations across Europe (Visser et al. 2003), with advances attributed to an effort to maintain synchrony with peak prey abundance (Charmantier et al. 2008), though some populations have shown no advance at all (Gienapp et al. 2006; Ahola et al. 2009).

1.2.2 Predator and prey phenology

The emergence of winter moth (*Operophtera brumata*) caterpillars in spring, timed to coincide with oak tree budburst is a crucial event in the life cycle of great tits in Europe, which rely on them during the breeding season (McCleery & Perrins 1998; Naef-Daenzer et al. 2000; Visser et al. 2006; Visser, te Marvelde, et al. 2011; Reed, Jenouvrier, et al. 2013). The consequences of a mismatch between budburst and emergence for the caterpillars is so severe (van Asch et al. 2007) that the lifecycle of the winter moth has been experimentally proven to retain synchrony with oak tree budburst even when exposed to elevated temperatures (Buse et al. 1999).

Consistently cited as 'the most important prey for nestling tits' (van Balen 1973), the significance of caterpillar emergence for great tit reproductive success is well documented with lower chick mass, body size and fledgling success all being the consequences for females that breed late relative to the food peak (Reed, Jenouvrier, et al. 2013; Visser et al. 2006; Buse et al. 1999; Gienapp, van Noordwijk, et al. 2013). In comparison, females capable of responding more strongly to temperature and consistently lay earlier confer a greater survival advantage to their offspring and successfully recruit more chicks into the population (Nussey et al. 2005; Perrins 1970; Visser et al. 2006; Reed, Jenouvrier, et al. 2013).

1.2.3 Wytham

Data on lay date come from a long-term study of great tits in Wytham Woods (Lack 1964), a Site of Special Scientific Interest in Oxfordshire, UK (51°46'N, 1°20'W). Designated an SSSI in 1950, the 375 acres of largely deciduous woodland that comprise Wytham Woods have formed the backdrop for numerous ecological studies (Watts 1968; Krebs 1971; King 1976;

Krebs 1982; Kirby & Thomas 2000; Perrins 2001; Macdonald et al. 2002; Wilkin et al. 2006; Savill et al. 2010; Culina et al. 2015).

Monitoring of the great tit population began in 1947 (Lack 1958; Lack 1964) and continues to the present day, overseen and managed by the University of Oxford. About 1020 nest boxes are provided for nesting in Wytham Woods, of which 250-450 are used annually (Garroway et al. 2013). Metal rings with a unique identifier code are affixed to the legs of every adult caught and chick hatched in the site.

Morphological measurements and behavioural observations recorded annually enable the life-histories of individual birds to be documented in detail. About 400 pairs of great tits inhabit the wood with around 2,000 chicks hatching each year. A total of around 80,000 birds have been ringed and studied since records began (B. Sheldon, pers. comm). Morphometric measures such as wing length and characteristics such as age (where known) are recorded for every bird caught. For female birds information associated with each breeding attempt includes various parameters describing the physical location of the nest such as GPS coordinates, altitude and the area of the wood where the nest has been constructed. Wytham is divided into nine areas each with a unique code (Figure 2.1).

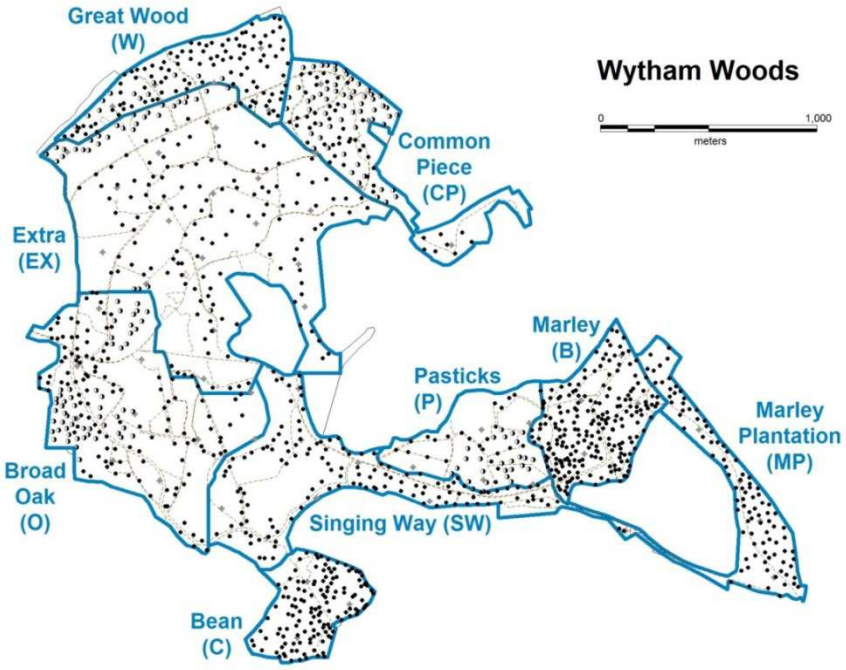


Figure 1.2. Map of Wytham Woods, Oxford (UK) highlighting the nine research areas and position of nest boxes (black dots) (with permission, Reinder Radersma).

The laying date of every breeding attempt is recorded as the number of days after April 1st (designated the official start of the breeding season) that the first egg was laid in each nest. Any breeding events that occurred more than 30 days after the first recorded breeding event of the year were removed as these are likely to represent second clutches (van Noordwijk et al. 1995).

1.2.4 Hoge Veluwe

Insights into great tit ecology in the Netherlands are derived from studies of populations inhabiting four main study sites; Liesbos (51°35'N, 04°40'E), Oosterhout (51°55'N, 05°50'E), Vlieland (53°15'N, 05°00'E) and Hoge Veluwe (52°05'N, 05°50'E). These sites represent some of the earliest and longest running models for long-term monitoring of any wild animal and were the inspiration for great tit and blue tit research as Wytham Woods (Lack 1964; Kluijver 1951).

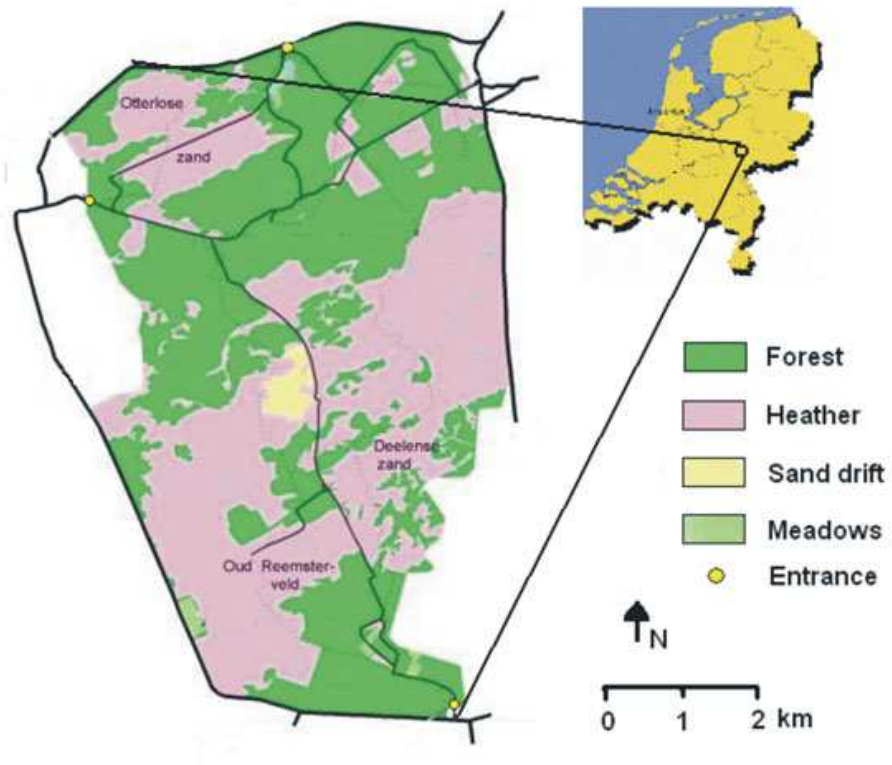


Figure 1.3. Location and map of Hoge Veluwe national park in the Netherlands (from (Hein 2011))

Of these sites the two largest, Vlieland and Hoge Veluwe, though comparable in pedigree size have fundamentally different population dynamics (van Noordwijk et al. 1981; Postma & van Noordwijk 2005; Gienapp, van Noordwijk, et al. 2013). Vlieland is an island; Hoge Veluwe represents 171 ha of mixed pine and oak woodland (van Noordwijk et al. 1981; Visser et al. 2006) (Figure 1.3), more similar in habitat type and ecosystem to Wytham Woods (Perrins 1965; Krebs 1971; Savill et al. 2010). Of the Netherlands populations the great tit population of Hoge Veluwe is the focus for analyses of patterns of reproductive behaviour in this study. Four hundred nest boxes are situated in the park and an average 125 pairs of great tits breed in the area (Gienapp et al. 2006).

The same principles used to collect phenotypic and genetic data from the great tits of Wytham Woods are applied to data collection in Hoge Veluwe. Birds are caught and ringed with metal and / or colour rings during the breeding season (chicks are ringed in the nest) to identify individuals. Assuming one egg is laid per day, laying date is calculated from the

number of eggs found in the nest and recorded as number of days from April 1st (Gienapp et al. 2006; Visser et al. 2006). Only first clutches are used in analyses.

1.3 Patterns of laying date in *Parus major*

The existence of several long term studies of the great tit *Parus major*, some of which are the longest in existence for any wild vertebrate (Krebs 1971; Krebs 1982; Visser et al. 2003; Garant et al. 2004; Matthysen et al. 2011; Garroway et al. 2013; Gienapp, van Noordwijk, et al. 2013), has ensured that there is an extensive record of laying date trends in this particular species. Variable responses to perceived mismatches have resulted in substantial variation between populations of the species.

A comparative study of the breeding responses of 13 great tit populations identified 5 sites where laying date has significantly advanced with marked variation in the responses of populations inhabiting different locations (Table 1.1) (Visser et al. 2003). Finnish great tit populations consistently respond to variation in breeding time temperature by laying earlier in warm springs, though the shifts are non-significant (Ahola et al. 2007; Ahola et al. 2009). In contrast temperature increases of 0.091˚C over a 29 year period resulted in a significant advance of 0.42 days / year in the laying dates of great tits in Belgium, the breeding phenology of birds in this location able to maintain synchrony with the peak in caterpillar abundance (Table 1.1) (Matthysen et al. 2011).

The laying date of the Wytham Woods great tits has advanced by about 14 days over 47 years in response to the earlier emergence of their winter moth larvae prey (Charmantier et al. 2008). The best predictor for laying date in this population was estimated to be temperatures between 15th March - 25th April (Husby et al. 2010) and mean temperatures in Wytham in this period increased by 0.048˚C / year between 1965-2008 (Garant et al.

2008), with a concurrent shift in laying date of 0.37 days / year (Table 1.1) (Husby et al. 2010).

Conversely early studies documenting the laying date of the Hoge Veluwe great tits found no advancement, resulting in asynchrony between chick hatching and caterpillar abundance (Visser et al. 1998; Visser et al. 2006). Between 1973-2003 laying date advanced by 5.4 days in response to temperature increases of 1.7˚C over this period but the shift was non-significant (p 0.06) and insufficient to maintain synchrony with prey phenology (Gienapp et al. 2006). However given strong directional selection and mean laying dates closely associated with mean half fall (peak caterpillar abundance) the expectation was that this population would eventually show an adaptive trend for earlier laying (Gienapp et al. 2006; Visser et al. 2006).

Laying date is closely related to temperature in both the Wytham and Hogue Veluwe populations, with pre-breeding temperatures increasing at similar rates and similar slopes in the reactions of both populations in each location (Husby et al. 2010). The relationship is significantly weaker in Hogue Veluwe with laying date shifting at half the rate of that observed in the Wytham population (HV; 0.2 days / year, WY; 0.37 days / year) (Husby et al. 2010). However recent evidence of the first significant early laying trend in the Hogue Veluwe population (Table 1.1) (Husby et al. 2010) suggests that phenological changes may have been masked in earlier studies by insufficient data coupled with a small magnitude of response.

Changes in environmental conditions and the phenology of other organisms at fine spatial scales may provide some explanation for the variation in laying date trends between populations. The absence of significant temperature changes in the most northern latitudes and evergreen (and thus unchanging vegetation / prey phenology) in the most southerly latitudes renders shifts in laying date unnecessary (Visser et al. 2003), reflected in the

absence of a response (Table 1.1). The defining patterns exhibited by Wytham and Hogue Veluwe are an adaptive population-level response in Wytham with females relatively invariant in their own responses to temperature, contrasted by high levels of individual plasticity but an insufficient population response in Hogue Veluwe (Nussey et al. 2005; Charmantier et al. 2008).

Contrary to other studies that have attributed trends directly to climate change, Visser et al. (2003) found temperature alone insufficient to explain variation in laying date, but instead found evidence that a decreasing incidence of second broods was counterbalancing selection for earlier laying, partially accounting for failures to synchronise laying date (and subsequently hatching date) with prey abundance. While laying date has a significant heritable component, the response to selection may be too small to be detected by current statistical methods (Gienapp et al. 2006).

1.3.1 A genetic component to laying date

With the development of advanced evolutionary genetic and statistical techniques, real-world studies of avian species have drawn a general conclusion that laying date is primarily determined by environmental as opposed to genetic factors (Thorley & Lord 2015; Perdeck & Cavé 1992; Hakkarainen et al. 1996; Charmantier et al. 2008; Wiggins 1991). Among many species, the heritability of laying date is described as modest, and not the largest source of variation, relative to phenotypic plasticity (Thorley & Lord 2015; Nussey et al. 2005; Charmantier et al. 2008; Husby et al. 2010; Brommer 2013; Garant et al. 2008; Gienapp et al. 2006).

However the literature consistently reports a significant additive genetic component of variation for laying date in great tits. From as early as 1981 significant heritable variation was detected for two out of four populations from long term study sites in the Netherlands

(Noordwijk et al. 1981), with support coming from estimates made in subsequent decades (Nussey et al. 2005; Gienapp et al. 2006). The highest estimate (0.30) was obtained in 2005 (Nussey et al. 2005), while the most recent estimate, of approximately 0.17, concurs with the earliest estimate (van Noordwijk et al. 1981; Gienapp et al. 2006).

Heritability estimates for the Wytham population have varied more widely. Both the time period considered, the methodological approach taken, and the abiotic / biotic factors taken into account in the calculation have resulted in estimates ranging from 0.03 to 0.24 (van der Jeugd & McCleery 2002; McCleery et al. 2004; Garant et al. 2008; Liedvogel et al. 2012). In general it appears laying date has a modest, but significant genetic component in most great tit populations including the ones studied in this thesis.

Evidence for the role of genetics in determining laying date has been illustrated by experiments in the Netherlands on captive great tits. Visual stimuli alone were shown to be insufficient to generate an early breeding response (Schaper et al. 2011) and female birds from early-laying maternal lineages laid their eggs earlier, suggesting that the timing of egg-laying is heritable (Visser, Schaper, et al. 2011). This was further supported by subsequent experiments that found that birds from early or late laying maternal lineages responded differently to temperature cues during certain periods, with a cold February / March followed by a temperature increase advancing the onset of laying in early layers and delaying it in late layers (Schaper et al. 2012).

However studies with the capacity to investigate microevolution for laying date in response to climate change have found it either to be absent (Gienapp et al. 2008) or attribute variation to phenotypic plasticity (Charmantier et al. 2008; Nussey et al. 2005; Gienapp et al. 2006).

1.4 Gene mapping and quantitative genetics

Gaining an understanding of the genotypes that explain phenotypes and the effect of naturally occurring genetic variation on the fitness of populations are some of the most fundamental pursuits of evolutionary genetics (Falconer & Mackay 1996; Stinchcombe & Hoekstra 2008; Mackay et al. 2009). To understand how an organism can adapt by means of a micro-evolutionary response requires a method by which we can determine the extent to which variation in certain traits is influenced by genetic variation. Mathematical modelling tools exist to enable us to determine whether the potential for animals to adapt is in their genes. With pedigree and phenotypic information for animal populations quantitative genetics can enable us to elucidate the mechanisms of inheritance and their evolutionary potential (Wilson et al. 2010).

Quantitative genetics is the study of continuously varying traits, those that do not fall into discrete classes, human height being the classic example (Galton 1886; Fisher 1918; Yang et al. 2010). Quantitative genetics is a statistical method that uses so called 'animal models' to separate phenotypic variation into its genetic and environmental components (Wilson et al. 2010). The science of quantitative genetics has been around for around a century, and is primarily used to improve yields in plants and animals under artificial selection (Nezer et al. 1999; Venuprasad et al. 2009; Hayes et al. 2010; Dekkers 2012). The potential of quantitative genetics as a tool in the study of evolutionary biology was realised in the 80s, and has since been used to study the inheritance of phenotypic traits in a wide variety of organisms (Kruuk 2004; Slate 2005; Beraldi, McRae, Gratten, Pilkington, et al. 2007; Kruuk & Hill 2008; Hernández-Sánchez et al. 2010).

1.4.1 Identifying genes responsible for heritable variation in wild animals

Identifying the genes underpinning variation in phenotypic traits is a fundamental goal of evolutionary biology (Ellegren & Sheldon 2008; Slate et al. 2010). The number and magnitude of loci can enable predictions to be made about potential responses to selection, with traits likely to respond faster if variation is mediated by loci of large effect (Charmantier et al. 2014). Classical quantitative genetic theory assumes that continuously varying or quantitative traits are likely to have complex (polygenic) genetic architectures with many genes of small effect contributing to overall phenotypic expression (Falconer & Mackay 1996; Lynch & Walsh 1998). This has largely been substantiated in the human and domestic animal scientific literature with variation in quantitative traits such as human height ascribed a polygenic basis (Yang et al. 2010; Visscher 2008; Lango Allen et al. 2010; Yang, T. Lee, et al. 2013; Franke et al. 2010; Dekkers 2012; Do et al. 2013; Carneiro et al. 2014; Daetwyler et al. 2008; Kolbehdari et al. 2008; Goddard & Hayes 2009; Buckler et al. 2009).

The challenge of describing the genetic architecture of traits in wild animals is typical exacerbated by small sample sizes. Genomic analyses performed on domestic animals and humans benefit from ease of access to their subjects and datasets can comprise hundreds to hundreds of thousands of subjects (Yang, T. Lee, et al. 2013; Visscher et al. 2007; Pimentel et al. 2011; Do et al. 2013; Vaysse et al. 2011). A consequence of the complexity of studying wild organisms is that few long-term studies have accumulated datasets with sufficient individuals and generations to make the detection of genetic variants of small effect viable. Despite some occasional evidence for single genes of genuinely large effect (Paterson et al. 1991), the magnitude of QTL tends to be overestimated when small sample sizes are used, a phenomenon known as the ‘Beavis effect’ (Beavis 1994; Slate 2013). For polygenic quantitative traits, where many small effect QTL contribute to trait variation,

small sample sizes may result in reduced (or even zero) statistical power to identify those regions harbouring genetic variants (Kroymann & Mitchell-Olds 2005).

In wild animal populations linkage mapping has identified QTL for a variety of morphological traits. The first QTL were mapped in a pedigreed population of red deer on the Isle of Rum (Slate et al. 2002) followed by linkage mapping of phenotypic traits in other ungulates such as Soay and bighorn sheep (Hernández-Sánchez et al. 2010; Beraldi, McRae, Gratten, Slate, et al. 2007; Beraldi, McRae, Gratten, Pilkington, et al. 2007; Gratten et al. 2010; Gratten et al. 2007; Poissant et al. 2012).

Contrary to the widely observed polygenicity of complex human traits (Yang, T. Lee, et al. 2013; Lango Allen et al. 2010; Yang, Manolio, et al. 2011) a number of linkage mapping studies of wild animals have identified loci of large effect, for horn size in the Soay sheep (Johnston et al. 2010; Johnston et al. 2011), wing length in the great reed warbler (Tarka et al. 2010) and morphology between benthic and limnetic forms of the three-spine stickleback (Colosimo et al. 2004).

Replication is the surest way of distinguishing whether traits in wild animals depart from the polygenic model (Schielzeth & Husby 2014; Charmantier et al. 2014). The locus on chromosome 10 responsible for horn-type in Soay sheep was replicated using a variety of statistical techniques (Johnston et al. 2010; Johnston et al. 2011) but the locus does not affect the same phenotype of many other bovid species (Johnston et al. 2011).

Direct comparisons between populations of the same species are even more challenging as cases where sufficient genomic and phenotype data are available for multiple distinct populations are rare. Colosimo et al. (2004) were able to determine that a single locus of large effect explains variation in the same trait in geographically separated populations of

threespine sticklebacks (*Gasterosteus aculeatus*) but few other such examples exist for wild animals.

Genome-wide analyses provide support for many loci of small effect contributing to variation in complex behavioural and physiological traits across wild animals as a result of increasingly sophisticated genotyping technologies and denser marker maps (Houde et al. 2013; Johnston et al. 2014; Lotterhos & Schaal 2014; Weber et al. 2013; Greenwood et al. 2015). In great tits, studies of wing length (Robinson et al. 2013), clutch size and egg mass (Santure et al. 2013) employing range of statistical methods found no evidence for QTL of large effect for any trait. A polygenic genetic architecture for clutch size was supported by evidence from collared flycatchers where a significant association detected on chromosome 18 explained only a small fraction of the phenotypic variance (3.9%), leading the authors to conclude that variation in clutch size is most likely determined by many genes of small effect (Husby et al. 2015).

Few studies have succeeded in identifying the genetic variants associated with phenotypic variation in wild organisms. Of those studies to attribute a genetic component to phenotypic variation (Réale et al. 2003; Møller 2001; Pulido & Berthold 2003) the conclusions in most cases have been based on purely phenotypic, as opposed to genetic data (Gienapp et al. 2008), and in a number of cases describe the non-genetic component (responses to environmental cues and phenotypic plasticity) as explaining a relatively larger proportion of the phenotypic variance (Tarka et al. 2015).

Environment-induced changes in an individual's behaviour, morphology or physiology, are collectively termed 'phenotypic plasticity' (Price et al., 2003) and are considered essential for survival in a heterogeneous environment (Lind and Johansson, 2011) and for enabling rapid short-term responses to increase the chances of survival (Visser et al. 2004; Yeh & Price 2004; Vedder et al. 2013). Phenotypic plasticity as opposed to microevolution is

responsible for adaptation in a number of traits in vertebrates (Canale & Henry 2010; Gienapp et al. 2008) including laying date and migration in avian species (Tarka et al. 2015; Charmantier et al. 2008; Nussey et al. 2005; Przybylo et al. 2000; Dingemanse et al. 2012; Vedder et al. 2013; Reed, Grotan, et al. 2013; Gienapp et al. 2006).

However, the reaction norms that allow individual's flexibility in their environment are unlikely to be optimal outside of the conditions in which they evolved (Ghalambor et al. 2007; Charmantier et al. 2014; Visser et al. 2004), resulting in reduced survival and increased chance of extinction (Lynch & Lande 1993; Burger & Lynch 1995). A genetic component to trait variation, upon which selection can act, would enable adaptation by means of microevolution and therefore could enable the persistence of populations.

1.4.2 *Mapping QTL*

QTL (quantitative trait loci) are regions of the genome that explain some of the variation in continuously varying traits (Slate et al. 2010). Linkage mapping of QTL can be performed subject to three basic requirements; a genetic map of variable markers, a pedigree with which to follow segregation of those markers and phenotypic data for life history traits such as reproductive success for each individual included in the pedigree (Slate 2005).

Linkage and association (linkage disequilibrium) mapping are two methods commonly employed to dissect the genetic architecture underpinning quantitative traits. Linkage mapping using an animal model / variance components approach partitions potential sources of trait variation into their genetic and non-genetic components using available pedigree information (Slate et al. 2010; Hernández-Sánchez et al. 2010). Genome Wide Association Mapping (GWAS) scores thousands of SNPs simultaneously to detect statistical associations based on the amount of LD between causal variants and markers (Slate et al. 2010; Pritchard & Przeworski 2001; Korte & Farlow 2013).

Linkage mapping exploits recent recombination events to examine the co-segregation of marker and trait alleles in a population of genotyped individuals (Hirschhorn & Daly 2005; Charmantier et al. 2014), and has considerable power to detect loci of large effect (Jimenez-Sanchez et al. 2001; Cardon & Bell 2001; Hirschhorn & Daly 2005). However, linkage analysis has much less power to identify causal variants in the absence of clear segregation (Cardon & Bell 2001), maps QTL with limited resolution and struggles to accurately estimate effect sizes (Hirschorn & Daly 2005; Slate et al. 2010; Korte & Farlow 2013).

Using ancestral recombination to detect statistical associations between marker and trait alleles (Cardon & Bell 2001; Charmantier et al. 2014), association mapping can detect QTL at much higher resolution and estimate effect sizes with a greater degree of accuracy (Korte & Farlow 2013). The limitations of association analyses include difficulty in detecting variants of low penetrance or small effect (Wang et al. 2005; Korte & Farlow 2013), and a tendency to overestimate effect sizes in the absence of sufficient sample sizes (Cardon & Bell 2001; Göring et al. 2001). Linkage mapping may lack precision but is also less prone to Type I error (Slate et al. 2010; Charmantier et al. 2014).

Strong physical linkage between a SNP marker and QTL increases the likelihood of detecting a statistically significant association between that SNP and the trait of interest (Slate 2005). GWAS analyses have the potential to detect QTL at much higher resolution than linkage mapping, but the requirement of high marker densities meant that for a long time their application was beyond the scope of wild animal genomic studies. The increasing availability of genomic information, together with more advanced and cheaper genotyping technologies is increasingly enabling the genetic basis of trait variation to be described for a much wider variety of species (Bush & Moore 2012; Hill 2012; Schielzeth & Husby 2014).

Early mapping studies suffered from being unable to pinpoint QTLs to a precise location (Slate et al. 2009). Constraints were imposed by limited genomic information, few typed individuals, and insufficient marker densities to detect marker alleles in linkage disequilibrium with genes responsible for phenotypic variation (Slate et al. 2010). The development of new more efficient and cost-effective genotyping technologies (Metzker 2010; Hudson 2008; Narum et al. 2013), together with the discovery of a suite of new genetic markers such as single nucleotide polymorphisms, or SNPs (Lynch & Walsh 1998), greatly facilitated the rise of molecular quantitative genetics as one of the primary tools for our understanding of evolution in wild populations. Next-generation sequencing technologies have the capacity to produce extensive quantities of data cheaply, in some cases performing billions of sequencing reactions simultaneously (Metzker 2010).

1.4.3 Partitioning variance

Complimentary to both linkage analyses and association mapping, genome or chromosome partitioning is a statistical technique that does not identify individually associated SNPs but estimates the cumulative contribution to phenotypic variation of markers in distinct (normally autosomal) blocks. The method works by comparing the similarity between relatives as an exact proportion of the genes they have in common (Visscher et al. 2006). Relationship matrices consisting of estimates of marker based relatedness between individuals are constructed for each chromosome and fitted within an animal model (Robinson et al. 2013). Variance is then partitioned across chromosomes and statistical testing used to determine whether any contribute a significant proportion of the additive genetic variation.

Under a polygenic mode of inheritance it would be expected that the amount of variation explained by each chromosome would positively scale with features such as chromosome size or length as bigger blocks harbour more genes and subsequently are the source of a

greater amount of variation. The theoretical expectations of the method have been substantiated by descriptions of the polygenicity of complex traits in a number of organisms. Relationships where variation explained was proportional to chromosome size or length has been observed for height and body mass index (Visscher et al. 2007; Yang et al. 2010; Yang, Manolio, et al. 2011; Yang, T. Lee, et al. 2013), as well as neurobehavioural disorders OCD, Tourette syndrome and schizophrenia (Lee et al. 2012; Davis et al. 2013) in humans. Similar observations have been made for commercially important production traits and disease resistance in cattle, sheep and chickens (Pimentel et al. 2011; Al-Kalaldeh et al. 2013; Yi et al. 2015), as well as clutch size and wing length in great tits (Robinson et al. 2013; Santure et al. 2013), the first instance of genome partitioning being applied to a wild animal population.

1.4.4 SNP markers

Modern quantitative studies utilise single nucleotide polymorphism (SNP) markers (Morin et al. 2004; Mackay et al. 2009; Slate et al. 2009; van Bers et al. 2010; Speed et al. 2012). SNPs are point mutations that result in a single nucleotide alteration in the genome sequence (Brumfield et al. 2003) and are the most abundant form of genetic variation to occur in most, if not all, organisms. The high frequency at which SNPs occur throughout the genome mean it is highly likely to find SNPs in close proximity to genes / genomic regions of interest (Seddon et al. 2005), which predisposes them to studies of additive genetic variation. The accumulation of genomic information for a variety of species adds to the suitability of SNPs for analysis of ecological and evolutionary processes (Morin et al. 2004).

Their simple mutational nature and the ease with which they can be screened, makes SNPs suitable for rapid and cost-effective gene mapping (Brumfield et al. 2003). SNP maps have accelerated genetic mapping in a number of species including the zebra-fish (*Danio rerio*) (Stickney et al. 2002) and mouse (*Mus musculus domesticus*) (Lindblad-Toh et al. 2000) with

new technologies able to infer genotypes from under-sequenced animal populations and thus more effectively map QTL (Baird et al. 2008). Twenty thousand novel SNPs detected before the existence of a sequenced genome in the great tit (*Parus major*) (van Bers et al. 2010) provided one of the largest frameworks for molecular analysis of the evolutionary genetics of a wild bird.

1.4.5 Great tit genetic resources

Extensive efforts have been made to generate a comprehensive set of resources that can be used for analyses of the genetic architecture of phenotypic variation in the great tit (van Bers et al. 2010; Santure et al. 2011; van Bers et al. 2012; van Oers et al. 2014). In the absence of a sequenced genome for the great tit, following a methodology proposed by Kerstens et al. (2009) for the development of genetic resources for unsequenced species, the mapping of short sequence reads into contigs enabled the discovery of over 20,000 novel SNP markers (van Bers et al. 2010).

Linkage maps were then constructed for application to linkage and association mapping. Based on alignment of contigs to the Zebra finch genome, the total genome coverage for two great tit populations from the Netherlands and the UK was 97.1% for the NL map and 98.3% for the UK map (Santure et al. 2011). The framework linkage maps for both the UK and NL populations are constructed for 32 of the great tit chromosomes (1-15, 17-24, 26-28, 1A, 4A, 25A, 25B, LGE22 and Z) (van Oers et al. 2014).

For the UK population the framework map is based on 1656 individuals and consists of 1674 single nucleotide polymorphism (SNP) markers covering 1893cM. A denser 'parsimonious' linkage map was constructed using 5591 markers, consisting of 4878 'chromosome-assigned' SNP markers and 713 markers putatively assigned a chromosomal location based on comparison with the Zebra finch genome (Santure et al. 2013). The NL

framework map contains 1674 markers, and the number of SNP markers placed on the NL parsimonious linkage map was 5582 spanning a length of 2010cM (van Oers et al. 2014).

The difference and significance between framework and parsimonious linkage maps relates to their marker density and the downstream analyses that utilise them. The framework linkage map only includes those markers that can be ordered with support that is statistically significantly stronger than any alternative order, while the parsimonious map includes less informative markers ordered in the most likely order, but where alternative orders have almost equally strong support (Tarka et al. 2010). The parsimonious map increases the power of GWAS by giving greater genome coverage.

1.5 Aims and objectives

The complexity of mechanisms underlying variation in quantitative traits ensures that there are many unanswered questions in the field of evolutionary biology. Primary among the goals of quantitative genetics is the potential to be able to identify causative genes and alleles for fitness related traits. We can begin to address the substantial gaps in our knowledge of this area by employing quantitative genetic and gene mapping techniques to pedigreed wild populations.

The purpose of this thesis was, for two geographically distinct populations of the great tit *Parus major* **1)** to determine whether loci responsible for variation in laying date can be mapped, by employing linkage mapping, association analyses and chromosome partitioning**, 2)** to provide a description of the underlying genetic architecture of an important life-history trait and **3)** compare the results between populations, to discern whether the same genomic regions underlie variation in laying date in two populations of the same species.

Chapter 2

Molecular quantitative genetic analyses for loci contributing to variation in laying date

Jenny C. Armstrong[1], Ben Sheldon[2], Anna Santure[1,3], Isabelle De Cauwer [1,4], Jon Slate[1]

[1]Department of Animal & Plant Sciences, Alfred Denny Building, University of Sheffield, Western Bank, Sheffield S10 2TN, UK; [2]Edward Grey Institute, Department of Zoology, University of Oxford, Oxford OX1 3PS,UK; [3]School of Biological Sciences, University of Auckland, Auckland, New Zealand; [4]Laboratoire de Génétique et Evolution des Populations Végétales, UMR CNRS 8198, Bâtiment SN2, Université des

2.1 Abstract

Anthropogenic climate change is predicted to impact ecosystems globally but it is the capacity of organisms to adapt to changes at a smaller localised scale that will ultimately determine the persistence of populations and potentially entire species. One of the most important factors influencing survival in wild animals is the timing of breeding in relation to food availability, with the survival of offspring dependent on the availability of ample resources around the time of birth. So crucial is the timing of this event that the evolution of breeding behaviour of many species has resulted in intricate synchronisation with food abundance or prey availability. As a warming climate alters the phenology of species at lower trophic levels, organisms dependent on such resources will have to modify their own phenology to ensure the survival of their offspring. Whilst rapid environmental changes may be expected to elicit rapid evolutionary responses the actual rate of response exhibited by organisms is likely to be influenced by many factors. Chief among these is presence of genetic variation. It is generally considered that the rate of response of any given phenotype to selection is proportional to the amount of additive genetic variation that exists for the trait. Detecting the amount of genetic variation attributable to a trait is dependent on extensive phenotypic and genomic information the absence of which is notable for wild animal species. Using data from one of the few long-term studies of a wild vertebrate, the great tit *Parus major*, I employ linkage and association statistical approaches to describe the genetic architecture of a behavioural life history trait and attempt to identify genetic variants associated with observed phenotypic variation. No regions or genetic markers of genome-wide significance were observed from genome scans. Evidence suggests variation in laying date is influenced by a small but significant genetic component consisting of many genes of small effect.

2.2 Introduction

With a changing climate predicted to have a range of impacts on animal populations the ability and speed with which populations can adapt may prove a deciding factor in their survival and persistence. In line with the concept that for any given trait the rate of response to selection is conditional on the amount of additive genetic variation of the trait (Fisher 1930), theoretically a genetic component to life history or fitness associated behavioural traits would enable a micro-evolutionary response. Though directional selection would be expected to erode additive genetic variation (Falconer & Mackay 1996), a heritable basis for phenotypic variation of fitness-related traits has been documented for several wild animal species (Mousseau & Roff 1987; Réale et al. 1999; Merilä, Sheldon, et al. 2001; Slate 2005; Ellegren & Sheldon 2008; Teplitsky et al. 2009).

Despite a consensus that behavioural variation is to a large extent mediated by rapid responses to environmental cues, so-called 'plasticity' (Przybylo et al. 2000; Price et al. 2003; Yeh & Price 2004; Nussey et al. 2007; Brommer 2013; Vedder et al. 2013; Charmantier & Gienapp 2014) evidence for the role of a genetic component in mediating behavioural responses also exists (Dingemanse et al. 2002; Stirling et al. 2002; Drent et al. 2003; Réale et al. 2009; Visser, Schaper, et al. 2011). However, quantifying the genetic component of variation through the identification of causative loci has proven challenging with many studies constrained by insufficient phenotype and genomic information to conduct genomic analyses (Slate 2005).

The great tit *Parus major* represents one of the few wild organisms for which several long-term studies provide a substantial resource for studying the factors influencing behavioural variation. This common passerine synchronises its breeding to the emergence of the winter moth caterpillar *Operophtera brumata*, which form such an important component of the nestlings diet that chick hatching roughly coincides with peak abundance of this essential

food source (Lack 1958; van Balen 1973; Perrins & McCleery 1989; van Noordwijk et al. 1995; Eeva et al. 2011) (Chapter 1).

Caterpillar emergence is strongly correlated with oak (*Quercus robur*) bud burst, itself determined by temperature , and caterpillars are only present in large quantities for a short period of time (Gibb 1950; Smith et al. 2011; Visser et al. 2006). Subsequently the window for great tits to take advantage of this resource is small. Increasing spring temperatures in Wytham Woods, Oxford have caused a shift to earlier emergence of the winter moth caterpillar, consequently resulting in a corresponding shift to earlier lay dates in great tits.

A laying date advancement of 14 days over 47 years has been documented in this population, attributed to the birds closely 'tracking' the emergence of their prey (Charmantier et al. 2008). This change in the breeding phenology is not unique to this population, with a cumulative assessment of patterns of lay date in 13 European great tit populations detecting significant advances in lay date for 5 of the populations, including Wytham and a second UK site (Visser et al. 2003).

Given the suitability of this species as a model system for studying the genomics of morphological and behavioural trait variation, research has already begun to investigate the potential for a genetic component to various behaviours. A heritable basis for risk-taking and exploratory behaviours in great tits from the Netherlands (Dingemanse et al. 2002; van Oers et al. 2003) form some of the first studies to demonstrate a genetic component to behavioural variation in a wild organism.

Several studies have identified a significant heritable component of laying date from long-term studies in the Netherlands (van Noordwijk et al. 1981; Nussey et al. 2005; Gienapp et al. 2006) and in the Wytham population, with estimates in the range of 0.03-0.24 (van der Jeugd & McCleery 2002; McCleery et al. 2004; Garant et al. 2008; Liedvogel et al. 2012).

The most recent study to estimate heritability for lay date in the Wytham great tits used a model accounting for fixed and random effects (factors such as female age, which influence lay date in individuals) and showed low but significant levels of additive genetic variation for the trait with a h^2 of 0.03 (SE $\pm$ 0.01) (Liedvogel et al. 2012).

Whilst research has indicated the presence of a significant genetic component to avian laying date (Perdeck & Cavé 1992; Hakkarainen et al. 1996; Wiggins 1991; McCleery et al. 2004; Nussey et al. 2005; Gienapp et al. 2006; Brommer et al. 2008; Garant et al. 2008; Liedvogel et al. 2012; Thorley & Lord 2015) few studies have had the capacity to attempt a description of the trait at the genomic level. A candidate gene approach detected significant associations between circadian clock gene CLOCK and breeding phenology in the blue tit and barn swallow (Liedvogel et al. 2009; Caprioli et al. 2012) but no such association with timing variables has ever been detected in the great tit (Liedvogel & Sheldon 2010; Liedvogel et al. 2012).

Using a phenotypic dataset spanning 26 years and the recently constructed great tit linkage map (van Bers et al. 2010; van Bers et al. 2012; van Oers et al. 2014) and genome assembly (Laine et al., manuscript in revision), I employ linkage and association statistical analyses to conduct scans of the great tit genome and describe the genetic architecture underlying variation in laying date.

2.3 Methods

2.3.1. Phenotype data

The phenotypic dataset for QTL analyses contained 1142 breeding events for 671 females and 767 males. GWAS analyses were able to include additional genotyped individuals that were previously excluded from QTL analyses because they were uninformative / absent from the pedigree. Therefore the GWAS phenotypic dataset consisted of 1523 breeding events for 951 females. Breeding event records span a 26 year period from 1985 to 2011.

2.3.2. Pedigree

Long-term observation of the Wytham great tit population has enabled the life histories of birds across many generations to be recorded in depth, and is recognised as rare example of a wild animal population for which a comprehensive pedigree exists. Genetic information is obtained by extracting a blood sample from each of the birds as they are caught and ringed during each breeding season. The QTL pedigree contains 2709 individuals linking 1733 genotyped individuals of know parentage. A second dataset consisting of 2497 genotyped birds was used for GWAS, from which genotype information for 951 females was extracted for use in the analysis.

2.3.3. Marker data

From 1985 blood samples have been obtained from individual birds as they are caught for the purposes of ringing / to make morphological measurements. Genetic information obtained from such samples enabled over two thousand UK birds to be genotyped using a 10K SNP chip (described in detail elsewhere (van Bers et al. 2010; van Bers et al. 2012)). Each of the chromosomes (linkage groups) range in length between 3.68 - 139.88cM (van Oers et al. 2014). The different downstream analyses required linkage maps of varying

marker densities. The QTL and GWAS analyses were performed using the framework and parsimonious linkage maps respectively (van Oers et al. 2014) (Chapter 1).

Identity-by-descent (IBD) matrices were then generated describing the relatedness of pairs of birds at each position on the genome using the software LOKI (ver2.4.5) (Heath 1997; Heath et al. 1997; as described in Santure et al. 2013).

2.3.4. Statistical analysis

Constructing an animal model

Prior to QTL mapping, animal models were employed to identify variables (known as 'terms') with a significant influence on phenotypic trait variation and to estimate the amount of heritable variation for laying date. Factors influencing trait variation and subsequently incorporated into the animal model were identified by statistical analysis using two statistical softwares. Fixed factors were determined by linear modelling using computing software R (version 3.0.0, (2013)) and random factors determined by comparison of models using a likelihood ratio test in ASReml (version 3.0) (Gilmour et al. 2002).

All variables with the potential to affect lay date were included as terms in an initial model (model 1). Variables outlined in blue were categorised as 'fixed' and variables outlined in purple categorised as 'random' (Table 2.1);

Layd = age + age^2 + altitude + femaleID + femalePE + maleID + area + year + density_poly

(model 1 – full model prior to removal of non-significant terms)

Table 2.1. Description of the variables included as terms in the initial version of the animal model and whether they were designated as 'fixed' or 'random' effects

Term	Description	Classification
age	True age in years. Where age was unknown, estimated 2 years old from first capture	Fixed
age^2	Age quadratic	Fixed
altitude	Height of nest box above sea level (m). Altitude varies by 106 m within the study site	Fixed
femaleID / maleID	Term capturing the additive genetic component of variation (VA); relies on estimating the relationship matrix between pedigreed individuals	Random
femalePE	Female 'permanent environment' effect. Term which takes into account repeated measures on individual females (necessary where multiple breeding events for individual females may be recorded over the course of their lifetime). This term estimates additional non-genetic sources of variance among individuals (Wilson et al. 2010)	Random
area	Wytham is divided into 9 areas, each with a unique code	Random
year	Year of data collection	Random
density_poly	Area around each nest box transformed into polygon dependent on density of nests in immediate area; large polygon indicates nest sits in sparsely occupied region, small polygon indicates close proximity of neighbouring nests	Random

A two-step process of elimination was then employed to determine which terms should be retained in the final model. The extent to which each term improved the fit of the model to the data was determined by significance testing, by comparing hierarchical models of increasing complexity (Wilson et al. 2009). The method of significance testing was dependent on each term's classification as either 'fixed' or 'random'.

Designating a term as fixed indicates explicit interest in how each level of the term affects the trait. Terms assigned as random effects account for other sources of non-independence in the data (Wilson et al. 2010). Fixed terms were decided by linear modelling (lm function) in R (stepwise backward method) (Appendix 2.1), random terms by significance of the likelihood ratio tests (LRT) derived by comparison of models excluding (null model) and including the term being tested, conducted in ASReml, and subsequent chi-square significance testing (Table 2.2).

Table 2.2. Statistics associated with the random effects included in the animal model, the likelihood ratio test (LRT) score obtained by comparison of the logarithm of the odds scores of the models excluding (LOD1) and including (LOD2) the relevant term. Variance components of the final polygenic model are highlighted

Term	**LRT**	**p value**	**Component**	**Comp/SE**	**D.O.F**
Year	586.32	<0.01	52.84	2.80	1
Area	34.78	<0.01	1.30	1.63	1
FemaleID	6.06	0.01	3.84	2.17	1
FemalePE	67.38	<0.01	3.36	1.82	1
Residual variance			14.73	15.91	
Total variance			76.07		1138

Using a chi-square test of significance the model incorporating an additive genetic term ('polygenic' model) (3) proved a better fit to the data than the model with no term capturing the additive genetic variation (2) ($p < 0.05$) indicating a significant additive genetic component (Table 2.2).

Layd = age + age^2 + altitude + year + area + femalePE

(model 2 – estimating variance components without additive genetic component)

The non-significant terms were sequentially deleted with each modification of the model. At the end of the process, only those terms found to be statistically significant ($p<0.05$) in contributing to trait variation remained (model 3).

Layd = age + age^2 + altitude + femaleID + femalePE + area + year

(model 3 – polygenic model)

This model, the 'polygenic' model, enables the estimation of the additive genetic component of variation (V_A), permanent environment variance (V_{PE}) and residual variance (V_R) from which total phenotypic variation (V_P), heritability (h^2), permanent environment effect (pe^2) and residual variance (r^2) can be calculated.

Complex interactions in the form of regression trees were modelled in R by means of R package 'MASS' (Venables & Ripley 2002). Regression trees, or tree models, were constructed to visualise interactions between fixed effects and laying date. This statistical analysis recursively partitions response variables into subsets based on their relationship with any number of predictor variables (Prasad et al. 2006). Splits in the tree are determined by an exhaustive search procedure, maximising between group sum of squares as in analyses of variance (ANOVA analyses) (De'ath & Fabricius 2000; Prasad et al. 2006). The result is a diagram illustrating a succession of splits from the top or 'root' of the tree, with mean values of the response variable at the terminal nodes (Prasad et al. 2006).

Of the terms considered in the original animal model (model 1), density-poly and male identity were dropped on the basis they do not significantly contribute to variation in lay date. Density-poly, a value given to each nest describing the proximity of other nests, was initially included based on a rationalisation that birds would nest at higher densities in habitat ideally suited to the lifecycle of the winter moth. Subsequently lay dates in the nests of these birds would be earlier due to increased prey availability and closer tracking of the emergence of their caterpillar prey.

Male identity was initially included to account for any indirect genetic effects of the male partner on lay date through resource provisioning or territory acquisition, but consistent with the findings of similar studies the contribution of the male to timing of breeding events (Pettifor R. A., Sheldon, B., Browne, CW. J., Rasbash, J. & McCleery, R. H., unpublished, from McCleery et al. 2004; Wilson et al. 2010) was found to have no such influence. Laying date is hereafter modelled as a female-only trait.

Scanning for QTL

To identify the presence of QTL identity by descent (IBD) matrices constructed at regularly spaced (1cM) intervals across the genome were incorporated into the polygenic model as an additional random effect (George et al. 2000). These models are termed 'QTL models' (4). Matrices were constructed for 31 chromosomes at intervals of 1cM using the pedigree, consisting of 2709 individuals, and framework map (van Oers et al. 2014), and generated using the program Loki (Heath 1997). A QTL model with an IBD matrix specific to the genomic region was created for each of 2077 chromosomal locations.

Layd = age + age^2 + altitude + femaleID + femalePE + area + year + loki0001

(model 4 – example QTL model)

Running the polygenic and QTL models was performed using ASReml software (Gilmour et al. 2002), employing restricted maximum likelihood (REML) statistics to estimate the probability of markers at each chromosomal location being in the proximity of genes associated with variation in the focal trait (Lander & Kruglyak 1995). The polygenic model assumes variation is attributed to many genes of small effect and provides the log-likelihood against which to test the alternate hypothesis of a segregating QTL at the specific location (Beraldi, McRae, Gratten, Slate, et al. 2007). A statistical comparison of polygenic and QTL models is conducted by calculating the ratio of the QTL effect component to the total phenotypic variance in ASReml.

LRT significance thresholds of 7.38 and 14.11, representative of suggestive (expected to occur once in a genome scan by chance) and genome-wide significance (false positive expected once in every 20 genome scans) respectively (Lander & Kruglyak 1995; Nyholt 2000), were used to ascertain the presence of any QTL. The calculation of these thresholds was based on the number of chromosomes and size (Morgans) of the great tit genome using the equation for calculating linkage thresholds determined by Lander & Kruglyak (1995).

Genome-wide association analysis

Repeated measures cannot be routinely accounted for in the genome-wide association analysis software GenABEL (Aulchenko, Ripke, et al. 2007). GWAS phenotype data therefore took the form of a 'phenotype value' generated for each individual from the running of a modified version of the aforementioned animal model where the additive genetic term was excluded leaving six terms including femalePE.

Quality control was performed in GenABEL to eliminate poor quality / uninformative markers and individuals (thresholds: individual call rate: 0.50, SNP call rate: 0.70, Hardy-

Weinberg equilibrium: P < 1e-5 (removal of markers out of Hardy-Weinberg equilibrium at a p-value of less than 0.00001), minor allele frequency: 0.05). A single run of quality control eliminated 279 markers and 1 individual. In total 5313 markers and 950 individuals passed quality control criteria and were used in subsequent analyses. A kinship matrix constructed using marker data was used to estimate breeding values and correct for population stratification, associated with Type I error (Price et al. 2006; Aulchenko, de Koning, et al. 2007; Svishcheva et al. 2012).

Association analyses assume phenotype data comes from unrelated individuals in the same population. However the large quantities of data required for such analyses increase the chance of relatedness in the dataset. Calculating an estimate of lambda, the slope of expected test statistics against an expected uniform distribution, gives an indication of whether there is stratification in the dataset, with inflation of >1 indicating potential population stratification in the data. Population stratification refers to subdivision (genetically distinct groups / individuals) in the dataset, which in association-mapping studies has the potential to cause spurious (or false-positive) associations between phenotypes and markers with no link to causative loci (Pritchard & Rosenberg 1999). Fitting the kinship matrix is expected to reduce the amount of inflation.

Prior to construction of the kinship matrix lambda was estimated at = 1.08 $\pm$ 4.9×10^{-4} (Appendix 2.2a) After correction for population stratification inflation was reduced (λ = 0.98 $\pm$ 6.5×10^{-4}) (Appendix 2.2b). Genome-wide association analysis was carried out using the function 'mmscore' function in R package GenABEL (Aulchenko, Ripke, et al. 2007; Chen & Abecasis 2007). A polygenic model was subsequently estimated to partition breeding values and residuals, and mmscore was used to test for associations between the SNPs and residuals.

Accounting for the false discovery rate inherent in statistical tests (Storey 2002), based on a p-value of 0.05 for 5592 SNPs, the number of false positives is estimated to be 280. The first of two significance thresholds (nominal significance) multiplying the lowest threshold for statistical significance (p = 0.05) by -log10 to give 1.3. The second (genome-wide) significant threshold, after Bonferroni correction for multiple testing (Bland & Altman 1995) and multiplying by -log10 was taken to be 5.05. Significance testing of associations was further determined by a permutation test, of 1000 iterations, whereby quantitative trait values are "shuffled" between individuals over the course of many permutations to generate a reference sample against which to test the probabilities of associations occurring by chance (Fisher 1935; Churchill & Doerge 1994; Ernst 2004).

Spatial auto-correlation

Physical distance between individual organisms is one of the biggest drivers of important biological processes such as extinction, speciation and species interactions (Legendre & Fortin 1989; Legendre 1993; Dormann et al. 2007). This is demonstrated by the phenotypic and genetic variation that can be observed across landscapes as organisms respond to fine-scale environmental conditions (Dormann et al. 2007). The consequence of this is a phenomenon known as spatial autocorrelation whereby values are not independent of other values taken in close proximity (Tobler 1970; Cliff & Ord 1981; Fortin & Dale 2005). Initially recognised as a source of bias in forestry and agricultural studies (Legendre 1993; Gilmour et al. 1997; Burgueño et al. 2000; Legendre et al. 2002; Silva et al. 2001) the non-independence of phenotypes is also an issue for ecological studies of animal populations. Only a handful of studies on animals have attempted to incorporate a direct measure of spatial autocorrelation and found it to be biologically informative (Peakall et al. 2003; Neville et al. 2006; Stopher et al. 2012; Marrot et al. 2015).

One of these studies (Stopher et al. 2012) assessed the effect of incorporating two measures, spatial autocorrelation and home-range overlap on the heritability of various life-history traits in red deer. In line with Stopher et al. (2012) and following the same method, spatial autocorrelation was also modelled in this study. To account for spatial heterogeneity and the potential for phenotypic similarity between individuals / relatives sharing the same environment, spatial autocorrelation was modelled as a first-order separable autoregressive process such that the SAC value between any two phenotypic values is modelled as a power function of their distance apart (Gilmour et al. 1997; Stopher et al. 2012). GPS coordinates of each nest were fitted in the animal model with a covariance structure, denoted by AR1 x AR1 (Gilmour et al. 1997) to account for spatial dependence. SAC was then fitted as an additional random term within the model (3).

Layd = age + age^2 + altitude + year + area + femaleID + femalePE + x.y

(model 5)

2.4 Results

2.4.1. Complex Interactions

A tree model constructed to visualise complex interactions showed females breeding for the first time consistently lay later than their older counterparts (Figure 2.1). Females younger than 2 years old consistently lay later than their older counterparts, irrespective of altitude (Figure 2.1). A significant interaction is observed between the laying dates of females older than 1.5 years and altitude with a difference of approximately three days between females at lower and higher altitudes (Figure 2.1).

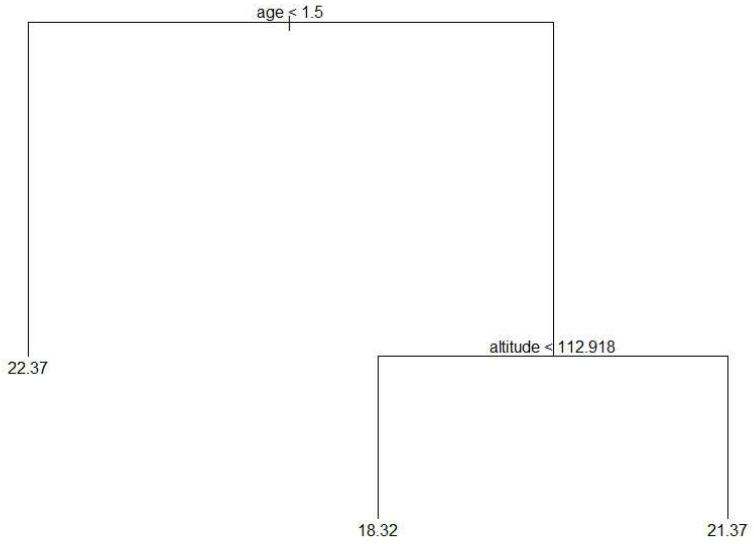


Figure 2.1. A tree diagram illustrating the interaction between female age (years) and altitude (m) and subsequent laying dates associated with each group

2.4.2. Variance components analysis

Heritability (h^2) of laying date, calculated as the ratio of additive genetic variance to total phenotypic variance (V_A / V_P) was estimated from this model ($\pm$SE) to be 0.05 $\pm$ 0.03. The permanent environment effect (pe^2) was estimated as 0.04 $\pm$ 0.03.

When tested for significance spatial autocorrelation (SAC) was found not to significantly contribute to variation in laying date by LRT or z-ratio tests. Examination of the z-ratios produced by a model incorporating SAC and excluding area, versus a model excluding SAC and incorporating area showed that spatially dependent phenotypic variation is captured by area. When area is excluded the spatial component of variation is redistributed to SAC. SAC was excluded from the model prior to QTL and GWAS analyses.

2.4.3. Genome scan and Genome-wide association analysis

No QTL peaks passed the genome-wide significant threshold (Figure 2.2a). One peak, a region of chromosome 10 represented by five SNPs passed the suggestive threshold, no more than would be expected by chance. Indications that no specific chromosomal regions / underlying genes significantly explained the majority of additive genetic variation in lay date was further substantiated by GWAS analyses. No SNPs breached the genome-wide significant threshold, with 291 SNPs surpassing the nominally significant threshold, which was about the number that would be expected by chance (Figure 2.2b). The absence of true associations was confirmed by permuting the data, as none of the the markers achieved statistical significance, being no higher than would be expected to occur by chance (Table 2.3).

The occurrence of no more positive associations than would be expected by chance was confirmed by means of a binomial test ($p > 0.1$). A comparison also reveals little accordance between the methods for regions or markers of nominal significance, with no obvious

similarities in the results. The highest QTL peaks on chromosomes 10, 11 and 1 did not correspond to the highest associations detected by GWAS which were located on chromosomes 8, 27 and 3, with the seventh highest association on chromosome 11.

Table 2.3. The top 10 highest associations from GWAS analysis following a permutation test of 1000 iterations to test for significance. Number of females for which complete phenotypic information was available (N), chromosome (Chr) and position (cM). Effect sizes (effB) and their standard errors (se_effB) of allelic substitutions in an additive models with 1df. Nominal p value (pre-permutation Pc1df) and genome-wide p value from permutations tests before and after lambda correction (P1df, Pc1df). Genotype effect sizes relative to AA and test statistics from a model where genotypes are fitted as a three level factor are described by effAB and effBB.

Marker	Chr	Position	A1	A2	N	effB	se_effB	chi2.1df	Pre-permutation Pc1df	P1df	Pc1df	effAB	effBB
M5243	8	14	A	G	951	0.252718	0.073132	11.94159	0.000339	0.943	0.778	0.346443	0.465262
M4088	27	29	A	G	951	0.274015	0.086193	10.10648	0.000852	1	0.991	0.327914	0.420842
M5547	3	54	G	A	949	0.276957	0.089532	9.569087	0.001266	1	0.998	0.324016	0.367933
M1348	1A	42	C	A	951	-0.26897	0.088528	9.231042	0.001351	1	1	-0.25948	-0.56662
M8791	Z	24	A	G	821	0.219958	0.073385	8.98388	0.001639	1	1	NA	0.439915
M7140	11	9	G	A	948	0.283902	0.094877	8.954039	0.001761	1	1	0.317232	0.434368
M9091	Z	25	G	A	757	0.191027	0.065296	8.558987	0.001522	1	1	NA	0.382055
M5073	6	77	G	A	951	-0.20988	0.072125	8.467465	0.002309	1	1	-0.27904	-0.36514
M9075	Z	40	G	A	733	-0.1802	0.062786	8.237147	0.002175	1	1	NA	-0.36039
M1745	2	78	G	A	951	-0.20863	0.072741	8.22635	0.002717	1	1	-0.21896	-0.41273

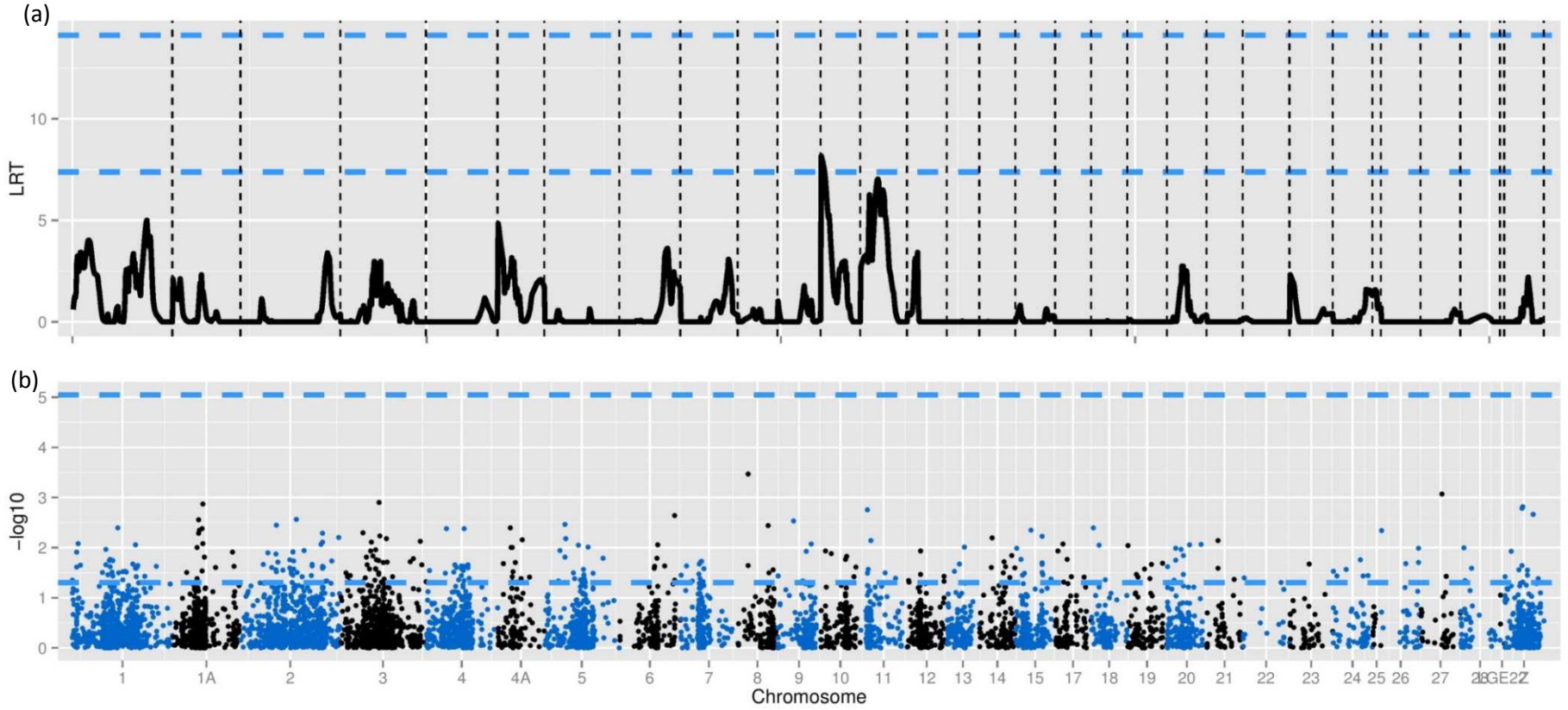


Figure 2.2. A QTL (a) and GWAS (b) scan for variants contributing to variation in laying date. Dashed blue lines represent (a) suggestive (7.38) and genome wide significant (14.11) linkage, and (b) nominal (1.3) and genome-wide significant (5.05) associations.

2.5 Discussion

Timing of breeding is a life history trait for which the presence of additive genetic variation has been identified from several analyses of laying date in the Wytham great tit population (van der Jeugd & McCleery 2002; McCleery et al. 2004; Garant et al. 2008; Liedvogel et al. 2012). By conducting statistical testing to search for genetic variants contributing to variation in breeding phenology, the research presented here compliments studies which have identified a genetic component of variation to laying date. In accordance with the findings of van der Jeugd & McCleery (2002), McCleery et al. (2004), Garant et al. (2008) and Liedvogel et al. (2012) we confirm the presence of a significant genetic component of variation. The heritability estimate obtained here (h^2 0.05) is markedly lower than the h^2 of 0.16 presented by the earliest studies (van der Jeugd & McCleery 2002; McCleery et al. 2004) but in concordance with the estimate of 0.03 reported by Liedvogel et al. (2012).

In comparison to higher estimates obtained from parent offspring regression (van der Jeugd & McCleery 2002; McCleery et al. 2004) the application of the animal model and, in particular, accounting for factors influencing phenotypic variation in the form of fixed and random effects, ultimately resulted in a lower and more precise heritability estimate in both cases (Liedvogel et al. 2012, current study). Mirroring Liedvogel et al. (2012), the removal of fixed and random effects restored our heritability estimate to within a similar range of all previous estimates (h^2 = 0.21). Estimates of heritability may vary for several reasons, such as the size and subset of the data used in the analyses, changes to genetic variance or variation in phenology in response to environmental perturbation in subsequent years (Liedvogel et al. 2012).

Potentially, selection favouring earlier laying that operated in the years succeeding the earliest analyses reduced the genetic variance associated with the trait. Charmantier et al. (2008) argue however that the observed behavioural shift bears more resemblance to a

plastic as opposed to a micro-evolutionary response, with the rate of response too fast to be accounted for by evolutionary adaptation. The similarity between the two estimates obtained from parent-offspring regression, and similar concordance between the two estimates obtained by means of an animal model, may also indicate that variation in estimates is an artefact of the statistical methods employed. Because they utilise a mean trait value and fail to account for additional sources of variation, estimates obtained by parent offspring regression are generally prone to inflate estimates of heritability (Kruuk 2004; Akesson et al. 2008).

Building on previous research, quantitative trait loci analyses were employed to identify the genes / genomic regions underlying the observed genetic component of variation. Behaviours are complex quantitative traits and as such are assumed to be largely polygenic with genetic components comprised of many small genes each of small effect. Together with the fact that as a result of the low heritability, the variance contributed by each of the multiple genetic determinants will be a fraction of the already small genetic component, it is perhaps not unsurprising that QTL scans failed to identify any SNPs of genome-wide significance.

QTL analyses have been used to identify genomic regions associated with behavioural variation in domestic / laboratory animals such as exploratory and anxiety-related behaviours in mice and rats (Gershenfeld et al. 1997; Flint 2003; Henderson et al. 2004), temperament in cattle (Gutiérrez-Gil et al. 2008; Glenske et al. 2010; Glenske et al. 2011; Haskell et al. 2014) and tame / aggressive behaviours in silver foxes (Kukekova et al. 2010; Johnson et al. 2013) but due to the inherent difficulties of studying wild animals and the dearth of long-term studies with sufficient data to enable their application fewer studies exist for wild animals.

Studies to identify QTL for behavioural traits in wild animals have typically used a series of backcross experiments, involving the controlled breeding of captured wild individuals to identify QTL associated with foraging (honeybees (Rüppell et al. 2004)), migratory (rainbow trout (Hecht et al. 2015)) and schooling (sticklebacks (Greenwood et al. 2015)) behaviours. Quantitative genetic analyses of personality on bighorn sheep represent rare examples of research conducted on a truly wild and free ranging organism (Réale et al. 2009; Poissant et al. 2013). In birds, QTL analyses for behavioural variation have until now only been conducted on domesticated or economically important species (Recoquillay et al. 2015; Schütz et al. 2002).

Relying solely on multigenerational phenotype and pedigree data for a truly wild animal with minimal interference and no (anthropogenic) induced selection for variation in laying date could be considered a major limitation in terms of reduced power to detect genetic variants. However the best way to understand processes operating in the wild is to study organisms in their natural environment and in this sense this study, together with studies such as those on the bighorn sheep (Réale et al. 2009; Poissant et al. 2013) have a distinct advantage.

The application of GWAS to identify genomic regions associated with behavioural variation is typically associated with large-scale human studies of personality and behavioural disorders (Terracciano et al. 2010; Calboli et al. 2010; Salvatore et al. 2015; Sokolowska & Hovatta 2013). Due to the necessity of an extensive database of genotyped individuals for viable and accurate GWAS, domestic and model animal species are the strongest candidates for such analyses, after humans, and GWAS has only recently been applied to such systems (Do et al. 2013).

GWAS analyses for morphological and behavioural traits in wild animals are proliferating, as genomic tools become increasingly accessible and affordable (Johnston et al. 2011; Bush

& Moore 2012; Hecht et al. 2013; Johnston et al. 2014; Bérénos et al. 2015). Genomic dissection of life-history traits such as clutch size and egg mass in Wytham great tits using genome-wide methods suggest the suitability of this system for genomic analyses of a wild vertebrate (Robinson et al. 2013; Santure et al. 2013). This study represents one of the first to conduct a GWAS for a behavioural trait.

In accordance with the QTL analysis, no SNPs reached genome-wide significance in the genome scan. No more SNPs passed the p value < 0.05 level of significance than would be expected by chance and the regions of highest association in the GWAS, on chromosomes 8, 27 and 3 do not correspond to the regions with the highest association as observed from QTL analyses, on chromosomes 10,11 and 1. This would tend to suggest a polygenic basis for heritability of lay date, with many genes contributing to variation each constituting such a small fraction of the total heritability, for a trait with a small genetic component, as to make their detection by genomic analyses a difficult prospect.

Our results further highlight the complexity of breeding behaviour being determined by multiple interacting factors. Even with a significant genetic component, variation in lay date is predominantly a consequence of external abiotic and biotic factors of which temperature is thought to be one of the most influential. This is supported by the fact the term 'year', which will include temperature variation between years, explaining more of the variance than any of the other fixed or random model terms including the additive genetic component. The effect of temperature on breeding behaviour has been documented in dozens of birds with significant advances in laying date occurring in numerous species and with predictions that early laying will be observed in many others as temperatures continue to increase (Crick & Sparks 1999; Dunn 2004).

Fine-scale variation in environmental conditions also has a significant effect on breeding phenology as demonstrated by the significance of altitude with birds consistently breeding

earlier at lower altitudes (Wilkin et al. 2006; Liedvogel et al. 2012; Garroway et al. 2013) (Figure 2.2). With poorer natal quality at higher altitudes (Wilkin & Sheldon 2009) the complex interaction with age (Figure 2.2) may also be the result of the acquisition of the best territories by experienced females leaving inexperienced females only those nest boxes at higher altitudes.

Research on lay date in several great tit populations and other species has determined phenotypic plasticity to be the most significant, if not sole, factor influencing variation in breeding time (Przybylo et al. 2000; Brommer et al. 2008; Matthysen et al. 2011; Charmantier & Gienapp 2014). Using the same method of regressing the mean within-individual variation against temperature, and comparing the results to the population level regression, two independent studies of Wytham and Belgian great tits found that plasticity explains most of the observed advance in lay date (Charmantier et al. 2008; Matthysen et al. 2011).

Further studies highlighting the sensitivity of birds to the immediate environment, with laying date largely determined by localised variation in vegetation and prey phenology, give strong supporting evidence for the importance of plasticity (Vedder et al. 2013; Hinks et al. 2015). The ability to pick up on cues and respond plastically is likely to be a species greatest defence against changing environmental conditions as the response is immediate, compared to micro-evolutionary change which requires the passing of several generations before the desired adaptation can be readily expressed. In Wytham great tits the absence of a lag between the shift in caterpillar biomass and laying date suggests an immediate plastic response (Charmantier et al. 2008).

With strong evidence that external factors play the largest role in determining breeding time in great tits, our ability to detect the genetic mechanisms contributing to variation, even where a significant genetic component exists, are impeded. For traits with low to

moderate heritability, detecting genetic variants is exceptionally difficult, especially when there is incomplete linkage disequilibrium between causal variants and typed SNPs (Wang et al. 2005; Mackay et al. 2009; Hayes et al. 2010; Yang et al. 2010). Even for traits with moderate to high heritabilities detecting and localising QTL is a difficult prospect (Mackay et al. 2009). Traits of low heritability or regions of low LD require unrealistically high, if not total, coverage of the genome to be genotyped for GWAS to be able to identify regions of interest (Risch 2000; Wang et al. 2005; Johnston et al. 2014) and variants of large effect are able to be detected more reliably in any case than those of small effect (Hirschorn & Daly 2005; Hayes et al. 2010).

Coupled with the importance of a detailed genetic map the issue of sample size presents further challenges for those seeking to describe the genetic architecture of a trait. Research into the efficacy of genomic analyses have determined extensive phenotypic datasets are required in order for genome scans to be able to detect genes or genomic regions associated with variation particularly for traits of low heritability (Beavis 1994; Risch & Merikangas 1996; Risch 2000). Sample size has been proposed to be even more influential in determining the success and accuracy of genomic analyses than dense SNP coverage (Hirschhorn & Daly 2005; Wang et al. 2005).

Despite using one of the largest and most comprehensive datasets for any wild animal the Wytham great tit dataset is not on the scale of the human or domestic organism datasets with which QTL and GWAS analyses are more commonly associated, in which phenotyped individuals typically number into the thousands or even hundreds of thousands (Hayes et al. 2010; Lango Allen et al. 2010; Yang, T. Lee, et al. 2013; Wood et al. 2014). The lack of long-term studies and extensive datasets available, particularly for wild populations, continues to be the main limiting factor in our ability to describe the genetic architecture of phenotypic variation (Hill 2012); something which can only be addressed by the continued

efforts of researchers working on model wild systems such as the great tit to generate the quantity of data that is essential for genomic analyses.

Chapter 3

A comparative and combined genome mapping analysis of variation in laying date in two populations of *Parus major*

Jenny C. Armstrong[1], Anna Santure[1,2], Isabelle de Cauwer[1,3], Kees van Oers[4], Nikkie van Bers[4, 5], Martien Groenen[5], Ben Sheldon[6], Marcel Visser[4], Jon Slate[1]

[1]Department of Animal & Plant Sciences, Alfred Denny Building, University of Sheffield, Western Bank, Sheffield S10 2TN, UK; [2]School of Biological Sciences, University of Auckland, Auckland, New Zealand; [3]Laboratoire de Génétique et Evolution des Populations Végétales, UMR CNRS 8198, Bâtiment SN2, Université des Sciences et Technologies de Lille – Lille 1, Villeneuve d'Ascq Cedex, France; [4]Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW),POBox 50, 6700ABWageningen, the Netherlands; [5]Animal Breeding and Genomics Centre, Wageningen University, De Elst 1, Wageningen, 6708 WD, the Netherlands; [6]Edward Grey Institute, Department of

3.1 Abstract

The greatest challenge facing investigations into the evolution and genetic architecture of phenotypic traits in wild animals is the notable lack of datasets with the scope and depth necessary to enable the application of genomic tools. The great tit *Parus major* is exemplary in being one of very few wild animal species where several geographically distinct populations have been monitored with enough intensity to enable direct comparisons of loci contributing to phenotypic variation. Following on from analyses of one population at Wytham Woods (Oxford, UK) (Chapter 2), here I examine the genetic architecture of lay date using data from a second long-term study of great tits in the Netherlands. The availability of phenotypic and genomic data from two populations of the same species meant that combining the datasets increased overall sample size and subsequently the power of gene mapping analyses. Analysis of the Netherlands population independently also enabled a direct comparison of the genetic architectures of laying date between the two populations. Applying QTL mapping and GWAS methods we found no evidence for distinct genomic regions contributing to phenotypic variation in laying date, in either the Netherlands population in isolation, or using the combined dataset from both populations. There was no correspondence in the genomic regions of strongest statistical association between populations, implying either the lack of a shared genetic architecture, or a lack of power with which to detect one.

3.2 Introduction

Differences in environmental conditions and selection pressures operating between locations can cause variation in life history traits between populations of the same species (Visser et al. 2003). The sensitivity of organisms to their environment, and subsequently the capacity for different populations of the same species to exhibit phenotypic variation, has consequences for our understanding of the genetics of trait evolution. Geographically distinct populations may not only differ outwardly in their phenotypic responses but, as the result of micro-evolutionary processes, different populations may show signatures of genetic differentiation (Wright 1943; El Mousadik & Petit 1996; Estoup et al. 1996; Merilä & Crnokrak 2001; Turchetto et al. 2014; Funfstuck et al. 2014).

Similarities in the genetic architecture phenotypic variation between populations of the same species is demonstrable from research on humans and model organisms (Hirschorn & Daly 2005; Flint & Mackay 2009) and increasingly from genomic analyses of wild and non-model organisms as extensive phenotypic and genetic data becomes available for a much wider variety of species (Colosimo et al. 2004; Neale & Savolainen 2004; Bonin et al. 2006; Protas et al. 2006; Wright et al. 2006; Franck et al. 2007; Stinchcombe & Hoekstra 2008; Houde et al. 2013; Christensen et al. 2014). However identifying specific genetic variants associated with phenotypic variation between populations remains a significant challenge (Ellegren & Sheldon 2008). Whilst a number of studies have managed to identify genetic variants in multiple populations of a wild organism (Colosimo et al. 2004; Greenwood et al. 2015) few studies have been able to quantify in as much detail the genetic structure of phenotypic variation in multiple wild populations of a species.

The great tit *Parus major* is one of the few wild organisms for which multiple long-term studies enable direct comparisons of phenotypic variation between geographically distinct populations of a species, and for which genetic resources are available to help identify the

determinants of variation between populations. The great tit's suitability as a model organism has facilitated some of the most advanced studies into the genetics of not only morphology but behaviour and personality, being one of the few wild animals for which a genetic basis for behavioural variation has been described from both captive and wild populations (Dingemanse et al. 2002; Drent et al. 2003; Fidler et al. 2007; Riyahi et al. 2015).

Phenotypic variation between great tit populations can be observed in varying patterns of laying date across the species range with some populations responding to increasing temperatures by laying earlier, with others showing no advancement of laying date at all, even where there is selection for early laying (Visser et al. 2003; Matthysen et al. 2011). In Finland laying date has not significantly advanced despite selection for early laying during breeding seasons with high breeding time temperatures and densities (Ahola et al. 2009). In contrast studies of a Belgian population found a significant advance in laying date of 12 days over 30 years (Matthysen et al. 2011).

Genetic analyses of populations of great tits in the Netherlands have consistently revealed a heritable component to laying date (Noordwijk et al. 1981; Nussey et al. 2005; Gienapp et al. 2006). Heritability estimates of around 0.17 have been consistently calculated for the Hoge Veluwe population from multiple studies over the past three decades (van Noordwijk et al. 1981; Gienapp et al. 2006). Despite evidence for a significant genetic component the pattern of lay date in Hoge Veluwe is strikingly different from Wytham with no significant advance in laying date despite selection for earlier laying (Visser et al. 1998; Gienapp et al. 2006).

In contrast the laying date of great tits in Wytham Woods has advanced by an average 14 days over a 47 year period (Charmantier et al. 2008). A significant advance in the date of peak caterpillar biomass of approximately 14 days from 1985 to 2004 (Visser et al. 2006) is

not reflected in the breeding phenology of the Hoge Veluwe great tit population which has seen a (non-significant) advancement of lay date of 5.4 days over roughly the same period (Gienapp et al. 2006).

The mismatch has been attributed to a number of factors including a reduction in the incidence of double broods (Visser et al. 2003) and the possibility of physiological constrains preventing earlier breeding (Stevenson & Bryant 2000). However the effect of temperature on predator / prey phenology provides the most convincing evidence as to why laying date and subsequently hatch date is lagging behind peak caterpillar abundance in Hoge Veluwe.

Several studies have demonstrated the importance of spring temperature as a cue for the onset of laying in birds (Visser et al. 2009; Conway & Martin 2000; Matthysen et al. 2011; Dolenec & Dolenec 2011; Both et al. 2004) and in determining the rate of development of caterpillar pupae (Smith et al. 2011; van Asch et al. 2007; Buse et al. 1999). The mismatch between peak food requirement and peak food abundance in Hoge Veluwe appears to be a consequence of birds and caterpillars taking their temperature cues for these crucial life-history events from different time periods.

Birds in Hoge Veluwe take cues on when to lay their eggs from temperatures during the period 1st March to 15th April (van Balen 1973; Visser et al. 2006). Between 1973 and 1995 temperatures during this early spring period showed no significant increase (Visser et al. 1998). In contrast, caterpillar development and hatching reacts to temperatures in the 30 days following April 15th, and average temperatures in this later spring period between 1973 and 1995 have shown a significant increase (Visser et al. 1998). This has resulted in a greater shift in caterpillar phenology than in the breeding phenology of their great tit predators (Visser et al. 1998; Visser et al. 2006).

In this chapter I use linkage (QTL) and association (GWAS) analyses to scan for genetic variants associated with laying date on two great tit datasets. I first conduct analyses on data from Hoge Veluwe to generate an estimate of heritability. Then I apply genomics tools to identify the genetic variants associated with variation in laying date in this population in isolation, the first study to attempt to do so.

A rare opportunity is taken to perform a comparison of the underlying genetic architectures of an important life-history trait in geographically distinct populations of the same species, as the results obtained for Hoge Veluwe are subsequently compared to those already obtained for Wytham Woods. Finally by combining the phenotypic and genetic datasets of Hoge Veluwe and Wytham, a dataset of much greater size and depth than that for either population in isolation is generated, lending greater power to subsequent statistical testing. I use this 'master' dataset to attempt to identify the determinants of variation in laying date in the great tit.

3.3 Methods

3.3.1. Phenotype data

Laying date phenotype data takes the form of number of days after April 1st. Breeding events that occurred more than 30 days from the first recorded breeding event of the season are likely to be second breeding attempts (van Noordwijk et al. 1995), and were therefore removed to limit the analysis to first breeding attempts.

Hoge Veluwe (NL)

Phenotype information for Hoge Veluwe (hereafter NL) birds comes from a long-term study of great tits at Hoge Veluwe (Balen 1973). Data collection methods mirror those practised in Wytham (Chapter 1), with metal rings affixed to the legs of great tits so that subsequent breeding events / maternities / paternities can be attributed to specific individuals and a blood sample obtained as the birds are ringed for genotyping purposes.

The subset of phenotype data available for the purposes of this study represents data gathered over the period 1995 to 2012. For the QTL analysis, taking into consideration only those females present in the pedigree, phenotype data takes the form of 774 breeding event records representing 409 female birds. The requirements of GWAS analyses do not include use of a pedigree. Therefore genotyped females absent from the pedigree, but for which phenotype information was available could be included in the analysis. The phenotypic dataset for NL GWAS analyses consisted of 1615 breeding event records for 947 females.

Merged Hoge Veluwe / Wytham Woods (NLUK) dataset

Merging the QTL phenotypic datasets for both populations resulted in 1916 recorded breeding events for 1080 females. For GWAS analysis, the addition of genotyped and

phenotyped females that were excluded from QTL analyses resulted in a combined NLUK phenotype file containing records of 3137 breeding events from 1898 females.

3.3.2. Pedigree data (A matrix)

For the purposes of the NL specific and the combined NLUK analyses an A matrix constructed from SNP markers (a realised relationship matrix) was used in lieu of a pedigree. Of 1407 genotyped NL birds, 666 related individuals were extracted to form the A matrix. Four hundred additional genotyped individuals were included for the purposes of incorporating informative links between the 666 individuals, with the resultant relationship matrix comprising of 1066 individuals. The genetic relatedness of each pairwise combination of birds is detailed in the subsequent A matrix. For GWAS analyses the number of individuals could be extended to include genotyped birds that were previously excluded from QTL analyses. Genetic information for 947 genotyped female birds was utilised in GWAS analyses.

To construct an A matrix for the combined NLUK analyses, NL birds retained their original numeric IDs where WY numeric IDs were modified by the addition of 1066 to their numeric IDs. The files were then combined to create a matrix detailing the genetic relatedness of each pairwise combination of 3775 individual birds. Using the full complement of genotyped female individuals for both populations genetic information for 1898 birds was used in GWAS analyses, with NL retaining their original numeric IDs, and WY bird IDs modified by the addition of 947 to their own numeric IDs.

3.3.3. Marker data

With genetic information available for thousands of individuals from both populations in the form of blood samples obtained whilst ringing, an extensive database of genetic information has been made available through the detection of SNPs by means of next

generation sequencing (van Bers et al. 2010), mapping of SNPs to the Zebra finch chromosome and subsequent development of great tit linkage maps (van Oers et al. 2014) and a full genome sequence assembly (Laine et al., manuscript in revision).

A total 4702 birds from populations of great tits in the Netherlands and the UK were genotyped using a 9193 SNP chip (described in detail in van Bers et al. (2012)). The NL subset of birds was drawn from 3 populations, 510 from a captive and experimental line (descendants of wild caught NL birds), 553 from de Hoge Veluwe and 937 from Westerheide. Of the UK birds 2652 individual samples came from Wytham, the remaining 50 from nearby Bagley woods. The two wild populations in the Netherlands as well as the two UK populations are geographically close and so were treated as if each were from one continuous population within each country (K. van Oers, personal communication from Van Bers et al. 2012; Verhulst et al. 1997).

Markers were mapped to 32 chromosomes for the great tit (1-15, 17-24, 26-28, 1A, 4A, 25A, 25B, LGE22 and Z). Two different maps, framework and parsimonious, were used for each analysis respectively (Chapter 1).

Building on the animal model principle, existing identity-by-descent (IBD) matrices for each of the 31 chromosomes, at multiple positions (loci) along the chromosome at 5cM intervals were obtained for use in QTL analyses. IBD matrices detail the genetic relatedness of every possible pair combination at the relevant chromosomal location and were constructed at 5cM intervals using available marker information for both populations (van Oers et al. 2014) together with respective population pedigree information and using the program LOKI ver2.4.5 (Heath 1997) (as described in (Santure et al. 2013)). The number of IBD matrices per chromosome reflects chromosome length (ranging from 4.58cM – 138.14cM (NL) and 3.68-139.88cM (UK)), with the largest chromosomes sporting the greatest number of IBD matrices.

For the combined QTL analysis, IBD matrices from NL and WY were merged together, NL birds retaining their original numeric IDs and the numeric IDs of WY birds modified by the addition of 1066 (the number of NL birds in the NL pedigree) to each. For GWAS marker data for the 5582 SNPs scored in both populations were used, with genotype information for WY individuals for the 10 SNPs not scored in the NL individuals removed from the genotype files.

3.3.4. Statistical analyses

Constructing the animal model

In order to accurately estimate the additive genetic component of variation it is important to distinguish causal factors influencing trait variation to minimise the possibility that the relationships we observe are driven by environmental covariances (Gienapp et al. 2006). Following the same protocol as Chapter 2, a univariate animal model in the form of a linear mixed effects model with fixed and random effects was constructed in order to estimate the additive genetic component of variation in laying date. Additional sources of variation in to laying date were identified by linear modelling in R ver3.0 (R Core Team, 2013) (fixed effects) and linear mixed modelling in ASReml ver3.0 (Gilmour et al. 1996). Accounting for variables found to significantly influence variation in lay date enables the true size of the additive genetic component to be more accurately estimated.

Less environmental and spatial information was available to accompany breeding event records for the NL birds than those from the UK population. Only those explanatory variables present for both populations were extracted from the UK dataset to contribute to the NLUK analysis. The subsequent model selection process was the same for both NL and NLUK analyses. Age and quadratic age (taking the form of a curvilinear relationship, as biological relationships with age may not be completely linear) were significant factors

contributing to variation in laying date ($p < 0.01$ both cases), determined by means of linear modelling in R statistical software (Chapter 2) and were included as fixed effects in both animal models (Appendix 3.1). 'Population' was also included as a fixed effect in the NLUK animal model.

Random effects (year of data collection, permanent environment effects i.e. from repeated measures and additive genetic effect) were determined by hierarchical significance testing of models of increasing complexity using restricted maximum likelihood (REML) testing (Wilson et al. 2009) in ASReml ver3.0. Likelihoods were obtained for models excluding and including each term. A comparison of the two models was then performed by likelihood ratio testing to test the significance of each term (Chapter 2) (Table 3.1).

All terms were significant in contributing to variation in laying date for both NL and NLUK datasets ($p < 0.01$). The final model, identical for both NL and NLUK analyses, is the 'polygenic' model. This type of mixed effects model partitions the variance associated with lay date into each of the genetic and environmental factors. This enables the estimation of the additive genetic component of variation (V_A), permanent environment variance (V_{PE}) and residual variance (V_R) from which total phenotypic variation (V_P), heritability (h^2), permanent environment effect (pe^2) and residual variance (r^2) can be calculated.

Table 3.1. Statistical testing of random effect terms included in the final versions of the polygenic animal models for the Netherlands (a) and combined (b) datasets, using logarithm of the odds (LOD) scores obtained from ASReml, and subsequent likelihood ratio test (LRT) of significance. Variance components statistics for the final polygenic models are highlighted.

(a) NL					
Term	**LRT**	**p value**	**Component**	**Comp/SE**	**D.O.F**
Year	417.1	<0.01	20.22	2.60	1
FemaleID	9.96	<0.01	7.22	2.86	1
FemalePE	76.86	<0.01	2.37	1.01	1
Residual variance			12.56	13.58	
Total variance			42.37		771
(b) NLUK					
Term	**LRT**	**p value**	**Component**	**Comp/SE**	**D.O.F**
Year	858.68	<0.01	34.72	2.99	1
FemaleID	10.48	<0.01	4.85	2.97	1
FemalePE	126.46	<0.01	4.00	2.41	1
Residual variance			16.41	21.05	
Total variance			59.93		1912

Scanning for QTL

For every position at 5cM intervals on each chromosome a modified version of the original (NL / NLUK) animal model, known as the QTL model (Chapter 2) was written with the relevant chromosomal location IBD matrix fitted as an additional random effect. Animal models were run using ASReml ver3.0 software which employs maximum likelihood to partition trait variance into discrete components. I was able to estimate the probability of there being a QTL at each genomic location by comparing the likelihood ratio test statistic of the polygenic model (assuming many genes of small effect) with each of the QTL models. If, after accounting for additional sources of variance (fixed and random effects), there is a difference between the polygenic and QTL models, such that the IBD matrix at a particular genomic location explains significant variation, there is evidence that the location contains genes that contribute to variation in the focal trait (Beraldi, McRae, Gratten, Slate, et al. 2007).

Likelihood ratio test significance thresholds of 7.38 and 14.11, representative of suggestive and genome-wide significance, were calculated according to the method of Lander & Kruglyak (1995), which is based on the size and number of chromosomes of the relevant organism's genome. Based on the likelihood of encountering false positive results, one false positive would be expected to exceed the suggestive threshold and genome-wide significant thresholds once every genome scan and once in every twenty genome scans respectively (Nyholt 2000).

Genome-wide association analysis

Quality control was performed in GenABEL, eliminating individuals and SNPs that fail to meet certain thresholds (individual call rate: 0.50, SNP call rate: 0.70, Hardy-Weinberg equilibrium: $P < 1e\text{-}5$, minor allele frequency: 0.05). To reduce the potential for Type I error

a relationship matrix based on genome-wide marker relatedness (Price et al. 2006) was used to correct for population stratification (Chapter 2).

For the NL dataset quality control eliminated 347 markers. No individuals were identified as being of inadequate quality, with 5234 markers and 947 individuals kept for subsequent analyses. For the combined NLUK dataset quality control eliminated 324 markers and 1 individual with 5258 markers and 1897 individuals retained for association analyses. Neither dataset showed much evidence of inflation indicative of population stratification with estimates of lambda <1 following steps to account for relatedness between individuals (Appendices 3.2 & 3.3).

Genome-wide association analyses was performed in the R package GenABEL (Aulchenko, Ripke, et al. 2007) using the function mmscore (Chen & Abecasis 2007) to account for population stratification and conduct statistical tests for association between each of 5582 SNP markers and phenotypic variation. As repeated phenotypic measures for individuals cannot be accounted for in this type of analysis, phenotype for each individual female took the form of a residual value estimated from a version of the aforementioned animal model which excludes the additive genetic term (Chapter 2).

Two thresholds were used to discriminate between significant and non-significant SNPs, the first representing a p value < 0.05 level of significance, the second a genome-wide level of significance taking into account Bonferroni's correction for multiple testing (Holm 1979). Given the thousands of different association tests being performed, a false discovery rate (FDR) was calculated to guard against type I errors using an equation derived from the method proposed by Storey (2002) using our defined probability threshold of $p < 0.05$ and the number of SNPs (Bolormaa et al. 2011; Storey & Tibshirani 2003). For 5582 SNPs this works out at approximately 280 SNPs, or 280 false positives, expected to cross the first of the significance thresholds in any one genome scan. As in Chapter 1, significance testing

took the form of a permutation test of 1000 iterations (Churchill & Doerge 1994; Ernst 2004).

3.4 Results

3.4.1. Variance components analysis (NL)

In the NL population, complex interactions modelled by the 'MASS' package (Venables & Ripley 2002) in R indicated female great tits younger than two and a half years old lay approximately 2 days later than their older counterparts with averages of 19 and 16.62 days from April 1st respectively. Heritability (h^2) of lay date, calculated as the ratio of additive genetic variance to total phenotypic variance (V_A / V_P) was estimated for the NL population ($\pm$SE) to be 0.15 $\pm$ 0.06 with permanent environment effect (pe^2) estimated as 0.08 $\pm$ 0.06. Combining the two populations and estimating variance components resulted in an h^2 estimate of 0.08 $\pm$ 0.03 and a pe^2 of 0.07 $\pm$ 0.03.

3.4.2. QTL scan and GWAS results (NL)

Results obtained for the NL population from both the QTL scan and GWAS indicate no genomic regions associated with variation in laying date. QTL peaks failed to pass either the suggestive or genome-wide significant thresholds (Figure 3.1a). Similarly results from the GWAS saw no SNPs pass the genome-wide significant threshold, with 265 passing the suggestive threshold, no more than would be expected by chance (Figure 3.1b). Permuting the data further confirmed the absence of significant results (Table 3.2).

Table 3.2. The top 10 associations from GWAS analyses of laying date in the Hoge Veluwe (NL) great tit population, following a permutation test of 1000 iterations to test for significance. Number of females for which complete phenotypic information was available (N), chromosome (Chr) and position (cM). Effect sizes (effB) and their standard errors (se_effB) of allelic substitutions with 1df. Nominal p value (pre-permutation Pc1df) and genome-wide p value from permutations tests before and after lambda correction (P1df, Pc1df). Genotype effect sizes relative to AA and test statistics from a model fitting genotypes as a three level factor described by effAB and effBB.

Marker	Chr	Position	A1	A2	N	effB	se_effB	chi2.1df	Pre-permutation Pc1df	P1df	Pc1df	effAB	effBB
M8363	2	84	A	G	947	0.288525	0.084818	11.57159	0.000380	0.955	0.845	0.337309	0.487702
M7462	21	14	G	A	947	0.253897	0.077463	10.74306	0.000678	0.998	0.951	0.307336	0.462446
M8748	28	48	A	C	947	0.240431	0.073439	10.71822	0.000671	0.998	0.953	0.268375	0.478541
M5035	5	10	A	G	947	0.23959	0.075065	10.18729	0.000918	1	0.988	0.307858	0.463719
M7628	3	84	G	A	943	0.289165	0.090795	10.14291	0.000966	1	0.991	0.399883	0.279369
M1122	12	5	A	C	946	0.234904	0.075156	9.769034	0.001302	1	0.998	0.106295	0.530105
M8706	5	45	A	G	947	0.267627	0.08624	9.630348	0.001339	1	0.999	0.226595	0.611727
M1495	14	45	A	C	947	0.327104	0.107963	9.179533	0.001630	1	1	0.166838	1.648011
M7100	1	59	G	A	947	0.24891	0.082379	9.129603	0.001832	1	1	0.25861	0.484432
M7391	14	5	G	A	947	-0.25121	0.083627	9.023682	0.001976	1	1	-0.3463	-0.35722

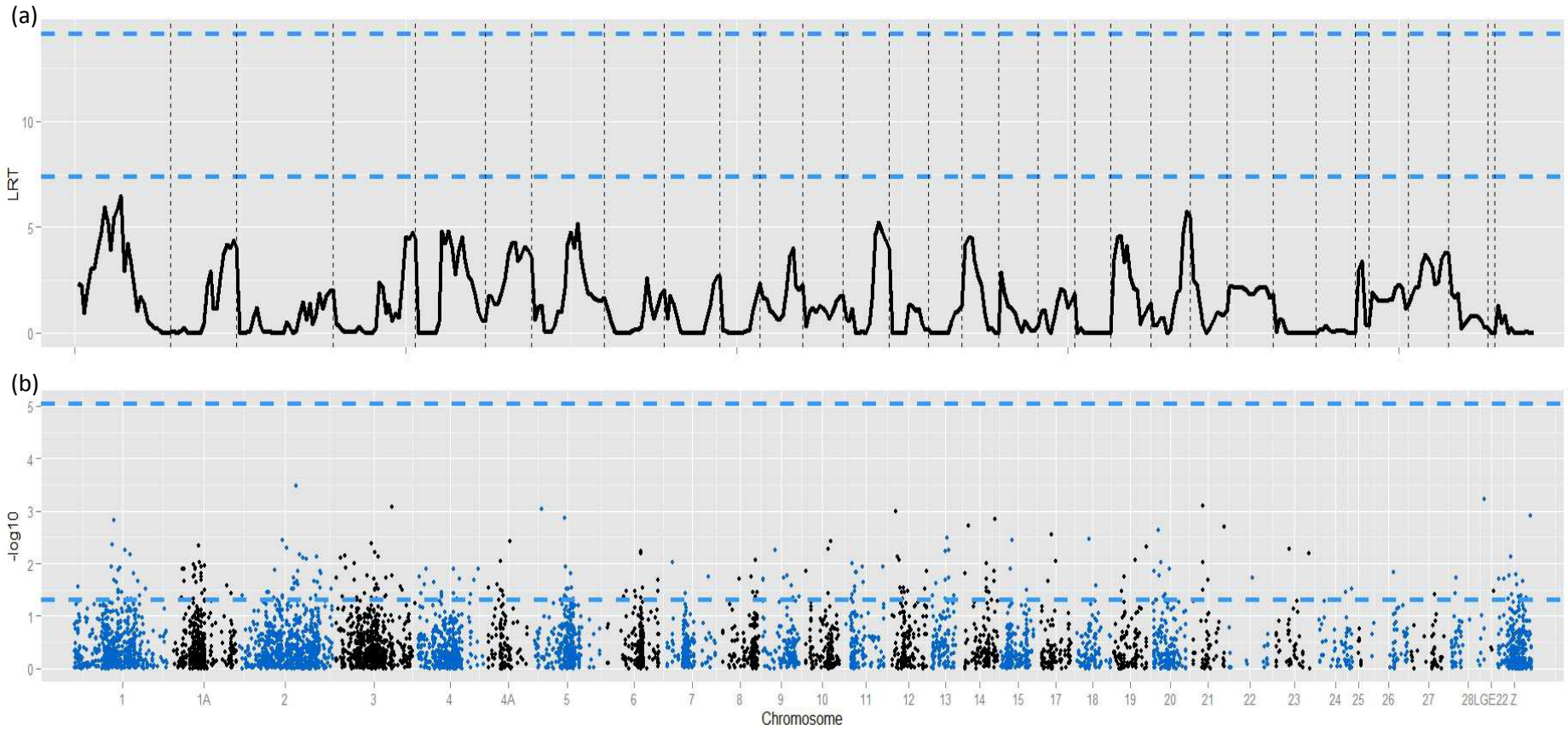


Figure 3.1. Genome scans using (a) linkage and (b) association mapping to detect genetic variants contributing to variation in laying date in the Hoge Veluwe (NL) great tit population. Dashed lines represent nominal (a) 7.38; b) 1.3) and genome-wide significant (a) 14.11; b) 5.05) thresholds.

3.4.3. QTL scan and GWAS results (NLUK)

Two QTL peaks passing the suggestive threshold were detected when QTL scans were performed on the NLUK dataset, on chromosomes 4A (representing two SNPs) and 10 (1 SNP). No QTL peaks passed the genome-wide significant threshold (Figure 3.3a). Two hundred and eighty three SNPs passed the suggestive threshold when GWAS was performed, again approximately the number that would be expected to exceed the threshold by chance. No SNPs breached the genome-wide significant threshold (Figure 3.3b). In line with results for the NL population in isolation, a permutation test of significance determined the highest associations for the combined dataset could statistically be expected to occur by chance and are not significant (Table 3.3).

A binomial test was applied in R to test for a shared genetic architecture; calculating the proportion of SNPs significant at $p<0.05$ for each population and multiplying them generated a proportion 0.0025 SNPs (or a total or ~13 SNPs) expected to be significant in both analyses. None of the 265 (NL) or 291 (WY) SNPs that exceeded the nominally significant GWAS threshold were common to both populations. The binomial test indicated that the null hypothesis of distinct genetic architectures for laying date between populations could not be rejected. Fewer SNPs were shared between the populations than would be expected by chance though this is unlikely to be biologically meaningful.

Table 3.3. The top 10 associations from GWAS analyses of laying date in the combined NL / UK great tit population dataset, following a permutation test of 1000 iterations to test for significance. Number of females for which complete phenotypic information was available (N), chromosome (Chr) and position (cM). Effect sizes (effB) and their standard errors (se_effB) of allelic substitutions with 1df. Nominal p value (pre-permutation Pc1df) and genome-wide p value from permutations tests before and after lambda correction (P1df, Pc1df). Genotype effect sizes relative to AA and test statistics from a model where genotypes are fitted as a three level factor described by effAB and effBB.

Marker	Chr	Position	A1	A2	N	effB	se_effB	chi2.1df	Pre-permutation Pc1df	P1df	Pc1df	effAB	effBB
M7391	14	5	G	A	1896	-0.194095	0.057457	11.411409	0.000429	0.978	0.838	-0.217444	-0.356669
M8073	2	50	A	G	1896	-0.165312	0.053077	9.700746	0.001031	1	0.996	-0.124355	-0.356850
M7935	1A	42	A	G	1893	0.163197	0.052794	9.555577	0.001068	1	0.998	0.212334	0.312123
M8426	2	94	G	A	1896	0.187656	0.061819	9.214727	0.001359	1	0.999	0.181071	0.389667
M8556	2	71	C	A	1891	0.168357	0.055816	9.097964	0.001600	1	1	0.128087	0.384251
M7750	4A	24	G	A	1897	-0.197062	0.066139	8.877477	0.001734	1	1	-0.250497	-0.201249
M4958	21	15	G	A	1895	-0.180656	0.060726	8.850314	0.001792	1	1	-0.054332	-0.623191
M6210	20	6	A	G	1892	0.156078	0.052674	8.779874	0.001886	1	1	0.133405	0.319018
M6296	19	14	C	A	1880	0.173324	0.058587	8.752127	0.002051	1	1	0.163786	0.362584
M7289	2	52	C	A	1891	0.172866	0.058599	8.702432	0.002054	1	1	0.223542	0.258746

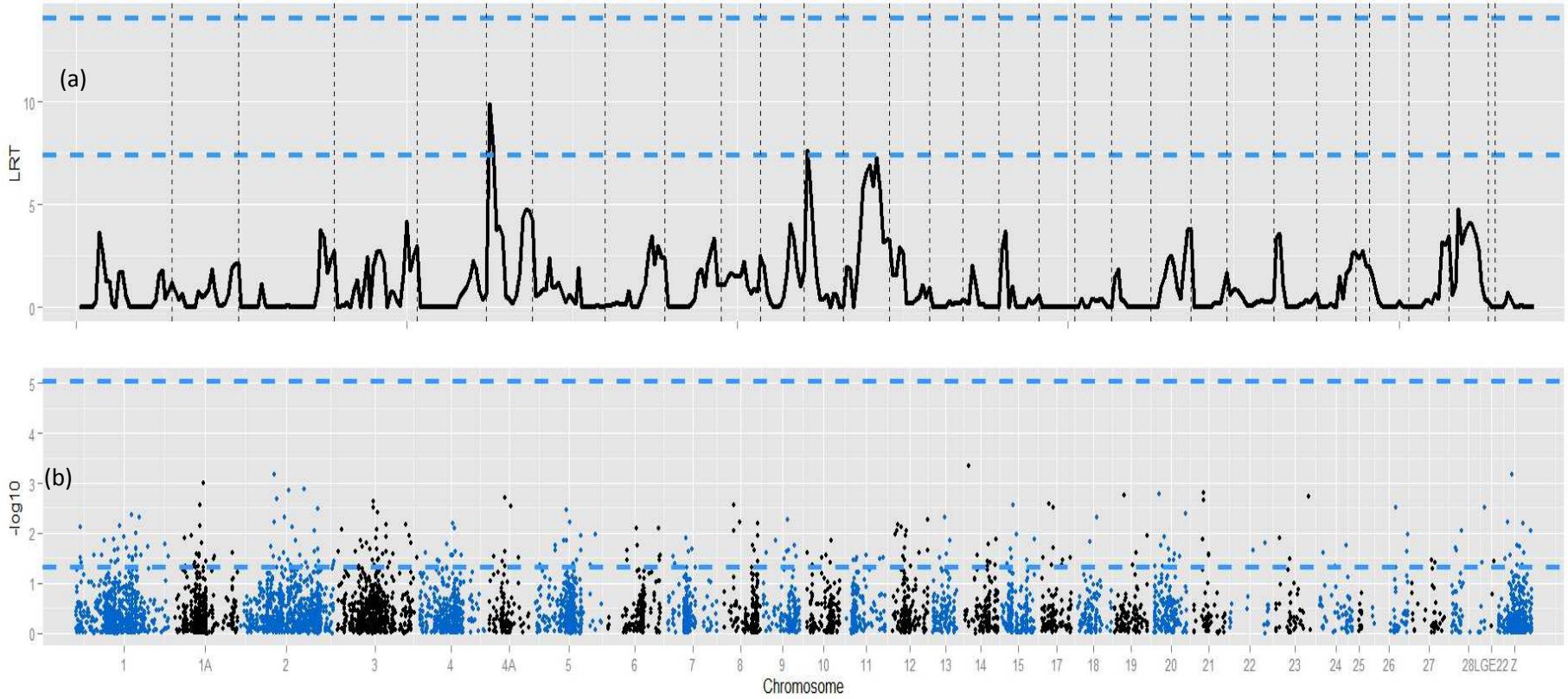


Figure 3.2. Genome scans using linkage (a) and association (b) mapping to detect genetic variants contributing to variation in laying date for the combined Hoge Veluwe (NL) and Wytham Woods (UK) great tit population dataset. Dashed blue lines represent (a) suggestive (7.38) and genome wide significant (14.11) linkage, and (b) nominal (1.3) and genome-wide significant (5.05) associations.

3.5 Discussion

The small number of studies to investigate the genetic basis of behaviour in wild (or captive wild) vertebrates have demonstrated the difficulties associated with attributing behavioural variation to underlying genetic loci, and they have often produced inconsistent or non-significant results (Korsten et al. 2010; Mueller et al. 2013; Poissant et al. 2013; Johnston et al. 2014; Riyahi et al. 2015). Combining some of the largest and most comprehensive phenotypic and genetic datasets for any wild animal I tackle the challenges associated with describing the genetic architecture of a polygenic trait.

The significant (yet modest) heritability estimate obtained by linkage analysis, undermined by the complete absence of any significant genomic regions detected by GWAS, suggests that there is little power in the dataset to detect loci underlying genetic variation in laying date, unless they are of large effect. Many studies emphasise the importance of sample size in being able to successfully detect causative genes, particularly for traits of low heritability (Beavis 1994; Hirschorn & Daly 2005; Risch 2000; Risch & Merikangas 1996). The number of individuals for which we have phenotypic data, even after augmenting the dataset by combining two populations, may be too few to enable effective localisation of genomic regions and detection of genes associated with phenotypic variation.

If, as is probable, different mechanisms contribute to variation in laying date between populations, resulting in distinct genetic architectures, increasing the sample size by combining the populations will fail to have the desired effect of amplifying our ability to detect causal variants. Initial associations between the DRD4 gene and exploratory behaviour in a single Dutch population, contrasted with the absence of such an association in other European populations (Korsten et al. 2010) exemplifies the potential for phenotypic variation between populations of the same species to be influenced by fundamentally different mechanisms, or at the very least, different genetic loci.

Adaptation by geographically distinct populations of a species to localised environmental conditions resulting in distinct phenotypic variation has been reflected in the breeding time of European bird populations, with observable variation in patterns of laying date in populations of great tits in response to temperatures experienced in different regions (Matthysen et al. 2011; Visser et al. 2003). Such fine-scale variation is a consequence of the sensitivity of the birds caterpillar prey to regional variation in temperature and topographic variation in vegetation phenology (van Asch et al. 2007; Stevenson & Bryant 2000; Hinks et al. 2015; Wilkin et al. 2007). The nature of laying date and the biology of *Parus major* are important factors to consider when attempting to determine the extent to which phenotypic variation has a genetic or plastic origin, and evidence from this study would support current consensus that plasticity and rapid responses to environmental cues play a larger role than genetic variation (Przybylo et al. 2000; Nussey et al. 2005; Charmantier et al. 2008; Visser et al. 2009; Charmantier & Gienapp 2014; Thorley & Lord 2015; Bourret et al. 2015).

Individual variation, even with little plasticity, can facilitate successful adaptation to a changing environment, with short lived species with high reproductive rates likely to be more resilient (Vedder et al. 2013). Research has shown that in years of population mismatch relative fitness differences between females are large but average absolute fitness is not dissimilar to years where the birds are in synchrony with their prey (Reed, Jenouvrier, et al. 2013), as fitness losses in mismatched broods are offset by increased fitness in other broods as a result of relaxed competition. Such factors may explain why the demography of Netherlands great tit populations have remained consistently stable despite mismatches between lay date the food peak (Reed, Grotan, et al. 2013; Reed, Jenouvrier, et al. 2013).

Linkage and association mapping represent two of the most widely used statistical tools available to dissect the genetic architecture of phenotypic variation. However a wide range of statistical methods exist and are continually being developed to estimate relatedness and partition genetic variance. To enable a more comprehensive description of the genetic component of variation contributing to variation in laying date, chromosome partitioning analyses were subsequently applied to each of the datasets (Chapter 4).

Chapter 4

Partitioning additive genetic variation for a reproductive trait in an ecological model organism

Jenny C. Armstrong[1], Jon Slate[1], Anna Santure[2], Matthew Robinson[3], Ben Sheldon[4], Marcel Visser[5], Kees van Oers[5], Nikkie van Bers[5, 6], Martien Groenen[6]

[1]Department of Animal & Plant Sciences, Alfred Denny Building, University of Sheffield, Western Bank, Sheffield S10 2TN, UK, [2]School of Biological Sciences Research, University of Auckland, New Zealand, [3]The Queensland Brain Institute, The University of Queensland, St Lucia, QLD 4072, Australia, [4]Edward Grey Institute, Department of Zoology, University of Oxford, Oxford OX1 3PS,UK, [5]Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW),POBox 50, 6700ABWageningen, the Netherlands, [6]Animal Breeding and Genomics Centre, Wageningen University, De Elst 1, Wageningen, 6708 WD, the Netherlands

4.1 Abstract

New statistical techniques are continually being developed to enable a description of the genetic architecture of phenotypes. Originally developed for genomic analyses on large human datasets, partitioning additive genetic variance associated with a trait across chromosomes has more recently been adapted for genomic analyses of wild organisms. For three datasets I applied chromosome partitioning to describe the genetic architecture of laying date in the passerine *Parus major.* Long-term studies of the great tits of Wytham Woods (UK) and Hoge Veluwe (NL) have generated substantial phenotypic and genetic data enabling the application of genomic statistical tools to describe the genetic architecture of phenotypic traits. The availability of such information for two geographically distinct populations of the same species, a relatively rare occurrence, also enabled the creation of a third combined dataset. A significant proportion of the genetic variance was attributed to single chromosomes in both the Wytham and Hoge Veluwe (NL) populations (chromosomes 6 and 13 respectively). Contrary to the predicted outcome of variance partitioning on complex quantitative traits, no statistically significant correlations were observed between the amount of variance explained by chromosomes and their size or number of genes, for any of the datasets. This is likely to be the result of a combination of two factors; low additive genetic variance associated with the trait, and low power in the dataset with which to accurately partition variance. However, it is possible that most of the variance is attributable to few chromosomes.

4.2 Introduction

Following the advent of high throughput genomic technologies, genome-wide analyses have underpinned the basis of investigations into the genetic architecture of complex traits (Slate et al. 2010; Kruglyak & Lander 1995), yet many studies of complex traits that find variants underlying QTL of small effect size cumulatively fail to explain all of the genetic variation attributed to the trait (Manolio et al. 2009; Eichler et al. 2010; Hirschhorn & Daly 2005; Vinkhuyzen et al. 2013). For example early methods to detect variants associated with human height were able to identify SNPs cumulatively explaining only ~5% of a narrow sense heritability of around 80% (Gudbjartsson et al. 2008; Lettre et al. 2008; Visscher 2008; Weedon et al. 2008; Manolio et al. 2009). Analyses of this phenomenon, termed 'missing heritability', have served to highlight some of the limitations that common methods employed to detect QTL are subject to.

Sample size is consistently regarded as one of the most influential factors with regards genomic analyses (Brito et al. 2011; Beavis 1994), with the estimated size of QTL likely to be inflated in studies with smaller samples sizes (Slate 2013) and the ability to localise QTL to precise genomic locations difficult even when they are detected (Beavis 1994; Mackay 2001). Smaller sample sizes may also exacerbate factors associated with difficulty in detecting variants such as incomplete linkage disequilibrium (LD) in the dataset (Yang et al. 2010), rendering variants not associated with typed markers undetectable.

The issue of being able to detect variants accurately or at all is increasingly reflected in a growing consensus that the ability of QTL analyses to detect genetic variants of small effect / rare occurrence is limited, that in fact heritability is not 'missing' only that the effects of many variants are too small to be detected by many of the current methods (Yang et al. 2010). Analyses have shown that the greater the number of variants contributing to genetic variance the smaller of number of QTLs that will actually be detected (Otto & Jones 2000).

Variation in many complex traits is thought to be maintained either by many common variants each of small effect (Yang, T. Lee, et al. 2013; Yang, Manolio, et al. 2011; Eyre-Walker 2010; Lee et al. 2012), or rare variants (Eyre-Walker 2010). Therefore the power to detect them may be low and the goals of simultaneously localising and estimating the effect size of QTL cannot be performed to the precision we increasingly require i.e. to identify the effects of causal mutations (Göring et al. 2001). In disciplines such as evolutionary biology, medicine and agriculture the importance of understanding the genetic architecture of phenotypic variation is recognised (Hill et al. 2008). Driven by the aforementioned limitations, new methods are continually being sought to improve ways to which we can accurately estimate the additive genetic component of variation and further identify the genomic regions / genes contributing to it.

In an attempt to increase the accuracy of genomic analyses, a method built on marker based QTL methodologies, chromosome partitioning (otherwise known as genome partitioning) is increasingly being applied to describe the genetic architecture of complex phenotypic traits (Visscher et al. 2007; Yang, Manolio, et al. 2011). Calculating actual relatedness (usually referred to as the 'realised relatedness') between individuals, as opposed to pedigree-based estimates of relatedness such as expected identity by descent (IBD) have been used in human studies to identify the genetic structure of traits such as height (Visscher et al. 2007; Yang, Manolio, et al. 2011).

Pedigree-based relatedness estimates produce IBD matrices that represent the expected proportions of alleles shared between individuals (Blouin 2003; Slate et al. 2009; Visscher 2009). Chromosome partitioning relies on using realised relatedness to calculate the covariance between the proportion of the genome shared and phenotypic similarity between individuals (Visscher et al. 2006). Adapting the method for a wild organism, Robinson et al. (2013) partitioned variance across the genome by calculating identity by

state (IBS) and constructing single chromosome genome-wide relationship matrices (GWRMs). These were then run in a familiar restricted maximum likelihood (REML) framework, the variance attributable to each genomic segment quantified by means of likelihood ratio testing (Robinson et al. 2013).

Conclusions drawn from chromosome partitioning studies have implied polygenicity is ubiquitous in complex human traits (Yang, T. Lee, et al. 2013), with observations that multiple chromosomes contribute to additive genetic variation. Variation is not typically localised to specific regions but spread across the entire genome, with strong correlations between the length of genomic segments and the amount of variation explained (Yang, T. Lee, et al. 2013; Yang, Manolio, et al. 2011; Visscher et al. 2007). Larger chromosomes possess a greater number of sites potentially harbouring genetic variants contributing to phenotypic variation, therefore theoretically we would expect a positive correlation between the proportion of variance explained by chromosomes and their size.

Given the much greater availability of large-scale datasets and practical / economic interest, genome partitioning has thus far been applied most widely to studies of humans, livestock and crop plants (Yang, T. Lee, et al. 2013; Dekkers 2012; Riedelsheimer et al. 2012). Due to inherent difficulties in studying animals in their natural environment, the lack of extensive genetic and phenotype data for any wild animal population remains the greatest limitation when interested in understanding the genetic architecture of traits in natural populations (Hill 2012; Slate 2005).

The great tit *Parus major*, introduced in Chapter 1, provides a rare example of a wild organism for which multiple long-term studies have yielded extensive amounts of phenotype and genetic data for multiple populations. Therefore, it provides a unique opportunity to apply genomic tools to describe and compare the genetic components of phenotypic trait variation. Robinson et al. (2013) applied chromosome partitioning analysis

to data available for the great tits of Wytham Woods, the first study to use such an approach on a wild population. Consistent with studies on complex traits in humans, the results of chromosome partitioning found multiple genomic regions contributing to variance in wing length, with the contribution of each chromosome relative to its gene content (Robinson et al. 2013; Santure et al. 2013).

QTL and genome-wide analyses presented in previous chapters have consistently been unable to detect variants associated with laying date. In this chapter I employ the method developed by Robinson et al. (2013) to partition genetic variance across the genome. This analysis represents the first time genetic variation will be partitioned across chromosomes for multiple populations of *Parus major*. The analysis should increase the understanding of the genetic architecture of life-history traits in an ecological model organism, and is especially relevant as laying date is a reproductive life-history trait, that potentially enables local adaptation in a wild animal.

4.3 Materials & Methods

4.3.1. Phenotype information

As detailed in Chapters 2 & 3 birds in the Wytham Woods (UK) and Hoge Veluwe (NL) populations are regularly caught and ringed to enable detailed accounts of individual life histories, morphological and behavioural characteristics to be made. Laying date is a numerical value taken as the number of days from April 1st. Nests subject to experimental manipulation were excluded and all breeding events (hereafter BE's) that exceeded 30 days from the first recorded BE of same year in which the event occurred (considered likely second broods) were removed so the analysis considered first broods only. Breeding events that predated 1985 were also excluded as no genetic information was available for birds before this period. Females for which age of birth was not known were assumed to be age '2' at first capture.

Data for Wytham (WY) consists of 1523 BE's representing 951 females for the period 1985-2011. For the Netherlands (NL) 1614 lay date records exist for 947 females spanning the period 1993-2008. Combining the populations to generate a master phenotype file subsequently resulted in a dataset consisting of 3137 BE's for 1898 females.

4.3.2. Pedigree & marker information

Based on a SNP chip consisting of 9193 SNP markers, 4702 birds from the two long-term GT study populations were genotyped (Van Bers et al., 2012). Two thousand samples were taken from NL birds with the remaining 2702 samples obtained from the UK. As with the marker data utilised in GWAS analyses (Chapters 2 & 3), chromosome partitioning utilises parsimonious linkage map marker data, taking the form of 5582 and 5592 SNPs placed on the NL and UK maps spanning 2010cM and 1917cM respectively. Given the small number of markers on some of the chromosomes, the 32 chromosomes were modified into 23

chromosomes and 'chromosome sets' of smaller chromosomes. Chromosomes 1-15, 17-20, 1A, 4A, LGE22 and the Z chromosome were analysed as individual chromosomes, with chromosomes 21-28 merged into a set known as the 'micro-chromosomes' (Santure et al. 2013).

Genotype information was available for 1407 and 2497 individuals from the NL and UK populations respectively. For the combined analysis, consistent with the modifications made for QTL and GWAS analyses (Chapter 3) NL individuals retained their original numeric IDs, WY individual IDs were modified with the addition of 1407 to each of their numeric IDs, creating genotype files for 3904 NLUK individuals.

4.3.3. Chromosome partitioning model method & statistical testing

Partitioning genetic variation across chromosomes was invented for application to human genetic studies (Yang et al. 2010; Yang, Manolio, et al. 2011), and was recently adapted for use in wild animal systems (Robinson et al. 2013). The method presented here follows the methodology presented by Robinson et al., (2013), using a different subset of phenotype data on the same study organism.

To generate IBS matrices for chromosome partitioning analysis genotype data first has to be converted from the format describing the alleles making up each individual's genotype at each SNP marker, to a -1, 0, 1 format, representing homozygote, heterozygote, and the other homozygote. Relationship matrices (GWRMs) were generated for 22 of the chromosomes / chromosome sets (1-15, 17-20, 1A, 4A, 21-28). The Z chromosome was excluded from chromosome partitioning analysis on the basis that SNP alleles are coded differently on the sex chromosome because it is heterogametic (Robinson et al. 2013), and for the purposes of describing the genetic architecture of laying date analyses of the autosomes would be sufficient.

For each of the 22 chromosomes (which includes the set of micro chromosomes) three different types of relationship matrices were constructed; one for all markers excluding those on the focal chromosome, one constructed using all markers and one containing only those markers on the specific chromosome (Robinson et al. 2013). Four mixed models using different combinations of relationship matrices were specified as follows, where 'X' is a focal chromosome;

GWRM (minus chromosome 'X') (model i)

GWRM (minus chromosome 'X') + GWRM (only chromosome 'X') (model ii)

GWRM (all markers) (model iii)

GWRM (all markers) + GWRM (only chromosome 'X') (model iv)

Using the principle of the animal model to calculate heritability, the four mixed model sets were then used to partition variance across each of the genomic regions. Two likelihood ratio tests (LRT) were used to investigate the contribution of each chromosome to the additive genetic variation attributed to laying date. LRT1 contrasts model (ii) against model (i) to determine the contribution of each chromosome to phenotypic variation. LRT2 contrasts model (iv) against model (iii) to assess whether variation exhibited by chromosome(s) is greater than expected given the region's size and gene content. The null hypothesis for the first test is that the chromosome explains no V_A; a significant result indicates the chromosome does contribute to V_A. The second test has a null hypothesis that the chromosome explains the amount of V_A that would be expected for its size, while a significant result indicates it explains more than expected for its size.

Linear models were constructed to test whether the trait has a polygenic architecture, by testing whether the V_A explained by each chromosome scales with its size (Mbp) and number of genes. The null hypothesis is that there is no relationship between chromosome

parameters and the proportion of variance explained while a positive relationship is consistent with the pattern of variation that would be expected from a polygenic pattern of genetic variation.

4.4 Results

4.4.1. Likelihood ratio testing

The twenty-two chromosome sets failed to account for any phenotypic variation in laying date in the UK great tit population, with heritability of laying date, based on an animal model incorporating a genome-wide relationship matrix estimated as zero (Table 4.1). LRT testing identified four chromosomes contributing to V_A but only one (chromosome 6) was determined to be significant (Table 4.2).

Table 4.1. Estimates of heritability for laying date in the UK, NL and combined NLUK great tit datasets, as obtained from animal modelling and chromosome partitioning (G_{matrix}), showing standard error.

	h^2		
	UK	**NL**	**NLUK**
Animal model	0.05 (±0.03)	0.15 (±0.06)	0.08 (±0.03)
G_{matrix}	0.00 (±0.00)	0.17 (± 0.03)	0.10 (± 0.03)

For the Netherlands population a significant genetic component of variation was detected using marker derived estimates of relatedness with an estimated h^2 of 0.17 (Table 4.1) divided between 11 chromosomes (Table 4.2). Likelihood ratio testing identified two chromosomes (4A and 13) as explaining significant amounts of additive genetic variation (*p* = 0.05, 0.02 respectively) (Table 4.2) though chromosome 4A explained no more than would be expected given its size (LRT2).

For the NLUK dataset a heritability of 0.10 (Table 4.1) was calculated using marker data alone. Ten chromosomes each explaining a h^2 of 0.01 (with the exception of the micro chromosome set; h^2 = 0.02) explained the total additive genetic variance of the trait (Table

4.2). Likelihood ratio testing identified a single chromosome (12) as explaining a significant amount of variation (p = 0.03) but did not explain more than expected given its size (LRT2). No chromosome set explained more V_A than expected.

One of the main purposes of these analyses was to determine if the same chromosomes contributed significantly to V_A in both populations. A direct comparison of the variance explained by the chromosomes between the UK and NL populations was largely uninformative and not significant (p = 0.08) (Figure 4.1), as all but four chromosomes in the UK explain no variance at all. There was also no correspondence between populations using likelihood ratio testing, with both LRT tests attributing significant V_A to different chromosomes; chromosome 6 (UK) and chromosome 13 (NL) (Table 4.2).

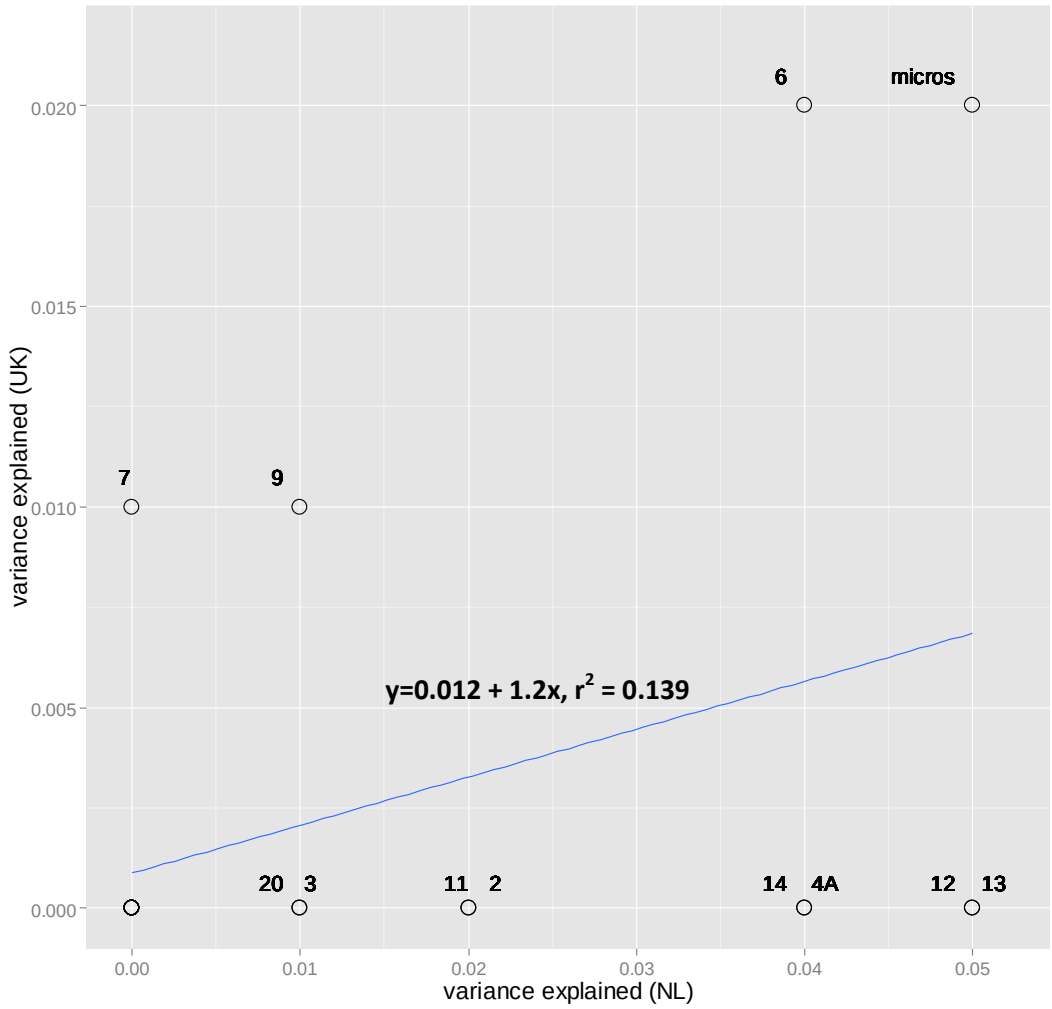


Figure 4.1. Results of the chromosome partitioning analysis, showing the relationship between the variance (h^2) explained by chromosomes in the UK and NL great tit populations (p = 0.09). For clarity chromosomes that explain no variance are not labelled.

Table 4.2 Likelihood ratio test (LRT) results for chromosome partitioning analysis performed on the Wytham Woods (UK), Hoge Veluwe (NL) and combined population (NLUK) datasets, for each chromosome (Chr) including micro chromosomes (M). Statistically significant results are highlighted in bold (continued on the next page)

						UK			NL			NLUK		
Chr	**No. SNPs (WY)**	**No. SNPs (NL / NLWY)**	**Length (cM)**	**Number genes**	**Size (Mbp)**	**h^2 (+SE)**	**LRT1**	**LRT2**	**h^2 (+SE)**	**LRT1**	**LRT2**	**h^2 (+SE)**	**LRT1**	**LRT2**
1	579	579	139.9	1254	119.6	0	0	0	0	0	0	0.01 (0.01)	0.4	0
1A	415	415	93.6	972	73.7	0	0	0	0	0	0	0	0	0
2	700	699	139.7	1450	156.4	0	0	0	0.02 (0.05)	0.16	0	0.01 (0.02)	0.2	0
3	596	596	114.9	1290	112.6	0	0	0	0.01 (0.04)	0.04	0	0.01 (0.01)	0.86	0
4	356	355	97.6	811	69.8	0	0.06	0.06	0	0	0	0	0.02	0
4A	103	103	59.4	39	20.7	0	0	0	0.04 (0.03)	**3.9**	3.16	0.01 (0.02)	1.18	0.52
5	346	346	98.6	998	62.4	0	0	0	0	0	0	0.01 (0.01)	0.66	0.02
6	177	177	78	596	36.3	0.02 (0.01)	**4.56**	**4.56**	0.04 (0.03)	1.88	1.34	0	0.38	0.04
7	176	176	72.6	562	39.8	0.01 (0.01)	0.68	0.68	0	0	0	0	0	0
8	134	134	53.8	575	28	0	0	0	0	0	0	0.01 (0.01)	3.28	1.84
9	130	130	54.2	497	27.2	0.01 (0.01)	0.32	0.32	0.01 (0.02)	0.12	0.002	0	0.18	0
10	148	148	50.5	444	20.8	0	0	0	0	0	0	0	0	0
11	135	135	58.2	397	21.4	0	0.04	0.04	0.02 (0.02)	0.78	0.48	0	0.16	0
12	152	151	51.9	369	21.6	0	0	0	0.05 (0.03)	3.68	2.82	0.01 (0.01)	**4.44**	2.74
13	117	115	40.9	379	17	0	0	0	0.06 (0.03)	**5.3**	**4.46**	0	0.3	0.06

Table 4.2 (continued)

						UK			NL			NLUK		
Chr	**No. SNPs (WY)**	**No. SNPs (NL / NLWY)**	**Length (cM)**	**Number genes**	**Size (Mbp)**	**h^2 (+SE)**	**LRT1**	**LRT2**	**h^2 (+SE)**	**LRT1**	**LRT2**	**h^2 (+SE)**	**LRT1**	**LRT2**
14	126	126	49.2	426	16.4	0	0	0	0.04 (0.03)	2.44	1.86	0.01 (0.01)	0.72	0.24
15	173	173	49.1	381	14.4	0	0	0	0	0	0	0	0.02	0
17	96	96	45.4	336	11.6	0	0	0	0	0	0	0	0	0
18	93	93	49.9	334	11.2	0	0	0	0	0	0	0	0	0
19	97	97	49.4	348	11.6	0	0.10	0.10	0	0	0	0	0	0
20	155	155	49.4	356	15.7	0	0	0	0.01 (0.02)	0.18	0.04	0.01 (0.01)	1.38	0.6
M	308	307	411	1808	40.2	0.02 (0.01)	1.96	1.96	0.05 (0.04)	1.96	1.22	0.02 (0.01)	3.56	1.68

4.4.2. Linear modelling

Linear modelling for variance explained by chromosomes in relation to their size parameters using the UK dataset largely confirmed LRT results of no significant V_A being attributed to laying date in this population, with no significant relationships between number of genes (Figure 4.2a) or size (Mbp) (Figure 4.2b). Using the NL dataset linear modelling of the number of genes and chromosome size (Mbp) against the variance (h^2) explained found no significant correlations with either of these parameters (Figure 4.3).

Linear modelling of the relationships between the two parameters and the amount of variance explained by chromosomes using the NLUK dataset revealed positive relationships in both cases (Figures 4.4a, 4.3b), with the number of genes (p = 0.05, r^2 = 0.17) (Figure 4.4a) explaining significant variation. To ascertain whether this result was purely a consequence of the cumulative size of the micro chromosomes, the micro chromosomes were removed to compare the effect of their inclusion on the pattern and statistical significance. The relationship between variance explained and number of genes remained positive but was non-significant (p > 0.1, r^2 = 0.03)

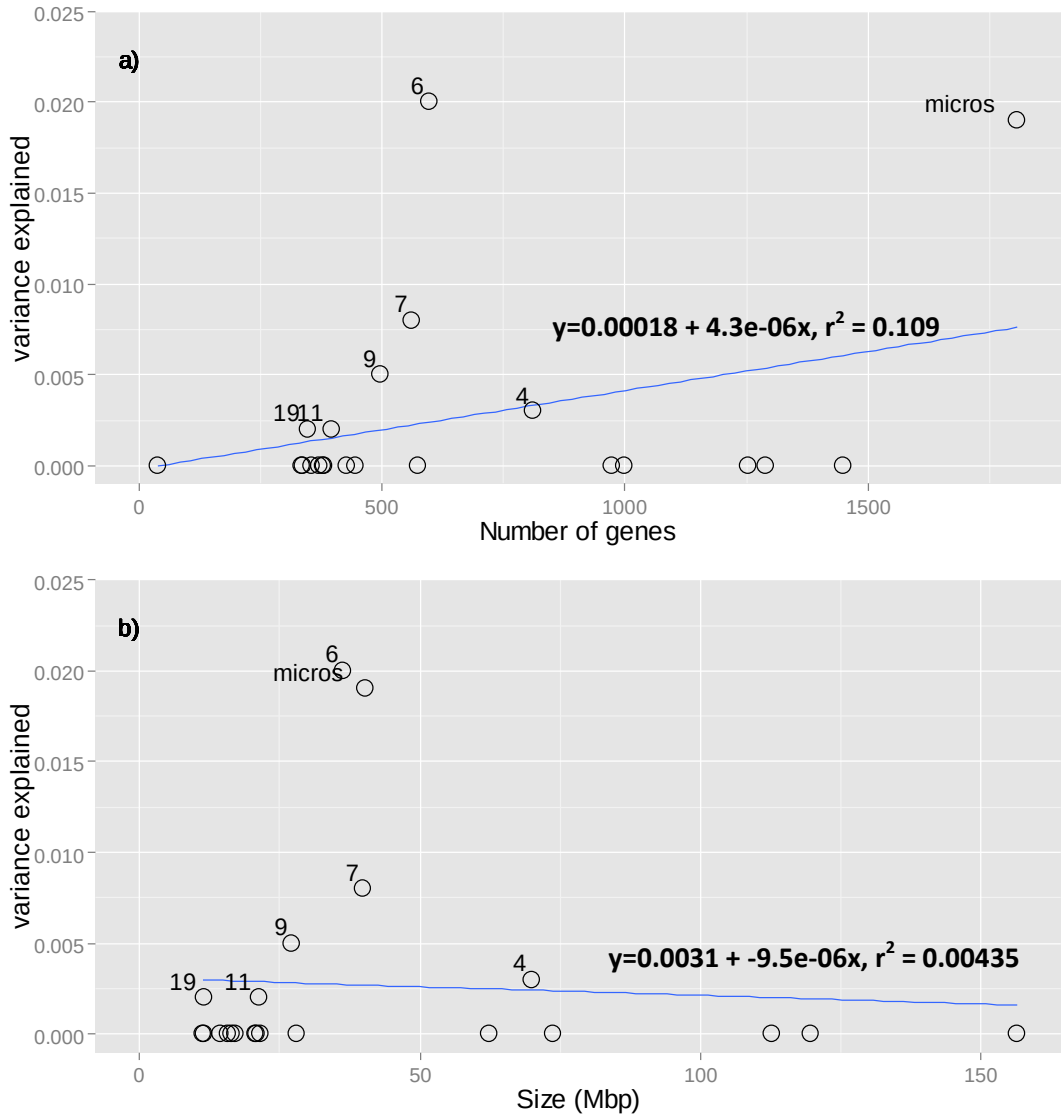


Figure 4.2. Linear modelling for variance attributed to chromosomes in relation to (a) number of genes ($p = 0.13$) and (b) size (Mbp) ($p = 0.77$) for the UK great tit population. For clarity chromosomes that explain no variance are not labelled.

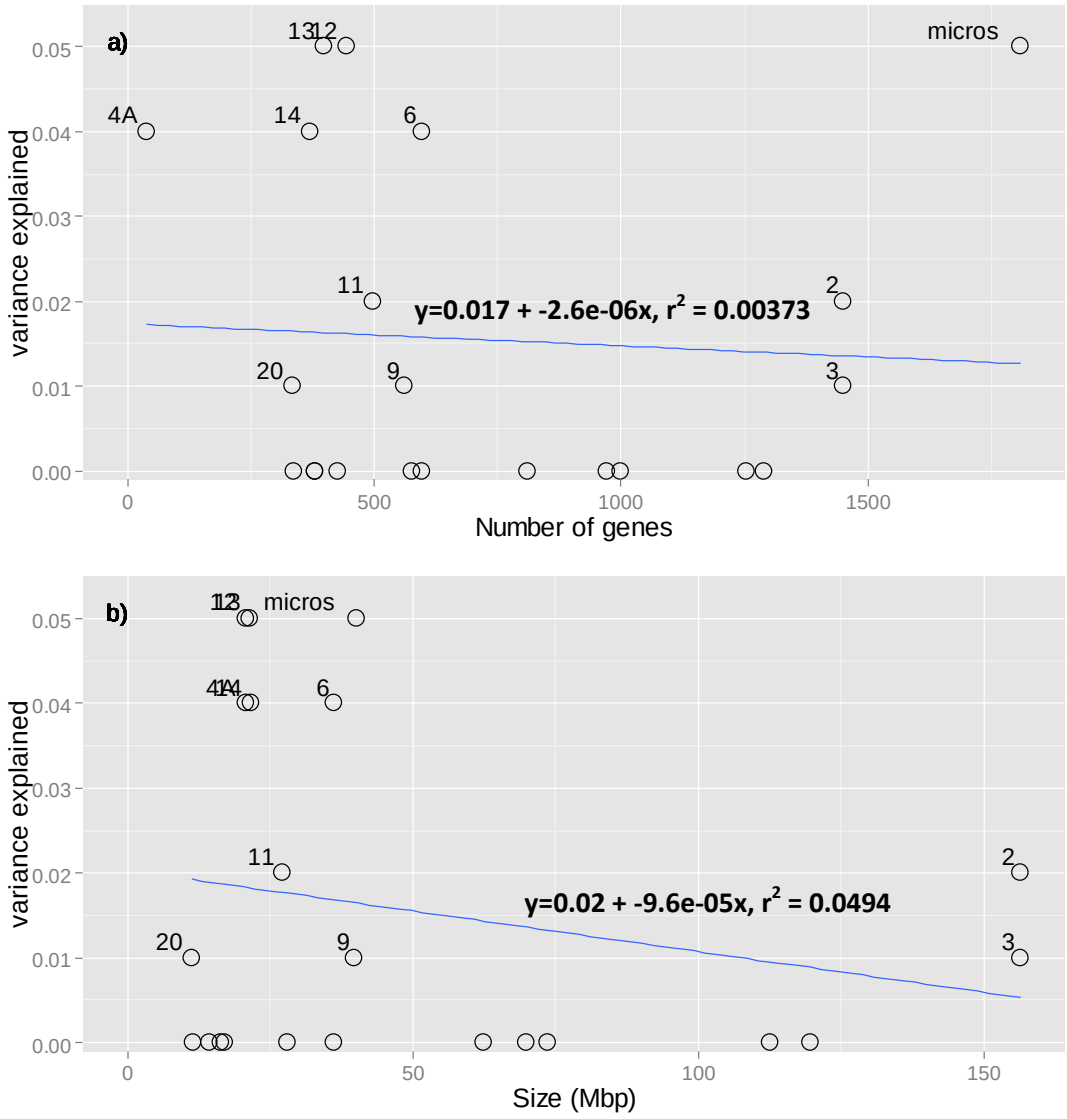


Figure 4.3. Linear modelling for variance attributed to chromosomes in relation to (a) number of genes ($p = 0.79$) and (b) size (Mbp) ($p = 0.32$) for the NL great tit population. For clarity chromosomes that explain no variance are not labelled.

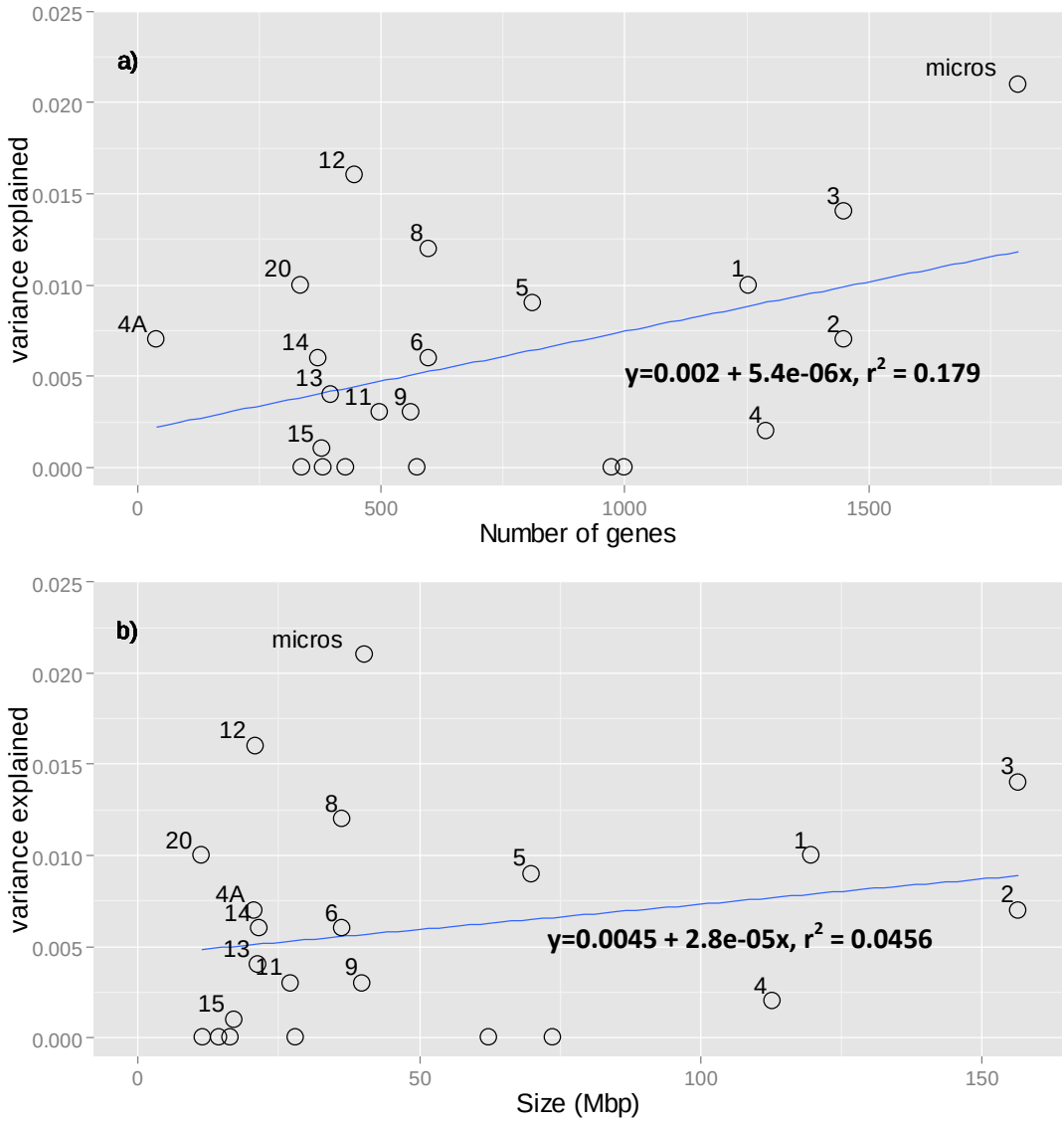


Figure 4.4. Linear modelling for variance attributed to chromosomes in relation to (a) number of genes ($p = 0.05$) and (b) size (Mbp) ($p = 0.34$) for the NLUK dataset. For clarity chromosomes that explain no variance are not labelled.

4.5 Discussion

As a comparison of the genetic architecture of a trait in multiple populations of a species, chromosome partitioning analyses revealed little correspondence between the UK and NL populations. Two different chromosomes with a significant contribution to V_A (LRT1), above that expected given their size (LRT2) were detected in the UK (chromosome 6) and NL (chromosome 13) populations (Table 4.2). The variance attributed to the significant chromosomes is small, explaining 2% (UK) and 5% (NL) of the additive genetic component.

With little additive genetic variation for laying date in the UK little could be inferred from a comparison of the chromosomes contributing to variance between populations. Unlike the UK, where most of the additive genetic variation was captured by a few chromosomes, statistical testing on the NL and combined datasets indicated that additive genetic variance was distributed more widely across chromosomes.

Contrary to expectations for both the NL and combined datasets, no significant relationships were observed between chromosome parameters and variance explained. Trends of the amount of variance explained in relation to chromosome size and number of genes for the NL population were predominantly negative. Larger chromosomes explained little to no variation, with contributions to additive genetic variation distributed among the smaller chromosomes (Figure 4.2). While trends for the combined dataset were positive, they were also non-significant. These results are contrary to what would be predicted assuming a polygenic mode of inheritance, where positive trends would be expected to be observed between measures and the amount of variance explained (Yang, Manolio, et al. 2011; Santure et al. 2013).

These results would suggest either an oligogenic genetic architecture for laying date, with variance attributed to single genes of large effect, or a lack of statistical power with which

to accurately partition variance resulting in the observed departure of patterns from those that were hypothesised. If variance has been partitioned accurately, and an oligogenic basis for variation is assumed, statistical testing has revealed a different genetic architecture for laying date between populations, with two different statistically significant chromosomes associated with phenotypic variation in the UK and NL.

The existence of single genes of large effect associated with phenotypic variation has been observed in wild organisms (Johnston et al. 2010; Tarka et al. 2010; Johnston et al. 2011; Poissant et al. 2012) but there is evidence that this effect may in some cases be a consequence of small sample size. Smaller sample sizes can lead to inflated estimates of V_A (Slate 2013), an effect that is known as the Beavis effect (Beavis 1994; Slate 2013).

The genetic architecture of complex traits where many common variants of small effect exist are notoriously difficult to describe, as the effective contribution of each genomic region is so small that statistical testing often fails to identify causal variants (Yang et al. 2010; Risch 2000; Cardon & Bell 2001). With other analyses of the UK great tit dataset having suggested inadequate power to detect genetic variants (Santure et al. 2013), a lack of power in the datasets analysed here, due to smaller sample sizes than is common in human or domestic animal datasets (Visscher et al. 2007; Weedon et al. 2008; Pimentel et al. 2011), may have limited the capacity of statistical analyses to accurately partition variance.

Chapter 5

The genetic architecture of laying date; results from a high density chip

Jenny C. Armstrong[1], Jun-Mo Kim[1,2], Anna Santure[1,3], Ben Sheldon[4], Jon Slate[1]

[1]Department of Animal & Plant Sciences, Alfred Denny Building, University of Sheffield, Western Bank, Sheffield S10 2TN, UK; [2]Centre for Nutrition and Food Sciences, Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Australia; [3]School of Biological Sciences Research, University of Auckland, New Zealand; [4]Edward Grey Institute, Department of Zoology, University of Oxford, Oxford OX1 3PS,UK

5.1 Abstract

The recent development of a new high density single nucleotide polymorphism (SNP) chip, with 500,000 SNP markers mapped to the great tit genome, has provided more powerful genomic tools to accurately identify regions and genes associated with heritable phenotypic variation. Previous work in this thesis and elsewhere has described a small yet heritable component of phenotypic variation in laying date, and provided a description of the genetic architecture based on a genomic map composed of fewer than 8000 markers. I now take advantage of the substantially greater number of markers to employ genome wide statistical analyses in an attempt to increase the precision with which regions associated with variation in this trait can be identified. GWAS analyses failed to find any SNPs of genome-wide significance. This finding is consistent with numerous genes of small effect, in line with what would be expected for a complex polygenic behavioural trait. However chromosome partitioning analyses attributed a significant proportion of V_A to chromosome 3, consistent with a SNP that almost reached genome-wide significance on the same chromosome detected by GWAS. Based on the results obtained from these analyses I further suggest that the dominant limiting factor in the ability to detect genetic variants is sample size itself and not marker density.

5.2. Introduction

In birds, the synchronicity of breeding time with peak food availability is known to be a crucial determining factor in the reproductive success of individual females. In great tits lower chick weight and fledging success are some of the consequences of mistimed breeding events (Perrins 1970; Nussey et al. 2005; Gebhardt-Henrich 1990; Visser et al. 2006). Given the importance of laying date on individual fitness, mechanisms are expected to have evolved to enable females to correctly time the laying of their eggs to ensure their offspring have the best possible chances of survival on hatching.

Most of the variation in lay date can be explained by sensitivity to environmental cues (Visser et al. 2009; Conway & Martin 2000), and it is likely that a great number of avian breeding and migratory behaviours are instinctive reactions to these cues, particularly temperature (Conway & Martin 2000; Torti & Dunn 2005; Ardia et al. 2006; Visser et al. 2009; Garroway et al. 2013; Charmantier & Gienapp 2014; Tarka et al. 2015). However evidence from populations of the great tit *Parus major* have suggested a genetic mechanism contributing to variation in laying date (van Noordwijk et al. 1981; van der Jeugd & McCleery 2002; McCleery et al. 2004; Gienapp et al. 2006; Garant et al. 2008; Schaper et al. 2011; Visser, Schaper, et al. 2011; Liedvogel et al. 2012).

Using several different statistical methods previous Chapters in this thesis described attempts to understand the genetic architecture of laying date, but there was little correspondence between the results. Linkage analyses of the Wytham Woods (UK) and Hoge Veluwe (NL) great tit populations failed to identify any regions of genome-wide significance (Chapters 2 & 3), though a region of chromosome 10 reached a nominally significant threshold using the UK dataset (Chapter 2).

Association (GWAS) analyses did not detect any SNPs for either population at a genome-wide significant level, with no more passing nominally significant thresholds than would be expected by chance (Chapters 2 & 3). Combining the datasets to increase overall sample size and statistical power resulted in the detection of two nominally significant SNPs using linkage analyses, on chromosomes 4A and 10, but again no genome-wide significant results were detected using either linkage or association methods (Chapter 3).

Finally genetic variance attributed to laying date was partitioned across the genome for each population in isolation and the combined dataset (Chapter 4). Contrary to an overall hypothesis and the results of previous analyses that would predict a polygenic architecture, variance for the UK and NL populations was attributed to single (different) chromosomes (6 and 13 respectively). Relatively large proportions of variance were further attributed to chromosomes 4A (NL) and chromosome 12, (NLUK) which though statistically significant, did not explain more variance than expected given their size.

Limited evidence for an oligogenic architecture was further supported by linear modelling and the absence of positive relationships between chromosome size and variance explained by chromosomes. However, low additive genetic variation may have masked a polygenic architecture, with variance approximately equally partitioned between chromosomes too small to reach statistical significance.

Assuming a polygenic mode of inheritance, as variance is distributed equally across the genome, larger chromosomes harbouring a greater number of genes would be expected to explain a greater proportion of variance (Yang, Manolio, et al. 2011). This would generate positive relationships between chromosome size and variance explained. A lack of statistically significant relationships between variance explained and chromosome size suggests either that few genes of large effect contribute to variation in laying date, or a lack of statistical power with which to accurately partition variance.

Complex traits are usually considered by their nature to be polygenic (Reif & Lesch 2003; Yang, Manolio, et al. 2011). In humans and domestic animals the polygenicity of quantitative traits have been largely substantiated by genomic analyses (Visscher 2008; Lango Allen et al. 2010; Yang et al. 2010; Pimentel et al. 2011; Lee et al. 2012; Yang, T. Lee, et al. 2013; Carneiro et al. 2014). In wild organisms however, analyses attempting to describe the architecture underlying complex traits have produced more varied results.

The results of some studies of wild organisms support the hypothesis of variance explained by many genes of small effect (Robinson et al. 2013; Santure et al. 2013; Husby et al. 2015; Bérénos et al. 2015) whilst others have detected genes of large effect contributing to phenotypic variation (Colosimo et al. 2004; Beraldi et al. 2007; Gratten et al. 2007; Johnston et al. 2010; Tarka et al. 2010). Whether variance in these cases is truly represented by few genes of large effect, or whether these results are an artefact of small sample sizes inflating QTL effect sizes (Beavis 1994; Slate 2013; Bérénos et al. 2015) is uncertain.

It is difficult to ascertain whether suggestive QTL (in the absence of genome-wide QTL) for traits in wild animals (Beraldi, McRae, Gratten, Slate, et al. 2007; Beraldi, McRae, Gratten, Pilkington, et al. 2007; Poissant et al. 2012; Poissant et al. 2013; Wenzel et al. 2015) (Chapters 2, 3 and 4) are true associations or represent a similar lack of power with which to accurately partition variance. Increasing the power of statistical analyses may provide a way to determine whether the genetic architecture of some quantitative traits in wild animals truly depart from a polygenic model.

Low additive genetic variance (V_A), low marker density and limited phenotypic data are all factors that make detecting genetic variants and partitioning variation difficult and may explain why previous analyses have failed to converge on genomic regions associated with variation in laying date. Genome-wide analyses on humans typically utilise genomic

information obtained from marker chips comprising hundreds of thousands of markers for thousands of subjects (Duerr et al. 2006; Franke et al. 2010; Kappen et al. 2015; Yang et al. 2010; Yang, Manolio, et al. 2011; Vinkhuyzen et al. 2013). Such resources for wild organisms are becoming increasingly available, enabling more accurate and precise descriptions of the genetic basis of variation (Johnston et al. 2011; Robinson et al. 2013; Santure et al. 2013).

Increasing marker density has the potential to increase the resolution of genomic analyses in a number of ways. The detection of QTL in GWAS studies relies on the extent of linkage disequilibrium (LD) between markers and causal loci (Pritchard & Przeworski 2001; Li & Merilä 2010; Slate et al. 2010; Bush & Moore 2012; Hill 2012; Bérénos et al. 2015) which can be low if causal variants have lower minor allele frequency (MAF) than genotyped SNPs (Wray 2005; Tabangin et al. 2009; Vinkhuyzen et al. 2012).

If the genetic variance associated with a trait is composed of rare variants, with lower minor allele frequencies than markers, statistical analyses may fail to identify the genes contributing to V_A (Manolio et al. 2009; Yang et al. 2010; Hill 2012; Speed et al. 2012). Increasing marker density not only increases the likelihood of SNPs being in LD with causal loci, but with wider SNP frequency distributions, including some with low minor allele frequencies, the identification of variants that previously would have gone undetected becomes possible (Xu 2003; Wang et al. 2005; Hill 2012).

Extensive efforts have been made to develop genomic resources for the great tit, one of a handful of wild vertebrates for which long-term observations and DNA samples are available for multiple populations. Over 4000 birds were genotyped on a 10K SNP chip, to develop a genetic map comprised of 32 linkage groups (van Bers et al. 2010; Santure et al. 2011; van Oers et al. 2014). More recently, as the result of collaborations between several research institutions DNA samples from over a dozen populations of great tits have been

used in the development of a high density SNP chip mapping around half a million markers to the great tit genome (in press).

Using genotype information obtained from the new SNP chip I revise my attempt to identify genomic regions associated with variation in laying date in the Wytham Woods great tit population, applying association mapping and chromosome partitioning statistical techniques to seek genetic variants for laying date.

5.3. Methods

5.3.1. Processing the Axiom array

As part of the collaborative great tit HapMap project DNA, samples representing 2010 individual great tits from 27 populations inhabiting Europe, Russia and Asia were selected for inclusion on a 500K SNP chip, the largest effort of its kind generating perhaps the most extensive genomic resource available for any wild vertebrate (unpublished). SNP genotyping was performed with an Affymetrix Axiom custom chip at Edinburgh Genomics, with output in the form of CEL files detailing raw probe intensity data for each sample.

Genotype calling for individuals was performed using the AffyPipe genotyping pipeline (Nicolazzi et al. 2014) (Figure 5.1) in accordance with Affymetrix's default best practice guidelines (using "Axiom genotyping solution data analysis guide" and "Best practice supplement to Axiom genotyping solution data analysis user guide"). Poor quality samples were first identified using a single sample metric, Dish-QC (DQC), followed by sample QC call test rate.

The DQC test is based on intensities of probe sequences for non-polymorphic genome locations. The default DQC threshold for Axiom arrays is 0.82. Twenty five individuals were excluded on the basis of low DQC. To achieve the highest genotyping performance, as not all poor quality samples are detectable by DQC, an additional step was applied (sample QC) whereby samples passing DQC are genotyped using a subset of probe sets, and those with a call rate of <0.97 excluded. No samples failed this test.

Two further tests for plate quality control were then performed. For Axiom genotyping projects samples are processed together on 96- or 384- array plates (in this case 96- array plates). All samples on plates that fail these tests are excluded from further analysis. The metrics used are plate pass rate and average QC call rate (details in Axiom genotyping

solution data analysis guide), with thresholds of 0.95 and 0.97 respectively. Two plate files failed to pass these quality control steps, eliminating 181 (94, 87) individuals. In total 1804 individuals passed the AffyPipe quality control steps.

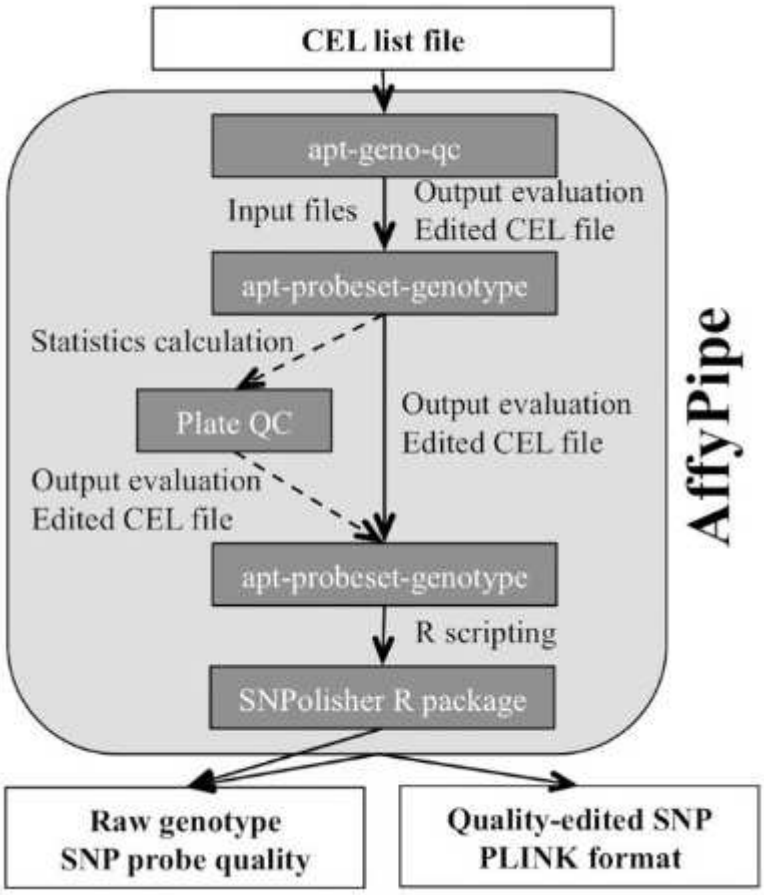


Figure 5.1. Schematic view of the Affymetrix genotyping 'AffyPipe' protocol for standard and 'best practice' (indicated by the dashed line) workflows (from (Nicolazzi et al. 2014))

The R package SNPolisher was then installed in AffyPipe to calculate QC metrics for each SNP. Used for post-processing the results of Axiom arrays, SNPolisher evaluates the quality of marker signals and classifies SNP probes into classes (Nicolazzi et al. 2014). Of 610,970 probes read from the remaining samples, 494,705 SNPs were kept. Finally all best probeset SNPs were converted into genotypes of PLINK ACGT format.

Of the 1804 individuals for which genotypes were successfully called, 861 were from Wytham Woods, UK. Only 42 individuals from the Hoge Veluwe population in the Netherlands were present in the post-processing genotype file, and as this sample size is insufficient for genomic analyses, only the UK great tit population was subject to further downstream analyses. Using the same method presented in Chapter 2, phenotype information was extracted for those females for which breeding events have been recorded. The resultant phenotype file consisted of 629 breeding events for 385 females.

5.3.2. GWAS analysis

Before running a genome-wide association analysis (GWAS), to account for repeated measures (i.e. where more than one breeding event has been recorded for a single female) a 'phenotype' in the form of a residual value obtained by running a version of the animal model in ASReml (Gilmour et al. 2002) was generated for each female (Chapters 2 & 3). As a different subset of females was used in this analysis compared to Chapter 2, tests of significance for fixed and random effects contributing to variation in laying date were re-run in R ver3.0 (Team 2013) and ASReml ver30 (Gilmour et al. 1996) respectively (methods, Chapters 2 & 3).

Linear modelling determined both female age and altitude to be significant ($p < 0.01$) and were fitted as fixed effects (Appendix 5.1). Year and area of the wood were statistically significant and fitted as random effects together with female permanent environment effect (Table 5.1). The full mixed effects model with statistically significant fixed (blue) and random (purple) terms was then;

Layd = age + altitude + femalePE + area + year

Table 5.1. Statistical testing of random effect terms included in the final versions of the mixed effects model for the UK great tit population using a high density SNP chip, using logarithm of the odds (LOD) scores obtained from ASReml, and subsequent likelihood ratio test (LRT) of significance. Variance components statistics for the final polygenic models are highlighted.

Term	**LRT**	**p value**	**Component**	**Comp/SE**	**D.O.F**
Year	351.64	<0.01	51.69	2.83	1
Area	14.12	<0.01	1.27	1.43	1
FemalePE	5.8	0.02	2.40	1.99	1
Residual variance			19.88	12.73	
Total variance			75.24		626

Genome-wide association analysis was performed using GenABEL (Aulchenko, Ripke, et al. 2007) implemented in R (version 3.0.0 (2013)). Quality control using default parameters was performed in GenABEL to remove markers and individuals of poor genotyping quality (Chapter 1). In total 48672 SNPs and 18 individuals were removed for failing to pass individual and marker call rate, Hardy-Weinberg equilibrium and minor allele frequency thresholds (individual call rate: 0.50, SNP call rate: 0.70, Hardy-Weinberg equilibrium: $P < 1e\text{-}5$, minor allele frequency: 0.05). A total 446,033 SNPs and 367 individuals were retained for association analysis.

Typically GWAS analyses assume phenotypic data comes from unrelated individuals from a single population (Svishcheva et al. 2012) but even with carefully constructed datasets there exists the potential for relatedness between individuals and population stratification. Population stratification can be problematic in that it is known to increase the incidence of Type I error (Aulchenko, de Koning, et al. 2007; Svishcheva et al. 2012). Tests implemented

in GenABEL correct for population stratification by fitting a relationship matrix based on genome-wide marker relatedness to account for the presence of related individuals (Chen & Abecasis 2007). Lambda before correction for population stratification was estimated as 1.04 (estimates of lambda >1 indicate inflation) (Appendix 5.2a).

A polygenic model applied using function 'mmscore' (Chen & Abecasis 2007) in GenABEL was used to test for associations between SNP markers and the trait. Interpretation of the association scan to identify regions of significance was determined by Bonferroni correction for multiple testing to set the genome-wide significance threshold and a permutation test subsequently to substantiate the results of the association scan and set the genome-wide significance threshold.

5.3.3. Chromosome partitioning analysis

A second statistical analysis was then employed to partition variance between individual chromosomes. Consistent with former analyses (Chapter 4) the smaller chromosomes (21-28, LGE22) were grouped to form a micro chromosome set. Results were then obtained for 22 chromosomes (1-15, 17-20, 1A and 4A, also including the micro chromosome set). In contrast to Chapter 4 which utilised the chromosome partitioning method of Robinson et al. (2013), the construction of genetic relatedness matrices and subsequent partitioning of genetic variation was performed using the programme GCTA (Yang, Lee, et al. 2011; Yang, S. H. Lee, et al. 2013).

The objective of this method is to estimate genetic variation captured by all SNPs (Yang, Lee, et al. 2011). GCTA uses information captured by markers to estimate the genomic relationships between pairs of individuals, and performs association analyses to correlate genome-wide SNP similarity with phenotypic similarity (Yang et al. 2010; Yang, Lee, et al. 2011; Deary et al. 2012; Yang, S. H. Lee, et al. 2013). Two sets of genetic relatedness

matrices (GRMs) were generated for each chromosome; one multi chromosome GRM consisting of marker information for all chromosomes with the exception of the focal chromosome, and one constructed from the markers on the focal chromosome. Partitioning variance across the autosomes was then performed using a familiar restricted maximum likelihood (REML) approach (Yang, Lee, et al. 2011). A comparison between the two sets of genetic relatedness matrices in the form of a likelihood ratio test (where one model fits both GRMs and the other omits the focal chromosome's GRM (Chapter 4)) was the performed for each chromosome to ascertain whether any of the chromosomes contributed a significance proportion of variance.

Linear modelling was then used to test for relationships between chromosome size (Mbp) and number of genes, to ascertain whether larger chromosomes explain a greater proportion of the additive genetic variance (V_A). If variance explained is proportional to these parameters, this would provide evidence for a polygenic mode of inheritance (Yang, Manolio, et al. 2011). The null hypothesis predicts no relationship between the amount of variance explained by chromosomes and either of the size parameters.

5.4. Results

5.4.1. GWAS

A heritability estimate of 0.27 was obtained from GenABEL. Correction for population stratification was successful with an estimate of lambda of 1.00 (± 1.11) (Appendix 5.2b). The association scan detected no genome-wide significant SNPs with 24100 passing a nominally significant threshold of $p<0.05$ (Figure 1). Assuming a false discovery rate of 5%, for 495705 markers we would expect 24735 markers to achieve nominal significance by chance. The results of the permutation test where an empirical distribution of test statistics was derived from 1000 permutations (Bush & Moore 2012) validated the absence of genome-wide significant SNP markers giving the SNP with the highest association, located on chromosome 3, a p-value of 0.13 (Table 5.2).

Table 5.2. The top 10 associations from GWAS analysis using genetic information from a 500K SNP chip, following a permutation test of 1000 iterations to test for significance. Number of females for which complete phenotypic information was available (N), chromosome (Chr) and position (cM). Effect sizes (effB) and their standard errors (se_effB) of allelic substitutions with 1df. Nominal p value (pre-permutation Pc1df) and genome-wide p value from permutations tests before and after lambda correction (P1df, Pc1df). Genotype effect sizes relative to AA and test statistics from a model fitting genotypes as a three level factor described by effAB and effBB.

Marker	Chr	Position	A1	A2	N	effB	se_effB	chi2.1df	Pre-permutation Pc1df	P1df	Pc1df	effAB	effBB
Affx-100103406	3	17523790	T	C	364	0.1673	0.0341	24.0337	7.07E-07	0.231	0.14	0.1352	0.3532
Affx-100794769	7	57838616	G	C	359	-0.2346	0.0526	19.8640	4.88E-06	0.921	0.795	-0.2519	-0.4103
Affx-100991634	4A	1.07E+08	C	G	364	0.1583	0.0357	19.7019	1.07E-05	0.936	0.818	0.1334	0.3259
Affx-100497585	3	31092231	A	G	363	0.1538	0.0347	19.6655	7.74E-06	0.94	0.822	0.1795	0.3048
Affx-100429736	3	14007486	A	T	359	0.1870	0.0427	19.2119	6.70E-06	0.971	0.895	0.1345	0.5393
Affx-100461164	3	67660605	C	T	366	0.1649	0.0385	18.3249	1.35E-05	0.996	0.973	0.1641	0.3310
Affx-100343672	20	58991237	G	A	366	-0.1685	0.0395	18.1699	1.68E-05	0.997	0.975	-0.2098	-0.2482
Affx-100775219	3	51021709	T	G	367	0.2264	0.0544	17.3217	1.90E-05	1	0.998	0.2653	0.2670
Affx-100256884	7	1436465	C	G	365	-0.2238	0.0541	17.1377	2.47E-05	1	0.998	-0.2263	-0.4321
Affx-100670379	3	31092283	G	A	363	0.2368	0.0573	17.1077	2.68E-05	1	0.998	0.2647	0.3283

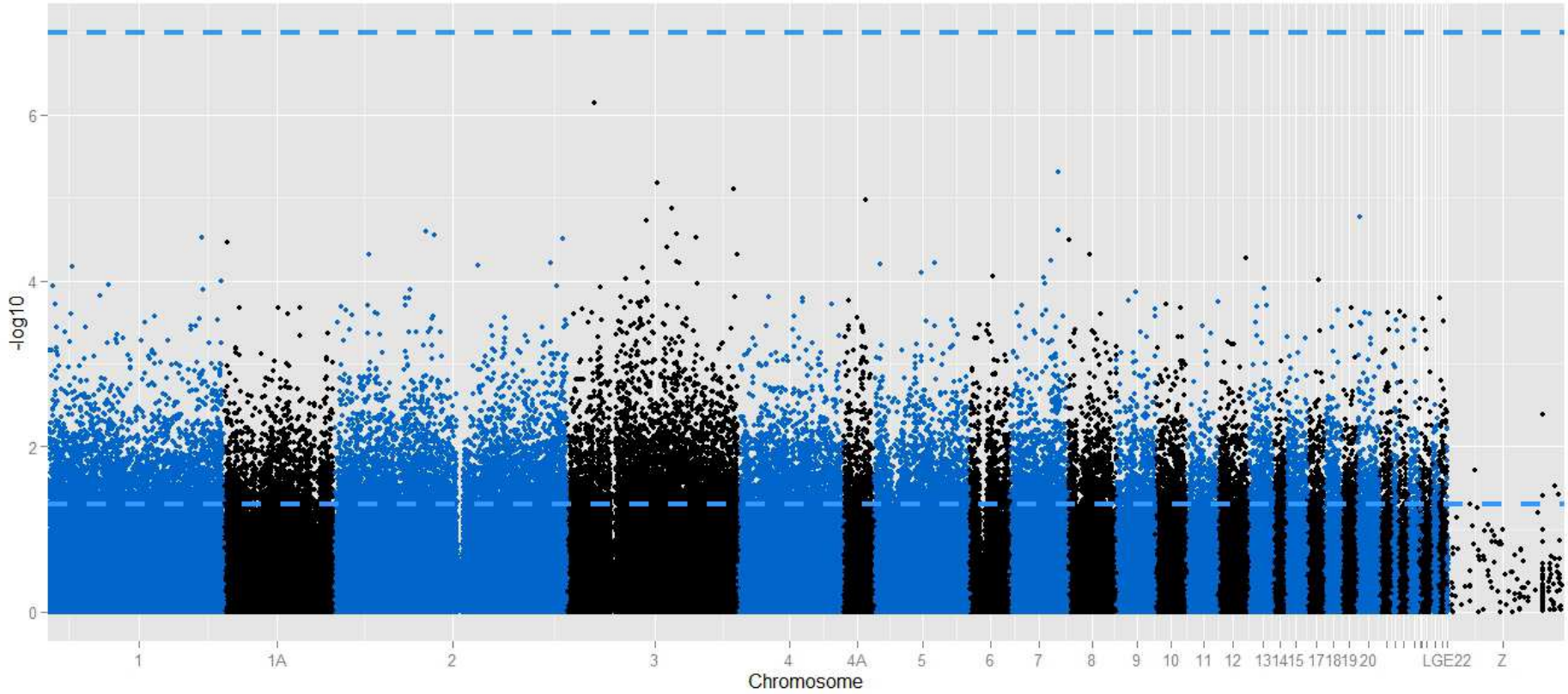


Figure 5.2. GWAS scan for regions associated with variation in lay date. The blue dashed lines represent significance thresholds of nominal (1.30) and genome-wide (7.00) significance

5.4.2. Chromosome partitioning

Likelihood ratio testing revealed V_A was distributed over 10 chromosomes of which chromosome 3 explained the most variation and was the only chromosome to explain a statistically significant proportion of variation overall ($p < 0.05$) (Table 5.3). No significant relationships were detected either between the amount of variance explained by each chromosome in relation to chromosome size (Mbp) or the number of genes ($p > 0.1$ all cases) (Figure 2).

Table 5.3. Likelihood ratio test (LRT) results for chromosome partitioning analysis performed on the Wytham Woods (UK) population using genotype information from a 500K SNP chip. Statistically significant results are highlighted in bold.

Chromosome / set	No. SNPs	Length (cM)	Number genes	Size (Mbp)	UK	
					h^2 (+SE)	LRT1
1	68300	139.9	1254	114.1	0	0.00
1A	36660	93.6	972	71.4	0	0.00
2	85089	139.7	1450	150.3	0.07 (0.15)	0.28
3	62097	114.9	1290	111.6	0.3 (0.15)	**7.87**
4	39833	97.6	811	68	0	0.00
4A	7969	59.4	39	19.9	0.12 (0.13)	0.85
5	29395	98.6	998	61.9	0	0.00
6	10116	78	596	33.3	0	0.00
7	18014	72.6	562	37.7	0.09 (0.12)	0.75
8	12611	53.8	575	31.3	0.17 (0.13)	1.50
9	10873	54.2	497	25.1	0	0.00
10	9600	50.5	444	20.2	0	0.00
11	9024	58.2	397	20.3	0	0.00
12	9659	51.9	369	20.5	0	0.00
13	7882	40.9	379	16.5	0.08 (0.11)	0.64
14	3374	49.2	426	16.2	0	0.00
15	5445	49.1	381	13.8	0.13 (0.11)	1.67
17	3512	45.4	336	10.5	0.04 (0.11)	0.16
18	4404	49.9	334	11.6	0.08 (0.10)	0.89
19	3565	49.4	348	9.9	0	0.00
20	6212	49.4	356	14.7	0.12 (0.13)	0.80
micros	11271	411	1808	44.2	0	0.00

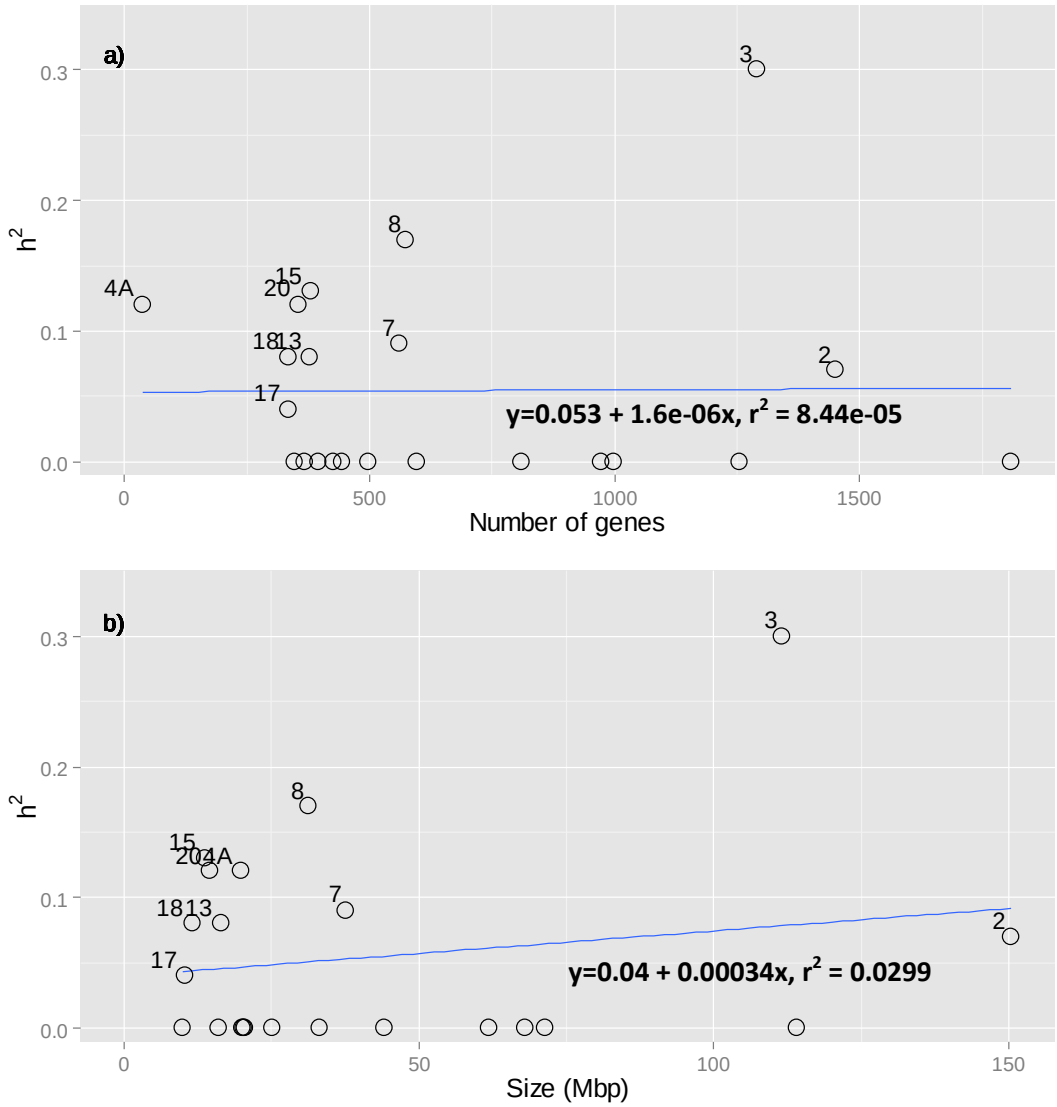


Figure 5.3. Results of the chromosome partitioning analysis where the amount of variance explained by each chromosome (labelled directly on the figure) is plotted relative to (a) number of genes (p = 1) and (b) size (Mbp) (p = 0.44).

5.5. Discussion

Increasing the density of markers mapped on the great tit genome did not result in the detection of any genome-wide significant genetic variants associated with variation in laying date. However, the position of the SNP with the highest association on chromosome 3 is in accordance with the findings of the chromosome partitioning analysis which found chromosome 3 the only chromosome to explain a significant amount of variation. This is as expected, based on the fact both analyses use the same markers and individuals. Whether there is a true causal variant in this region of chromosome 3 remains unresolved.

There is no consistency between the analyses performed on the Wytham, Dutch and combined populations and those obtained for Wytham using the new chip. None of the association analyses on the other datasets identified significant variants, or consistently found that the same locations produced the strongest associations. Markers on chromosomes 8, 2 and 14 represented the top associations for the UK, NL and combined datasets respectively (Chapters 2&3) and (different) markers on chromosome 3 appeared in the top 10 associations for both the UK (Chapter 2) and NL datasets (Chapter 3). However permutation tests of the significance of associations drawn from all datasets suggested that none of the association scores were stronger than what would be expected to occur by chance.

The results of chromosome partitioning analyses between datasets also failed to consistently show the same chromosome(s) contributing the most variance. Chromosome 3 was not found to contribute a significant amount of variance using the other datasets, which identified chromosomes 6, 13 and 12 from the UK, NL and combined datasets respectively (Chapter 4). The only consistency between analyses on different datasets came from linear modelling, when the amount of variance attributed to chromosomes was plotted relative to their size and gene content.

Identifying whether quantitative traits adhere to a polygenic model, or whether in fact the genetic component of variation is composed of few genes of large effect is one of the fundamental aims of evolutionary biology (Charmantier et al. 2014). Having data from multiple populations of the same species, enabling replication of analyses and a direct comparison of results, is a useful tool when attempting to describe the genetic architecture underlying trait variation. If a pattern of many genes of small effect is observed between populations evidence of a polygenic architecture may be substantiated.

Contrary to theory for assumed polygenic traits, which would predict the amount of variance explained by chromosomes should be proportional to their size (Yang, Manolio, et al. 2011), results using three previous three datasets (UK, NL, NLUK (Chapter 4)) found no significant associations between variance explained by chromosomes and their size (Mbp) or number of genes. The only exception was a positive association with number of genes using the NLUK dataset, though this was largely driven by the cumulative size of the micro chromosome set (Chapter 4). The results presented here using the high density chip similarly found no statistically significant relationships between size / number of genes and the amount of variance explained in the UK population. These results suggest either a lack of power in the dataset to accurately partition variance, as a consequence of low additive genetic variance or small sample size, or a departure from the polygenic model laying date is assumed to have.

Human genetic studies on traits such as height, BMI and certain medical conditions have found strong positive correlations between the amount of variance explained by chromosomes relative to chromosome length (Yang, Manolio, et al. 2011; Visscher et al. 2007; Lee et al. 2012) and gene content (Yang, Manolio, et al. 2011). However examples of large effect loci have been described for complex quantitative traits in both domestic and wild organisms (Rich 1990; Nezer et al. 1999; Colosimo et al. 2004; Bernier et al. 2007;

Venuprasad et al. 2009; Johnston et al. 2010; Tarka et al. 2010; Yang, Manolio, et al. 2011). Variation in armour plate pattern and number in multiple populations of threespine sticklebacks (Colosimo et al. 2004), horn size and length in soay sheep (Johnston et al. 2010) and wing length in great reed warblers (Tarka et al. 2010) were attributed to large effect single loci. The detection of several large effect QTL in wild animal populations may be a consequence of the sample sizes used in analyses.

Sample size is a key determinant not only of the ability to detect QTL but also influences QTL effect size (Beavis 1994; Slate 2013). With an observed correlation between sample and QTL effect sizes (Slate 2013) many large effect QTL may in fact be the result of small sample size masking the true genetic architecture. Utilising the first great tit SNP chip research into the genetic architecture of clutch size and egg mass found significant positive correlations between chromosome size (Mbp) and both traits (Robinson et al. 2013; Santure et al. 2013) suggestive of a polygenic basis for phenotypic variation.

Whilst the patterns observed between datasets in this thesis suggest variants of small effect, the absence of positive correlations between the variance explained by chromosomes in relation to their size is an unexpected departure from the pattern that would be expected from a highly polygenic trait. The lack of consistency between QTL identified by means of different statistical analyses, and between populations and datasets, suggests a lack of power as opposed to an oligogenic architecture may explain the results. A potential factor, and one of the greatest limitations of these analyses, is small sample size. While the phenotypic data available to us is, one of the most extensive available for a wild animal (see Husby et al. 2012; Bérénos et al. 2015 for other examples), the power of genomic analyses is dependent, even proportional, to sample size (Beavis 1994; Slate 2013).

Increased marker density and sample sizes are often cited as being essential for the success of genomic analyses (Risch & Merikangas 1996; Manolio et al. 2009; Poissant et al. 2012; Poissant et al. 2013). Sample size and low heritability, together with many variants of small effect are likely all contributing factors limiting the effectiveness and accuracy of the statistical analyses that have been applied here. One conclusion that may be drawn from the analyses presented in this chapter is that increased statistical power from increased marker density was likely countered by a reduction in power due to a reduced sample size, compromising the results. The power of genomic statistical analyses relies on the availability of both extensive phenotypic and genetic data. Insufficient quantity or quality of just one of these data sources effectively limits the power of statistical tests to identify the variants associated with phenotypic variation.

Chapter 6

Discussion

Advances in both the sophistication and accessibility of genotyping technologies are increasingly providing an avenue for researchers studying a wide range of organisms to dissect the genetic architecture of phenotypic traits. Understanding the amount of genetic variation associated with traits has significant implications for our understanding of how traits may evolve and the potential for organisms to adapt. In this thesis, I took advantage of privileged access to data from two of the longest running individual-based studies of wild animal populations to understand the genetic basis of a behavioural life history trait.

Focusing on the Wytham Woods population of great tits, Chapter 2 described how linkage and association mapping were used to scan for genetic variants associated with variation in laying date. A suggestive peak on chromosome 10 identified by a QTL scan failed to be replicated when the data was subjected to GWAS analysis where the highest associations were located on chromosomes 8, 27 and 3. In Chapter 3, the same methods were used to explore the genetic architecture of laying date in the Dutch Hoge Veluwe population. No genome-wide significant QTL were detected using either method. In fact neither linkage nor association mapping identified any QTL that were significant at the less stringent suggestive threshold.

The datasets for the UK and Dutch great tit populations were then combined, increasing the overall size of the dataset to determine whether more data leads to increased resolution and / or power. Applying the same statistical techniques, suggestive QTL peaks were identified on chromosomes 10 and 4A. No genome-wide significant regions were identified and, mirroring the results of each population analysed independently, the highest GWAS associations did not co-localise with the suggestive QTL peaks.

In Chapter 4 a more recently developed statistical technique, genome or chromosome partitioning was employed to partition the genetic variance associated with variation in laying date between autosomal blocks. This approach was applied to each population and

the combined dataset. There was little evidence for variation in laying date in the Wytham population being attributed to any specific region, with the vast majority of chromosomes explaining no variation. However Chromosome 6 explained a significant amount of the variance and significantly more than would be expected given its size, but still its contribution represented just 2% (h^2 0.02) of the total phenotypic variance. For the Netherlands population two chromosomes (4A and 13) explained a significant amount of the variance with chromosome 13 explaining more than would be expected. One chromosome (12) explained a significant proportion of variation in the combined dataset.

Using the results of the chromosome partitioning analysis, linear modelling was then used to elucidate the relationships between the amount of variance explained by each chromosome relative to the parameters of chromosome size (Mbp) and number of genes. For each of the three datasets no significant relationships were detected for variance contributed by each chromosome relative to either of these parameters.

In Chapter 5 genotype information was extracted from the most recently developed and much larger SNP chip and used in association and chromosome partitioning analyses. The hope was that the increased marker density would increase the power of statistical tests to detect regions harbouring genetic variants associated with laying date, and to map them with a greater degree of precision. Evidence for a SNP marker approaching the genome-wide significant threshold was detected on chromosome 3, and was substantiated by a subsequent chromosome partitioning analysis where the same chromosome explained a significant amount of variance. All correlations between the additive genetic variance explained by a chromosome and its size and number of genes were positive, but again were non-significant.

Variation in laying date has a significant genetic component

Analyses on the three datasets using genomic information from the 8K SNP chip, and on the Wytham Woods dataset using the 500K SNP chip supported the findings of several other studies (van Noordwijk et al. 1981; van der Jeugd & McCleery 2002; McCleery et al. 2004; Nussey et al. 2005; Gienapp et al. 2006; Garant et al. 2008; Liedvogel et al. 2012), that laying date is heritable with a small but significant genetic component contributing to variation. The lower density SNP chip gave low heritability estimates for Wytham of 0.05 (± 0.03) (Chapter 2) and 0.00 (Chapter 4) which were consistent with the previously most recent pedigree based estimate of 0.03 (± 0.01) (Liedvogel et al. 2012).

The Hoge Veluwe heritability estimates were also consistent with estimates published in the literature, with h^2 of 0.15 (± 0.06) (Chapter 3) and 0.17 (± 0.03) (Chapter 4) closely matching a previously published estimate of 0.17 (Gienapp et al. 2006). The estimate calculated for the Wytham Woods population, using GenABEL and data from using the new SNP chip, was 0.27 and is the largest calculated for the species, close to an earlier estimate of 0.24 before accounting for spatial autocorrelation (van der Jeugd & McCleery 2002).

The genetic architecture of laying date does not adhere to patterns commonly associated with a polygenic mode of inheritance

An absence of genome-wide significant QTL, no more significant SNPs passing a suggestive threshold than would be expected by chance and a visual interpretation of association scans showing a fairly homogenous distribution across the genome point to a polygenic genetic architecture. However the absence of significant or positive correlations identified from chromosome partitioning analyses on any of the datasets is contrary to the expected trends that would be exhibited by a polygenic trait.

The alternate scenario would be an oligogenic architecture, with few genes of large effect contributing to phenotypic variation, but evidence for this is also scarce. A SNP on chromosome 3 close to genome-wide significance was deemed non-significant following permutation tests (Chapter 5), and on further inspection the marker is not in a region currently known to harbour genes associated with seasonal timing (Veronika Laine, pers comm).

The genetic architecture of complex traits in wild organisms has been observed to vary substantially. Major loci explaining a large proportion of additive genetic variation associated with phenotypic variation have been described in several wild species (Slate et al. 2002; Beraldi, McRae, Gratten, Slate, et al. 2007; Johnston et al. 2010; Poissant et al. 2012; Tarka et al. 2010), though there is some evidence that many of these results are an artefact of small sample sizes causing overestimates of the genetic effect (Beavis 1994; Slate 2013).

In wild avian species polygenicity of phenotypic variation has been demonstrated for traits such as clutch size in the collared flycatcher (Husby et al. 2015) as well as clutch size, egg mass and wing length in the great tit (Santure et al. 2013; Robinson et al. 2013). Based on the evidence accumulated in this thesis it is difficult to ascribe laying date an oligogenic or polygenic architecture. It remains highly likely that laying date is polygenic, influenced by many genes of small effect, but that low heritability and low statistical power hindered the ability of statistical tests to identify QTL or accurately partition variance.

Observed breeding time shifts in response to climate change are largely influenced by non-genetic effects

One of the greatest problems that has faced biologists attempting to distinguish between micro-evolutionary and plastic responses to climate change is a lack of phenotypic /

genomic information and subsequently statistical power to detect a genetic component, leading them to infer micro-evolutionary responses from phenotypic data alone (Parmesan & Yohe 2003; Root et al. 2003). Whilst there is mounting evidence from long-term studies that support a genetic component to several avian life-history traits (Pulido & Berthold 2003; Gienapp et al. 2006; Garant et al. 2008; Liedvogel et al. 2012; Husby et al. 2015; Tarka et al. 2015; Thorley & Lord 2015), there is undoubtedly a consensus that much of the observed variation can be attributed to behaviour within the range of existing reaction norms (Visser et al. 2004; Ghalambor et al. 2007) and phenotypic plasticity (Husby et al. 2015; Tarka et al. 2015; Charmantier & Gienapp 2014; Vedder et al. 2013; Gienapp et al. 2008).

The reaction norm concept describes how phenotypes change across an environmental gradient (Falconer 1990; Weis & Gorman 1990; Gavrilets & Scheiner 1993; Via et al. 1995; Schlichting & Pigliucci 1998). Organisms are adapted to the varying conditions that may be experienced in their environment, as a result of seasonal fluctuations in temperature for example. The range of conditions within which they are able to function (for example the upper and lower temperature limits the species has evolved under and is expected to experience) is defined as the reaction norm, enabling phenotypic variation in response to environmental variation (Weis & Gorman 1990; Gavrilets & Scheiner 1993; Schlichting & Pigliucci 1998). The capacity for organisms to cope with environmental perturbation depends on the extent of existing reaction norms. If environmental conditions fall within the range of the reaction norm adaptive responses may occur without the aid of additive genetic variation. If however conditions exceed these thresholds, organisms would be expected to evolve new mechanisms to cope, or perish.

Of the studies that have attempted to explain variation in laying date there is a consensus that environmental cues are the dominating factor in determining breeding time, with

plasticity the main mechanism enabling responses to changes in climate (Przybylo et al. 2000; Brommer et al. 2005; Nussey et al. 2005; Gienapp et al. 2008; Husby et al. 2010; Thorley & Lord 2015; Perdeck & Cavé 1992; Hakkarainen et al. 1996). The rapidity of the responses suggest plasticity as the most likely mechanism as genetic-based adaptation would require generations before a phenotypic shift could be observed (Charmantier et al. 2008). Even studies with the capacity to test for a genetic based response to advances in laying date attributed most of the variation to plasticity, in response to temperature cues, as opposed to a microevolutionary responses (Sheldon et al. 2003; Gienapp et al. 2006).

Temperature has been suggested to be one of the most important cues dictating behavioural responses in birds, especially with regard to migratory behaviour (Tarka et al. 2015), onset and length of incubation (Ardia et al. 2006; Vedder 2012) and laying date (Winkler et al. 2002; Thorley & Lord 2015; Matthysen et al. 2011; McCleery & Perrins 1998). In great tits temperature is consistently identified as one of the most influential factors determining the onset of laying (Visser et al. 2009; Schaper et al. 2011; Ahola et al. 2009).

Whilst temperature was not explicitly modelled in the analyses presented in this thesis, the year in which the breeding event was recorded was consistently the most important random effect, explaining the highest proportion of variation in laying date (Chapters 2 & 3). This is possibly a reflection of increases in spring temperature and subsequent shift in the timing of emergence of their caterpillar food prey (Charmantier et al. 2008). Altitude was also significant in accordance with research demonstrating that great tits time their breeding events in response to food availability at a very localised scale (Wilkin et al. 2007; Hinks et al. 2015).

Limitations

One of the major limitations of molecular quantitative genetic analyses is that the small effect sizes typically attributed to each gene make it extremely difficult for statistical methods to detect variants associated with the trait. Power analyses performed on the Wytham great tit dataset have concluded that even for traits of moderate heritability / QTL of moderate effect size, the power to detect variants in this dataset is very low (Santure et al. 2013, Poissant pers comm).

Given the complexity of laying date being determined by many factors, with a much smaller genetic component than traits such as clutch size, the inability of statistical methods to find significant associations or accurately partition variation between autosomal blocks in this case is a consequence of the inherent lack of power in the dataset, which is smaller and consists of fewer individuals than the vast majority of datasets associated with similar analyses on humans and domestic animals (Meyer 1989; Dekkers 2012; Visscher et al. 2007).

This thesis attempted to identify any of the numerous small effect genes underlying laying date. Given the ability to detect QTL relies on the magnitude of effect size (Lynch & Walsh 1998; Hill 2012; Korte & Farlow 2013; Bush & Moore 2012), and for association analyses the amount of LD between markers and causal variants (Pritchard & Przeworski 2001; Hill 2012), the small genetic component and size of the dataset ultimately limited the power of statistical tests detect the variants associated with the focal trait. Statistical methods, particularly GWAS, typically struggle to identify variants of small effect size (Korte & Farlow 2013), which is a major issue when attempting to analyse highly polygenic traits, and is exacerbated by small sample sizes.

Sample size has been demonstrated to inflate effects sizes of variants associated with phenotypic variation and affect the precision with which genetic variants can be mapped (Beavis 1994; Slate 2013). The amount of variation explained by QTL is found to diminish with increasing sample size, indicating that small sample sizes result in upwardly biased effect sizes (Slate 2013). Linked chromosomal regions tend to be inherited together, only becoming separated by rare recombination events, which means that the more progeny and generations contained within a dataset the higher the probability that the sample will include rare recombinants which will help us to interpret their individual effects and increase the precision of QTL estimation (Beavis 1994).

Such is the impact of sample size on the estimation and identification of quantitative genetic variance components that Beavis (1994) cited it as the singular most important factor in our ability to detect and describe the genetic component of variation of phenotypic traits, even above increased computational power, improved statistical techniques and higher marker density (Beavis 1994; Kearsey & Farquhar 1998).

Implications & future directions

An understanding of the genetic architecture underlying complex traits may ultimately be used to measure the capacity of animal populations to respond to changes in the environment. Current theories suggest the reaction norms that enable organisms a degree of flexibility in response to environmental perturbation may enable populations to adapt in the short-term with individuals capable of modifying their behaviour (Yeh & Price 2004). Long-term irreversible changes may have far more serious consequences for population persistence as traits which have evolved to provide specific advantages become maladaptive (Visser et al. 2004; Ghalambor et al. 2007; Gienapp, Lof, et al. 2013).

Currently it appears breeding time birds is occurring within the existing reaction norms and as such plasticity is sufficient, with micro-evolution absent or too small to detect (Gienapp et al. 2006; Sheldon et al. 2003). With persistent changes in environmental conditions, selection may act on the available genetic component of variation resulting in a significant shift to earlier laying by means of micro-evolution. Yet heritability and selection for earlier laying may not culminate in the patterns we would predict, given the ecology and demography of the population.

Research has found that while mismatches may have a strong effect on relative individual fitness, the overall effects on mean demographic rates are negligible as the fitness losses encountered by late breeding females in the form of fewer recruits are offset by the benefits accrued by individuals and offspring of early breeding females as a result of reduced competition (Reed, Grotan, et al. 2013; Reed, Jenouvrier, et al. 2013). Great tits seem amply capable of modifying their behaviour to localised environmental conditions. The effect of increased temperatures and climate change induced mismatches on the fitness of populations and the role of a genetic component of variation will only be resolved by taking advantage of annually expanding datasets to run increasingly powerful statistical models.

The power of genomic statistical analyses relies heavily on the size of the phenotypic dataset, with the number of cases and generations enabling statistical methods to more accurately detect genetic variants. Continued long-term monitoring is essential for advancing our understanding of the genetic architecture of complex traits in wild vertebrates by providing ever more detailed pedigrees and enabling association studies to utilise complex genomic information and patterns of LD.

Increasingly sophisticated statistical methods can also be used to elucidate the mechanisms underlying phenotypic variation. Given that loci contributing to phenotypic variation may

contain multiple different alleles, regional heritability mapping uses genome-wide SNP data to estimate distant and previously uncaptured familial relationships, increasing statistical power, to estimate whole genome and regional genome heritability (Nagamine et al. 2012; Uemoto et al. 2013; Riggio & Pong-Wong 2014). Advances in our computing capability and the development of programmes and methods such as regional heritability mapping (Uemoto et al. 2013; Riggio & Pong-Wong 2014; Riggio et al. 2013) will also enable future studies to make the best use of the phenotypic datasets that exist to detect genetic variants.

The rapidly advancing field of epigenetics represents a future avenue of research for studies attempting to elucidate the mechanisms underlying behavioural and morphological variation. Observations made in rodents that environmentally-induced changes in gene expression manifest phenotypically in subsequent generations (Morgan et al. 1999) have been attributed to a heritable pattern of DNA methylation that occurs without modification of the DNA sequence (Holliday 2006; Bird 2007; Bollati & Baccarelli 2010).

Experiments focused on modifying expression of the agouti gene in mice, resulting in yellow fur, obesity and diabetes, found offspring phenotype to be directly related to the phenotype of the mother and not a consequence of postpartum maternal environment (Wolff et al. 1998; Morgan et al. 1999). Subsequent studies confirmed maternal phenotype for traits such as obesity and diabetes could affect the phenotype of the offspring (Reifsnyder 2000; Li et al. 2013) and mice born from dams exposed to high levels of stress predisposed to display behaviours, such as hyperactivity and deficits in social interaction associated with psychiatric disorders (Holmes et al. 2005; Franklin et al. 2010; Matrisciano et al. 2013; Dong et al. 2016).

Given the ubiquity of conditions such as diabetes and mental health disorders in human populations much research has focused on the potential for epigenetics to aid in the

diagnosis, prognosis and treatment of such conditions, notably cancer (Jones & Laird 1999; Egger et al. 2004; Esteller 2008), Alzheimer's (Mastroeni et al. 2011; Rao et al. 2012; Traynor & Renton 2015; Zhu et al. 2015; Watson et al. 2016) and obesity (Campión et al. 2009; Campión et al. 2010; Soubry et al. 2013; Thomas 2015) in humans. However controversy surrounding the extent and accuracy with which phenotypic variation can be assigned an epigenetic mode of inheritance is pervasive in the field of genetics and genomics (Jablonka & Raz 2009; Bollati & Baccarelli 2010; Hughes 2014; Meloni & Testa 2014).

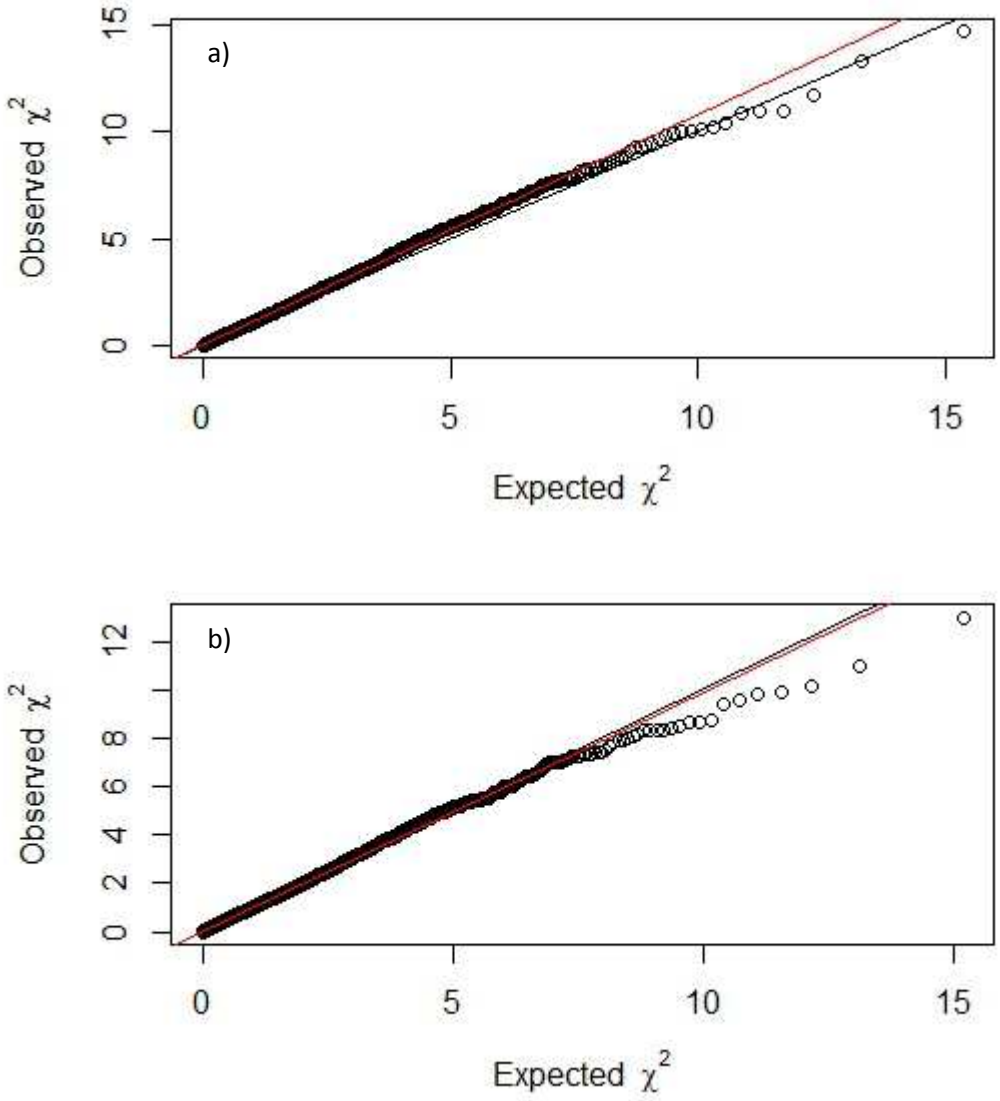
Phenotypic expression has the potential to be affected by a complex interaction of epigenetics and genetics (Jiang et al. 2004; Chibnik et al. 2015). In human cancer research epimutations linked to predispositions for colon cancer were subsequently found to be a consequence of a genetic mutation in a neighbouring gene (Chan et al. 2006; Ligtenberg et al. 2009; van Engeland et al. 2011; Meloni & Testa 2014). There is also difficulty in establishing whether epimutations are truly inherited by offspring as opposed to being triggered after birth (Daxinger & Whitelaw 2012; Meloni & Testa 2014). The continual development of statistical techniques and research in fields such as epigenetics can only aid our efforts and increase our understanding of the mechanisms underlying phenotypic variation.

Appendices

Appendix 2.1 Coefficients and standard errors for fixed effects included in the final animal model, for heritability of laying date in the Wytham Woods great tit population

Factor	Coefficient	SE	p value
age	-3.53	1.02	<0.01
age quadratic	0.44	0.12	<0.01
altitude	0.04	0.01	<0.01

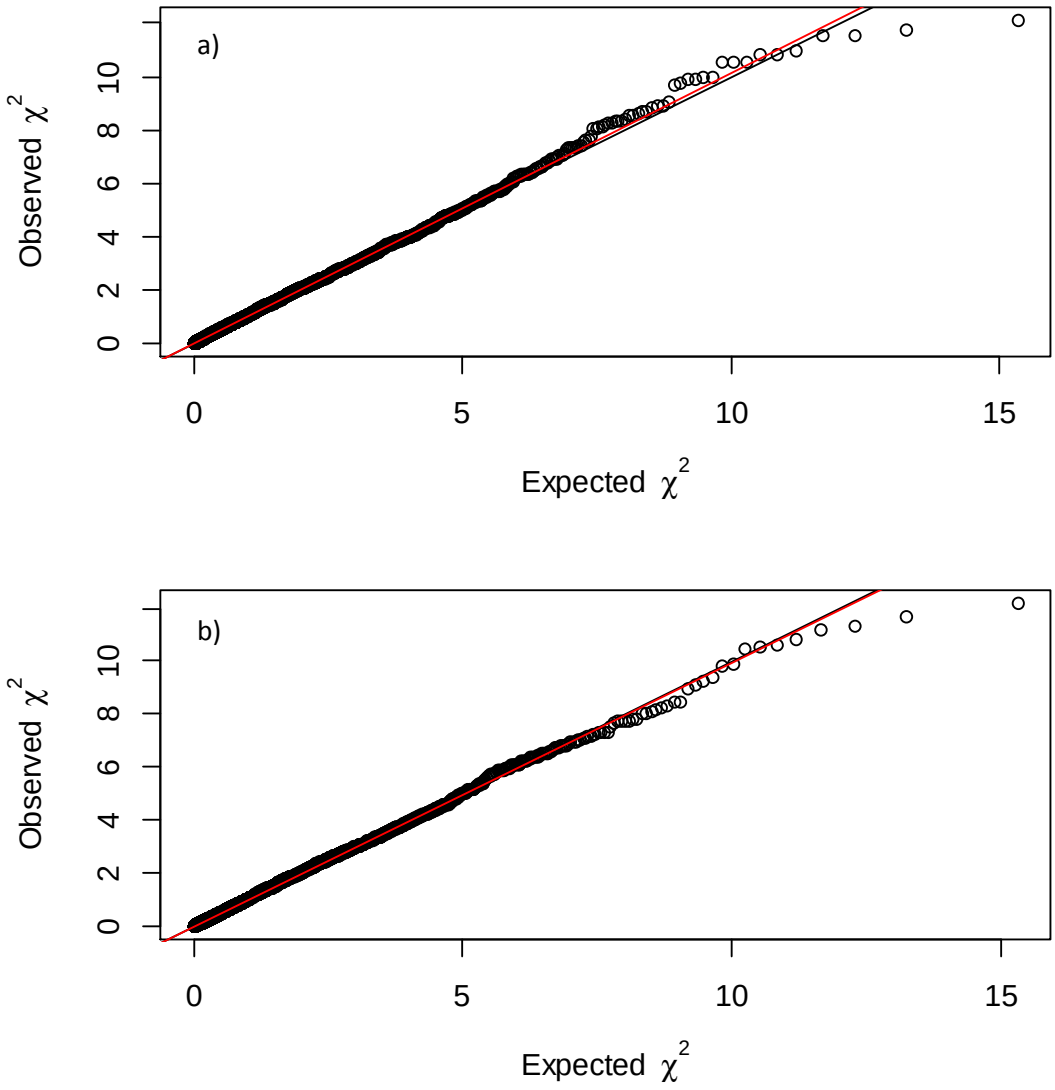
Appendix 2.2 Quantile-Quantile (QQ) plot of an association scan for laying date in the Wytham Woods (UK) great tit population before ($\lambda = 1.08 \pm 4.9 \times 10^{-4}$) (a) and after ($\lambda = 0.98 \pm 6.5 \times 10^{-4}$) (b) correction for population stratification



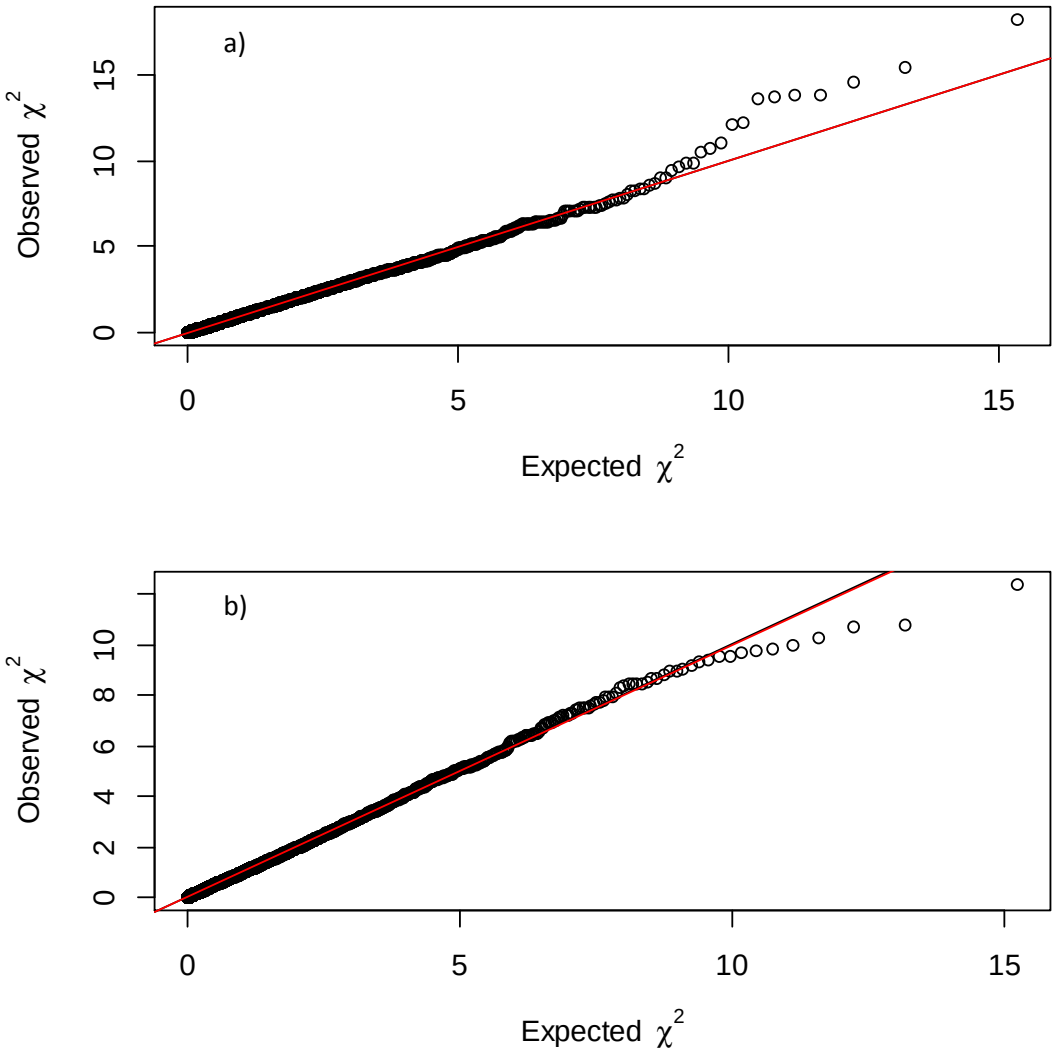
Appendix 3.1. Coefficients and standard errors for fixed effects included in the final animal model for Hoge Veluwe (NL) and combined (NLUK) dataset

NL			
Factor	**Coefficient**	**SE**	**p value**
age	-2.32	0.61	<0.01
age quadratic	0.28	0.10	<0.01
NLUK			
Factor	**Coefficient**	**SE**	**p value**
age	-2.99	0.43	<0.01
age quadratic	0.36	0.07	<0.01
population	2.18	0.30	<0.01

Appendix 3.2. Quantile-Quantile (QQ) plot of an association scan for laying date in the Hoge Veluwe (NL) great tit population before ($\lambda = 1.01 \pm 4.9 \times 10^{-4}$) (a) and after ($\lambda = 0.99 \pm 4.6 \times 10^{-4}$) (b) correction for population stratification



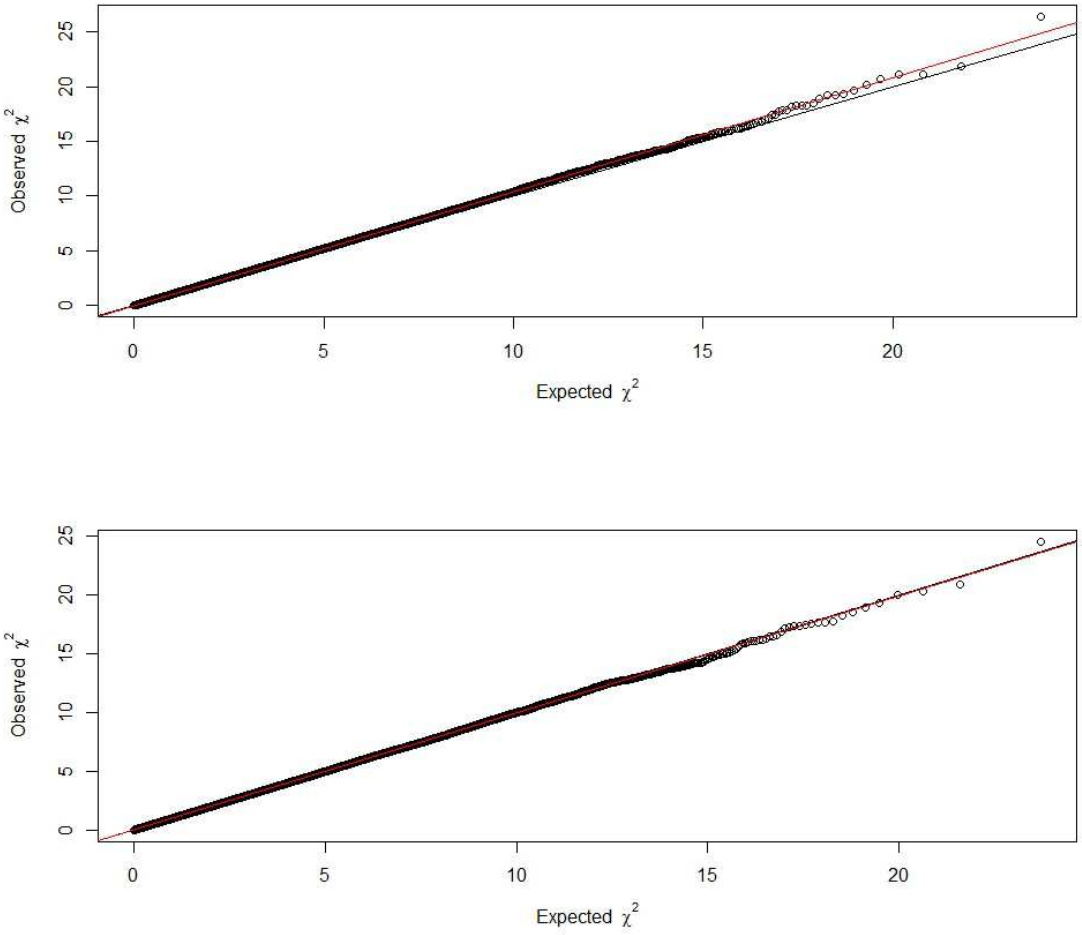
Appendix 3.3. Quantile-Quantile (QQ) plot of an association scan for laying date in the combined Wytham Woods (UK) / Hoge Veluwe (NL) dataset before ($\lambda = 1.00 \pm 8.5 \times 10^{-4}$) (a) and after ($\lambda = 0.99 \pm 5.6 \times 10^{-4}$) (b) correction for population stratification



Appendix 5.1 Coefficients and standard errors for fixed effects included in the final mixed effects model, for heritability of laying date in the Wytham Woods (UK) great tit population

Factor	Coefficient	SE	p value
Age	-0.64	1.17	<0.01
Altitude	0.03	0.01	<0.01

Appendix 5.2 Quantile-Quantile (QQ) plot of an association scan for laying date in the Wytham Woods (UK) great tit population before ($\lambda = 1.04 \pm 4.74 \times 10^{-6}$) (a) and after ($\lambda = 1.00 \pm 5.65 \times 10^{-6}$) (b) correction for population stratification using a 500K SNP chip.



Bibliography

Ahola, M.P., Laaksonen, T., Eeva, T. & Lehikoinen, E., 2007. Climate change can alter competitive relationships between resident and migratory birds. *The Journal of Animal Ecology*, 76(6), pp.1045–52.

Ahola, M.P., Laaksonen, T., Eeva, T. & Lehikoinen, E., 2009. Great tits lay increasingly smaller clutches than selected for: a study of climate- and density-related changes in reproductive traits. *The Journal of Animal Ecology*, 78(6), pp.1298–306.

Ahola, M.P., Laaksonen, T., Eeva, T. & Lehikoinen, E., 2012. Selection on laying date is connected to breeding density in the pied flycatcher. *Oecologia*, 168(3), pp.703–10.

Akesson, M., Bensch, S., Hasselquist, D., Tarka, M. & Hansson, B., 2008. Estimating heritabilities and genetic correlations: comparing the "animal model" with parent-offspring regression using data from a natural population. *PLoS ONE*, 3(3), p.e1739.

Al-Kalaldeh, M., Gibson, J.P., Werf, J.H.J. van der, Gondro, C. & Villalobos, N.L., 2013. Partitioning the genetic variance into genomic and pedigree components for parasite resistance in sheep. In *Proceedings of the Twentieth Conference of the Association for the Advancement of Animal Breeding and Genetics, Translating Science into Action, Napier, New Zealand, 20th-23rd October 2013.* Association for the Advancement of Animal Breeding and Genetics, pp. 412–415.

Ardia, D.R., Cooper, C.B. & Dhondt, A.A., 2006. Warm temperatures lead to early onset of incubation, shorter incubation periods and greater hatching asynchrony in tree swallows Tachycineta bicolor at the extremes of their range. *Journal of Avian Biology*, 37(2), pp.137–142.

Aulchenko, Y.S., de Koning, D.-J. & Haley, C., 2007. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait loci association analysis. *Genetics*, 177(1), pp.577–85.

Aulchenko, Y.S., Ripke, S., Isaacs, A. & van Duijn, C.M., 2007. GenABEL: an R library for genome-wide association analysis. *Bioinformatics (Oxford, England)*, 23(10), pp.1294–6.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z. a, Selker, E.U., Cresko, W. a & Johnson, E. a, 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, 3(10), p.e3376.

Beavis, W.D., 1994. The power and deceit of QTL experiments: lessons from comparative qtl studies. In *Proceedings of the Forty-Ninth Annual Corn and Sorghum Industry Research Conference*. pp. 250–266.

Beraldi, D., McRae, A.F., Gratten, J., Pilkington, J.G., Slate, J., Visscher, P.M. & Pemberton, J.M., 2007. Quantitative trait loci (QTL) mapping of resistance to strongyles and coccidia in the free-living Soay sheep (Ovis aries). *International Journal for Parasitology*, 37(1), pp.121–9.

Beraldi, D., McRae, A.F., Gratten, J., Slate, J., Visscher, P.M. & Pemberton, J.M., 2007. Mapping quantitative trait Loci underlying fitness-related traits in a free-living sheep population. *Evolution*, 61(6), pp.1403–16.

Bérénos, C., Ellis, P.A., Pilkington, J.G., Lee, S.H., Gratten, J. & Pemberton, J.M., 2015. Heterogeneity of genetic architecture of body size traits in a free-living population. *Molecular Ecology*, 24(8), pp.1810–30.

Bérénos, C., Ellis, P.A., Pilkington, J.G. & Pemberton, J.M., 2014. Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. *Molecular Ecology*, pp.3434–3451.

Bergmann, F., 1999. Long-term increase in numbers of early-fledged reed warblers (Acrocephalus scirpaceus) at Lake Constance (Southern Germany). *Journal Für Ornithologie*, 140(1), pp.81–86.

Bernier, J., Kumar, A., Ramaiah, V., Spaner, D. & Atlin, G., 2007. A Large-Effect QTL for Grain Yield under Reproductive-Stage Drought Stress in Upland Rice. *Crop Science*, 47(2), p.507.

Bird, A., 2007. Perceptions of epigenetics. *Nature*, 447(7143), pp.396–8.

BirdLife International, 2015. Species factsheet: Parus major. *BirdLife International*. Available at: http://www.birdlife.org [Accessed September 22, 2015].

Bland, J.M. & Altman, D.G., 1995. Statistics notes: Multiple significance tests: the Bonferroni method. *BMJ*, 310(6973), pp.170–170.

Blouin, M.S., 2003. DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology & Evolution*, 18(10), pp.503–511.

Bollati, V. & Baccarelli, A., 2010. Environmental epigenetics. *Heredity*, 105(1), pp.105–12.

Bolormaa, S., Hayes, B.J., Savin, K., Hawken, R., Barendse, W., Arthur, P.F., Herd, R.M. & Goddard, M.E., 2011. Genome-wide association studies for feedlot and growth traits in cattle. *Journal of Animal Science*, 89(6), pp.1684–97.

Bonin, A., Taberlet, P., Miaud, C. & Pompanon, F., 2006. Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (Rana temporaria). *Molecular Biology and Evolution*, 23(4), pp.773–83.

Both, C., Artemyev, A. V, Blaauw, B., Cowie, R.J., Dekhuijzen, A.J., Eeva, T., Enemar, A., Gustafsson, L., Ivankina, E. V, Järvinen, A., Metcalfe, N.B., Nyholm, N.E.I., Potti, J., Ravussin, P.-A., Sanz, J.J., Silverin, B., Slater, F.M., Sokolov, L. V, Török, J., Winkel, W., Wright, J., Zang, H. & Visser, M.E., 2004. Large-scale geographical variation confirms that climate change causes birds to lay earlier. *Proceedings. Biological sciences / The Royal Society*, 271(1549), pp.1657–62.

Both, C., Bouwhuis, S., Lessells, C.M. & Visser, M.E., 2006. Climate change and population declines in a long-distance migratory bird. *Nature*, 441(7089), pp.81–3.

Both, C. & Visser, M.E., 2001. Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. *Nature*, 411(6835), pp.296–8.

Bourret, A., Bélisle, M., Pelletier, F. & Garant, D., 2015. Multidimensional environmental influences on timing of breeding in a tree swallow population facing climate change. *Evolutionary Applications*, 8, pp.933–944.

Brito, F. V, Neto, J.B., Sargolzaei, M., Cobuci, J.A. & Schenkel, F.S., 2011. Accuracy of genomic selection in simulated populations mimicking the extent of linkage disequilibrium in beef cattle. *BMC Genetics*, 12(1), p.80.

Brommer, J.E., 2013. Phenotypic plasticity of labile traits in the wild. *Current Zoology*, 59(4), pp.485–505.

Brommer, J.E., Merilä, J., Sheldon, B.C. & Gustafsson, L., 2005. Natural selection and genetic variation for reproductive reaction norms in a wild bird population. *Evolution*, 59(6), pp.1362–71.

Brommer, J.E., Rattiste, K. & Wilson, A.J., 2008. Exploring plasticity in the wild: laying date-temperature reaction norms in the common gull Larus canus. *Proceedings. Biological sciences / The Royal Society*, 275(1635), pp.687–93.

Brumfield, R.T., Beerli, P., Nickerson, D. a. & Edwards, S. V., 2003. The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology & Evolution*, 18(5), pp.249–256.

Buckler, E.S., Holland, J.B., Bradbury, P.J., Acharya, C.B., Brown, P.J., Browne, C., Ersoz, E., Flint-Garcia, S., Garcia, A., Glaubitz, J.C., Goodman, M.M., Harjes, C., Guill, K., Kroon, D.E., Larsson, S., Lepak, N.K., Li, H., Mitchell, S.E., Pressoir, G., Peiffer, J.A., Rosas, M.O., Rocheford, T.R., Romay, M.C., Romero, S., Salvo, S., Sanchez Villeda, H., da Silva, H.S., Sun, Q., Tian, F., Upadyayula, N., Ware, D., Yates, H., Yu, J., Zhang, Z., Kresovich, S. & McMullen, M.D., 2009. The genetic architecture of maize flowering time. *Science*, 325(5941), pp.714–8.

Burger, R. & Lynch, M., 1995. Evolution and extinction in a changing environment: a quantitative-genetic analysis. *Evolution*, 49, pp.151–163.

Burgueño, J., Cadena, A., Crossa, J., Banziger, M., Gilmour, A.R. & Cullis, B., 2000. *User's guide for spatial analysis of field variety trials using ASREML*, Mexico: CIMMYT.

Burthe, S., Butler, A., Searle, K.R., Hall, S.J.G., Thackeray, S.J. & Wanless, S., 2011. Demographic consequences of increased winter births in a large aseasonally breeding mammal (Bos taurus) in response to climate change. *The Journal of Animal Ecology*, 80(6), pp.1134–44.

Buse, A., Dury, S.J., Woodburn, R.J.W., Perrins, C.M. & Good, J.E.G., 1999. Effects of elevated temperature on multi-species interactions: the case of Pedunculate Oak, Winter Moth and Tits. *Functional Ecology*, 13(s1), pp.74–82.

Bush, W.S. & Moore, J.H., 2012. Chapter 11: Genome-wide association studies. *PLoS Computational Biology*, 8(12), p.e1002822.

Calboli, F.C.F., Tozzi, F., Galwey, N.W., Antoniades, A., Mooser, V., Preisig, M., Vollenweider, P., Waterworth, D., Waeber, G., Johnson, M.R., Muglia, P. & Balding,

D.J., 2010. A genome-wide association study of neuroticism in a population-based sample. *PLoS ONE*, 5(7), p.e11504.

Campión, J., Milagro, F. & Martínez, J.A., 2010. Epigenetics and obesity. *Progress in Molecular Biology and Translational Science*, 94, pp.291–347.

Campión, J., Milagro, F.I. & Martínez, J.A., 2009. Individuality and epigenetics in obesity. *Obesity Reviews*, 10(4), pp.383–92.

Canale, C.I. & Henry, P.-Y., 2010. Adaptive phenotypic plasticity and resilience of vertebrates to increasing climate unpredictability. *Climate Research*, 43, pp.135–147.

Caprioli, M., Ambrosini, R., Boncoraglio, G., Gatti, E., Romano, A., Romano, M., Rubolini, D., Gianfranceschi, L. & Saino, N., 2012. Clock gene variation is associated with breeding phenology and maybe under directional selection in the migratory barn swallow. *PLoS ONE*, 7(4), p.e35140.

Cardon, L.R. & Bell, J.I., 2001. Association study designs for complex diseases. *Nature reviews. Genetics*, pp.91–99.

Carneiro, M., Rubin, C.-J., Di Palma, F., Albert, F.W., Alföldi, J., Barrio, A.M., Pielberg, G., Rafati, N., Sayyab, S., Turner-Maier, J., Younis, S., Afonso, S., Aken, B., Alves, J.M., Barrell, D., Bolet, G., Boucher, S., Burbano, H.A., Campos, R., Chang, J.L., Duranthon, V., Fontanesi, L., Garreau, H., Heiman, D., Johnson, J., Mage, R.G., Peng, Z., Queney, G., Rogel-Gaillard, C., Ruffier, M., Searle, S., Villafuerte, R., Xiong, A., Young, S., Forsberg-Nilsson, K., Good, J.M., Lander, E.S., Ferrand, N., Lindblad-Toh, K. & Andersson, L., 2014. Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. *Science*, 345(6200), pp.1074–9.

Chan, T.L., Yuen, S.T., Kong, C.K., Chan, Y.W., Chan, A.S.Y., Ng, W.F., Tsui, W.Y., Lo, M.W.S., Tam, W.Y., Li, V.S.W. & Leung, S.Y., 2006. Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nature genetics*, 38(10), pp.1178–83.

Charmantier, A., Garant, D. & Kruuk, L.E.B. eds., 2014. *Quantitative Genetics in the Wild*, Oxford: Oxford University Press.

Charmantier, A. & Gienapp, P., 2014. Climate change and timing of avian breeding and migration: evolutionary versus plastic changes. *Evolutionary Applications*, 7(1), pp.15–28.

Charmantier, A., McCleery, R.H., Cole, L.R., Perrins, C., Kruuk, L.E.B. & Sheldon, B.C., 2008. Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, 320(5877), pp.800–3.

Chen, W. & Abecasis, G., 2007. Family-based association tests for genome-wide association scans. *The American Journal of Human Genetics*, 81(5), pp.913–26.

Chibnik, L.B., Yu, L., Eaton, M.L., Srivastava1, G., Schneider, J.A., Kellis, M., Bennett, D.A. & Jager, P.L.D., 2015. Alzheimer's loci: epigenetic associations and interaction with genetic factors. *Annals of Clinical and Translational Neurology*, 2(6), pp.636–647.

Christensen, K.A., Brunelli, J.P., Wheeler, P.A. & Thorgaard, G.H., 2014. Antipredator behavior QTL: differences in rainbow trout clonal lines derived from wild and hatchery populations. *Behavior Genetics*, 44(5), pp.535–46.

Churchill, G.A. & Doerge, R.W., 1994. Empirical threshold values for quantitative trait mapping. *Genetics*, 138(3), pp.963–971.

Cliff, A.D. & Ord, J.K., 1981. *Spatial Processes: Models and Applications*, London: Pion Limited.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. & Kingsley, D.M., 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology*, 2(5), p.E109.

Conway, C.J. & Martin, T.E., 2000. Evolution of passerine incubation behavior: influence of food, temperature, and nest predation. *Evolution; International Journal of Organic Evolution*, 54(2), pp.670–85.

Cresswell, W. & McCleery, R., 2003. How great tits maintain synchronization of their hatch date with food supply in response to long-term variability in temperature. *Journal of Animal Ecology*, 72(2), pp.356–366.

Crick, H.Q.P., Dudley, C., Glue, D.E. & Thomson, D.L., 1997. UK birds are laying eggs earlier. *Nature*, 388(6642), p.526.

Crick, H.Q.P. & Sparks, T.H., 1999. Climate change related to egg-laying trends. *Nature*, 399(6735), pp.423–424.

Culina, A., Lachish, S. & Sheldon, B.C., 2015. Evidence of a link between survival and pair fidelity across multiple tit populations. *Journal of Avian Biology*, 46(5), pp.507–515.

D'Alba, L., Monaghan, P. & Nager, R.G., 2010. Advances in laying date and increasing population size suggest positive responses to climate change in Common Eiders Somateria mollissima in Iceland. *Ibis*, 152(1), pp.19–28.

Daetwyler, H.D., Schenkel, F.S., Sargolzaei, M. & Robinson, J.A.B., 2008. A genome scan to detect quantitative trait loci for economically important traits in Holstein cattle using two methods and a dense single nucleotide polymorphism map. *Journal of Dairy Science*, 91(8), pp.3225–36.

Davis, L.K., Yu, D., Keenan, C.L., Gamazon, E.R., Konkashbaev, A.I., Derks, E.M., Neale, B.M., Yang, J., Lee, S.H., Evans, P., Barr, C.L., Bellodi, L., Benarroch, F., Berrio, G.B., Bienvenu, O.J., Bloch, M.H., Blom, R.M., Bruun, R.D., Budman, C.L., Camarena, B., Campbell, D., Cappi, C., Cardona Silgado, J.C., Cath, D.C., Cavallini, M.C., Chavira, D.A., Chouinard, S., Conti, D. V, Cook, E.H., Coric, V., Cullen, B.A., Deforce, D., Delorme, R., Dion, Y., Edlund, C.K., Egberts, K., Falkai, P., Fernandez, T. V, Gallagher, P.J., Garrido, H., Geller, D., Girard, S.L., Grabe, H.J., Grados, M.A., Greenberg, B.D., Gross-Tsur, V., Haddad, S., Heiman, G.A., Hemmings, S.M.J., Hounie, A.G., Illmann, C., Jankovic, J., Jenike, M.A., Kennedy, J.L., King, R.A., Kremeyer, B., Kurlan, R., Lanzagorta, N., Leboyer, M., Leckman, J.F., Lennertz, L., Liu, C., Lochner, C., Lowe, T.L., Macciardi, F., McCracken, J.T., McGrath, L.M., Mesa Restrepo, S.C., Moessner, R., Morgan, J.,

Muller, H., Murphy, D.L., Naarden, A.L., Ochoa, W.C., Ophoff, R.A., Osiecki, L., Pakstis, A.J., Pato, M.T., Pato, C.N., Piacentini, J., Pittenger, C., Pollak, Y., Rauch, S.L., Renner, T.J., Reus, V.I., Richter, M.A., Riddle, M.A., Robertson, M.M., Romero, R., Rosàrio, M.C., Rosenberg, D., Rouleau, G.A., Ruhrmann, S., Ruiz-Linares, A., Sampaio, A.S., Samuels, J., Sandor, P., Sheppard, B., Singer, H.S., Smit, J.H., Stein, D.J., Strengman, E., Tischfield, J.A., Valencia Duarte, A. V, Vallada, H., Van Nieuwerburgh, F., Veenstra-Vanderweele, J., Walitza, S., Wang, Y., Wendland, J.R., Westenberg, H.G.M., Shugart, Y.Y., Miguel, E.C., McMahon, W., Wagner, M., Nicolini, H., Posthuma, D., Hanna, G.L., Heutink, P., Denys, D., Arnold, P.D., Oostra, B.A., Nestadt, G., Freimer, N.B., Pauls, D.L., Wray, N.R., Stewart, S.E., Mathews, C.A., Knowles, J.A., Cox, N.J. & Scharf, J.M., 2013. Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. *PLoS Genetics*, 9(10), p.e1003864.

Dawson, A., 2008. Control of the annual cycle in birds: endocrine constraints and plasticity in response to ecological variability. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1497), pp.1621–33.

Daxinger, L. & Whitelaw, E., 2012. Understanding transgenerational epigenetic inheritance via the gametes in mammals. *Nature reviews. Genetics*, 13(3), pp.153–62.

de Jong, M., Ouyang, J.Q., Da Silva, A., van Grunsven, R.H.A., Kempenaers, B., Visser, M.E. & Spoelstra, K., 2015. Effects of nocturnal illumination on life-history decisions and fitness in two wild songbird species. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 370(1667), p.8.

De'ath, G. & Fabricius, K.E., 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81(11), pp.3178–3192.

Deary, I.J., Yang, J., Davies, G., Harris, S.E., Tenesa, A., Liewald, D., Luciano, M., Lopez, L.M., Gow, A.J., Corley, J., Redmond, P., Fox, H.C., Rowe, S.J., Haggarty, P., McNeill, G., Goddard, M.E., Porteous, D.J., Whalley, L.J., Starr, J.M. & Visscher, P.M., 2012. Genetic contributions to stability and change in intelligence from childhood to old age. *Nature*, 482(7384), pp.212–5.

Dekkers, J.C.M., 2012. Application of genomics tools to animal breeding. *Current Genomics*, 13(3), pp.207–12.

Deviche, P., 1997. Seasonal Reproductive Pattern of White-Winged Crossbills in Interior Alaska. *Journal of Field Ornithology*, 68(4), pp.613–621.

Deviche, P. & Sharp, P.J., 2001. Reproductive endocrinology of a free-living, opportunistically breeding passerine (white-winged crossbill, Loxia leucoptera). *General and Comparative Endocrinology*, 123(3), pp.268–79.

Deviche, P., Small, T., Sharp, P. & Tsutsui, K., 2006. Control of luteinizing hormone and testosterone secretion in a flexibly breeding male passerine, the Rufous-winged Sparrow, Aimophila carpalis. *General and Comparative Endocrinology*, 149(3), pp.226–35.

Dingemanse, N.J., Both, C., Drent, P.J., van Oers, K. & van Noordwijk, A.J., 2002. Repeatability and heritability of exploratory behaviour in great tits from the wild. *Animal Behaviour*, pp.929–938.

Dingemanse, N.J., Bouwman, K.M., van de Pol, M., van Overveld, T., Patrick, S.C., Matthysen, E. & Quinn, J.L., 2012. Variation in personality and behavioural plasticity across four populations of the great tit Parus major. *Journal of Animal Ecology*, pp.116–126.

Do, D.N., Strathe, A.B., Ostersen, T., Jensen, J., Mark, T. & Kadarmideen, H.N., 2013. Genome-wide association study reveals genetic architecture of eating behavior in pigs and its implications for humans obesity by comparative mapping. *PLoS ONE*, 8(8), p.e71509.

Dolenec, Z. & Dolenec, P., 2011. Influence of the local spring warming on the breeding phenology in blackcap (Sylvia atricapilla) in Croatia. *Journal of Environmental Biology*, 32(5), pp.625–7.

Dong, E., Tueting, P., Matrisciano, F., Grayson, D.R. & Guidotti, A., 2016. Behavioral and molecular neuroepigenetic alterations in prenatally stressed mice: relevance for the study of chromatin remodeling properties of antipsychotic drugs. *Translational Psychiatry*, 6, p.e711.

Donnelly, A., Caffarra, A. & O'Neill, B.F., 2011. A review of climate-driven mismatches between interdependent phenophases in terrestrial and aquatic ecosystems. *International Journal of Biometeorology*, 55(6), pp.805–17.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., G. Davies, R., Hirzel, A., Jetz, W., Daniel Kissling, W., Kühn, I., Ohlemüller, R., R. Peres-Neto, P., Reineking, B., Schröder, B., M. Schurr, F. & Wilson, R., 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30(5), pp.609–628.

Drent, P.J., van Oers, K. & van Noordwijk, A.J., 2003. Realized heritability of personalities in the great tit (Parus major). *Proceedings. Biological sciences / The Royal Society*, 270(1510), pp.45–51.

Duerr, R.H., Taylor, K.D., Brant, S.R., Rioux, J.D., Silverberg, M.S., Daly, M.J., Steinhart, A.H., Abraham, C., Regueiro, M., Griffiths, A., Dassopoulos, T., Bitton, A., Yang, H., Targan, S., Datta, L.W., Kistner, E.O., Schumm, L.P., Lee, A.T., Gregersen, P.K., Barmada, M.M., Rotter, J.I., Nicolae, D.L. & Cho, J.H., 2006. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science*, 314(5804), pp.1461–3.

Dunn, P., 2004. Breeding Dates and Reproductive Performance. In A. Møller, W. Fiedler, & L. Yiqi, eds. *Birds and Climate Change*. Advances in Ecological Research. Elsevier, pp. 69–87.

Eeva, A.T., Ryömä, M. & Riihimäki, J., 2011. International Association for Ecology Pollution-Related Changes in Diets of Two Insectivorous Passerines Pollution-related changes in diets of two insectivorous passerines. *Oecologia*, 145(4), pp.629–639.

Egger, G., Liang, G., Aparicio, A. & Jones, P.A., 2004. Epigenetics in human disease and prospects for epigenetic therapy. *Nature*, 429(6990), pp.457–463.

Eichler, E.E., Flint, J., Gibson, G., Kong, A., Leal, S.M., Moore, J.H. & Nadeau, J.H., 2010. Missing heritability and strategies for finding the underlying causes of complex disease. *Nature reviews. Genetics*, 11(6), pp.446–50.

Ellegren, H. & Sheldon, B.C., 2008. Genetic basis of fitness differences in natural populations. *Nature*, 452(7184), pp.169–75.

Ernst, M.D., 2004. Permutation Methods: A Basis for Exact Inference. *Statistical Science*, 19(4), pp.676–685.

Esteller, M., 2008. Epigenetics in cancer. *The New England Journal of Medicine*, 358(11), pp.1148–59.

Estoup, A., Solignac, M., Corbuet, J.M., Goudet, J. & Scholl, A., 1996. Genetic differentiation of continental and island populations of Bombus terrestris (Hymenoptera: Apidae) in Europe. *Molecular Ecology*, 5(1), pp.19–31.

Eyre-Walker, A., 2010. Evolution in health and medicine Sackler colloquium: Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proceedings of the National Academy of Sciences of the United States of America*, 107 Suppl (suppl_1), pp.1752–6.

Falconer, D.S., 1990. Selection in different environments: effects on environmental sensitivity (reaction norm) and on mean performance. *Genetical Research*, 56(1), pp.57–70.

Falconer, D.S. & Mackay, T.F.C., 1996. *Introduction to Quantitative Genetics*, New York: Longman.

Fidler, A.E., van Oers, K., Drent, P.J., Kuhn, S., Mueller, J.C. & Kempenaers, B., 2007. Drd4 gene polymorphisms are associated with personality variation in a passerine bird. *Proceedings of the Royal Society of London B*, pp.1685–1691.

Fisher, R.A., 1918. The Correlation Between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh*, 52, pp.399–433.

Fisher, R.A., 1935. *The Design of Experiments*, Edinburgh: Oliver and Boyd.

Fisher, R.A., 1930. *The Genetical Theory of Natural Selection*, Oxford: Clarendon Press.

Flint, J., 2003. Analysis of quantitative trait loci that influence animal behavior. *Journal of Neurobiology*, 54(1), pp.46–77.

Flint, J. & Mackay, T.F.C., 2009. Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Research*, 19(5), pp.723–33.

Forchhammer, M.C., Post, E. & Stenseth, N.C., 1998. Breeding phenology and climate... *Nature*, 391(6662), pp.29–30.

Fortin, M.J. & Dale, M.R.T., 2005. *Spatial Analysis: A guide for ecologists*, Cambridge: Cambridge University Press.

Fossøy, F., Stokke, B.G., Kåsi, T.K., Dyrset, K., Espmark, Y., Hoset, K.S., Wedege, M.I. & Moksnes, A., 2014. Reproductive success is strongly related to local and regional climate in the Arctic snow bunting (Plectrophenax nivalis). *Polar Biology*, 38(3), pp.393–400.

Franck, P., Reyes, M., Olivares, J. & Sauphanor, B., 2007. Genetic architecture in codling moth populations: comparison between microsatellite and insecticide resistance markers. *Molecular Ecology*, 16(17), pp.3554–64.

Franke, A., McGovern, D.P.B., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R., Anderson, C.A., Bis, J.C., Bumpstead, S., Ellinghaus, D., Festen, E.M., Georges, M., Green, T., Haritunians, T., Jostins, L., Latiano, A., Mathew, C.G., Montgomery, G.W., Prescott, N.J., Raychaudhuri, S., Rotter, J.I., Schumm, P., Sharma, Y., Simms, L.A., Taylor, K.D., Whiteman, D., Wijmenga, C., Baldassano, R.N., Barclay, M., Bayless, T.M., Brand, S., Büning, C., Cohen, A., Colombel, J.-F., Cottone, M., Stronati, L., Denson, T., De Vos, M., D'Inca, R., Dubinsky, M., Edwards, C., Florin, T., Franchimont, D., Gearry, R., Glas, J., Van Gossum, A., Guthery, S.L., Halfvarson, J., Verspaget, H.W., Hugot, J.-P., Karban, A., Laukens, D., Lawrance, I., Lemann, M., Levine, A., Libioulle, C., Louis, E., Mowat, C., Newman, W., Panés, J., Phillips, A., Proctor, D.D., Regueiro, M., Russell, R., Rutgeerts, P., Sanderson, J., Sans, M., Seibold, F., Steinhart, A.H., Stokkers, P.C.F., Torkvist, L., Kullak-Ublick, G., Wilson, D., Walters, T., Targan, S.R., Brant, S.R., Rioux, J.D., D'Amato, M., Weersma, R.K., Kugathasan, S., Griffiths, A.M., Mansfield, J.C., Vermeire, S., Duerr, R.H., Silverberg, M.S., Satsangi, J., Schreiber, S., Cho, J.H., Annese, V., Hakonarson, H., Daly, M.J. & Parkes, M., 2010. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nature Genetics*, 42(12), pp.1118–25.

Franklin, T.B., Russig, H., Weiss, I.C., Gräff, J., Linder, N., Michalon, A., Vizi, S. & Mansuy, I.M., 2010. Epigenetic transmission of the impact of early stress across generations. *Biological psychiatry*, 68(5), pp.408–15.

Funfstuck, T., Arandjelovic, M., Morgan, D.B., Sanz, C., Breuer, T., Stokes, E.J., Reed, P., Olson, S.H., Cameron, K., Ondzie, A., Peeters, M., Kuhl, H.S., Cipolletta, C., Todd, A., Masi, S., Doran-Sheehy, D.M., Bradley, B.J. & Vigilant, L., 2014. The genetic population structure of wild western lowland gorillas (Gorilla gorilla gorilla) living in continuous rain forest. *American Journal of Primatology*, 76(9), pp.868–878.

Galton, F., 1886. Hereditary Stature. *Nature*, 33(849), pp.317–317.

Garant, D., Hadfield, J.D., Kruuk, L.E.B. & Sheldon, B.C., 2008. Stability of genetic variance and covariance for reproductive characters in the face of climate change in a wild bird population. *Molecular Ecology*, 17(1), pp.179–88.

Garant, D., Kruuk, L.E.B., McCleery, R.H. & Sheldon, B.C., 2004. Evolution in a changing environment: a case study with great tit fledging mass. *The American Naturalist*, 164(5), pp.E115–29.

Garroway, C.J., Radersma, R., Sepil, I., Santure, A.W., De Cauwer, I., Slate, J. & Sheldon, B.C., 2013. Fine-scale genetic structure in a wild bird population: the role of limited dispersal and environmentally-based selection as causal factors. *Evolution*, 67(12), pp.3488–3500.

Gavrilets, S. & Scheiner, S.M., 1993. The genetics of phenotypic plasticity. V. Evolution of reaction norm shape. *Journal of Evolutionary Biology*, 6(1), pp.31–48.

Gebhardt-Henrich, S.G., 1990. Temporal and spatial variation in food availability and its effects on fledgling size in the great tit. In J. Blondel et al., eds. *Population Biology of Passerine Birds*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 175–185.

George, a W., Visscher, P.M. & Haley, C.S., 2000. Mapping quantitative trait loci in complex pedigrees: a two-step variance component approach. *Genetics*, 156(4), pp.2081–92.

Gershenfeld, H.K., Neumann, P.E., Mathis, C., Crawley, J.N., Li, X. & Paul, S.M., 1997. Mapping Quantitative Trait Loci for Open-Field Behavior in Mice. *Behavior Genetics*.

Ghalambor, C.K., McKay, J.K., Carroll, S.P. & Reznick, D.N., 2007. Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, 21(3), pp.394–407.

Gibb, J., 1950. The breeding biology of the great and blue titmice. *Ibis*, 92(4), pp.507–539.

Gienapp, P., Lof, M., Reed, T.E., McNamara, J., Verhulst, S. & Visser, M.E., 2013. Predicting demographically sustainable rates of adaptation: can great tit breeding time keep pace with climate change? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 368(1610), p.20120289.

Gienapp, P., Postma, E., Visser, M.E. & Url, S., 2006. Why breeding time has not responded to selection for earlier breeding in a songbird population. *Evolution*, 60(11), pp.2381–2388.

Gienapp, P., Teplitsky, C., Alho, J.S., Mills, J.A. & Merilä, J., 2008. Climate change and evolution: disentangling environmental and genetic responses. *Molecular ecology*, 17(1), pp.167–78.

Gienapp, P., van Noordwijk, A.J. & Visser, M.E., 2013. Genetic background, and not ontogenetic effects, affects avian seasonal timing of reproduction. *Journal of Evolutionary Biology*, pp.1–7.

Gilmour, A.R., Cullis, B.R. & Verbyla, A.P.V., 1997. Accounting for Natural and Extraneous Variation in the Analysis of Field Experiments. *Journal of Agricultural, Biological and Environmental Statistics*, 2(3), pp.269–293.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. & Thompson, R., 2002. ASReml user guide release 1.0.

Gilmour, A.R., Thompson, R., Cullis, B.R. & Welham, S., 1996. *ASREML*, Orange, NSW Australia: NSW Agriculture.

Glenske, K., Brandt, H., Prinzenberg, E.M., Gauly, M. & Erhardt, G., 2010. Verification of a QTL on BTA1 for temperament in German Simmental and German Angus calves. *Archiv Tierzucht*, 53(4), pp.388–392.

Glenske, K., Prinzenberg, E.-M., Brandt, H., Gauly, M. & Erhardt, G., 2011. A chromosome-wide QTL study on BTA29 affecting temperament traits in German Angus beef cattle and mapping of DRD4. *Animal*, 5(2), pp.195–7.

Goddard, M.E. & Hayes, B.J., 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature reviews. Genetics*, 10(6), pp.381–91.

Goodenough, A.E., Elliot, S.L., Maitland, D.P. & Hart, A.G., 2009. Variation in the Relationship between Lay Date and Clutch Size in Three Cavity-Nesting Woodland Passerines. *Acta Ornithologica*, 44(1), pp.27–36.

Gordo, O. & Sanz, J.J., 2010. Impact of climate change on plant phenology in Mediterranean ecosystems. *Global Change Biology*, 16(3), pp.1082–1106.

Göring, H.H., Terwilliger, J.D. & Blangero, J., 2001. Large upward bias in estimation of locus-specific effects from genomewide scans. *American Journal of Human Genetics*, 69(6), pp.1357–69.

Gosler, A., Clement, P. & Christie, D.A., 2013. Great tit (Parus major). In J. del Hoyo et al., eds. *Handbook of the Birds of the World Alive*. Barcelona: Lynx Edicions.

Gratten, J., Beraldi, D., Lowder, B. V, McRae, a F., Visscher, P.M., Pemberton, J.M. & Slate, J., 2007. Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings. Biological sciences / The Royal Society*, 274(1610), pp.619–26.

Gratten, J., Pilkington, J.G., Brown, E. a, Beraldi, D., Pemberton, J.M. & Slate, J., 2010. The genetic basis of recessive self-colour pattern in a wild sheep population. *Heredity*, 104(2), pp.206–14.

Greenwood, A.K., Ardekani, R., McCann, S.R., Dubin, M.E., Sullivan, A., Bensussen, S., Tavaré, S. & Peichel, C.L., 2015. Genetic mapping of natural variation in schooling tendency in the threespine stickleback. *G3 (Bethesda, Md.)*, 5(5), pp.761–9.

Gregory, B., Christophe, L. & Martin, E., 2009. Rapid biogeographical plankton shifts in the North Atlantic Ocean. *Global Change Biology*, 15(7), pp.1790–1803.

Gudbjartsson, D.F., Walters, G.B., Thorleifsson, G., Stefansson, H., Halldorsson, B. V, Zusmanovich, P., Sulem, P., Thorlacius, S., Gylfason, A., Steinberg, S., Helgadottir, A., Ingason, A., Steinthorsdottir, V., Olafsdottir, E.J., Olafsdottir, G.H., Jonsson, T., Borch-Johnsen, K., Hansen, T., Andersen, G., Jorgensen, T., Pedersen, O., Aben, K.K., Witjes, J.A., Swinkels, D.W., den Heijer, M., Franke, B., Verbeek, A.L.M., Becker, D.M., Yanek, L.R., Becker, L.C., Tryggvadottir, L., Rafnar, T., Gulcher, J., Kiemeney, L.A., Kong, A., Thorsteinsdottir, U. & Stefansson, K., 2008. Many sequence variants affecting diversity of adult human height. *Nature Genetics*, 40(5), pp.609–15.

Gutiérrez-Gil, B., Ball, N., Burton, D., Haskell, M., Williams, J.L. & Wiener, P., 2008. Identification of quantitative trait loci affecting cattle temperament. *The Journal of Heredity*, 99(6), pp.629–38.

Haftorn, S., 2012. Incubation during the Egg-Laying Period in Relation to Clutch-Size and Other Aspects of Reproduction in the Great Tit Parus major. *Ornis Scandinavica*, 12(3), pp.169–185.

Hahn, T.P., 1998. Reproductive seasonality in an opportunistic breeder, the red crossbill, Loxia curvirostra. *Ecology*, 79(7), pp.2365–2375.

Hakkarainen, H., Korpimäki, E., Ryssy, J. & Vikström, S., 1996. Low heritability in morphological characters of Tengmalm's owls: The role of cyclic food and laying date? *Evolutionary Ecology*, 10(2), pp.207–219.

Haskell, M.J., Simm, G. & Turner, S.P., 2014. Genetic selection for temperament traits in dairy and beef cattle. *Frontiers in Genetics*, 5, p.368.

Hau, M., Wikelski, M., Gwinner, H. & Gwinner, E., 2004. Timing of reproduction in a Darwin's finch: temporal opportunism under spatial constraints. *Oikos*, 106(3), pp.489–500.

Hayes, B.J., Pryce, J., Chamberlain, A.J., Bowman, P.J. & Goddard, M.E., 2010. Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. *PLoS Genetics*, 6(9), p.e1001139.

Heath, S.C., 1997. Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *American Journal of Human Genetics*, 61(3), pp.748–60.

Heath, S.C., Snow, G.L., Thompson, E.A., Tseng, C. & Wijsman, E.M., 1997. MCMC segregation and linkage analysis. *Genetic Epidemiology*, 14(6), pp.1011–6.

Hecht, B.C., Campbell, N.R., Holecek, D.E. & Narum, S.R., 2013. Genome-wide association reveals genetic basis for the propensity to migrate in wild populations of rainbow and steelhead trout. *Molecular Ecology*, 22(11), pp.3061–76.

Hecht, B.C., Hard, J.J., Thrower, F.P. & Nichols, K.M., 2015. Quantitative genetics of migration-related traits in rainbow and steelhead trout. *G3 (Bethesda, Md.)*, 5(5), pp.873–89.

Hein, L.G., 2011. Economic benefits generated by protected areas: the case of the Hoge Veluwe forest, the Netherlands. *Ecology and Society*, 16(2), p.Art. 13.

Henderson, N.D., Turri, M.G., DeFries, J.C. & Flint, J., 2004. QTL analysis of multiple behavioral measures of anxiety in mice. *Behavior Genetics*, 34(3), pp.267–93.

Hernández-Sánchez, J., Chatzipli, A., Beraldi, D., Gratten, J., Pilkington, J.G. & Pemberton, J.M., 2010. Mapping quantitative trait loci in a wild population using linkage and linkage disequilibrium analyses. *Genetics Research*, 92(4), pp.273–81.

Hill, W.G., 2012. Quantitative genetics in the genomics era. *Current Genomics*, 13(3), pp.196–206.

Hill, W.G., Goddard, M.E. & Visscher, P.M., 2008. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics*, 4(2), p.e1000008.

Hinks, A.E., Cole, E.F., Daniels, K.J., Wilkin, T.A., Nakagawa, S. & Sheldon, B.C., 2015. Scale-Dependent Phenological Synchrony between Songbirds and Their Caterpillar Food Source. *The American Naturalist*, 186(1), pp.84–97.

Hirschhorn, J.N. & Daly, M.J., 2005. Genome-wide association studies for common diseases and complex traits. *Nature reviews. Genetics*, 6(2), pp.95–108.

Hirschorn, J.N. & Daly, M.J., 2005. Genome-wide association studies for common diseases and complex traits. *Nature reviews. Genetics*, pp.95–108.

Holliday, R., 2006. Epigenetics: a historical overview. *Epigenetics*, 1(2), pp.76–80.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), pp.65–70.

Holmes, A., le Guisquet, A.M., Vogel, E., Millstein, R.A., Leman, S. & Belzung, C., 2005. Early life genetic, epigenetic and environmental factors shaping emotionality in rodents. *Neuroscience & Biobehavioral Reviews*, 29(8), pp.1335–1346.

Houde, A.L., Wilson, C.C. & Neff, B.D., 2013. Genetic architecture of survival and fitness-related traits in two populations of Atlantic salmon. *Heredity*, 111(6), pp.513–9.

Hudson, M.E., 2008. Sequencing breakthroughs for genomic ecology and evolutionary biology. *Molecular Ecology Resources*, 8(1), pp.3–17.

Hughes, V., 2014. Epigenetics: The sins of the father. *Nature*, 507(7490), pp.22–4.

Husby, A., Gustafsson, L. & Qvarnström, A., 2012. Low genetic variance in the duration of the incubation period in a collared flycatcher (Ficedula albicollis) population. *The American Naturalist*, 179(1), pp.132–6.

Husby, A., Kawakami, T., Rönnegård, L., Smeds, L., Ellegren, H. & Qvarnström, A., 2015. Genome-wide association mapping in a wild avian population identifies a link between genetic and phenotypic variation in a life-history trait. *Proceedings. Biological sciences / The Royal Society*, 282(1806), p.20150156.

Husby, A., Nussey, D.H., Visser, M.E., Wilson, A.J., Sheldon, B.C. & Kruuk, L.E., 2010. Contrasting patterns of phenotypic plasticity in reproductive traits in two great tit (Parus major) populations. *Evolution*, 64(8), pp.2221–2237.

Jablonka, E. & Raz, G., 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *The Quarterly review of biology*, 84(2), pp.131–76.

Van der Jeugd, H.P. & McCleery, R., 2002. Effects of spatial autocorrelation, natal philopatry and phenotypic plasticity on the heritability of laying date. *Journal of Evolutionary Biology*, 15(3), pp.380–387.

Jiang, Y.-H., Bressler, J. & Beaudet, A.L., 2004. Epigenetics and human disease. *Annual Review of Genomics and Human Genetics*, 5, pp.479–510.

Jimenez-Sanchez, G., Childs, B. & Valle, D., 2001. Human disease genes. *Nature*, pp.853–855.

Johnson, J.L., Wittgenstein, H., Mitchell, S.E., Hyma, K.E., Temnykh, S. V., Kharlamova, A. V., Gulevich, R.G., Vladimirova, A. V., Fong, H.W.F., Acland, G.M., Trut, L.N. & Kukekova, A. V., 2013. Genotyping-By-Sequencing (GBS) Detects Genetic Structure and Confirms Behavioral QTL in Tame and Aggressive Foxes (Vulpes vulpes) W. J. Murphy, ed. *PLoS ONE*, 10(6), p.e0127013.

Johnston, S.E., Beraldi, D., McRae, A.F., Pemberton, J.M. & Slate, J., 2010. Horn type and horn length genes map to the same chromosomal region in Soay sheep. *Heredity*, 104(2), pp.196–205.

Johnston, S.E., McEwan, J.C., Pickering, N.K., Kijas, J.W., Beraldi, D., Pilkington, J.G., Pemberton, J.M. & Slate, J., 2011. Genome-wide association mapping identifies the genetic basis of discrete and quantitative variation in sexual weaponry in a wild sheep population. *Molecular Ecology*, 20(12), pp.2555–66.

Johnston, S.E., Orell, P., Pritchard, V.L., Kent, M.P., Lien, S., Niemelä, E., Erkinaro, J. & Primmer, C.R., 2014. Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a wild population of Atlantic salmon (Salmo salar). *Molecular Ecology*, 23(14), pp.3452–68.

Jones, P.A. & Laird, P.W., 1999. Cancer epigenetics comes of age. *Nature Genetics*, 21(2), pp.163–7.

Källander, H. & Karlsson, J., 1993. Supplemental food and laying date in the Eurupean starling. *The Condor*, 95(4), pp.1031–1034.

Kappen, J.H., Medina-Gomez, C., van Hagen, P.M., Stolk, L., Estrada, K., Rivadeneira, F., Uitterlinden, A.G., Stanford, M.R., Ben-Chetrit, E., Wallace, G.R., Soylu, M. & van Laar, J.A.M., 2015. Genome-wide association study in an admixed case series reveals IL12A as a new candidate in Behçet disease. *PLoS ONE*, 10(3), p.e0119085.

Kearsey, M.J. & Farquhar, A.G.L., 1998. QTL analysis in plants; where are we now? *Heredity*, 80, pp.137–142.

Kerby, J. & Post, E., 2013. Capital and income breeding traits differentiate trophic match-mismatch dynamics in large herbivores. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 368(1624), p.20120484.

Kerstens, H.H.D., Crooijmans, R.P.M.A., Veenendaal, A., Dibbits, B.W., Chin-A-Woeng, T.F.C., den Dunnen, J.T. & Groenen, M.A.M., 2009. Large scale single nucleotide

polymorphism discovery in unsequenced genomes using second generation high throughput sequencing technology: applied to turkey. *BMC Genomics*, 10(1), p.479.

King, C.M., 1976. The fleas of a population of weasels in Wytham Woods, Oxford. *Journal of Zoology*, 180, pp.525–535.

Kirby, K.J. & Thomas, R.C., 2000. Changes in the ground flora in Wytham Woods, southern England from 1974 to 1991 - implications for nature conservation. *Journal of Vegetation Science*, 11(6), pp.871–880.

Kluijver, H.N., 1951. *The population ecology of the Great Tit, Parus m. major L.*, Leiden: Brill.

Koike, S. & Higuchi, H., 2002. Long-term trends in the egg-laying date and clutch size of Red-cheeked Starlings Sturnia philippensis. *Ibis*, 144(1), pp.150–152.

Kolbehdari, D., Wang, Z., Grant, J.R., Murdoch, B., Prasad, A., Xiu, Z., Marques, E., Stothard, P. & Moore, S.S., 2008. A whole-genome scan to map quantitative trait loci for conformation and functional traits in Canadian Holstein bulls. *Journal of Dairy Science*, 91(7), pp.2844–56.

Korsten, P., Mueller, J.C., Hermannstädter, C., Bouwman, K.M., Dingemanse, N.J., Drent, P.J., Liedvogel, M., Matthysen, E., van Oers, K., van Overveld, T., Patrick, S.C., Quinn, J.L., Sheldon, B.C., Tinbergen, J.M. & Kempenaers, B., 2010. Association between DRD4 gene polymorphism and personality variation in great tits: a test across four wild populations. *Molecular Ecology*, 19(4), pp.832–43.

Korte, A. & Farlow, A., 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, 9.

Kowalczyk, N.D., Reina, R.D., Preston, T.J. & Chiaradia, A., 2015. Environmental variability drives shifts in the foraging behaviour and reproductive success of an inshore seabird. *Oecologia*, 178(4), pp.967–79.

Krebs, J.R., 1982. Territorial Defence in the Great Tit (Parus major): Do Residents Always Win ? *Behavioural Ecology and Sociobiology*, 11(3), pp.185–194.

Krebs, J.R., 1971. Territory and Breeding Density in the Great Tit , Parus Major L . *Ecology*, 52(1), pp.2–22.

Kroymann, J. & Mitchell-Olds, T., 2005. Epistasis and balanced polymorphism influencing complex trait variation. *Nature*, 435(7038), pp.95–8.

Kruglyak, L. & Lander, E.S., 1995. A Nonparametric Approach for Mapping Quantitative Trait Loci. *Genetics*, 139, pp.1421–1428.

Kruuk, L.E.B., 2004. Estimating genetic parameters in natural populations using the “animal model”. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 359(1446), pp.873–90.

Kruuk, L.E.B. & Hill, W.G., 2008. Introduction. Evolutionary dynamics of wild populations: the use of long-term pedigree data. *Proceedings. Biological sciences / The Royal Society*, 275(1635), pp.593–6.

Kukekova, A. V., Trut, L.N., Chase, K., Kharlamova, A. V., Johnson, J.L., Temnykh, S. V., Oskina, I.N., Gulevich, R.G., Vladimirova, A. V., Klebanov, S., Shepeleva, D. V., Shikhevich, S.G., Acland, G.M. & Lark, K.G., 2010. Mapping Loci for Fox Domestication: Deconstruction/Reconstruction of a Behavioral Phenotype. *Behavior Genetics*, 41(4), pp.593–606.

Lack, D., 1964. A long-term study of the great tit (Parus major). *Journal of Animal Ecology*, 33, pp.159–173.

Lack, D., 1958. A quantitative breeding study of British tits. *Ardea*, 46(1), pp.91–124.

Lander, E. & Kruglyak, L., 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics*, 11(3), pp.241–7.

Lango Allen, H., Estrada, K., Lettre, G., Berndt, S.I., Weedon, M.N., Rivadeneira, F., Willer, C.J., Jackson, A.U., Vedantam, S., Raychaudhuri, S., Ferreira, T., Wood, A.R., Weyant, R.J., Segrè, A. V, Speliotes, E.K., Wheeler, E., Soranzo, N., Park, J.-H., Yang, J., Gudbjartsson, D., Heard-Costa, N.L., Randall, J.C., Qi, L., Vernon Smith, A., Mägi, R., Pastinen, T., Liang, L., Heid, I.M., Luan, J., Thorleifsson, G., Winkler, T.W., Goddard, M.E., Sin Lo, K., Palmer, C., Workalemahu, T., Aulchenko, Y.S., Johansson, A., Zillikens, M.C., Feitosa, M.F., Esko, T., Johnson, T., Ketkar, S., Kraft, P., Mangino, M., Prokopenko, I., Absher, D., Albrecht, E., Ernst, F., Glazer, N.L., Hayward, C., Hottenga, J.-J., Jacobs, K.B., Knowles, J.W., Kutalik, Z., Monda, K.L., Polasek, O., Preuss, M., Rayner, N.W., Robertson, N.R., Steinthorsdottir, V., Tyrer, J.P., Voight, B.F., Wiklund, F., Xu, J., Zhao, J.H., Nyholt, D.R., Pellikka, N., Perola, M., Perry, J.R.B., Surakka, I., Tammesoo, M.-L., Altmaier, E.L., Amin, N., Aspelund, T., Bhangale, T., Boucher, G., Chasman, D.I., Chen, C., Coin, L., Cooper, M.N., Dixon, A.L., Gibson, Q., Grundberg, E., Hao, K., Juhani Junttila, M., Kaplan, L.M., Kettunen, J., König, I.R., Kwan, T., Lawrence, R.W., Levinson, D.F., Lorentzon, M., McKnight, B., Morris, A.P., Müller, M., Suh Ngwa, J., Purcell, S., Rafelt, S., Salem, R.M., Salvi, E., Sanna, S., Shi, J., Sovio, U., Thompson, J.R., Turchin, M.C., Vandenput, L., Verlaan, D.J., Vitart, V., White, C.C., Ziegler, A., Almgren, P., Balmforth, A.J., Campbell, H., Citterio, L., De Grandi, A., Dominiczak, A., Duan, J., Elliott, P., Elosua, R., Eriksson, J.G., Freimer, N.B., Geus, E.J.C., Glorioso, N., Haiqing, S., Hartikainen, A.-L., Havulinna, A.S., Hicks, A.A., Hui, J., Igl, W., Illig, T., Jula, A., Kajantie, E., Kilpeläinen, T.O., Koiranen, M., Kolcic, I., Koskinen, S., Kovacs, P., Laitinen, J., Liu, J., Lokki, M.-L., Marusic, A., Maschio, A., Meitinger, T., Mulas, A., Paré, G., Parker, A.N., Peden, J.F., Petersmann, A., Pichler, I., Pietiläinen, K.H., Pouta, A., Ridderstråle, M., Rotter, J.I., Sambrook, J.G., Sanders, A.R., Schmidt, C.O., Sinisalo, J., Smit, J.H., Stringham, H.M., Bragi Walters, G., Widen, E., Wild, S.H., Willemsen, G., Zagato, L., Zgaga, L., Zitting, P., Alavere, H., Farrall, M., McArdle, W.L., Nelis, M., Peters, M.J., Ripatti, S., van Meurs, J.B.J., Aben, K.K., Ardlie, K.G., Beckmann, J.S., Beilby, J.P., Bergman, R.N., Bergmann, S., Collins, F.S., Cusi, D., den Heijer, M., Eiriksdottir, G., Gejman, P. V, Hall, A.S., Hamsten, A., Huikuri, H. V, Iribarren, C., Kähönen, M., Kaprio, J., Kathiresan, S., Kiemeney, L., Kocher, T., Launer, L.J., Lehtimäki, T., Melander, O., Mosley, T.H., Musk, A.W., Nieminen, M.S., O'Donnell, C.J., Ohlsson, C., Oostra, B., Palmer, L.J., Raitakari, O., Ridker, P.M., Rioux, J.D., Rissanen, A., Rivolta, C., Schunkert, H., Shuldiner, A.R., Siscovick, D.S., Stumvoll, M., Tönjes, A.,

Tuomilehto, J., van Ommen, G.-J., Viikari, J., Heath, A.C., Martin, N.G., Montgomery, G.W., Province, M.A., Kayser, M., Arnold, A.M., Atwood, L.D., Boerwinkle, E., Chanock, S.J., Deloukas, P., Gieger, C., Grönberg, H., Hall, P., Hattersley, A.T., Hengstenberg, C., Hoffman, W., Lathrop, G.M., Salomaa, V., Schreiber, S., Uda, M., Waterworth, D., Wright, A.F., Assimes, T.L., Barroso, I., Hofman, A., Mohlke, K.L., Boomsma, D.I., Caulfield, M.J., Cupples, L.A., Erdmann, J., Fox, C.S., Gudnason, V., Gyllensten, U., Harris, T.B., Hayes, R.B., Jarvelin, M.-R., Mooser, V., Munroe, P.B., Ouwehand, W.H., Penninx, B.W., Pramstaller, P.P., Quertermous, T., Rudan, I., Samani, N.J., Spector, T.D., Völzke, H., Watkins, H., Wilson, J.F., Groop, L.C., Haritunians, T., Hu, F.B., Kaplan, R.C., Metspalu, A., North, K.E., Schlessinger, D., Wareham, N.J., Hunter, D.J., O'Connell, J.R., Strachan, D.P., Wichmann, H.-E., Borecki, I.B., van Duijn, C.M., Schadt, E.E., Thorsteinsdottir, U., Peltonen, L., Uitterlinden, A.G., Visscher, P.M., Chatterjee, N., Loos, R.J.F., Boehnke, M., McCarthy, M.I., Ingelsson, E., Lindgren, C.M., Abecasis, G.R., Stefansson, K., Frayling, T.M. & Hirschhorn, J.N., 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature*, 467(7317), pp.832–8.

Lattin, C.R., Breuner, C.W. & Michael Romero, L., 2016. Does corticosterone regulate the onset of breeding in free-living birds?: The CORT-Flexibility Hypothesis and six potential mechanisms for priming corticosteroid function. *Hormones and Behavior*, 78, pp.107–120.

Lee, S.H., DeCandia, T.R., Ripke, S., Yang, J., Sullivan, P.F., Goddard, M.E., Keller, M.C., Visscher, P.M. & Wray, N.R., 2012. Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature Genetics*, 44(3), pp.247–50.

Legendre, P., 1993. Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74, pp.1659–1673.

Legendre, P., Dale, M.R.T., Fortin, M.-J., Gurevitch, J., Hohn, M. & Myers, D., 2002. The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, 25(5), pp.601–615.

Legendre, P. & Fortin, M.J., 1989. Spatial pattern and ecological analysis. *Vegetatio*, 80(2), pp.107–138.

Lettre, G., Jackson, A.U., Gieger, C., Schumacher, F.R., Berndt, S.I., Sanna, S., Eyheramendy, S., Voight, B.F., Butler, J.L., Guiducci, C., Illig, T., Hackett, R., Heid, I.M., Jacobs, K.B., Lyssenko, V., Uda, M., Boehnke, M., Chanock, S.J., Groop, L.C., Hu, F.B., Isomaa, B., Kraft, P., Peltonen, L., Salomaa, V., Schlessinger, D., Hunter, D.J., Hayes, R.B., Abecasis, G.R., Wichmann, H.-E., Mohlke, K.L. & Hirschhorn, J.N., 2008. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nature Genetics*, 40(5), pp.584–91.

Li, C.C.Y., Young, P.E., Maloney, C.A., Eaton, S.A., Cowley, M.J., Buckland, M.E., Preiss, T., Henstridge, D.C., Cooney, G.J., Febbraio, M.A., Martin, D.I.K., Cropley, J.E. & Suter, C.M., 2013. Maternal obesity and diabetes induces latent metabolic defects and widespread epigenetic changes in isogenic mice. *Epigenetics*, 8(6), pp.602–11.

Li, M.H. & Merilä, J., 2010. Extensive linkage disequilibrium in a wild bird population. *Heredity*, 104(6), pp.600–10.

Liedvogel, M., Cornwallis, C.K. & Sheldon, B.C., 2012. Integrating candidate gene and quantitative genetic approaches to understand variation in timing of breeding in wild tit populations. *Journal of Evolutionary Biology*, 25(5), pp.813–23.

Liedvogel, M. & Sheldon, B.C., 2010. Low variability and absence of phenotypic correlates of Clock gene variation in a great tit Parus major population. *Journal of Avian Biology*, 41(5), pp.543–550.

Liedvogel, M., Szulkin, M., Knowles, S.C.L., Wood, M.J. & Sheldon, B.C., 2009. Phenotypic correlates of Clock gene variation in a wild blue tit population: evidence for a role in seasonal timing of reproduction. *Molecular Ecology*, 18(11), pp.2444–56.

Ligtenberg, M.J.L., Kuiper, R.P., Chan, T.L., Goossens, M., Hebeda, K.M., Voorendt, M., Lee, T.Y.H., Bodmer, D., Hoenselaar, E., Hendriks-Cornelissen, S.J.B., Tsui, W.Y., Kong, C.K., Brunner, H.G., van Kessel, A.G., Yuen, S.T., van Krieken, J.H.J.M., Leung, S.Y. & Hoogerbrugge, N., 2009. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. *Nature Genetics*, 41(1), pp.112–7.

Lindblad-Toh, K., Winchester, E., Daly, M.J., Wang, D.G., Hirschhorn, J.N., Laviolette, J.P., Ardlie, K., Reich, D.E., Robinson, E., Sklar, P., Shah, N., Thomas, D., Fan, J.B., Gingeras, T., Warrington, J., Patil, N., Hudson, T.J. & Lander, E.S., 2000. Large-scale discovery and genotyping of single-nucleotide polymorphisms in the mouse. *Nature Genetics*, 24(4), pp.381–6.

Lotterhos, K.E. & Schaal, S.M., 2014. Genome scans for the contemporary response to selection in quantitative traits. *Molecular Ecology*, 23(18), pp.4435–7.

Lynch, M. & Lande, R., 1993. Evolution and extinction in reponse to environmental change. In P. M. Kareiva, J. G. Kingsolver, & R. B. Huey, eds. *Biotic Interactions and Global Change*. Sunderland, Massachusetts: Sinauer Associates, Inc, pp. 251–266.

Lynch, M. & Walsh, B., 1998. *Genetics and Analysis of Quantitative Traits*, MA, U.S.A: Sinauer Associates, Inc.

Lyon, B.E., Chaine, A.S. & Winkler, D.W., 2008. A matter of timing. *Science*, 321(5892), pp.1051–2.

Macdonald, D.W., Newman, C., Stewart, P.D., Domingo-Roura, X. & Johnson, P.J., 2002. Density-dependent regulation of body mass and condition in badgers (Meles meles) from Wytham Woods. *Ecology*, 83(7), pp.2056–2061.

Mackay, T., Stone, E. & Ayroles, J., 2009. The genetics of quantitative traits: challenges and prospects. *Nature reviews. Genetics*, pp.565–577.

Mackay, T.F., 2001. The genetic architecture of quantitative traits. *Annual Review of Genetics*, 35, pp.303–39.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R., Chakravarti, A., Cho, J.H., Guttmacher, A.E., Kong, A., Kruglyak, L., Mardis, E., Rotimi, C.N., Slatkin, M., Valle, D., Whittemore, A.S., Boehnke, M., Clark, A.G., Eichler, E.E., Gibson, G., Haines, J.L., Mackay, T.F.C., McCarroll, S.A. & Visscher, P.M., 2009. Finding the missing heritability of complex diseases. *Nature*, 461(7265), pp.747–53.

Marrot, P., Garant, D. & Charmantier, A., 2015. Spatial autocorrelation in fitness affects the estimation of natural selection in the wild. *Methods in Ecology and Evolution*.

Mastroeni, D., Grover, A., Delvaux, E., Whiteside, C., Coleman, P.D. & Rogers, J., 2011. Epigenetic mechanisms in Alzheimer's disease. *Neurobiology of Aging*, 32(7), pp.1161–1180.

Matrisciano, F., Tueting, P., Dalal, I., Kadriu, B., Grayson, D.R., Davis, J.M., Nicoletti, F. & Guidotti, A., 2013. Epigenetic modifications of GABAergic interneurons are associated with the schizophrenia-like phenotype induced by prenatal stress in mice. *Neuropharmacology*, 68, pp.184–94.

Matthysen, E., Adriaensen, F. & Dhondt, A. a., 2011. Multiple responses to increasing spring temperatures in the breeding cycle of blue and great tits (Cyanistes caeruleus, Parus major). *Global Change Biology*, 17(1), pp.1–16.

McCleery, R.H. & Perrins, C.M., 1998. .....temperature and egg-laying trends. *Nature*, 391(6662), pp.30–31.

McCleery, R.H., Pettifor, R. a, Armbruster, P., Meyer, K., Sheldon, B.C. & Perrins, C.M., 2004. Components of variance underlying fitness in a natural population of the great tit Parus major. *The American Naturalist*, 164(3), pp.E62–72.

Meijer, T. & Langer, U., 1995. Food availability and egg-laying of captive European starlings. *The Condor*, 97(3), pp.718–728.

Meloni, M. & Testa, G., 2014. Scrutinizing the epigenetics revolution. *BioSocieties*, 9(4), pp.431–456.

Memmott, J., Craze, P.G., Waser, N.M. & Price, M. V, 2007. Global warming and the disruption of plant–pollinator interactions. *Ecology Letters*, 10(8), pp.710–717.

Merilä, J. & Crnokrak, P., 2001. Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, 14(6), pp.892–903.

Merilä, J., Kruuk, L.E. & Sheldon, B.C., 2001. Cryptic evolution in a wild bird population. *Nature*, 412(6842), pp.76–9.

Merilä, J., Sheldon, B.C. & Kruuk, L.E., 2001. Explaining stasis: microevolutionary studies in natural populations. *Genetica*, 112-113(2), pp.199–222.

Metzker, M.L., 2010. Sequencing technologies - the next generation. *Nature reviews. Genetics*, 11(1), pp.31–46.

Meyer, K., 1989. Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. *Genetics, Selection, Evolution*, 21, pp.317–340.

Miller-Rushing, A.J., Høye, T.T., Inouye, D.W. & Post, E., 2010. The effects of phenological mismatches on demography. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 365(1555), pp.3177–86.

Miller-Struttmann, N.E., Geib, J.C., Franklin, J.D., Kevan, P.G., Holdo, R.M., Ebert-May, D., Lynn, A.M., Kettenbach, J.A., Hedrick, E. & Galen, C., 2015. Functional mismatch in a bumble bee pollination mutualism under climate change. *Science*, 349(6255), pp.1541–1544.

Møller, A.P., 2001. Heritability of arrival date in a migratory bird. *Proceedings. Biological sciences / The Royal Society*, 268(1463), pp.203–6.

Møller, A.P., Flensted-Jensen, E. & Mardal, W., 2006. Rapidly advancing laying date in a seabird and the changing advantage of early reproduction. *The Journal of Animal Ecology*, 75(3), pp.657–65.

Morgan, H.D., Sutherland, H.G., Martin, D.I. & Whitelaw, E., 1999. Epigenetic inheritance at the agouti locus in the mouse. *Nature Genetics*, 23(3), pp.314–8.

Morin, P. a., Luikart, G. & Wayne, R.K., 2004. SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, 19(4), pp.208–216.

El Mousadik, A. & Petit, R.J., 1996. High level of genetic differentiation for allelic richness among populations of the argan tree [Argania spinosa (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, 92(7), pp.832–9.

Mousseau, T.A. & Roff, D.A., 1987. Natural selection and the heritability of fitness components. *Heredity*, 59(2), pp.181–197.

Moussus, J.-P., Clavel, J., Jiguet, F. & Julliard, R., 2011. Which are the phenologically flexible species? A case study with common passerine birds. *Oikos*, 120(7), pp.991–998.

Moyes, K., Nussey, D.H., Clements, M.N., Guinness, F.E., Morris, A., Morris, S., Pemberton, J.M., Kruuk, L.E.B. & Clutton-Brock, T.H., 2011. Advancing breeding phenology in response to environmental change in a wild red deer population. *Global Change Biology*, 17(7), pp.2455–2469.

Mueller, J.C., Korsten, P., Hermannstaedter, C., Feulner, T., Dingemanse, N.J., Matthysen, E., van Oers, K., van Overveld, T., Patrick, S.C., Quinn, J.L., Riemenschneider, M., Tinbergen, J.M. & Kempenaers, B., 2013. Haplotype structure, adaptive history and associations with exploratory behaviour of the DRD4 gene region in four great tit (Parus major) populations. *Molecular Ecology*, 22(10), pp.2797–2809.

Naef-Daenzer, L., Naef-Daenzer, B. & Nager, R.G., 2000. Prey selection and foraging performance of breeding Great Tits Parus major in relation to food availability. *Journal of Avian Biology*, 31(2), pp.206–214.

Nagamine, Y., Pong-Wong, R., Navarro, P., Vitart, V., Hayward, C., Rudan, I., Campbell, H., Wilson, J., Wild, S., Hicks, A.A., Pramstaller, P.P., Hastie, N., Wright, A.F. & Haley, C.S., 2012. Localising loci underlying complex trait variation using Regional Genomic Relationship Mapping. *PLoS ONE*, 7(10), p.e46501.

Narum, S.R., Buerkle, C.A., Davey, J.W., Miller, M.R. & Hohenlohe, P.A., 2013. Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, 22(11), pp.2841–7.

Neale, D.B. & Savolainen, O., 2004. Association genetics of complex traits in conifers. *Trends in plant science*, 9(7), pp.325–30.

Neville, H.M., Isaak, D.J., Dunham, J.B., Thurow, R.F. & Rieman, B.E., 2006. Fine-scale natal homing and localized movement as shaped by sex and spawning habitat in Chinook salmon: insights from spatial autocorrelation analysis of individual genotypes. *Molecular Ecology*, 15(14), pp.4589–602.

Nezer, C., Moreau, L., Brouwers, B., Coppieters, W., Detilleux, J., Hanset, R., Karim, L., Kvasz, A., Leroy, P. & Georges, M., 1999. An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs. *Nature Genetics*, 21(2), pp.155–6.

Nicolazzi, E.L., Iamartino, D. & Williams, J.L., 2014. AffyPipe: an open-source pipeline for Affymetrix Axiom genotyping workflow. *Bioinformatics (Oxford, England)*, 30(21), pp.3118–9.

Nooker, J.K., Dunn, P.O. & Whittingham, L.A., 2005. Effects of food abundance, weather and female condition on reproduction in Tree Swallows (Tachycineta bicolor). *The Auk*, 122(4), p.1225.

Noordwijk, a. J., Balen, J.H. & Scharloo, W., 1981. Genetic variation in the timing of reproduction in the Great Tit. *Oecologia*, 49(2), pp.158–166.

Nussey, D.H., Postma, E., Gienapp, P. & Visser, M.E., 2005. Selection on heritable phenotypic plasticity in a wild bird population. *Science*, 310(5746), pp.304–6.

Nussey, D.H., Wilson, a J. & Brommer, J.E., 2007. The evolutionary ecology of individual phenotypic plasticity in wild populations. *Journal of Evolutionary Biology*, 20(3), pp.831–44.

Nyholt, D.R., 2000. All LODs are not created equal. *American Journal of Human Genetics*, 67(2), pp.282–8.

Otto, S.P. & Jones, C.D., 2000. Detecting the undetected: estimating the total number of loci underlying a quantitative trait. *Genetics*, 156(4), pp.2093–107.

Parmesan, C. & Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918), pp.37–42.

Paterson, a H., Damon, S., Hewitt, J.D., Zamir, D., Rabinowitch, H.D., Lincoln, S.E., Lander, E.S. & Tanksley, S.D., 1991. Mendelian factors underlying quantitative traits in

tomato: comparison across species, generations, and environments. *Genetics*, 127(1), pp.181–97.

Peakall, R., Ruibal, M. & Lindenmayer, D.B., 2003. Spatial autocorrelation analysis offers new insights into gene flow in the Australian bush rat, Rattus fuscipes. *Evolution*, 57(5), pp.1182–1195.

Pearce-Higgins, J.W., Yalden, D.W. & Whittingham, M.J., 2005. Warmer springs advance the breeding phenology of golden plovers Pluvialis apricaria and their prey (Tipulidae). *Oecologia*, 143(3), pp.470–6.

Perdeck, A.C. & Cavé, A.J., 1992. Laying Date in the Coot: Effects of Age and Mate Choice. *Journal of Animal Ecology*, 61, pp.13–19.

Perrins, C.M., 2001. Effect of increasing numbers of deer on bird populations in Wytham Woods, central England. *Forestry*, 74(3), pp.299–309.

Perrins, C.M., 1965. Population fluctuations and clutch-size in the great tit, Parus major, L. *Journal of Animal Ecology*, 34(3), pp.601–647.

Perrins, C.M., 1970. The timing of birds breeding seasons. *Ibis*, 112(2), pp.242–255.

Perrins, C.M. & McCleery, R.H., 1989. Laying dates and clutch size in the great tit. *The Wilson Bulletin*, pp.236–253.

Perry, A.L., Low, P.J., Ellis, J.R. & Reynolds, J.D., 2005. Climate Change and Distribution Shifts in Marine Fishes. *Science*, 308(5730), pp.1912–1915.

Pimentel, E. da C.G., Erbe, M., König, S. & Simianer, H., 2011. Genome partitioning of genetic variation for milk production and composition traits in holstein cattle. *Frontiers in Genetics*, 2, p.19.

Poissant, J., Davis, C.S., Malenfant, R.M., Hogg, J.T. & Coltman, D.W., 2012. QTL mapping for sexually dimorphic fitness-related traits in wild bighorn sheep. *Heredity*, 108(3), pp.256–63.

Poissant, J., Réale, D., Martin, J., Festa-Bianchet, M. & Coltman, D., 2013. A quantitative trait locus analysis of personality in wild bighorn sheep. *Ecology and Evolution*, 3(3), pp.474–81.

Porlier, M., Charmantier, A., Bourgault, P., Perret, P., Blondel, J. & Garant, D., 2012. Variation in phenotypic plasticity and selection patterns in blue tit breeding time: between- and within-population comparisons. *Journal of Animal Ecology*, 81(5), pp.1041–1051.

Post, E. & Forchhammer, M.C., 2008. Climate change reduces reproductive success of an Arctic herbivore through trophic mismatch. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1501), pp.2369–75.

Postma, E. & van Noordwijk, A.J., 2005. Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature*, 433(7021), pp.65–8.

Potti, J., 2009. Advanced breeding dates in relation to recent climate warming in a Mediterranean montane population of Blue Tits Cyanistes caeruleus. *Journal of Ornithology*, 150(4), pp.893–901.

Pozsgai, G. & Littlewood, N. a., 2011. Changes in the phenology of the ground beetle Pterostichus madidus (Fabricius, 1775). *Insect Science*, 18(4), pp.462–472.

Prasad, A.M., Iverson, L.R. & Liaw, A., 2006. Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. *Ecosystems*, 9(2), pp.181–199.

Price, A.L., Patterson, N.J., Plenge, R.M., Weinblatt, M.E., Shadick, N.A. & Reich, D., 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8), pp.904–9.

Price, T.D., Qvarnström, A. & Irwin, D.E., 2003. The role of phenotypic plasticity in driving genetic evolution. *Proceedings. Biological sciences / The Royal Society*, 270(1523), pp.1433–40.

Pritchard, J.K. & Przeworski, M., 2001. Linkage disequilibrium in humans: models and data. *American Journal of Human Genetics*, 69(1), pp.1–14.

Pritchard, J.K. & Rosenberg, N.A., 1999. Use of unlinked genetic markers to detect population stratification in association studies. *American Journal of Human Genetics*, 65(1), pp.220–8.

Protas, M.E., Hersey, C., Kochanek, D., Zhou, Y., Wilkens, H., Jeffery, W.R., Zon, L.I., Borowsky, R. & Tabin, C.J., 2006. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics*, 38(1), pp.107–11.

Przybylo, R., Sheldon, B.C. & Merila, J., 2000. Climatic effects on breeding and morphology: evidence for phenotypic plasticity. *Journal of Animal Ecology*, 69(3), pp.395–403.

Pulido, F. & Berthold, P., 2003. Quantitative genetic analysis of migratory behaviour. In P. Berthold, E. Gwinner, & E. Sonnenschein, eds. *Avian Migration*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 53–77.

Rao, J.S., Keleshian, V.L., Klein, S. & Rapoport, S.I., 2012. Epigenetic modifications in frontal cortex from Alzheimer's disease and bipolar disorder patients. *Translational Psychiatry*, 2, p.e132.

Réale, D., Festa-Bianchet, M. & Jorgenson, J.T., 1999. Heritability of body mass varies with age and season in wild bighorn sheep. *Heredity*, 83(5), pp.526–532.

Réale, D., Martin, J., Coltman, D.W., Poissant, J. & Festa-Bianchet, M., 2009. Male personality, life-history strategies and reproductive success in a promiscuous mammal. *Journal of Evolutionary Biology*, 22(8), pp.1599–607.

Réale, D., McAdam, A.G., Boutin, S. & Berteaux, D., 2003. Genetic and plastic responses of a northern mammal to climate change. *Proceedings. Biological sciences / The Royal Society*, 270(1515), pp.591–6.

Recoquillay, J., Pitel, F., Arnould, C., Leroux, S., Dehais, P., Moréno, C., Calandreau, L., Bertin, A., Gourichon, D., Bouchez, O., Vignal, A., Fariello, M.I., Minvielle, F., Beaumont, C., Leterrier, C. & Bihan-Duval, E. Le, 2015. A medium density genetic map and QTL for behavioral and production traits in Japanese quail. *BMC Genomics*.

Reed, T.E., Grotan, V., Jenouvrier, S., Saether, B.E. & Visser, M.E., 2013. Population Growth in a Wild Bird Is Buffered Against Phenological Mismatch. *Science*, 340(6131), pp.488–491.

Reed, T.E., Jenouvrier, S. & Visser, M.E., 2013. Phenological mismatch strongly affects individual fitness but not population demography in a woodland passerine. *The Journal of Animal Ecology*, 82(1), pp.131–44.

Reed, T.E., Warzybok, P., Wilson, A.J., Bradley, R.W., Wanless, S. & Sydeman, W.J., 2009. Timing is everything: flexible phenology and shifting selection in a colonial seabird. *The Journal of Animal Ecology*, 78(2), pp.376–87.

Reif, A. & Lesch, K.-P., 2003. Toward a molecular architecture of personality. *Behavioural Brain Research*, 139(1-2), pp.1–20.

Reifsnyder, P.C., 2000. Maternal Environment and Genotype Interact to Establish Diabesity in Mice. *Genome Research*, 10(10), pp.1568–1578.

Rich, S.S., 1990. Mapping Genes in Diabetes: Genetic Epidemiological Perspective. *Diabetes*, 39(11), pp.1315–1319.

Riedelsheimer, C., Czedik-Eysenberg, A., Grieder, C., Lisec, J., Technow, F., Sulpice, R., Altmann, T., Stitt, M., Willmitzer, L. & Melchinger, A.E., 2012. Genomic and metabolic prediction of complex heterotic traits in hybrid maize. *Nature Genetics*, 44(2), pp.217–20.

Riggio, V., Matika, O., Pong-Wong, R., Stear, M.J. & Bishop, S.C., 2013. Genome-wide association and regional heritability mapping to identify loci underlying variation in nematode resistance and body weight in Scottish Blackface lambs. *Heredity*, 110(5), pp.420–9.

Riggio, V. & Pong-Wong, R., 2014. Regional Heritability Mapping to identify loci underlying genetic variation of complex traits. *BMC Proceedings*, 8(Suppl 5), p.S3.

Risch, N., 2000. Searching for genetic determinants in the new millenium. *Nature*, 405(6788), pp.847–856.

Risch, N. & Merikangas, K., 1996. The future of genetic studies of complex human diseases. *Science*, pp.1516–1517.

Riyahi, S., Sánchez-Delgado, M., Calafell, F., Monk, D. & Senar, J.C., 2015. Combined epigenetic and intraspecific variation of the DRD4 and SERT genes influence novelty seeking behavior in great tit Parus major. *Epigenetics*, 10(6), pp.516–25.

Robinson, M.R., Santure, A.W., DeCauwer, I., Sheldon, B.C. & Slate, J., 2013. Partitioning of genetic variation across the genome using multimarker methods in a wild bird population. *Molecular Ecology*, 22(15), pp.3963–3980.

Romanov, M.N., Tuttle, E.M., Houck, M.L., Modi, W.S., Chemnick, L.G., Korody, M.L., Mork, E.M.S., Otten, C. a, Renner, T., Jones, K.C., Dandekar, S., Papp, J.C., Da, Y., Green, E.D., Magrini, V., Hickenbotham, M.T., Glasscock, J., McGrath, S., Mardis, E.R. & Ryder, O. a, 2009. The value of avian genomics to the conservation of wildlife. *BMC Genomics*, 10 Suppl 2, p.S10.

Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig, C. & Pounds, J.A., 2003. Fingerprints of global warming on wild animals and plants. *Nature*, 421(6918), pp.57–60.

Rüppell, O., Pankiw, T. & Page, R.E., 2004. Pleiotropy, epistasis and new QTL: the genetic architecture of honey bee foraging behavior. *The Journal of Heredity*, 95(6), pp.481–91.

Saino, N., Ambrosini, R., Rubolini, D., von Hardenberg, J., Provenzale, A., Hüppop, K., Hüppop, O., Lehikoinen, A., Lehikoinen, E., Rainio, K., Romano, M. & Sokolov, L., 2010. Climate warming, ecological mismatch at arrival and population decline in migratory birds. *Proceedings. Biological sciences / The Royal Society*, 278(1707), pp.835–42.

Salvatore, J.E., Edwards, A.C., McClintick, J.N., Bigdeli, T.B., Adkins, A., Aliev, F., Edenberg, H.J., Foroud, T., Hesselbrock, V., Kramer, J., Nurnberger, J.I., Schuckit, M., Tischfield, J.A., Xuei, X. & Dick, D.M., 2015. Genome-wide association data suggest ABCB1 and immune-related gene sets may be involved in adult antisocial behavior. *Translational Psychiatry*, 5, p.e558.

Santure, A.W., De Cauwer, I., Robinson, M.R., Poissant, J., Sheldon, B.C. & Slate, J., 2013. Genomic dissection of variation in clutch size and egg mass in a wild great tit (Parus major) population. *Molecular Ecology*, 22(15), pp.3949–3962.

Santure, A.W., Gratten, J., Mossman, J.A., Sheldon, B.C. & Slate, J., 2011. Characterisation of the transcriptome of a wild great tit Parus major population by next generation sequencing. *BMC Genomics*, 12(1), p.283.

Sanz, J.J., Potti, J., Moreno, J., Merino, S. & Frias, O., 2003. Climate change and fitness components of a migratory bird breeding in the Mediterranean region. *Global Change Biology*, 9(3), pp.461–472.

Savill, P.S., Perrins, C.M., Kirby, K.J. & Fisher, N. eds., 2010. *Wytham Woods: Oxford's Ecological Laboratory*, New York: Oxford University Press.

Schaper, S. V, Dawson, A., Sharp, P.J., Gienapp, P., Caro, S.P. & Visser, M.E., 2012. Increasing temperature, not mean temperature, is a cue for avian timing of reproduction. *The American Naturalist*, 179(2), pp.E55–69.

Schaper, S. V, Rueda, C., Sharp, P.J., Dawson, A. & Visser, M.E., 2011. Spring phenology does not affect timing of reproduction in the great tit (Parus major). *The Journal of Experimental Biology*, 214(Pt 21), pp.3664–71.

Schielzeth, H. & Husby, A., 2014. Challenges and prospects in genome-wide quantitative trait loci mapping of standing genetic variation in natural populations. *Annals of the New York Academy of Sciences*, 1320(1), pp.35–57.

Schlichting, C.D. & Pigliucci, M., 1998. *Phenotypic evolution: a reaction norm perspective.*, Sinauer Associates Incorporated.

Schütz, K., Kerje, S., Carlborg, Örjan, Jacobsson, L., Andersson, L. & Jensen, P., 2002. QTL analysis of a red junglefowl x white leghorn intercross reveals trade-off in resource allocation between behaviour and production traits. *Behavior Genetics*, 32(6).

Seddon, J.M., Parker, H.G., Ostrander, E. a & Ellegren, H., 2005. SNPs in ecological and conservation studies: a test in the Scandinavian wolf population. *Molecular Ecology*, 14(2), pp.503–11.

Sheldon, B.C., Kruuk, L.E.B. & Merila, J., 2003. Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution*, 57(2), pp.406–420.

Shultz, M., Piatt, J., Harding, A., Kettle, A. & Van Pelt, T., 2009. Timing of breeding and reproductive performance in murres and kittiwakes reflect mismatched seasonal prey dynamics. *Marine Ecology Progress Series*, 393, pp.247–258.

Silva, J.C. e, Dutkowski, G.W. & Gilmour, A.R., 2001. Analysis of early tree height in forest genetic trials is enhanced by including a spatially correlated residual. *Canadian Journal of Forest Research*, 31, pp.1887–1893.

Silverin, B., Wingfield, J., Stokkan, K.-A., Massa, R., Järvinen, A., Andersson, N.-A., Lambrechts, M., Sorace, A. & Blomqvist, D., 2008. Ambient temperature effects on photo induced gonadal cycles and hormonal secretion patterns in Great Tits from three different breeding latitudes. *Hormones and Behavior*, 54(1), pp.60–8.

Slate, J., 2013. From Beavis to Beak colour: a simulation study to examine how much QTL mapping can reveal about the genetic architecture of quantitative traits. *Evolution*, 67(5), pp.1251–1262.

Slate, J., 2005. Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Molecular Ecology*, 14(2), pp.363–79.

Slate, J., Gratten, J., Beraldi, D., Stapley, J., Hale, M. & Pemberton, J.M., 2009. Gene mapping in the wild with SNPs: guidelines and future directions. *Genetica*, 136(1), pp.97–107.

Slate, J., Santure, A.W., Feulner, P.G.D., Brown, E. a, Ball, A.D., Johnston, S.E. & Gratten, J., 2010. Genome mapping in intensively studied wild vertebrate populations. *Trends in Genetics*, 26(6), pp.275–84.

Slate, J., Visscher, P.M., MacGregor, S., Stevens, D., Tate, M.L. & Pemberton, J.M., 2002. A genome scan for quantitative trait loci in a wild population of red deer (Cervus elaphus). *Genetics*, 162(4), pp.1863–73.

Smith, K.W., Smith, L., Charman, E., Briggs, K., Burgess, M., Dennis, C., Harding, M., Isherwood, C., Isherwood, I. & Mallord, J., 2011. Large-scale variation in the temporal patterns of the frass fall of defoliating caterpillars in oak woodlands in Britain: implications for nesting woodland birds. *Bird Study*, 58(4), pp.506–511.

Snow, D.W. & Perrins, C.M., 1998. *The Birds of the Western Palearctic vol. 2: Passerines*, Oxford: Oxford University Press.

Sokolowska, E. & Hovatta, I., 2013. Anxiety genetics – findings from cross-species genome-wide approaches. *Biology of Mood and Anxiety Disorders*, 3(1), p.9.

Solonen, T., 2014. Timing of breeding in rural and urban Tawny Owls Strixaluco in southern Finland: effects of vole abundance and winter weather. *Journal für Ornithologie*, 155(1), pp.27–36.

Soubry, A., Schildkraut, J.M., Murtha, A., Wang, F., Huang, Z., Bernal, A., Kurtzberg, J., Jirtle, R.L., Murphy, S.K. & Hoyo, C., 2013. Paternal obesity is associated with IGF2 hypomethylation in newborns: results from a Newborn Epigenetics Study (NEST) cohort. *BMC Medicine*, 11, p.29.

Speed, D., Hemani, G., Johnson, M.R. & Balding, D.J., 2012. Improved heritability estimation from genome-wide SNPs. *American Journal of Human Genetics*, 91(6), pp.1011–21.

Stenseth, N.C. & Mysterud, A., 2002. Climate, changing phenology, and other life history traits: nonlinearity and match-mismatch to the environment. *Proceedings of the National Academy of Sciences of the United States of America*, 99(21), pp.13379–81.

Stevenson, I.R. & Bryant, D.M., 2000. Climate change and constraints on breeding. *Nature*, 406(6794), pp.366–7.

Stickney, H.L., Schmutz, J., Woods, I.G., Holtzer, C.C., Dickson, M.C., Kelly, P.D., Myers, R.M. & Talbot, W.S., 2002. Rapid mapping of zebrafish mutations with SNPs and oligonucleotide microarrays. *Genome Research*, 12(12), pp.1929–34.

Stinchcombe, J.R. & Hoekstra, H.E., 2008. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, 100, pp.158–170.

Stireman, J.O., Dyer, L.A., Janzen, D.H., Singer, M.S., Lill, J.T., Marquis, R.J., Ricklefs, R.E., Gentry, G.L., Hallwachs, W., Coley, P.D., Barone, J.A., Greeney, H.F., Connahs, H., Barbosa, P., Morais, H.C. & Diniz, I.R., 2005. Climatic unpredictability and parasitism of caterpillars: implications of global warming. *Proceedings of the National Academy of Sciences of the United States of America*, 102(48), pp.17384–7.

Stirling, D.G., Réale, D. & Roff, D.A., 2002. Selection, structure and the heritability of behaviour. *Journal of Evolutionary Biology*, 15(2), pp.277–289.

Stoleson, S.H. & Beissinger, S.R., 1995. Hatching Asynchrony and the Onset of Incubation in Birds, Revisited. In D. M. Power, ed. *Current Ornithology*. Boston, MA: Springer US, pp. 191–270.

Stopher, K. V, Walling, C.A., Morris, A., Guinness, F.E., Clutton-brock, T.H., Pemberton, J.M. & Nussey, D.H., 2012. Shared spatial effects on quantitative genetic parameters : accounting for spatial autocorrelation and home range overlap reduces estimates of heritability in wild red deer. *Evolution*, 66-8, pp.2411–2426.

Storey, J.D., 2002. A direct approach to false discovery rates. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(3), pp.479–498.

Storey, J.D. & Tibshirani, R., 2003. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences of the United States of America*, 100(16), pp.9440–5.

Svishcheva, G.R., Axenovich, T.I., Belonogova, N.M., van Duijn, C.M. & Aulchenko, Y.S., 2012. Rapid variance components-based method for whole-genome association analysis. *Nature Genetics*, 44(10), pp.1166–70.

Tabangin, M.E., Woo, J.G. & Martin, L.J., 2009. The effect of minor allele frequency on the likelihood of obtaining false positives. *BMC Proceedings*, 3 Suppl 7(Suppl 7), p.S41.

Tarka, M., Akesson, M., Beraldi, D., Hernández-Sánchez, J., Hasselquist, D., Bensch, S. & Hansson, B., 2010. A strong quantitative trait locus for wing length on chromosome 2 in a wild population of great reed warblers. *Proceedings. Biological sciences / The Royal Society*, 277(1692), pp.2361–9.

Tarka, M., Hansson, B. & Hasselquist, D., 2015. Selection and evolutionary potential of spring arrival phenology in males and females of a migratory songbird. *Journal of Evolutionary Biology*, 28(5), pp.1024–38.

Team, R.C., 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Teplitsky, C., Mills, J. a, Yarrall, J.W. & Merilä, J., 2009. Heritability of fitness components in a wild bird population. *Evolution*, 63(3), pp.716–26.

Terracciano, A., Sanna, S., Uda, M., Deiana, B., Usala, G., Busonero, F., Maschio, A., Scally, M., Patriciu, N., Chen, W.-M., Distel, M.A., Slagboom, E.P., Boomsma, D.I., Villafuerte, S., Sliwerska, E., Burmeister, M., Amin, N., Janssens, A.C.J.W., van Duijn, C.M., Schlessinger, D., Abecasis, G.R. & Costa, P.T., 2010. Genome-wide association scan for five major dimensions of personality. *Molecular Psychiatry*, 15(6), pp.647–56.

Thomas, D.W., Blondel, J., Perret, P., Lambrechts, M.M. & Speakman, J.R., 2001. Energetic and fitness costs of mismatching resource supply and demand in seasonally breeding birds. *Science*, 291(5513), pp.2598–600.

Thomas, H., 2015. Pregnancy: Maternal obesity affects neonatal adipogenesis. *Nature reviews. Endocrinology*, 12(2), p.63.

Thorley, J.B. & Lord, A.M., 2015. Laying Date is a Plastic and Repeatable Trait in a Population of Blue Tits Cyanistes caeruleus. *Ardea*, 103(1), pp.69–78.

Tobler, 1970. A computer movie simulating urban growth in the detroit region. *Economic Geography*, pp.234–240.

Torti, V.M. & Dunn, P.O., 2005. Variable effects of climate change on six species of North American birds. *Oecologia*, 145(3), pp.486–95.

Traynor, B.J. & Renton, A.E., 2015. Exploring the epigenetics of Alzheimer disease. *JAMA neurology*, 72(1), pp.8–9.

Turchetto, C., Segatto, A.L.A., Telles, M.P.C., Diniz-Filho, J.A.F. & Freitas, L.B., 2014. Infraspecific classification reflects genetic differentiation in the widespread Petunia axillaris complex: A comparison among morphological, ecological, and genetic patterns of geographic variation. *Perspectives in Plant Ecology, Evolution and Systematics*, 16(2), pp.75–82.

Uemoto, Y., Pong-Wong, R., Navarro, P., Vitart, V., Hayward, C., Wilson, J.F., Rudan, I., Campbell, H., Hastie, N.D., Wright, A.F. & Haley, C.S., 2013. The power of regional heritability analysis for rare and common variant detection: simulations and application to eye biometrical traits. *Frontiers in Genetics*, 4, p.232.

van Asch, M., van Tienderen, P.H., Holleman, L.J.M. & Visser, M.E., 2007. Predicting adaptation of phenology in response to climate change, an insect herbivore example. *Global Change Biology*, 13(8), pp.1596–1604.

van Asch, M. & Visser, M.E., 2007. Phenology of forest caterpillars and their host trees: The importance of synchrony. *Annual Review of Entomology*, 52, pp.37–55.

van Balen, J.H., 1973. *A comparative study of the breeding ecology of the great tit, Parus major, in different habitats*. Wageningen University.

van Bers, N.E.M., Santure, A.W., Van Oers, K., De Cauwer, I., Dibbits, B.W., Mateman, C., Crooijmans, R.P.M.A., Sheldon, B.C., Visser, M.E., Groenen, M.A.M. & Slate, J., 2012. The design and cross-population application of a genome-wide SNP chip for the great tit Parus major. *Molecular Ecology Resources*, 12(4), pp.753–70.

van Bers, N.E.M., van Oers, K., Kerstens, H.H.D., Dibbits, B.W., Crooijmans, R.P.M. a, Visser, M.E. & Groenen, M. a M., 2010. Genome-wide SNP detection in the great tit Parus major using high throughput sequencing. *Molecular Ecology*, 19 Suppl 1, pp.89–99.

van Engeland, M., Derks, S., Smits, K.M., Meijer, G.A. & Herman, J.G., 2011. Colorectal cancer epigenetics: complex simplicity. *Journal of Clinical Oncology*, 29(10), pp.1382–91.

van Noordwijk, A.J., McCleery, R.H. & Perrins, C.M., 1995. Selection for the timing of great tit breeding in relation to caterpillar growth and temperature. *Journal of Animal Ecology*, 64(4), pp.451–458.

van Noordwijk, A.J., van Balen, J.H. & Scharloo, W., 1981. Genetic variation in the timing of reproduction in the great tit. *Oecologia*, 49(2), pp.158–166.

van Oers, K., Drent, P.J., de Goede, P. & van Noordwijk, A.J., 2003. Realized heritability and repeatability of risk-taking behaviour in relation to avian personalities. *Proceedings of the Royal Society of London B*, pp.65–73.

van Oers, K., Santure, A.W., De Cauwer, I., van Bers, N.E.M., Crooijmans, R.P.M.A., Sheldon, B.C., Visser, M.E., Slate, J. & Groenen, M.A.M., 2014. Replicated high-density genetic maps of two great tit populations reveal fine-scale genomic departures from sex-equal recombination rates. *Heredity*, 112(3), pp.307–16.

Vaysse, A., Ratnakumar, A., Derrien, T., Axelsson, E., Rosengren Pielberg, G., Sigurdsson, S., Fall, T., Seppälä, E.H., Hansen, M.S.T., Lawley, C.T., Karlsson, E.K., Bannasch, D., Vilà, C., Lohi, H., Galibert, F., Fredholm, M., Häggström, J., Hedhammar, A., André, C., Lindblad-Toh, K., Hitte, C. & Webster, M.T., 2011. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. *PLoS Genetics*, 7(10), p.e1002316.

Vedder, O., 2012. Individual birds advance offspring hatching in response to increased temperature after the start of laying. *Oecologia*, 170(3), pp.619–28.

Vedder, O., Bouwhuis, S. & Sheldon, B.C., 2013. Quantitative assessment of the importance of phenotypic plasticity in adaptation to climate change in wild bird populations. *PLoS Biology*, 11(7), p.e1001605.

Venables, W.N. & Ripley, B.D., 2002. *Modern Applied Statistics with S* Fourth Edi., Springer, New York.

Venuprasad, R., Dalid, C.O., Del Valle, M., Zhao, D., Espiritu, M., Sta Cruz, M.T., Amante, M., Kumar, A. & Atlin, G.N., 2009. Identification and characterization of large-effect quantitative trait loci for grain yield under lowland drought stress in rice using bulk-segregant analysis. *Theoretical and Applied Genetics*, 120(1), pp.177–90.

Verhulst, S., Perrins, C.M. & Riddington, R., 1997. Natal Dispersal of Great Tits in a Patchy Environment. *Ecology*, pp.864–872.

Via, S., Gomulkiewicz, R., De Jong, G., Scheiner, S.M., Schlichting, C.D. & Van Tienderen, P.H., 1995. Adaptive phenotypic plasticity: consensus and controversy. *Trends in Ecology & Evolution*, 10(5), pp.212–217.

Vidyasagar, T.R., 1999. Climate change related to egg-laying trends. *Nature*, 399, pp.2–3.

Vinkhuyzen, A.A.E., Pedersen, N.L., Yang, J., Lee, S.H., Magnusson, P.K.E., Iacono, W.G., McGue, M., Madden, P.A.F., Heath, A.C., Luciano, M., Payton, A., Horan, M., Ollier, W., Pendleton, N., Deary, I.J., Montgomery, G.W., Martin, N.G., Visscher, P.M. & Wray, N.R., 2012. Common SNPs explain some of the variation in the personality dimensions of neuroticism and extraversion. *Translational Psychiatry*, 2, p.e102.

Vinkhuyzen, A.A.E., Wray, N.R., Yang, J., Goddard, M.E. & Visscher, P.M., 2013. Estimation and partition of heritability in human populations using whole-genome analysis methods. *Annual Review of Genetics*, 47, pp.75–95.

Visscher, P.M., 2008. Sizing up human height variation. *Nature Genetics*, 40(5), pp.489–90.

Visscher, P.M., 2009. Whole genome approaches to quantitative genetics. *Genetica*, 136(2), pp.351–8.

Visscher, P.M., Macgregor, S., Benyamin, B., Zhu, G., Gordon, S., Medland, S., Hill, W.G., Hottenga, J.-J., Willemsen, G., Boomsma, D.I., Liu, Y.-Z., Deng, H.-W., Montgomery, G.W. & Martin, N.G., 2007. Genome partitioning of genetic variation for height from 11,214 sibling pairs. *American Journal of Human Genetics*, 81(5), pp.1104–10.

Visscher, P.M., Medland, S.E., Ferreira, M.A.R., Morley, K.I., Zhu, G., Cornes, B.K., Montgomery, G.W. & Martin, N.G., 2006. Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *PLoS Genetics*, 2(3), p.e41.

Visser, M.E., 2008. Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings. Biological sciences / The Royal Society*, 275(1635), pp.649–59.

Visser, M.E., Adriaensen, F., Van Balen, J.H., Blondel, J., Dhondt, A. a, Van Dongen, S., Du Feu, C., Ivankina, E. V, Kerimov, A.B., De Laet, J., Matthysen, E., McCleery, R., Orell, M. & Thomson, D.L., 2003. Variable responses to large-scale climate change in European Parus populations. *Proceedings. Biological sciences / The Royal Society*, 270(1513), pp.367–72.

Visser, M.E. & Both, C., 2005. Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings. Biological sciences / The Royal Society*, 272(1581), pp.2561–9.

Visser, M.E., Both, C. & Lambrechts, M.M., 2004. Global Climate Change Leads to Mistimed Avian Reproduction. *Advances in Ecological Research*, 35, pp.89–110.

Visser, M.E., Holleman, L.J.M. & Caro, S.P., 2009. Temperature has a causal effect on avian timing of reproduction. *Proceedings. Biological sciences / The Royal Society*, 276(1665), pp.2323–31.

Visser, M.E., Holleman, L.J.M. & Gienapp, P., 2006. Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia*, 147(1), pp.164–72.

Visser, M.E., te Marvelde, L. & Lof, M.E., 2011. Adaptive phenological mismatches of birds and their food in a warming world. *Journal of Ornithology*, 153(S1), pp.75–84.

Visser, M.E., Schaper, S. V., Holleman, L.J.M., Dawson, A., Sharp, P., Gienapp, P. & Caro, S.P., 2011. Genetic variation in cue sensitivity involved in avian timing of reproduction. *Functional Ecology*, 25(4), pp.868–877.

Visser, M.E., van Noordwijk, A.J., Tinbergen, J. & Lessells, C., 1998. Warmer springs lead to mistimed reproduction in great tits (Parus major). *Proceedings of the National Academy of Sciences of the United States of America*, 265(1408), pp.1867–1870.

Vors, L.S. & Boyce, M.S., 2009. Global declines of caribou and reindeer. *Global Change Biology*, 15(11), pp.2626–2633.

Votier, S.C., Hatchwell, B.J., Mears, M. & Birkhead, T.R., 2009. Changes in the timing of egg-laying of a colonial seabird in relation to population size and environmental conditions. *Marine Ecology Progress Series*, 393, pp.225–233.

Walther, G., Post, E., Convey, P., Menzel, A., Parmesank, C., Beebee, T.J.C., Fromentin, J., I, O.H. & Bairlein, F., 2002. Ecological responses to recent climate change. *Nature*, 416, pp.389–395.

Wang, W.Y.S., Barratt, B.J., Clayton, D.G. & Todd, J.A., 2005. Genome-wide association studies: theoretical and practical concerns. *Nature reviews. Genetics*, pp.109–118.

Watanuki, Y., Ito, M., Deguchi, T. & Minobe, S., 2009. Climate-forced seasonal mismatch between the hatching of rhinoceros auklets and the availability of anchovy. *Marine Ecology Progress Series*, 393, pp.259–271.

Watson, C.T., Roussos, P., Garg, P., Ho, D.J., Azam, N., Katsel, P.L., Haroutunian, V. & Sharp, A.J., 2016. Epigenetic Modification and Its Role in Alzheimer's Disease. *Genome Medicine*, 8(1), pp.1–14.

Watts, C.H.S., 1968. The Foods Eaten by Wood Mice (Apodemus sylvaticus) and Bank Voles (Clethrionomys glareolus) in Wytham Woods, Berkshire. *Journal of Animal Ecology*, 37(1), pp.25–41.

Weber, J.N., Peterson, B.K. & Hoekstra, H.E., 2013. Discrete genetic modules are responsible for complex burrow evolution in Peromyscus mice. *Nature*, 493(7432), pp.402–5.

Weedon, M.N., Lango, H., Lindgren, C.M., Wallace, C., Evans, D.M., Mangino, M., Freathy, R.M., Perry, J.R.B., Stevens, S., Hall, A.S., Samani, N.J., Shields, B., Prokopenko, I., Farrall, M., Dominiczak, A., Johnson, T., Bergmann, S., Beckmann, J.S., Vollenweider, P., Waterworth, D.M., Mooser, V., Palmer, C.N.A., Morris, A.D., Ouwehand, W.H., Zhao, J.H., Li, S., Loos, R.J.F., Barroso, I., Deloukas, P., Sandhu, M.S., Wheeler, E., Soranzo, N., Inouye, M., Wareham, N.J., Caulfield, M., Munroe, P.B., Hattersley, A.T., McCarthy, M.I. & Frayling, T.M., 2008. Genome-wide association analysis identifies 20 loci that influence adult height. *Nature Genetics*, 40(5), pp.575–83.

Weis, A.E. & Gorman, W.L., 1990. Measuring selection on reaction norms : an exploration of the Eurosta-solidago system. *Evolution*, 44(4), pp.820–831.

Wenzel, M.A., James, M.C., Douglas, A. & Piertney, S.B., 2015. Genome-wide association and genome partitioning reveal novel genomic regions underlying variation in gastrointestinal nematode burden in a wild bird. *Molecular Ecology*, 24(16), pp.4175–4192.

Wiggins, D.A., 1991. Natural Selection on Body Size and Laying Date in the Tree Swallow. *Evolution*, 45(5), pp.1169–1174.

Wikelski, M., Hau, M. & Wingfield, J.C., 2000. Seasonality of reproduction in a neotropical rain forest bird. *Ecology*, 81(9), pp.2458–2472.

Wilkin, T. a, Garant, D., Gosler, A.G. & Sheldon, B.C., 2006. Density effects on life-history traits in a wild population of the great tit Parus major: analyses of long-term data with GIS techniques. *The Journal of Animal Ecology*, 75(2), pp.604–15.

Wilkin, T.A., King, L.E. & Sheldon, B.C., 2009. Habitat quality, nestling diet, and provisioning behaviour in great tits Parus major. *Journal of Avian Biology*, 40(2), pp.135–145.

Wilkin, T.A., Perrins, C.M. & Sheldon, B.C., 2007. The use of GIS in estimating spatial variation in habitat quality: a case study of lay date in the great tit Parus major. *Ibis*, 149, pp.110–118.

Wilkin, T.A. & Sheldon, B.C., 2009. Sex differences in the persistence of natal environmental effects on life histories. *Current Biology*, 19(23), pp.1998–2002.

Williams, T.D., Bourgeon, S., Cornell, A., Ferguson, L., Fowler, M., Fronstin, R.B. & Love, O.P., 2015. Mid-winter temperatures, not spring temperatures, predict breeding phenology in the European starling Sturnus vulgaris. *Royal Society Open Science*, 2(1), p.140301.

Wilson, A.J., Gelin, U., Perron, M.-C. & Réale, D., 2009. Indirect genetic effects and the evolution of aggression in a vertebrate system. *Proceedings. Biological sciences / The Royal Society*, 276(1656), pp.533–41.

Wilson, A.J., Réale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C. a, Kruuk, L.E.B. & Nussey, D.H., 2010. An ecologist's guide to the animal model. *The Journal of Animal Ecology*, 79(1), pp.13–26.

Winkler, D.W., Dunn, P.O. & McCulloch, C.E., 2002. Predicting the effects of climate change on avian life-history traits. *Proceedings of the National Academy of Sciences of the United States of America*, 99(21), pp.13595–9.

Wolff, G.L., Kodell, R.L., Moore, S.R. & Cooney, C.A., 1998. Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. *FASEB J*, 12(11), pp.949–57.

Wood, A.R., Esko, T., Yang, J., Vedantam, S., Pers, T.H., Gustafsson, S., Chu, A.Y., Estrada, K., Luan, J., Kutalik, Z., Amin, N., Buchkovich, M.L., Croteau-Chonka, D.C., Day, F.R., Duan, Y., Fall, T., Fehrmann, R., Ferreira, T., Jackson, A.U., Karjalainen, J., Lo, K.S., Locke, A.E., Mägi, R., Mihailov, E., Porcu, E., Randall, J.C., Scherag, A., Vinkhuyzen, A.A.E., Westra, H.-J., Winkler, T.W., Workalemahu, T., Zhao, J.H., Absher, D., Albrecht, E., Anderson, D., Baron, J., Beekman, M., Demirkan, A., Ehret, G.B., Feenstra, B., Feitosa, M.F., Fischer, K., Fraser, R.M., Goel, A., Gong, J., Justice, A.E., Kanoni, S., Kleber, M.E., Kristiansson, K., Lim, U., Lotay, V., Lui, J.C., Mangino, M., Leach, I.M., Medina-Gomez, C., Nalls, M.A., Nyholt, D.R., Palmer, C.D., Pasko, D., Pechlivanis, S., Prokopenko, I., Ried, J.S., Ripke, S., Shungin, D., Stancáková, A., Strawbridge, R.J., Sung, Y.J., Tanaka, T., Teumer, A., Trompet, S., van der Laan, S.W., van Setten, J., Van Vliet-Ostaptchouk, J. V, Wang, Z., Yengo, L., Zhang, W., Afzal, U., Arnlöv, J., Arscott, G.M., Bandinelli, S., Barrett, A., Bellis, C., Bennett, A.J., Berne, C., Blüher, M., Bolton, J.L., Böttcher, Y., Boyd, H.A., Bruinenberg, M., Buckley, B.M., Buyske, S., Caspersen, I.H., Chines, P.S., Clarke, R., Claudi-Boehm, S., Cooper, M., Daw, E.W., De Jong, P.A., Deelen, J., Delgado, G., Denny, J.C., Dhonukshe-Rutten, R., Dimitriou, M., Doney, A.S.F., Dörr, M., Eklund, N., Eury, E., Folkersen, L., Garcia, M.E., Geller, F., Giedraitis, V., Go, A.S., Grallert, H.,

Grammer, T.B., Gräßler, J., Grönberg, H., de Groot, L.C.P.G.M., Groves, C.J., Haessler, J., Hall, P., Haller, T., Hallmans, G., Hannemann, A., Hartman, C.A., Hassinen, M., Hayward, C., Heard-Costa, N.L., Helmer, Q., Hemani, G., Henders, A.K., Hillege, H.L., Hlatky, M.A., Hoffmann, W., Hoffmann, P., Holmen, O., Houwing-Duistermaat, J.J., Illig, T., Isaacs, A., James, A.L., Jeff, J., Johansen, B., Johansson, A., Jolley, J., Juliusdottir, T., Junttila, J., Kho, A.N., Kinnunen, L., Klopp, N., Kocher, T., Kratzer, W., Lichtner, P., Lind, L., Lindström, J., Lobbens, S., Lorentzon, M., Lu, Y., Lyssenko, V., Magnusson, P.K.E., Mahajan, A., Maillard, M., McArdle, W.L., McKenzie, C.A., McLachlan, S., McLaren, P.J., Menni, C., Merger, S., Milani, L., Moayyeri, A., Monda, K.L., Morken, M.A., Müller, G., Müller-Nurasyid, M., Musk, A.W., Narisu, N., Nauck, M., Nolte, I.M., Nöthen, M.M., Oozageer, L., Pilz, S., Rayner, N.W., Renstrom, F., Robertson, N.R., Rose, L.M., Roussel, R., Sanna, S., Scharnagl, H., Scholtens, S., Schumacher, F.R., Schunkert, H., Scott, R.A., Sehmi, J., Seufferlein, T., Shi, J., Silventoinen, K., Smit, J.H., Smith, A.V., Smolonska, J., Stanton, A. V, Stirrups, K., Stott, D.J., Stringham, H.M., Sundström, J., Swertz, M.A., Syvänen, A.-C., Tayo, B.O., Thorleifsson, G., Tyrer, J.P., van Dijk, S., van Schoor, N.M., van der Velde, N., van Heemst, D., van Oort, F.V.A., Vermeulen, S.H., Verweij, N., Vonk, J.M., Waite, L.L., Waldenberger, M., Wennauer, R., Wilkens, L.R., Willenborg, C., Wilsgaard, T., Wojczynski, M.K., Wong, A., Wright, A.F., Zhang, Q., Arveiler, D., Bakker, S.J.L., Beilby, J., Bergman, R.N., Bergmann, S., Biffar, R., Blangero, J., Boomsma, D.I., Bornstein, S.R., Bovet, P., Brambilla, P., Brown, M.J., Campbell, H., Caulfield, M.J., Chakravarti, A., Collins, R., Collins, F.S., Crawford, D.C., Cupples, L.A., Danesh, J., de Faire, U., den Ruijter, H.M., Erbel, R., Erdmann, J., Eriksson, J.G., Farrall, M., Ferrannini, E., Ferrières, J., Ford, I., Forouhi, N.G., Forrester, T., Gansevoort, R.T., Gejman, P. V, Gieger, C., Golay, A., Gottesman, O., Gudnason, V., Gyllensten, U., Haas, D.W., Hall, A.S., Harris, T.B., Hattersley, A.T., Heath, A.C., Hengstenberg, C., Hicks, A.A., Hindorff, L.A., Hingorani, A.D., Hofman, A., Hovingh, G.K., Humphries, S.E., Hunt, S.C., Hypponen, E., Jacobs, K.B., Jarvelin, M.-R., Jousilahti, P., Jula, A.M., Kaprio, J., Kastelein, J.J.P., Kayser, M., Kee, F., Keinanen-Kiukaanniemi, S.M., Kiemeney, L.A., Kooner, J.S., Kooperberg, C., Koskinen, S., Kovacs, P., Kraja, A.T., Kumari, M., Kuusisto, J., Lakka, T.A., Langenberg, C., Le Marchand, L., Lehtimäki, T., Lupoli, S., Madden, P.A.F., Männistö, S., Manunta, P., Marette, A., Matise, T.C., McKnight, B., Meitinger, T., Moll, F.L., Montgomery, G.W., Morris, A.D., Morris, A.P., Murray, J.C., Nelis, M., Ohlsson, C., Oldehinkel, A.J., Ong, K.K., Ouwehand, W.H., Pasterkamp, G., Peters, A., Pramstaller, P.P., Price, J.F., Qi, L., Raitakari, O.T., Rankinen, T., Rao, D.C., Rice, T.K., Ritchie, M., Rudan, I., Salomaa, V., Samani, N.J., Saramies, J., Sarzynski, M.A., Schwarz, P.E.H., Sebert, S., Sever, P., Shuldiner, A.R., Sinisalo, J., Steinthorsdottir, V., Stolk, R.P., Tardif, J.-C., Tönjes, A., Tremblay, A., Tremoli, E., Virtamo, J., Vohl, M.-C., Amouyel, P., Asselbergs, F.W., Assimes, T.L., Bochud, M., Boehm, B.O., Boerwinkle, E., Bottinger, E.P., Bouchard, C., Cauchi, S., Chambers, J.C., Chanock, S.J., Cooper, R.S., de Bakker, P.I.W., Dedoussis, G., Ferrucci, L., Franks, P.W., Froguel, P., Groop, L.C., Haiman, C.A., Hamsten, A., Hayes, M.G., Hui, J., Hunter, D.J., Hveem, K., Jukema, J.W., Kaplan, R.C., Kivimaki, M., Kuh, D., Laakso, M., Liu, Y., Martin, N.G., März, W., Melbye, M., Moebus, S., Munroe, P.B., Njølstad, I., Oostra, B.A., Palmer, C.N.A., Pedersen, N.L., Perola, M., Pérusse, L., Peters, U., Powell, J.E., Power, C., Quertermous, T., Rauramaa, R., Reinmaa, E., Ridker, P.M., Rivadeneira, F., Rotter, J.I., Saaristo, T.E., Saleheen, D., Schlessinger, D., Slagboom, P.E., Snieder, H., Spector, T.D., Strauch, K., Stumvoll, M., Tuomilehto, J., Uusitupa, M., van der Harst, P., Völzke, H., Walker, M., Wareham, N.J., Watkins, H., Wichmann, H.-E., Wilson, J.F., Zanen, P., Deloukas, P., Heid, I.M., Lindgren, C.M., Mohlke, K.L., Speliotes, E.K., Thorsteinsdottir, U., Barroso, I., Fox, C.S., North, K.E., Strachan, D.P., Beckmann, J.S., Berndt, S.I., Boehnke, M., Borecki, I.B.,

McCarthy, M.I., Metspalu, A., Stefansson, K., Uitterlinden, A.G., van Duijn, C.M., Franke, L., Willer, C.J., Price, A.L., Lettre, G., Loos, R.J.F., Weedon, M.N., Ingelsson, E., O'Connell, J.R., Abecasis, G.R., Chasman, D.I., Goddard, M.E., Visscher, P.M., Hirschhorn, J.N. & Frayling, T.M., 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature Genetics*, 46(11), pp.1173 – 1186.

Wray, N.R., 2005. Allele frequencies and the r2 measure of linkage disequilibrium: impact on design and interpretation of association studies. *Twin Research and Human Genetics*, 8(2), pp.87–94.

Wright, D., Nakamichi, R., Krause, J. & Butlin, R.K., 2006. QTL analysis of behavioral and morphological differentiation between wild and laboratory zebrafish (Danio rerio). *Behavior Genetics*, 36(2), pp.271–84.

Wright, S., 1943. Isolation by Distance. *Genetics*, 28(2), pp.114–38.

Xu, S., 2003. Estimating polygenic effects using markers of the entire genome. *Genetics*, 163(2), pp.789–801.

Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., Martin, N.G., Montgomery, G.W., Goddard, M.E. & Visscher, P.M., 2010. Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42(7), pp.565–9.

Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M., 2011. GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, 88(1), pp.76–82.

Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M., 2013. Genome-wide complex trait analysis (GCTA): methods, data analyses, and interpretations. *Methods in Molecular Biology*, 1019, pp.215–36.

Yang, J., Lee, T., Kim, J., Cho, M.-C., Han, B.-G., Lee, J.-Y., Lee, H.-J., Cho, S. & Kim, H., 2013. Ubiquitous polygenicity of human complex traits: genome-wide analysis of 49 traits in Koreans. *PLoS Genetics*, 9(3), p.e1003355.

Yang, J., Manolio, T.A., Pasquale, L.R., Boerwinkle, E., Caporaso, N., Cunningham, J.M., de Andrade, M., Feenstra, B., Feingold, E., Hayes, M.G., Hill, W.G., Landi, M.T., Alonso, A., Lettre, G., Lin, P., Ling, H., Lowe, W., Mathias, R.A., Melbye, M., Pugh, E., Cornelis, M.C., Weir, B.S., Goddard, M.E. & Visscher, P.M., 2011. Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics*, 43(6), pp.519–25.

Yeh, P.J. & Price, T.D., 2004. Adaptive phenotypic plasticity and the successful colonization of a novel environment. *The American Naturalist*, 164(4), pp.531–542.

Yi, G., Shen, M., Yuan, J., Sun, C., Duan, Z., Qu, L., Dou, T., Ma, M., Lu, J., Guo, J., Chen, S., Qu, L., Wang, K. & Yang, N., 2015. Genome-wide association study dissects genetic architecture underlying longitudinal egg weights in chickens. *BMC Genomics*, 16(1), p.746.

Zhu, Y.-P., Feng, Y., Liu, T. & Wu, Y.-C., 2015. Epigenetic Modification and Its Role in Alzheimer's Disease. *Integrative Medicine International*, 2, pp.63–72.


BROOKLYN ROAD