Exploring adaptation in the introduced house sparrow to Australian climates and environments

Samuel C. Andrew, BSc. (Hons)

Department of Biological Sciences

Faculty of Science and Engineering

Macquarie University

NSW Australia

MACQUARIE University

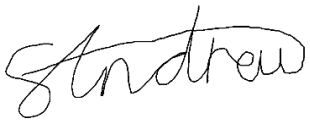
This thesis is presented for the degree of Doctor of Philosophy

Submitted for examination: 18th of August 2017

Declaration

I do declare that this thesis constitutes my original work, and has not been submitted or accepted for the award of a higher degree at another institution. No persons have contributed to this work that have not been given credit, and all the work of previous scholars influencing this work are cited in the appropriate manner.

The protocols to collect blood samples and morphological measurements from House Sparrows used in this thesis were approved by ethics committees at Macquarie University (ARA 2014/248) and the approval of Otago University (Animal ethics reference number: 87/08). Zebra Finch morphological measurements were collected under the approval of Macquarie University (Animal Ethics Approval: ARA2007/038, ARA 2013/029).



Samuel C. Andrew

17th August 2017

‘‘Dr. Johnson's rule, aim at the eagle though you may only hit a sparrow.’’

— Hon. Dr Davies, 17th December 1864,
South Australian Weekly Chronicle,
Advice for young students.

Dedication

I dedicate this thesis to my parents, Louise and Phillip Andrew, for supporting me in every way possible throughout my education and development as a researcher during my post-graduate studies.

Table of Contents

List of Original Publications

This thesis by publication is based on the following original manuscripts that are at various stages of publication, they are referred to in the text as Chapters Two to Seven (See bookmarks in the PDF version for links to chapters). Original publications are reproduced with permission from their copyright holders.

Two - **Andrew SC** and Griffith SC (2016) Inaccuracies in the history of a well-known introduction: a case study of the Australian House Sparrow (*Passer domesticus*). *Avian Research*, **7**, 9.

Three- **Andrew SC**, Awasthy M, Griffith AD, Nakagawa S and Griffith S C (in review: 07/07/2017) Clinal variation in avian body size is better explained by summer maximum temperatures during development and not cold winter temperatures. *The Auk*.

Four - **Andrew SC**, Hurley LL, Mariette MM and Griffith SC (submitted: 16/08/2017) Higher temperatures during development constrain body size in the zebra finch in the wild and in experimental conditions in the laboratory. *Journal of Evolutionary Biology*.

Five - **Andrew SC**, Awasthy M, Bolton PE, Rollins LA, Nakagawa S and Griffith SC (in review: 16/05/2017) The genetic structure of the introduced house sparrow populations in Australia and New Zealand is consistent with historical descriptions. *Biological Invasions*.

Six - **Andrew SC**, Jensen H, Hagen IJ, Lien S and Griffith SC (in prep) Signatures of genetic adaptation to extremely varied Australian environments in introduced European house sparrows. (target journal *Molecular Ecology*)

Seven- **Andrew SC**, Taylor MP, Lundregan S, Lien S, Jensen H and Griffith SC (in prep) Living in a heavy metal environment: Evidence for selection in the house sparrow to varied urban environments contaminated with trace elements. (target journal *Molecular Ecology*)

List of Contributors

Supervisor	Professor Simon Griffith Department of Biological Sciences Macquarie University, Australia

The following associate supervisors contributed to different parts of the thesis, see table below.

Associate supervisor	Dr. Lee Ann Rollins School of Life and Environmental Sciences Deakin University, Australia (Microsatellite population genetics)
Associate supervisor	Associate Professor Henrik Jensen Centre for Biodiversity Dynamics Department of Biology NTNU, Trondheim, Norway (Genomics)
Associate supervisor	Associate Professor Shinichi Nakagawa School of Biological, Earth and Environmental Sciences University of New South Wales, Australia (Statistics)

Contribution by collaborators and co-authors of the thesis, my own percent contribution in brackets. Collaborator key: SCA – Samuel C. Andrew, SCG – Simon C. Griffith, LAR – Lee Ann Rollins, SN – Shinichi Nakagawa, MA – Monica Awasthy, AG – Amanda Griffith, MMM – Mylene M. Mariette, LLH – Laura L. Hurley, PEB – Peri E. Bolton, HJ – Henrik Jensen, SL – Sarah Lundregan, MPT – Mark P. Taylor.

	CHAPTER ONE	**CHAPTER TWO**	**CHAPTER THREE**	**CHAPTER FOUR**	**CHAPTER FIVE**	**CHAPTER SIX**	**CHAPTER SEVEN**	**CHAPTER EIGHT**
CONCEPTION AND PLANNING	SCA (100%)	SCA (70%), SCG	SCA (70%), SCG	SCA (50%), SCG	SCA (70%), SCG, LAR	SCA (80%), SCH, HJ	SCA (80%), SCG, MPT	SCA (100%)
FIELD WORK	NA	NA	SCA (30%), SCG, AG, MA, SN	SCA (0%), MMM, LLH	SCA (30%), SCG, AG, MA, SN	SCA (40%), SCG, MA	SCA (40%), SCG, MA	NA
SAMPLE/DATA PROCESSING	NA	NA	SCA (100%)	SCA (100%)	SCA (100%)	SCA (100%)	SCA (100%)	NA
DATA ANALYSIS	NA	SCA (100%)	SCA (95%), SN	SCA (95%), SN	SCA (90%), PEB, LAR	SCA (100%)	SCA (100%)	NA
INTERPRETATION AND WRITING	SCA (100%)	SCA (80%), SCG	SCA (80%), SCG	SCA (75%), SCG	SCA (75%), SCG, LAR, PEB	SCA (90%), SCG, HJ	SCA (85%), SCG, MPT, SL	SCA (100%)

Acknowledgements

I would like to thank my primary supervisor Simon Griffith for giving me the incredible opportunity to work in the Griffith Ecology Lab and for his invaluable mentorship and efforts to resource the lab so we can all continue to do research on the birds we love. My secondary supervisor Lee Ann Rollins for helping me to break through into the world of population genetics, something I have wanted to do since undergrad, and for pushing me to be a better researcher and writer. My unofficial statistics supervisor Shinichi Nakagawa for teaching me so many valuable lessons about statistics and how to answer the important questions. Henrik Jensen was also extremely important for making the genomics portion of my PhD possible and has also been a fantastic collaborator and mentor. Finally, my 'field work supervisor' Monica Awasthy for teaching me everything I know about catching sparrows and for travelling with me around Australia.

The Griffith Ecology Lab group has been amazing and I benefited so much from their support at work and friendship outside of work. Firstly, I would like to thank Peri for teaching me everything from how to do lab work to referencing software and everything in between. Daisy and Fumi for helping me with stats and R coding. Lori and Ondi for teaching me bird handling skills. Marianne for proof reading parts of my thesis. Ellie for working in the lab and the field with me on the house sparrow project and all the other members of the lab who have been around during my PhD.

The biological sciences department at Macquarie University has also been an integral part of my development as a researcher during my undergrad studies and into post-grad with the Genes to Geoscience's workshops, department seminars and the R users group. I would also like to make special mention to the mentors and supervisors in the department that helped me get to the stage of starting a PhD: Andrew Barron, Darrell Kemp and Ken Cheng, without them I don't think I could have made it this far. I would also like to thank my fellow PhD students from the department who I have made friends with and shared this experience.

I have really enjoyed doing this research but none of it would have been possible without government funding allocated to Simon Griffith from the Australian Research Council and support from Macquarie University through the Macquarie University Research Excellence Scholarship. In addition to financial support the field work to collect data from house sparrow populations would not have been possible without the countless people around Australia and New Zealand who provided access for us to do our work. I would like to thank Laura Hurley and Mylene Mariette for providing the data to write the fourth chapter of the thesis on the developmental effects of temperature on growth, this work was a great accompaniment to data collected on the house sparrow from the third chapter. I also thank Shinichi Nakagawa for generously supplying house sparrow samples from New Zealand.

Another important part of my PhD was a trip to Norway in June 2016, funded by the Post-Graduate Research Fund (PGRF) at Macquarie University. The trip started a great collaboration which made possible the sixth and seventh chapters of my thesis. I would like to thank my hosts Glenn-Peter Sætre at Oslo University and Henrik Jensen at NTNU who were extremely hospitable and helpful during my trip. I would also like to thank the rest of their lab groups for making the trip such a great learning experience.

Finally, I would like to thank friends and family who have always been there for me in the background. My mother, father, brother, sister and whole extended family have been amazingly supportive throughout my life and this period during my PhD, I am extremely lucky to have all of them. All my friends have also been very supportive during my PhD and never forget to ask me how it is going and when I will be finished. In particular I would like to thank Bryan, Cam, Mitch, Morgen, Alli, Duy and James for their friendship and support.

General Abstract

Invasive species have always been excellent models for understanding local adaptation through their colonization of new and varied environments. The house sparrow (*Passer domesticus*) has been a key model invasive species around the world and in this thesis, I have expanded this research to consider the Australian population across the current distribution over eastern Australia. Given that we believe most sparrows were introduced to Australia from England in the 1860's they have colonized temperate biomes that are like their home range and very different climates in the tropical and arid biomes. I have found that genetic differentiation across this region is relatively strong, and this population structure over a range of climates can allow independent populations to adapt to local climates, over the past 150 years. In line with this idea I have found evidence of significant differentiation in morphology across climatic gradients, in this species. In another passerine species, I demonstrated that this is likely largely explained by phenotypic plasticity. Additionally, using genomic SNP data and genome scan methods I explored links between differentiation in allele frequencies and climatic variation across populations. These models found many significant outlier loci, some of which were linked to genes on the reference genome relating to traits important to local adaptation, such as thermal tolerance, immune function and morphology. In addition to exploring climate as a selective force I consider the effects of heavy metal pollution on urban sparrow populations and found outlier genes that could be relevant to responding to lead contamination. These functional loci provide targets for future research describing the mechanisms behind local adaptation. Understanding how species adapt to different environmental conditions through both phenotypic plasticity and selection will be important for predicting the response of wildlife to the rapid changes in habitat and climate caused by humans.

Chapter One

General Introduction

Invasive species are considered to be one of the most common threatening influences on endangered species and natural ecosystems (Gurevitch & Padilla 2004). An invasive species can be defined as a species which has been introduced, intentionally or due to human activity, outside of its natural range and is successfully reproducing and spreading its distribution (Blackburn *et al.* 2011). Invasive species are often thought of as 'pests' because their population size and density often become extremely high, leading to the exploitation of the ecosystem and the decline of native species or the exploitation of the agriculture industry (Pimentel *et al.* 2005; Saunders *et al.* 2010). Some well known examples of highly damaging invasive species in Australia include rabbits (*Oryctolagus cuniculus,* Cooke & Fenner 2003), Cane toads (*Rhinella marina*, Phillips *et al.* 2007), foxes (*Vulpes vulpes*, Fleming *et al.* 2006) and lantana (*Lantana camara*, Turner & Downey 2010). Invasive species can pose a significant threat to biodiversity. However, many introduced species do not become a serious threat to ecosystems and instead become a regular part of the community (Gurevitch & Padilla 2004). Even in less threatening cases invasive species will always have some impact to the community due to competition for resources (Ehrenfeld 2010) and their inclusion in the food web (Blackburn 2008).

The importance of managing the impacts of invasive species has made them important targets of research investigating invasion biology and more broadly ecology and evolution (Blackburn 2008). Invasive species provide a valuable research opportunity because their introductions act as 'natural' experiments. A natural experiment is a study in which clusters of individuals are exposed to the experimental and control conditions via processes that are out of the control of the investigator. Natural experiments are still considered observational studies because experimental conditions are not controlled like traditional experiments. In invasion biology, the habitat pre-introduction or habitat in unaffected areas acts as a control for studying colonization dynamics and the climate of the source population can act as a control for studies on local adaptation and selection (Lee 2002; Sax *et al.* 2007). The field of invasion biology spans topics from the fundamentals of ecological interactions to modern studies of molecular evolution.

From an evolutionary genetics point of view, invasive species are interesting because they overcome some primary threats to endangered species without any apparent difficulties (Estoup *et al.* 2016). For example, endangered species are often affected by demographic bottlenecks that can result in reduced genetic diversity and inbreeding depression which are detrimental to populations and a major threat to their survival (Frankham 2002, 2010; O'Grady *et al.* 2006). In contrast, invasive species are often affected by genetic bottlenecks when they are introduced but in many cases they still manage to adapt to novel environments

(Dlugosch & Parker 2008; Lima *et al.* 2012; Rollins *et al.* 2013). Many endangered species are also threatened by habitat destruction and climate change (Deutsch & Tewksbury 2008; Hoffmann & Sgrò 2011), while invasive species often live in highly disturbed habitats and across a broad range of climates (Lever 1985, 2005). For example, the house sparrow (*Passer domesticus*) was introduced to one area of North America and spread its distribution to cover a majority of the continent in around 100 generations (Selander & Johnston 1967; Johnston & Selander 1971; Moulton *et al.* 2010). This kind of rapid adjustment to highly varied climates is dependent on the ability of populations to change phenotypes through genetic adaptations and phenotypic plasticity (Lee 2002). How invasive species adapt to drastically different conditions to those of their native habitat is the primary question for researchers studying local adaptation at both genetic and epigenetic levels in invasive species.

Because of its rapid spread across many different habitats, the house sparrow has been used as a model to demonstrate natural selection and local adaptation in invasive species for over 100 years (Anderson 2006; Liebl *et al.* 2015). A landmark study for natural selection investigated house sparrows. Bumpus (1899) compared a collection of 136 house sparrows that had either survived or perished after a severe storm in Rhode Island, USA. He found without the aid of statistical tests that surviving birds were shorter, lighter and less variable than the birds that perished. This indicated that these birds possessed favourable characters to survive the storm, which acted as a selection force. This study was one of the first reported observations of natural selection. Bumpus' data set went on to be one of the most reanalysed data sets in evolutionary biology (Anderson 2006). The results of Bumpus' work have remained mostly unchanged but the conclusions from these data have been debated over many decades (Johnston *et al.* 1972; Lande & Arnold 1983; Buttemer 1992; Pugesek & Tomer 1996; Janzen & Stern 1998). This work was extended on in North America where the house sparrow was introduced in the 1850's and 60's from England and Germany and rapidly expanded its distribution (Grinnell 1919; Wing 1943). Lack (1940) used the house sparrow as a comparison for his work on the Galapagos finch to support the theory of natural selection. The species gave him the opportunity to "test the interesting evolutionary point of whether or not it had changed significantly since its introduction into the United States" (Lack 1940). He found low variation between introduced populations and that introduced populations were larger than the current populations in England but not Germany (Lack 1940). A later study by Calhoun (1947) used museum skins from North America to show house sparrows were larger in colder climates and this pattern was established prior to 1908. This data was collected only a couple of decades after the species had been established in western regions (Wing 1943). A series of studies by Johnston and Selander (Johnston & Selander 1964, 1971; Selander &

Johnston 1967) collected extremely high quality and comprehensive morphological data that demonstrated rapid differentiation between house sparrow populations within a period of around 100 generations. They interpreted these results as showing the emergence of new 'house sparrow races' in a period of around 100 years, a phenomenon which was previously predicted to take around 4000 years (Johnston & Selander 1964; Selander & Johnston 1967).

Since then, the house sparrow has been sampled across nearly all of its native and introduced distribution to examine the effect of climate on morphological variation (Murphy 1985). These studies have used a mixture of skins, skeletons and live birds to collect measurements to support Bergmann's rule (Translation in James 1970) which predicts that larger 'races' of a species will occur in colder climates. Bergmann's rule is often shown by a cline in body size relative to climate or latitude. Bergmann's rule has been demonstrated in the introduced populations of North America (Calhoun 1947; Johnston & Selander 1973; Murphy 1985). However, this pattern is not supported in the introduced populations of New Zealand (Baker 1980) and South America (Johnston & Selander 1973) possibly due to insufficient variation in temperature across the sampling area (Murphy 1985). Using skeletal measurements and latitude as a predictor of seasonality, Bergmann's rule was initially not supported across the native range in Europe (Johnston 1973). Later, the data were reanalysed with winter temperature used as a predictor of seasonality and this factor was positively correlated with body size, revealing consistent support for Bergmann's rule in both North America and Europe (Murphy 1985).

Winter temperature is predicted to be the strongest force driving the morphological differentiation between house sparrow populations (Lowther 1977). Selective pressure resulting from winter temperature has also been shown to correlate with thermal conductance which is inversely related to measures of climatic variation (Blem 1974). This correlation illustrates that birds from colder climates lose body heat at slower rates due to better insulation (Blem 1974). A number of studies have monitored selection acting on house sparrow populations during winter periods in the colder climates where selection is strongest (Parkin 1987). Two studies on birds collected in Kansas, USA found evidence for colder winter temperatures selecting for larger males and smaller females (Johnston & Fleischer 1981; Fleischer & Johnston 1984). To support the hypothesis that large males and smaller females have higher survival during the winter months the smaller males and larger females should be in worse condition (i.e. lower fat deposits) during the winter months but this was not supported by empirical evidence (Fleischer & Murphy 1992). This study instead found that average fat deposits were positively correlated with body size in both sexes and positively correlated with allozyme heterozygosis in fat tissue (Fleischer & Murphy 1992). There is

however a sexual size dimorphism in sparrows in which males are larger than females, that is consistent across the distribution (Packard 1967; Johnston & Selander 1973; Murphy 1985).

Another pattern of morphological differentiation in house sparrow populations across hot and cold climates is consistent with Allen's rule (Allen 1877). This rule predicts clinal variation in core-to-limb ratios that trends toward longer limbs and smaller core size in warmer climates. This pattern has been demonstrated using core-to-limb ratio measurements in a number of species (Mayr 1956; Symonds & Tattersall 2010). Strong support for Allen's rule has also been found in European house sparrow populations where sparrows near the Mediterranean have smaller core-to-limb ratios which increases the surface area for heat loss compared to northern populations (Johnston 1969, 1973). This selection affects both limb length and core size and this pattern was demonstrated using skeletal measurements (Johnston 1973). Allen's rule is supported also in North American populations but the pattern is secondary to the effect of larger body size in colder climates (Johnston & Selander 1971). Fleischer and Johnston (1982) compared samples collected in autumn and the next spring in North America and found that there was an increased core-to-limb ratio that causes a lower surface-to-volume ratio which, in turn, reduces thermal conductivity (Blem 1974; Fleischer & Johnston 1982). The selective forces and mechanisms behind these clinal patterns are still not fully understood even after decades of extensive research (Ashton 2002; Meiri & Dayan 2003; Symonds & Tattersall 2010; Teplitsky & Millien 2014).

Although variation in morphology is well demonstrated for this species it does not prove that this variation has an adaptive advantage and is a result of natural selection and not simple a by-product of environmental effects on phenotype or adaptive phenotypic plasticity. Future research should aim to partition the variation explained by genetic effects from the environmental effects on phenotypic plasticity. Morphological traits have high estimates of narrow sense heritability in intensively studied sparrow populations from Norway (Jensen *et al.* 2003) and there have also been genetic associations with variation in morphology in the same populations (Silva *et al.* 2017). This indicates that these global patterns in morphological variation could be linked to genetic variation. However, low genetic diversity in introduced populations as a result of small founder populations can limit the potential of populations to adapt to new climates (Hoffmann & Sgrò 2011).

The house sparrow has again been an important model species for studying founder effects and inbreeding depression in introduced populations. One of the first studies to measure allozyme diversity in a bird species used house sparrow populations from North America, almost 100 years after the species was introduced (Klitz 1973). The study found that

the populations were monomorphic and had no variation for 19 out of the 20 allozymes surveyed (Klitz 1973). Allozyme hetrozygosity can be highly variable and sometimes very low (Nevo 1978). Later studies on the house sparrow have found no variation in other allozyme loci (e.g. in Norway (Väisänen & Lehväslaiho 1984) and England (Cole & Parkin 1981)). Klitz (1973) concluded that his result was due to the enzymes being highly conserved in sparrows because even if a rare allele had a frequency of q = 0.05 the allele would still have a 98.5% chance of being represented in a founding population of only 20 individuals. Therefore, bottlenecks suffered by introduced house sparrows should change allele frequency and not necessarily wipe out all rare alleles. This hypothesis was supported by a comparative study in South Australia on two native species and the introduced house sparrow, the sparrow had the highest estimates of alloyzme hetrozygosity over the native species (Manwell & Baker 1975). Using a more polymorphic set of allozymes, Cole and Parkin (1981) were unable to find genetic differentiation across the connected landscape of the East midlands in England (Parkin & Cole 1984). Using the same allozymes, differences between England, Western Europe, Australia and New Zealand were detected (Parkin & Cole 1985). This differentiation was possibly due to the introduced populations of Australia and New Zealand having fewer rare alleles. The New Zealand populations also had lower levels of hetrozygosity. Surprisingly the highest level of genetic differentiation within a region was in Australia. Differentiation between house sparrow populations has also been linked to population size. Significant differentiation was detected between five small rural populations in Norway but, not between 6 larger populations (Bjordal *et al.* 1986). However, allozymes can have low amounts of variation, which makes identifying independent populations difficult.

Population genetics research on house sparrows expanded rapidly with the development of species specific microsatellite primers (Neumann & Wetton 1996; Griffith *et al.* 1999, 2007; Dawson *et al.* 2012) and other conserved passerine microsatellite markers (Richardson *et al.* 2000; Dawson *et al.* 2006, 2010). These markers have been used in house sparrows to estimate: effective population size (Engen *et al.* 2007; Baalsrud *et al.* 2014), inbreeding coefficients (Jensen *et al.* 2007; Billing *et al.* 2012) and kin structure (Liker *et al.* 2009; Vangestel *et al.* 2011) with the majority of the population genetics work on house sparrows focused on population structure (Liebl *et al.* 2015). The previously described population structure can be partly explained by the species biology. Sparrows have a highly fragmented distribution because they only exist with human settlements and the species also has a low natural dispersal ability and is very philopatric (Anderson 2006). However, with human influence, sparrows can disperse long distances (Schrey *et al.* 2014). In Finland,

limited dispersal has maintained connectivity between populations but the sea between Finland and Sweden acts as a dispersal barrier that has resulted in genetic population structure between the two populations (Kekkonen *et al.* 2011). This result is analogous to the alloyzme results from Norway, were they also found strong differentiation between Norway and mainland Europe (Bjordal *et al.* 1986). However, population structure has also been detected over a fine scale with microsatellites (Hole *et al.* 2002; Brommer *et al.* 2014). A recent study on mainland and island populations in Norway found the more isolated island population were more strongly affected by genetic drift and bottlenecks (Jensen *et al.* 2013), as we would expect in a fragmented meta-population. In a broader study on populations from North America, Europe and Kenya, differentiation between the invasive populations from America and the source populations in Europe was not large but more recently established populations did have lower genetic diversity (Schrey *et al.* 2011). The Kenyan population was totally distinct from the other populations with lower allelic richness but also higher private allele richness indicating this invasive population had a source population outside of Europe (Schrey *et al.* 2011, 2014). In general the house sparrow shows genetic differentiation across moderate distances and this has been compared to morphological differentiation in a limited number of cases (Kekkonen *et al.* 2012; Lima *et al.* 2012; Brommer *et al.* 2014). The next stage for genetic research on this model invasive species is to use genomic methods to understand the traits that are important to local adaptation and selection in the house sparrow.

A number of genomic resources have recently been developed for the house sparrow including a transcriptome (Ekblom *et al.* 2013), a reference genome (Elgvin *et al.* 2017), a 10K SNP array (Hagen *et al.* 2013) and even more recently a 200K SNP array has been developed for the house sparrow (Lundregan *et al. in prep*). These SNP arrays allow for thousands of Single Nucleotide Polymorphisms (SNPs) to be consistently genotyped across the house sparrow's genome to observe changes in allele frequency that could be the result of selection. A number of analytical techniques have recently been developed to use genomic data to identify outlier loci that are physically linked to regions of the genome being affected by selection (de Villemereuil *et al.* 2014; Lotterhos & Whitlock 2014, 2015; Rellstab *et al.* 2015). Outlier loci show signatures of selection that are characterised by having much higher than expected differentiation between populations or variation in allele frequencies that is highly correlated with environmental variation (Rellstab *et al.* 2015; Francois *et al.* 2016; Hoban *et al.* 2016). These outlier SNPs can then be associated with protein coding regions on the house sparrow's reference genome. These current technological and analytical developments have brought us to a critical period for expanding our understanding of how species adapt to local environments. Understanding the molecular mechanisms behind local

adaptation to novel environments is fundamental to research on invasion biology and continues the work on the sparrow that was started by the pioneering work by Bumpus (1899) over a century ago.

Aims and Thesis Outline

The house sparrow has been used as a model to study many questions in evolutionary biology already and this thesis aims to add to this body of research. Introduced populations of house sparrows in Australia have been largely unstudied, and this is a major missing element to the global study of this model species. To fill this gap, I have collected comprehensive data from this region. Additionally, despite the extensive population and evolutionary genetics work investigating the house sparrow, modern genomic techniques have been little used in these investigations. Genomic approaches provide a new opportunity to identify and understand the molecular mechanisms behind local adaptation. The ultimate goal of this thesis is to identify some of the possible reasons why the house sparrow has been successful across a broad range of environments in Australia.

Chapter Two: Detailed descriptions of introduction history are a fundamental part of a species natural history and will be beneficial to studies using invasive species as models for colonization biology and local adaptation. Many sources have reported details about the house sparrow's introduction to Australia but none are comprehensive and many sources report contradicting details. My first aim is to make a systematic and comprehensive review of the house sparrows introduction history to create a complete picture of the events comprising the introduction and to address conflicting reports.

Chapter Three and Four: The house sparrow has been used as a global model to look at morphological differentiation that is consistent with Bergmann's rule. However, the house sparrow in Australia has not been surveyed to see if it fits with the global trend. It has also not been resolved if this pattern of larger body size in colder climates is a result of selection to winter temperatures or phenotypic plasticity during development. A growing number of studies suggest that the ultimate body size attained by avian species is influence by temperatures experienced during development. We aim to test this hypothesis using morphological data from house sparrow populations across Australia and New Zealand (Chapter Three). Observational data from across the distribution of the sparrow replicates classic studies but there is a lack of experimental data looking at the effects of temperature on avian development. Using the zebra finch (*Taeniopygia guttata*) as a model system we also analyse morphological data collected during development from finches in a wild population and from a temperature controlled experiment (Chapter Four). The zebra finch was used to

test whether the temperature during development directly affected body size because this species can be more easily breed in nest boxes and captivity than Australian sparrows.

Chapter Five: Despite the global effort to study the species, house sparrow population structure has not been studied in Australia and New Zealand using microsatellite data. We aim to describe the population structure and identify how it relates to the species introduction history. In this region, we aim to use one of the most comprehensive sampling designs applied to house sparrow populations across a large geographic scale. This sampling aims to give a high resolution to describe population structure and connectivity.

Chapter Six: House sparrows have colonized a range of climatic conditions in Australia but the genetic mechanisms, if any, the species uses to adapt to varied climates is still not understood. We use genomic SNP data from sampling localities across a range of climates to identify regions of the genome that are putative targets of selection, using genome scan analyses.

Chapter Seven: Sparrows live in urban environments that are often contaminated by metal pollutants such as lead. In Australia, urban environments are generally affected by contamination from lead paint and petrol but this contamination is minor in comparison to contamination observed in some mining towns. It is still not known if sparrows can adapt to environmental contamination using genetic mechanisms. We use a genome scan approach known as an ecological association analysis to identify regions of the genome that show higher than expected differentiation that covaries with variation in lead pollution across Australian sampling sites.

Chapter Eight: Is a general discussion summarising our results and the links between our findings from different chapters. as well as discussing a number of targets for future research.

References

Allen JA (1877) The influence of physical conditions in the genesis of species. *Radical Review*, **1**, 108–140.

Anderson TR (2006) Biology of the Ubiquitous House Sparrow: from Genes to Populations. In: *University Press New York.*, Oxford University Press,Inc, New York.

Ashton KG (2002) Patterns of within-species body size variation of birds: Strong evidence for Bergmann's rule. *Global Ecology and Biogeography*, **11**, 505–523.

Baalsrud HT, Saether B-E, Hagen IJ *et al.* (2014) Effects of population characteristics and structure on estimates of effective population size in a house sparrow metapopulation. *Molecular Ecology*, **23**, 2653–2668.

Baker AJ (1980) Morphometric differentiation in New Zealand populations of the house sparrow (*Passer domesticus*). *Evolution*, **34**, 638–653.

Billing AM, Lee AM, Skjelseth S *et al.* (2012) Evidence of inbreeding depression but not inbreeding avoidance in a natural house sparrow population. *Molecular Ecology*, **21**, 1487–99.

Bjordal H, Cole SR, Parkin DT (1986) Genetic differentiation among some populations of the House Sparrow, *Passer domesticus*, from southwestern Norway. *Hereditas*, **101**, 107–114.

Blackburn TM (2008) Using aliens to explore how our planet works. *Proceedings of the National Academy of Sciences*, **105**, 9–10.

Blackburn TM, Pyšek P, Bacher S *et al.* (2011) A proposed unified framework for biological invasions. *Trends in ecology & evolution*, **26**, 333–9.

Blem C (1974) Geographic variation of thermal conductance in the house sparrow *Passer domesticus*. *Comparative Biochemistry and Physiology Part A: …*, **47**, 101–108.

Brommer JE, Hanski IK, Kekkonen J, Väisänen R a (2014) Size differentiation in Finnish house sparrows follows Bergmann's rule with evidence of local adaptation. *Journal of evolutionary biology*, **27**, 737–747.

Bumpus HC (1899) The elimination of the unfit as illustrated by the introduced sparrow, *Passer domesticus*. *Biological Lectures, Wood Hole Marine Biological …*, 3–27.

Buttemer W (1992) Differential overnight survival by Bumpus' house sparrows: an alternate interpretation. *Condor*, 944–954.

Calhoun J (1947) The role of temperature and natural selection in relation to the variations in the size of the English Sparrow in the United States. *American Naturalist*, **81**, 203–228.

Cole SR, Parkin DT (1981) Enzyme polymorphisms in the House Sparrow, *Passer domesticus*. *Biological Journal of the Linnean Society*, **15**, 13–22.

Cooke B, Fenner F (2003) Rabbit haemorrhagic disease and the biological control of wild rabbits, *Oryctolagus cuniculus*, in Australia and New Zealand. *Wildlife Research*, **29**, 689–706.

Dawson DA, Burke T, Hansson B *et al.* (2006) A predicted microsatellite map of the passerine genome based on chicken-passerine sequence similarity. *Molecular Ecology*, **15**, 1299–320.

Dawson DA, Horsburgh GJ, Krupa AP *et al.* (2012) Microsatellite resources for Passeridae species: a predicted microsatellite map of the house sparrow *Passer domesticus*. *Molecular ecology resources*, **12**, 501–23.

Dawson DA, Horsburgh GJ, Küpper C *et al.* (2010) New methods to identify conserved microsatellite loci and develop primer sets of high cross-species utility - as demonstrated for birds. *Molecular ecology resources*, **10**, 475–94.

Deutsch C, Tewksbury J (2008) Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the …*, **105**, 6668–6672.

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431–49.

Ehrenfeld JG (2010) Ecosystem consequences of biological invasions. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 59–80.

Ekblom R, Wennekes P, Horsburgh GJ, Burke T (2013) Characterization of the house sparrow (*Passer domesticus*) transcriptome: a resource for molecular ecology and immunogenetics. *Molecular ecology resources*. **14**, 636–646.

Elgvin TO, Trier CN, Tørresen OK *et al.* (2017) The genomic mosaicism of hybrid speciation. *Science Advances*, **3**, e1602996.

Engen S, Ringsby TH, Saether B-E *et al.* (2007) Effective size of fluctuating populations with two sexes and overlapping generations. *Evolution; international journal of organic evolution*, **61**, 1873–85.

Estoup A, Ravign V, Hufbauer R *et al.* (2016) Is there a genetic paradox of biological invasion? *Annual Review of Ecology Evolution and Systematics*, **47**, 51–72.

Fleischer R, Johnston R (1982) Natural selection on body size and proportions in house sparrows. *Nature*, **298**, 747–749.

Fleischer RC, Johnston RF (1984) The relationships between winter climate and selection on body size of house sparrows. *Canadian Journal of Zoology*, **62**, 405–410.

Fleischer RC, Murphy MT (1992) Relationships among allozyme heterozygosity, morphology and lipid levels in house sparrows during winter. *Journal of Zoology*, **226**, 409–419.

Fleming PJS, Allen LR, Lapidge SJ *et al.* (2006) A strategic approach to mitigating the impacts of wild canids: proposed activities of the Invasive Animals Cooperative Research Centre. *Animal Production Science*, **46**, 753–762.

Francois O, Martins H, Caye K, Schoville S (2016) Controlling False discoveries in genome scans for selection. *Molecular Ecology*, **25**, 454–469.

Frankham R (2002) Introduction to Conservation Genetics, p161-206. Cambridge University Press, Cambridge.

Frankham R (2010) Inbreeding in the wild really does matter. *Heredity*, **104**, 124.

Griffith SC, Dawson DA, Jensen H *et al.* (2007) Fourteen polymorphic microsatellite loci characterized in the house sparrow Passer domesticus (Passeridae, Aves). *Molecular Ecology Notes*, **7**, 333–336.

Griffith SC, Stewart IRK, Dawson DA, Owens IPF, Burke T (1999) Contrasting levels of extral pair paternity in mainland and island populations of the house sparrow (*Passer domesticus*): is there an "island effect"? *Biological Journal of the Linnean Society*, **68**, 303–316.

Grinnell J (1919) The english sparrow has arrived in death valley: an experiment in nature. *American Naturalist*, **53**, 468–472.

Gurevitch J, Padilla DK (2004) Are invasive species a major cause of extinctions? *Trends in Ecology & Evolution*, **19**, 470–474.

Hagen IJ, Billing AM, Rønning B *et al.* (2013) The easy road to genome-wide medium density SNP screening in a non-model species: development and application of a 10 K SNP-chip for the house sparrow (*Passer domesticus*). *Molecular ecology resources*, **13**, 429–39.

Hoban S, Kelley JL, Lotterhos KE *et al.* (2016) Finding the Genomic Basis of Local Adaptation: Pitfalls, Practical Solutions, and Future Directions. *The American Naturalist*, **188**, 379–397.

Hoffmann A a, Sgrò CM (2011) Climate change and evolutionary adaptation. *Nature*, **470**, 479–85.

Hole DG, Whittingham MJ, Bradbury RB *et al.* (2002) Widespread local house-sparrow extinctions. *Nature*, **418**, 931–2.

James FC (1970) Geographic size variation in birds and its relationship to climate. *Ecology*, **51**, 365–390.

Janzen FJ, Stern HS (1998) Logistic Regression for Empirical Studies of Multivariate Selection. *Evolution*, **52**, 1564–1571.

Jensen H, Bremset EM, Ringsby TH, Saether B-E (2007) Multilocus heterozygosity and inbreeding depression in an insular house sparrow metapopulation. *Molecular Ecology*, **16**, 4066–78.

Jensen H, Moe R, Hagen IJ *et al.* (2013) Genetic variation and structure of house sparrow populations: is there an island effect? *Molecular Ecology*, **22**, 1792–805.

Jensen H, Saether B-E, Ringsby TH *et al.* (2003) Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (*Passer domesticus*). *Journal of Evolutionary Biology*, **16**, 1296–1307.

Johnston RF (1969) Character Variation and Adaptation in European Sparrows. *Systematic Zoology*, **18**, 206.

Johnston RF (1973) Evolution in the House Sparrow. IV. Replicate studies in phenetic covariation. *Systematic Biology*, 219–226.

Johnston RF, Fleischer RC (1981) Overwinter mortality and sexual size dimorphism in the house sparrow. *The Auk*, **98**, 503–511.

Johnston R, Niles D, Rohwer S (1972) Hermon Bumpus and natural selection in the house sparrow *Passer domesticus*. *Evolution*, **26**, 20–31.

Johnston RF, Selander RK (1964) House sparrows: rapid evolution of races in North America. *Science*, **144**, 548–50.

Johnston RF, Selander RK (1971) Evolution in the house sparrow. II. Adaptive differentiation in North American populations. *Evolution*, **25**, 1–28.

Johnston RF, Selander RK (1973) Evolution in the house sparrow. III. Variation in size and sexual dimorphism in Europe and North and South America. *American Naturalist*, **107**, 373–390.

Kekkonen J, Jensen H, Brommer J (2012) Morphometric differentiation across House Sparrow *Passer domesticus* populations in Finland in comparison with the neutral expectation for divergence. *Ibis*, **154**, 846–857.

Kekkonen J, Seppä P, Hanski IK *et al.* (2011) Low genetic differentiation in a sedentary bird: house sparrow population genetics in a contiguous landscape. *Heredity*, **106**, 183–90.

Klitz W (1973) Empirical population genetics of the North American house sparrow. *Ornithological Monographs*, 34–38.

Lack D (1940) Variation in the introduced English sparrow. *The Condor*, **42**, 239–241.

Lande R, Arnold S (1983) The measurement of selection on correlated characters. *Evolution*, **37**, 1210–1226.

Lee CEE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, **17**, 386–391.

Lever C (1985) Naturalized Mammals of the World. Longman, London.

Lever C (2005) Naturalised Birds of the world. pp. 204–218. T & A D Poyser, London.

Liebl AL, Schrey AW, Andrew SC, Sheldon EL, Griffith SC (2015) Invasion genetics: Lessons from a ubiquitous bird, the house sparrow *Passer domesticus*. *Current Zoology*, **61**, 465–476.

Liker A, Bókony V, Kulcsár A *et al.* (2009) Genetic relatedness in wintering groups of house sparrows (Passer domesticus). *Molecular Ecology*, **18**, 4696–706.

Lima MR, Macedo RHF, Martins TLF *et al.* (2012) Genetic and Morphometric Divergence of an Invasive Bird: The Introduced House Sparrow (*Passer domesticus*) in Brazil. *PLoS One*, **7**. e53332.

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, **23**, 2178–2192.

Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, **24**, 1031–1046.

Lowther PE (1977) Selection intensity in North American house sparrows (*Passer domesticus*). *Evolution*, **31**, 649–656.

Lundregan, S., Jensen, H., Hagen, I., Husby, A. & Gohli, J. Genetic architecture of bill morphology in a free-living house sparrow metapopulation. **In prep.**

Manwell C, Baker CMA (1975) Molecular genetics of avian proteins XIII. Protein polymorphism in three species of Australian passerines. *Australian journal of biological sciences*, **28**, 545–557.

Mayr E (1956) Geographical character gradients and climatic adaptation. *Evolution*, **10**, 105–108.

Meiri S, Dayan T (2003) On the validity of Bergmann's rule. *Journal of Biogeography*, **30**, 331–351.

Moulton MP, Cropper WP, Avery ML, Moulton LE (2010) The earliest House Sparrow introductions to North America. *Biological Invasions*, **12**, 2955–2958.

Murphy EC (1985) Bergmann's rule, seasonality, and geographic variation in body size of house sparrows. *Evolution*, **39**, 1327–1334.

Neumann K, Wetton JH (1996) Highly polymorphic microsatellites in the house sparrow *Passer domesticus*. *Molecular Ecology*, **5**, 307–9.

Nevo E (1978) Genetic variation in natural populations: patterns and theory. *Theoretical population biology*, **13**, 121–177.

O'Grady JJ, Brook BW, Reed DH *et al.* (2006) Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation*, **133**, 42–51.

Packard GC (1967) House sparrows: evolution of populations from the great plains and colorado rockies. *Systematic Zoology*, **16**, 73–89.

Parkin D (1987) Evolutionary genetics of House Sparrows. In: *Avian genetics* (eds Cooke F, Buckley P), pp. 381–406. Academic Press, London.

Parkin DT, Cole SR (1984) Genetic variation in the house sparrow, *Passer domesticus*, in the East Midlands of England. *Biological Journal of the Linnean Society*, **23**, 287–301.

Parkin DT, Cole SR (1985) Genetic differentiation and rates of evolution in some introduced populations of the House Sparrow, *Passer domesticus* in Australia and New Zealand. *Heredity*, **54**, 15–23.

Phillips BL, Brown GP, Greenlees M, Webb JK, Shine R (2007) Rapid expansion of the cane toad (*Bufo marinus*) invasion front in tropical Australia. *Austral Ecology*, **32**, 169–176.

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, **52**, 273–288.

Pugesek BH, Tomer A (1996) The Bumpus house sparrow data: a reanalysis using structural equation models. *Evolutionary Ecology*, **10**, 387–404.

Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015) A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, **24**, 4348–4370.

Richardson DS, Dawson DA, Salgueiro P, Komdeur J, Burke T (2000) Fifty Seychelles warbler (*Acrocephalus sechellensis*) microsatellite loci polymorphic in Sylviidae species and their cross-species amplification in other passerine birds. *Molecular Ecology*, **9**, 2226–31.

Rollins LA, Moles AT, Lam S *et al.* (2013) High genetic diversity is not essential for successful introduction. *Ecology and Evolution*, **3**, 4501–4517.

Saunders G, Cooke B, McColl K, Shine R, Peacock T (2010) Modern approaches for the biological control of vertebrate pests: an Australian perspective. *Biological Control*, **52**, 288–295.

Sax DF, Stachowicz JJ, Brown JH *et al.* (2007) Ecological and evolutionary insights from species invasions. *Trends in ecology & evolution*, **22**, 465–71.

Schrey AW, Grispo M, Awad M *et al.* (2011) Broad-scale latitudinal patterns of genetic diversity among native European and introduced house sparrow (*Passer domesticus*) populations. *Molecular Ecology*, **20**, 1133–1143.

Schrey AW, Liebl AL, Richards CL, Martin LB (2014) Range expansion of house sparrows (*Passer domesticus*) in Kenya: evidence of genetic admixture and human-mediated dispersal. *The Journal of heredity*, **105**, 60–9.

Selander RK, Johnston RF (1967) Evolution in the House Sparrow. I. Intrapopulation Variation in North America. *The Condor*, **69**, 217–258.

Silva CNS, McFarlane SE, Hagen IJ *et al.* (2017) Insights into the genetic architecture of morphological traits in two passerine bird species. *Heredity*, 1–9.

Symonds MRE, Tattersall GJ (2010) Geographical variation in bill size across bird species provides evidence for Allen's rule. *The American naturalist*, **176**, 188–97.

Teplitsky C, Millien V (2014) Climate warming and Bergmann's rule through time: Is there any evidence? *Evolutionary Applications*, **7**, 156–168.

Turner P, Downey P (2010) Ensuring invasive alien plant management delivers biodiversity conservation: insights from an assessment of *Lantana camara* in Australia. *Plant Protection Quarterly*, **25**, 102–111.

Väisänen RA, Lehväslaiho H (1984) Absence of genic polymorphism in a northern population of the house sparrow, *Passer domesticus*. *Hereditas*, **100**, 161–162.

Vangestel C, Mergeay J, Dawson D a, Vandomme V, Lens L (2011) Spatial heterogeneity in genetic relatedness among house sparrows along an urban-rural gradient as revealed by individual-based analysis. *Molecular Ecology*, **20**, 4643–53.

de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.

Wing L (1943) Spread of the starling and English sparrow. *The Auk*, **60**, 74–87.

Chapter Two

Inaccuracies in the history of a well-known introduction: a case study of the Australian House Sparrow (*Passer domesticus*)

Chapter Two Vignette

Information on colonization history is important background knowledge for research on invasion biology. Details on the origin of the source population, the time of arrival and the number of individuals introduced can all be useful for framing future research. We knew at the start of this project that historical information on the introduction of the house sparrow to Australia would give the research on population and evolutionary genetics context. When we reviewed secondary sources that reported information on the house sparrow's introduction to Australia we began to find many inconsistences. As a result, we began a systematic review of the literature and primary sources such as newspaper articles and the records of Acclimatisation Societies that introduced sparrows and a number of other species to Australia. This review turned out to uncover many plausible invasion scenarios for the introduction that will be the focus of our research on genetic population structure in this thesis.

This manuscript is published in Avian Research and the final publication is presented here.

See supplementary material for Chapter 2 at:
https://figshare.com/s/dbc245f85f4cc3cdd5b1

Andrew and Griffith *Avian Res* (2016) 7:9
DOI 10.1186/s40657-016-0044-3

Avian Research

RESEARCH Open Access

Inaccuracies in the history of a well-known introduction: a case study of the Australian House Sparrow (*Passer domesticus*)

Samuel C. Andrew* and Simon C. Griffith

Abstract

Background: Modern ecosystems contain many invasive species as a result of the activity of acclimatisation societies that operated in the second half of the nineteenth century, and these species provide good opportunities for studying invasion biology. However, to gain insight into the ecological and genetic mechanisms that determine the rate of colonization and adaptation to new environments, we need a good understanding of the history of the introduced species, and a knowledge of the source population, timing, and number of individuals introduced is particularly important. However, any inaccuracies in the history of an introduction will affect subsequent assumptions and conclusions.

Methods: Focusing on a single well-known species, the House Sparrow (*Passer domesticus*), we have documented the introduction into Australia using primary sources (e.g. acclimatisation records and newspaper articles).

Results: Our revised history differs in a number of significant ways from previous accounts. Our evidence indicates that the House Sparrow was not solely introduced from source populations in England but also from Germany and most strikingly also from India—with the latter birds belonging to a different race. We also clarify the distinction between the number released and the number of founders, due to pre-release captive breeding programs, as well as identifying inaccuracies in a couple of well-cited sources with respect to the range expansion of the introduced populations.

Conclusions: Our work suggests that caution is required for those studying introductions using the key sources of historical information and ideally should review original sources of information to verify the accuracy of published accounts.

Keywords: Acclimatisation, Bottlenecks, Propagule size, Range expansion, Hybridisation

Background

Species that were deliberately introduced through the efforts of the acclimatisation societies in the late 1800s have been used as a 'natural' experiment to provide insight into a range of questions in ecology (Duncan et al. 2003; Cassey et al. 2004), evolution (Johnston and Selander 1964; Sax et al. 2007; Moran and Alexander 2014), and population and conservation genetics (Briskie and Mackintosh 2004; Congdon and Briskie 2010). One reason for the focus on these human-assisted introductions is that the temporal and spatial scale enables us to study competition and adaptation by organisms in very different environments (to the natural range) over an appropriate timescale. Another key reason is that there is a sense that there are good data available on important 'experimental' variables such as propagule size, the source of the founders, and the sites of introduction. Certainly for vertebrates, the existence of a few key texts that summarise the introductions by the acclimatisation societies such as the works by Long (1981) and Lever (1985, 1987, 2005), suggest great promise for studies of invasive biology, and they have been well used as key sources

*Correspondence: samuel.andrew@students.mq.edu.au
Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia

for example by Cassey et al. (2004), Sol et al. (2008) and Blackburn et al. (2009). However, whilst these summaries are potentially valuable sources and represent an impressive effort in summarising the introduction histories of hundreds of species across the world, there have been very few efforts to evaluate the quality of the data summarised in them, and there are some reasons to suspect that the numbers are not always entirely accurate.

Many of the data summarised in accounts of introduced species (e.g. Long 1981; Lever 1985, 1987, 1992, 2005) are derived from the records of acclimatisation societies themselves, and are potentially compromised by inadequate reporting by these societies. Certainly in Australia animals were being introduced by private individuals before the establishment of the acclimatisation societies, and indeed the main rationale given for founding the first acclimatisation society (in Victoria, in 1857) was to provide better support (politically, logistically and infrastructure) for the already ongoing attempts to introduce animals (Courcy 2003). A central role of the societies was to house arrivals for a period of acclimatisation before they were released, and for many species they were also bred in captivity before release (Jenkins 1977; Leishman 1997; Courcy 2003). Following the initial burst of enthusiasm from the late 1850s to mid-1860s the acclimatisation societies often became administratively dysfunctional and marred by in-fighting as the enthusiasm for acclimatisation quickly passed, and in Australia they morphed into the first Zoological Parks (Lever 1992; Dunlap 1997; Leishman 1997; Courcy 2003). As a result, the record keeping and the reports vary in quality across the different Australian acclimatisation societies and over the years (the history of the acclimatisation societies of South Australia, New South Wales, Tasmania, Queensland, and Victoria, is dealt with in detail elsewhere: Rolls 1969; Jenkins 1977; Lever 1992; Courcy 2003).

The principal cause for concern over the reliability of historical data is the inconsistencies in the numbers reported across different sources. This is nicely exemplified by the work that has examined the effects of propagule pressure on invasion success of passerine birds introduced to Australia, New Zealand and the United States (Blackburn et al. 2011a, 2013; Moulton et al. 2011, 2012a, b; Moulton and Cropper 2014). In their study, Moulton et al. (2012b) highlighted the discrepancies between four different key sources with respect to propagule size and used the four different estimates to model the effect of propagule pressure on the success of introductions. For example the total number of House Sparrows (*Passer domesticus*) reported to have been introduced into Australia by these four sources were 345 (Ryan 1906), 130 (Jenkins 1977), 65 (Balmford 1981) and an unknown number greater than 100 (Newsome and Noble 1986). Similar differences were found across other species, although for the question addressed successful establishment versus propagule size—the results were relatively consistent using the variety of different data available for each species (Moulton et al. 2012b). Similar problems were encountered with records from New Zealand (Moulton et al. 2011), suggesting that there may be persistent problems with this kind of data. Moulton et al.'s work (2011, 2012b) demonstrates how variable the well-cited key sources of information can be. A single species case study on the Yellowhammer (*Emberiza citrinella*) by Pipek et al. (2015) in New Zealand shows the value of using additional primary sources of information. The problems do not just relate to the number of animals released and are therefore not just restricted to questions regarding propagule size.

Introduced and invasive populations have long been the focus of molecular study and again, here an understanding of the source of introduced populations is particularly important in interpreting outcomes. For example, significant genetic differences in allele frequency and diversity of microsatellites were found between the introduced House Sparrow populations of Kenya and North America (Schrey et al. 2011). This population differentiation is thought to be due to Kenyan populations being founded by birds from the Middle East or India and not from Europe which was the source of the North American House Sparrows (Lever 1987, 2005), rather than population differentiation being a result of stochastic factors (Schrey et al. 2011, 2014).

The available histories that we have of the introduced populations around the world (e.g. Long 1981; Lever 1985, 1987, 2005) will remain important sources of information that will help us to understand a range of questions in invasive biology. However, it is important to develop an appreciation of any deficiencies in summaries such as these and the likely sources of those errors (Moulton et al. 2011, 2012b) because a more detailed investigation of introduction history can unearth valuable information (Moulton et al. 2010). To that end, we have researched the introduction of the House Sparrow to Australia. This species is a good target for such work because, as a highly commensal species, it has been reasonably well documented over the past 150 years in both the scientific literature and popular accounts and was also well covered in the news media from the time of introduction (with the recent digitisation of these archives making them readily accessible).

Methods

Introduction and establishment records

Our starting references were species accounts that described the introduction of the House Sparrow into

Australia (Summers-Smith 1963; Anderson 2006; Higgins et al. 2006), and accounts of the introduction of birds into Australia (Long 1981; Lever 2005). We also searched all the volumes of the journal *Emu* (the long-running journal of ornithology for the Australian region) from 1901 to present using the key words of "sparrow", "house sparrow" and "*Passer domesticus*". From these starting references we followed up on all cross-referenced papers that appeared to be relevant to either the introduction, or the spread of the species. We also reviewed secondary sources of information in the form of journal articles and books that described the history of acclimatisation in Australia.

The inconsistencies that rapidly came to light in these secondary sources led us to review primary sources of information such as the annual proceedings of Australian acclimatisation societies; and the electronic archive of all Australian newspapers (Trove 2015) in an effort to establish details from contemporary sources from 1860 to the first decade of the twentieth century. We read through the annual proceedings of the Acclimatisation Society of New South Wales covering the key period 1864–1867, and the annual proceedings of the Zoological and Acclimatisation Society of Victoria from 1861 to 1875.

We conducted an online search of newspaper articles archived by the National Library of Australia using the Trove search engine (Trove 2015). We used the search term "sparrow" for all searches and refined our searches by selecting newspapers and dates to search within. We searched the Melbourne newspaper *The Argus* from 1855 to 1875 because there was discussion of the introduction for several years prior to the first shipments as the editor of this newspaper (Edward Wilson), was instrumental in the establishment of the acclimatisation movement in Australia. We searched other papers covering other regions for different periods as below and in line with the likely introduction timings for those places: *Sydney Morning Herald*, Tasmanian newspapers (including: *The Cornwall Chronicle*, *The Mercury*, *Launceston Examiner*), South Australia newspapers (including: *Border Watch*, *The South Australian Advertiser*, *The South Australian Register*, *South Australian Weekly Chronicle*) from 1860 to 1870, *The Brisbane Courier*, *The Queenslander*, *Queensland Times*, *The Telegraph* from 1863 to 1896, and *The Barrier Miner* (the local newspaper of Broken Hill, NSW) from 1860 to 1900. We also searched Western Australia newspapers (including: *The Beverley Times*, *The West Australian*, *Western Mail*) from 1900 to 1980 to find more details on the unsuccessful arrivals of sparrows in Western Australia. The majority of flagged articles were linked to the surname 'Sparrow' or did not contain information relevant to the introduction and history of the House Sparrow in Australia. However, we found 67 newspaper articles (see Additional file 1: Table S1, PDF's of these articles have been uploaded to figshare at https://dx.doi.org/10.6084/m9.figshare.2061150) that contained relevant information, not including reprints of the same story in other newspapers or articles that gave no new details. We have cited these newspaper articles (NA1) in the text by using the reference number given in Additional file 1: Table S1. Newspaper articles are ordered by date of publication and are cited in tables and Additional file 1: Tables S1, S2, S3, S4, S5 and S6 if not in the main text.

We also consulted the online Historical Atlas of Living Australia (HALA 2014) to identify the earliest records of House Sparrows observed in different local government regions of Australia (source of these entries is reported in Additional file 1: Table S4). The HALA collates records from a variety of sources such as museum specimens and government reports, and in this case the most valuable information we found through that route was a New South Wales Department of Agriculture report that summarised the results of a postal survey in 1905 that had been conducted between 1890 and 1905 to investigate the spread of the sparrow across the whole of NSW (Musson 1907). Using the HALA we also identified dates from bird surveys in Rockhampton, and Roma in Queensland and Tennant Creek in the Northern Territory. In addition to searching the HALA for arrival dates we used literature searches for bird lists published for regions with unknown arrival dates for the sparrow. In these cases the key terms were 'bird list' and the location name. For a number of towns (Innisfail, Mt Isa, Townsville, Atherton and Torres Strait) we found publications that specifically summarised the arrival of the species in the region (Additional file 1: Table S4).

The collection of primary evidence described above allowed us to characterise the spread of the House Sparrow across Australia. We often found unequivocal discussion in the newspaper articles of the arrival of the species in the location. Articles either reported the arrival of shipments of sparrows and details of the numbers and site of release, or in areas that were naturally colonised, such as Broken Hill, stories documented the incidence of sparrows on a certain date and made clear reference to the absence of them at an earlier time.

The information collected is not suitable for any quantitative analyses but is appropriate for description and visualisation through a map charting the temporal and spatial spread of the species across Australia, which can be usefully compared to earlier representations of the House Sparrow invasion (Summers-Smith 1963; Jones 1986).

Results

We found that important details about the House Sparrow introduction to Australia as often reported in well recognized and widely cited secondary sources (see Table 1 for Summary), are inaccurate. There are inconsistencies between these secondary sources and also between those and the information we constructed from primary sources (Table 1). For example, the total number of House Sparrows liberated in Melbourne varies widely between all sources, from as few as 65 to over 400. Whilst is it widely accepted that the introduced birds were from England (all secondary sources in Table 1), we found strong evidence that the first House Sparrows to breed in Australia were from India, with good indications that House Sparrows also came from other sources in Europe. Below we summarise the new details and clarify contradicting reports (Table 2 provides our revised summary). We have structured our findings by considering the different states/territories involved separately as at the time of the introduction (over 40 years before federation) these entities were economically and politically very isolated. Furthermore, from the perspective of a sparrow they were biologically isolated with vast tracts of inhospitable habitat separating the major settlements of these colonies and the immediately adjacent farmlands.

The initial introduction into Victoria

Most sources report that the first successful arrival of House Sparrows to Melbourne was on the *Princess Royal* in January 1863 (19 birds), followed by another shipment in June of the same year (see Table 2 for more details on all introductions). We confirm that House Sparrows were amongst the birds landed from the *Princess Royal* on 26 January 1863, but we have found clear evidence of House Sparrows already being in Australia before that. A newspaper article from 29 November 1862 reported a private shipment by a Mr. Landells, of a number of animals including 11 'Indian Sparrows' (NA 4). These were confirmed to be *P. domesticus* in another newspaper article that documented the grateful receipt of "four Indian house sparrows, from Mr. G. J. Landells" by the Acclimatisation Society of Victoria in December 1862 (NA 5). On 23 January 1863, another newspaper article (NA 8) reported on the breeding success of the earlier shipment of sparrows (that are presumably the Indian ones) in the aviary in Melbourne, before the *Princess Royal* shipment arrived. This article (from 23 January 1863 NA 8) leaves no room for doubt that these 'Indian sparrows' are *P. domesticus*, "Perhaps the most important inhabitants of this locality are the sparrows, about which there can be no mistake. They are the cock-sparrow of English city tradition or nothing, nor have they sacrificed a single characteristic to their change of country. Their impudence in inimitable, and their strut and inquisitive cocking of the head and impertinent twist of the tail, are as familiar to them as ever…, … and few will regret their appearance on the roofs and on the streets of our city and suburbs. This may be looked for, for they are breeding fast." This article therefore confirms that the House Sparrow was already present and breeding before the arrival of the *Princess Royal*. On 31 January 1863, a report in *The Argus* (NA 10) refers to the sparrows that have recently arrived on the *Princess Royal*, but also makes it quite clear again that there are sparrows from earlier imports and that they are breeding: "The young sparrows, born of the older importations, have tumbled out of their nest, and will soon set up for themselves, while other sparrows and linnets are again breeding, probably for the last time this season." We did not find any earlier mention of House Sparrows being brought in before 1862 and therefore suggest that the species arrived into Australia in 1862 initially from India. In September 1863 a female House Sparrow arrived in Ballarat (brought over from Europe by a Mr. Weber) as the sole survivor of a flock of 160 that were originally taken from Germany (NA 22). Therefore, in the space of less than a year House Sparrows were documented as arriving in Australia from source populations in India, England and Germany.

Whilst we were unable to find confirmation of their eventual arrival there are also reports of plans to ship sparrows from Ceylon (NA 14 *Argus* 30 April 1863), and Vienna (NA 37 *The Mercury* 19 June 1866). In March 1863 the ship 'Adam Sedgewick' left Calcutta, with more sparrows on board (from India) (NA 15), and two of these were recorded as landing and were confirmed as being from India (NA 17) whilst the ship 'Relief' arrived in July 1863 with an unspecified quantity of 'London sparrows' on board (NA 16), which were taken to an aviary in Pentridge where they were reported to be pairing well (NA 19). So House Sparrows were arriving from England as well as other being sent from other source populations.

The diversity of sources from which the House Sparrow was sent reflects its wide distribution at that time and the commercial nature of the acclimatisation operation. An article posted by the Acclimatisation Society of New South Wales, in the *Sydney Morning Herald* (NA 11), suggests that if people place their orders, Mr. Landells will ship more Indian House Sparrows at ten shillings each, presumably under-cutting the cost of shipping them from England. Shipments of the species from India were also more likely to survive the shorter voyage than those from Europe (it is apparent from the primary sources that the majority of House Sparrows loaded in Europe failed to survive the long sea journey from there to Australia. See Table 2 and Additional file 1: Table S2 for details).

Table 1 Comparison between recognized summaries of the acclimatisation of bird species

References	Location							
	Victoria				Adelaide	Queensland	Sydney	Tasmania
	Source	(1) Year (no. of releases)	Number of birds	(2) 1867 mass release	(3) 1863–1865 from England	(4) "Failed" 1869 introduction	(5) Sent 1863	(6) Unknown number sent in 1860s
Ryan (1906)	England	1863, 1864, 1866, 1867 and 1872	In total 345 released and one release of unknown size	Yes	No	No	Yes	Yes
Summers-Smith (1963)	England	1863, 1864–1872 (many)	A large number	No	No	Yes	Large number	Yes
Long (1981)	England and sparrows from China and Java	1863 (2), 1864, 1866, 1867 and 1872	Same as Ryan with additional numbers for the importations in 1863	Yes	Yes	Yes	No	Yes
HANZAB (Higgins et al. 2006)	Mostly England but also China and Java	1863, 1864, 1865, 1866, 1867 and 1872	In total 473 released and one unknown release	Yes	Yes	Yes	Yes	Yes
Jenkins (1977)	England	1863 (4)	In total 130 released	Yes	Yes	Yes	No	Yes
Balmford (1981)	England	1863 (2)	In total 65 released	No	NA	NA	NA	NA
Newsome and Noble (1986)	Not clear	1860s	>100	No	No	No	No	No
Lever (2005)	Suggests England	1863 (2), 1864, 1865, 1866, 1867 and 1872	In total 351 released and one unknown release	Yes	Yes	Yes	Yes	Yes
Anderson (2006)	England and western Europe	1863 (3), 1864–1872 (many)	Not mentioned	No	Yes	Yes	Yes	Yes
Andrew and Griffith (the present paper)	England, India and Germany	1863 (2), 1864 (2), 1865 and 1867 (many)	Approx. 90–130 birds imported and at least 161 released	Yes, possibly at 17 locations	1 bird in 1863 and one-third of a shipment in 1865	Did not fail, birds liberated in 1869 and 12 birds from VIC in 1868	Yes, at least 6 birds purchased by NSW society	At least 15 birds sent to Hobart in 1867

Here we summaries the facts reported by these sources. There are six commonly reported details: (1) the number and date of importation/liberations in Victoria, (2) the large number of liberations across Victoria in 1867, (3) an unknown number of birds shipped to Adelaide, (4) the "failed" introduction to Queensland in 1869, (5) the small number of birds sent to Sydney and (6) birds being sent to Tasmania. A summary of what sources these texts used is shown in Additional file 1: Table S6

NA non-applicable

Table 2 Summary of the House Sparrow's introduction to Australia

Import/liberation	Introduced to	Time	Origin of birds	Number of sparrows	Reference
Failed import	Melbourne Botanic Gardens	1860	England, sent by Edward Wilson	Failed shipment of unknown size	Courcy (2003)
Failed import	Melbourne Botanic Gardens	February 1861	England, sent by Edward Wilson on *The Norfolk*	108 birds perished	NA 2
Failed import	Melbourne Botanic Gardens	1862	England, on *The Suffolk*	60 birds perished	Jenkins (1977)
Import	Melbourne, VIC	November 1862	India, first sparrows from India	11 Indian sparrows out of many	NA 4
Import	Melbourne Botanic Gardens	January 1863	England, on the *Princess Royal*	19 sparrows out of many	Le Souef (1958); NA 9
Failed import	Melbourne, VIC	Sent April 1863	Ceylon, Sent by Mr. Layard	Unknown size possible failed	NA 14
Import	Melbourne, VIC	June 1863	Calcutter India, Sent by Mr. Butler	2 Indian sparrows out of many	NA 17
Import	Pentridge Stockade, Victoria	June 1863	England, on *The Relief*	130–160 small birds many sparrows, 30–40 sparrows given to Colonel Champ	Hardy (1928), Le Souef (1958) and Long (1981); NA 16
Failed import	Melbourne, VIC	July 1863	Unknown starting port, on *The Ceres*	250 small birds including sparrows drowned	NA 18
Import	Adelaide, South Australia	July 1863	England, by Captain Harris	1 survivor from the shipment	NA 21
Liberation	Melbourne Botanic Gardens	September 1863	Could include birds from all shipments as well as tree sparrows and java sparrows	80 birds released	Le Souef (1958)
Import	Melbourne, VIC	September 1863	Leipsic Germany, on *The Star*	1 survivor out of 160 birds	NA 22
Liberation	Pentridge Stockade, Victoria	October 1863	Sparrows given to Colonel Champ from *The Relief*	30–40	Hardy (1928), Le Souef (1958) and Long (1981)
Liberation	Boroondara, Victoria	1864	Acclimatisation Society of Victoria	A small number	Le Souef (1958)
Liberation	Ballarat, Victoria	April 1864	Annual meeting of the Victorian Acclimatisation Society 1864	22 birds	NA 28
Import/liberation	The Sydney Botanic Gardens	1863–1864	Acclimatisation Society of Victoria	Birds purchased: 2 in 1863 and 4 in 1864	Bennett (1864) and Leishman (1997); NA 24, 27 and 30
Liberation	Ballarat, Victoria	1865	Acclimatisation Society of Victoria	Unknown	Le Souef (1958)
Liberation	Murrundi, NSW	June 1865	Unknown	Three males and three females	NA 32
Import	Adelaide, South Australia	July 1865	England, on *The Orient*	One-third of the birds survived the trip.	Sutton (1935) and Condon (1951); NA 33
Failed import	Brisbane, Queensland	April 1866	London, from the British Acclimatisation Society	Failed shipment of 22 sparrows	NA 35
Failed import	Melbourne, VIC	Sent June 1866	Vienna, Austria	Failed shipment of unknown size	NA 37
Liberation	Across the State of Victoria (see Additional file 1: Table S3 for locations)	1867	Acclimatisation Society of Victoria	Unknown number of birds released at most sites. Five birds were released at Ballarat and 14 at Ararat (Sage 1957). Probably similar numbers at other locations.	Ryan (1906), Tarr (1950), Sage (1956), Le Souef (1958), Rolls (1969) and Long (1981); NA 43, 44

Table 2 continued

Import/liberation	Introduced to	Time	Origin of birds	Number of sparrows	Reference
Liberation	Hobart, Tasmania	June 1867	Acclimatisation Society of Victoria	15	Le Souef (1958); NA 40, 41
Import	Geelong, VIC	July 1867	Unknown	Unknown	NA 42
Liberation	Mount Gambier, SA	April 1868	Probable Adelaide	Unknown	NA 45
Liberation	Bowen Park, QLD	July 1868	Acclimatisation Society of Victoria	12	NA 46
Import	Queensland	May 1869	England, on *The Flying Cloud*	Unknown number released	Chisholm (1919) and Jenkins (1977); NA 49, 51
Liberation	Launceston, Tasmania	1870s	Unknown, Littler (1901) suggests the birds were from Adelaide	Unknown	Littler (1901); NA 60, 61, 62

This table shows a chronological order to events identified by our systematic review of the House Sparrow's introduction. The table includes failed importations, importations and liberations of sparrows (see full details in Additional file 1: Tables S2 and S3)

From our reading of the primary sources, we believe that the main effort to import the House Sparrow was concentrated in the years between 1862 and 1867, and that during that period the shipments that did arrive were probably still sufficiently newsworthy to have been covered. Certainly there are multiple articles documenting many of the arrivals of birds (suggesting that multiple papers felt them important enough to cover). Given how quickly the birds appear to have increased in number in Melbourne (see below), we suspect that there were no further significant imports after about 1869 given the difficulty of shipping them and the increasing consternation about the potential for them to become a pest. All of the arrivals by ship that are documented in the newspapers are listed in Additional file 1: Table S2. It remains difficult to judge how well this represents the actual number of House Sparrows imported from overseas, and news coverage may have declined over time and some shipments may have been missed. Therefore, unfortunately, whilst the primary sources provide important insight into the multiple source populations from which the House Sparrow arrived in Australia, it is difficult to confidently chart the growth of the population of House Sparrows in the first few years, or work out what proportion of adults introduced in the these first few years came from different sources.

From our reading of the primary and secondary sources reporting on the imports and releases of this species it appears that the sources of some of the later inaccuracies are due to a confusion between imports and releases. Ryan (1906) reported that "120 were first liberated in the Botanical Gardens in 1863; in 1864, 125 more; in 1866 another lot, and in 1867, many birds about Melbourne were caught and were distributed generally over the state of Victoria, and in 1872, 100 more were imported and liberated.". In the first coherent attempt to characterise the species' introduction into Australia, Sage (1957) reported the same numbers by Ryan (but does not cite Ryan). These records (which form the basis of many of the more recent estimates of the propagule size of the House Sparrow in Australia) are problematic for a few reasons. First we are unable to identify all the birds liberated in 1864, and we only have records for a small number released in Boroondara (Le Souef 1958). Alternatively, the 1864 liberations could refer to shipments sent in 1863 but for which we cannot find any further details of arrival (Table 2). The 1866 liberation could refer to a shipment of 100 birds sent from Vienna to Melbourne (NA 37) but we have no record of these birds successfully arriving. The 1872 shipment seems altogether unlikely. It is not mentioned in any newspaper article nor recorded in the 1872 or 1873 annual proceedings of the Zoological and Acclimatisation Society of Victoria. It could have been a private shipment; however it does seem unlikely that a private individual would have gone to the trouble and expense of importing new birds at this time given that House Sparrows were already abundant around the Victorian colony by this time (Additional file 1: Table S5).

Whilst these numbers reported by Ryan (1906) and Sage (1957) are typically interpreted as being the size of the population of House Sparrows imported and released, in fact Ryan (1906) just refers to the size of the release. On arrival into Melbourne, many, if not all, of the birds were established in aviaries managed by the Acclimatisation Society of Victoria before being later liberated locally in Victoria, or shipped elsewhere in Australia. There are no quantitative accounts of the number of birds breeding or the number of offspring produced in Australia in the first few years (either in aviaries or in the wild after release). There is also an absence of any records regarding escaped birds from captivity. We know that on arrival from the ships birds were taken to the aviaries, in which the House Sparrows were reported as breeding successfully and producing several broods in a season (NA 8, 10). It is therefore likely that a good number of the birds documented as being released were hatched in Australia and may have come from a much smaller group of breeding adults. In total we have found evidence for approximately 90–130 adults arriving on ships into Victoria, and suggest that the propagule size (and genetic bottleneck) is around 110 individuals because it seems likely that not every individual would have survived and produced offspring.

The 1872 annual proceedings of the Zoological and Acclimatisation Society of Victoria purportedly have a full list of all the animals imported and liberated by the society prior to 1872 (Black 1872). The records suggest that a total of 105 English sparrows (60 in Royal Park, 40 in Pentridge, and 5 in Ballarat), were released by the society before 1872 in the Melbourne area. However, the 1873 proceedings that we accessed and were cited by Balmford (1981) only report 65 sparrows being liberated, because a liberation in Pentridge Stockade was not included. Thus, the inconsistencies in the records appear to start very early on, with the 1872 summary being inconsistent with the 1873–1875 proceedings. It is unlikely that the society released birds in the Melbourne area after 1872 because the species was already very well established in Victoria by this point, with newspapers from as early as 1868 suggesting that there were already hundreds of sparrows in the city (NA 47). By the early 1870s the policy of the Acclimatisation Society was also changing having realised they were introducing a species that was fast being recognised as a pest in Australia (Le Souef 1958; Higgins et al. 2006). For example, in 1875 they rejected a request to send House Sparrows to the

Torres Strait for liberation (Le Souef 1958). With respect to the number of individuals that were initially brought into Australia from overseas, it is very difficult to come up with an accurate figure.

Introduction into South Australia

The first shipment to Adelaide in 1863 of 100 birds had only one surviving sparrow, which was sold at auction as the "only one in the colony" of South Australia in July 1863 (NA 21). The next documented landing of sparrows into Adelaide from England on *The Orient* (in July 1865) reported that one-third of the birds survived but the starting number was not reported (NA 33). These birds were delivered into captivity and there are later reports of sparrows being liberated in both Adelaide (NA 50), and Mount Gambier (NA 45). There is relatively little in the newspapers about further shipments or releases into South Australia, although we know that they were numerous in Adelaide by 1881 (Condon 1951). There is no evidence that birds were shipped from Victoria into South Australia and it is possible that the South Australian birds represent a separate and genetically independent source population.

Introduction into Queensland

In 1866, a report on a meeting of the Queensland Acclimatisation Society, confirms the loading of 22 sparrows onto a ship as a gift from the British Acclimatisation Society (NA 35); however there is no record of the arrival of these sparrows. The failure of these birds to have arrived is suggested by the observation that in a report on a meeting of the Queensland Acclimatisation Society in July 1868, 12 sparrows were recorded as having been sent up as a donation from the Acclimatisation Society of Victoria, and released (NA 46), and were still doing well and breeding the following year in February 1869 (NA 48). In April of 1869 a successful shipment of sparrows and other 'English birds' arrived on *The Flying Cloud* (into Brisbane from London), although unfortunately there is no detail of the number of arrivals beyond the fact that the numbers had been 'greatly reduced on the voyage' (NA 49, 51). These birds were released into the Botanic Gardens after they had recovered from the voyage (NA 51). In his summaries Chisholm (1919, 1926) maintains that this early introduction failed and that Brisbane and Queensland must have been populated by the natural spread of birds from the south in the early 1900s. For example Chisholm (1926) writes "While on a visit to Toowoomba in 1903 I noticed a small flock of about a dozen sparrows in the town and was told that only a few weeks before they had been seen for the first time". The picture presented by Chisholm (1919, 1926) was, not surprisingly cited and followed by Summers-Smith (1963), and Jones (1986), which is not surprising because it seems to be so clear. However, it is clear from a number of newspaper articles dating from the 1870s and 1880s that the House Sparrow was thriving in both Brisbane and Toowoomba (NA 56, 57, 58), and there is very little doubt that Chisholm (1919, 1926) was wrong. It is possible but fairly unlikely that a species like the House Sparrow could have died out after being described as being so common in Brisbane that it had become a 'grievous evil' in 1881 (NA 57), to be absent for several decades before being replaced by later arrivals spreading from the south.

Introduction into New South Wales

In April 1864 the Acclimatisation Society of New South Wales reported that two sparrows had been brought into NSW in the previous year (NA 27), but there were no details as to where these came from, although Le Souef (1958) reports these as a gift from the Acclimatisation Society of Victoria. In November 1863, in another report covered by a newspaper, the Acclimatisation Society of New South Wales reported that the pair of sparrows (presumably in captivity in the Sydney Botanic Gardens) had laid eggs and that the female was incubating them (NA 24). A report on 3 December 1864 (NA 29) states that another five English sparrows were added to the collection in the aviaries in the Botanic Gardens in November 1864 (but again no hint of their origin). A second report (NA 30) on 6 December 1864 refers to the purchase of two pairs of English sparrows but unfortunately there are no details on where they came from or when exactly they arrived and given the date they could be included in the five new additions reported in the same week (NA 29). These are the only details that we have on sparrows being brought to Sydney, so it is very difficult to know where they came from. The only evidence of the introduction of sparrows into the wild is made in the fourth annual report of the Acclimatisation Society of New South Wales (April 1865) where it is stated that four English sparrows were liberated in Sydney (Bennett 1872) and apparently in the Botanic Gardens and Elizabeth bay a few kilometres away (Leishman 1997). The number of sparrows received and held by the Acclimatisation Society of New South Wales remained low compared to the operation in Melbourne and in the sixth report (1867) they refer to just three sparrows being kept in their collection in the aviaries. There are only two other records of sparrows being released in NSW. In 1865 in a communication to the Acclimatisation Society of New South Wales a Dr. Gordon of Murrurundi (in the Upper Hunter Valley) reported releasing three pairs of English sparrows in March of 1865. It is not clear where these birds came from but the nature of his letter makes it clear that they did not come from the Acclimatisation

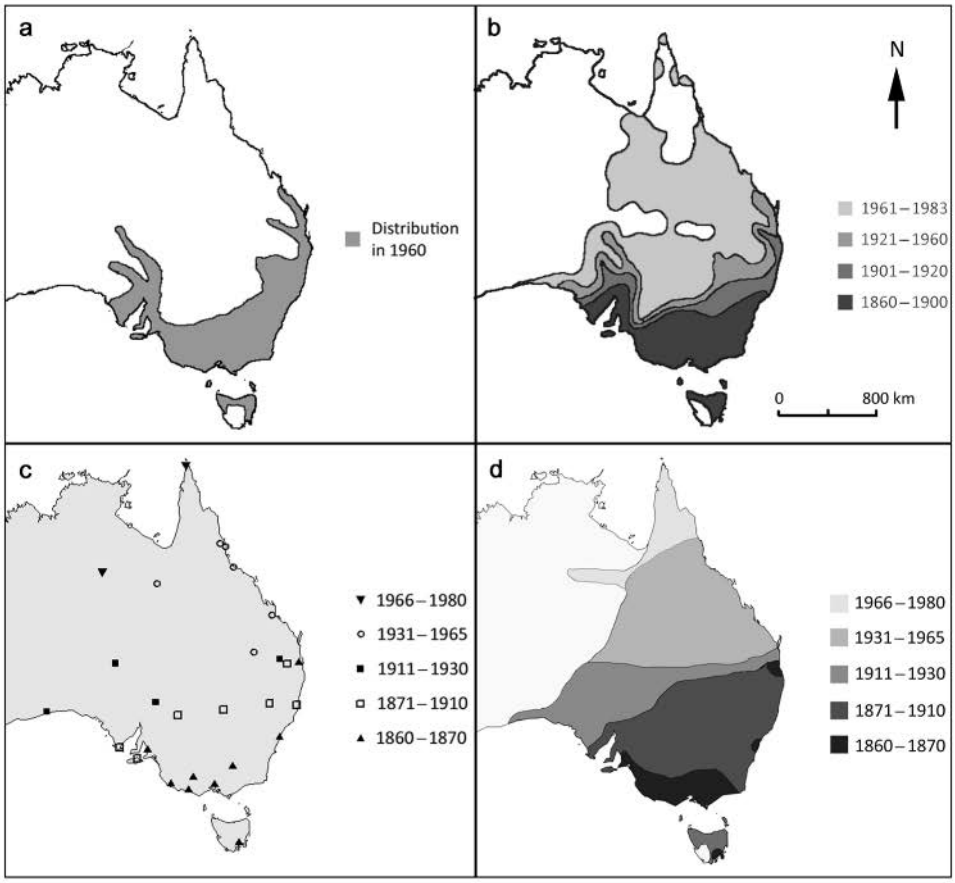


Fig. 1 The range expansion of the Australian House Sparrow. The map (**a**) is a redrawing of a distribution map published by Summers-Smith (1963) which estimated the distribution of Australian House Sparrows in 1960. The map (**b**) is a redrawing of a map by Jones (1986) which uses multiple sources of information to estimate the distribution of the House Sparrow in Australia at different time points and displays the range growth using contour lines. Jones' 1960 distribution line seems to follow Summers-Smith's distribution. The map (**c**) was generated from our research which collected the arrival time of House Sparrows in different towns and cities (see Additional file 1: Table S4 for full details). The map (**d**) uses our arrival points to draw our own contour lines which estimate the species range expansion between different time intervals in a comparable way to Jones (1986)

Society of NSW (NA 32). This letter suggests that individuals such as Dr. Gordon were acting independently to acquire and release sparrows. Dr. Gordon could have either imported them into Australia from overseas, or perhaps acquired them from Victoria where they were already quite numerous by this time. The other mention of an introduction is in a newspaper article from 1881 (NA 55) discussing the issue of the sparrow as an agricultural pest. In this there is a report that sparrows were introduced into Parramatta in 1875. As Parramatta is just 20 km from Sydney it is most likely that these birds either arrived on their own or were moved from the established population in Sydney. The sparrow was starting to become common in other parts of Sydney by around 1875 (Rolls 1969; Leishman 1997).

Introduction to Tasmania

The Victorian Acclimatization Society sent 15 birds to Hobart in 1867 and these birds were reportedly going to be housed over winter before release (reported in June NA 41), but we could find no further mention of them in the newspapers. It was later reported that House Sparrows were taken from Adelaide to Launceston in the 1860s or 1870s (Littler 1901). A newspaper article from 1891 (NA 61) also suggests that the House Sparrow was introduced into Launceston 'about 20 years ago' by a Mr.

Bennett from Launceston, with a slightly later article blaming the 'late Mr. Henry Bennell' (from Launceston, NA 62), so there is some disagreement over the spelling of the surname. Neither of these later articles gives any hint as to where they came from, although Adelaide, Melbourne and Hobart are the most likely sources and sparrows were doing well in both of these mainland colonies by the late 1860s.

Population establishment, range expansion, and current distribution

In 1906 the President's Address for the Australasian Ornithologist's Union reviewed the successful introduction of a number of introduced European bird species, identifying the House Sparrow and Starling (*Sturnus vulgaris*) as the most successful invaders (Ryan 1906). By the early part of the twentieth century the House Sparrow was known to have spread across New South Wales, the south-eastern part of South Australia, all over Victoria, all of the human occupied areas of Tasmania, and some of the southern towns in Queensland (Ryan 1906; Musson 1907). As they spread into new towns, newly established populations of House Sparrows quickly grew to be described as in plague proportions over the first 10–20 years (see examples in Additional file 1: Table S5), demonstrating their capacity for rapid population growth and establishment success.

A distribution map of the Australian House Sparrow indicating the spread of the species over time, was produced in the monograph of the species by Summers-Smith (1963) describing the distribution in 1960 (Fig. 1a). Summers-Smith's (1963) distribution was not significantly different to Ryan's (1906) description of the distribution, which is a little surprising since they were separated by over 50 years. Jones (1986) produced an updated version of the temporal spread of the species, which is largely based on Summers-Smith's (1963) figure with additional range expansion points for after 1960 (Fig. 1b). In our review of the primary and secondary material available, we identified inaccuracies in both of these earlier accounts of the range spread (Summers-Smith 1963; Jones 1986). Using arrival estimates for key towns across Australia (Fig. 1c; Additional file 1: Table S4) we produced a revised range expansion map for the House Sparrow in Australia (Fig. 1d). Our re-examination reveals a more extensive and quicker invasion across the southern states, as well as a quicker arrival in northern Queensland, with a more even rate of expansion overall (see Fig. 1 for comparison).

The House Sparrow was never intentionally introduced to either the Northern Territory or Western Australia. However the species has apparently made it by boat to Perth on many occasions before being detected and exterminated by the Department of Agriculture (Long 1988). Some of these arrivals were by a single individual and others were small groups with up to 10 individuals. The species has also expanded over land towards Western Australia across the South Australian border, and the earliest of these events was in 1914 (Long 1981) and also 1918 (NA 64). None of these range expansions have successfully breached the inhospitable Nullabour plain to reach the human settled areas of Western Australia. This failure is possible due to a consistent effort to eradicate expanding House Sparrows and other invasive species by the Western Australian Government (see examples NA 65, 66). In total Long (1988) found records for approximately 130 sparrows arriving in the docks of Perth or crossing the Western Australia border before extermination. A newspaper article from May 1969 (NA 67) linked the arrival of five sparrows in Perth to a boat that travelled directly from Lisbon, Portugal to Perth. This is an impressive distance for sparrows to hitch-hike by boat and suggests the possibility that sparrows could be continuing to enter Australia by boat from many countries and subsequently joining the already established populations in other cities (into which international freighters from around the world arrive).

The invasion of the Northern Territory by the House Sparrow has also been largely blocked by the harshness and scale of the arid habitat between the edges of the species range in South Australia to the south and Queensland to the east. The exception here is the small town of Tennant Creek, NT which is one of the last towns successfully invaded. The House Sparrow arrived in Tennant Creek in 1978 and we spoke to an eye witness who reported that they arrived from the east and had been colonising (in a stepping-stone process) the remote homesteads of cattle stations that are spread thinly from Camoweal in Queensland to Tennant Creek in the few years immediately prior to 1978. Whilst the major towns and cities of the Northern Territory (Katherine, Darwin and Alice Springs) are all likely to provide suitable habitat for the House Sparrow, dispersal to those from Tennant Creek remains blocked by vast stretches of desert with an extremely low density of human homesteads, most of which are not likely to support a viable population of House Sparrows.

Between 1977 and 1981 the distribution of the House Sparrow was recorded systematically across Australia as part of the effort to create the first *Atlas of Australian Birds* (Blakers et al. 1984). *The New Atlas of Australian Birds* was produced 20 years later and there were no significant changes in the recorded distribution of the House Sparrow (Barrett et al. 2003). This suggests again that the expansion of the House Sparrow has been stable since the late 1970s (see Fig. 1d for map).

Discussion

In conducting an intensive study of the history of the introduction of the House Sparrow into Australia, we discovered some very significant discrepancies with other well-used accounts. Both the source population and the number of individuals introduced appear to differ from the established record for this species. House Sparrows arrived in Australia from England, Germany and India, not just England. From a genetic viewpoint this is particularly important as those populations are genetically distinct and the Indian population is even a different race (*P. d. indicus*). The number of individuals introduced appears to have been much lower than previous estimates although we acknowledge that even our extensive historical research makes it difficult to ascertain an accurate number. Whilst these findings are of most interest to further work focused on this single species (e.g. Liebl et al. 2015), the discrepancies have broader implications with respect to the general accuracy of the well-cited references that are so well-used in invasive biology research.

The history of an introduced species provides an important foundation to a range of research questions relating to invasive biology (Duncan et al. 2003; Blackburn et al. 2011b) and population genetics (Dlugosch and Parker 2008). Much work in invasion biology has been based on the information provided by key references such as the books by Lever (1987, 2005) and Long (1981), that provide a wealth of data on hundreds of introductions across the world. However, the quality of analyses and conclusions based on these and similar sources may be jeopardised by any widespread inaccuracies in these sources. The introduction of the House Sparrow to Australia has provided a good opportunity to evaluate the accuracy of secondary sources on a single introduction and in many areas we find that both in the 'introduction literature' and 'species accounts' many important details that have become well-established were wrong. Our findings, certainly cast some doubt on the quality of these general sources, and suggest that caution needs to be applied to future work using data derived directly from them. The errors we have identified are very significant biologically. For example, the genetic composition of the Australian House Sparrow population today is likely to reflect the admixture of three genetically distinct populations and it is interesting to note that even an early study with allozymes (Parkin and Cole 1985) identified a higher level of gene diversity in House Sparrows in Australia versus New Zealand (where sparrows were only introduced directly from a single source—England). Furthermore, there is a key difference in a propagule size of approximately 110 against 400 (the number reported by some sources), although again the situation is undoubtedly complicated. Around 400 individuals may have been released into the wild around Melbourne so the ecologically relevant propagule may be around 400. However, if the majority of those 400 individuals were the progeny of a far smaller number of adults breeding in captivity, then the bottleneck and genetic propagule size are much smaller. This relatively subtle difference will be important depending on what particular question is being addressed.

We are reasonably confident that our revised history is accurate. It is unlikely that many song birds were successfully introduced into Australia prior to 1860 (Hardy 1928). Edward Wilson argued for the establishment of the first acclimatisation society in Australia because of the difficulties in keeping song birds alive on the long sea voyage (Courcy 2003). Perhaps because of Wilson's great interest in acclimatisation and his position as editor of one of Australia's key newspapers of the time (*The Melbourne Argus*), the local Australian newspapers in the 1860s–1870s regularly reported the activities of acclimatisation groups (see examples in Additional file 1: Table S1), and certainly seem to have covered all of the details that are covered by the proceedings of the societies themselves and the details that were picked up in other accounts of introductions. Importantly however, the newspapers go beyond these other sources and have provided significant missing details. The Australian newspapers therefore appear to be a very useful source, certainly with respect to the House Sparrow. It is possible that the coverage of other species would have been a little less thorough. The House Sparrow is a bird that people are very familiar with and it was newsworthy firstly because they were welcomed as an agent of biocontrol (to remove herbivorous insects on crops), and secondly as a familiar bird from home. Finally, within relatively few years the sparrow was being recognised as a pest and a number of newspaper articles focus on the debate about its status as a pest and efforts to eradicate it. The other key attribute of the sparrow is that they are a highly commensal species and they readily associate with people (Summers-Smith 1963; Anderson 2006), therefore they would have been very recognisable and an obvious member of the avifauna. Therefore as sparrows arrived in new locations, their presence was very quickly noticed and noted in newspapers.

It has been suggested that some bird species were arriving in undocumented private shipments (Balmford 1981), but there is no indication that sparrows were present in areas before the introductions and arrivals that we have documented here, and again they are a species that seems to attract comment when they are around. The birds that were introduced presumably found urban and rural settlements that were perfectly suited to them and multiplied fast. Within just a few years of initial introduction

(by 1870) there was an abundance of House Sparrows in the country (Additional file 1: Table S5), and it seems unlikely that people would have gone to great efforts to ship them from elsewhere in the world after this time (when they were still very much in the news). However it is certainly likely that small numbers would have arrived by hitch-hiking on ships as seems to have been the case in Perth, where they were detectable for the next century because of the ongoing effort to keep Western Australia sparrow-free (Long 1988; NA 67).

There are a number of reasons why the general sources such as the accounts by Lever (1987, 2005) and Long (1981) are erroneous on certain points. Perhaps a key one of these is that those sources are over-reliant on the records of the acclimatisation societies. Certainly in Australia the period from around 1860–1875 was the high point of the acclimatisation movement. After this period the societies started to decline, struggled to find funds and become administratively dysfunctional. The proceedings from the Acclimatisation Society of Victoria (which were the best of all of them) declined in quantity and quality over time and from 1872 to 1875 details were already starting to be lost when compared to the notes of earlier proceedings used by Le Souef (1958). Through old newspaper articles we have found many transcripts of regular meetings of acclimatisation societies that report details about correspondence with members of the society that was not being thoroughly minuted in the societies own records and provide important insight (Additional file 1: Table S1). The references given as source evidence also indicate that many of the authors of summary information for this species have failed to use all of the available sources (shown by limited references, Additional file 1: Table S6). There are also clear cases of incorrect interpretation and mis-citing of references in these texts that compound errors over time. For example Ryan (1906) reports 125 birds were liberated in 1864 (and he is referring to the total of releases across Victoria). Le Souef (1958) reports a small release of an unknown number of sparrows at Boroondara in 1864. Lever (2005) apparently puts the two together and reports 125 birds being liberated at Boroondara, Victoria in 1864, although we are unable to find any independent record of that number being released at that location. This is just one example of how different authors have published contradictory facts. The problem here is that a later author might now think that if 125 birds were released at Boroondara in 1864 then any birds released elsewhere in that year must have been additional ones which is not in fact the case, and thus the numbers get revised upwards.

Another possible source of confusion that arises from the proceedings of the Victorian Acclimatisation Society is that they also released 255 Java sparrows and 20 Chinese sparrows. These were not House Sparrows from Java and China as suggested by Hardy (1928) who is later referenced in HANZAB (Higgins et al. 2006). The former were Java Sparrows (*Lonchura oryzivora*), and the latter were Tree Sparrows (*P. montanus*, Le Souef 1958) that are the commensal sparrow in China. The Tree Sparrow is still found in Victoria and southern New South Wales to the west of the Great Dividing Range (Barrett et al. 2003), while the introduction of the Java Sparrow failed completely (Higgins et al. 2006). The adaptability and resilience of the House Sparrow is shown by its success on arrival compared to other failed introductions such as the Java Sparrow, Hedge Sparrow (*Prunella modularis*) and the Chaffinch (*Fringilla coelebs*). These differences likely relate to the ecology evolutionary potential of species to adapt to new environments (Cassey et al. 2004; Whitney and Gabler 2008). The House Sparrow did not really have to adapt to a new environment so much because they live in heavily modified urban and agrarian habitats, and had been associating closely with humans for 10,000 years (Anderson 2006). The environment available to the sparrow around Melbourne, Adelaide and Sydney in the 1860s was very similar to the one that they left in England, Germany and India. Food sources would have been primarily associated with horses, chickens and pigs and they would have nested in cavities under eaves in houses and farm buildings.

Conclusions

Through our searches we have found records for 16 attempts to ship House Sparrows to Australia, only nine of these shipments had birds that reached Australia alive. The acclimatisation societies intended to send large shipments but due to mortality on the voyage the founding populations were much smaller than intended but not impossibly small for the adaptable sparrow. An island population of House Sparrows has been successfully founded by only four individuals (Jensen et al. 2007) and the introduced populations to Australia rapidly increased from low numbers (Additional file 1: Table S5). Most of the birds shipped to Australia went to Melbourne in 1862–1863 with about 90–130 birds arriving alive over a 12 month period. The Melbourne population was founded by birds from England, Germany and also India, possibly leading to hybridisation between these two subspecies resulting in a population with a unique genetic composition. The findings that we have made from our intensive search of both primary and secondary sources, provide a picture of multiple introductions within a short space of time and a fairly rapid expansion across the whole eastern half of the Australian continent. This pattern of introduction and establishment should be detectable using molecular techniques and this work remains to be done. Nevertheless, our work

demonstrates clearly that previous summaries of the introduction of this single species are inconsistent and erroneous and this does cast some doubt on the accuracy of the well-cited key references for bird introductions around the world. This suggests that such works should be used cautiously and ideally further work will take a similar approach that evaluates the accuracy of introduction histories for their target species and locations.

Additional file

Additional file 1. Supplementary tables: **Table S1**. Newspaper quotes, **Table S2**. Importation events, **Table S3**. Liberation events, **Table S4**. Range growth events, **Table S5**. Examples of population growth and **Table S6**. Citation consistency between major summaries of introduction history.

Authors' contributions

Both SCA and SCG contributed to data collection and the writing of the manuscript. SCA prepared tables and the figure. Both authors read and approved the final manuscript.

Acknowledgements

We would like to thank the Australian Museum for providing access to their research library and their copies of the acclimatisation records of New South Wales and Victoria. We would also like to thank David Peacock for his notes on the House Sparrow's introduction to South Australia.

Availability of data and material

All new data produced by our research is provided in our Additional file file 1: Tables S1, S2, S3, S4, S5 and S6. Any PDF's of newspaper articles that were used have been uploaded to figshare (https://dx.doi.org/10.6084/m9.figshare.2061150) for accessibility.

Competing interests

The authors declare that they have no competing interests.

Funding

SCA was supported by Macquarie University Research Excellence Scholarships (No. 2013077). SCG was supported by an Australian Research Council Future Fellowship (FT130101253).

Received: 20 January 2016 Accepted: 2 May 2016
Published online: 18 May 2016

References

Anderson TR. Biology of the ubiquitous house sparrow: from genes to populations. New York: Oxford University Press; 2006.
Balmford R. Early introductions of birds to Victoria. Vic Nat. 1981;98:96–105.
Barrett G, Silcocks A, Barry S, Cunningham R, Poulter R. The new atlas of Australian Birds. Hawthorn East: Royal Australasian Ornithologist Union; 2003. p. 635–6.
Bennett G. The third annual report of the Acclimatisation Society of New South Wales, Sydney; 1864.
Bennett G. Papers on Acclimatisation by George Bennett, together with the 3rd–7th annual (1864–1868) report and proceedings of the Acclimatisation Society of New South Wales, Sydney; 1872.
Black J. Proceedings of the annual meeting of the Zoological and Acclimatisation Society of Victoria, Melbourne; 1872. p 25–28.
Blackburn TM, Lockwood JL, Cassey P. Avian invasions: the ecology and evolution of exotic birds. Oxford: Oxford University Press; 2009.
Blackburn TM, Prowse TAA, Lockwood JL, Cassey P. Passerine introductions to New Zealand support a positive effect of propagule pressure on establishment success. Biodivers Conserv. 2011a;20:2189–99.
Blackburn TM, Pyšek P, Bacher S, Carlton JT, Duncan RP, Jarošík V, Wilson JRU, Richardson DM. A proposed unified framework for biological invasions. Trends Ecol Evol. 2011b;26:333–9.
Blackburn TM, Prowse TAA, Lockwood JL, Cassey P. Propagule pressure as a driver of establishment success in deliberately introduced exotic species: fact or artefact? Biol Invasions. 2013;15:1459–69.
Blakers M, Davies SJJF, Reilly PN. The atlas of Australian Birds. Carlton: Melbourne University Press; 1984. p. 586–7.
Briskie JV, Mackintosh M. Hatching failure increases with severity of population bottlenecks in birds. PNAS. 2004;101:558–61.
Cassey P, Blackburn T, Sol D, Duncan R, Lockwood JL. Global patterns of introduction effort and establishment success in birds. Proc R Soc Lond B Biol Sci. 2004;271:S405–8.
Chisholm A. Introduced birds in Queensland. Emu. 1919;19:60–2.
Chisholm A. Birds introduced into New South Wales. Emu. 1926;25:276–9.
Condon HT. Notes on the birds of South Australia: occurrence distribution and taxonomy. S Aust Ornithol. 1951;8:26–68.
Congdon NM, Briskie JV. Effect of population bottlenecks on the egg morphology of introduced birds in New Zealand. Ibis. 2010;152:136–44.
Courcy C. Evolution of a zoo: a history of the Melbourne Zoological Gardens. Auburn: Quiddlers Press; 2003. p. 1–15.
Dlugosch KM, Parker IM. Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol Ecol. 2008;17:431–49.
Duncan R, Blackburn T, Sol D. The ecology of bird introductions. Annu Rev Ecol Evol. 2003;34:71–98.
Dunlap TR. Remaking the land: the acclimatization movement and anglo ideas of nature. J World Hist. 1997;8:303–19.
HALA. Atlas of living Australia; 2014. http://www.ala.org.au.
Hardy AD. Sky-larks and nightingales in Australia. Emu. 1928;27:300–1.
Higgins PJ, Peter JM, Cowling SJ. Handbook of Australian, New Zealand and Antarctic Birds: boatbill to starlings, vol. 7. Melbourne: Oxford University Press; 2006. p. 1070–88.
Jenkins C. The Noah's Ark Syndrome. Perth: The Zoological Gardens Board Western Australia; 1977. p. 73–95.
Jensen H, Bremset EM, Ringsby TH, Saether B-E. Multilocus heterozygosity and inbreeding depression in an insular house sparrow metapopulation. Mol Ecol. 2007;16:4066–78.
Johnston RF, Selander RK. House sparrows: rapid evolution of races in North America. Science. 1964;144:548–50.
Jones D. Exotic birds—selected examples. In: Kitching R, Jones RE, editors. The ecology of exotic animals and plants: some Australian case histories. Brisbane: Wiley; 1986. p. 93–106.
Le Souef J. The introduction of sparrows into Victoria. Emu. 1958;58:264–6.
Leishman AJ. Introduced birds in the Botanic Gardens, Sydney. Aust Birds. 1997;31:1–11.
Lever C. Naturalized mammals of the world. London: Longman; 1985.
Lever C. Naturalized birds of the world. London: Longman Scientific & Technical; 1987.
Lever C. They dined on eland. London: Quiller Press Ltd; 1992.
Lever C. Naturalised birds of the world. London: T&AD Poyser; 2005. p. 204–18.
Liebl AL, Schrey AW, Andrew SC, Sheldon EL, Griffith SC. Invasion genetics: lessons from a ubiquitous bird, the house sparrow *Passer domesticus*. Curr Zool. 2015;61:465–76.
Littler FM. European birds in Tasmania. Emu. 1901;1:121–4.
Long JL. Introduced birds of the world: the worldwide history, distribution and influence of birds introduced to new environments. Sydney: Universe Books; 1981. p. 372–81.
Long JL. Introduced birds and mammals in Western Australia. Perth: Agriculture Protection Board of Western Australia; 1988. p. 13–4.
Moran EV, Alexander JM. Evolutionary responses to global change: lessons from invasive species. Ecol Lett. 2014. doi:10.1111/ele.12262.
Moulton MP, Cropper WP. A comparison of success rates of introduced passeriform birds in New Zealand, Australia and the United States. Peer J. 2014;2:e509.
Moulton MP, Cropper WP, Avery ML, Moulton LE. The earliest House Sparrow introductions to North America. Biol Invasions. 2010;12:2955–8.

Moulton MP, Cropper WP, Avery ML. A reassessment of the role of propagule pressure in influencing fates of passerine introductions to New Zealand. Biodivers Conserv. 2011;20:607–23.
Moulton MP, Cropper WP, Avery ML. Historical records of passerine introductions to New Zealand fail to support the propagule pressure hypothesis. Biodivers Conserv. 2012a;21:297–307.
Moulton MP, Cropper WP, Moulton LE, Avery ML, Peacock D. A reassessment of historical records of avian introductions to Australia: no case for propagule pressure. Biodivers Conserv. 2012b;21:155–74.
Musson CT. The House Sparrow in New South Wales. Sydney: Department of Agriculture; 1907. p. 1–17.
Newsome A, Noble I. Ecological and physiological characters of invading species. In: Groves R, Burdon J, editors. Ecology of biological invasions: an Australian perspective. Cambridge: Cambridge University Press; 1986. p. 1–20.
Parkin DT, Cole SR. Genetic differentiation and rates of evolution in some introduced populations of the House Sparrow, *Passer domesticus* in Australia and New Zealand. Heredity. 1985;54:15–23.
Pipek P, Pysek P, Blackburn TM. How the Yellowhammer became a Kiwi: the history of an alien bird invasion revealed. Neobiota. 2015;24:1–31.
Rolls EC. They all ran wild. Sydney: Angus & Robertson LTD; 1969. p. 225–50.
Ryan CS. President's address: on European and other birds liberated in Victoria. Emu. 1906;5:110–9.
Sage B. Remarks on the racial status, history and distribution of the Tree Sparrow introduced into Australia. Emu. 1956;56:137–40.
Sage B. Remarks on the taxonomy, history and distribution of the House Sparrow introduced into Australia. Emu. 1957;67:349–52.
Sax DF, Stachowicz JJ, Brown JH, Bruno JF, Dawson MN, Gaines SD, Grosberg RK, Hastings A, Holt RD, Mayfield MM, O'Connor MI, Rice WR. Ecological and evolutionary insights from species invasions. Trends Ecol Evol. 2007;22:465–71.
Schrey AW, Grispo M, Awad M, Cook MB, McCoy ED, Mushinsky HR, Albayrak T, Bensch S, Burke T, Butler LK, Dor R, Fokidis HB, Jensen H, Imboma T, Kessler-Rios MM, Marzal A, Stewart IRK, Westerdahl H, Westneat DF, Zehtindjiev P, Martin LB. Broad-scale latitudinal patterns of genetic diversity among native European and introduced house sparrow (*Passer domesticus*) populations. Mol Ecol. 2011;20:1133–43.
Schrey AW, Liebl AL, Richards CL, Martin LB. Range expansion of house sparrows (*Passer domesticus*) in Kenya: evidence of genetic admixture and human-mediated dispersal. J Hered. 2014;105:60–9.
Sol D, Vilà M, Kühn I. The comparative analysis of historical alien introductions. Biol Invasions. 2008;10:1119–29.
Summers-Smith JD. The House Sparrow. London: Collins; 1963. p. 175–86.
Sutton J. Acclimatization in South Australia. S Aust Ornithol. 1935;8:92–103.
Tarr HE. The distribution of foreign birds in Australia. Emu. 1950;49:189–99.
Trove. National Library of Australia; 2015. http://trove.nla.gov.au/newspaper/search?adv=y.
Whitney KD, Gabler CA. Rapid evolution in introduced species, "invasive traits" and recipient communities: challenges for predicting invasive potential. Divers Distrib. 2008;14:569–80.

Chapter Three

Clinal variation in house sparrow body size is better explained by summer maximum temperatures during development, than cold winter temperatures

Chapter Three Vignette

The house sparrow is a classic model system for studying clinal variation in body size that is potentially a result of adaptation for thermal tolerance. Studies on introduced house sparrow populations in North America found that these clines were present in introduced populations that had been established for less than 100 years (Calhoun 1947; Johnston & Selander 1964). The clinal pattern of larger body size in colder climates, which is attributed to Bergmann's rule, was initially expected to be the result of selection to adapt to extreme winter temperatures. In Australia, we have described the variation in population age in Chapter Two, ranging from ca. 150 to 50 years. In Chapter Three we hope to replicate the studies from North America in Australian introduced populations to see if there is similar morphological differentiation across climate gradients for winter or summer temperatures. Some recent studies have highlighted the potential for clinal variation in body size to be explained by developmental plasticity that is potentially resulting from temperature stress during development. We explore if summer or winter temperatures are the best predictor of clinal variation in body size.

This manuscript is under review at The Auk, and is formatted as such.

References:

Calhoun J (1947) The role of temperature and natural selection in relation to the variations in the size of the English Sparrow in the United States. *American Naturalist*, **81,** 203–228.

Johnston RF, Selander RK (1964) House sparrows: rapid evolution of races in North America. *Science*, **144**, 548–50.

Authors: Samuel C. Andrew[1]*, Monica Awasthy[1], Amanda D. Griffith[1], Shinichi Nakagawa[2,3] and Simon C. Griffith[1]

Affiliations:

[1]Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia

[2]Department of Zoology, University of Otago, Dunedin 9054, New Zealand

[3]Evolution and Ecology Research Group and School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, NSW 2052, Australia

Short title: Summer temperatures effect house sparrow size

Abstract

Across many taxa clinal variation in body size has been observed to follow Bergmann's rule, which predicts larger body size in colder climates. For over a century, this pattern has typically been ascribed to selection for large body size in cold winter climates. Here, in spatially distributed observational data from 30 populations of house sparrow introduced into Australia and New Zealand, we show that this relationship appears to be explained by a negative relationship with high temperatures during the breeding season. Our results suggest temperature affects plasticity in growth, and suggest that a hotter climate during breeding periods could drive significant changes in morphology between populations (and potentially within populations as well if climate varies temporally across a breeding season). This idea, and our support for it, could account for much of the variation in body size that drives the well observed patterns first described by Bergmann, and that are still largely attributed to selection on adult body size during cold winters. Understanding the mechanisms behind any climate-dependent developmental plasticity could prove useful for understanding how endotherms may be affected by climate change in the future.

Keywords: Bergmann's rule, phenotypic plasticity, morphology, *Passer domesticus*.

Introduction

For over 100 years since the publication of Bergmann's Rule in 1847 (Bergmann 1847) a clinal pattern of animals having larger body size in colder climates has been observed in a majority of the hundreds of species that have been examined (Mayr 1956, James 1970, Ashton et al. 2000, Ashton 2002, Meiri and Dayan 2003, Millien et al. 2006, Clauss et al. 2013, Teplitsky and Millien 2014). To date, most studies still cite Bergmann's original explanation that larger body size is favored by natural selection in colder climates because of the thermo-regulatory benefits of a smaller surface area-to-volume ratio (Briscoe et al. 2015, Cardilini et al. 2016, Salewski and Watt 2017). A classic example of clinal variation in avian body size was previously demonstrated in North American populations of the introduced house sparrow, *Passer domesticus* (Johnston and Selander 1964, 1973; Murphy 1985). If winter temperatures are the selective force responsible for this clinal variation, as predicted by Bergmann's rule, then variation in body size between populations should be best explained by winter minimum temperatures. However, in hotter climates smaller body size can also be advantageous to an individuals' ability to thermo-regulate by dissipating heat (Partridge and Coyne 1997), even though the benefits of minor changes in body size within species have been questioned for over 40 years (Scholander 1955, McNab 1971).

Understanding the mechanisms creating the morphological differentiation described by Bergmann's rule has gained fresh impetus as part of the study of the effects of a changing climate on animal populations (Gardner et al. 2011). Indeed, declining body size in a number of avian species has been linked to increasing temperatures consistent with climate change (Gardner et al. 2009, Van Buskirk et al. 2010), and it has been suggested that higher temperatures during development may act as an influence on plasticity in growth (Merilä and Hendry 2014). The idea that clines in body size are a result of phenotypic plasticity in morphology that is mediated by the affects of high temperatures on growth, is now gaining traction (Teplitsky et al. 2008, Van Buskirk et al. 2010, Yom-Tov and Geffen 2011). Whilst

in hot climates nests can potentially act to buffer ambient conditions they can still get very hot. Recent work found that zebra finch (*Taeniopygia guttata*) nests in the Australian desert were typically several degrees warmer than ambient conditions and internal nest temperatures occasionally exceeded 50 ºC (Griffith et al. 2016). Nest microclimates may therefore be a significant determinant of variation in developmental plasticity and have the capacity to affect development and growth. Even short periods of high ambient conditions can be stressful, and promote the excessive production of reactive oxygen species (Lin et al. 2008) and changes in the oxidative status of an individual.

If the temperature during development is indeed important, then at the population level, summer maximum temperatures will be a better predictor of mean body size across locations than winter minimum temperatures. As with the house sparrows studied in North America (Johnston and Selander 1964, 1971; Murphy 1985), the species was deliberately introduced into Australia and New Zealand in the mid 19th Century from founders taken from north-western Europe (Andrew & Griffith 2016). Over the next century house sparrows expanded their range to occupy most of the urban areas across both North and South Islands of New Zealand and the eastern half of Australia (Andrew & Griffith 2016), and today are found in a range of climates that are far more variable and extreme than those in the area from which they were brought. The house sparrow populations in Australia and New Zealand therefore provide an opportunity to assess the extent to which a species responds to a changing climate in a relatively short period of time (around 160 years at most, and less than 50 years for those populations at the extreme end of the range in northern and central Australia; Andrew & Griffith 2016). Here we use these populations to test the extent to which any clinal variation in body size relates to both winter minimum and summer maximum temperatures. This will provide new insight into the extent to which body size is a response to the climate during development (due to plasticity or selection) rather than a response to selection over the winter.

Methods

Sampling

Adult house sparrows were sampled from 26 locations (see Figure. 1 and Table S1, total number of males = 636, females = 512) across Australia with approximately 40 birds measured at each location. The Australian samples were collected from April to September 2014 and in March 2015, under an Animal Research Authority of the Animal Ethics Committee at Macquarie University (ARA 2014/248). New Zealand samples were collected from four locations as part of earlier work (Table S1, total number males = 511, females = 242) between June and August 2005. New Zealand house sparrow measurements were collected by a separate team from Otago University (Animal Ethics reference number 87/08). Birds were captured using mist nets and placed in bird bags until they were measured. Birds were not held for more than 30 minutes and were released as soon as possible.

Measurements

The age and sex of birds was determined using plumage and bill color. All juvenile birds that could be still growing were removed from the morphological analyses. All individuals sampled in Australia and New Zealand had tarsus and body mass recorded. Tarsus length was measured on the right leg, from the bottom of the tarsus with the toes bent forward to the ankle joint. Body mass was taken to the nearest 0.1g using a Pesola spring scale.

All house sparrows sampled in Australia were measured by one of two measurers (S.C.G. and M.A.) and some birds were measured by both investigators to test the consistency of tarsus measurements (regression between tarsus length measurements was significant (R^2 = 0.89, t = 20.20, P < 0.001). In New Zealand, all measurements were taken by a single investigator (Karin Ludwig, employed as a research associate by S.N.).

Geographic and Climatic data.

Latitude and longitude and date of collection was recorded for each sampling site. We used the coordinates to extract Maximum Temperature of Warmest Month (BIO5), Minimum Temperature of Coldest Month (BIO6) and Temperature Seasonality (BIO4 = standard deviation *100) from Worldclim climate data (Hijmans et al. 2005). Maximum Temperature of Warmest Month (from now referred to as summer maximum) was our measure of high temperatures during the breeding season to test the developmental plasticity hypothesis. Minimum Temperature of Coldest Month (from now referred to as winter minimum) was used as a measure of winter extremes to test the cold adapted hypothesis.

To test for a relationship between variation in temperature during the breeding season and variation in body size within populations we looked at the change in mean maximum temperature across the breeding season. The house sparrow breeding season typically occurs across September to December (Duursma et al. 2017). Using weather stations closest to our 26 Australian sample sites we took the difference in mean maximum temperatures between December and September (BOM website: http://www.bom.gov.au, accessed 2015). Data for New Zealand was sourced from the National Institute of Water and Atmospheric Research (NIWA website: https://www.niwa.co.nz/education-and-training/schools/resources/climate). A second measure of climate variability in Australia used daily maximum temperatures from the three breeding seasons (September to December) prior to sampling (2011-2013). Daily maximum temperature for these 360 days were downloaded from Australian Water Availability Project (Jones et al. 2009) via http://www.bom.gov.au/jsp/awap/. Temperature variability for this period was calculated by adding the log of sample standard deviation to sample variance (Nakagawa et al. 2015). Temperature variability was found to be highly correlated with breeding season range in Australia (estimate = 0.11, $t_{24} = 6.17$, $P < 0.001$, $R^2 = 0.61$).

Data analysis

All statistical analyses were done using the R statistical program, version 3.3.1 (R Core Team 2017). All R code and data to replicate our main results is provided as supplementary material (see Figshare, link: https://figshare.com/s/dbc245f85f4cc3cdd5b1). The individual body size measurements for mass and tarsus showed a normal distribution. We calculated the mean body size and variability within the sample populations. Variability was calculated by adding the log of sample standard deviation to sample variance, this method was chosen because it will be linearly related to the mean (Nakagawa et al. 2015). If temperature during development is affecting body size then we would expect higher variability in body size in locations where the climate shows a higher degree of variation in temperature during the breeding season, due to different groups of nestlings being exposed to different temperatures throughout the breeding season. For the mean and variability data linear models were fitted using the standard "lm" function in R. For mean and variability data males and females were analyzed separately as well as all birds combined. Summer maximum and winter minimum were used as fixed effects in models for mean size. Summer maximum and temperature range across the breeding season were used as fixed effects in size variability models. To account for differences in sample size all linear models included weights (sample weight = 2(n–1), Nakagawa et al. 2015). All linear models using variability included the log of the mean size as a predictor to account for any relationship between mean body size and variability.

Linear Mixed Models (LMM) were used for fitting individual measurements of all 1901 birds. For these models sample population was used as a random factor. Linear Mixed Models were fitted using the R package *lme4* (Bates et al. 2015). For the LMM's P-values and degrees of freedom were calculated with the R package *lmerTest* (Kuznetsova et al. 2016). Summer maximum, winter minimum and sex were included as a fixed effect in the LMM's. Inter-class correlation coefficient (ICC) were also calculated for the random effect of sample

population to describe how much variation is partitioned between populations (Nakagawa and Schielzeth 2010). Inter-Class Correlation Coefficients (ICC), R^2 and narrow sense heritability (h^2) all estimate the proportion of the variance in the response variable that is explained by a fixed or random effect (Nakagawa and Schielzeth 2013). Because ICC is a proportion it can be compared between similar models such as our LMM that share the same fixed and random effects. To calculate the proportion of variance explained by random factors the residual variance and variance explained by fixed effects (known as Marginal R^2) need to be included. As a result the total variance explained by the model can also be calculated which is called the Conditional R^2 (Nakagawa and Schielzeth 2013). We report the Marginal (fixed effects) and Conditional (total model) R^2 for both mass and tarsus models.

To compare the predictive power of individual fixed effects we use semi-partial correlations for all our main models (Schielzeth 2010). Semi-partial correlations scale the response and predictor variables so the mean is 0 and the standard deviation is 1. This scaling results in estimates that are known as semi-partial correlations (from here on abbreviated to semi-partial r) and are relatable to estimates of other response variables within and between models. However, all *P* values and *t* values remain unchanged due to scaling. Scaling also allows for binary variables, such as sex here, to be coded as -1 and 1 which allows these binary factors to be directly compared to other continuous variables (Schielzeth 2010). We choose summer maximum and winter minimum as our two bioclimatic variables because they are the most relevant to our hypotheses and summer maximum is highly correlated with other bioclimatic variables such as Latitude and seasonality but not winter minimum (Table S2).

Results

Using the mean mass and tarsus measurements of house sparrow sample populations, we found that summer maximum was a better predictor of body size than winter minimum (for both tarsus length and body mass; Figure 2, Tables 1 and 2). Likewise, at the individual level

where both summer maximum and winter minimum were used as fixed effects in Linear Mixed Models (LMMs), summer maximum was a substantially better predictor for body mass (semi-partial r = -0.34 *vs.* semi-partial r = -0.01, respectively, see full LMM in Table 3). For tarsus length the fixed effect of summer maximum temperature was a stronger predictor than the winter minimum but they were similar to each other (semi-partial r = -0.13 *vs*. semi-partial r = -0.10, respectively, see Table 3). The random factor of location explained a similar amount of variance for both body mass and tarsus length, 10.7% and 9.3% respectively (Table 3).

We would expect that populations breeding across the relatively long breeding period (September – December in this region), to encounter a wider range of ambient temperatures. In locations with higher temperature ranges we would expect that individuals are developing in a wider range of temperatures resulting in more variation in body size in those populations with the highest temperature range across the breeding season. We did not find any significant relationships between temperature range and variability in mass or tarsus (Table S3 and S4). However, the variability of body mass in our sample populations was positively related to summer maximum temperature, this positive relationship was also significant for females but not for males when they were analyzed separately (Figure 3, Table S3). For tarsus length, there was a nonsignificant positive relationship between summer maximum and tarsus variability in males, females and all birds combined (Figure 3, Table S4). There was not a strong correlation between the temperature range across the breeding season and summer maximum (estimate = 0.18, t_{28} = 2.63, P = 0.01, R^2 = 0.20).

Discussion

Our observational work on the populations of house sparrow introduced into Australia and New Zealand essentially replicates earlier work in North America (Johnston and Selander

1964, 1973) and Europe (Murphy 1985), that has revealed latitudinal clines in body size in this species. As with most other similar studies across animal taxa, in these earlier papers, the clinal variation in body size was attributed to the selective effects of cold weather during the winter (Johnston and Fleischer 1981, Fleischer and Johnston 1984). However, there have been suggestions that a similar pattern may also be driven by developmental constraints affected by the climate experienced during development (Van Buskirk et al. 2010, Gardner et al. 2011, Cunningham et al. 2013). We found support for this idea through our observation that summer maximum temperatures better predicted body size variation than winter minimums. This observational finding, from house sparrow populations introduced into the range of climates found in Australia and New Zealand around 150 years ago (Andrew and Griffith 2016), supports the hypothesis that excessive environmental heat during development may be responsible for affecting growth (Van Buskirk et al. 2010, Gardner et al. 2011, Burness et al. 2013), or survival. We also explored this hypothesis by looking at the relationship between variability in body size within populations and climate variability. We did not find any significant relationship between the temperature range across the breeding season and the variation in the traits of mass or tarsus length (Table S3 and S4). For the summer maximum predictor we found the expected positive relationship however it was not always significant (Tables S3 and S4). The nonsignificant results could be due to low statistical power (only 30 populations) or a weaker effect on skeletal measurements (tarsus) than mass. The relationship between summer maximum and variability could be due to warmer climates being more likely to exceed possible "threshold temperatures" that significantly affect development. This new approach promises to be useful to future studies especially those with large numbers of sample populations, to test if this result can be replicated. Only a portion of the variation between sites was explained by temperature, ~10% of the variation in tarsus and mass was partitioned between locations by the random factor of location. This differentiation between populations could be linked to genetic differentiation or other environmental factors not included in the model such as the time of year that samples were measured.

Summer temperatures explained variation in both tarsus length and mass, with mass having a larger effect. A comparison of the three locations with the hottest, and coolest summer maximums revealed that sparrows in the hottest locations had tarsi that were 1.8% smaller and were 5.6% lighter. The greater magnitude of effect on mass than the skeletal measure of tarsus length is consistent with earlier studies showing greater plasticity of body mass and lower heritability (Alatalo et al. 1990, Jensen et al. 2003). Similarly, a study on North American migratory birds (249 species migrating during all four seasons) found increases in summer temperature caused a larger percentage decline in mass (0.34% per degree Celsius) than wing cord length (0.09% per degree Celsius; Van Buskirk et al. 2010). By comparison, in the house sparrow populations studied here mass declined by 0.33%, and tarsus declined by 0.11% per degree Celsius.

The results presented here show a proportion of the variation in body size is explained by the average climatic temperatures during the breeding season. Our findings are consistent with Bergmann's rule, but not the widely cited mechanism that cold temperatures select for large adults. Future studies should explore the mechanistic link between the climate experienced during development and the body size attained (or the size of extremities as predicted by Allen's rule; Allen 1877, Symonds and Tattersall 2010). Possible mechanisms include Temperature constraining growth (Gardner et al. 2009), constraints on parental provisioning (Cunningham et al. 2013) and parental effects (Mariette and Buchanan 2016). Our study leads to the prediction that increasing summer temperatures at a given site would drive down mean body size in the population, but it remains to be tested whether this will increase average survivability in a warming climate or is just the result of constraints during development (van Gils et al. 2016). Our conclusion that high ambient conditions affects growth and development may also lead to the prediction that breeding phenology in birds, and perhaps animals more generally, should not be timed just in relation to the abundance of

resources (Both 2010), but also in respect to the avoidance of periods of climatic extremes (Duursma et al. 2017).

Acknowledgments

We would like to thank the large number of people who helped us with access to urban catch sites for sampling the human commensal house sparrow. We would also like to thank Elizabeth L. Sheldon, Anna Feit and Peter Bird for their participation with field work in Australia and Karin Ludwig in New Zealand. We thank Daisy E. Duursma for assistance with climatic data. We would like to thank Laura L. Hurley, Mylene M. Mariette, Kate Buchanan, Sasha Dall, Peter Dunn, Janet Gardner, Tim Parker, and Andy Russell for comments on earlier drafts of the manuscript. Funding sources: S.C.A was supported by Macquarie University Research Excellence Scholarship (no. 2013077). S.C.G was supported by an Australian Research Council Future Fellowship (FT130101253). S.N is also supported by a Rutherford Discovery Fellowship (New Zealand) and an Australian Research Council Future Fellowship (FT130100268). Australian birds were sampled under an Animal Research Authority of the Animal Ethics Committee at Macquarie University (ARA 2014/248). New Zealand samples were collected through Otago University (Animal Ethics reference number 87/08). **Authors' contributions:** S.C.A. and S.C.G conceived the idea for the paper. All Authors contributed to data collection. S.C.A and S.N analyzed the data. All authors wrote and/or substantially edited the paper. S.C.G and S.N secured funding for this work.

References

Alatalo, R. V., L. Gustafsson, and A. Lundberg (1990). Phenotypic selection on heritable size traits: Environmental variance and genetic response. The American Naturalist 135:464–471.

Allen, J. A. (1877). The influence of physical conditions in the genesis of species. Radical Review 1:108–140.

Andrew, S. C., and S. C. Griffith (2016). Inaccuracies in the history of a well-known introduction: a case study of the Australian House Sparrow (*Passer domesticus*). Avian Research 7:9. doi: 10.1186/s40657-016-0044-3

Ashton, K. G. (2002). Patterns of within-species body size variation of birds: Strong evidence for Bergmann's rule. Global Ecology and Biogeography 11:505–523. doi: 10.1046/j.1466-822X.2002.00313.x

Ashton, K., M. Tracy, and A. De Queiroz (2000). Is Bergmann's rule valid for mammals? The American Naturalist 156:390–415.

Bates, D., M. Maechler, B. Bolker, and S. Walker (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8, http://CRAN.R-project.org/package=lme4.

Bergmann, C. (1847). Ueber die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. Gottinge studien 3:595–708.

Both, C. (2010). Food availability, mistiming, and climate change. In Effects of Climate Change on Birds (A. P. Møller, W. Fiedler and P. Berthold, Editors). OUP Oxford, Oxford, UK, pp. 129–149.

Briscoe, N. J., A. Krockenberger, K. a. Handasyde, and M. R. Kearney (2015). Bergmann meets Scholander: geographical variation in body size and insulation in the koala is related to climate. Journal of Biogeography 42:791–802. doi: 10.1111/jbi.12445

Burness, G., J. R. Huard, E. Malcolm, and G. J. Tattersall (2013). Post-hatch heat warms adult beaks: irreversible physiological plasticity in Japanese quail. Proceedings. Biological sciences / The Royal Society 280:20131436. doi: 10.1098/rspb.2013.1436

Van Buskirk, J., R. S. Mulvihill, and R. C. Leberman (2010). Declining body sizes in North American birds associated with climate change. Oikos 119:1047–1055. doi: 10.1111/j.1600-0706.2009.18349.x

Cardilini, A. P. A., K. L. Buchanan, C. D. H. Sherman, P. Cassey, and M. R. E. Symonds (2016). Tests of ecogeographical relationships in a non-native species: what rules avian morphology? Oecologia 181:783–93. doi: 10.1007/s00442-016-3590-9

Clauss, M., M. T. Dittmann, D. W. H. Müller, C. Meloro, and D. Codron (2013). Bergmann's rule in mammals: A cross-species interspecific pattern. Oikos 122:1465–1472. doi: 10.1111/j.1600-0706.2013.00463.x

Cunningham, S. J., R. O. Martin, C. L. Hojem, and P. a R. Hockey (2013). Temperatures in Excess of Critical Thresholds Threaten Nestling Growth and Survival in A Rapidly-Warming Arid Savanna: A Study of Common Fiscals. PLoS ONE 8. doi: 10.1371/journal.pone.0074613

Duursma, D. E., R. V. Gallagher, and S. C. Griffith (2017). Characterizing opportunistic breeding at a continental scale using all available sources of phenological data; an assessment of 337 species across the Australian continent. The Auk 134:509–519.

Fleischer, R. C., and R. F. Johnston (1984). The relationships between winter climate and selection on body size of house sparrows. Canadian Journal of Zoology 62:405–410. doi: 10.1139/z84-063

Gardner, J. L., R. Heinsohn, and L. Joseph (2009). Shifting latitudinal clines in avian body size correlate with global warming in Australian passerines. Proceedings. Biological sciences / The Royal Society 276:3845–52. doi: 10.1098/rspb.2009.1011

Gardner, J. L., A. Peters, M. R. Kearney, L. Joseph, and R. Heinsohn (2011). Declining body size: a third universal response to warming? Trends in ecology & evolution 26:285–91. doi: 10.1016/j.tree.2011.03.005

van Gils, J. A., S. Lisovski, T. Lok, W. Meissner, A. Ożarowska, J. de Fouw, E. Rakhimberdiev, M. Y. Soloviev, T. Piersma, and M. Klaassen (2016). Response to warming Arctic warming reduces red knot fitness in tropical wintering range. Science 352:819–821. doi: 10.1126/science.aad6351

Griffith, S. C., M. C. Mainwaring, E. Sorato, and C. Beckmann (2016). High atmospheric temperatures and “ambient incubation” drive embryonic development and lead to earlier hatching in a passerine bird. Royal Society open science 3:150371. doi: 10.1098/rsos.150371

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis (2005). Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965–1978. doi: 10.1002/joc.1276

James, F. C. (1970). Geographic size variation in birds and its relationship to climate. Ecology 51:365–390.

Jensen, H., B.-E. Saether, T. H. Ringsby, J. Tufto, S. C. Griffith, and H. Ellegren (2003). Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (*Passer domesticus*). Journal of Evolutionary Biology 16:1296–1307. doi: 10.1046/j.1420-9101.2003.00614.x

Johnston, R. F., and R. C. Fleischer (1981). Overwinter mortality and sexual size dimorphism in the house sparrow. The Auk 98:503–511.

Johnston, R. F., and R. K. Selander (1964). House sparrows: rapid evolution of races in North America. Science 144:548–50. doi: 10.1126/science.144.3618.548

Johnston, R. F., and R. K. Selander (1971). Evolution in the house sparrow. II. Adaptive differentiation in North American populations. Evolution 25:1–28.

Johnston, R. F., and R. K. Selander (1973). Evolution in the house sparrow. III. Variation in size and sexual dimorphism in Europe and North and South America. American Naturalist 107:373–390.

Jones, D. A., W. Wang, and R. Fawcett (2009). High-quality spatial climate data-sets for Australia. Australian Meteorological and Oceanographic Journal 58:233–248.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen (2016). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-32. https://CRAN.R-project.org/package=lmerTest.

Lin, H., D. De Vos, E. Decuypere, and J. Buyse (2008). Dynamic changes in parameters of redox balance after mild heat stress in aged laying hens (*Gallus gallus domesticus*). Comparative Biochemistry and Physiology 147:30–35. doi: 10.1016/j.cbpc.2007.07.005

Mariette, M. M., and K. L. Buchanan (2016). Prenatal acoustic communication programs offspring for high posthatching temperature in a songbird. Science 353:812–814.

Mayr, E. (1956). Geographical character gradients and climatic adaptation. Evolution 10:105–108.

McNab, B. (1971). On the ecological significance of Bergmann's rule. Ecology 52:845–854.

Meiri, S., and T. Dayan (2003). On the validity of Bergmann's rule. Journal of Biogeography 30:331–351. doi: 10.1046/j.1365-2699.2003.00837.x

Merilä, J., and A. P. Hendry (2014). Climate change, adaptation, and phenotypic plasticity: The problem and the evidence. Evolutionary Applications 7:1–14. doi: 10.1111/eva.12137

Millien, V., S. Kathleen Lyons, L. Olson, F. a. Smith, A. B. Wilson, and Y. Yom-Tov (2006). Ecotypic variation in the context of global climate change: Revisiting the rules. Ecology Letters 9:853–869. doi: 10.1111/j.1461-0248.2006.00928.x

Murphy, E. C. (1985). Bergmann's rule, seasonality, and geographic variation in body size of house sparrows. Evolution 39:1327–1334.

Nakagawa, S., R. Poulin, K. Mengersen, K. Reinhold, L. Engqvist, M. Lagisz, and A. M. Senior (2015). Meta-analysis of variation: ecological and evolutionary applications and beyond. Methods in Ecology and Evolution 6:143–152. doi: 10.1111/2041-210X.12309

Nakagawa, S., and H. Schielzeth (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. Biological Reviews 85:935–956. doi: 10.1111/j.1469-185X.2010.00141.x

Nakagawa, S., and H. Schielzeth (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods in Ecology and Evolution 4:133–142. doi: 10.1111/j.2041-210x.2012.00261.x

Partridge, L., and J. A. Coyne (1997). Bergmann's Rule in Ectotherms: Is It Adaptative. Evolution 51:632–635.

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. [Online.] Available at https://www.r-project.org/.

Salewski, V., and C. Watt (2017). Bergmann's rule: a biophysiological rule examined in birds. Oikos 126:161–172. doi: 10.1111/oik.03698

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. Methods in Ecology and Evolution 1:103–113. doi: 10.1111/j.2041-210X.2010.00012.x

Scholander, P. F. (1955). Evolution of Climatic Adaptation in Homeotherms. Evolution 9:15–26.

Symonds, M. R. E., and G. J. Tattersall (2010). Geographical variation in bill size across bird species provides evidence for Allen's rule. The American naturalist 176:188–97. doi: 10.1086/653666

Teplitsky, C., and V. Millien (2014). Climate warming and Bergmann's rule through time: Is there any evidence? Evolutionary Applications 7:156–168. doi: 10.1111/eva.12129

Teplitsky, C., J. a Mills, J. S. Alho, J. W. Yarrall, and J. Merilä (2008). Bergmann's rule and climate change revisited: disentangling environmental and genetic responses in a wild bird population. PNAS 105:13492–13496. doi: 10.1073/pnas.0800999105

Yom-Tov, Y., and E. Geffen (2011). Recent spatial and temporal changes in body size of terrestrial vertebrates: Probable causes and pitfalls. Biological Reviews 86:531–541. doi: 10.1111/j.1469-185X.2010.00168.x

Tables

Table 1. Multiple linear regression for house sparrow mean body mass. The mean mass was calculated for all 30 sample locations and for males and females separately. Significant effects are in bold. Summer maximum temperature had a significant negative relationship for all three models. Winter minimum was not a significant predictor in any of the three models.

Male mass					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	-0.092	0.116	-0.790	27	0.436
Summer max	-0.549	0.120	**-4.562**	27	**<0.001**
Winter min	-0.191	0.132	-1.450	27	0.159
Multiple R^2			Adjusted R^2 = **0.529**		R^2 = **0.561**
Female mass					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	-0.061	0.121	-0.504	27	0.619
Summer max	-0.728	0.129	**-5.644**	27	**<0.001**
Winter min	0.069	0.135	0.509	27	0.615
Multiple R^2			Adjusted R^2 = **0.530**		R^2 = **0.562**
All birds mass					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	-0.043	0.119	-0.365	27	0.718
Summer max	-0.686	0.124	**-5.513**	27	**< 0.001**
Winter min	-0.087	0.134	-0.646	27	0.524
Multiple R^2			Adjusted R^2 = **0.567**		R^2 = **0.597**

Table 2. Multiple linear regression for house sparrow mean tarsus length. The mean tarsus length was calculated for all 30 sample locations and for males and females separately. Significant effects are in bold. Summer maximum temperature was significantly negatively related to tarsus length for females and all birds combined however the same negative trend is not significant for males. Winter minimum was not a significant predictor in any of the three models.

Male tarsus length					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.230	0.172	1.335	27	0.193
Summer max	-0.301	0.178	-1.693	27	0.102
Winter min	-0.063	0.195	-0.321	27	0.751
Multiple R^2			Adjusted R^2 = 0.069		R^2 = 0.133
Female tarsus length					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.118	0.148	0.801	27	0.430
Summer max	-0.560	0.158	**-3.543**	27	**0.001**
Winter min	-0.100	0.166	-0.601	27	0.553
Multiple R^2			Adjusted R^2 = **0.337**		R^2 = **0.383**
All birds tarsus length					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.197	0.171	1.154	27	0.259
Summer max	-0.460	0.179	**-2.568**	27	**0.016**
Winter min	-0.082	0.193	-0.426	27	0.673
Multiple R^2			Adjusted R^2 = **0.196**		R^2 = **0.252**

Table 3. The LMM's for house sparrow individual body mass and tarsus length measurements. These LMM's use sampling location as a random factor and the measurements from 30 populations across Australia and New Zealand (n = 1901 individuals). Significant effects are in bold. Body mass had a significant negative relationship with summer max temperature but not winter minimum. For mass males were bigger than females but the difference between the sexes is only marginally significant. Tarsus length decreased with increasing summer maximum. For tarsus males were significantly bigger than females but the semi-partial correlation value is low indicating a small size difference between the sexes.

Body Mass					
Fixed effects	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.010	0.068	0.140	25.2	
Summer max	-0.339	0.071	**-4.804**	25.5	**<0.001**
Winter min	-0.007	0.066	-0.102	25.8	0.919
Sex	0.043	0.021	**1.998**	1885	**0.046**
	Variance	SD	n		ICC
Marginal R^2	0.120				0.116
Location	0.111	0.334	30		0.107
Residual	0.804	0.897	1901	Conditional R^2	0.223
Tarsus length					
Fixed effects	Semi-partial *r*	SE	*t*	df	*P*
Intercept	-0.085	0.064	-1.334	26.9	
Summer max	-0.134	0.066	**-2.037**	27.3	**0.051**
Winter min	-0.103	0.061	-1.675	27.8	0.105
Sex	0.076	0.022	**3.423**	1888.5	**<0.001**
	Variance	SD	n		ICC
Marginal R^2	0.046				0.046
Location	0.093	0.306	30		0.093
Residual	0.867	0.931	1901	Conditional R^2	0.138

Figures

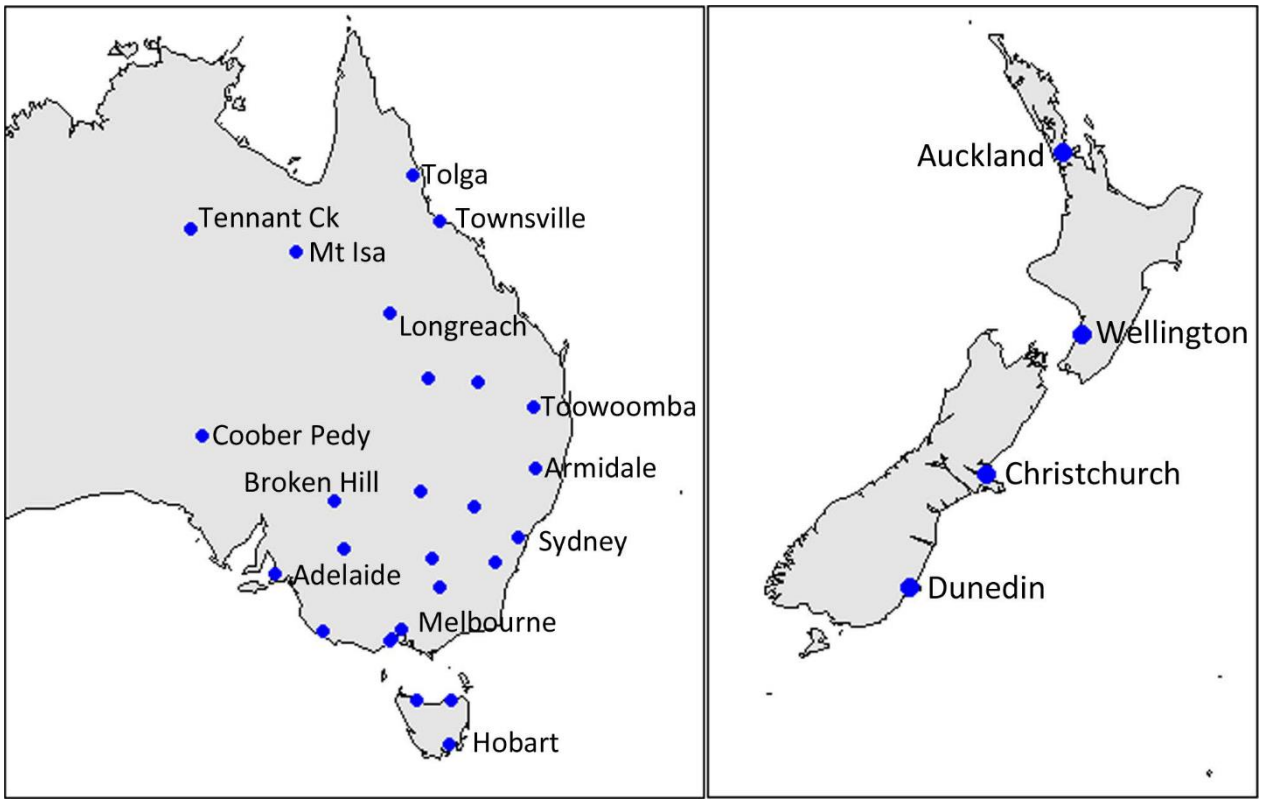


Figure 1. Map of house sparrow sample sites. The map plots 26 catch sites in Australia (there are three points near Melbourne) and 4 in New Zealand. Not all catch sites are labelled for a full list see Table S1.

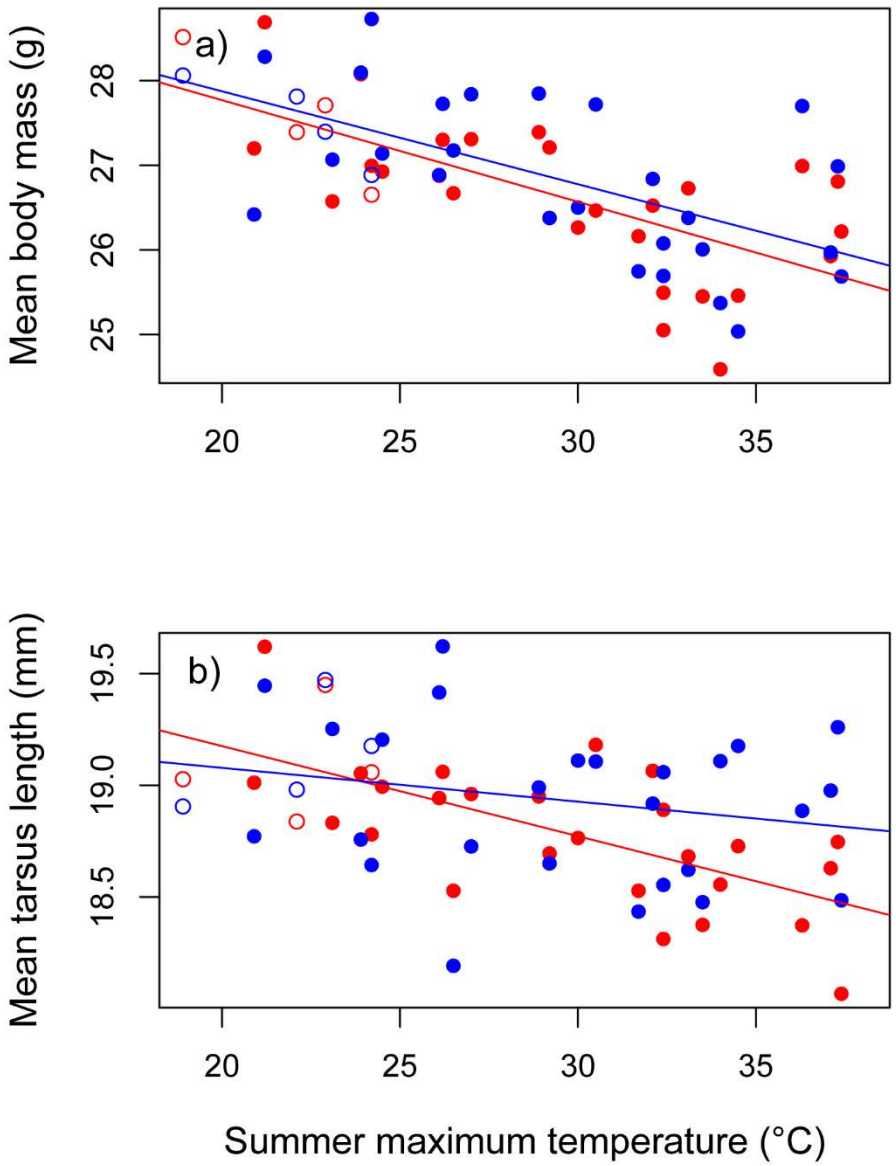


Figure 2. Mean body size for house sparrow populations. Dependent variables are plotted against the summer maximum temperature. Males and females are plotted separately with males as blue circles and females as red circles. The four New Zealand populations are marked as open circles. (a) plots mean mass against summer maximum, both males and females had significant negative relationships (Table 1). (b) plots mean tarsus length against summer maximum for male and female sparrows and there was a significant negative relationship for females (Table 2).

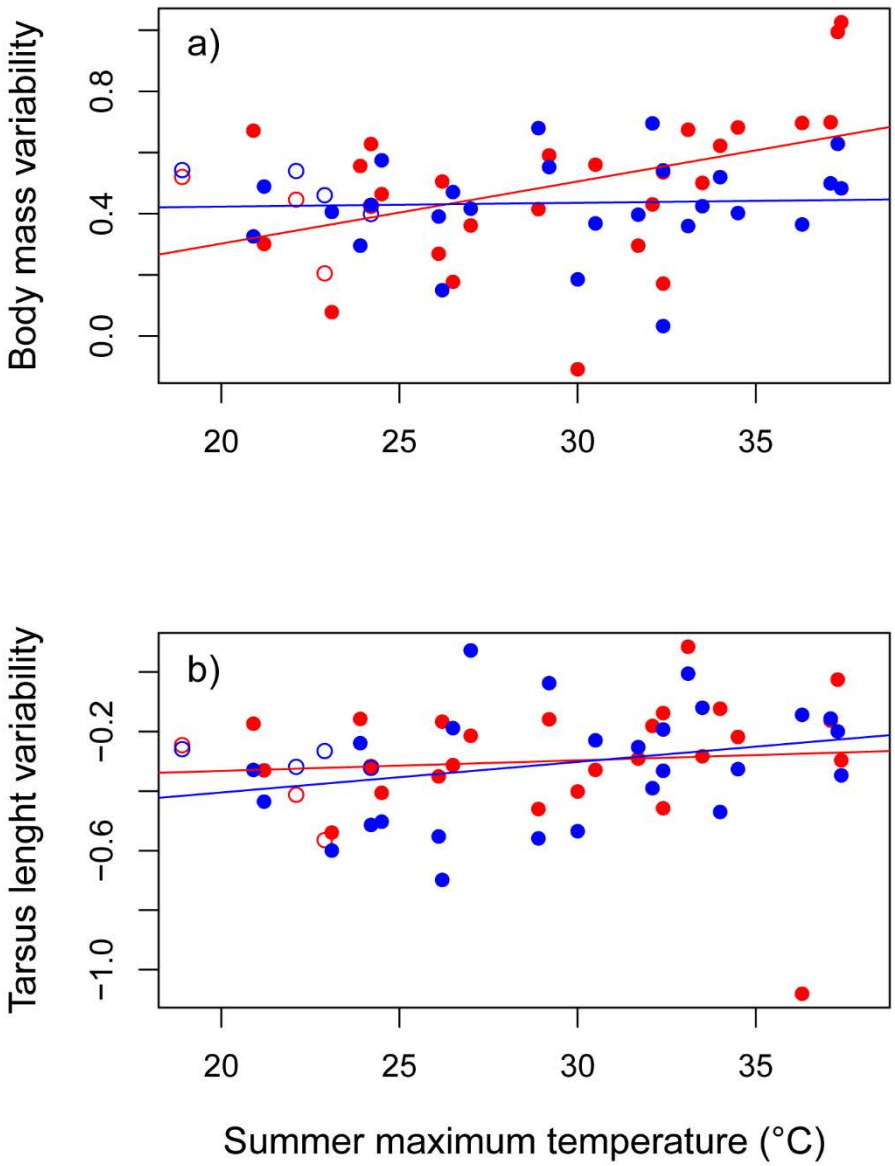


Figure 3. Body size variability for house sparrow populations. Variability is plotted against the summer maximum. Males and females are plotted separately with males as blue circles and females as red circles. The four New Zealand populations are marked as open circles. Panel (a) plots mass variability against summer max and there was a significant positive relationship for females but not for males. For mass variability there was also a significant positive relationship for all birds combined (Table S3). Panel (b) plots tarsus variability against summer maximum temperature and there was no significant positive relationship (Table S4).

Supplementary Materials for Chapter Three

Clinal variation in avian body size is better explained by summer maximum temperatures during development and not cold winter temperatures

Samuel C. Andrew, Monica Awasthy, Amanda D. Griffith, Shinichi Nakagawa, Simon C. Griffith

correspondence to: samuel.andrew@students.mq.edu.au

Table S1. Summary of house sparrow sampling populations. This table includes all the sample locations and sample sizes for all the house sparrow populations. This table also includes climatic variables.

Sample Population	Country	Latitude	Longitude	no. all birds	no. females	no. males	Summer max	Winter min	Season-ality
Dubbo	Aus	-32.2214	148.6262	39	17	22	32.1	2.8	5609
Cobar	Aus	-31.4923	145.8299	42	12	30	33.5	4.6	5873
Broken Hill	Aus	-31.9464	141.4654	43	17	26	32.4	4.1	5512
Wentworth	Aus	-34.1048	141.9166	25	11	14	33.1	4.4	5143
Leeton	Aus	-34.5619	146.4154	49	20	29	32.4	3.3	5686
Goulburn	Aus	-34.7633	149.6993	46	23	23	27	0.5	4838
Torquay	Aus	-38.317	144.299	34	13	21	23.9	5.6	3261
Geelong	Aus	-38.1783	144.3715	39	16	23	24.5	5.6	3357
Melbourne	Aus	-37.7888	144.9149	43	20	23	26.2	5.5	3775
Albury	Aus	-35.8373	146.8022	69	37	32	30.5	2.5	5487
Townsville	Aus	-19.3191	146.8238	44	18	26	31.7	13.1	3098
Atherton	Aus	-17.2145	145.4795	41	19	22	29.2	10.9	2687
Armidale	Aus	-30.5161	151.6735	47	21	26	26.1	0	4742
Toowoomba	Aus	-27.7218	151.6318	42	14	28	30	3.5	4719
Roma	Aus	-26.5617	148.791	42	18	24	34	3.7	5480
Charleville	Aus	-26.403	146.2511	47	22	25	34.5	3.9	5886
Longreach	Aus	-23.4358	144.2554	43	13	30	37.3	7	5363
Mount Isa	Aus	-20.7296	139.5025	104	49	55	37.1	8.3	4936
Tennant Creek	Aus	-19.6567	134.1924	39	12	27	37.4	11.1	4668
Hobart	Aus	-42.9358	147.3513	45	17	28	20.9	3.3	3014
Bridport	Aus	-41.0001	147.3875	43	22	21	23.1	4.3	3154
Wynyard	Aus	-40.9719	145.6532	44	20	24	21.2	4.6	2841
Coober Pedy	Aus	-29.0063	134.748	18	11	7	36.3	5.9	5715
Adelaide	Aus	-35.2324	138.4888	30	18	12	26.5	7.5	3507
Mt Gambier	Aus	-37.8565	140.8481	37	30	7	24.2	5.4	3053
Sydney	Aus	-33.6197	150.8163	53	22	31	28.9	3.7	4414
Auckland	NZ	-36.8649	174.7757	210	72	138	24.2	6.9	3106
Wellington	NZ	-40.6229	175.2877	257	71	186	22.9	4.9	3213
Christchurch	NZ	-43.5318	172.6269	205	77	128	22.1	1.9	3736
Dunedin	NZ	-45.8764	170.4957	81	22	59	18.9	1.7	3348

Table S2. Relationship between latitude and summer maximum and winter minimum temperature. Latitude was correlated with both summer and winter temperatures. The highest *t* value and most significant relationship was for summer maximum temperature. However, there was no significant correlation between summer and winter temperature for our 30 locations (estimate = 0.634, t_{28} = 1.951, P = 0.061, R^2 = 0.120). The summer maximum (of the hottest month) was also correlated with seasonality which is a metric for climate variability (estimate = 0.038, t_{28} = 6.505, P < 0.001, R^2 = 0.602). There was not as strong of a correlation between breeding season temperature range and summer maximum (estimate = 0.177, t_{28} = 2.631, P = 0.014, R^2 = 0.198).

	Estimate	SE	*Z*	df	*P*
Intercept	65.353	3.632	17.993	27	
Summer max	-0.956	0.132	**-7.221**	27	**<0.0001**
Winter min	-0.985	0.243	**-4.059**	27	**<0.001**
Multiple R^2			Adjusted R^2 = 0.789, R^2 = 0.773		

Table S3. Multiple regression for house sparrow mass variability per population. We found a significant positive relationship between body mass variability and summer maximum temperature for female birds and all birds combined. The positive slope for summer maximum for male birds was not significant. A positive slope shows that in warmer climates there is more variability in body mass at the population level. For the variable temperature range across the breeding season there was no significant effect.

Male Mass					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.097	0.168	0.577	26	0.569
Summer max	0.398	0.262	1.515	26	0.142
Temp range	-0.313	0.209	-1.498	26	0.146
Mean mass	0.456	0.270	1.691	26	0.103
Multiple R^2			Adjusted R^2 = 0.040		R^2 = 0.140
Female Mass					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.009	0.146	0.064	26	0.949
Summer max	0.640	0.221	**2.902**	26	**0.007**
Temp range	-0.015	0.172	-0.086	26	0.932
Mean mass	0.305	0.236	1.296	26	0.206
Multiple R^2			Adjusted R^2 = 0.200		R^2 = **0.283**
All birds Mass					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.070	0.156	0.447	26	0.659
Summer max	0.646	0.249	**2.594**	26	**0.015**
Temp range	-0.126	0.185	-0.681	26	0.502
Mean mass	0.456	0.248	1.837	26	0.078
Multiple R^2			Adjusted R^2 = 0.117		R^2 = 0.209

Table S4. Multiple regression for house sparrow tarsus variability per population. There was no significant relationship in all three models, for variability in tarsus length. There was a consistent nonsignificant positive slope for summer maximum temperature. A positive slope would mean in warmer climates there is more variability in tarsus length at a population level. For temperature range across the breeding season there was no significant effect.

Male Tarsus length					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.177	0.151	1.169	26	0.253
Summer max	0.020	0.166	0.123	26	0.903
Temp range	0.045	0.179	0.249	26	0.805
Mean tarsus	-0.321	0.163	-1.970	26	0.060
Multiple R^2			Adjusted R^2 = 0.069		R^2 = 0.165
Female Tarsus length					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	-0.042	0.156	-0.270	26	0.790
Summer max	0.267	0.216	1.232	26	0.229
Temp range	0.032	0.180	0.176	26	0.862
Mean tarsus	-0.010	0.203	-0.050	26	0.961
Multiple R^2			Adjusted R^2 = 0.018		R^2 = 0.120
All birds Tarsus length					
	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.085	0.154	0.552	26	0.586
Summer max	0.246	0.185	1.331	26	0.195
Temp range	0.054	0.179	0.299	26	0.768
Mean tarsus	-0.135	0.168	-0.800	26	0.431
Multiple R^2			Adjusted R^2 = 0.099		R^2 = 0.193

Chapter Four

Higher temperatures during development constrain body size in the zebra finch in the wild and under experimental conditions in the laboratory

Chapter Four Vignette

In Chapter Three we find that summer temperatures are a stronger predictor of clinal variation in house sparrow body size than winter temperatures. In Chapter Four we take the spatial observational data from the house sparrow one step further by look at temporal variation in the body size of zebra finch fledglings that are developing at different temperatures across the breeding season. Finally, we use a controlled captive experiment to look at the effects of temperature during development on the ultimate body size attained. This chapter aims to test the hypothesis that avian body size is negatively affected by higher temperatures during development.

This manuscript is submitted to the Journal of Evolutionary Biology, and is formatted as such.

Authors: Samuel C. Andrew[1]*, Laura L. Hurley[1], Mylene M. Mariette[1,2], Simon C. Griffith[1,3]

Affiliations:

[1]Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia

[2]Current address School of Life and Environmental Sciences, Deakin University, Waurn Ponds, Geelong, Victoria 3220, Australia

[3]School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, NSW 2052, Australia

Running head: Effects of temperature on body size

Corresponding author: Samuel C. Andrew

E-mail: samuel.andrew@students.mq.edu.au

phone: +61 2 9850 1305, fax: +61 2 9850 9231

Department of Biological Sciences

Macquarie University

Sydney, NSW, 2109, Australia.

Acknowledgments: We would like to thank Mitchell Francis for assistance with the collection of the captive zebra finch data at Macquarie University. We thank Daisy E. Duursma for assistance with climatic data. We would like to thank Monica Awasthy, Amanda D. Griffith, Kate Buchanan, Sasha Dall, Peter Dunn, Janet Gardner, Tim Parker, and Andy Russell for comments on earlier drafts of the manuscript. Funding sources: SCA was supported by Macquarie University Research Excellence Scholarship (no. 2013077). SCG was supported by an Australian Research Council Future Fellowship (FT130101253). Authors' contributions: SCA carried out the statistical analyses, participated in the design of the study and drafted the manuscript; LLH carried out captive zebra finch experiments; MMM collected zebra finch field data; SCG collected data, designed the study, coordinated the study and helped draft the manuscript. All authors reviewed the manuscript and gave final approval for publication.

Abstract

The most commonly documented morphological response across many taxa to climatic variation across their range follows Bergmann's rule, which predicts larger body size in colder climates. Using here observational data from wild zebra finches breeding across a range of temperatures in the spring and summer, we show that this relationship appears to be driven by the negative effect of high temperatures during development. This idea was then experimentally tested on zebra finches breeding in temperature controlled climates in the laboratory. These experiments confirmed that those individuals produced in a hotter environment (30°C) were smaller than those produced in colder conditions (18°C). Our results suggest a proximate causal link between temperature and body size, and suggest that a hotter climate during breeding periods could drive significant changes in morphology within and between populations. This effect could account for much of the variation in body size that drives the well observed patterns first described by Bergmann, and that are still largely attributed to selection on adult body size during cold winters. The climate-dependent developmental plasticity that we have demonstrated is an important component in understanding how endotherms may be effected by climate change.

Keywords: Bergmann's rule, phenotypic plasticity, morphology, *Taeniopygia guttata*.

Introduction

Since the publication of Bergmann's Rule in 1847 (Bergmann, 1847) over a century of observational studies have supported the pattern of endothermic animals having larger body size in colder climates in a majority of hundreds of species (Mayr, 1956; James, 1970; Ashton *et al.*, 2000; Ashton, 2002; Meiri & Dayan, 2003; Millien *et al.*, 2006; Clauss *et al.*, 2013; Teplitsky & Millien, 2014). To date, many studies still cite Bergmann's original explanation that larger body size is favored by natural selection in colder climates because of the thermo-regulatory benefits of a smaller volume-to-surface area ratio (Briscoe *et al.*, 2015; Cardilini *et al.*, 2016; Salewski & Watt, 2017). In hotter climates smaller body size can also be advantageous to thermo-regulation (Partridge & Coyne, 1997), even though the benefits of minor changes in body size within species have been questioned for over 40 years (Scholander, 1955; McNab, 1971). Understanding the phenomena underlying Bergmann's rule has gained fresh impetus as part of the study of the effects of a changing climate on animal populations (Millien *et al.*, 2006; Gardner *et al.*, 2011). There is abundant evidence of the intuitive direct effects of environmental temperatures on growth and development in ectotherms (Walters & Hassall, 2006; Zuo *et al.*, 2012). However, as endotherms maintain their body temperature within a fairly narrow thermal range, traditionally, it has seemed less intuitive that they would be directly effected by climatic conditions. As such, the most widely cited reason for clinal relationships in body size is selection (Ashton *et al.*, 2000).

The initially neglected alternative explanation, that temporal changes in body size are largely the result of temperature-mediated plasticity in growth rather than selection, is now gaining traction (Teplitsky *et al.*, 2008; Van Buskirk *et al.*, 2010; Yom-Tov & Geffen, 2011), but support remains limited by a lack of experimental evidence (Stillwell, 2010; Knudsen *et al.*, 2011; Boutin & Lane, 2014; Gienapp & Merilä, 2014; Merilä & Hendry, 2014). Although birds are endothermic, they can become heat stressed when they are unable to balance the production and loss of heat (Calder, 1964), and even short periods of high temperature can

promote the excessive production of reactive oxygen species (Lin *et al.*, 2008) and changes in the oxidative status of an individual (Costantini *et al.*, 2012). Nestlings of altricial species are particularly vulnerable to heat stress because they are at a sensitive stage in their development; are poikilothermic up to about five days of age; and confined to a nest that may become hotter than ambient conditions due to its position. For example, nests of the zebra finch, *Taeniopygia guttata* in the Australian desert were typically several degrees hotter than ambient conditions and internal nest temperatures occasionally exceeded 50 ºC (Griffith *et al.*, 2016). In a recent study, Wada *et al.* (2015) demonstrated long-lasting effects of different sub-optimal incubation temperatures on survival, physiological function and body mass in this same species in the laboratory, indicating a sensitivity to environmental temperatures.

A classic example of clinal variation in avian body size was previously demonstrated in both North America and Europe populations of the house sparrow, *Passer domesticus*, (Johnston & Selander, 1964, 1973; Murphy, 1985). If winter survival is the selective force responsible for this geographical pattern as predicted by Bergmann's rule, body size should be best explained by winter minimum temperature. In addition to clinal variation across latitudes, declining body size in a number of avian species has been linked to increasing temperatures consistent with climate change (Gardner *et al.*, 2009; Van Buskirk *et al.*, 2010). Although the changing climate will raise winter minimums, it will also lead to higher temperatures during avian breeding seasons (Knutti & Sedláček, 2013), and that could result in reduced body size during development.

Experimentally testing the effect of environmental temperature on development in endotherms has rarely been conducted, with two previous laboratory studies of mammals (Barnett & Dickson, 1984; Riek & Geiser, 2012), and none of post-natal birds (Teplitsky & Millien, 2014). Here we used an endemic Australian passerine that is amenable to intensive study in the field and laboratory, the zebra finch *Taeniopygia guttata*, to test the hypothesis that high temperatures during post-natal development impact growth.

The only previous examination of clinal variation in the zebra finch across a latitudinal gradient in Australia provides some weak evidence that is consistent with Bergmann's Rule (a nonsignificant cline in wing length across its extensive latitudinal distribution; $P = 0.07$, $R^2 = 0.09$, $n = 37$, see Higgins et al. [2006]). The zebra finch is highly mobile with no apparent genetic structure across the Australian continent (Balakrishnan & Edwards, 2009), and consequently, adults captured at any location might have developed under a different climate to that in which they are caught. However, at any particular location, zebra finches breed across many months of the year and at a population level the developing young from successive broods hatched from late winter to late summer are exposed to a substantial range of weather conditions with maximum daily nest temperatures ranging from below 15 ºC to above 40 ºC in a field study of nest temperature during breeding (Griffith *et al.*, 2016). The zebra finch therefore provides the opportunity to address the consequence of climatic variation during breeding through an observational study in the field, across a range of weather conditions, at a single site. Furthermore, they are relatively easy to breed in captivity and this permits the experimental study of climatic effects in controlled climate laboratory conditions, which can remove confounding effects of parental, seasonal and other environmental variation.

Materials and methods

Wild zebra finch nestlings (n=901 nestlings) were measured at Fowlers Gap Arid Zone Research Station, Western NSW, Australia, over three consecutive breeding seasons in 2007, 2008 and 2009. Nest-boxes were checked every one to three days to record laying, hatching and fledging dates, as well as clutch size and brood size. When nestlings were one to three days old, half of the nestlings in each brood were cross-fostered between two or three nests hatched within two days of each other to facilitate brood size manipulations (under Animal Ethics Approval ARA ARA2007/038; further details in Mariette and Griffith [2013]). All

surviving nestlings were then measured and weighed shortly before fledging, at about 12-days old (mean days old = 12.54 ± 1.00 SD). We were limited to measuring tarsus at that point because nestlings will emerge from the nest prematurely if disturbed after around day 14, and given their mobility after fledging, we rarely recapture offspring measured in the nest as adults.

Captive zebra finches were reared under controlled conditions at Macquarie University (under Animal Ethics approval ARA 2013/029). In short, 24 pairs of domestically bred zebra finches were force paired and housed as pairs (70x47x130cm cages) in temperature controlled rooms using a repeat breeding experimental design. Rooms were then set to 18ºC or 30ºC (the 12 pairs in each treatment were swapped to the opposite treatment after the first breeding attempt) for two weeks prior to giving the birds two 13.5x15cm rattan nests and nesting material (November grass, white cotton thread, and emu feathers). The first clutch that each pair laid was removed once it was completed (for other work), and each pair was then allowed to lay a second clutch that they were left to rear. Birds were exposed to the experimental temperature and supplemented food (green pea-spinach mash with hard-boiled egg, and sprouted seed) for four weeks before laying the clutch of eggs that produced chicks for this study. Chicks were reared and maintained at the experimental temperature for at least 35-days post hatch, at which time they were moved to a 25ºC holding room. Morphological measures were taken at post hatch day 12, 28, and over 90 days for the final adult offspring measure. Adult offspring measurements were taken for all birds at the same time in the 25ºC holding rooms. After the completion of a whole reproductive attempt at 18ºC or 30ºC, pairs were switched to the other temperature and the whole procedure was repeated. As usual for this species in captivity (Griffith *et al.*, 2017), some pairs failed to breed in either one of the trials and therefore only sixteen pairs successfully bred at both temperatures and were used in the models to test the effect of temperature on growth. The analysis is therefore a within-family balanced design. Nine of the pairs bred at 18ºC first and then 30ºC, and the other seven pairs bred in the hot conditions first.

Tarsus length and mass were recorded for both wild and captive zebra finches using the same method. Tarsus length was measured on the right leg, from the bottom of the tarsus with the toes bent forward to the ankle joint. Body mass was taken to the nearest 0.1g using a Pesola spring scale. Tarsus length is the most widely applied measure of skeletal size in passerine birds, and body mass relates to overall size as well as being more indicative of condition in passerine birds. The two measures, tarsus and body mass, therefore both provide information on body size, but in slightly different ways. All wild and captive zebra finch measurements were taken by a single measurer (wild: M.M.M.; captive L.L.H.).

Climate data for wild zebra finches

Daily temperature data for the Fowlers Gap research station were extracted from the Australian Water Availability Project (Jones *et al.*, 2009) website (www.csiro.au/awap). We estimated both the mean maximum and minimum temperatures experienced by each nestling during development, from hatching to 10-days old. To do so, we calculated the average daily maximum and average daily minimum temperatures for that 10-day period, from the hatching date of each individual nestling in the dataset. These two variables were highly correlated (R^2 = 0.78) so all further analyses just use maximum temperature as a predictor.

Data analysis

All statistical analyses were done using the R statistical program, version 3.3.1 (R Core Team 2017). All R code and data to replicate our main results is provided as supplementary material (see Figshare, link: https://figshare.com/s/dbc245f85f4cc3cdd5b1).

For the wild zebra finch data we predicted that body size would be smaller when nestlings developed during hot weather periods. To test this hypothesis we used Linear Mixed models (LMM) fitted using the R package *lme4* (Bates *et al.*, 2015). For the LMM's, P-values and degrees of freedom were calculated by the R package *lmerTest* (Kuznetsova *et al.*, 2016) when running models using *lme4*. Predictors in the model were mean maximum temperature

during development, brood size (number of chicks after cross fostering) and age (days since hatching) as continuous fixed effects (Breeding season was also trialled as a fixed effect). The continuous fixed effect of offspring age on the day of measurement was included to account for its confounding effect on size. The temperatures recorded for broods within months also increased from August to December as we would expect (estimate = 2.299, $t_9 = 3.915$, $P = 0.004$, $R^2 = 0.630$) meaning the two variables of month and temperature are conflated and were therefore not included in the same model. As broods were cross-fostered in this study, we used genetic clutch ID (broad-sense heritability) and foster nest box (common environment and parental care) as random factors with independent intercepts. These two random effects allowed the models to account for the genetic and environmental effects on growth that are not explained by the included fixed effects. Inter-class correlation coefficient (ICC) were also calculated for the random effects to describe how much variation is explained by them (Nakagawa & Schielzeth, 2010). Statistics such as ICC, R^2 and narrow sense heritability (h^2) all estimate the proportion of the variance in the response variable that is explained by a fixed or random effect (Nakagawa & Schielzeth, 2013). Because it is a proportion it can be compared between similar models such as our LMM that share the same fixed and random effects. To calculate the proportion of variance explained by random factors, the residual variance and variance explained by fixed effects (known as Marginal R^2) need to be included. As a result the total variance explained by the model can also be calculated which is called the Conditional R^2 (see Nakagawa and Schielzeth [2013]). We report the Marginal (fixed effects) and Conditional (total model) R^2 for all models.

We also ran models for tarsus and mass with a reduced data set that included only pairs that had two clutches recorded. These models had the same fixed effects and the random factor of clutch ID, which had random slopes and intercepts by using the interaction between the fixed effect of temperature and clutch ID. Sometimes both parents were not tagged (see methods Mariette and Griffith [2013]) so if an individual was observed breeding twice in the same season with an unknown partner we assumed it was the same pair of birds. Paternity

was not considered in this study because extra pair paternity is low in wild populations of zebra finch and accounts for only about 2% of nestlings (Griffith *et al.*, 2010).

For the captive zebra finch study temperature treatment and brood size were used as fixed effects. Breeding pair ID was used as a random effect with random intercepts and slopes by using the interaction between temperature and pair ID. The interaction can be used because each pair reared a clutch at both temperatures due to the repeat-breeding experimental design. Clutch order was not a significant factor in the model and did not change the temperature result so was excluded to simplify the model. Calculating ICC, Marginal R^2 and Conditional R^2 for random slope models is different from models with just random intercepts. See Johnson (2014) for a description of the method and the supplementary R code for examples.

Results

Wild zebra finch nestling size

In an analysis of approximately 900 zebra finch nestlings at a single site, the mean maximum temperature experienced across the first 10 days of each individuals' development was used to test the effect of temperature on body size. Mean maximum temperature experienced during growth negatively affected the body mass attained by around day 12 (Table 1; Fig. 1). Tarsus length showed a negative yet nonsignificant relationship (Table 1). Tarsus length has not typically finished growing at 12 days of age. In the wild birds measured here we found that day 12 tarsus length is typically around 95.8% of adult tarsus length (t-test, t_{253} = -8.49, $P <$ 0.001, fledglings were compared to a random sample of 108 wild adults).

The effects of common nest environment and relatedness also suggest higher plasticity in mass than tarsus in wild zebra finches (see random factors Table 1). This negative co-variation of body mass and maximum temperature was detected above the effect of a brood size manipulation, which corrected for seasonal variations in reproductive investment (brood size) and parental effort per offspring. In this overall dataset, zebra finches reared in the

hottest 20 % of conditions (mean 33.2 ºC, range 31.2- 41.4 ºC) were 4.1 % lighter in mass than finches reared during the coolest 20 % of conditions (mean 25.1 ºC, range 20.6 - 26.6 ºC). However, there may be variation in the quality of adults breeding at different times of the season and food quality. To address this, we used a reduced dataset of offspring produced by the same pair (n = 34 pairs; n = 242 offspring) in different broods (and therefore at different times of the season). As in the full data set there was a significant negative relationship between maximum temperature and nestling mass (Table 2) and similarly a negative relationship between nestling tarsus length and temperature, albeit a nonsignificant one (Table 2). For these pairs that produced two sets of offspring in different conditions, we divided broods between those that developed in the hottest (mean 30.9 ºC, range 26.0 - 34.9 ºC) and coolest temperatures (mean 26.7 ºC, range 20.9 – 31.2 ºC). For these groups mass was 7.9 % lighter, and tarsus was 0.9 % shorter for the nestlings reared in the hotter conditions, compared with their full siblings reared in the cooler conditions.

Wild zebra finches were sampled across three breeding seasons (2007-2009). The temperatures recorded for broods from 2009 were higher than 2007 (Tukey's test, diff = 1.39, $P < 0.001$) and 2008 (Tukey's test, diff = 1.37, $P < 0.001$). Interestingly when we add year as a factor in our models the 2009 cohort was also smaller (mass difference between 2007 vs 2009; estimate = -0.76, $t_{237} = -4.399$, $P < 0.001$), this effect was also true for tarsus (2007 vs 2009; estimate = -0.508, $t_{237} = -3.043$, $P < 0.001$). There was no significant difference between 2007 and 2008 for temperature, mass and tarsus. However, adding year to the models did not change how we would interpret the main results for temperature (Table 1) so it was excluded due to the collinearity between year and temperature, to improve the interpretability of the model (but can be evaluated using the supplementary data and R code). Due to collinearity, the two predictors would explain a shared portion of the variance in the model that is impossible to separate in this observational study.

Climate Controlled zebra finch experiment

Data from our climate controlled breeding experiment was consistent with the patterns observed in the wild. Zebra finch offspring reared at 30°C were significantly smaller at 28 days post-hatching, than their full siblings reared at 18°C (Fig. 2). In this case, temperature negatively impacted both mass and tarsus (Table 3). Zebra finches produced in the hot treatment had tarsi that were 1.2 % shorter and were 5% lighter than those produced in the cool treatment. The random factor of breeding pair in these models explained 34.4% of the variance in mass and 21.8% of the variance in tarsus length. The factor of brood size was non-significant in both models (Tables 3). In these captive bred birds, day 12 tarsus was 96.6% of the final tarsus reached when fully grown (paired t-test, $t_{105} = -7.45$, $P < 0.001$). Day 12 tarsus length measurements were not highly correlated with day 28 measurements (linear model, estimate = 0.93, $t_{97} = 6.07$, $P < 0.0001$, $R^2 = 0.28$) and the same was true for both day 12 and 28 mass measurements (linear model, estimate = 0.77, $t_{97} = 6.32$, $P < 0.0001$, $R^2 = 0.29$).

Discussion

Using temporal variation in weather conditions in a wild population of zebra finches, we found that fledglings that hatched during hotter periods attained less mass than those that developed during cooler periods, supporting the hypothesis that temperature negatively effects growth. However, as with other studies (Cunningham *et al.*, 2013; Kruuk *et al.*, 2015), this difference in offspring size across a season might have been confounded by unmeasured parameters such as foraging conditions, differential mortality of nestlings due to an interaction between size and temperature or adult body condition, and parental investment, that may also vary temporally. This limitation is common throughout the largely observational literature that has examined Bergmann's rule. We did not find that tarsus length was significantly smaller in those nestlings that developed in hot conditions, but in these wild nestlings we were unable to measure this skeletal measure when fully developed.

To address the limitations of work on wild birds (both in terms of not being able to measure offspring as adults, and uncontrolled sources of variation), we bred birds in experimental climates with a within-pair repeat-breeding design that enabled us to control for most genetic and parental variation, and found further support for smaller body size at higher temperatures. Offspring produced in the hot climate were smaller and lighter at 28 days (fully grown juvenile age), than their full siblings produced in the cooler climate. Interestingly, the differences were not statistically significant when the offspring were 12 days old (although the semi-partial r values were similar between the wild and captive models). This supports the idea that the effects of heat on tarsus length are perhaps less clear at this early stage of development compared to those on body mass.

Across both datasets, temperature effects were associated with reductions in size and mass, but the effects were of greater magnitude for body mass. This is consistent with much other work showing greater plasticity of body mass and lower heritability compared to skeletal measures (Alatalo *et al.*, 1990; Jensen *et al.*, 2003). Similarly, a study on North American migratory birds found increases in summer temperature caused a larger percentage decline in mass (0.34% per degree Celsius) than wing chord length (0.09% per degree Celsius) across 61 species measured in summer (Van Buskirk *et al.*, 2010).

The results presented here show a significant proportion of the variation in body size is explained by phenotypic plasticity influenced by increasing temperatures during growth. In this, our study extends the recent work by Wada *et al.* (2015) that demonstrated the effect of sub-optimal experimental effects of incubation temperature on development and growth. Our experimental hot treatment of 30°C was below the optimal incubation temperature of around 37 °C in this species (Zann & Rossetto, 1991), and the experimental temperatures used by Wada *et al*. (2015; 36.4-38.4°C). As a result, it is unlikely that our treatment changed the incubation thermal regime of the embryos in our experiment (which was provided by the heat transferred by incubating parents). As such our findings, are most likely to reflect either the

direct effect of the environmental temperature on the nestlings post-hatch, or perhaps indirectly through effects on their parents during the rearing period. It would be interesting for follow-up work to examine the extent to which the effects of temperature before and after hatching are additive or interact with each other.

Our findings are entirely consistent with the pattern described by Bergmann (1847), but in this case have not been produced by selection on adult body size. If we accept thermoregulatory advantages of large and small bodies in cold and hot environments respectively (Mayr, 1956; Ashton, 2002), then the plasticity observed here could be considered as adaptive because it is in the direction we would expect local adaptation to take (Ghalambor *et al.*, 2007). Our study did not examine selection on body size at all, and therefore is unable to determine the influence of selection on changes in body size over climatic gradients. However, our results indicate that even in the absence of such selection, or indeed fixed genetic differences between populations, clinal patterns are likely to exist due to the plasticity that we have demonstrated. Our study indicates that the mechanisms underlying Bergmann's rule and its adaptive advantages are still not fully understood (Scholander, 1955; McNab, 1971; Partridge & Coyne, 1997). Future studies should explore the mechanistic link between the climate experienced during development and the body size attained (or the size of extremities as predicted by Allen's rule (Allen, 1877; Symonds & Tattersall, 2010)). Possible mechanisms include physiological constraints (Gardner *et al.*, 2009), constraints on parental provisioning (Cunningham *et al.*, 2013) and parental effects (Mariette & Buchanan, 2016). Our study did not address any of these potential underlying mechanisms, but our data do suggest that hot temperatures affect development, and provide justification for further work addressing these possible mechanisms. Our study leads to the prediction that increasing global temperatures will generate smaller body sizes, but it remains to be tested whether this will improve their ability to adapt to a warming climate or is just the result of constraints during development (van Gils *et al.*, 2016). Our conclusion that high ambient conditions affect

growth and development may also lead to the prediction that breeding phenology in birds, and perhaps animals more generally, should not be timed just in relation to the abundance of resources (Both, 2010), but also in respect to the avoidance of periods of climatic extremes (Duursma et al. 2017).

References

Alatalo, R. V., Gustafsson, L. & Lundberg, A. 1990. Phenotypic selection on heritable size traits: Environmental variance and genetic response. *Am. Nat.* **135**: 464–471.

Allen, J.A. 1877. The influence of physical conditions in the genesis of species. *Radic. Rev.* **1**: 108–140.

Ashton, K., Tracy, M. & Queiroz, A. De. 2000. Is Bergmann's rule valid for mammals? *Am. Nat.* **156**: 390–415.

Ashton, K.G. 2002. Patterns of within-species body size variation of birds: Strong evidence for Bergmann's rule. *Glob. Ecol. Biogeogr.* **11**: 505–523.

Balakrishnan, C.N. & Edwards, S. V. 2009. Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* **181**: 645–60.

Barnett, S.A. & Dickson, R.G. 1984. Changes among wild House mice (*Mus musculus*) bred for ten generations in a cold environment, and their evolutionary implications. *J. Zool.* **203**: 163–180.

Bates, D., Maechler, M., Bolker, B. & Walker, S. 2015. lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8, http://CRAN.R-project.org/package=lme4.

Bergmann, C. 1847. Ueber die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Gottinge Stud.* **3**: 595–708.

Both, C. 2010. Food availability, mistiming, and climate change. In: *Effects of Climate Change on Birds* (A. P. Møller et al., eds), pp. 129–149. OUP Oxford, Oxford, UK.

Boutin, S. & Lane, J.E. 2014. Climate change and mammals: Evolutionary versus plastic responses. *Evol. Appl.* **7**: 29–41.

Briscoe, N.J., Krockenberger, A., Handasyde, K. a. & Kearney, M.R. 2015. Bergmann meets Scholander: geographical variation in body size and insulation in the koala is related to climate. *J. Biogeogr.* **42**: 791–802.

Calder, A.W. 1964. Gaseous Metabolism and Water Relations of the Zebra Finch, *Taeniopygia castanotis*. *Physiol. Zool.* **37**: 400–413.

Cardilini, A.P.A., Buchanan, K.L., Sherman, C.D.H., Cassey, P. & Symonds, M.R.E. 2016. Tests of ecogeographical relationships in a non-native species: what rules avian morphology? *Oecologia* **181**: 783–93. Springer Berlin Heidelberg.

Clauss, M., Dittmann, M.T., Müller, D.W.H., Meloro, C. & Codron, D. 2013. Bergmann's rule in mammals: A cross-species interspecific pattern. *Oikos* **122**: 1465–1472.

Costantini, D., Monaghan, P. & Metcalfe, N.B. 2012. Early life experience primes resistance to oxidative stress. *J. Exp. Biol.* **215**: 2820–2826.

Cunningham, S.J., Martin, R.O., Hojem, C.L. & Hockey, P. a R. 2013. Temperatures in excess of critical thresholds threaten nestling growth and survival in a rapidly-warming arid savanna: A study of common fiscals. *PLoS One* **8**. e74613

Duursma, D.E., Gallagher, R.V. & Griffith, S.C. 2017. Characterizing opportunistic breeding at a continental scale using all available sources of phenological data; an assessment of 337 species across the Australian continent. *Auk* **134**: 509–519.

Gardner, J.L., Heinsohn, R. & Joseph, L. 2009. Shifting latitudinal clines in avian body size correlate with global warming in Australian passerines. *Proc. Biol. Sci.* **276**: 3845–52.

Gardner, J.L., Peters, A., Kearney, M.R., Joseph, L. & Heinsohn, R. 2011. Declining body size: a third universal response to warming? *Trends Ecol. Evol.* **26**: 285–91.

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. *Funct. Ecol.* **21**: 394–407.

Gienapp, P. & Merilä, J. 2014. Disentangling plastic and genetic changes in body mass of Siberian jays. *J. Evol. Biol.* **27**: 1849–1858.

Griffith, S.C., Crino, O.L., Andrew, S.C., Nomano, F.Y., Adkins-Regan, E., Alonso-Alvarez, C., *et al.* 2017. Variation in Reproductive Success Across Captive Populations: Methodological Differences, Potential Biases and Opportunities. *Ethology* **123**: 1–29.

Griffith, S.C., Holleley, C.E., Mariette, M.M., Pryke, S.R. & Svedin, N. 2010. Low level of extrapair parentage in wild zebra finches. *Anim. Behav.* **79**: 261–264. Elsevier Ltd.

Griffith, S.C., Mainwaring, M.C., Sorato, E. & Beckmann, C. 2016. High atmospheric temperatures and "ambient incubation" drive embryonic development and lead to earlier hatching in a passerine bird. *R. Soc. open Sci.* **3**: 150371.

Higgins, P.J., Peter, J.M. & Cowling, S.J. 2006. Handbook of Australian, New Zealand and Antarctic birds: Volume 7 Boatbill to Starlings. pp. 1102–26. Oxford University Press, Melbourne.

James, F.C. 1970. Geographic size variation in birds and its relationship to climate. *Ecology* **51**: 365–390.

Jensen, H., Saether, B.-E., Ringsby, T.H., Tufto, J., Griffith, S.C. & Ellegren, H. 2003. Sexual variation in heritability and genetic correlations of morphological traits in house sparrow (*Passer domesticus*). *J. Evol. Biol.* **16**: 1296–1307.

Johnson, P.C.D. 2014. Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods Ecol. Evol.* **5**: 944–946.

Johnston, R.F. & Selander, R.K. 1973. Evolution in the house sparrow. III. Variation in size and sexual dimorphism in Europe and North and South America. *Am. Nat.* **107**: 373–390.

Johnston, R.F. & Selander, R.K. 1964. House sparrows: rapid evolution of races in North America. *Science (80-.).* **144**: 548–50.

Jones, D.A., Wang, W. & Fawcett, R. 2009. High-quality spatial climate data-sets for Australia. *Aust. Meteorol. Oceanogr. J.* **58**: 233–248.

Knudsen, E., Lindén, A., Both, C., Jonzén, N., Pulido, F., Saino, N., *et al.* 2011. Challenging claims in the study of migratory birds and climate change. *Biol. Rev.* **86**: 928–946.

Knutti, R. & Sedláček, J. 2013. Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Clim. Chang.* **3**: 369–373.

Kruuk, L.E.B., Osmond, H.L. & Cockburn, A. 2015. Contrasting effects of climate on juvenile body size in a Southern Hemisphere passerine bird. *Glob. Chang. Biol.* **21**: 2929–2941.

Kuznetsova, A., Brockhoff, P.B. & Christensen, R.H.B. 2016. lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-32. https://CRAN.R-project.org/package=lmerTest.

Lin, H., De Vos, D., Decuypere, E. & Buyse, J. 2008. Dynamic changes in parameters of redox balance after mild heat stress in aged laying hens (*Gallus gallus domesticus*). *Comp. Biochem. Physiol.* **147**: 30–35.

Mariette, M.M. & Buchanan, K.L. 2016. Prenatal acoustic communication programs offspring for high posthatching temperature in a songbird. *Science (80-.).* **353**: 812–814.

Mariette, M.M. & Griffith, S.C. 2013. Does coloniality improve foraging efficiency and nestling provisioning? A field experiment in the wild Zebra Finch. *Ecology* **94**: 325–335.

Mayr, E. 1956. Geographical character gradients and climatic adaptation. *Evolution (N. Y).* **10**: 105–108.

McNab, B. 1971. On the ecological significance of Bergmann's rule. *Ecology* **52**: 845–854.

Meiri, S. & Dayan, T. 2003. On the validity of Bergmann's rule. *J. Biogeogr.* **30**: 331–351.

Merilä, J. & Hendry, A.P. 2014. Climate change, adaptation, and phenotypic plasticity: The problem and the evidence. *Evol. Appl.* **7**: 1–14.

Millien, V., Kathleen Lyons, S., Olson, L., Smith, F. a., Wilson, A.B. & Yom-Tov, Y. 2006. Ecotypic variation in the context of global climate change: Revisiting the rules. *Ecol. Lett.* **9**: 853–869.

Murphy, E.C. 1985. Bergmann's rule, seasonality, and geographic variation in body size of house sparrows. *Evolution (N. Y).* **39**: 1327–1334.

Nakagawa, S. & Schielzeth, H. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol. Evol.* **4**: 133–142.

Nakagawa, S. & Schielzeth, H. 2010. Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biol. Rev.* **85**: 935–956.

Partridge, L. & Coyne, J.A. 1997. Bergmann's rule in ectotherms: is it adaptative. *Evolution (N. Y).* **51**: 632–635.

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Riek, A. & Geiser, F. 2012. Developmental phenotypic plasticity in a marsupial. *J. Exp. Biol.* **215**: 1552–1558.

Salewski, V. & Watt, C. 2017. Bergmann's rule: a biophysiological rule examined in birds. *Oikos* **126**: 161–172.

Scholander, P.F. 1955. Evolution of climatic adaptation in homeotherms. *Evolution (N. Y).* **9**: 15–26.

Stillwell, R.C. 2010. Are latitudinal clines in body size adaptive? *Oikos* **119**: 1387–1390.

Symonds, M.R.E. & Tattersall, G.J. 2010. Geographical variation in bill size across bird species provides evidence for Allen's rule. *Am. Nat.* **176**: 188–97.

Teplitsky, C. & Millien, V. 2014. Climate warming and Bergmann's rule through time: Is there any evidence? *Evol. Appl.* **7**: 156–168.

Teplitsky, C., Mills, J. a, Alho, J.S., Yarrall, J.W. & Merilä, J. 2008. Bergmann's rule and climate change revisited: disentangling environmental and genetic responses in a wild bird population. *PNAS* **105**: 13492–13496.

Van Buskirk, J., Mulvihill, R.S. & Leberman, R.C. 2010. Declining body sizes in North American birds associated with climate change. *Oikos* **119**: 1047–1055.

van Gils, J.A., Lisovski, S., Lok, T., Meissner, W., Ożarowska, A., de Fouw, J., *et al.* 2016. Response to warming Arctic warming reduces red knot fitness in tropical wintering range. *Science (80-.).* **352**: 819–821.

Wada, H., Kriengwatana, B., Allen, N., Schmidt, K.L., Soma, K.K. & MacDougall-Shackleton, S.A. 2015. Transient and permanent effects of suboptimal incubation temperatures on growth, metabolic rate, immune function, and adrenocortical responses in zebra finches. *J. Exp. Biol.* **218**: 2847–2855.

Walters, R.J. & Hassall, M. 2006. The Temperature-Size Rule in ectotherms: May a general explanation exist after all? *Am. Nat.* **167**: 510–523.

Yom-Tov, Y. & Geffen, E. 2011. Recent spatial and temporal changes in body size of terrestrial vertebrates: Probable causes and pitfalls. *Biol. Rev.* **86**: 531–541.

Zann, R. & Rossetto, M. 1991. Zebra finch incubation: Brood patch, egg temperature and thermal properties of the nest. *Emu* **91**: 107–120.

Zuo, W., Moses, M.E., West, G.B., Hou, C. & Brown, J.H. 2012. A general model for effects of temperature on ectotherm ontogenetic growth and development. *Proc. R. Soc. London B Biol. Sci.* **279**: 1840–1846.

Tables

Table 1. The LMMs for wild zebra finch mass and tarsus length. Measurements are from fledglings sampled from nest boxes at Fowlers Gap, mean fledgling age was 12.5 days. These LMMs used clutch ID (genetic relatedness) and nest ID (accounting for common environment due to cross fostering) as random factors. Significant effects are in bold. There was a significant negative relationship between mean maximum temperature during development and body mass as we would predict but not for tarsus. For both traits, older chicks were larger as expected. Larger brood size had a negative effect on mass but not tarsus, possible due to food competition affecting condition. Nest ID explained a ten times larger proportion of the variance in mass than clutch ID, 57.2% and 6.1% respectively (ICC = Inter-class Correlation Coefficient). For tarsus length, Nest ID and clutch ID explained a similar proportion of the variance, 20.7% and 20.5% respectively.

Total body mass					
Fixed effects	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.029	0.060	0.490	213.3	
Max temp	-0.189	0.053	**-3.555**	263.4	**<0.001**
Brood size	-0.201	0.053	**-3.802**	212.8	**<0.001**
Chick age	0.250	0.032	**7.761**	853.5	**<0.0001**
	Variance	SD	n		ICC
Marginal R^2	0.134				0.128
Nest ID	0.598	0.773	227		0.572
Clutch ID	0.065	0.254	257		0.062
Residual	0.249	0.499	878	Conditional R^2	0.762
Tarsus length					
Fixed effects	Semi-partial *r*	SE	*t*	df	*P*
Intercept	-0.005	0.053	-0.102	202.3	
Max temp	-0.030	0.050	-0.607	238.5	0.544
Brood size	0.054	0.044	1.246	155.8	0.214
Chick age	0.331	0.040	**8.302**	642.6	**<0.0001**
	Variance	SD	n		ICC
Marginal R^2	0.110				0.100
Nest ID	0.229	0.476	226		0.207
Clutch ID	0.226	0.478	255		0.205
Residual	0.538	0.734	873	Conditional R^2	0.512

Table 2. The LMMs for the wild zebra finch paired data for mass and tarsus. For this data mean fledgling age was 12.5 days. These models only used data from 34 pairs that reared two broods in monitored nest boxes. To account for the balanced design we used independent intercepts and slops by using an interaction between Temperature and the random factor of Pair ID. The significant fixed effects in this model are the same as the full data set (Table 1). Importantly, for mass we found a significant negative relationship with mean maximum temperature which is consistent with the full dataset but the trend is not significant for tarsus.

Total body mass					
Fixed effects	Semi-partial *r*	SE	*t*	df	*P*
Intercept	-0.006	0.095	-0.067	31.89	
Max temp	-0.394	0.088	**-4.481**	25.54	**<0.001**
Brood size	-0.269	0.053	**-5.107**	229.17	**<0.0001**
Chick age	0.259	0.060	**4.295**	174.67	**<0.0001**
	Variance	SD	Correlation	n	ICC
Marginal R^2	0.278				0.254
Pair ID	0.218	0.467		34	0.319
Temp interaction	0.132	0.363	0.28		Conditional R^2
Residual	0.469	0.685		241	0.572
Tarsus length					
Fixed effects	Semi-partial *r*	SE	*t*	df	*P*
Intercept	0.008	0.105	0.073	31.8	
Max temp	-0.096	0.075	-1.270	18.94	0.220
Brood size	0.017	0.053	0.326	225.8	0.745
Chick age	0.414	0.058	**7.129**	125.42	**<0.0001**
	Variance	SD	Correlation	n	ICC
Marginal R^2	0.180				0.173
Pair ID	0.288	0.537		34	0.347
Temp interaction	0.075	0.273	0.67		Conditional R^2
Residual	0.502	0.709		242	0.519

Table 3. The LMMs for the experimental effects of warm ambient temperatures on mass and tarsus at day 28. These LMMs used breeding pair as a random factor with independent slopes by using the interaction with the fixed effect of temperature. In this dataset all pairs successfully bred at both temperature treatments. There was a significant negative relationship between both traits and temperature as we would expect based on the results from the wild data.

Total body mass					
Fixed effects	Semi-partial *r*	SE	*t*	df	*P*
Intercept	-0.039	0.172	-0.225	15.4	
Temperature	-0.272	0.112	**-2.438**	14.2	**0.029**
Brood size	-0.067	0.139	-0.478	28.5	0.636
	Variance	SD	Correlation	n	ICC
Marginal R^2	0.064				0.062
Breeding pair	0.335	0.579		16	0.344
Temp interaction	0.034	0.184	0.4		Conditional R^2
Residual	0.613	0.783		100	0.407

Tarsus length					
Fixed effects	Semi-partial *r*	SE	*t*	df	*P*
Intercept	-0.026	0.145	-0.181	13.6	
Temperature	-0.221	0.103	**-2.152**	65.04	**0.035**
Brood size	-0.240	0.125	-1.914	48.78	0.062
	Variance	SD	Correlation	n	ICC
Marginal R^2	0.067				0.065
Breeding pair	0.192	0.438		16	0.218
Temp interaction	0.016	0.127	-1.00		Conditional R^2
Residual	0.743	0.862		100	0.283

Figures

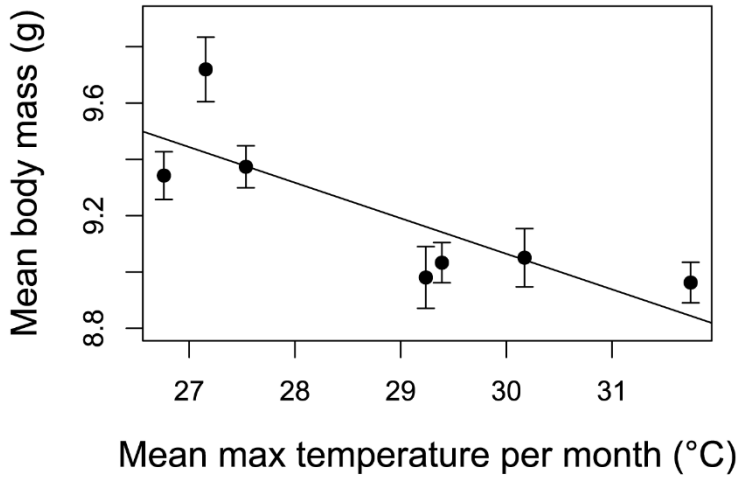


Figure 1. Mean body mass across 7 months of observations from a wild zebra finch colony. We calculated the mean mass and mean maximum temperature for all fledgling measured within a month (n > 70 per month) across the three breeding seasons. When we ran a linear model on these data, using weights to account for different sample sizes, we found a significant negative relationship (estimate = -0.12, t_5 = -3.23, P = 0.02, Adjusted R^2 = 0.61). When we included four extra months with low sample sizes (n < 30) we still find the same relationship (estimate = -0.10, t_9 = -3.42, P = 0.008, Adjusted R^2 = 0.52). The trend line shows the liner relationship.

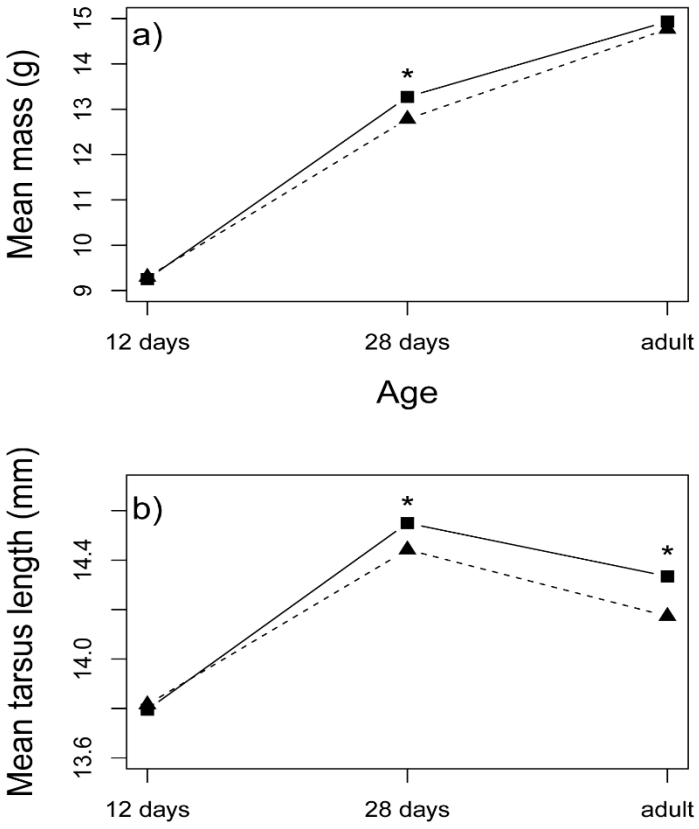


Figure 2. Captive zebra finch growth experiment. Birds were reared at two temperatures treatments. The mean body size for the 18ºC treatment are marked as squares (■) and the 30ºC treatment are marked as triangles (▲). Significant differences are indicated with a star, the effect of temperature on body size was tested using LMM's (see Table 3 and supplementary R code). The offspring were measured at three time points 12-days, 28-days and adult, adult measurements of offspring were taken after birds were removed from experimental conditions and keep in neutral ambient temperatures. At 28-days old both mass and tarsus length were significantly smaller in the high temperature treatment birds. Both are not significant at 12-days and only tarsus is significantly smaller in the high temperature birds when they were measured after the temperature treatment.

Supporting Information

Supplementary data for all primary data

Supplementary R code for analyses

Available at Figshare, link: https://figshare.com/s/dbc245f85f4cc3cdd5b1

Chapter Five

The genetic structure of the introduced house sparrow populations in Australia and New Zealand is consistent with historical descriptions

Chapter five Vignette

We detected some morphological differentiation between house sparrow populations from Australia and New Zealand, the next stage is to explore to what extent this differentiation is mirrored by genetic differentiation. Reproductive isolation can result in genetic differentiation at functional loci that are being acted on by selection for local adaptation. The population structure of invasive species is important information for the general biology of the species and is valuable to the design of genomic studies on local adaptation. Cost effective genetic markers (per sample) such as polymorphic microsatellites and mtDNA sequence are often used to identify population structure and reconstruct invasion histories (Cristescu 2015). These markers can provide similar information about population structure and the colonization history when compared to the more expensive population genomics methods that are growing in popularity (Jeffries et al. 2016). Due to the affordability of population genetics markers they can be useful in pilot studies or for studies with high resolution sampling designs. There are many valuable questions population genetics studies can answer. First, are population boundaries closely linked to physical barriers to dispersal? Alternately, do the borders of climatic regions define the boundaries of genetic populations (Orsini et al. 2013). It is also valuable to know what model of demographic history and dispersal the sampled populations fit? Simulations have shown populations that fit the equilibrium Isolation By Distance model (IBD-eq) or the Island model (IM) have lower false positive rates when identifying outlier loci affected by selection, than populations that fit the model of non-

equilibrium IBD due to expansion from multiple refugia (Lotterhos & Whitlock 2015). It can also be advantageous to estimate the level of inbreeding, effective population size and the size of genetic bottlenecks (Foll & Gaggiotti 2008).

Modern molecular techniques have been applied to native and introduced house sparrow populations around the world. Here we aim to gain a comprehensive understanding of the population structure and demographic history of Australian and New Zealand house sparrow populations in order to design a more focused genomic study of local adaption in an introduced bird species.

This manuscript is under review at Biological Invasions, and is formatted as such.

References:

Cristescu ME (2015) Genetic reconstructions of invasion history. *Molecular Ecology*, **24**, 2212–2225.

Foll M, Gaggiotti O (2008) A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics*, **180**, 977–993.

Jeffries DL, Copp GH, Handley LL et al. (2016) Comparing RADseq and microsatellites to infer complex phylogeographic patterns, a real data informed perspective in the Crucian carp, *Carassius carassius, L*. *Molecular Ecology*, **25**, 2997–3018.

Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, **24**, 1031–1046.

Orsini L, Vanoverbeke J, Swillen I et al. (2013) Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. *Molecular Ecology*, **22**, 5983–99.

Authors: Samuel C. Andrew[1]; Monica Awasthy[1]; Peri E. Bolton[1]; Lee A. Rollins[2]; Shinichi Nakagawa[3,4] and Simon C. Griffith[1]

[1]Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia.

[2]Centre for Integrative Ecology, School of Life & Environmental Sciences, Deakin University, Pigdons Road, Geelong, Vic. 3217, Australia.

[3]Department of Zoology, University of Otago, Dunedin 9054, New Zealand.

[4]Evolution and Ecology Research Centre and School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, NSW 2052, Australia.

Keywords: *Passer domesticus*; Invasive species; colonization history; population structure; founder effect; natural experiment.

Acknowledgements: For funding support: SCA was supported by Macquarie University Research Excellence Scholarships (no. 2013077). SCG was supported by an Australian Research Council Future Fellowship (FT130101253). SN was supported by a Rutherford Discovery Fellowship (New Zealand) and an Australian Research Council Future Fellowship (FT130100268). Fragment analysis was conducted at Macrogen, Inc. We would like to thank the large number of people who helped us with access to urban catch sites for sampling the human commensal house sparrow. We would also like to thank, Amanda D. Griffith, Elizabeth L. Sheldon, Anna Feit, Ondi Crino and Peter Bird for their participation with field work in Australia and Karin Ludwig in New Zealand. We would also like to thank Elizabeth L. Sheldon for their assistance with lab work.

Abstract

The house sparrow is one of the most widely introduced vertebrate species around the world, making it an important species for the study of invasion ecology. Population genetic studies of these invasions provide important insights into colonisation processes and adaptive responses occurring during invasion. Here we use microsatellite data to infer the population structure and invasion history of the introduced house sparrow (*Passer domesticus*) in Australia and New Zealand. Our results identify stronger population structure within Australia in comparison to New Zealand and patterns are consistent with historical records of multiple introduction sites across both countries. Within the five population clusters identified in Australia, we find declines in genetic diversity as we move away from the reported introduction site within each cluster. This pattern is consistent with sequential founder events. Interestingly, an even stronger decline in genetic diversity is seen across Australia as we move away from the Melbourne introduction site; secondary historical reports suggest this site imported a large number of sparrows and was possible the source of a single range expansion across Australia. However, private allele numbers are highest in the north, away from Melbourne, which could be a result of drift increasing the frequency of rare alleles in areas of smaller population size or due to an independent introduction that seeded or augmented the northern population. This study highlights the difficulties of elucidating population dynamics in introduced species with complex introduction histories and suggests that a combination of historical and genetic data can be useful.

Introduction

The recent and rapid colonization history of invasive species can make them valuable systems for the study of colonization events and their ecological impacts (Barrett 2014). Exotic introductions also provide excellent 'natural' experiments to test hypotheses about local adaptation in species that have been introduced to different environmental conditions than are found in their native range (Lee 2002; Sax et al. 2007; Blackburn 2008). Studies of colonization ecology or the extent to which introduced species have become locally adapted benefit from background knowledge such as the time of arrival and the identity of the original source population. To an extent this information can be reconstructed or confirmed using population genetic approaches (Estoup and Guillemaud 2010). Introduced and/or invasive populations are regularly affected by multiple introductions, population bottlenecks and successive founder effects as they colonize new areas (Dlugosch and Parker 2008). These colonization processes leave genetic signatures that can be used to reconstruct invasion histories (Benazzo et al. 2015; Cristescu 2015). A challenge for studies of local adaptation is that these 'genetic signatures' of colonization history and population structure can hamper the use of genetic methods to identify true signals of local adaptation (Excoffier et al. 2009b; Günther and Coop 2013; de Villemereuil et al. 2014; Lotterhos and Whitlock 2014). As a result, population genetic data are valuable in helping us to characterize and limit the confounding effect of these phenomena, before the successful implementation of genomic studies of local adaptation (De Mita et al. 2013; Rellstab et al. 2015; Francois et al. 2016).

Without the benefit of population genetic data, invasion history may be based on assumptions derived from historical records. These are inherently prone to error because their sources may be unreliable or may lack a global understanding of the invasion. Potential complexities that can result in inaccurate accounts include multiple unreported introductions to the same, or different localities. Due to direct and indirect anthropogenic influences, the colonization history of invasive species is likely to be more complex than naturally distributed

species (Miller et al. 2005). In many cases these anthropogenic colonization events have even been organised and intensive, with some of the most notable examples of introductions being from global chapters of Acclimatisation Society in the mid 1800's (Lever 1985; Lever 1992; Lever 2005). At the time when Acclimatisation Societies were active, private introductions were also taking place, that were motivated by similar philosophies, but were probably less well documented (Balmford 1981).

Acclimatisation Societies focused on a number of species but a few, including the house sparrow (*Passer domesticus*), were introduced widely across the world. Currently, the house sparrow is one of the most broadly distributed bird species in the world (Anderson 2006), largely due to human introductions starting in the mid 1800's to North America, South America, Australasia and Africa (Lever 2005). The species' native distribution covers most of the Palaearctic; this distribution was probably established after the species formed a commensal relationship with humans about 10 000 years ago and spread throughout Eurasia, concurrent with the spread of agriculture from the Middle East (Sætre et al. 2012). The house sparrow was introduced to Australia and New Zealand in the 1860's mostly by Acclimatisation Societies (Lever 2005). The species was a very successful colonizer and expanded its distribution to cover almost all the climatic conditions across Australasia (Higgins et al. 2006). Although Acclimatisation Societies kept detailed records of their introductions, many details may have been lost or miscommunicated through time. This is demonstrated by an investigation of primary literature for house sparrow introductions to Australia, which uncovered repeated introductions, new successful introductions and source populations that were previously unrecognised in the scientific literature (Andrew and Griffith 2016). The house sparrow was also introduced to New Zealand from England in the 1860's with five reported introduction points (see Table S1. and Thomson [1922]).

The house sparrow's broad distribution and its close proximity to humans has made this species an excellent and broadly studied species for invasion genetics (Liebl et al. 2015).

However, to date the introduction history and local adaptation in Australasian house sparrow populations has not been studied using modern molecular techniques. Early work on Australian and New Zealand populations using allozymes found that introduction events reduced genetic diversity and increased differentiation in introduced populations compared to native populations in Europe (Parkin and Cole 1985; Cole and Parkin 1986). More recent molecular techniques have been applied to native and introduced house sparrow populations around the world. These studies cover a number of topics including population structure (Schrey et al. 2011; Kekkonen et al. 2011b; Jensen et al. 2013; Schrey et al. 2014) demographic factors (Vangestel et al. 2011; Kekkonen et al. 2011a; Baalsrud et al. 2014) and observing the link between phenotypic and genetic differentiation (Lima et al. 2012). The Australasian house sparrow populations provide a nice opportunity to replicate findings from other introductions and examine the evolution of the species in the context of the Australasian landscape and climate.

Genetic population structure and demographic history are unique for each introduction and should be characterised as part of the basic biology of an introduced species. Here we investigated the introduced house sparrow populations of both Australia and New Zealand using neutral genetic markers (microsatellites). We predict: i) independent introduction events with varied propagule size and origins will have caused genetic differentiation and population structure between introduction sites due to initial differences in allele frequency and genetic diversity; ii) range expansions from the original sites of introduction will have caused declines in genetic diversity with successive founder events (Peter and Slatkin 2015); iii) successive founder events will have resulted in population differentiation that is strongest at the range edge, due to genetic drift; iv) Across the broad geographic sampling range in Australia we expect to see a pattern like Isolation by Distance (IBD) due to the recent colonization and the species low natural dispersal ability across the large distances of uninhabitable habitat between isolated human settlements (in this highly commensal species).

We discuss the relevance of our findings to historical records of the introduction of these house sparrow populations and related systems.

Materials and Methods

Sampling

Adult house sparrows were collected from 25 urban localities across Australia with approximately 40 birds sampled at each locality (total number of birds genotyped = 1027, Table 1, Fig. 1). The Australian samples were collected during April to September 2014 and March 2015 under the Animal Research Authority of the Animal Ethics Committee at Macquarie University (ARA 2014/248). Samples from New Zealand were collected from four localities between June and August 2005, under the ethics approval of Otago University (Animal ethics reference number: 87/08). Approximately 40 individuals were genotyped from each of the four localities in New Zealand (n = 170, Table 1). Blood samples from a population in Morthen, South Yorkshire, UK were also sourced from a previous study to be included as a comparison for an English source population (n = 40) (Ockendon et al. 2009). Birds were captured using mist nets and placed in bird bags until a blood sample could be taken. Blood was taken from the brachial vein with a capillary tube (ca. 40 µl) and was stored in 800 µl of absolute ethanol in a 1.5ml microcentrifuge tube. Birds were not held for more than 30 minutes and were released as soon as they had been sampled and banded. In total, 1237 birds were genotyped from 30 localities (Table 1, Fig. 1).

Molecular methods

DNA was extracted using a Gentra PureGene tissue kit (Qiagen, Valencia, CA, USA) following manufacturer's instructions. Samples were genotyped using two multiplexes developed by Dawson et al. (2012) which included 13 polymorphic loci and a sexing locus

(Multiplex 1: Ase18, Pdoμ1, Pdoμ3, Pdoμ5, Pdoμ6, Pdo9, Pdo10, P2D/P8; Multiplex 2: Pdo16A, Pdo17, Pdo19, Pdo22, Pdo27, Pdo40A). PCRs were carried out using 5 µl reactions. For each reaction 1 µl of genomic DNA (ca. 100 ng/µl) was added to 2.5 µl of Master Mix (Qiagen, Valencia, CA, USA), 0.5 µl of primer mix (see concentrations in Dawson et al. [2012]) and 1 µl of Milli-Q water. Both multiplex reactions used the same PCR thermal cycle with a hot-start denaturing phase of 10 min at 95 °C followed by 33 cycles of 94 °C for 30 s, 57 °C for 90 s and 72 °C for 90 s, before a final extension at 72 °C for 10 min. The post PCR product was diluted and genotyped on the ABI 3730XL DNA analyzer using GS500 (LIZ) as a size standard for Multiplex 1 and GS1200 for Multiplex 2. Microsatellite alleles were scored using the GeneMapper program version 3.7 (data included with online supplementary material on Figshare link: https://figshare.com/s/dbc245f85f4cc3cdd5b1).

Genetic analyses

Missing data percentages were calculated in Microsoft Excel, and loci with more that 5% missing data were excluded. Deviations from Hardy-Weinberg Equilibrium (HWE) were tested in ARLEQUIN version 3.5.2.2 (Excoffier et al. 2005). Null allele frequency estimates were calculated using CERVUS (Marshall et al. 1997) and loci with null allele frequencies greater than 10% were excluded. GENEPOP (Raymond and Rousset 1993) was used to test linkage disequilibrium (LD) between loci within sampling localities. Allelic richness, number of alleles and genetic diversity (expected heterozygosity) were calculated using FSTAT version 2.9 (Goudet 1995). ARLEQUIN was used to produce a pairwise F_{ST} matrix and to run Analysis of Molecular Variance (AMOVA) to estimate among population differentiation (F_{ST}), for all sample sites and for pooled Australian and pooled New Zealand samples separately. Using the *mmod* package in R, Hedrick's G'st was also calculated. GENALEX version 6.503 (Peakall and Smouse 2006; Peakall and Smouse 2012) was used to calculate private allele frequencies. Using R version 3.3.1 (R Core Development Team, 2017), allele

frequency heat maps were drawn for each locus to show differences between population clusters, as a graphical aid to describe population genetic diversity.

STRUCTURE (Pritchard et al. 2000) was used to identify populations using a model-based Bayesian clustering method that calculates the probability an individual belongs to a cluster when a given number of clusters (K) is specified. To infer the number of clusters with the most support in each analysis we used the delta K method (Evanno et al. 2005). Delta K and mean LnP(K) and other summary statistics were calculated in STRUCTURE HARVESTER (Earl and VonHoldt 2012). The delta K method will identify the upper most level of structure, we also looked at hierarchical population structure within the genetic groups defined by our initial models (Rollins et al. 2009). For our final STRUCTURE analyses we used a MCMC length of 1 000 000 iterations and a burn-in period of 100 000 and 10 repeats for each value of K using a admixture model. The range of K values used in our final runs were chosen based on shorter preliminary runs with fewer iterations (200 000 - 400 000) that included values of K equal to our number of sample localities. We confirmed the peak in delta K was a true signal by checking that the variance in mean LnP(K) was stable between values of K because some repeats were not converging (see Table S2). We determined there was greatest support for one cluster (K=1) when mean LnP(K) was highest for K = 1. Q-plots for the most highly supported value of K were drawn using the results of all 10 repeats with the programs CLUMPP (Jakobsson and Rosenberg 2007) and DISTRUCT (Rosenberg 2004).

The STRUCTURE method relies on demographic assumptions about the study populations that are rarely met in the real world (e.g. no pattern of IBD). For this reason, we also used a second method to look at population structure that does not make assumptions about demographic models. The R package *adegenet* (Jombart 2008) was used for a Correspondence Analysis (CoA) of microsatellite data for all 30 localities; this multivariate approach uses a summary of sample site allele frequencies to create a distance matrix that is

used to generate Principal Component (PC) values for each locality, similar to a Principal Coordinate Analysis (PCoA) of individuals. Our CoA used 5 PC's because we only had 30 sample localities (PC's must be less than *n*), and were enough to explain almost all the variance in the data. To visualise patterns of population structure which can be compared with the results from the STRUCTURE clustering analyses, we used the PC values from the CoA in a Discriminant Analysis of Principal Components (DAPC) in *adegenet* (Jombart et al. 2009; Jombart et al. 2010). The number of genetic clusters was inferred using the "find.clusters" function and the optimal number of clusters was decided based on BIC reduction. The cluster labels were then used in the first DAPC of the PCs from the CoA to make a scatter plot. Using the cluster labels defined by the "find.clusters" function, we ran a second DAPC using the individual data to calculate the percentage of individuals that were correctly assigned to their population clusters defined by the first DAPC of localities. We choose to use the PC from a CoA of sample site (that uses allele frequencies, also accounting for the presence and absence of alleles) to define the main genetic clusters, because we predict founder effects will have had the clearest effect on allele frequency and allelic diversity between localities.

Neighbour-joining trees are often used to reconstruct invasion histories by using the branching pattern to describe invasion routes (Estoup and Guillemaud 2010). We tested if there were any clear invasion routs across Australia by drawing a neighbour-joining tree based on the genetic distance metric of Cavalli-Sforza & Edwards (1967) with the program POPULATIONS (Langella 2002). Bootstrap values for the tree were calculated over 1000 iterations that used different subsets of loci.

To test for Isolation by Distance (IBD) we used Mantel tests, in R using the package *adegenet* (Jombart 2008). We used the function "mantel.randtest" which used 999 replications for the tests. The Mantel tests used Edwards' Euclidean distance (2nd method option) to calculate genetic distance. We ran this analysis for all Australian sampling sites and for the

larger population clusters independently to look for differences in connectivity at different special scales. We also modelled the effects of sequential founder effects on genetic differentiation (pairwise F_{ST}) within population clusters. For this question, the response variable was the F_{ST} between each sample site and the sample site nearest the original historic introduction site (this removed introduction sites from the model). The predictor was "year established" (see methods below) and we expected F_{ST} to be higher for more recently established populations.

To infer the relationship between colonization history and genetic diversity we used Linear Mixed Models (LMM). The LMMs were run in R using the package *lme4* (Bates et al. 2015) with the package *lmerTest* (Kuznetsova et al. 2016) to calculate degrees of freedom and P values. To model the effects of range expansion and founder events on genetic diversity we collected data on the "year established" (the year populations were first recorded as being present in a locality, as previously described in Andrew & Griffith [2016]) and the "distance from source" (distance from the proposed original introduction site location (km)) to be used as fixed effects. The two fixed effects of "year established" and "distance from source" were found to be highly correlated (estimate = 0.072, $t_{27} = 7.664$, $P < 0.0001$, $R^2 = 0.685$), so could not be used together in the same model. We chose to use year established as a proxy for range expansion and sequential founder events in the LMMs. Demographic and genetic non-independence was accounted for by using the "population" clusters from the DAPC analysis as a random factor. The response variables used to measure genetic diversity were allelic richness and expected heterozygosity; these response variables were used in separate models with the same structure as described above. We calculated marginal R^2 and Intra-class Correlation Coefficients (ICC) values for all LMM's using the method described in Nakagawa & Schielzeth (2013). To visualise the patterns in the data we used scatter plots for all combinations of the two response variables vs the two predictor variables, and a third

predictor variable “Sample site distance from Melbourne” was also plotted. Lines of best fit were plotted using a linear regression between the two variables.

Results

We used 11 polymorphic loci (Ase18, Pdoμ1, Pdoμ3, Pdoμ6, Pdo10, Pdo16A, Pdo17, Pdo19, Pdo22, Pdo27, Pdo40A) for analyses after removing Pdo9 due to more than 5 % missing data. Pdoμ5 was also removed for being out of HWE in more than 20 % of localities, which may be owing to a high null allele frequency (15.1%, Table S3). The remaining loci were not out of HWE in more than 5 % of the sample sites. No loci were found to be consistently affected by LD within sampling localities. In total, we genotyped 1237 individuals from 30 sample sites, summary genetic diversity statistics for each sampling locality are presented in Table 1.

Tests of genetic differentiation using an AMOVA found significant differentiation among sampling localities. For all 30 localities, among population F_{ST} = 5.60% (df = 29, P < 0.001); for only Australian samples, among populations F_{ST} = 6.01% (df = 24, P < 0.001); and for only New Zealand samples, among population F_{ST} = 1.90% (df = 3, P < 0.001, see Table S4 for all AMOVA results). Pairwise F_{ST} comparisons also found strong evidence for genetic differentiation between sampling localities, with significant differentiation in over 95% of pair-wise comparisons (420 out of 435) after Bonferroni corrections (Fig. S1). The Global Hedrick’s G’st for all loci was also calculated for all sample sites (G’st = 0.265), Australia (G’st = 0.274) and New Zealand (G’st = 0.109), to compare to amoung population F_{ST}.

Population structure was visualised using a DAPC approach (Fig. 2). A multivariate analyses (CoA) of population structure that used a distance matrix for the 30 sample sites based on allele frequencies, was used in a find clusters analysis as part of the DAPC to identify the number of meaningful clusters in the data that provided the lowest BIC. We found BIC levelled off and stopped improving at eight clusters (Fig. 2a, 2b). The eight cluster ID’s

for the sample sites were then applied to the individual genotype data and used to run a second DAPC. We found that greater than 80% of individuals (range 80 – 96%) were correctly assigned to their population ID's from the first DAPC, giving support for these population clusters (Fig. 2c). The names for the eight clusters in Fig. 2 are based on the capital city within a polygon drawn around the clusters sample sites or the geographic region. Apart from these eight clusters the DAPC also visualises a very clear divide between the "northern Australia" localities and all the other localities (including southern Australia, New Zealand and England).

The results from the model-based clustering analysis in STRUCTURE (Fig. S2) were similar to the results from the DAPC approach. After accounting for substructure we found 10 clusters using STRUCTURE which are summarised in Fig. 2b. The only inconsistencies between the STRUCTURE and the DAPC method was that the northern Australia localities are broken into four rather than two clusters respectively and STRUCTURE grouped Cobar with the Melbourne cluster rather than with the Sydney cluster (Fig. 2b). The scatter plot in Fig. 2a, however, shows that the eight northern Australia localities (hence referred to as Brisbane cluster) are closely grouped, indicating weak structure within that region; The Melbourne and Sydney clusters were also relatively close together. A neighbour-joining tree of Australian localities found localities from the same cluster were grouped within clades but the branching pattern of the tree did not have any significant support (most boot strap values < 70%) for possible invasion roots across Australia (Fig. 3).

Population structure established by founder effects could have been eroded since the original colonization events due to gene flow. If gene flow is low then regions that had independent shipments of sparrows into Australia could have maintained private alleles. We visualise allele presence absence between genetic populations using heat maps (Fig. S3). Predictably, New Zealand and England had alleles that were not observed in Australia for most loci. Between the five Australian population clusters, the total number of private alleles

(including those with minor frequencies) were: Melbourne = 9, Brisbane = 8, Sydney = 4, Adelaide = 4, Hobart = 1. The number of private alleles with a frequency > 1% (to avoid falsely identifying rare alleles that are easily missed without comprehensive sampling) are much lower, the Brisbane cluster had three, the Melbourne cluster had two and the Hobart cluster had one private allele.

To further look at patterns of dispersal, we tested for IBD. The Mantel test found strong support for IBD across sample sites within Australia ($R^2 = 0.758$, n = 25, P = 0.001). This pattern is consistent within the two main population clusters of Brisbane (northern Australia, $R^2 = 0.545$, n = 8, P = 0.024), southern Australia ($R^2 = 0.767$, n = 17, P = 0.001) and for the Melbourne cluster within southern Australia ($R^2 = 0.520$, n = 6, P = 0.020) when analysed separately (see Fig. S4 for details). Other sub-clusters had a smaller number of sample localities so were not also analysed separately. However, the relationships were less strong within the Brisbane and Melbourne clusters (Fig. S4).

In invasions, sequential founder events can reduce genetic diversity and increase F_{ST}. We predict that sequential founder effects have created population differentiation, where pairwise F_{ST} between sample sites (and the putative source) should be positively correlated with the "year established". Using linear models, we found a significant positive relationship between the "year established" and the "pairwise F_{ST} comparison with the sample site nearest the putative site of introduction" (estimate = 0.0003, $t_{23} = 4.651$, $P < 0.001$, $R^2 = 0.485$).

The LMMs for the effects of range expansion (year established) on genetic diversity also found allelic richness and expected heterozygosity declined in more recently established populations (Table 2). We also found that the random factor of "population" explained a large proportion of the variance in the data (Table 2), where lower intercepts corresponded to lower levels of genetic diversity (Fig. S5). We have used "year established" as a proxy measure of sequential founder events, but distance from the putative introduction site is also correlated with this variable and genetic diversity (Fig. 4). The distance from the putative introduction

site was negatively correlated with genetic diversity (Fig. 4b, 4e). The negative correlation with genetic diversity is even stronger when we calculate the distance of each site from Melbourne (Fig. 4c, 4f). However, Melbourne is in the south of Australia and genetic diversity is higher in the south versus the north, allelic richness ($t_{23} = 4.506$, $P < 0.001$, $R^2 = 0.469$) and expected hetrozygosity ($t_{23} = 7.007$, $P < 0.0001$, $R^2 = 0.681$) are both positively correlated with latitude. Therefore, this relationship for Melbourne could be true for any southern locality. A summary for the linear models for the lines of best fit in these scatter plots in Fig. 4 are given in Table S5.

Discussion

Across our Australian and New Zealand sampling locations and the founding population (England), we identified significant population structuring with eight population clusters identified by the DAPC method (Fig. 2). The main population structure in Australia was found between northern Australia (Brisbane cluster) and the rest of the sampling localities (southern Australia). The two population 'sub-clusters' within the Brisbane cluster were very genetically similar and this separation is likely due to sequential colonization events and isolation. Therefore, we refer to the eight northern localities as the Brisbane cluster. The English locality clustered with the South Island of New Zealand although they are clearly independent reproductive populations but have similar genetic compositions. This population structure was also supported using a Bayesian clustering analysis (Fig. S2) which suggests these results are repeatable and relatively robust.

We find evidence of IBD across our Australian sampling localities and this violates one of the assumptions of the widely used STRUCTURE analyses that could potentially influence our results (Frantz et al. 2009; Jombart et al. 2009). Therefore, we have focused on a DAPC method to identify population structure. We suggest the DAPC method will be useful in future studies of population structure in invasive species with complex introduction histories

because the method does not rely on demographic assumptions (Jombart et al. 2009; Jombart et al. 2010). This method also has flexibility in its application because the genetic data can be used to define population clusters or to test the accuracy of predefined clusters. Predefined populations could be based on sampling design, the species biology or results from related analyses, as was done here in the second DAPC.

The high level of genetic differentiation found among populations (AMOVA and pairwise F_{ST}) across both Australia and New Zealand is expected for a relatively sedentary species such as the house sparrow that has gone through sequential colonization events. Levels of population differentiation (pairwise F_{ST}) reported for house sparrow populations around the world report some similar patterns (Schrey et al. 2011; Lima et al. 2012; Jensen et al. 2013). However, genetic differentiation was found to be very low (F_{ST} among Finnish populations = 0.004 ± 0.001 s.e.) across a broad distribution of native house sparrow populations in Finland (Kekkonen et al. 2011b). Although house sparrows are generally sedentary, individuals have been shown to naturally disperse up to 50 km in an island archipelago off of Norway (Tufto et al. 2005). The sparrow, can also disperse long distances across uninhabitable landscape to reach new human settlements by hitchhiking on human modes of transport such as trains, trucks and boats (Long 1988). This mode of range expansion has been proposed in introduced sparrow populations in Africa (Schrey et al. 2014) as well as in other invasions that show evidence of long distance unintentional anthropogenic introductions (Miller et al. 2005; Pascual et al. 2007; Preuss et al. 2015). Long distance dispersal events would have been necessary for the colonization of remote Australian towns. The isolation between rural towns in Australia has led to independent populations and high population differentiation that is characteristic of a meta-population (pairwise F_{ST}, Fig. S1). In contrast, the highly populated areas sampled around Melbourne show lower levels of differentiation (Fig. S1) and the neighbouring populations of Albury and Burrumbuttock show no significant differentiation at this much smaller spatial scale (ca. 30km apart, F_{ST} = 0.004).

The pairs of sample sites within the north and south islands of New Zealand also show low differentiation (F_{ST} = 0.006 and 0.001 respectively, both n.s.).

We find across all our 25 Australian sampled localities a clear pattern of IBD (Fig. S4a). However, within smaller regions linked to the same population cluster the trend is less strong, possibly because there is more connectivity and gene flow (Fig. S4b and S4d). Another explanation of this result is Simpson's Paradox which describes instances where if data is divided into known categories and analysed separately the original result using the full dataset is no longer supported (Wagner 1982). When we look at IBD across all of Australia the two main categories are comparisons within and between population clusters. There is strong population structure between northern and southern Australia so these pairwise comparisons have high F_{ST} and are also far apart but this genetic differentiation is not necessarily explained by distance alone but also potentially by independent introduction events and low subsequent gene flow between the two populations. Northern Australian sample sites also have lower genetic diversity than southern sites and differences in genetic diversity will affect estimates of genetic differentiation (F_{ST}) with lower diversity increasing F_{ST} (Jakobsson et al. 2013). These differences in genetic diversity can be explained by differences in founder population size as well as genetic drift. There is also the possibility we see genetic diversity decline towards the range edges after a single range expansion from a single introduction point. In Australia, the single point could be Melbourne, which has the highest genetic diversity; we do see a decline in genetic diversity as we move away from the Melbourne area in the South of Australia (Fig. 4). Although the neighbour-joining tree found no significant support for an invasion originating from Melbourne and expanding across Australia (Fig. 3). We also see an overall decline in genetic diversity going from the south to the north of Australia (Table S5). In this study, it is most likely that the latitudinal pattern in genetic diversity is best explained by introduction history.

Acclimatisation Societies have documented five main introduction sites in Australia and they were all found to be linked to distinct genetic populations (Fig. 2). These five sites were the capital cities that supported Acclimatisation Societies: Melbourne, Sydney, Hobart, Adelaide and Brisbane. The localities in the Melbourne population show the highest levels of genetic diversity (Table 1 and Fig. S5), which is consistent with the historical records that report the largest numbers of birds being imported into Melbourne (Andrew and Griffith 2016). The Sydney population is reported as being founded by individuals imported in the 1860's from the newly established population in Melbourne. This event would explain why these populations are close to each other in Fig. 2a, but population structure due to founder effects has been maintained possible due to limited gene flow between the two regions. Similarly, if the Hobart population on the island of Tasmania was originally founded by birds sent from Melbourne in 1867 (Andrew and Griffith 2016) then the subsequent genetic drift due to founder effects could have created genetic differentiation with the mainland. The sea barrier between the two islands of New Zealand has also maintained differentiation between the two island populations, that were likely established by separate introductions (see Thomson [1922] and Table S1). Interestingly the English sample has grouped with the South Island of New Zealand which is also genetically similar to the North Island. The New Zealand populations could be more similar to England than Australian populations because sparrows were imported from England and India to Melbourne with potential successful introgression between sparrows from these two sub-species (*P. d. domesticus* and *P. d. indicus,* [Andrew and Griffith 2016]).

Independent shipments from England were also reported to have successfully introduced sparrows to Brisbane and Adelaide. The sparrows contributing to the primary introduction into Brisbane may include those arriving by ship in 1869 and 12 sparrows sent from Melbourne in 1868 (Andrew and Griffith 2016). There are also clear primary reports that house sparrow populations were established in Brisbane as well as Adelaide in the 1870's,

well before sparrows spread naturally out of Victoria without human intervention (Andrew and Griffith 2016). The DAPC (Fig. 2) shows a clear gap between all the sample sites in the Brisbane cluster and the rest of Australia. The Brisbane population also had the most private alleles (three) with a frequency greater than 1%. This is surprising since the Brisbane population also had the lowest number of alleles. These results suggest that the Brisbane population is more than just a subset of the genetic diversity found in the south. However, sequential founder events and genetic drift can also make rare alleles much more common so this population structure could be explained by demographic bottlenecks during colonization (Excoffier et al. 2009a). The Melbourne and Adelaide population clusters are not as differentiated, indicating more connectivity within the southern half of the species distribution in Australia, if there was a successful independent introduction to Adelaide in the 1860's. It is also possible that both the Adelaide and Melbourne introductions had similar source populations.

The complexity of invasion histories and the demographic bottlenecks experienced by invasive species can provide unique challenges for studies of local adaptation. In parallel with genetic changes due to local adaptation, there are a number of neutral processes driving population subdivision such as sequential colonization events (Peter and Slatkin 2015), small population size/inbreeding (Keller and Waller 2002) and admixture (Orsini et al. 2013). The recent development of genome scan methods and the identification of confounding genetic drift add new value to the study of genetic population structure in invasive species (Excoffier et al. 2009b; Günther and Coop 2013; de Villemereuil et al. 2014; Lotterhos and Whitlock 2014). Information on population structure and demographic history that is gathered through high resolution genetic sampling, can improve the sampling design of genome scan projects and the informative power of results (De Mita et al. 2013; Rellstab et al. 2015; Francois et al. 2016). A more general benefit of describing population structure in invasive populations is to propose general patterns of genetic differentiation in invasive populations. these observations

will help us to draw conclusions about the origin and dynamics of biological invasions that have limited historical information.

The large number of reported introduction events of the house sparrows in Australia and New Zealand is probably not unusual for species that were introduced by Acclimatisation Societies to many locations around the world in the mid 1800's (Long 1981; Lever 1992; Lever 2005). The house sparrow has also been used as a model species to study invasion genetics globally (Liebl et al. 2015). Here we have described the population structure of the house sparrow within the last major region of the world that had not been previously subjected to study by genetic markers. The population structure that we have characterised within Australia and New Zealand is consistent with our expectations for this relatively sedentary species that has gone through reasonably well documented sequential colonization events. Furthermore, the relatively strong structuring that we have characterised suggests that there is reasonable scope for local adaptation to have occurred, even in the relatively short period of time since the introduction, just over 150 years ago. The populations of introduced house sparrows around the world remain a good target for further studies of evolution and ecology (Liebl et al. 2015), particularly given the genomic resources that are coming online for this species (Hagen et al. 2013). The population structure that we have described provides a very useful foundation for further work in this area.

References

Anderson TR (2006) Biology of the Ubiquitous House Sparrow: from Genes to Populations. Univ. Press New York.

Andrew SC, Griffith SC (2016) Inaccuracies in the history of a well-known introduction: a case study of the Australian House Sparrow (*Passer domesticus*). Avian Res 7:9. doi: 10.1186/s40657-016-0044-3

Baalsrud HT, Saether B-E, Hagen IJ, et al (2014) Effects of population characteristics and structure on estimates of effective population size in a house sparrow metapopulation. Mol Ecol 23:2653–2668. doi: 10.1111/mec.12770

Balmford R (1981) Early introductions of birds to Victoria. Vic Nat 98:96–105.

Barrett SCH (2014) Foundations of invasion genetics: the Baker and Stebbins legacy. Mol Ecol 24:1927–1941. doi: 10.1111/mec.13014

Bates D, Maechler M, Bolker B, Walker S (2015) lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-8, http://CRAN.R-project.org/package=lme4.

Benazzo A, Ghirotto S, Vilaça ST, Hoban S (2015) Using ABC and microsatellite data to detect multiple introductions of invasive species from a single source. Heredity (Edinb) 115:262–272. doi: 10.1038/hdy.2015.38

Blackburn TM (2008) Using aliens to explore how our planet works. Proc Natl Acad Sci 105:9–10.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic Analysis Models and Estimation Procedures. Am J Hum Genet 19:233–57. doi: 10.2307/2406616

Cole SR, Parkin DT (1986) Adenosine-deaminase polymorphism in House Sparrow, *Passer-domesticus*. Anim Genet 17:77–88.

Cristescu ME (2015) Genetic reconstructions of invasion history. Mol Ecol 24:2212–2225. doi: 10.1111/mec.13117

Dawson DA, Horsburgh GJ, Krupa AP, et al (2012) Microsatellite resources for Passeridae species: a predicted microsatellite map of the house sparrow *Passer domesticus*. Mol Ecol Resour 12:501–23. doi: 10.1111/j.1755-0998.2012.03115.x

De Mita S, Thuillet A-C, Gay L, et al (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. Mol Ecol 22:1383–1399. doi: 10.1111/mec.12182

de Villemereuil P, Frichot É, Bazin É, et al (2014) Genome scan methods against more complex models: when and how much should we trust them? Mol Ecol 23:2006–2019. doi: 10.1111/mec.12705

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. Mol Ecol 17:431–49. doi: 10.1111/j.1365-294X.2007.03538.x

Earl DA, VonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. doi: 10.1007/s12686-011-9548-7

Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: Why, how and so what? Mol Ecol 19:4113–4130. doi: 10.1111/j.1365-294X.2010.04773.x

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Excoffier L, Foll M, Petit RJ (2009a) Genetic Consequences of Range Expansions. Annu Rev Ecol Evol Syst 40:481–501. doi: 10.1146/annurev.ecolsys.39.110707.173414

Excoffier L, Hofer T, Foll M (2009b) Detecting loci under selection in a hierarchically structured population. Heredity (Edinb) 103:285–298. doi: 10.1038/hdy.2009.74

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47.

Francois O, Martins H, Caye K, Schoville S (2016) A tutorial on Controlling False Discoveries in Genome Scans for Selection. Mol Ecol 25:454–469. doi: 10.1111/mec.13513

Frantz AC, Cellina S, Krier A, et al (2009) Using spatial Bayesian methods to determine the genetic structure of a continuously distributed population: Clusters or isolation by distance? J Appl Ecol 46:493–505. doi: 10.1111/j.1365-2664.2008.01606.x

Goudet J (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. J Hered 86:485–486. doi: 10.1093/jhered/esh074

Günther T, Coop G (2013) Robust identification of local adaptation from allele frequencies. Genetics 195:205–20. doi: 10.1534/genetics.113.152462

Hagen IJ, Billing AM, Rønning B, et al (2013) The easy road to genome-wide medium density SNP screening in a non-model species: development and application of a 10 K SNP-chip for the house sparrow (*Passer domesticus*). Mol Ecol Resour 13:429–39. doi: 10.1111/1755-0998.12088

Higgins PJ, Peter JM, Cowling SJ (2006) Handbook of Australian, New Zealand and Antarctic birds: Volume 7 Boatbill to Starlings. Oxford University Press, Melbourne, pp 1102–26

Jakobsson M, Edge MD, Rosenberg NA (2013) The relationship between FST and the frequency of the most frequent allele. Genetics 193:515–528. doi: 10.1534/genetics.112.144758

Jakobsson M, Rosenberg NA (2007) CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806. doi: 10.1093/bioinformatics/btm233

Jensen H, Moe R, Hagen IJ, et al (2013) Genetic variation and structure of house sparrow populations: is there an island effect? Mol Ecol 22:1792–805. doi: 10.1111/mec.12226

Jombart T (2008) Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. doi: 10.1093/bioinformatics/btn129

Jombart T, Devillard S, Balloux F, et al (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:94. doi: 10.1186/1471-2156-11-94

Jombart T, Pontier D, Dufour A-B (2009) Genetic markers in the playground of multivariate analysis. Heredity (Edinb) 102:330–41. doi: 10.1038/hdy.2008.130

Kekkonen J, Hanski IK, Jensen H, et al (2011a) Increased genetic differentiation in house sparrows after a strong population decline: From panmixia towards structure in a common bird. Biol Conserv 144:2931–2940. doi: 10.1016/j.biocon.2011.08.012

Kekkonen J, Seppä P, Hanski IK, et al (2011b) Low genetic differentiation in a sedentary bird: house sparrow population genetics in a contiguous landscape. Heredity (Edinb) 106:183–90. doi: 10.1038/hdy.2010.32

Keller LF, Waller DM (2002) Inbreeding effects in wild populations. Trends Ecol Evol 17:230–241.

Kuznetsova A, Brockhoff PB, Christensen RHB (2016) lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-32. https://CRAN.R-project.org/package=lmerTest.

Langella O (2002) POPULATIONS. http://bioinformatics.org/project/ ?group_id=84. Bioinformatics

Lee CEE (2002) Evolutionary genetics of invasive species. Trends Ecol Evol 17:386–391. doi: 10.1016/S0169-5347(02)02554-5

Lever C (1985) Naturalized Mammals of the World. Longman, London

Lever C (1992) They Dined on Eland. Quiller Press Ltd, London

Lever C (2005) Naturalised Birds of the world. T & A D Poyser, London, pp 204–218

Liebl AL, Schrey AW, Andrew SC, et al (2015) Invasion genetics: Lessons from a ubiquitous bird, the house sparrow *Passer domesticus*. Curr Zool 61:465–476. doi: http://dx.doi.org/10.1093/czoolo/61.3.465

Lima MR, Macedo RHF, Martins TLF, et al (2012) Genetic and Morphometric Divergence of an Invasive Bird: The Introduced House Sparrow (*Passer domesticus*) in Brazil. PLoS One. doi: e53332 10.1371/journal.pone.0053332

Long JL (1988) Introduced birds and mammals in Western Australia. Agric. Prot. Board West. Aust. Agriculture Protection Board, Perth, pp 13–14

Long JL (1981) Introduced Birds of the World: The Worldwide History, Distribution and Influence of Birds Introduced to New Environments. Universe Books, New South Wales, pp 372–381

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. Mol Ecol 23:2178–2192. doi: 10.1111/mec.12725

Marshall TC, Slate J, Kruuk LEB and Pemberton JM (1997) Statistical confidence for likelihood-based paternity. Mol Ecol 7:639–655.

Miller N, Estoup A, Toepfer S, et al (2005) Multiple Transatlantic Introductions of the Western Corn Rootworm. Science 310:992–992. doi: 10.1126/science.1115871

Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods Ecol Evol 4:133–142. doi: 10.1111/j.2041-210x.2012.00261.x

Ockendon N, Griffith SC, Burke T (2009) Extrapair paternity in an insular population of house sparrows after the experimental introduction of individuals from the mainland. Behav Ecol 20:305–312. doi: 10.1093/beheco/arp006

Orsini L, Vanoverbeke J, Swillen I, et al (2013) Drivers of population genetic differentiation in the wild: isolation by dispersal limitation, isolation by adaptation and isolation by colonization. Mol Ecol 22:5983–99. doi: 10.1111/mec.12561

Parkin DT, Cole SR (1985) Genetic differentiation and rates of evolution in some introduced populations of the House Sparrow, *Passer domesticus* in Australia and New Zealand. Heredity (Edinb) 54:15–23. doi: 10.1038/hdy.1985.4

Pascual M, Chapuis MP, Mestres F, et al (2007) Introduction history of *Drosophila subobscura* in the New World: A microsatellite-based survey using ABC methods. Mol Ecol 16:3069–3083. doi: 10.1111/j.1365-294X.2007.03336.x

Peakall R, Smouse PE (2006) GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288–295. doi: 10.1111/j.1471-8286.2005.01155.x

Peakall R, Smouse PE (2012) GenALEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537–2539. doi: 10.1093/bioinformatics/bts460

Peter BM, Slatkin M (2015) The effective founder effect in a spatially expanding population. Evolution 69:721–734. doi: 10.1111/evo.12609.This

Preuss S, Berggren Å, Cassel-Lundhagen A (2015) Genetic patterns reveal an old introduction event and dispersal limitations despite rapid distribution expansion. Biol Invasions 17:2851–2862. doi: 10.1007/s10530-015-0915-2

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–59.

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.

Raymond M, Rousset F (1993) GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. J Hered 86:248–249.

Rellstab C, Gugerli F, Eckert AJ, et al (2015) A practical guide to environmental association analysis in landscape genomics. Mol Ecol 24:4348–4370. doi: 10.1111/mec.13322

Rollins LA, Woolnough AP, Wilton AN, et al (2009) Invasive species can't cover their tracks: using microsatellites to assist management of starling (*Sturnus vulgaris*) populations in Western Australia. Mol Ecol 18:1560–73. doi: 10.1111/j.1365-294X.2009.04132.x

Rosenberg NA (2004) Distruct: a Program for the Graphical Display of Population Structure. Mol Ecol Notes 4:137–138. doi: 10.1046/j.1471-8286.2003.00566.x

Sætre G-P, Riyahi S, Aliabadian M, et al (2012) Single origin of human commensalism in the house sparrow. J Evol Biol 25:788–96. doi: 10.1111/j.1420-9101.2012.02470.x

Sax DF, Stachowicz JJ, Brown JH, et al (2007) Ecological and evolutionary insights from species invasions. Trends Ecol Evol 22:465–71. doi: 10.1016/j.tree.2007.06.009

Schrey AW, Grispo M, Awad M, et al (2011) Broad-scale latitudinal patterns of genetic diversity among native European and introduced house sparrow (*Passer domesticus*) populations. Mol Ecol 20:1133–1143. doi: 10.1111/j.1365-294X.2011.05001.x

Schrey AW, Liebl AL, Richards CL, Martin LB (2014) Range expansion of house sparrows (*Passer domesticus*) in Kenya: evidence of genetic admixture and human-mediated dispersal. J Hered 105:60–9. doi: 10.1093/jhered/est085

Thomson GM (1922) The Naturalisation of Animals and Plants in New Zealand. Cambridge: University Press

Tufto J, Ringsby T, Dhondt AA, et al (2005) A Parametric Model for Estimation of Dispersal Patterns Applied to Five Passerine Spatially Structured Populations. Am Nat 165:E13--E26.

Vangestel C, Mergeay J, Dawson D a, et al (2011) Spatial heterogeneity in genetic relatedness among house sparrows along an urban-rural gradient as revealed by individual-based analysis. Mol Ecol 20:4643–53. doi: 10.1111/j.1365-294X.2011.05316.x

Wagner CH (1982) Simpson's paradox in real life. Amercian Stat 36:46–48.

Tables

Table 1. Population information including: Locality ID number (No.); Location name; State within country; Name of population cluster (Population cluster); sample size (*n*); The year the population was reported as established (year est.); Distance from the proposed introduction site (Dist. (km)); Mean allelic richness across loci (mean Ar); mean observed hetrozygosity (mean H_O); mean expected hetrozygosity (mean H_E).

No.	Location	State	Population cluster	*n*	Lat	Long	year est.	Dist. (km)	mean Ar	mean H_O	mean H_E
1	Tennant Ck	NT	Brisbane	40	-19.66	134.19	1978	1986	6.64	0.70	0.68
2	Mt Isa	QLD	Brisbane	77	-20.73	139.50	1965	1454	10.25	0.71	0.74
3	Longreach	QLD	Brisbane	43	-23.44	144.26	1970	880	9.31	0.73	0.72
4	Atherton	QLD	Brisbane	42	-17.22	145.48	1965	1328	9.39	0.73	0.73
5	Townsville	QLD	Brisbane	42	-19.32	146.82	1965	1055	9.78	0.74	0.73
6	Charleville	QLD	Brisbane	43	-26.40	146.25	1960	553	9.55	0.70	0.72
7	Roma	QLD	Brisbane	42	-26.56	148.79	1955	309	9.28	0.69	0.71
8	Toowoomba	QLD	Brisbane	42	-27.72	151.63	1875	0	10.83	0.74	0.77
9	Armidale	NSW	Sydney	44	-30.52	151.67	1905	354	11.27	0.81	0.82
10	Dubbo	NSW	Sydney	39	-32.22	148.63	1905	257	10.63	0.79	0.79
11	Sydney	NSW	Sydney	42	-33.62	150.82	1865	0	10.32	0.80	0.80
12	Goulburn	NSW	Sydney	44	-34.76	149.70	1905	163	12.19	0.78	0.78
13	Cobar	NSW	Syd/Melb	38	-31.49	145.83	1905	524	12.11	0.81	0.82
14	Wentworth	NSW	Melbourne	39	-34.11	141.92	1905	490	13.62	0.81	0.83
15	Burrumbuttock	NSW	Melbourne	25	-35.84	146.80	1905	274	11.20	0.81	0.83
16	Albury	NSW	Melbourne	35	-36.07	146.94	1905	274	12.36	0.86	0.83
17	Melbourne	VIC	Melbourne	42	-37.79	144.92	1863	0	12.97	0.81	0.82
18	Geelong	VIC	Melbourne	40	-38.18	144.37	1863	64	14.27	0.82	0.85
19	Mt Gambier	SA	Adel/Melb	41	-37.86	140.85	1868	360	11.49	0.78	0.79
20	Broken Hill	NSW	Adelaide	41	-31.95	141.47	1897	458	11.84	0.81	0.81
21	Adelaide	SA	Adelaide	42	-35.23	138.50	1865	0	10.97	0.78	0.81
22	Coober Pedy	SA	Adelaide	20	-29.01	134.75	2000	777	7.69	0.82	0.77
23	Wynyard	TAS	Hobart	38	-40.97	145.65	1875	260	12.36	0.83	0.82
24	Bridport	TAS	Hobart	43	-41.00	147.39	1875	215	11.04	0.83	0.82
25	Hobart	TAS	Hobart	43	-42.94	147.35	1867	0	11.30	0.77	0.80
26	Auckland	North Is	North NZ	41	-36.87	174.78	1867	0	13.18	0.79	0.81
27	Wellington	North Is	North NZ	43	-40.62	175.29	1866	0	13.36	0.79	0.82
28	Christchurch	South Is	South NZ	42	-43.53	172.63	1867	0	11.73	0.84	0.83
29	Dunedin	South Is	South NZ	44	-45.88	170.50	1868	0	13.27	0.83	0.84
30	Morthen	UK	South NZ	40	51.75	-1.25	NA	NA	12.98	0.81	0.85

Table 2. Effects of range expansion on allelic richness and genetic diversity (H_E). Year of population establishment and distance from source introduction could not be used as fixed effects in the same model due to high collinearity. We choose to use "year established" as a proxy for sequential colonization events. There is a consistent significant negative relationship between the year populations were established and genetic diversity. These relationships are plotted in Fig. 4. Mean allelic richness and expected heterozygosity was calculated using the 11 loci for the 29 populations in Australian and New Zealand. The marginal R^2 is the variance explained by the fixed effect of year established and the Conditional R^2 is for the total model.

a) Mean allelic richness					
	Estimate	SE	df	t	P
Intercept	64.32	10.61	16.1	6.06	<0.0001
Year established	-0.028	0.006	15.9	-5.00	**<0.0001**
	Variance	SD	n		ICC
Marginal R^2	1.473				0.527
Population	0.227	0.477	6		0.081
Residual	1.095	1.047	29	Conditional R^2	0.608
b) Mean genetic diversity					
	Estimate	SE	df	t	P
Intercept	1.667	0.23	26.4	7.33	<0.0001
Year established	-0.0004	0.0001	26.4	-3.84	**0.0007**
	Variance	SD	n		ICC
Marginal R^2	0.0004				0.301
Population	0.0006	0.024	6		0.437
Residual	0.0003	0.019	29	Conditional R^2	0.739

Figures

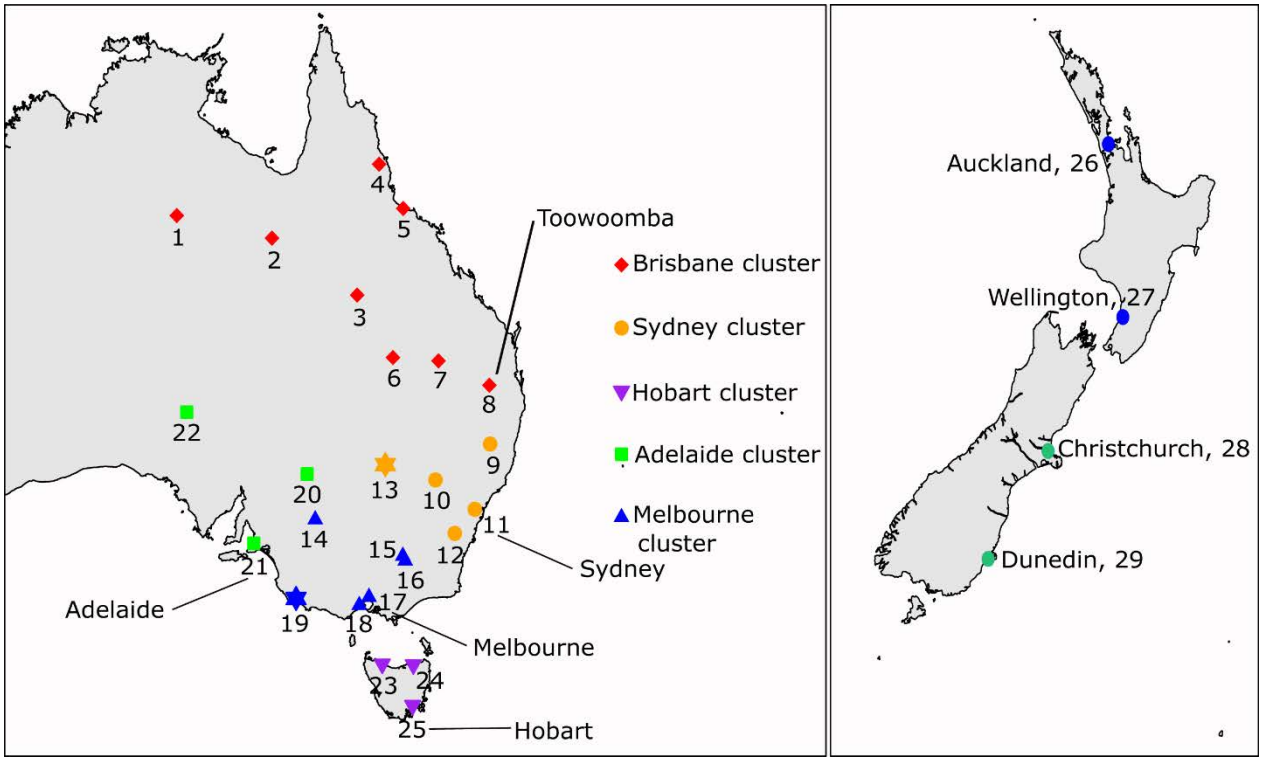


Figure 1. Map of sampling sites. The numbers next to the points for the sampling localities are the same as the ID numbers in Table 1. The colour coding is linked to the genetic population clusters described in Fig. 2. The five labelled sites on Australia are those closest to the major cites that housed acclimatisation society chapters that are link to the sparrows' introduction. The two stars represent two sample sites with uncertain population allocations possible due to admixture.

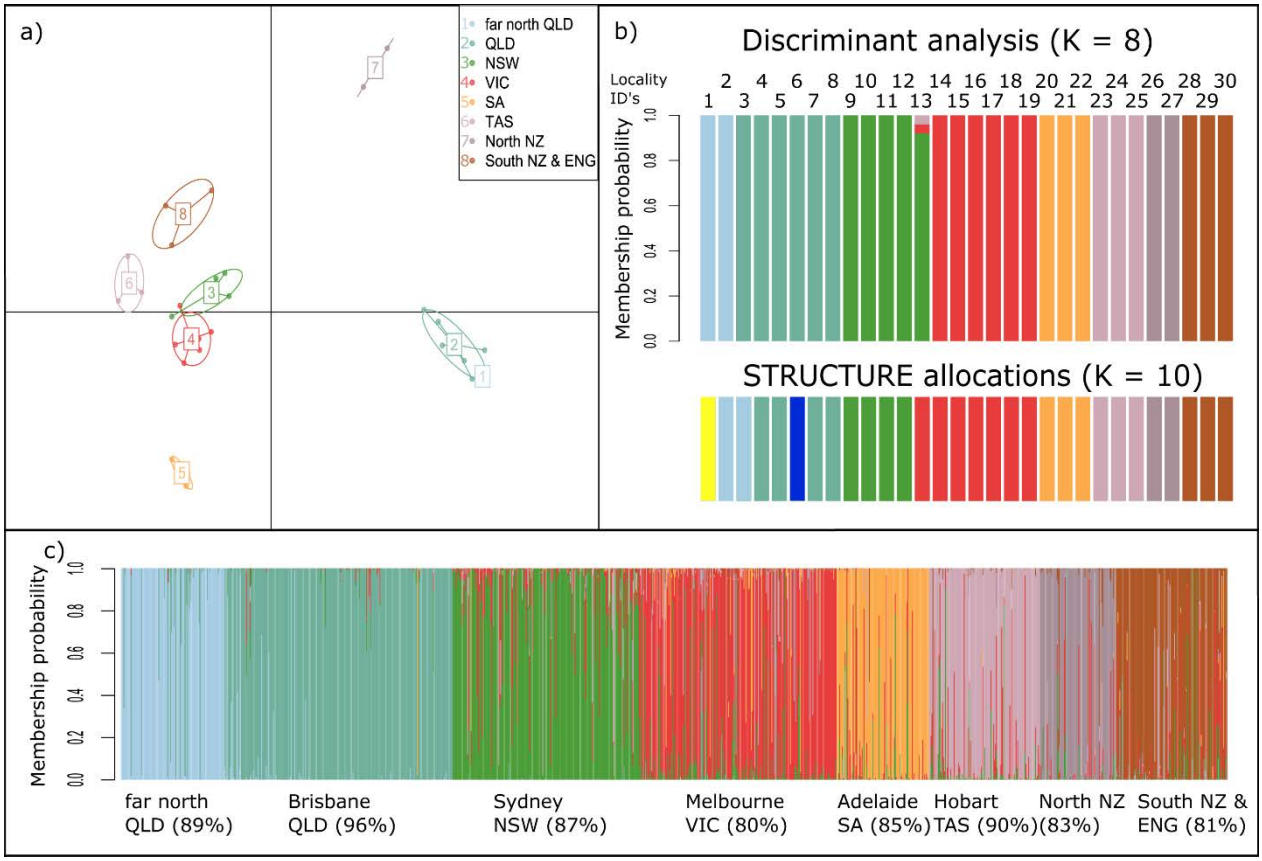


Figure 2. Discriminant Analysis of Principle Components. Panel a) shows the scatter plot for the DAPC of the 30 sample localities that found 8 clusters (see figure key). Panel b) shows the membership probability of each locality to the clusters, only 13 (Cobar) had mixed membership. Below the membership probabilities is a summary of how localities were allocated to clusters using STRUCTURE (Fig. S2). Panel c) uses the 8 genetic population ID's from panel a) and calculates the membership probabilities of individuals using a second DAPC. The percentages show the proportion of individuals correctly assigned to their predicted cluster.

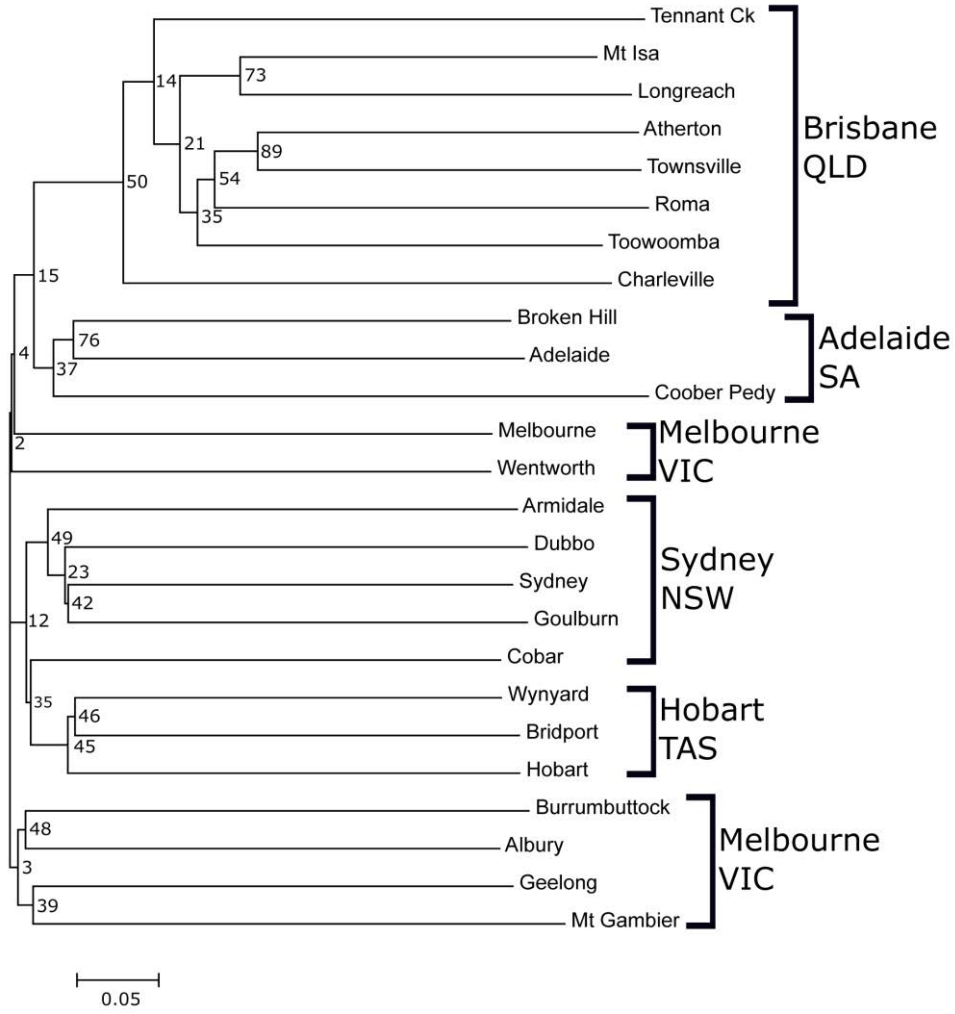


Figure 3. Neighbour joining tree of Australian localities. Labels to the right link localities in the same population cluster from the DAPC (Fig. 2). The Tree puts the sample sites linked to the Brisbane cluster in a clade with a large separation from the other sites. Sample sites linked to the Sydney, Adelaide and Hobart clusters are consistently grouping in their own independent clades. Cobar is on a branch between the Sydney and Hobart clades showing again that it is not consistently grouping with the same localities in different analyses. The localities linked to the Melbourne population are mixed across multiple clades. Bootstrap percentages for nodes are also included on the tree, calculated using subsets of loci. In general sister branches on the tree are sample sites with low pair-wise F_{ST} (Fig. S1). Bootstrap values are relatively high within clades but there is no clear relationship between clades illustrating invasion roots across Australia.

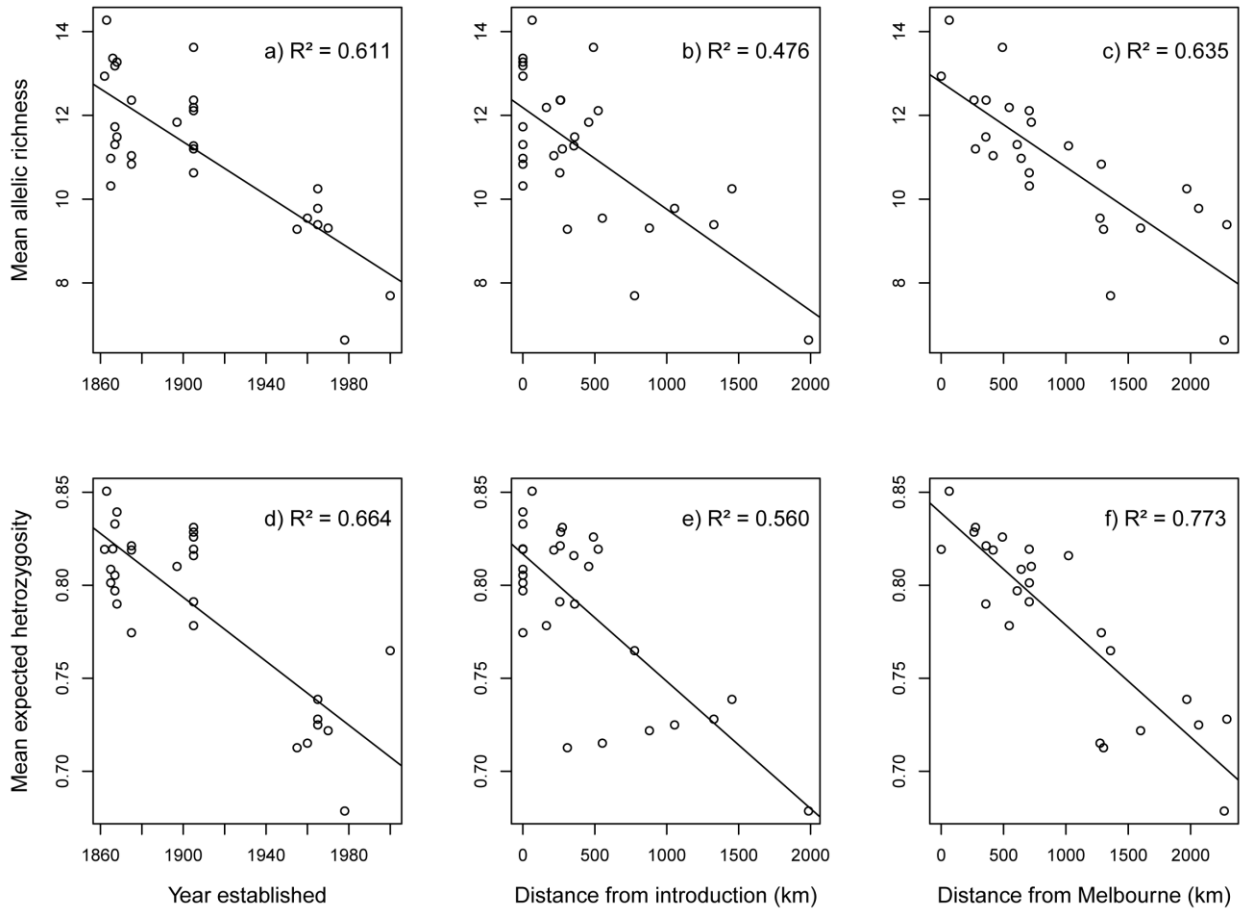


Figure 4 Genetic diversity has a negative relationship with time and distance from introduction sites. Both allelic richness and expected hetrozygosity decline with the recorded year the population was established (panel a and d) and the distance from the proposed introduction site (panel b and e). However, the same negative relationship exists between these diversity metrics and the distance a population is from the Melbourne introduction site (panel c and f). Melbourne was probably the release point for the largest number of sparrows imported to Australia and currently has relatively high levels of genetic diversity. These graphs plot the raw data for each locality, lines of best fit and R^2 values are from linear regressions using the two variables (see Table S5 for details).

Supplementary Material Chapter Five

The genetic structure of the introduced house sparrow populations in Australia and New Zealand is consistent with historical descriptions

Authors: Samuel C. Andrew; Monica Awasthy; Peri E Bolton; Lee Ann Rollins; Shinichi Nakagawa and Simon C. Griffith

Journal: *Biological Invasions*

Corresponding Author: Samuel C. Andrew

e-mail: samuel.andrew@students.mq.edu.au

Supplementary Figures

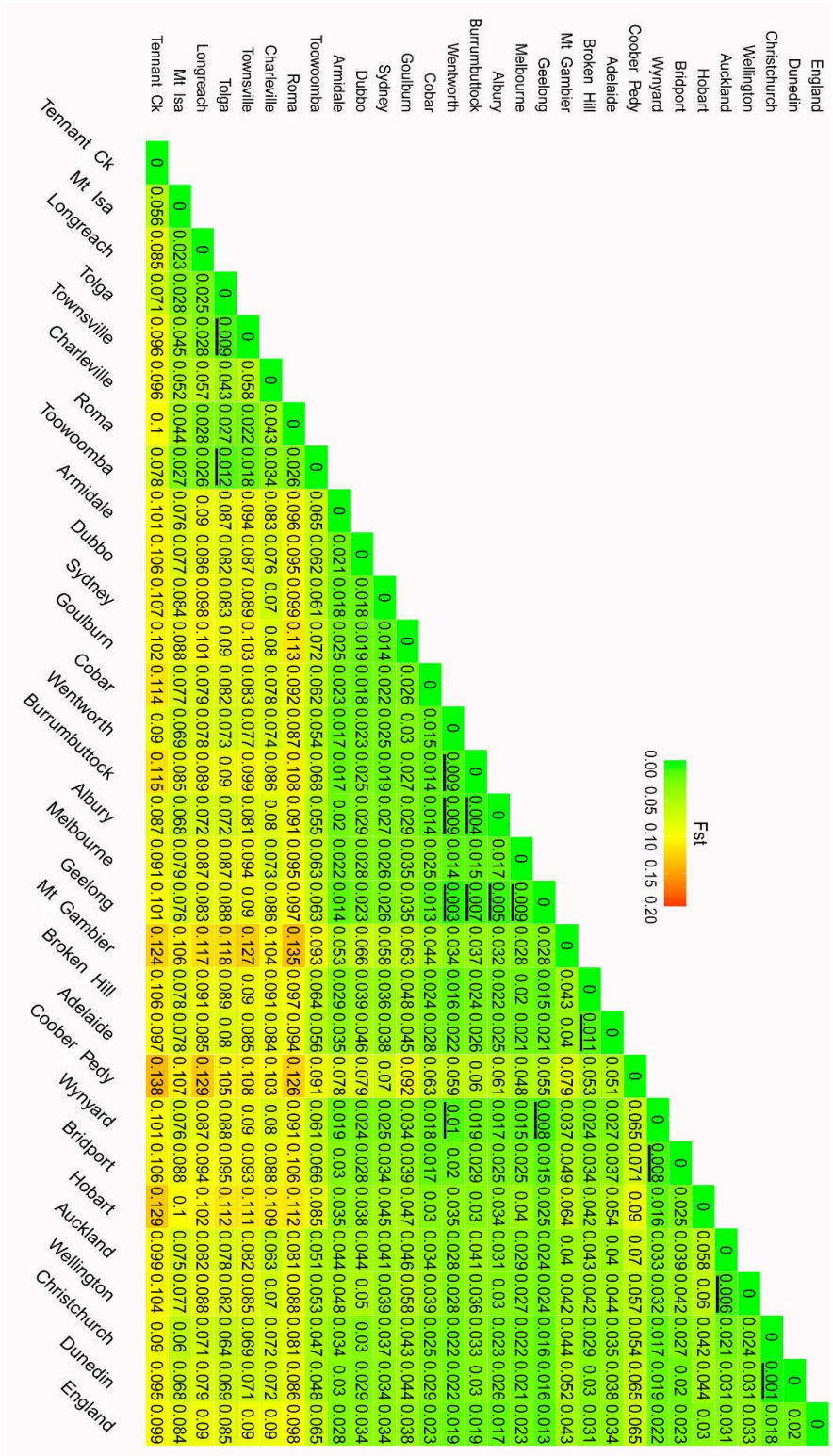


Figure S1 Pair-wise F_{ST} Matrix. This matrix includes all thirty localities from Australia, New Zealand and England. After Bonferroni correction significant *P* values are less than 0.0001 (equivalent to a F_{ST} of less than 0.013). Nonsignificant pair-wise comparisons are underlined. Nonsignificant *P* values ranged from 0.0001 to 0.318. 420 out 435 pair-wise comparisons were significant.

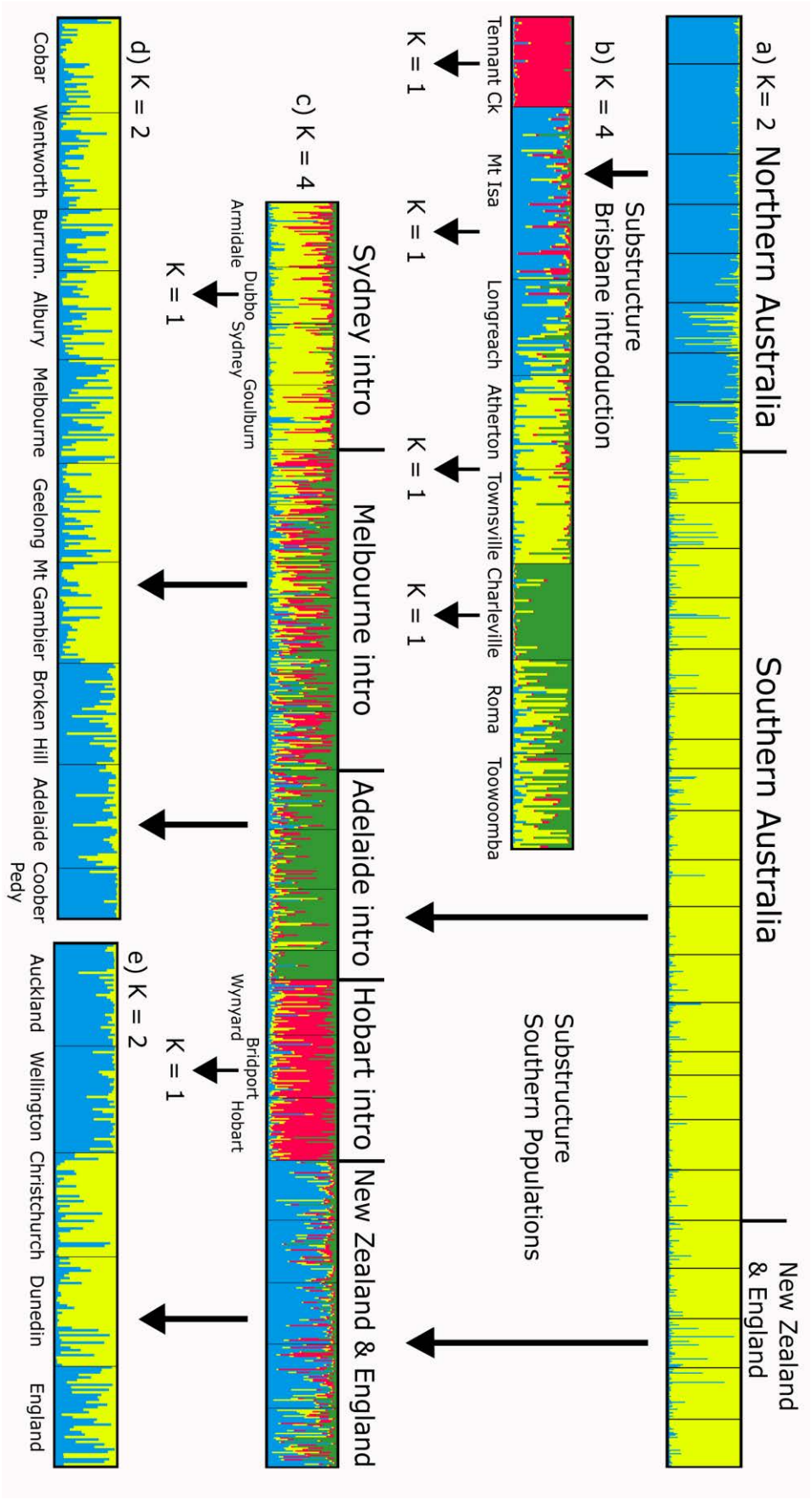


Figure S2. (see legend below)

Figure S2 (Above). Q-plots from STRUCTURE Analyses. Q-plots were drawn using the summary of 10 iterations for the most supported value of K (Table S2). Populations are ordered so they are next to sample sites from the same geographic region. Population clusters are labelled by proposed introduction sites in the region ("intro"). After accounting for substructure we have identified 10 clusters (summary Figure 2b). This first level of population structure clearly separated the populations north of the NSW/QLD border from the populations in southern Australia, New Zealand and England (Figure S2a). We also found clear substructure within these two main clusters. For the north Australian sample sites, a separate analysis found the most supported value of K was 4 (Figure S2b, n = 8 sites). Within these four northern clusters, we found no substructure (results not presented). In the second main cluster including southern Australia, we found substructure that supported a K of four (Figure S2c, n = 22 sites). Individuals from the localities near Melbourne had mainly mixed allocations (see Figure S2c). The one inconsistence between Figure S2c and S2d is Mt Gambier appears to no longer be grouping with Adelaide.

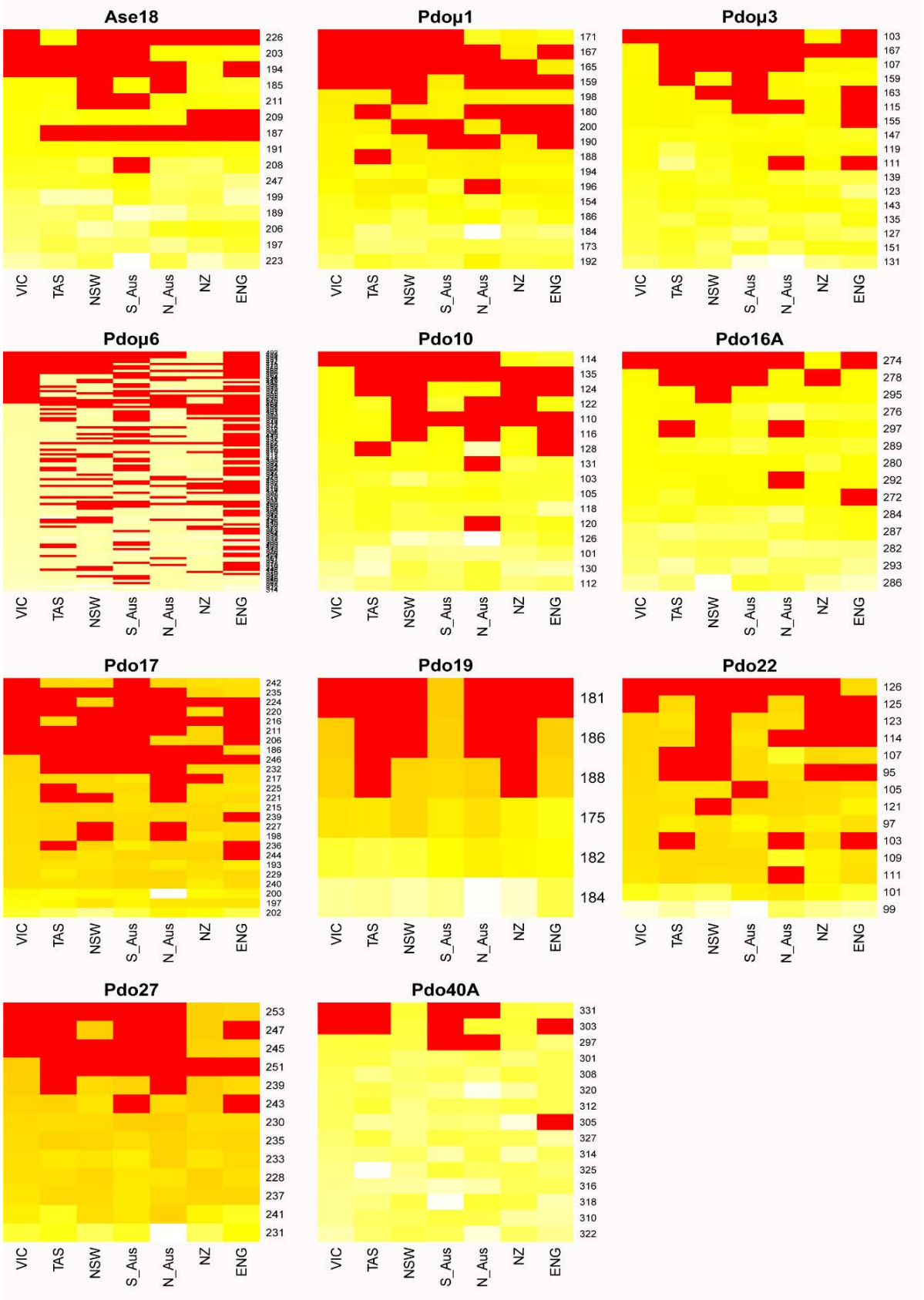


Figure S3. (See legend below)

Figure S3. Private allele heatmaps comparing all major populations. This figure plots the allele frequencies for each locus separately. Allele frequencies are plotted for the 7 major groups of localities (Table 1). The alleles are ordered by their frequency in the Melbourne cluster (VIC) so that it is easy to identify Alleles that are not present in this population but are present in tother regions that are reported to have had independent introductions: Brisbane (N_Aus), Adelaide (S_Aus) and New Zealand (NZ). There were also reported introductions to Hobart (Tas) and Sydney (NSW) from the recently established population in Victoria in the 1860's. The home range of England (Eng) is also a group but this group had only one sampling locality. Localities are assigned to groups based on genetic populations, Cobar and Mt Gambier were removed from the analysis because they had mixed allocations between methods. Dark red indicates an Allele frequency of 0 and loci with light yellow to white shading indicate alleles with the highest frequencies observed within a population.

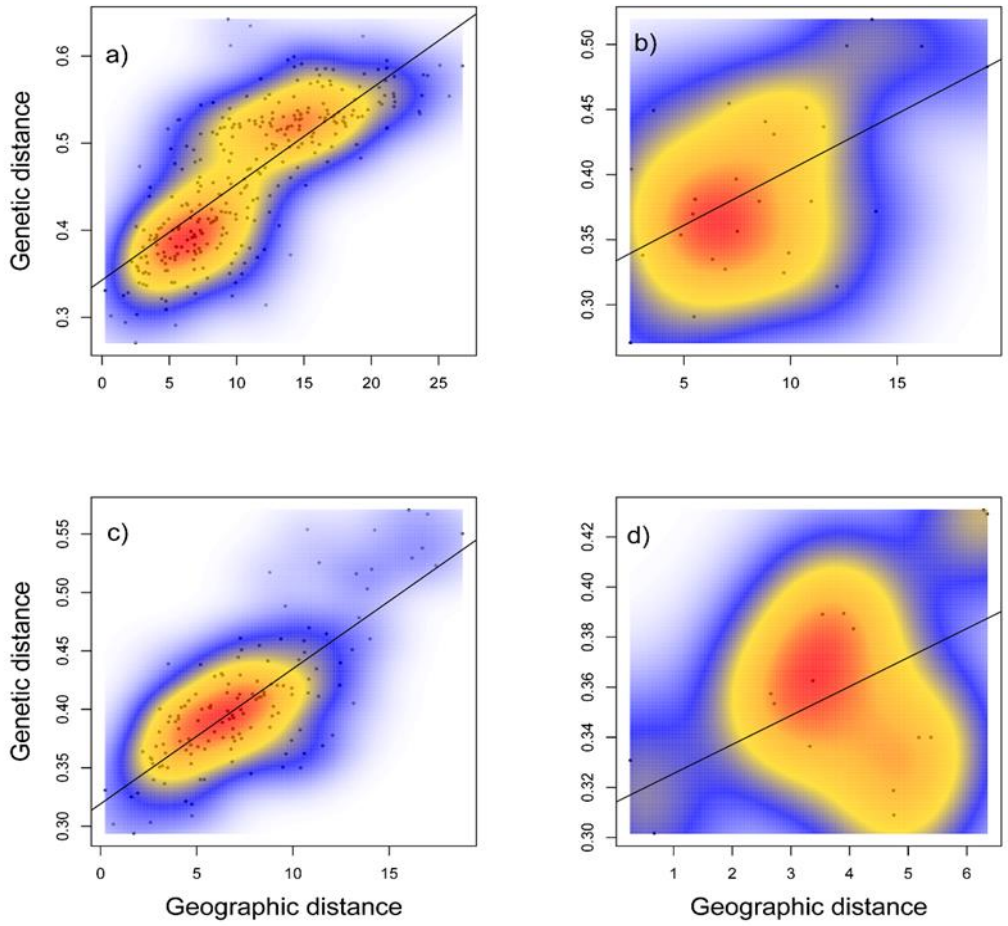


Figure S4. Isolation by distance plots**.** Panel a) is a IBD plot for all 25 of the Australian sampling sites, there is a highly significant relationship that indicates IBD ($R^2 = 0.758$, n = 25, P = 0.001). Panel b) Uses the eight localities linked to the Brisbane cluster and shows a significant relationship ($R^2 = 0.545$, n = 8, P = 0.024). Panel c) Uses the 17 southern Australian localities and there is also a highly significant effect of IBD ($R^2 = 0.767$, n = 17, P = 0.001). Panel d) uses only the six localities that cluster together with the Melbourne introduction in Figure S2d ($R^2 = 0.520$, n = 6, P = 0.020).

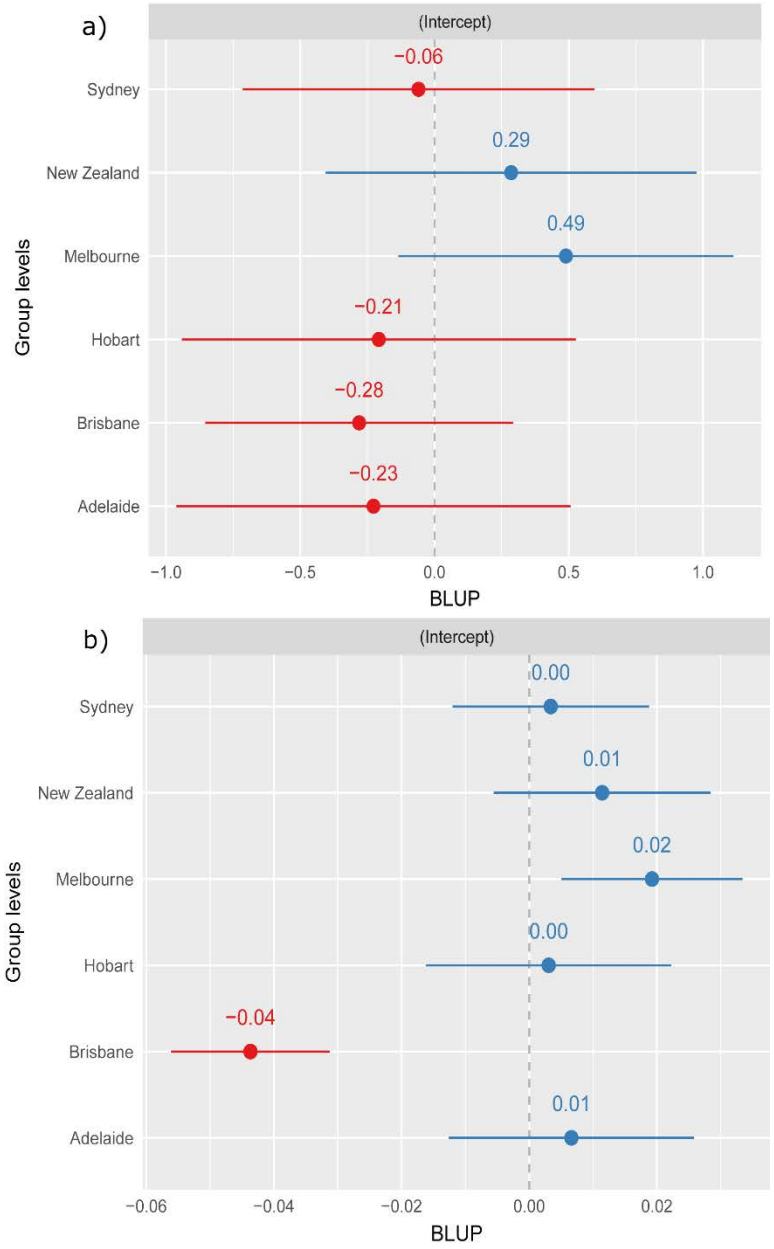


Figure S5. Differences in genetic diversity between populations. Panel a) plots the intercepts for the random factor of population in the LMM for Allelic richness (Table 2a). Panel b) plots the intercepts for the random factor of population in the LMM for expected hetrozygosity (Table 2b). This visualizes differences in average genetic diversity between population clusters from different regions rather than as individual localities. These intercepts are also from models that include the estimated year populations at localities were established. The Melbourne cluster had the highest intercept for both models with New Zealand second. Brisbane has the lowest intercept for Allelic richness and expected hetrozygosity. Expected hetrozygosity is significantly lower than all other clusters for Brisbane. The x-axis is the Best Linear Unbiased Prediction (BLUP) of the intercept of the different random factor levels with 95% C.I.

Supplementary Tables

Table S1. Notes on house sparrow introduction history to New Zealand. We found no indication that sparrows were imported from any location other than Great Britain to New Zealand. Thomson (1922) gives multiple examples of advertisements in English newspapers for travellers to bring sparrows to sell in New Zealand to Acclimatisation Societies. However, it is still possible sparrows were imported from other countries.

Introduced to	Date	Origin of birds	Number of birds	Quotes from the source text	Reference	Our Comments
Auckland	August 1859	From England on the "*Swordfish*"	300 liberated by Mr Brodie	"Already helping with the Caterpillar plague by September 1859"	Hargreaves A (1943) Introduction of Sparrows. *Victorian Naturalist*, 60, 96.	This source is highly contradictory to a newspaper article that reports all birds perished on the voyage (See below).
Auckland	July 1859	From England on the "*Swordfish*"	300 died at sea	"The attempt to import sparrows from England has been very unfortunate: out of 300 put on board the Swordfish, which arrived at Auckland on the 11th, not one survived the excessive bad weather."	The Sydney Morning Herald, NSW, 30 July 1859, page 7)	
Nelson Society	1862 and 1864	Shipped from England	1 bird arrived in each attempt	"But the Nelson Society forestalled Wanganui, for they succeeded in bringing in one sparrow in 1862." "In 1864 the Nelson Society imported a number of sparrows, but only one was landed alive."	Thomson GM (1922) The Naturalisation of Animals and Plants in New Zealand. Cambridge: University Press.	
Auckland	1865	From Glasgow, Great Britain on the "Viola"	2	"landed two sparrows out of six dozen which were shipped"	Thomson (1922)	
Wanganui Society	1866	Shipped from England	Possible as many as 200	"According to Sir Walter Buller the Wanganui Society introduced sparrows in 1866,"	Thomson (1922) and Williams G (1953) The dispersal from New Zealand and Australia of some introduced European passerines. *IBIS*, 95, 676–692.	
Canterbury Society	1867	Shipped from England	40	"The Canterbury Society in 1864 printed a list of prices which they offered to immigrants for each Pair... the same Society liberated forty sparrows in 1867, and the annual report for 1871 states that they are "thoroughly established and need no further importation""	Thomson (1922) and Lever C (2005) Naturalised Birds of the world. In:, pp. 204–218. T & A D Poyser, London.	By 1871 the species was fully established

Auckland Society	1867	Shipped from England	47	"This Society liberated 47 sparrows in 1867 and in the annual report for 1868, "consider them thoroughly acclimatised."'	(Thomson 1922)	A letter by Mr T. B. Hill (see Thomson (1922)) reports that he purchased sparrows and successful breed them and even gave some to "friends in the country who were anxious to get them".
Otago	1868, 1869	Shipped from England	3, 11	"The Otago Society liberated three in 1868 and 11 in 1869."	(Thomson 1922)	
Stoke	1871	Shipped from England	6	"In 1871 six were introduced, and were liberated at Stoke, where they soon increased."	(Thomson 1922)	These birds soon increased in number.

Table S2. Delta K and LnPD summaries from STRUCTURE analyses shown in Figure S2. Values of K with the most support are in bold.

	All localities (N = 30 samples)			**Northern Aus cluster (N = 8)**			**Southern cluster (N = 22)**			**Sydney cluster (N = 4)**		
K	Delta K	mean LnP	SD LnP	Delta K	mean LnP	SD LnP	Delta K	mean LnP	SD LnP	Delta K	mean LnP	SD LnP
1	—	-61255.1	0.282	—	-15373.9	0.288	—	-43389.3	0.314	—	**-7692.78**	**0.286**
2	**2601.687**	**-58855**	**0.773**	244.682	-14937.9	1.238	4.124	-43012	14.356	12.674	-7699.55	10.110
3	5.918	-58467.3	10.134	19.042	-14804.8	4.571	16.817	-42694	4.251	5.553	-7834.45	131.265
4	1.653	-58139.5	49.616	**298.117**	**-14584.7**	**2.381**	**39.142**	**-42447.4**	**4.046**	1.073	-8698.28	716.076
5	52.136	-57893.7	4.476	6.408	-15074.5	145.583	7.309	-42359.2	13.840	1.783	-10330.5	1264.093
6	—	-57881.3	13.930	27.275	-14631.4	22.497	0.494	-42372.1	540.100	—	-9708.3	777.412
7				0.618	-14801.9	299.324	0.572	-42651.8	747.141			
8				0.680	-15157.4	101.852	6.794	-42503.8	97.350			
9				0.709	-15582.1	118.515	0.933	-43017.2	663.412			
10				—	-16090.8	388.377	0.510	-44149.7	1241.911			
11							0.599	-44649.4	390.279			
12							—	-45382.6	1070.279			
	Victoria and South Aus (N = 10)			**Tasmania (N = 3)**			**New Zealand & England (N = 5)**			**New Zealand (N = 4)**		
K	Delta K	mean LnP	SD LnP	Delta K	mean LnP	SD LnP	Delta K	mean LnP	SD LnP	Delta K	mean LnP	SD LnP
1	—	-17900	0.291	—	**-5690.83**	**0.574**	—	-10526.6	0.457	—	-8384.35	0.690
2	**193.924**	**-17847.9**	**5.399**	83.466	-5701.1	13.446	**24.293**	**-10444.9**	**10.424**	**9.985**	**-8250.84**	**47.417**
3	4.291	-18842.8	281.713	3.134	-6833.62	485.541	12.184	-10616.4	69.143	2.308	-8590.77	58.205
4	0.865	-18628.8	379.701	1.503	-6444.48	248.724	0.235	-11630.3	376.436	0.460	-8796.36	88.395
5	10.895	-18086.3	106.545	1.073	-6429.19	232.690	1.959	-12555.8	729.449	3.444	-9042.57	146.410
6	0.111	-18704.7	316.560	—	-6164.32	89.259	4.505	-12052.6	139.881	1.173	-9793.07	462.776
7	0.563	-19358.2	548.674				0.200	-12179.6	380.784	3.277	-10000.7	155.139
8	0.478	-19702.9	292.100				—	-12382.6	327.555	—	-9699.83	249.358
9	0.817	-20187.1	684.014									
10	—	-20112.1	611.983									

Table S3. Microsatellite loci summary. Including: the number of individuals genotyped (*n*); number of alleles genotyped for the locus; observed hetrozygosity (H_O); expected hetrozygosity (H_E); allelic range; and estimates of null allele frequencies (Null freq.). Loci in bold were removed from further analyses because of the parameter in bold and italics.

Locus	*n*	No. Alleles	H_O	H_E	Allelic range	Null freq.
Ase18	1222	15	0.81	0.87	62	0.035
Pdoμ1	1236	16	0.78	0.85	46	0.041
Pdoμ3	1237	17	0.83	0.87	64	0.028
Pdoμ5	1208	15	0.58	0.78	36	***0.151***
Pdoμ6	1234	106	0.94	0.98	224	0.020
Pdo9	***1054***	19	0.72	0.82	56	0.067
Pdo10	1231	16	0.81	0.87	34	0.037
Pdo16A	1236	14	0.83	0.87	25	0.023
Pdo17	1228	25	0.77	0.83	60	0.038
Pdo19	1237	7	0.58	0.59	64	0.014
Pdo22	1237	14	0.68	0.76	31	0.055
Pdo27	1236	13	0.69	0.76	25	0.055
Pdo40A	1216	15	0.84	0.91	34	0.038

Table S4. AMOVA's for population differentiation among sample localities. The first AMOVA calculates among population F_{ST} for all 30 localities. Separate AMOVA's for Australian and New Zealand sample localities were also run.

AMOVA for all 30 localities					
Source of Variation	d.f.	Sum of squares	variance components	Percentage of variation	P
Among populations	29	734.5	0.255	5.6	<0.001
Among individuals within populations	1207	5226.4	0.032	0.7	0.037
Within individuals	1237	5277.5	4.266	93.7	<0.001
Total	2473	11238.4	4.553		
All Australian localities					
Source of Variation	d.f.	Sum of squares	variance components	Percentage of variation	P-Value
Among populations	24	636.9	0.272	6.01	<0.001
Among individuals within populations	1002	4273.2	0.018	0.41	0.182
Within individuals	1027	4342	4.228	93.58	<0.001
Total	2053	9252.1	4.518		
New Zealand localities					
Source of Variation	d.f.	Sum of squares	variance components	Percentage of variation	P-Value
Among populations	3	36.2	0.088	1.9	<0.001
Among individuals within populations	166	766	0.079	1.71	0.046
Within individuals	170	757.5	4.456	96.39	<0.001
Total	339	1559.7	4.623		

Table S5. Summary of linear models for lines of best fit in Fig. 4 (see main text). Ar stands for mean allelic richness and He is for mean expected hetrozygosity.

Test	Estimate	d.f.	*t* value	*P* value	R^2
Ar vs Year established	-0.032	27	-6.512	< 0.0001	0.611
He vs Year established	-0.001	27	-7.298	< 0.0001	0.664
Ar vs Distance from intro	-0.002	27	-4.951	< 0.0001	0.476
He vs Distance from intro	-0.0001	27	-5.866	< 0.0001	0.560
Ar vs Distance from Melbourne	-0.002	23	-6.319	< 0.0001	0.635
He vs Distance from Melbourne	-0.0001	23	-8.860	< 0.0001	0.773
Ar vs Latitude	0.164	23	4.506	< 0.0001	0.469
He vs Latitude	0.005	23	7.007	< 0.0001	0.681

Chapter Six

Signatures of genetic adaptation to extremely varied Australian environments in introduced European house sparrows

Chapter Six Vignette

The house sparrow has expanded its distribution to cover a wide range of Australian climates that contrast the conditions experienced by the primary source populations in Western Europe. The population structure created by the colonisation of the fragmented habitat of town and cities across Australia (Chapter Five), provides a good system for studying local adaptation in an invasive species. The relatively high isolation between populations in different climates allows for allele frequencies to change at specific loci targeted by selection. These changes in allele frequencies are the signatures of selection that are expected to be present in locally adapted populations. Using genome wide SNP data from the house sparrow 200K SNP array we hope to identify these signatures of selection using genome scan methods. These genome scan methods identify outlier loci with levels of differentiation above the background levels of differentiation across the genome that is a result of the population structure created by neutral genetic drift. The development of advanced analytical techniques that account for the underlying population structure, make the study of local adaptation in introduced species with complex demographic histories more viable.

This manuscript is prepared for submission to Molecular Ecology, and is formatted as such.

Authors: Samuel C. Andrew[1]; Henrik Jensen[2]; Ingerid J. Hagen[2,3]; Sigbjørn Lien[4]; and Simon C. Griffith[1]

[1]Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia

[2]Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway.

[3]Norwegian Institute for Nature Research, P.O. Box 5685 Torgard, NO-7485 Trondheim, Norway.

[4]Centre for Integrative Genetics (CIGENE), Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway.

Keywords: *Passer domesticus*, invasive species, local adaptation, ecological association, population differentiation, natural experiment.

* To whom correspondence should be addressed

Samuel C. Andrew

Department of Biological Sciences, Macquarie University, New South Wales, 2109, AUSTRALIA

E-mail: samuel.andrew@students.mq.edu.au

Running title: Local adaptation in introduced sparrows

Abstract

Due to its history of multiple introductions to novel environments world-wide, the house sparrow has been used as a model species to study local adaption in invasive avian species. New genomic resources such as a custom 200K SNP array and a house sparrow reference genome provide great prospects for studying rapid local adaptation in this invasive species. Here we analyse high-density genome-wide genetic data collected across an extensive range of temperate, arid and tropical climates, in Australian populations that were introduced from Europe 150 years ago. We used both Population Differentiation (PD) and Ecological Association (EA) methods to identify putative loci subject to selection across these varied climates. A majority of the outlier SNPs thus identified were through the use of the Latent Factor Mixed Models (LFMM) EA method but the BayeScEnv EA method had the strongest overlap with the outliers from the two PD methods. Out of all the 971 outliers identified across the different methods, 38.3 % were physically linked (within 20 kbps) to 575 known protein coding regions in the house sparrow reference genome. Interestingly, some outlier genes had been previously identified in genome scan studies of broadly distributed species or had strong links to traits that are expected to be important to local adaptation e.g. heat shock proteins, immune response and HOX genes. However, many outliers still have unknown relevance or could be false positives. Our results identify an opportunity to use the house sparrow to further study outlier genes that are potentially linked to local adaptation in an invasive species.

Introduction

Phenotypic variation is critical to natural selection and local adaptation, and evidence of selection acting across varied environmental conditions is often provided by observing phenotypic differentiation between populations (Ashton 2002; Arthur *et al.* 2008). Clinal phenotypic variation can be explained by a combination of both phenotypic plasticity linked to the environment and selection acting on molecular variation linked to both protein coding genes and gene regulation (Price *et al.* 2003). Genome scan studies aim to identify regions of the genome that are strongly differentiated between populations due to selection for adaptation to local conditions (Savolainen *et al.* 2013). Genome wide SNP markers can be used to identify the putative loci that are influenced by selection due to them being physically linked with loci responsible for variation in fitness and natural history traits (Hoban *et al.* 2016). An increasing number of studies in evolutionary ecology have studied the genetic regulation of a variety of traits such as: thermal tolerance (Hoffmann *et al.* 2003, 2012; Thomas *et al.* 2017), immune function (Zueva *et al.* 2014; Sin *et al.* 2014; Kim *et al.* 2015; Shultz *et al.* 2016) and reproductive life history (Barson *et al.* 2015). Since 2010 there have been an increasing number of genome scan studies focused on non-model organisms (Haasl & Payseur 2016), but those that have linked outlier loci to phenotypic and fitness variation, or genes related to fitness traits, are still very limited (Jensen *et al.* 2016).

Currently the most common genome scan methods use thousands of SNP markers across the genome to identify the loci that are subject to selection. There are two primary methods for identifying signatures of local adaptation in wild populations. Population Differentiation (PD) methods identify outlier loci across the genome with levels of differentiation among populations that are significantly above the background levels across all loci (Beaumont & Nichols 1996). Loci can be outliers due to higher than expected inter-population variation in allele frequency or heterozygosity due to directional selection on the linked trait, or lower than average differentiation due to balancing selection on the trait.

Selection can also drive significant 'clinal' phenotypic differences between populations due to minor changes in allele frequency that are correlated with the environment, without there needing to be significant overall inter-population differences in allele frequencies. Accordingly, Ecological Association (EA) models scan genomic markers for those that show changes in allele frequency that have the strongest clinal relationships with the environmental variables of interest (De Mita *et al.* 2013). Commonly targeted environmental variables for EA studies are the extreme temperatures experienced during both summer and winter as well as other important factors affecting survival, such as aridity, water availability and vegetation cover (Frichot *et al.* 2013; de Villemereuil & Gaggiotti 2015; Gautier 2015; Benestan *et al.* 2016).

Introduced species offer unique opportunities to observe genetic adaptation to drastically different environments than those they were adapted to in their natural home range (Colautti & Lau 2015). Another advantage of studying local adaptation in invasive species is that they provide a unique opportunity to observe evolutionary consequences of natural selection over short time scales. Introductions to environments that contrast with those in the source population(s) act as a kind of 'natural' experiment for local adaptation. Some introduced populations have been established for over 100 generations which is a reasonable amount of time for species to respond to dramatically different climate conditions and cause the changes in allele frequencies that are the signatures of local adaptation.

Here we aim to test if local adaptation has left strong signals of selection across the introduced house sparrow populations of Australia, that were established from introductions over 150 years ago (Andrew & Griffith 2016). The house sparrow (*Passer domesticus*) is an obligate commensal species with humans (Anderson 2006), and although the urban habitat and food sources that it derives from humans and their livestock are broadly similar across the world, they are exposed to considerable climatic variation. Australian house sparrow populations that were primarily introduced from temperate climates in Western Europe now

cover a wide range of climates with varied temperature extremes (Table 1). The distribution of the species now covers all of eastern Australia including temperate, tropical and arid biomes. The temperatures experienced in these hot arid regions are known to be near the physiological limits of even arid adapted avian species such as the zebra finch (Calder 1964). Climatic variables are therefore the focus in our own EA models. Climate is known to be linked in house sparrows to traits such as reproductive timing, feather density, moulting (Anderson 2006) and potentially thermal tolerance (Feder & Hofmann 1999). Genes with functional annotations relating to known fitness and life history traits are the primary target of genome scan studies. Studies in other species have previously identified some loci associated with adaptation to climate. For example genes associated with thermal tolerance in domestic ungulates (Kim *et al.* 2015), climatic variability in grey wolves (*Canis lupus*, Schweizer *et al.* 2016) and water temperature tolerance for coral reefs (Thomas *et al.* 2017). A group of genes that have broadly been associated with responding to heat stress are Heat Shock Proteins (HSP), these proteins are also expected to be related to local adaptation to extreme temperatures (Feder & Hofmann 1999; Sørensen *et al.* 2003; Bentley *et al.* 2017).

The house sparrow Affymetrix 200K SNP array (Lundregan *et al.* in prep) was used to generate high density genomic SNP data from introduced house sparrow populations across Australia. Using these data we aimed to identify loci that show levels of differentiation significantly above the general background levels across the genome by using methods that account for the existing population structure. These outlier loci will then be linked to genes on the annotated house sparrow reference genome (Elgvin *et al.* 2017). We test the replicability of our results by using two PD and two EA methods. We predict that in this introduced species significant outliers identified with PD methods will be of very low frequency across the genome due to high neutral differentiation caused by genetic drift (because of founder and bottleneck events, and relatively small population size in general), masking these signatures of selection. We also predict the EA methods will identify clinal changes in gene frequency

that are correlated with climatic variation due to selection, even in these recently established populations. Furthermore, HSP were selected as candidate functional loci identified *a priori* in the house sparrow reference genome to make focused observations of changes in allele frequency that covary with climate.

Methods

Sampling design

Background knowledge on genetic population structure will benefit the sampling design of genomic studies. For example, a paired sampling design uses pairs of sample sites that are taken within a genetic population but maximises their environmental differentiation. This design with multiple pairs of sites can increase the statistical power of models to detect consistent changes in allele frequency due to selection rather than stochastic changes due to drift and founder effects (Lotterhos & Whitlock 2015; Rellstab *et al.* 2015). When determining sampling design, population genetic information can also be used to avoid including the most inbred and bottlenecked populations to reduce the false discovery of outliers (Foll & Gaggiotti 2008). Previous descriptions of the Australian house sparrows' introduction history (Chapter Two [Andrew & Griffith 2016]) and population structure (Chapter Five) have helped inform the sampling design of this study. Specifically in earlier work we identified independent genetic clusters reflecting the introductions of the species into the Melbourne, Adelaide, Hobart, Sydney and Brisbane regions (Chapter Five). We used samples from each of the main climatic regions that the species occupies in Australia and avoided sample sites that had potentially been through the most recent demographic bottlenecks. We also used pairs of sample sites within the same genetic cluster and tried to maximise their climatic differentiation.

Adult house sparrows were sampled in urban areas from 11 towns/cities across Australia, we genotyped 16 male individuals from each of our sampling localities (Table 1, Figure 1). Only males were genotyped because of other research goals and males also have the advantage of being the homogametic sex for consistent analysis of Z chromosome loci. The sampling was carried out during April to September 2014 and in March 2015 under the Animal Research Authority of the Animal Ethics Committee at Macquarie University (ARA 2014/248). Birds were captured using mist nets and placed in bird bags until a blood sample could be taken. Blood was taken from the brachial vein with a capillary tube (ca. 40 µl) and was stored in 800 µl of absolute ethanol in a 1.5ml Eppendorf tube.

Environmental data

The latitude and longitude of sample sites was used to extract climatic data from the Worldclim bioclimatic variables data base (Hijmans *et al.* 2005). We extracted three variables for analysis from Worldclim: Maximum Temperature of Warmest Month (BIO5, hence referred to as "Summer maximum"), Minimum Temperature of Coldest Month (BIO6, hence referred to as "Winter minimum") and Temperature Seasonality (BIO4 = standard deviation *100). Seasonality uses the standard deviations of mean monthly temperatures across the year to quantify temperature change throughout the year. A fourth climatic variable was average Vapour Pressure Deficit (VPD) which is the difference between the current water vapour in the air and the saturated vapour pressure at the current temperature. We calculated average daily VPD using the daily maximum temperature and daily vapour pressure at 3pm. These climatic variables were downloaded from Australian Water Availability Project (Jones *et al.* 2009) via http://www.bom.gov.au/jsp/awap/. The esat function in the R library *plantecophys* (Duursma 2015) was used to convert daily maximum temperature to saturated vapour pressure. The VPD was averaged across the period of 1950 to 2016 for each site, and was used as a standard measure of each site's "dryness/aridity" (Jones 1992).

Molecular methods

DNA was extracted using a Gentra PureGene tissue kit (Qiagen, Valencia, CA, USA) following manufacturer's instructions. DNA concentrations were standardised to ca.12ng/μL (in 0.1 mM EDTA/10mM Tris buffer) and 100 μL of DNA in buffer was sent off for genotyping on the house sparrow Affymetrix 200K SNP array (Lundregan *et al.* in prep). DNA from 176 individuals were sent to CIGENE (Ås, Norway) for genotyping. Genotype clustering and initial quality checking of SNPs was done by CIGENE: 185 249 SNPs were found to be polymorphic high quality SNPs, and in total 14 751 SNPs where identified as low quality due to factors such as low call rate and/or being monomorphic. Using these 185 249 high quality SNPs we did further quality control and filtering. We removed SNPs with a minor allele frequency of less than 5% and SNPs that had been genotyped in less than 91% of individuals. After this filtering only one individual had over 5% missing data but it was only marginally over (5.3%) so all individuals were kept for future analyses. Finally, there were 162 299 SNPs used in further analysis and 176 individuals from eleven Australian sample sites. Calculating minor allele frequencies and missing data was done using the program PLINK (Purcell *et al.* 2007). Data conversion between PED, TPED and RAW formats was also done using PLINK.

Statistical Analyses

We began by describing the genetic population structure for our sample sites using the SNP data. The number of population clusters was visualised using a Discriminant Analysis of Principle Components (DAPC) in adegenet (Jombart 2008; Jombart *et al.* 2010). Due to the large size of the SNP dataset we split our data into 32 portions (ca. 5 000 SNPs in each) to be used in replicate analyses. Using the "optim.a.score" function, we identified 8 Principle

Components (PCs) as the best number for our DAPCs. Too many or too few PCs can lead to low repeatability of results and over or under fitting the data. The first DAPC was run using the sample site IDs as cluster labels. We also used the "find.clusters" function to identify the best number of meaningful clusters in our data (number of clusters with the lowest BIC). These cluster labels were then used in the DAPC to make a scatter plot.

Markers physically linked to loci affected by selection can show much higher than expected differentiation between populations and hence be a signal of directional selection (Whitlock 2015). However, not all differentiated loci are true signals of natural selection on genetic variation. The demographic history and population structure of study populations has been show to affect the False Discovery Rates (FDR) of outliers in genome scan analyses (de Villemereuil *et al.* 2014; Lotterhos & Whitlock 2014, 2015; Francois *et al.* 2016). Simulations have shown that populations that fit different demographic models such as Isolation By Distance, have different FDR estimates (Lotterhos & Whitlock 2014). Information on demographic history can therefore be important in selecting the most appropriate statistical methods for the genomic data (de Villemereuil *et al.* 2014; Lotterhos & Whitlock 2014). For example some methods account for the genetic population structure observed in the data or perform well when the population has been through a range expansion (de Villemereuil & Gaggiotti 2015; Rellstab *et al.* 2015). We choose two Population Differentiation methods (PD, outlier loci) and two Ecological Association methods (EA, Allele frequencies correlated with environmental variance) that have been shown in simulations to be most suitable for working on populations with similar demographic histories to invasive species. OutFLANK (Whitlock & Lotterhos 2015) and PCAdapt (Luu *et al.* 2017) were selected as the two PD methods because they were shown to have the lowest false discovery rate in range expansion simulations (Luu *et al.* 2017). A separate study also found that OutFLANK was a strong performer in simulations of meta-populations with complex demographic histories (Whitlock & Lotterhos 2015). Our two EA methods were LFMM (Frichot *et al.* 2013) and BayeScEnv

(de Villemereuil & Gaggiotti 2015). LFMM had a good balance of power and a low FDR in simulations with different demographic models (de Villemereuil *et al.* 2014). The newly developed ecological association method of BayeScEnv also appears to perform well in simulations of complex demographic scenarios and was intentionally designed to have a capability to account for underlying population structure (de Villemereuil & Gaggiotti 2015).

The PD methods of OutFLANK and PCAdapt were run in the statistical program R (R Core Team, 2017). All figures were also plotted in R and Venn diagrams were drawn using the package *VennDiagram* (Chen 2016). The options used for the OutFLANK analyses were as follows: the left and right trim used before the likelihood function was applied was 0.05; the minimum hetrozygosity required for inclusion was 0.1 and the desired false discovery rate threshold was 0.05 (q-threshold). The same q-threshold was used to identify outliers from the PCAdapt analyses using the R package *qvalue* (Dabney *et al.* 2010). PCAdapt is a PCs method for identifying outliers (Duforet-Frebourg *et al.* 2016; Luu *et al.* 2017). To work out the correct number of PCs to use in our PCAdapt analyses we started with K = 20 PCs, and identified 4 useful PCs using the method described in the tutorial for the package (Link: http://membres-timc.imag.fr/Michael.Blum/PCAdapt.html). The final analyses presented here used these 4 PCs. We also ran a Communality analysis with PCAdapt, this method option uses the first PC and generates a differentiation statistic more similar to F_{ST} (Luu *et al.* 2017).

The EA methods were run using the same genetic data and four separate environmental variables (see above). The LFMM analyses were run in R using a single repeat for each environmental variable, 10 000 iterations and a burn-in of 5 000. The specified K was 6 (number of population clusters, see methods above). For this method we choose the relatively strict q-threshold of 0.001 to define outliers. Preliminary runs using LFMM produced consistent results with the final repeat (more than 95% of all significant SNPs remained unchanged).

The BayeScEnv models were run separately for each environmental variable using the same options: 10 pilot runs with 2 000 iterations; a thinning interval of 10; 5 000 outputted iterations and a burn-in length of 10 000. All other options were kept as the default. Environmental variables were given in the form of environmental differentiation by using the difference of site values to the mean of all sample sites. The difference/distance was then standardised by dividing by the standard deviation across all sites. This is the recommended method for preparing environmental data for BayeScEnv (de Villemereuil & Gaggiotti 2015).

Linked genes and gene ontologies

The physical position of significant SNPs from each analysis were used to identify if the SNPs that were physically linked to functional regions on the reference genome. The house sparrow reference genome has 14 260 know protein coding genes (Elgvin *et al.* 2017). The average LD across the house sparrow genome drops to approximately half the background level at ca. 20 kbps, so this was used as a window for physical linkage between SNPs and genes (Lundregan *et al.* in prep). We recorded all genes within 20 kbps of each SNP and if the SNP was within a gene or within a known exon of a gene. The proportion of SNPs linked to genes was calculated for all analyses.

In many cases the genes from the house sparrow reference genome have been linked to analogous genes in other species with annotations for their function. These gene names are used across many species and have been given general functional descriptions called Gene Ontology terms (GO terms [Ashburner *et al.* 2000]). To assess if our lists of outlier genes included groups of genes with shared GO terms that are more common than we would expect if they were being sampled at random, we used the PANTHER (Mi *et al.* 2013, 2016) gene ontology analysis. Analyses were run with both the *Homo sapiens* and *Gallus gallus* (chicken) reference genomes.

Results

Out of the 200 000 SNPs on the house sparrow SNP array 185 249 were identified as high quality SNPs by CIGENE. This was a very similar number to the original filtering of Norwegian sparrow data (Lundregan *et al.* in prep). Within our Australian data we found 21 491 loci with a minor allele frequency of less than 5%. After this filtering, we removed a further 1459 loci for having a low call rate (less than 91% of individuals genotyped) which gave a final overall missing data frequency of less than 1%. There were 162 299 loci remaining for further analysis, that had been genotyped in 176 individuals from 11 sampling locations.

Across our sampling localities we found similar patterns of population structure, using the DAPC method, to previous descriptions of the house sparrow's population structure using microsatellite data (Chapter Five). Based on multiple iterations using different subsets of ca. 5000 SNPs we found the most support for 6 'meaningful' clusters across our 11 sampling locations (See details in supplementary Figure S1). As expected from the genetic structuring results based on microsatellites, our 11 sampling locations (Figure 1) fall within independent genetic populations that experience the same climate, e.g. the Arid environments of Broken Hill and Mt Isa and the Temperate climates of Melbourne, Mt Gambier and Adelaide. Some sampling localities were within the same genetic population but experience contrasting climates, e.g. Adelaide and Broken Hill, Toowoomba and Townsville (Tropical). This climatic variation within genetic populations fits with our intention for a paired sampling design (see methods).

OutFLANK and PCAdapt Population Differentiation results

The OutFLANK analysis found one significant outlier, this SNP (SNPa492110) was located inside an exon of a gene analogous to RNF20, on the Z chromosome. This SNP is potentially a target of balancing selection because all but one of the 176 males genotyped were heterozygous at this locus. The PCAdapt analysis using Mahalanobis Distance found 41 significant outliers. Of these SNP's 13 were physically linked to 21 protein coding regions identified on the house sparrow reference genome (Table 2). The Communality analysis method in PCAdapt found a more similar result to OutFLANK with no significant outliers. We extracted the top 100 OutFLANK loci with the smallest *P* values ("Suggestive outlier SNPs") to test if they were consistent with the PCAdapt significant outliers or those from the EA models. Manhattan plots for the OutFLANK and PCAdapt analyses are presented in Figure S2.

F_{ST} values from OutFLANK were correlated with the Mahalanobis Distance differentiation statistics from PCAdapt (estimate = 58.04, $t_{161094} = 446.20$, $P < 0.0001$, $R^2 = 0.55$, Figure S3a). The correlation between F_{ST} from OutFLANK and the Communality analysis from PCAdapt was even stronger (estimate = 1.30, $t_{161094} = 594.14$, $P < 0.0001$, $R^2 = 0.69$, Figure S3b). However, only two SNPs were represented in both the lists of the OutFLANK top 100 suggestive outliers and the 41 PCAdapt outliers (Figure 2a). These two methods show some repeatability but they also show an ability to identify outliers using different patterns of differentiation. For example, the RNF20 outlier was significant using OutFLANK because it showed much lower than expected levels of differentiation across populations but, was not identified by PCAdapt which only identifies differentiated loci.

LFMM and BayeScEnv Ecological Association results

The LFMM models for all four environmental variables identified 856 SNPs in total and 644 unique SNPs that were significantly linked to the environmental gradients. The overlap

between the significant SNPs for each environmental variable is shown in Figure 2b. A summary of the number of significant SNPs for each environmental variable is given in Table 2. The proportion of outlier SNPs physically linked to genes ranged from 33 to 42 per cent between environmental variables (Table 2). Manhattan plots of the log transformed *P* values are presented in Figure S4.

The BayeScEnv models for all four environmental variables identified 287 significant SNPs in total and 216 unique SNPs that were significantly linked to the environmental differentiation using the q-value method. Using the Posterior Error Probability (PEP) there were 143 significant SNPs in total and 114 unique SNPs. The overlap between the significant SNPs for each environmental variable is shown in Figure 2c. A summary of the number of significant SNPs for each environmental variable is given in Table 2. The proportion of significant SNPs physically linked to genes ranged from 33 to 40 percent for q-values and 38 to 42 percent for PEP (Table 2). Manhattan plots of log transformed q-values are presented in Figure S5.

Overall, there was not a strong overlap between the significant SNPs from the LFMM and BayeScEnv methods with only three shared SNPs (Figure 2a). However, there was an overlap between significant SNPs from correlated environmental variables (Figure 2b-c). The mean and SD of H_E did not fluctuate much between populations or with summer maximum temperature (Figure 3a). However, mean H_E was significantly negatively related to the estimated year of colonisation (estimate = -0.0002, t_9 = -5.211, $P < 0.001$, $R^2 = 0.75$, colonization dates from Andrew & Griffith (2016) Chapter Two). The mean H_E also showed a very strong negative relationship with the SD of H_E (estimate = -1.277, t_9 = -23.05, $P < 0.001$, $R^2 = 0.98$).

Heat shock Proteins

For a visual illustration of the changes in SNP frequency at a few target regions of the house sparrow genome we used the heat shock proteins. There are 25 genes on the house sparrow reference genome that are analogues to heat shock proteins (Table S3). Of these 19 had physically linked SNPs (within ca. 20 kbps) in our final dataset. Two pairs of genes were linked to the same SNP. The SNP linked to the HSPD1 and hspe1 genes on the house sparrow genome, was significantly associated with the Vapour Pressure Deficit environmental variable in the LFMM analysis. The allele frequencies of this SNP and one other linked SNPs were significantly correlated with summer maximum temperature using linear models (Table S3 and Figure 3b).

GO analysis

A summary of the significant outlier SNPs from the OutFLANK (top 100 suggestive SNPs), PCAdapt, LFMM and BayeScEnv methods are listed in Table S1. A full list of the physically linked genes to these SNPS, that are used in GO term analyses is given in Table S2. A GO analysis was run using PANTHER to test if the same GO terms were overrepresented for the OutFLANK and PCAdapt outliers. Both groups of outliers had no significant GO terms (Table S4).

For the LFMM outliers no significant GO terms were identified for individual environmental variables (data not presented) but for genes linked to all four variables combined there were some significant terms (Table S4). Similarly, for the BayeScEnv linked genes (using q-value significant SNPs), no significant GO terms were identified for individual environmental variables (data not presented) but for all four variables combined there were some significant terms (Table S4). Most of the significant GO terms appear to relate to cell metabolism and not directly to any of the traits we predicted to be involved in local adaptation in the Australian house sparrows, such as life history, morphology and thermal tolerance.

Discussion

All four genome scan methods identified significant putative SNPs that are potentially linked to local adaptation in introduced Australian house sparrow populations. However, between methods the number of outlier SNPs varied and the number of SNPs that were significant across multiple methods was limited (Figure 2a). Interestingly the strongest overlap between methods was between the top 100 suggestive SNPs from OutFLANK and the BayeScEnv outliers. This result shows that some of the loci with the highest F_{ST} values had differentiation in allele frequencies that also covaried with climatic differentiation. There is a large contrast between these four methodologies and as a result they appear to identify outliers using different signatures of selection. Although there are always going to be some false positives when using genome scan methods (Francois *et al.* 2016), we do find some supporting evidence that there could be some true signals of local adaptation in these introduced populations. This adaptation has occurred relatively rapidly over periods ranging from 26 to 75 generations since the establishment of sampled populations across the invasive range in Australia (generation time in the species being around two years [Jensen *et al.* 2013]).

The OutFLANK PD method, identified only one outlier that was significant because of extremely low levels of differentiation due to nearly all individuals being heterozygous for this SNP. This outlier was inside an exon of a gene analogous to RNF20 which has been linked to the epigenetic regulation of gene expression and chromatin structure (Nakamura *et al.* 2011). Four other Ring Finger Protein's were also linked to outliers (RNF10, RNF123, RNF13 and RNF41, there are 101 RNF genes on reference genome and they are 1.23 times over represented in the list of outliers) as well as other proteins that have functions relating to epigenetic gene regulation (e.g. TRMR13 and NR2C1). The next 99 suggestive SNPs had relatively high F_{ST} (mean ± SD = 0.344 ± 0.080) but were not found to be significantly different from the background differentiation across the genome by OutFLANK (mean F_{ST} ±

SD for all SNPs = 0.060 ± 0.048). The other PD method, PCAdapt found 41 outliers but only 2 overlapped with the suggestive OutFLANK loci.

The LFMM method identified the largest number of significant loci and also found the largest number of significant GO terms (Table S4). The BayeScEnv method found the highest proportion of SNPs linked to genes and found a number of consistent outliers with the PD methods (Figure 2a). Accordingly, 9 of the PCAdapt loci, and 16 of the suggestive OutFLANK loci overlapped with the BayeScEnv outliers. The BayeScEnv loci were more consistent with the PD methods than the second EA method (LFMM had only 3 shared loci with BayeScEnv). The BayeScEnv Outliers over the LFMM outliers, also showed stronger linear relationships with the raw allele frequencies and environmental variables when compared to random subsets of loci (see Figure S6 and S7). Across the BayeScEnv models for the four climatic variables the strongest liner relationships were between the climatic variables and the allele frequencies for the summer maximum and VPD outliers. These two variables are highly correlated ($t_9 = 8.98$, $P < 0.001$, $R^2 = 0.90$) and a majority of the outlier SNPs overlapped for both variables (Figure 2c). The strong relationships between clinal variation in summer maximum and VPD (a metric of aridity) could indicate that extreme high temperatures are a strong selective force in the Australian context. the lack of any relationship for the LFMM outliers is potentially explained by the method identifying significant associations between SNPs frequencies and climatic variation after accounting for genetic population structure, which is not corrected for in these linear models that were used to visualise changes in SNP frequency (Figure S6).

At present, we have too many loci to describe in depth their potential adaptive function relating to the house sparrow. We have identified some key genes of interest (Table 3) and significant GO terms (Table S4) that could be the target of future research that collects phenotypic and gene expression data. The genes listed in Table 3 were physically linked to our outlier loci and were significant loci in other studies on local adaptation and functional

genetics. Out of all the genes linked to our outlier loci the family of genes with the highest representation was the Solute Carrier Family (SLC) with 17 linked genes (there are 326 SLC genes on reference genome and are 1.29 times over represented in the outliers). These SLC genes are responsible for transporting charged and uncharged molecules across cell membranes and could be important to a number of traits (César-Razquin *et al.* 2015). Follow up research on hundreds of outlier loci is not viable in most cases, however, we would suggest the prioritising of those such as the few genes relating to the heat shock proteins that were linked to an outlier SNP here (Table 3 and Figure 3b). These genes would be the best candidates for future research aimed at explaining phenotypic variation in thermal tolerance. However the complex interactions between genes and environment plus the interactions between genes regulating polygenic traits has made it difficult to do follow up research on outlier loci (Haasl & Payseur 2016; Jensen *et al.* 2016)

Introduced populations are influenced by several factors that will cause changes in allele frequencies due to genetic drift, occurring simultaneously with any selection (Shultz *et al.* 2016). Introduced species are potentially difficult targets for studying local adaptation due to their complex demographic histories resulting in high levels of neutral genetic drift between invasive populations (Estoup & Guillemaud 2010; Cristescu 2015). Despite the fact that genetic drift is expected to affect the whole genome consistently (Messer *et al.* 2016), when we screen thousands of loci it is difficult to verify if an outlier SNP is a result of selection or extremely unlikely patterns of genetic drift. One way to validate if a SNP is linked to adaptation is to identify if the function of any linked genes is consistent with any predicted or observed patterns of local adaptation. However, some SNPs may not be linked to annotated or protein coding regions but still may be important to selection because they are linked to variants of promoter regions (Grubert *et al.* 2015) and non-coding RNA (Mercer *et al.* 2009) loci that regulate the gene expression patterns that drives phenotypic variation. As we develop our understanding of the complexity of trait regulation and use that information to

identify important: promoter regions, epigenetic markers and non-coding RNA on the house sparrow reference genome our interpretation of outlier loci should improve. the outlier loci found here may provide clear links to traits that are important to adapting to climatic variation in the future once we annotate the loci regulating gene expression in house sparrows.

Many molecular ecology studies that use genomic SNP data have used RADseq and genotyping by sequencing methods to identify SNP mutations (Puritz *et al.* 2014). However, many questions have been asked about the ability of these methods to sample evenly across the genome (Lowry *et al.* 2017). If only a few regions of the genome are influenced by selection then RADseq approaches will have a high probability of missing important results (Tiffin & Ross-Ibarra 2014). The house sparrow is one of the only invasive species to have a SNP array specifically developed for it, and this genotyping method overcomes some of the biggest weaknesses of the RADseq methods. The house sparrow 200K SNP array was developed in Norway using populations from Scandinavia (Lundregan *et al.* in prep). For the first time the SNP array has been successfully used on populations outside of that region (although the founders of the Australian population came from Western Europe). This success highlights the potential of the SNP array to be a useful tool for studying evolution at a global scale in this species. Previously a review on the study of invasion genetics in the house sparrow by Liebl et al (2015) identified that the future development of research on this model species relied on asking questions on a global scale. The consistence of the SNP array allows for the same informative loci to be genotyped consistently across most of the genome with minimal missing data which allows for questions to be asked coherently at a global scale. Important traits and loci identified using the house sparrow model system will be valuable to our understanding of their biology and the biology of other related systems that will benefit the conservation of diversity and evolutionary potential within species.

Acknowledgements: For funding support: SCA was supported by Macquarie University Research Excellence Scholarships (no. 2013077). SCG was supported by an Australian Research Council Future Fellowship (FT130101253). SNP array genotyping was conducted at CIGENE Norway. We would like to thank the large number of people who helped us with access to urban catch sites for sampling the human commensal house sparrow. We thank, Monica Awasthy, Amanda D. Griffith and Elizabeth L. Sheldon for their participation with field work. We would also like to thank Lee Ann Rollins for their assistance with lab work.

References

Anderson TR (2006) Biology of the Ubiquitous House Sparrow: from Genes to Populations. In: *University Press New York.*, pp. 31–56. Oxford University Press,Inc, New York.

Andrew SC, Griffith SC (2016) Inaccuracies in the history of a well-known introduction: a case study of the Australian House Sparrow (*Passer domesticus*). *Avian Research*, **7**, 9.

Arthur AL, Weeks AR, Sgrò CM (2008) Investigating latitudinal clines for life history and stress resistance traits in *Drosophila simulans* from eastern Australia. *Journal of evolutionary biology*, **21**, 1470–9.

Ashburner M, Ball CA, Blake JA *et al.* (2000) Gene Ontology: Tool for The Unification of Biology. *Nature Genetics*, **25**, 25–29.

Ashton KG (2002) Patterns of within-species body size variation of birds: Strong evidence for Bergmann's rule. *Global Ecology and Biogeography*, **11**, 505–523.

Barson NJ, Aykanat T, Hindar K *et al.* (2015) Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature*, **528**, 405–408.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B*, **263**, 1619–1626.

Benestan L, Quinn BK, Maaroufi H *et al.* (2016) Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology*, **25**, 5073–5092.

Bentley BP, Haas BJ, Tedeschi JN, Berry O (2017) Loggerhead sea turtle embryos (*Caretta caretta*) regulate expression of stress response and developmental genes when exposed to a biologically realistic heat stress. *Molecular Ecology*, **26**, 2978–2992.

Calder AW (1964) Gaseous Metabolism and Water Relations of the Zebra Finch, *Taeniopygia castanotis*. *Physiological Zoology*, **37**, 400–413.

César-Razquin A, Snijder B, Frappier-Brinton T *et al.* (2015) A Call for Systematic Research on Solute Carriers. *Cell*, **162**, 478–487.

Chen H (2016) VennDiagram: Generate High-Resolution Venn and Euler Plots. R package version 1.6.17. https://CRAN.R-project.org/package=VennDiagram.

Coccurello R, D'Amato FR, Moles A (2009) Chronic social stress, hedonism and vulnerability to obesity: Lessons from Rodents. *Neuroscience and Biobehavioral Reviews*, **33**, 537–550.

Colautti RI, Lau JA (2015) Contemporary evolution during invasion: evidence for differentiation, natural selection, and local adaptation. *Molecular Ecology*, **24**, 1999–2017.

Cristescu ME (2015) Genetic reconstructions of invasion history. *Molecular Ecology*, **24**, 2212–2225.

Dabney A, Storey JD, Warnes GR (2010) qvalue: Q-value estimation for false discovery rate control. *R package version 1.0*.

Doyle JM, Katzner TE, Roemer GW *et al.* (2016) Genetic structure and viability selection in the golden eagle (*Aquila chrysaetos*), a vagile raptor with a Holarctic distribution.

Conservation Genetics, **17**, 13071322.

Duforet-Frebourg N, Luu K, Laval G, Bazin E, Blum MGB (2016) Detecting genomic signatures of natural selection with principal component analysis: application to the 1000 Genomes data. *Molecular Biology and Evolution*, **33**, 1082–1093.

Duursma RA (2015) Plantecophys - An R package for analysing and modelling leaf gas exchange data. *PLoS ONE*, **10**, e0143346.

Elgvin TO, Trier CN, Tørresen OK *et al.* (2017) The genomic mosaicism of hybrid speciation. *Science Advances*, **3**, e1602996.

Estoup A, Guillemaud T (2010) Reconstructing routes of invasion using genetic data: Why, how and so what? *Molecular Ecology*, **19**, 4113–4130.

Feder ME, Hofmann GE (1999) Heat shock proteins, molecular chaperones, and the stress response: evolutionary and ecological physiology. *Annual Review of Physiology*, **61**, 243–282.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A bayesian perspective. *Genetics*, **180**, 977–993.

Francois O, Martins H, Caye K, Schoville S (2016) Controlling false discoveries in genome scans for selection. *Molecular Ecology*, **25**, 454–469.

Frichot E, Schoville SD, Bouchard G, François O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, **30**, 1687–1699.

Gautier M (2015) Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics*, **201**, 1555–1579.

Grubert F, Zaugg JB, Kasowski M *et al.* (2015) Genetic control of chromatin states in humans

involves local and distal chromosomal interactions. *Cell*, **162**, 1051–1065.

Grueber CE, Wallis GP, Jamieson IG (2013) Genetic drift outweighs natural selection at toll-like receptor (TLR) immunity loci in a re-introduced population of a threatened species. *Molecular Ecology*, **22**, 4470–4482.

Haasl RJ, Payseur BA (2016) Fifteen years of genomewide scans for selection: Trends, lessons and unaddressed genetic sources of complication. *Molecular Ecology*, **25**, 5–23.

Hartmann S a, Schaefer HM, Segelbacher G (2014) Genetic depletion at adaptive but not neutral loci in an endangered bird species. *Molecular Ecology*, **23**, 5712–5725.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hoban S, Kelley JL, Lotterhos KE *et al.* (2016) Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *The American Naturalist*, **188**, 379–397.

Hoffmann AA, Blacket MJ, McKechnie SW *et al.* (2012) A proline repeat polymorphism of the Frost gene of *Drosophila melanogaster* showing clinal variation but not associated with cold resistance. *Insect molecular biology*, **21**, 437–45.

Hoffmann AA, Sorensen JG, Loeschcke V (2003) Adaptation of Drosophila to temperature extremes: bringing together quantitative and molecular approaches. *Journal of Thermal Biology*, **28**, 175–216.

Jensen JD, Foll M, Bernatchez L (2016) The past, present and future of genomic scans for selection. *Molecular Ecology*, **25**, 1–4.

Jensen H, Moe R, Hagen IJ *et al.* (2013) Genetic variation and structure of house sparrow populations: is there an island effect? *Molecular Ecology*, **22**, 1792–805.

Jombart T (2008) Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, **24**, 1403–1405.

Jombart T, Devillard S, Balloux F *et al.* (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*, **11**, 94.

Jones HG (1992) *Plants and Microclimate: A Quantitative Approach to Environmental Plant Physiology*. Cambridge University Press, Cambridge.

Jones DA, Wang W, Fawcett R (2009) High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal*, **58**, 233–248.

Kim E-S, Elbeltagy AR, Aboul-Naga AM *et al.* (2015) Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. *Heredity*, **116**, 1–10.

Kuo LE, Kitlinska JB, Tilan JU *et al.* (2007) Neuropeptide Y acts directly in the periphery on fat tissue and mediates stress-induced obesity and metabolic syndrome. *Nature Medicine*, **13**, 803–811.

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, **23**, 2178–2192.

Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, **24**, 1031–1046.

Lowry DB, Hoban S, Kelley JL *et al.* (2017) Breaking RAD: An evaluation of the utility of restriction site associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, **17**, 142–152.

Lundregan S, Jensen H, Hagen I, Husby A, Gohli J Genetic architecture of bill morphology in

a free-living house sparrow metapopulation. *in prep*.

Luu K, Bazin E, Blum MGB (2017) pcadapt: An R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, **17**, 67–77.

Martin LB, Kidd L, Liebl AL, Coon CAC (2011) Captivity induces hyper-inflammation in the house sparrow (*Passer domesticus*). *Journal of Experimental Biology*, **214**, 2579–2585.

Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. *Nature Reviews Genetics*, **10**, 155–159.

Messer PW, Ellner SP, Hairston NG (2016) Can Population Genetics Adapt to Rapid Evolution? *Trends in Genetics*, **32**, 408–418.

Mi H, Muruganujan A, Casagrande JT, Thomas PD (2013) Large-scale gene function analysis with the PANTHER classification system. *Nature protocols*, **8**, 1551–1566.

Mi H, Poudel S, Muruganujan A, Casagrande JT, Thomas PD (2016) PANTHER version 10: Expanded protein families and functions, and analysis tools. *Nucleic Acids Research*, **44**, D336–D342.

De Mita S, Thuillet A-C, Gay L *et al.* (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, **22**, 1383–1399.

Montavon T, Soshnikova N (2014) Hox gene regulation and timing in embryogenesis. *Seminars in Cell and Developmental Biology*, **34**, 76–84.

Nakamura K, Kato A, Kobayashi J *et al.* (2011) Regulation of Homologous Recombination by RNF20-Dependent H2B Ubiquitination. *Molecular Cell*, **41**, 515–528.

Price TD, Qvarnström A, Irwin DE (2003) The role of phenotypic plasticity in driving genetic evolution. *Proceedings. Biological sciences / The Royal Society*, **270**, 1433–40.

Purcell S, Neale B, Todd-Brown K *et al.* (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**, 559–575.

Puritz JB, Matz M V, Toonen RJ *et al.* (2014) Demystifying the RAD fad. *Molecular ecology*, **23**, 5937–42.

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. , URL https://www.R-project.org.

Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015) A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, **24**, 4348–4370.

Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. *Nature reviews. Genetics*, **14**, 807–20.

Schweizer RM, Robinson J, Harrigan R *et al.* (2016) Targeted capture and resequencing of 1040 genes reveal environmentally driven functional variation in grey wolves. *Molecular Ecology*, **25**, 357–379.

Shultz AJ, Baker AJ, Hill GE, Nolan PM, Edwards S V. (2016) SNPs across time and space: population genomic signatures of founder events and epizootics in the House Finch (*Haemorhous mexicanus*). *Ecology and Evolution*, **6**, 7475–7489.

Sin YW, Annavi G, Dugdale HL *et al.* (2014) Pathogen burden, co-infection and major histocompatibility complex variability in the European badger (*Meles meles*). *Molecular Ecology*, **23**, 5072–5088.

Sørensen JG, Kristensen TN, Loeschcke V (2003) The evolutionary and ecological role of heat shock proteins. *Ecology Letters*, **6**, 1025–1037.

Thomas L, Kennington W, Evans R, Kendrick G, Stat M (2017) Restricted gene flow and local adaptation highlight the vulnerability of high-latitude reefs to rapid environmental change. *Global Change Biology*, **23**, 2197–2205.

Tiffin P, Ross-Ibarra J (2014) Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution*, **29**, 673–680.

de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.

de Villemereuil P, Gaggiotti OE (2015) A new FST-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*, **6**, 1248–1258.

Whitlock MC (2015) Modern approaches to local adaptation. *The American Naturalist*, **186**, S1–S4.

Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of FST. *The American Naturalist*, **186**, S24–S36.

Zappavigna V (2017) Special issue on HOX genes in development. *Journal of Developmental Biology*, **5**, 5.

Zueva KJ, Lumme J, Veselov AE *et al.* (2014) Footprints of directional selection in wild atlantic salmon populations: Evidence for parasite-driven evolution? *PLoS ONE*, **9**, e91672.

Tables

Table 1. Sample site information. Locality ID number (No.) and population cluster names are the same as used in figures. Information on sample site coordinates and environmental variables are included. Mean expected hetrozygosity (H_E) and the Standard deviation for H_E (SD) are also reported.

No.	Location	Population cluster	Latitude	Longitude	Summer max	Winter min	Season-ality	VPD	H_E	SD
1	Hobart	VIC & Tasmania	-42.936	147.351	20.9	3.3	3014	1.07	0.330	0.150
2	Melbourne	VIC & Tasmania	-37.789	144.915	26.2	5.5	3775	1.44	0.341	0.138
3	Mt Gambier	Mt Gambier	-37.857	140.848	24.2	5.4	3053	1.22	0.327	0.152
4	Adelaide	South Aus	-35.232	138.489	26.5	7.5	3507	1.51	0.329	0.151
5	Broken Hill	South Aus	-31.946	141.465	32.4	4.1	5512	2.46	0.334	0.145
6	Sydney	NSW	-33.62	150.816	28.9	3.7	4414	1.87	0.327	0.153
7	Armidale	NSW	-30.516	151.674	26.1	0	4742	1.47	0.326	0.152
8	Toowoomba	East QLD	-27.722	151.632	30	3.5	4719	2.06	0.322	0.155
9	Townsville	East QLD	-19.319	146.824	31.7	13.1	3098	1.94	0.314	0.161
10	Longreach	North west QLD	-23.436	144.255	37.3	7	5363	3.82	0.300	0.171
11	Mt Isa	North west QLD	-20.73	139.503	37.1	8.3	4936	3.95	0.305	0.168

Table 2. Proportion of outlier SNPs linked to Protein coding regions. The four analysis methods are included in this table. The LFMM and BayeScEnv analyses used four environmental variables and BayeScEnv had two statistics for significance (q-value and PEP). Less than half of the significant SNPs were physically linked (within 20 Kbp) to known genes but many SNPs were also linked to multiple genes. A summary of all unique SNPs is also given, out of 1427 outliers across all analyses there were 971 unique SNPs.

Analysis	Total SNPs	No. SNPs linked to genes	No. of genes linked to SNPs	Percentage of SNPs linked to genes	Percentage of SNPs within genes	Percentage of SNPs within exons
All SNPs	971	373	575	38.29	16.29	1.93
OutFLANK	100	39	57	39.00	15.00	2.00
PCAdapt	41	13	21	31.71	12.20	0.00
LFMM						
Summer max	137	58	92	42.34	18.25	0.00
Winter min	261	109	171	41.76	17.24	2.30
Seasonality	223	73	125	32.74	13.00	0.45
VPD	235	96	154	40.85	17.87	1.28
BayeScEnv						
Summer max: q-value	51	17	24	33.33	17.65	3.92
PEP	24	9	13	37.50	16.67	4.17
Winter min: q-value	107	38	64	35.51	13.08	1.87
PEP	56	21	34	37.50	16.07	1.79
Seasonality: q-value	40	16	28	40.00	17.50	5.00
PEP	19	8	16	42.11	26.32	5.26
VPD: q-value	89	36	51	40.45	20.22	4.49
PEP	44	18	28	40.91	15.91	2.27

Table 3. Previously identified candidate genes linked to selection. Our outlier loci were physically linked to 575 genes (Table S2). Due to FDR and the fact that some outlier SNPs were linked to multiple genes not all of these genes are likely to be targets of selection. Some of the linked genes we identify here have been described in similar studies on local adaptation and fitness traits in vertebrates so are the most likely candidates for local adaptation in sparrows.

Gene name	Analysis that found outlier	Function	Previously identified gene	Reference	Note
HSPD1, Hspe1	LFMM VPD	Heat Shock Proteins	Multiple HSP genes	(Bentley *et al.* 2017)	Up regulation of HSP genes in stressed turtle embryos
GNA11	BayeScEnv winter min	Thermal-tolerance	GNAI3	(Kim *et al.* 2015)	Found to be associated with thermal-tolerance in goats and sheep. Both genes are G Proteins.
APOD	LFMM VPD and Summer max	cholesterol binding and control	APOB	(Schweizer *et al.* 2016)	Associated with local adaptation to climatic variation in broadly distributed grey wolves
IL22RA2	BayeScEnv winter min	Inflammatory response	IL1R1	(Kim *et al.* 2015)	Linked to nervous and autoimmune response for goats and sheep across climates.
TLR4	BayeScEnv winter min, PCAdapt, OutFLANK	Inflammatory response	TLR4	(Martin *et al.* 2011), (Grueber *et al.* 2013) and (Hartmann *et al.* 2014)	Expression increases in captive house sparrows, affects fitness in Stewart Island Robin and survival probability in Ecuadorian Brushfinch.
NPY2R	LFMM seasonality, BayeScEnv VPD	Feeding behaviour and body fat	NPY2R	(Kuo *et al.* 2007; Coccurello *et al.* 2009)	Linked to stress dependent fat acquisition in humans and rodents.
BMP4	LFMM VPD	Growth and development	BMP4	(Kim *et al.* 2015) and (Doyle *et al.* 2016)	Associated with adaptation in goats/sheep and golden eagles.
HOXA4, HOXA5, HOXA7, HOXA9	LFMM Winter min	Embryonic development	All HOX genes are linked to regulating development	(Montavon & Soshnikova 2014; Zappavigna 2017)	How HOX genes regulate embryonic development and limb formation has been studied broadly. 57 HOX genes on sparrow reference genome (3.05 times over represented)
HOXC6	LFMM seasonality	Embryonic development			
hoxd3, HOXD8	BayeScEnv seasonality	Embryonic development			

Figures

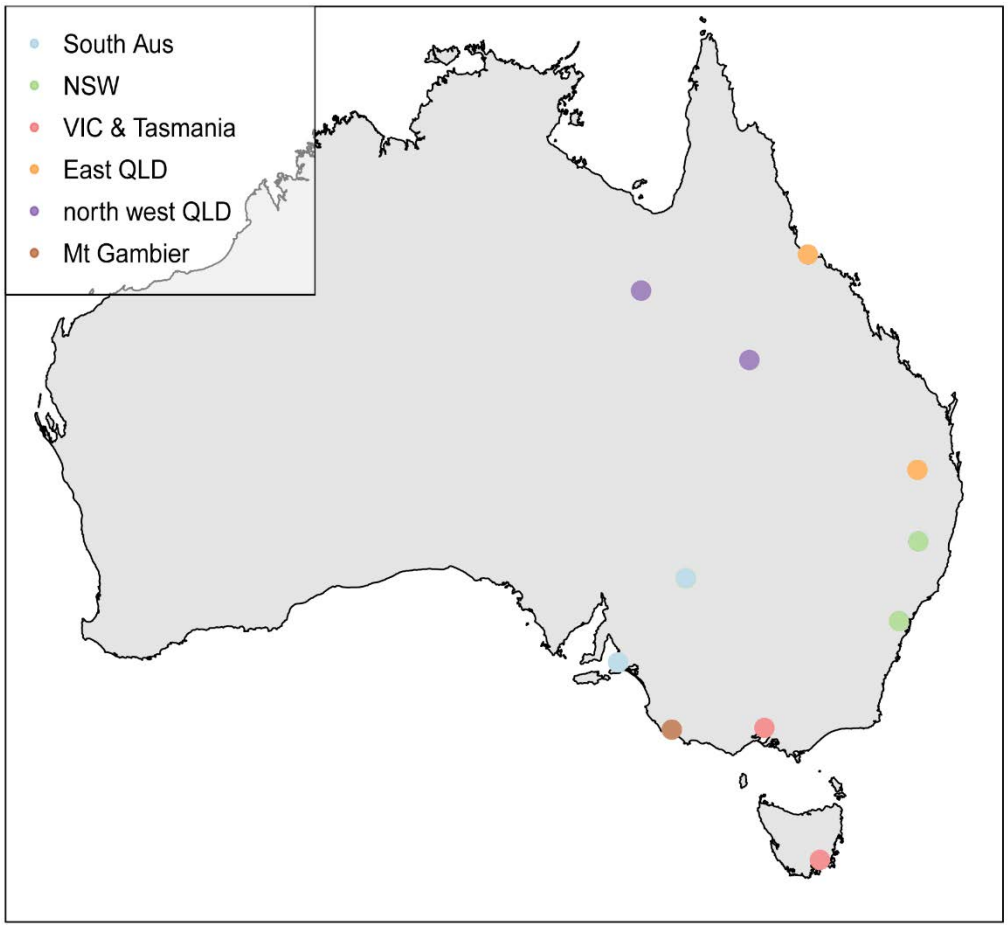


Figure 1. Map of sample sites across Australia. The points are colour coded based on the genetic populations defined in Figure S1.

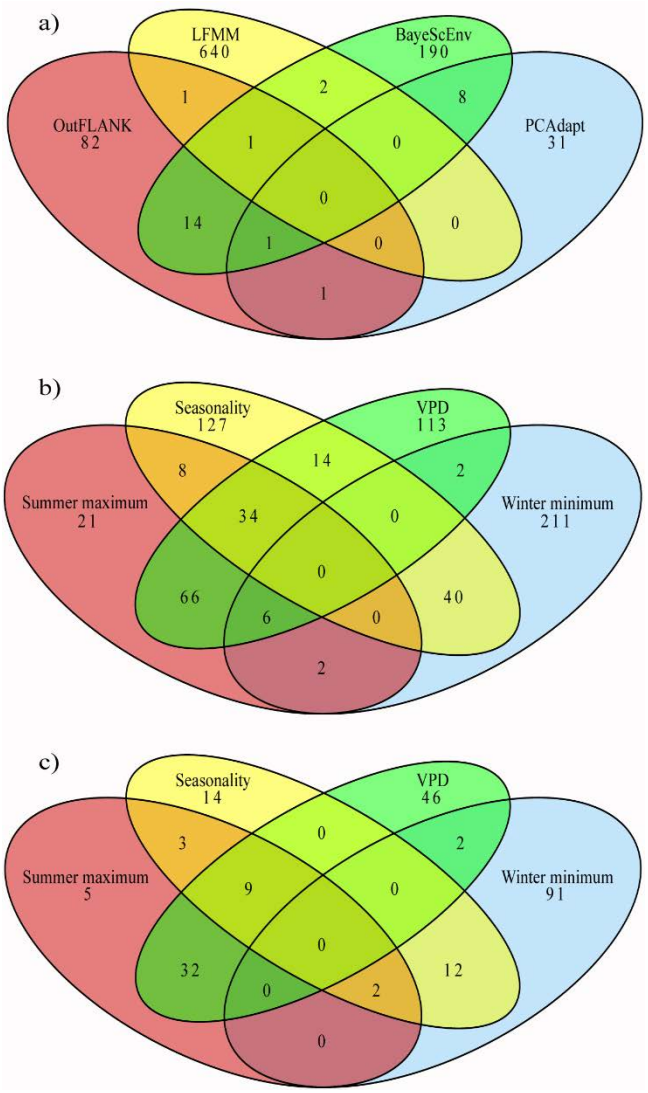


Figure 2. Venn diagrams of Outlier loci. Panel a) shows a comparison between the outliers of the OutFLANK, PCAdapt, LFMM and BayeScEnv methods. Here we used the top 100 loci with the smallest P values from OutFLANK and the 41 significant outliers from PCAdapt. For the two EA methods, LFMM and BayeScEnv, all duplicates of loci that were significant for multiple environmental variables were removed. Panel b) is a Venn diagram showing the overlap of significant loci identified for the four environmental variables using the LFMM method. Panel c) shows the overlap between the significant loci identified using the BayeScEnv method. For this Venn diagram significant loci were identified by the q-value (q-value < 0.05). The PEP is a more conservative statistic for significance from BayeScEnv, as a result there were more significant loci using the q-value than for PEP method, 287 and 143 respectively (totalled across the four environmental variables).

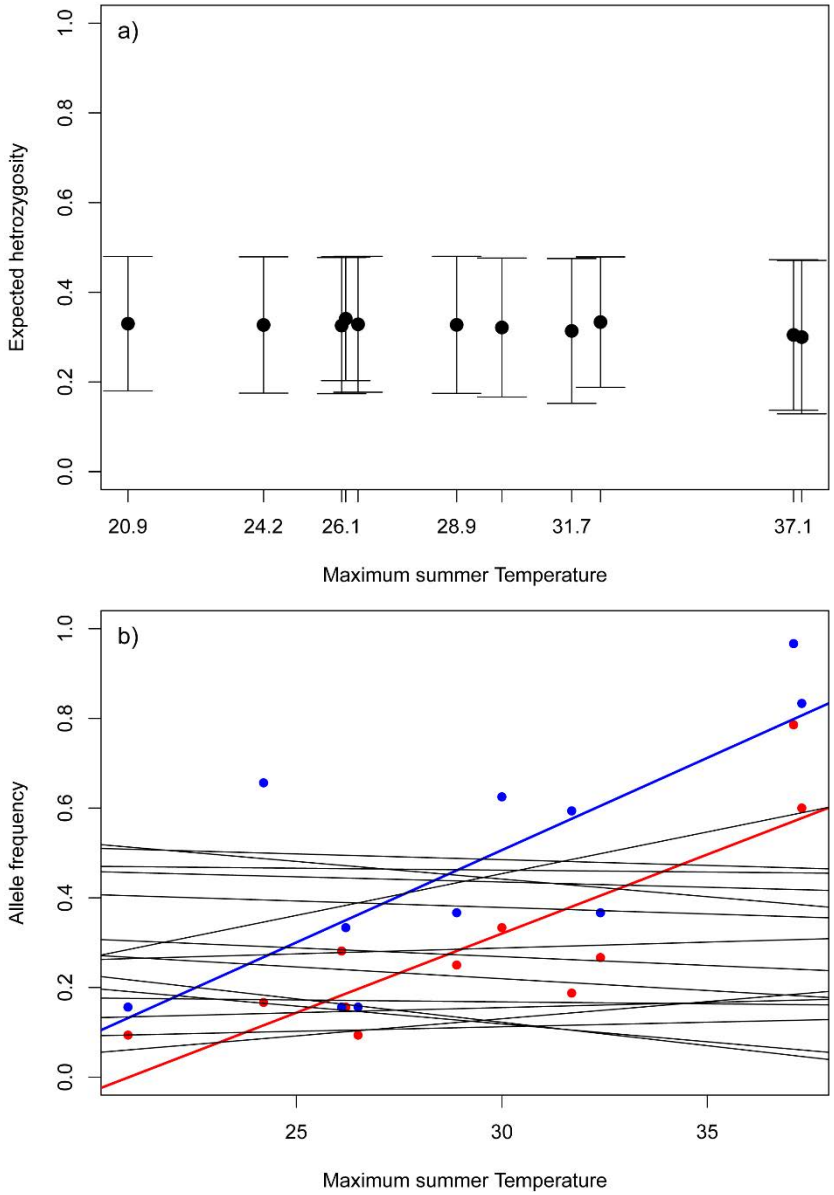


Figure 3. Genetic diversity and functional loci. Panel a) plots the mean H_E of different sampling populations against the mean maximum summer temperature. The error bars show the standard deviation. Panel b) plots the linear relationship for the allele frequencies of SNPs linked to heat shock proteins against the maximum summer temperature. Significant relationships have lines of best fit in colour and points for the sample sites are also plotted. Non-significant relationships are plotted as black lines of best fit (See full results in Table S3). The red line of best fit is for the SNP linked to HSPD1 and Hspe1, the blue line is for the SNP linked to two genes analogues to hsp30c.

Supplementary Material Chapter Six

Supplementary tables on Figshare (link: https://figshare.com/s/dbc245f85f4cc3cdd5b1)

Table S1. List of all significant SNPs from each analysis. See summary information in table 2.

Table S2. List of all linked genes to outlier SNPs used in GO analyses.

Table S3. Table of Heat Shock proteins in the house sparrow reference genome plotted in Figure 3b.

Table S4. Summary of GO term Analyses using PANTHER.

Supplementary Figures

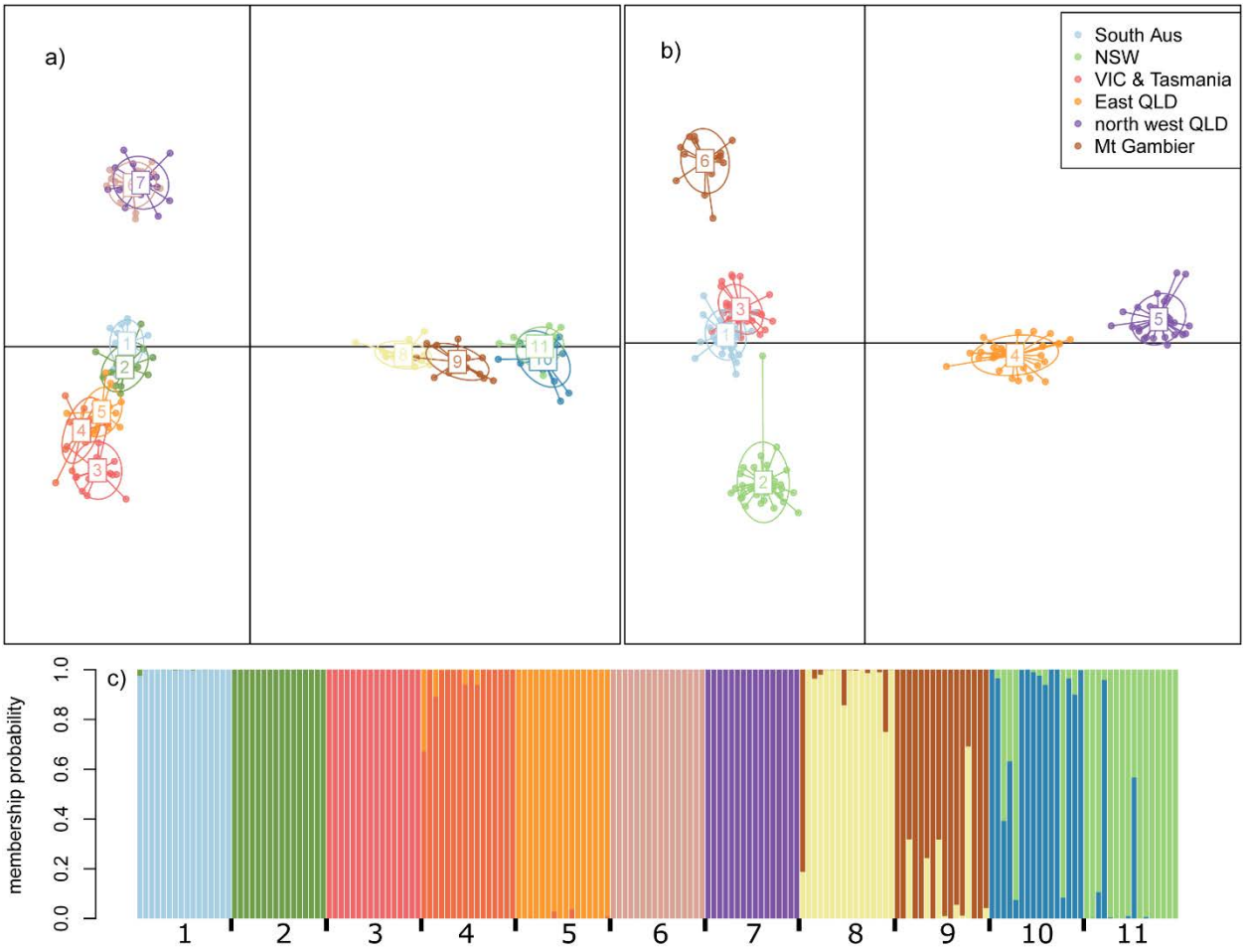


Figure S1. Population structure across sampling localities. Panel a) is a DAPC using the 11 sample site labels for the individuals. The colour coded and numbered clusters correspond to the following sample sites: 1 Hobart, 2 Melbourne, 3 Mt Gambier, 4 Adelaide, 5 Broken Hill, 6 Sydney, 7 Armidale, 8 Toowoomba, 9 Townsville, 10 Longreach, 11 Mt Isa. Panel b) is a scatter plot of a DAPC that used the cluster labels from the "find clusters" method to find the best number of meaningful clusters for the data. The "find clusters" method did not find a consistent best number of clusters using different subsets of the SNP loci. Based on the first DAPC in panel a) and previous population genetics studies of Australian house sparrows six appeared to be the best number of meaningful clusters for our data. Using a K of 6 for the best number of clusters, individuals from the same sample site were consistently grouped within a cluster and membership probability was 100% for individuals. The sample sites included in the six clusters were: South Aus (sites 4 & 5); NSW (sites 6 & 7); VIC & Tasmania (sites 1, 2); East QLD (sites 8 & 9); North west QLD (sites 10 &11) and Mt Gambier (site 3). Panel c) shows the membership probability of individuals to their sample site ID (11 sites, 16 per sample site, n = 176 columns) used in the first DAPC in panel a). Most individuals had 100% membership probability for their sample site.

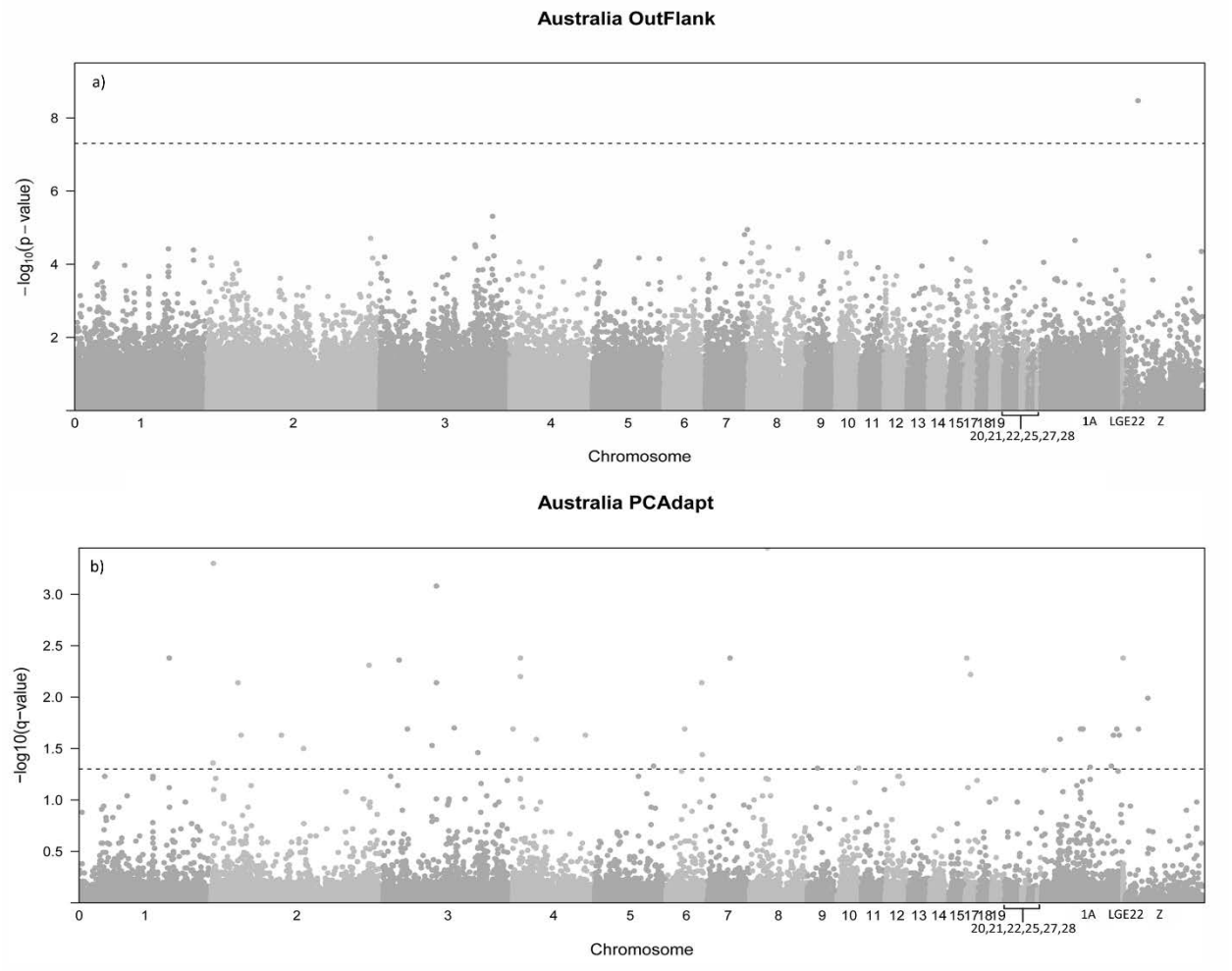


Figure S2. Manhattan plots for Population Differentiation methods. Panel a) shows the OutFLANK results and panel d) show PCAdapt result. Significant loci are plotted above the dashed line.

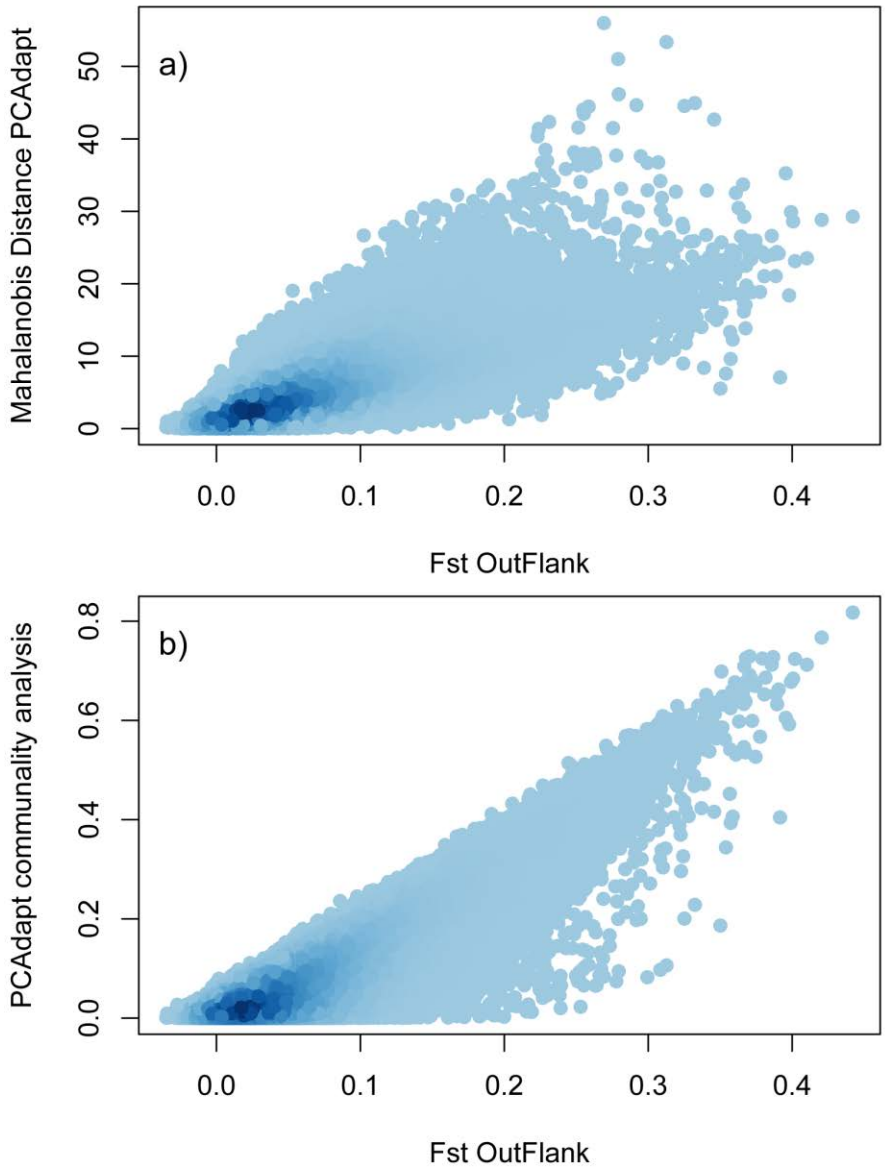


Figure S3. Comparison between population differentiation methods. Panel a) shows the correlation between PCAdapt Mahalanobis Distance and OutFLANK F_{ST} values from the analyses shown in Figure S2. Panel b) shows the correlation between the PCAdapt communality values (a PC version of F_{ST}) and OutFLANK F_{ST} values. Darker shading shows where there is a higher density of overlapping points.

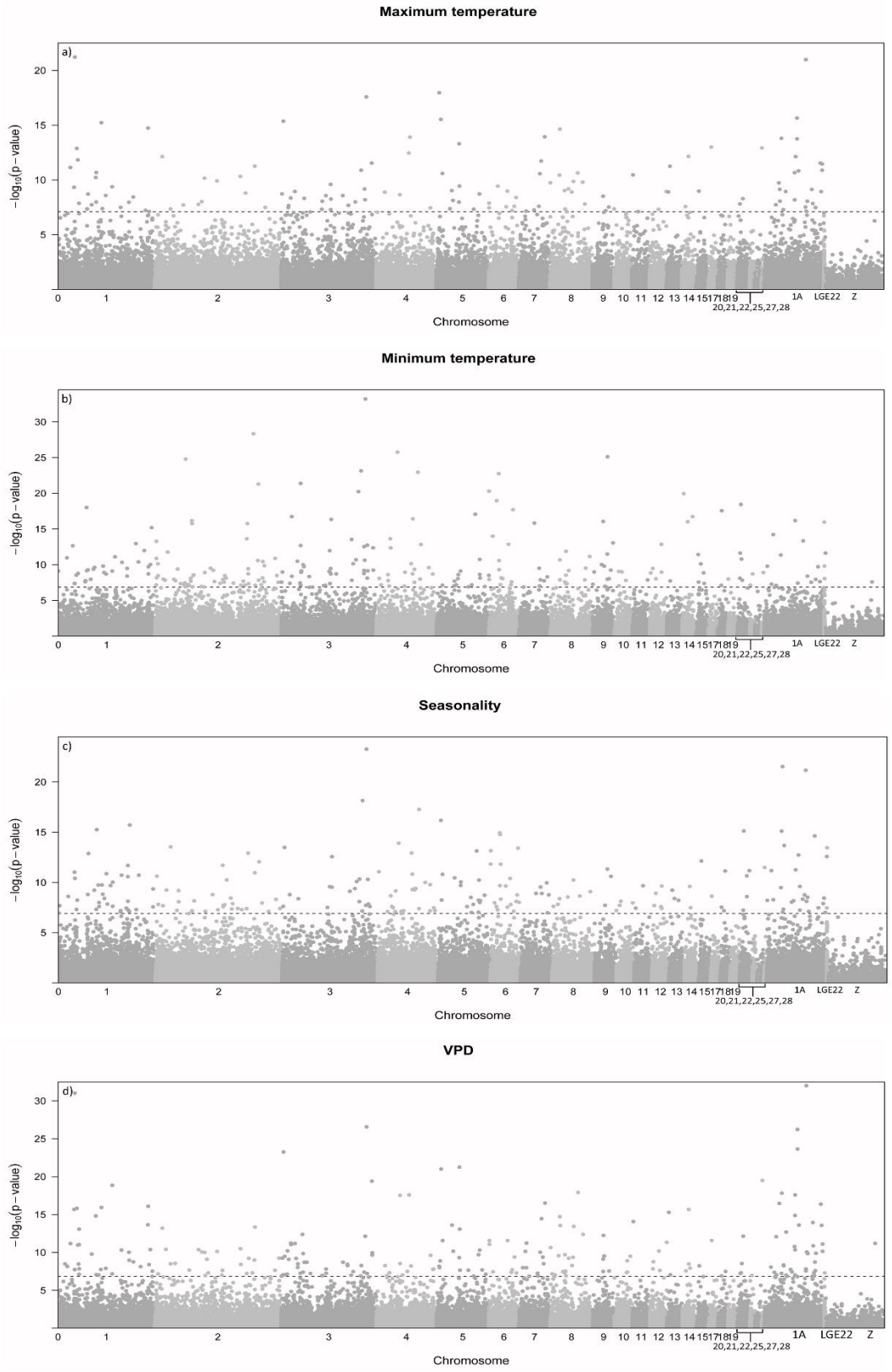


Figure S4. (see figure legend below)

Figure S4. LFMM Manhattan plot results. Panel a) is a for the model using the summer maximum temperature as a predictor. Panel b) has the Manhattan plot for the model using the winter minimum temperature. Panel c) is the Manhattan plot for the Seasonality predictor. Panel d) shows the Manhattan plot for the VPD model. Significant loci are plotted above the dashed line.

Figure S5. BayeScEnv Manhattan plots (See below). Panel a) is a Manhattan plot for the model using the summer maximum temperature as a predictor tor. Panel b) shows the Manhattan plot for the model using the winter minimum temperature. Panel c) is the Manhattan plot for the Seasonality predictor. Panel d) shows the Manhattan plot for the VPD model. Significant loci are plotted above the dashed line.

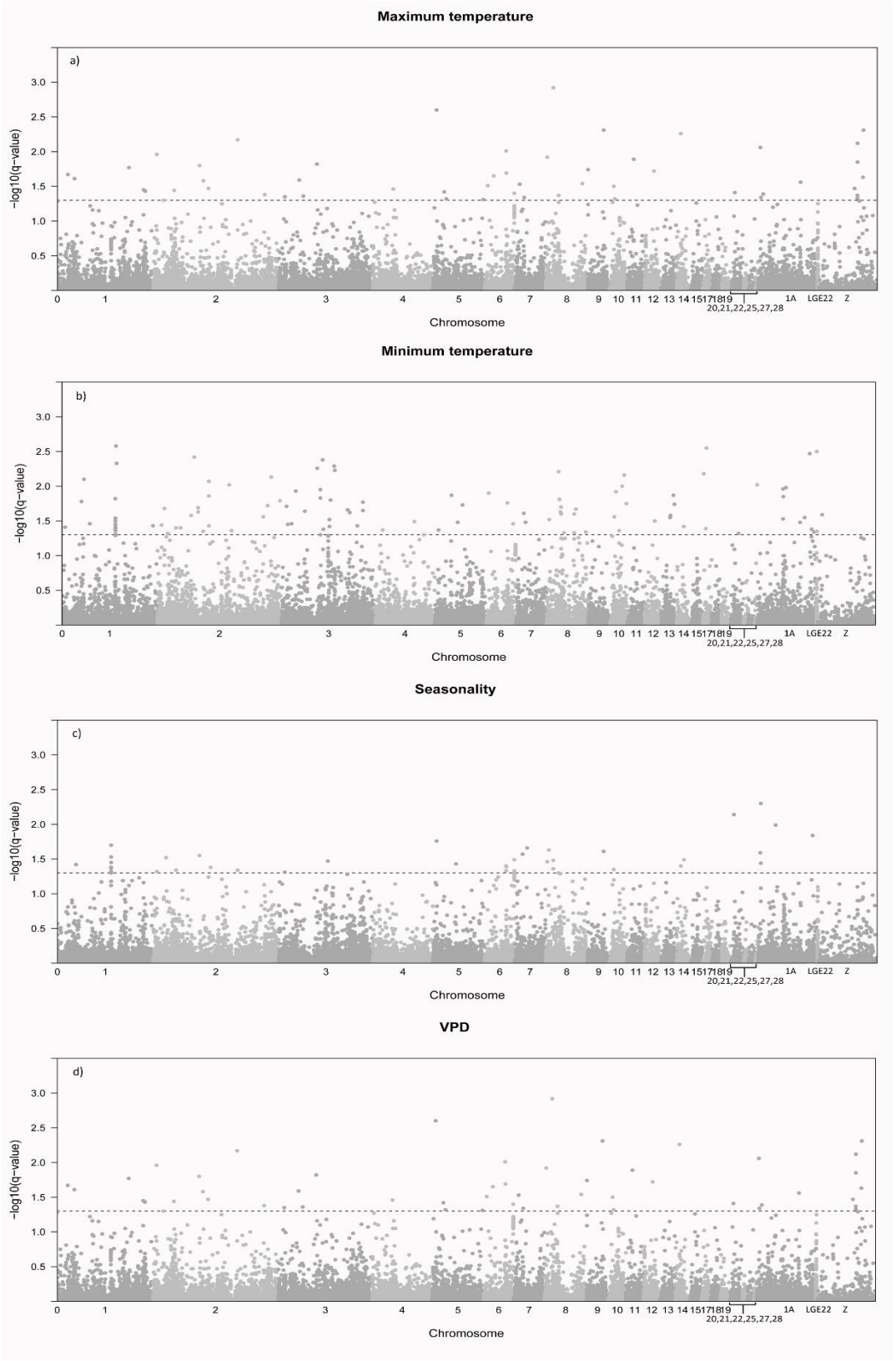


Figure S5. (See figure legend above)

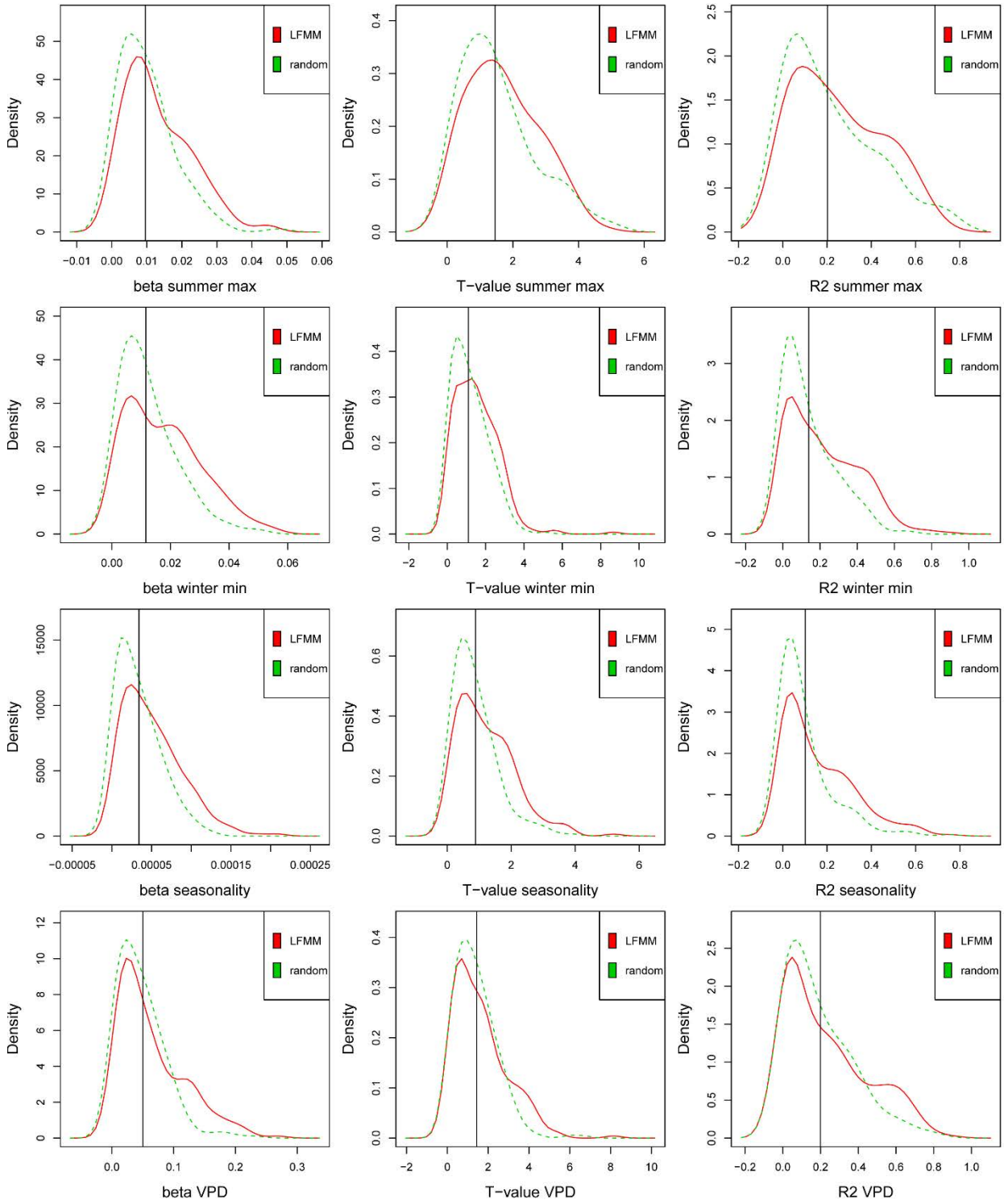


Figure S6. (see figure legend below)

Figure S6. Summaries of linear regressions of raw SNP frequencies against climatic variables. Each plot is the density distributions for the significant SNPs from the LFMM models and a random subset of SNPs (equal in number to the significant SNPs). The mean value for all loci is plotted as a vertical black line. The top row is for summer maximum temperature (LFMM = 137 SNPs), second row is for winter minimum (261 SNPs), third row is for seasonality (223 SNPs) and the forth row is for VPD (235 SNPs). The beta coefficients are plotted in column 1, the T-values in column 2 and the R^2 values are plotted in column 3. For these plots, we expect the beta coefficients to be higher than the random subset of SNPs. We also expect the T-values and the R^2 values to be much higher for the significant SNPs over the random subset. We do not see the expected pattern for beta coefficients but the T-values and R^2 Values have more high values than some of the random subsets. Beta coefficients and T-values are all converted to be positive for consistency because all SNPs are binary so the direction of the slope is not meaningful.

Figure S7. BayeScEnv density distributions plots (see below). These plots are density distributions of summary statistics from linear regressions between climatic variables and the frequency of the significant SNPs from the BayeScEnv models and a random subset of SNPs (equal in number to the significant SNPs). The mean value for all loci is plotted as a vertical black line. The top row is for summer maximum temperature (BayeScEnv = 51 SNPs), second row is for winter minimum (107 SNPs), third row is for seasonality (40 SNPs) and the forth is for VPD (89 SNPs). Beta coefficient plots are in column 1, T-value plots in column 2 and R^2 plots are in column 3. We have the same predictions as for the LFMM SNPs. For summer max, seasonality and VPD the beta coefficients were much higher than the random distribution and the overall mean. This pattern is also true for the T-values and R^2 for summer max and VPD but not as clear for seasonality. There is no clear expected pattern for winter minimum.

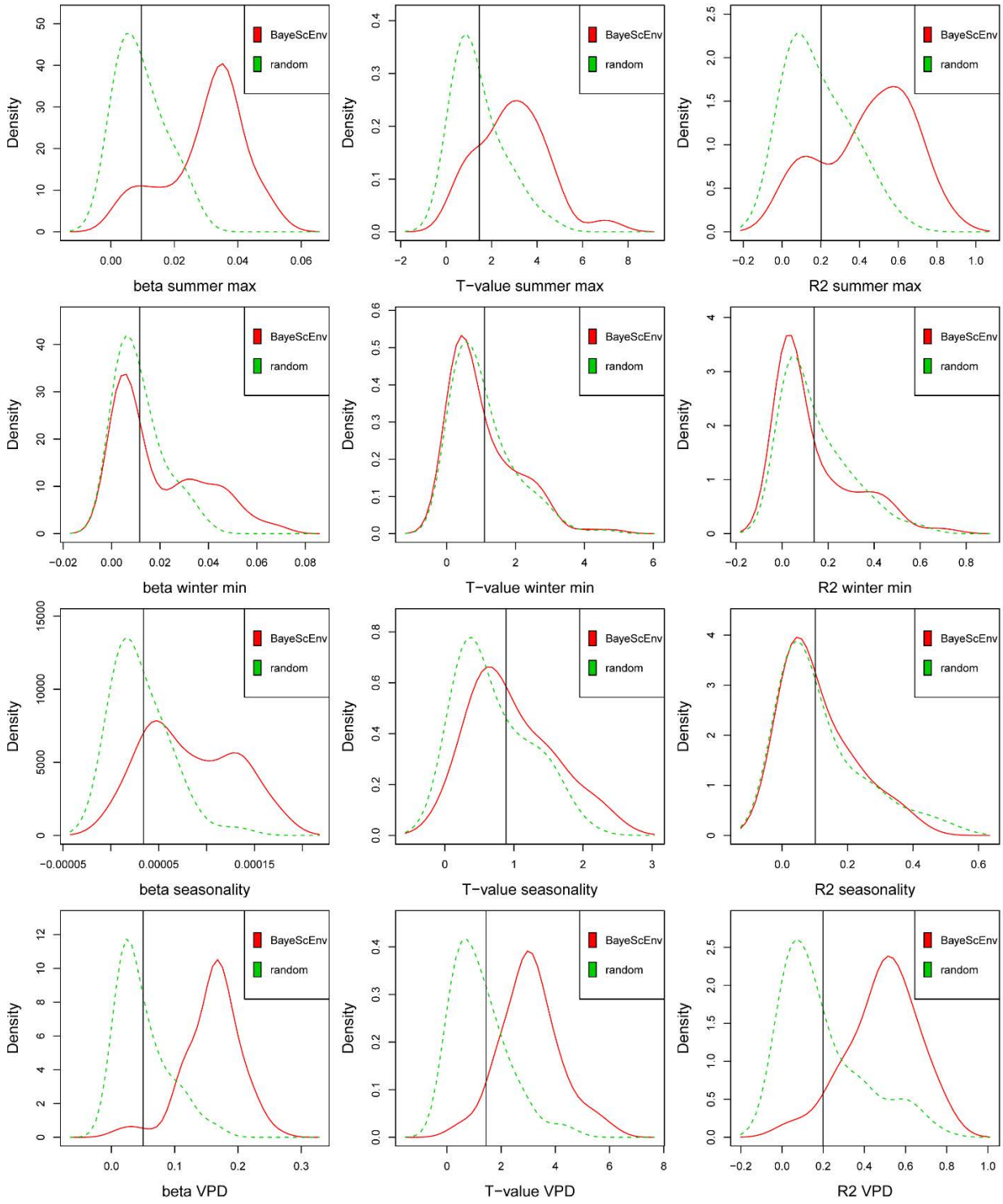


Figure S7. (see figure legend above)

Chapter Seven

Living in a heavy metal environment: Evidence for selection in the house sparrow to varied urban environments contaminated with trace elements

Chapter Seven Vignette

House sparrows live in a variety of climates as well as a variety of urban environments. Human activities create heterogeneity in the environment that house sparrows must respond to. A primary example of this is pollution that is toxic to humans as well as wildlife. In particular, heavy metals such as lead are known to be hazardous to the health of both humans and wildlife. Environmental contamination is also known to be a stressor to wildlife so is potentially a significant selective force for species living in highly contaminated environments. Lead is a naturally occurring trace element that has no biological function but due to human activity it now occurs in dangerously high concentrations in some environments. The negative effects of lead on health has been identified as a serious problem in heavily contaminated mining towns around Australia. The house sparrows' presence in these mining towns and other urban environments across Australia creates a 'natural' experiment for observing local adaptation to heavy metal contamination. This chapter aims to explore another aspect of local adaptation that is more specific to the urban environment this species specialises in due to it relationships with humans.

This manuscript is prepared for submission to Molecular Ecology, and is formatted as such.

Authors: Samuel C. Andrew[1]; Mark P. Taylor[1]; Sarah Lundregan[2]; Sigbjørn Lien[3]; Henrik Jensen[2]; and Simon C. Griffith[1]

[1]Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia

[2]Centre for Biodiversity Dynamics, Department of Biology, Norwegian University of Science and Technology, NO-7491 Trondheim, Norway.

[3]Centre for Integrative Genetics (CIGENE), Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway.

Key words: lead, avian, ecotoxicity, local adaptation, *Passer domesticus*

* To whom correspondence should be addressed

Samuel C. Andrew

Department of Biological Sciences, Macquarie University, New South Wales, 2109, AUSTRALIA

E-mail: samuel.andrew@students.mq.edu.au

Abstract

Metals and metalloids at elevated concentrations can be toxic to both humans and wildlife. In particular, lead exposure can act as a stressor to wildlife and cause negative effects on fitness. The ability of species to adapt to stress caused by the negative effects of exposure, should be beneficial for species living in contaminated environments. However, the mechanisms for responding adaptively to metal contamination are not fully understood in free-living organisms. The Australian house sparrow (*Passer domesticus*) provides an excellent opportunity to study adaptation to lead contamination because they have a commensal relationship with humans and are distributed broadly across Australian settlements including many long-term mining and smelting communities. To examine the potential for an evolutionary response to long-term metal exposure, we collected genomic SNP data using the house sparrow 200K SNP array, from 11 localities across the Australian distribution including two mining towns (Broken Hill and Mount Isa, that are two genetically independent populations) that have well characterised and high levels of lead contamination (and other heavy metals). We contrast these known contaminated locations with other lesser-contaminated environments. Using an ecological association genome scan method to identify genomic differentiation associated with estimates of lead contamination we identified 26 outlier loci across two tests. We were able to identify physical links between some of these outliers and a total of 13 genes on the house sparrow reference genome. The significant and suggestive candidate genes included genes relevant to lead exposure, such as two metal transporters that can transport metals including lead and zinc across cell membranes. Candidate genes such as these provide targets for future studies using common garden experiments to observe the phenotypic adaptations between populations with varied lead exposure.

Introduction

Environmental contaminants such as lead and other metals are dangerous for humans and wildlife when found in high concentrations (Ray *et al.* 2014; Peterson *et al.* 2017). The alarming effects of these pollutants on health has led to a large body of research aimed towards understanding how humans and other species respond to contaminants in their environment (Baccarelli & Bollati 2009; Pierron *et al.* 2011; Lattin *et al.* 2014; Varian-Ramos *et al.* 2014; González *et al.* 2017; Li *et al.* 2017). These pollutants are not always lethal and can have subtle negative effects on health. For example, the accumulation of contaminates inside individuals living in contaminated environments is known to be a stressor for animals (Romero & Wikelski 2001; Wikelski *et al.* 2002; Lattin *et al.* 2014). Stressors result in physiological responses that can have negative effects on fitness and reproduction (Romero 2004; Cyr & Romero 2007; Bonier *et al.* 2009). The negative effects of stress are known to be an important threatening process due to a loss of fitness that can cause population decline (Wikelski & Cooke 2006). Physiological responses to stressors can also alleviate stress and return the individual to equilibrium. Due to the negative effects of stress on fitness there should be positive section for any adaptations that reduce the impact of environmental contaminates on individuals living in polluted environments. It is still unknown if many species possess traits that can respond to alleviate the impacts environmental contamination caused by humans.

One of the most studied metal contaminants is lead (Pb) because of its toxicity and its known adverse impacts on human health (National Toxicology Program 2012; Lanphear 2015). Environmental lead in aerosols, dusts and soils inter alia are elevated in most global urban city environments due to the former massive emissions of lead from industrial emissions dominated by leaded petrol (Olszowy *et al.* 1995; Laidlaw & Taylor 2011; Mielke *et al.* 2011; Kristensen 2015; Ericson *et al.* 2016; Kristensen *et al.* 2017; Laidlaw *et al.* 2017; Rouillon *et al.* 2017). These industrial emissions and depositions remain present in the

environment in a highly bioavailable state (Mackay *et al.* 2013; Laidlaw *et al.* 2017) and consequently present a risk of harm to organisms living in urban environments. In Australia, there are a number of specific locations that have a protracted history of environmental lead emission from lead mining and smelting practices (Morrison 2003; Taylor *et al.* 2010, 2014a; Mackay *et al.* 2013; Dong *et al.* 2015; Kristensen & Taylor 2016). Environmental exposures in lead producing communities as well as those impacted by former leaded petrol depositions are typically via the ingestion of soils and dusts (Gulson *et al.* 2013, 2014). Unlike dust, soil metal concentrations have national guidelines values promulgated under the NEPM (2013). The most relevant NEPM (2013) soil lead value is the NEPM (2013) guideline of 300 mg/kg, which is formally associated with standard residential dwellings and is designed to ensure blood lead levels remain below 7.5 µg/dL. However, it is well-accepted that there is no safe lower threshold for blood lead exposure (Lanphear *et al.* 2005) and moreover, exposure to blood lead as low as 2 µg/dL is considered deleterious to human health (National Toxicology Program 2012). Therefore, we use the lowest upper acceptable limit (300 mg/kg for soil lead) provided in the NEPM (2013) as a benchmark for assessing the level of contamination that the studied sparrow populations are exposed to (Table 1).

Understanding how species adapt to metal contamination in urban environments and more broadly in the environment could be important to urban ecosystem management. The introduced house sparrow (*Passer domesticus*) in Australia provides a good model system to investigate evolutionary adaptation to polluted urban environments. House sparrows are an obligate commensal species with humans and they are typically constrained to the urban environment (Anderson 2006). Sparrows also have sedentary populations with relatively strong genetic population structure and low gene flow between broadly spaced human settlements across Australia (Chapter Five). This species also has a broad distribution that covers a range of urban environments and a number of mining communities in Australia. The invasion of the house sparrow to different mining communities provides an opportunity to

observe how the populations are responding today to the environmental stressor and whether these populations have acquired a capacity to cope with the stressor through the process of local adaptation over a relatively limited, and known number of generations. For example, sparrows have been present in the mining communities of Broken Hill (New South Wales) and Mount Isa (Queensland) for approximately 100 and 50 years respectively (Chapter Two, Andrew & Griffith 2016). Long term mining practises, and relatively high levels of contamination have potentially allowed for selection to take place over dozens of generations (50 and 25 in the two towns respectively [Jensen *et al.* 2013]) to reduce the deleterious effects of lead and/or heavy metal pollution on fitness. The genetic structuring that we have previously demonstrated and the independent origins of the Broken Hill and Mount Isa populations (Chapter Five) means that these two populations are to an extent genetically independent replicates in which genetic changes over the time since the populations were founded are likely to have occurred independently. As such, the study is based on the premise that any patterns of selection that are seen in these two populations and not in the other nine populations are likely to be associated with the high levels of pollution experienced by these two populations. Whilst there may be selection at other loci in either one of the populations (that we are not detecting through the approach used here), we will pick up the signal of selection that has occurred twice, once in each location.

There are many potential physiological mechanisms sparrows could employ to adaptively respond to lead contamination in their environment. Lead (Pb) is not essential to life and is a non-biodegradable element, meaning there are no efficient pathways of metabolization and elimination, thus it tends to accumulate in the different organs and tissues of individuals exposed to lead (Peakall & Burger 2003). The primary line of defence for species against lead would be through trying to reduce the unintentional uptake of lead into the body through external surfaces like the inner ear (Ding *et al.* 2014), the respiratory system (Ribeiro *et al.* 2014), or the alimentary canal (Madigosky *et al.* 1991). Another suggested trait

involved in adapting to lead contamination is increased exclusion through the accumulation of lead in the kidneys and other tissues to reduce the amount lead bioavailable in the body (Ribeiro *et al.* 2014). Lead is also known to be a stressor to the endoplasmic reticulum of cells, affecting protein production, however this effect can also be countered through different molecular pathways that remove misfolded proteins (Qian & Tiffany-Castiglioni 2003; Shinkai *et al.* 2010). Alternatively, sparrows could also be responding to lead by trying to avoid contamination through behavioural modifications that reduce exposure.

We have collected genomic data from 11 sparrow populations across Australia including the mining towns of Broken Hill and Mount Isa, using a 200K SNP array. This is an observational study, in which we aim to explore genetic differentiation across loci, and relate it to geographical variation in lead contamination across localities in Australia, and are particularly interested in the contrast between the heavily polluted towns of Broken Hill and Mount Isa, and the other nine populations. We predict that some of the loci that are significantly associated with lead contamination will be physically linked to genes on the house sparrow reference genome that are related to known traits associated with lead contamination. We will also discuss how candidate loci for adapting to lead pollution can be used in future research to fully understand the mechanisms behind this form of local adaptation.

Methods

Sampling and genetic data collection

Adult house sparrows were sampled in urban areas from 11 towns/cities across Australia and 16 male individuals from each of our sampling localities were used for genotyping in a previous study (see details in Chapter Six). The sampling was carried out under the Animal

Research Authority of the Animal Ethics Committee at Macquarie University (ARA 2014/248).

The individuals were genotyped on the house sparrow Affymetrix 200K SNP array (Lundregan *et al.* in prep). After quality filtering (See Chapter Six) there were 162 299 SNPs used in further analyses and 176 individuals. SNP filtering for minor alleles (removing loci with less than 5%) and missing data (loci with more than 91% missing data) was done using the program PLINK (Purcell *et al.* 2007). Data conversion between PED and BayeScEnv formats was performed using PGDSpider (Lischer & Excoffier 2012).

Environmental lead estimations

Soil lead concentrations are characteristically heterogenous, even over small areas (Rouillon *et al.* 2017). However, mean soil lead levels in most urban areas around Australia are typically higher than background levels (ca. less than 30 mg/kg (Callender 2014)). A report by Olszowy *et al.* (1995) illustrates this pattern clearly by describing mean lead levels in environments with different levels of urbanisation across the main states of Australia. Empirical data has shown that Mount Isa and Broken Hill have much higher than typical average concentrations of lead and other heavy metals, and are elevated above most of the state capital cities of Australia (Taylor *et al.* 2010, 2014a; b; Mackay *et al.* 2013; Kristensen & Taylor 2016). Lead levels in inner Sydney are also elevated (74% of homes > 300 mg/kg) relative to outer suburbs (< 99 mg/kg) and background values (20-30 mg/kg) due to the former use of leaded paint and petrol historically (Kristensen 2015; Rouillon *et al.* 2017). This study relies on available soil lead data to characterise concentrations at the different sample sites. We have focused on using soil lead measurements because these deposits are stable over time (Semlali *et al.* 2004) and house sparrows spend a large amount of time on the ground foraging and dust bathing. Relevantly, atmospheric lead emission from automotive vehicles,

industry or mining typically accumulate in the uppermost section of the soil profile, i.e. the top 2 cm (Taylor *et al.* 2010), which would form the most significant component of the soil horizon exposed to the sparrows. We found relevant soil lead measurements for all our sampling localities (Table 1).

The lead data were prepared for use in the genomic analysis method BayeScEnv (de Villemereuil & Gaggiotti 2015). This method was chosen because it compares genetic differentiation to environmental differentiation to find loci across the genome that have the strongest covariation with the environmental variable of interest. This covariation is likely caused by selection in many cases. The BayeScEnv method which associates environmental differentiation with genetic differentiation is preferable because our estimates of lead contamination do not have a continues distribution which would be more important in a linear ecological association analysis such as LFMM (Frichot *et al.* 2013). The best way to pass environmental data to the BayeScEnv model is to use standardised distances of environmental differentiation. This requires defining a mean environmental condition and calculating how many standard deviations each location is above or below that mean for a given environmental variable. These values should not be more than 3 standard deviations above or below the mean to avoid over stressing the model (de Villemereuil & Gaggiotti 2015). Here we define our mean environmental condition as the average lead contamination in non-mining towns. We then standardised the distances from this mean by dividing by the standard deviation of non-mining towns (See Table 1). The two towns of Mount Isa and Broken Hill were considered as outliers so were not used to calculate the mean environment or standard deviation. Consequently, the locations of Mount Isa and Broken Hill had environmental differentiation scores that were greater than 3 so all of our environmental differentiation scores were scaled down by dividing by three (Table 1). For Broken Hill and Mount Isa we still had to round the environmental differentiation scores to 3 to minimise supplying the model with excessive outliers (Mount Isa only changed from 3.1). A full summary of the soil

lead values applied to the different locations is detailed in Table 1. To try to account for inaccuracies in estimates of environmental differentiation scores we replicated our analysis with predicted estimates of lead contamination that assumed the mining towns had extreme lead contamination and were given a score of 3 and all other towns had an average score of 0. This provided an exploratory model to compare our results from the estimated contamination model.

Statistical analyses

The Program BayeScEnv was used to do Ecological Association (EA) analyses with our estimates of environmental lead levels. The model was run separately using exploratory values and estimates of environmental lead contamination (See Table 1). Models were run using the same options: 10 pilot runs with 2 000 iterations; a thinning interval of 10; 5 000 outputted iterations and a burn-in length of 10 000. All other options were kept as the default. Lead estimates were given in the form of environmental differentiation (see methods above). This is the recommended method for preparing environmental data for BayeScEnv (de Villemereuil & Gaggiotti 2015). All figures and additional analyses were done using R (R core team, 2017).

Linked genes and gene ontologies

The physical position of significant SNPs from each analysis were used to identify if the SNPs that were physically linked to functional regions on the reference genome. The house sparrow reference genome has 14 260 know protein coding genes (Elgvin *et al.* 2017). The average LD across the house sparrow genome drops to approximately half the background level at ca. 20 kbps, so this was used as a window for physical linkage between SNPs and genes (Lundregan *et al.* in prep). We recorded all genes within 20 kbps of each SNP and if the

SNP was within a gene or within a known exon of a gene. The proportion of SNPs linked to genes was calculated for all analyses. Supplementary tables with lists of outliers SNPs and linked genes are available on Figshare (link: https://figshare.com/s/dbc245f85f4cc3cdd5b1)

Results

The BayeScEnv EA models test if there is a significant association between the differentiation in allele frequencies of our 162 299 SNP loci and lead contamination. The two BayeScEnv models identified 26 significant SNPs in total (exploratory model = 19 and estimated contamination model = 10 SNPs, using q-values, Figure 1 and Table S1), there were three SNPs significant in both models (Figure 2). The 26 Significant SNPs were found to be physically linked to 13 genes, within 20kbps of the SNP (see details Table 2 and S2). Of the three SNPs significant in both models: SNP (SNPi41730) on chromosome 3 was linked to the FAM167A gene which is associated with endoplasmic reticulum membrane structure; the second SNP (SNPa160965) on chromosome 5, was within 42 kbps of the TMEM251 gene and the UBR7 gene, this is still a plausible distance for physical linkage (Lundregan *et al. in prep*); The closest gene to the third SNP (SNPa395289) on chromosome 1A was GATA3 (ca. 326 kbps away), a GATA-binding factor linked to transcription regulation (Table S2). Out of the genes within 20 kbps in the exploratory model one was a metal ion transmembrane transporter (solute transporter protein SLC39A9) associated with the transport of zinc and other metals. Interestingly, the 25th SNP (SNPa222796) with the smallest q-value (q-value = 0.092) in the estimated contamination model was within 20kbps of another zinc transporter (Slc39a8). In the estimated contamination model, the SNP linked to the SLC39A9 gene was also nearing significance (q-value = 0.061) and was the SNP with the 15th smallest q-values out of all 162 299 SNPs. There were no other analogous outlier genes (Table S2).

The exploratory model used average lead levels (score of 0) for all localities except the mining towns of Mount Isa and Broken Hill which were given the maximum lead score of 3.

This run identified five significant SNPs on the 1A chromosome that were within 40kbps of each other. This was the only peak of significant SNPs that were linked to each other and interestingly the nearest downstream gene was a third solute transporter protein (SLC6A15, ca. 400kbps away, Table S2).

Discussion

After scanning 162 299 SNPs across the house sparrow genome we found 26 significant outlier SNPs associated with estimates of lead pollution in the 11 locations investigated (two with high levels of lead). Some of the significant SNPs identified here were physically linked to genes that were previously described to be linked to traits related to lead pollution in other species. For example, the SLC39A9 gene that was a significant in the exploratory mode, is a known zinc transporter in chicken B cells (Taniguchi *et al.* 2013). In the house sparrow this transporter could also be involved in the transport of lead or other metals that are common in the mining towns of Broken Hill and Mount Isa such as zinc itself (Taylor *et al.* 2010, 2014b; Dong *et al.* 2015). Another zinc transporter that was linked to a SNP nearing significance in the estimated contamination model was Slc39a8. This highly conserved gene codes for the ZIP8 zinc transporter which is known to be one of the three primary transporters of lead in many taxa (Liu *et al.* 2008; Nebert *et al.* 2012; Ding *et al.* 2014), only two of the three transporter genes are present on the house sparrow reference genome (ZIP8 and ZIP14 but not DMT1). The Slc39a8 transporter can also transport cadmium which is another toxic metal pollutant (Liu *et al.* 2008; Prozialeck *et al.* 2008). The Slc39a8 gene has been shown to be expressed in the inner ear of rats (*Rattus norvegicus*, Ding *et al.* 2014) and selection for the down regulation of these metal transporters could result in less lead and other trace elements entering the body. For example, in a transcriptomics study Slc39a8 was down regulated in rats exposed to high lead concentrations (Schneider *et al.* 2012). The SLC39 family of zinc transporters are generally involved in bringing metals into the body rather than excreting them

(Jeong & Eide 2013). Therefore, perhaps the most likely mechanism for responding to lead contamination is to down regulate transporters on external surfaces to reduce the amount of lead unintentionally entering the body.

Lead is known to be a stressor/suppressor of the functioning of the endoplasmic reticulum inside cells (Shinkai *et al.* 2010). In both BayeScEnv models a SNP linked to the FAM167A gene was found to be significant and this gene is connected to the structural formation of the endoplasmic reticulum (Table S2). Another SNP Significant in both models was linked to the UBR7 gene, this genes ontology is ubiquitin protein ligase activity, this process is important for breaking down incorrectly folded proteins on the endoplasmic reticulum (Smith *et al.* 2011; Teixeira & Reed 2013) and more specifically E3 ubiquitin ligases like UBR7 (Sriram *et al.* 2011; Christianson & Ye 2014). For another example UBE2J1 has been linked to the task of breaking down terminally misfolded proteins on the endoplasmic reticulum (Burr *et al.* 2011). The SNPa291647 SNP is physical linked to the UBE2J1 gene on the house sparrow reference genome, this SNP has the 32nd lowest q-value (out of 162 229) in the estimated contamination model (q-value = 0.108). The outlier SNPs that are physically linked to genes relating to lead transport and endoplasmic reticulum function give support for the presence of local adaptation in Australian sparrows to lead contamination. For all four genes that appear to be relevant to lead contamination the SNP frequency reached fixation in both mining towns (Figure 3). As Broken Hill and Mount Isa are in separate genetic populations this suggest two independent instances of selection.

In total 3 of our 26 significant SNPs and 2 suggestive SNPs were physically linked (within ca. 20 - 40 kbps) to genes that had prior associations to traits relating to lead pollution (Shinkai *et al.* 2010; Burr *et al.* 2011; Nebert *et al.* 2012; Jeong & Eide 2013; Teixeira & Reed 2013). It is unclear at this stage if other significant loci are associated with traits with unknown relationships to mechanisms relating to how sparrows might be responding to contamination from lead and other metals. Many loci may be part of a gene network that is

responsible for regulating a polygenic trait. Loci that are not physically linked to protein coding regions can also be important to gene networks because they are associated with loci important to gene regulation e.g. promoter regions, epigenetic markers and non-coding RNA segments (Gutierrez-arcelus *et al.* 2015; Koch 2015; Novère 2015; Allis & Jenuwein 2016). As we develop our understanding of gene regulation it may become easier to identify links between candidate loci and the target traits which are predicted to respond to natural selection. This development can happen with improvements to the annotation of the house sparrow reference genome, to include non-protein coding loci that are important to gene expression (Elgvin *et al.* 2017). However not all significant loci are going to be true signals of natural section. false discovery rates are almost unavoidable for genome scan methods (de Villemereuil *et al.* 2014; Whitlock & Lotterhos 2015; Francois *et al.* 2016). Demographic history and high neutral genetic drift in introduced species can also increase false discovery rates (de Villemereuil *et al.* 2014; Hoban *et al.* 2016; Shultz *et al.* 2016).

The samples used in this study were primarily collected to study invasion biology and adaptation to climate but, adapting to pollutants in the environment will also be important for this species. Future work can help expand our understanding of how sparrows adapt to lead in more detail. The first step would be to quantify blood lead levels in sparrow populations across Australia. We predict that sparrows living in mining towns will have higher than average blood lead levels but maybe not as elevated relative to levels in the soil, due to adaptations to mitigate lead contamination. To test this hypothesis a 2x2 common garden experiment could be used. Sparrows from mining and non-mining towns could be collected from the wild and brought to controlled captive environments. After a holding period blood lead levels could be measured before introducing the two groups of birds to two treatments of low (average levels across Australia) and high lead concentrations (levels equivalent to those observed in mining towns). After an exposure period birds could be returned to a controlled zero lead environment for a recovery period and a third blood measurement. We would expect

the birds from the mining communities to have a lower accumulation of lead and a faster recovery. Follow up studies from this could look at mechanisms relating to lead tolerance further by measuring gene expression and physiological variation.

Some of the strongest examples of evolution have been observed when species are exposed to drastically different conditions to what they have originally occupied (Reid et al. 2016). Human caused pollution has arisen over a relatively short evolutionary time scale and we still need to understand how it is affecting wildlife using a wide range of model species. Here we have found evidence that suggests local adaptation is taking place at a genetic level in introduced Australian house sparrow populations, even after a relatively short period of time. We propose that the house sparrow will be a good model system to study these mechanisms further.

Acknowledgements: For funding support: SCA was supported by Macquarie University Research Excellence Scholarships (no. 2013077). SCG was supported by an Australian Research Council Future Fellowship (FT130101253). SNP array genotyping was conducted at CIGENE Norway. We would like to thank the large number of people who helped us with access to urban catch sites for sampling the human commensal house sparrow. We thank, Monica Awasthy, Amanda D. Griffith and Elizabeth L. Sheldon for their participation with field work. We would also like to thank Lee Ann Rollins for their assistance with lab work.

References

Allis CD, Jenuwein T (2016) The molecular hallmarks of epigenetic control. *Nature Reviews Genetics*, **17**, 487–500.

Anderson TR (2006) Biology of the Ubiquitous House Sparrow: from Genes to Populations. In: *University Press New York.*, pp. 31–56. Oxford University Press,Inc, New York.

Andrew SC, Griffith SC (2016) Inaccuracies in the history of a well-known introduction: a case study of the Australian House Sparrow (*Passer domesticus*). *Avian Research*, **7**, 9.

Baccarelli A, Bollati V (2009) Epigenetics and environmental chemicals. *Curr Opin Pediatr.*, **21**, 243–251.

Bonier F, Moore IT, Martin PR, Robertson RJ (2009) The relationship between fitness and baseline glucocorticoids in a passerine bird. *General and Comparative Endocrinology*, **163**, 208–213.

Burr ML, Cano F, Svobodova S *et al.* (2011) HRD1 and UBE2J1 target misfolded MHC class I heavy chains for endoplasmic reticulum-associated degradation. *PNAS*, **108**, 2034–2039.

Callender E (2014) Heavy Metals in the Environment – Historical Trends. In: *Treatise on Geochemistry*, pp. 67–105.

Christianson JC, Ye Y (2014) Cleaning up in the endoplasmic reticulum: ubiquitin in charge. *Nature structural & molecular biology*, **21**, 325–35.

Cyr NE, Romero ML (2007) Chronic stress in free-living European starlings reduces corticosterone concentrations and reproductive success. *General and Comparative Endocrinology*, **151**, 82–89.

Ding D, Salvi R, Roth JA (2014) Cellular localization and developmental changes of Zip8, Zip14 and transferrin receptor 1 in the inner ear of rats. *BioMetals*, **27**, 731–744.

Dong C, Taylor MP, Kristensen LJ, Zahran S (2015) Environmental contamination in an Australian mining community and potential influences on early childhood health and behavioural outcomes. *Environmental Pollution*, **207**, 345–356.

Elgvin TO, Trier CN, Tørresen OK *et al.* (2017) The genomic mosaicism of hybrid speciation. *Science Advances*, **3**, e1602996.

Ericson B, Landrigan P, Taylor MP *et al.* (2016) The Global Burden of Lead Toxicity Attributable to Informal Used Lead-Acid Battery Sites. *Annals of Global Health*, **82**, 686–699.

Francois O, Martins H, Caye K, Schoville S (2016) Controlling false discoveries in genome scans for selection. *Molecular Ecology*, **25**, 454–469.

Frichot E, Schoville SD, Bouchard G, François O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, **30**, 1687–1699.

González F, López I, Suarez L, Moraleda V, Rodríguez C (2017) Levels of blood lead in Griffon vultures from a Wildlife Rehabilitation Center in Spain. *Ecotoxicology and Environmental Safety*, **143**, 143–150.

Gulson B, Anderson P, Taylor A (2013) Surface dust wipes are the best predictors of blood leads in young children with elevated blood lead levels. *Environmental Research*, **126**, 171–178.

Gulson B, Mizon K, Taylor A *et al.* (2014) Pathways of Pb and Mn observed in a 5-year longitudinal investigation in young children and environmental measures from an urban setting. *Environmental Pollution*, **191**, 38–49.

Gutierrez-arcelus M, Ongen H, Lappalainen T (2015) Tissue-Specific Effects of Genetic and Epigenetic Variation on Gene Regulation and Splicing. *PLOS Genetics*, 1–25.

Hoban S, Kelley JL, Lotterhos KE *et al.* (2016) Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist*, **188**, 379–397.

Jensen H, Moe R, Hagen IJ *et al.* (2013) Genetic variation and structure of house sparrow populations: is there an island effect? *Molecular Ecology*, **22**, 1792–805.

Jeong J, Eide DJ (2013) The SLC39 family of zinc transporters. *Molecular Aspects of Medicine*, **34**, 612–619.

Koch L (2015) Genomics: Adding another dimension to gene regulation. *Nature Reviews Genetics*, **16**, 563–563.

Kristensen LJ (2015) Quantification of atmospheric lead emissions from 70 years of leaded petrol consumption in Australia. *Atmospheric Environment*, **111**, 195–201.

Kristensen LJ, Taylor MP (2016) Unravelling a "miner"s myth' that environmental contamination in mining towns is naturally occurring. *Environmental Geochemistry and Health*, **38**, 1015–1027.

Kristensen LJ, Taylor MP, Evans AJ (2016) Tracing changes in atmospheric sources of lead contamination using lead isotopic compositions in Australian red wine. *Chemosphere*, **154**, 40–47.

Kristensen LJ, Taylor MP, Flegal AR (2017) An odyssey of environmental pollution: The rise, fall and remobilisation of industrial lead in Australia. *Applied Geochemistry*, **83**, 3–13.

Laidlaw MAS, Mohmmad SM, Gulson BL *et al.* (2017) Estimates of potential childhood lead exposure from contaminated soil using the US EPA IEUBK Model in Sydney, Australia. *Environmental Research*, **156**, 781–790.

Laidlaw MAS, Taylor MP (2011) Potential for childhood lead poisoning in the inner cities of Australia due to exposure to lead in soil dust. *Environmental Pollution*, **159**, 1–9.

Lanphear BP (2015) The Impact of Toxins on the Developing Brain. *Annual Review of Public Health*, **36**, 211–230.

Lanphear BP, Hornung R, Khoury J *et al.* (2005) Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. *Environmental Health Perspectives*, **113**, 894–899.

Lattin CR, Ngai HM, Romero LM (2014) Evaluating the Stress Response as a Bioindicator of Sub-Lethal Effects of Crude Oil Exposure in Wild House Sparrows (*Passer domesticus*). *PLoS ONE*, **9**, e102106.

Li X, Xing M, Chen M *et al.* (2017) Effects of selenium-lead interaction on the gene expression of inflammatory factors and selenoproteins in chicken neutrophils. *Ecotoxicology and Environmental Safety*, **139**, 447–453.

Lischer HEL, Excoffier L (2012) PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, **28**, 298–299.

Liu Z, Li H, Soleimani M *et al.* (2008) Cd2+ versus Zn2+ uptake by the ZIP8 HCO3--dependent symporter: Kinetics, electrogenicity and trafficking. *Biochemical and Biophysical Research Communications*, **365**, 814–820.

Lundregan S, Jensen H, Hagen I, Husby A, Gohli J Genetic architecture of bill morphology in a free-living house sparrow metapopulation. *in prep*.

Mackay AK, Taylor MP, Munksgaard NC, Hudson-Edwards KA, Burn-Nunes L (2013) Identification of environmental lead sources and pathways in a mining and smelting town: Mount Isa, Australia. *Environmental Pollution*, **180**, 304–311.

Madigosky SR, Alvarez-hernandez X, Glass J (1991) Lead, Cadmium, and Aluminum Accumulation in the Red Swamp Crayfish *Procambarus clarkii G.* Collected from Roadside Drainage Ditches in Louisiana. *Arch. Environ. Contam. Toxicol.*, **20**, 253–258.

Mielke HW, Laidlaw MAS, Gonzales CR (2011) Estimation of leaded (Pb) gasoline's continuing material and health impacts on 90 US urbanized areas. *Environment International*, **37**, 248–257.

Morrison AL (2003) An assessment of the effectiveness of lead pollution reduction strategies in North Lake Macquarie, NSW, Australia. *Science of the Total Environment*, **303**, 125–138.

National Toxicology Program (2012) *health effects of low-level lead. NTP Monograph,* 1-148

Nebert DW, Gálvez-Peralta M, Hay E Ben *et al.* (2012) ZIP14 and ZIP8 zinc/bicarbonate symporters in *Xenopus oocytes*: characterization of metal uptake and inhibition. *Metallomics*, **4**, 1218–25.

NEPM (2013) Schedule B1: Guideline on Investigation for Soil and Groundwater. In: *National Environmental Protection Measure (NEPM)*, p. 57.

Novère N Le (2015) Quantitative and logic modelling of molecular and gene networks. *Nature Reviews Genetics*, **16**, 146–158.

Olszowy H, Torr P, Imray P *et al.* (1995) *Trace Element Concentrations in Soils from Rural and Urban Areas of Australia. Contaminated Sites Monograph Series. No.4.*

Peakall D, Burger J (2003) Methodologies for assessing exposure to metals: Speciation, bioavailability of metals, and ecological host factors. *Ecotoxicology and Environmental Safety*, **56**, 110–121.

Peterson EK, Buchwalter DB, Kerby JL *et al.* (2017) Integrative behavioral ecotoxicology: bringing together fields to establish new insight to behavioral ecology, toxicology, and conservation. *Current Zoology*, **63**, 185–194.

Pierron F, Normandeau E, Defo MA *et al.* (2011) Effects of chronic metal exposure on wild fish populations revealed by high-throughput cDNA sequencing. *Ecotoxicology*, **20**, 1388–1399.

Prozialeck WC, Edwards JR, Nebert DW *et al.* (2008) The vascular system as a target of metal toxicity. *Toxicological Sciences*, **102**, 207–218.

Purcell S, Neale B, Todd-Brown K *et al.* (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, **81**, 559–575.

Qian Y, Tiffany-Castiglioni E (2003) Lead-induced endoplasmic reticulum (ER) stress responses in the nervous system. *Neurochemical Research*, **28**, 153–162.

R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, URL https://www.R-project.org.

Ray PD, Yosim A, Fry RC (2014) Incorporating epigenetic data into the risk assessment process for the toxic metals arsenic, cadmium, chromium, lead, and mercury: Strategies and challenges. *Frontiers in Genetics*, **5**, 1–26.

Reid NM, Proestou DA, Clark BW et al. (2016) The genomic landscape of rapid repeated evolutionary adaptation to toxic pollution in wild fish. *Science*, **354**, 1305–1308.

Ribeiro AM, Risso WE, Fernandes MN, Martinez CBR (2014) Lead accumulation and its effects on the branchial physiology of *Prochilodus lineatus*. *Fish Physiology and Biochemistry*, **40**, 645–657.

Romero LM (2004) Physiological stress in ecology: Lessons from biomedical research. *Trends in Ecology and Evolution*, **19**, 249–255.

Romero LM, Wikelski M (2001) Corticosterone levels predict survival probabilities of Galapagos marine iguanas during El Nino events. *Proceedings of the National Academy of Sciences*, **98**, 7366–7370.

Rouillon M, Gore DB, Taylor MP (2013) The nature and distribution of Cu, Zn, Hg, and Pb in urban soils of a regional city: Lithgow, Australia. *Applied Geochemistry*, **36**, 83–91.

Rouillon M, Harvey PJ, Kristensen LJ, George SG, Taylor MP (2017) VegeSafe: A community science program measuring soil-metal contamination, evaluating risk and providing advice for safe gardening. *Environmental Pollution*, **222**, 557–566.

Schneider JS, Anderson DW, Talsania K, Mettil W, Vadigepalli R (2012) Effects of developmental lead exposure on the hippocampal transcriptome: Influences of sex, developmental period, and lead exposure level. *Toxicological Sciences*, **129**, 108–125.

Semlali RM, Dessogne JB, Monna F *et al.* (2004) Modeling Lead Input and Output in Soils Using Lead Isotopic Geochemistry. *Environ. Sci. Technol*, **38**, 1513–1521.

Shinkai Y, Yamamoto C, Kaji T (2010) Lead induces the expression of endoplasmic reticulum chaperones GRP78 and GRP94 in vascular endothelial cells via the JNK-AP-1 pathway. *Toxicological Sciences*, **114**, 378–386.

Shultz AJ, Baker AJ, Hill GE, Nolan PM, Edwards S V. (2016) SNPs across time and space: population genomic signatures of founder events and epizootics in the House Finch (*Haemorhous mexicanus*). *Ecology and Evolution*, **6**, 7475–7489.

Smith MH, Ploegh HL, Weissman JS (2011) Road to Ruin: Targeting Proteins for Degradation in the Endoplasmic Reticulum. *Science*, **334**, 1086–1090.

Sriram SM, Kim BY, Kwon YT (2011) The N-end rule pathway: emerging functions and molecular principles of substrate recognition. *Nat Rev Mol Cell Biol*, **12**, 735–747.

Taniguchi M, Fukunaka A, Hagihara M *et al.* (2013) Essential Role of the Zinc Transporter ZIP9/SLC39A9 in Regulating the Activations of Akt and Erk in B-Cell Receptor Signaling Pathway in DT40 Cells. *PLoS ONE*, **8**, e58022.

Taylor MP (2015) Atmospherically deposited trace metals from bulk mineral concentrate port operations. *Science of the Total Environment*, **515**–**516**, 143–152.

Taylor MP, Davies PJ, Kristensen LJ, Csavina JL (2014a) Licenced to pollute but not to poison: The ineffectiveness of regulatory authorities at protecting public health from atmospheric arsenic, lead and other contaminants resulting from mining and smelting operations. *Aeolian Research*, **14**, 35–52.

Taylor MP, Mackay AK, Hudson-Edwards KA, Holz E (2010) Soil Cd, Cu, Pb and Zn contaminants around Mount Isa city, Queensland, Australia: Potential sources and risks to human health. *Applied Geochemistry*, **25**, 841–855.

Taylor MP, Mould SA, Kristensen LJ, Rouillon M (2014b) Environmental arsenic, cadmium and lead dust emissions from metal mine operations: Implications for environmental management, monitoring and human health. *Environmental Research*, **135**, 296–303.

Teixeira LK, Reed SI (2013) Ubiquitin Ligases and Cell Cycle Control. *Annual Review of Biochemistry*, **82**, 387–414.

Varian-Ramos CW, Swaddle JP, Cristol DA (2014) Mercury reduces avian reproductive success and imposes selection: An experimental study with adult- or lifetime-exposure in zebra finch. *PLoS ONE*, **9**, e95674.

de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.

de Villemereuil P, Gaggiotti OE (2015) A new FST-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*, **6**, 1248–1258.

Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of FST. *The American Naturalist*, **186**, S24–S36.

Wikelski M, Cooke SJ (2006) Conservation physiology. *Trends in Ecology and Evolution*, **21**, 38–46.

Wikelski M, Wong V, Chevalier B, Rattenborg N, Snell HL (2002) Marine iguanas die from trace oil pollution. *Nature*, **417**, 607–608.

Tables

Table 1. Mean lead data and Environmental differentiation calculations (see below). The "Environmental differentiation" statistic was used for BayeScEnv analyses. The conservative threshold for safe lead levels in areas with urban dwelling is 300 (mg/kg) (NEPM 2013). The two towns above the threshold are heavily affected by mining (Mount Isa and Broken Hill, see Mean/threshold column). Environmental differentiation was calculated using the mean (112 mg/kg) and standard deviation (56 mg/kg) of mean lead contamination at non-mining sites. We scaled these values by dividing by 3 to account for very high lead levels in the two mining towns (Environmental differentiation = ((site value – mean)/SD)/3).

Table 1. (see table legend above)

Pop No.	Location	State	Mean Soil lead (mg/kg)	Citation	Mean/ threshold	Environmental differentiation	Comments/Notes
1	Hobart	TAS	92	(Olszowy *et al.* 1995)	0.31	-0.119	See Table 4 data in Olszowy *et al.* (1995) for Hobart.
2	Melbourne	VIC	173	Laidlaw et al. (in press)	0.58	0.363	Similar estimate to Olszowy *et al.* (1995) for Victoria.
3	Mt Gambier	SA	54	(Olszowy *et al.* 1995)	0.18	-0.345	Used mean for low traffic areas in SA no data for Mount Gambier specifically.
4	Adelaide	SA	144	(Olszowy *et al.* 1995)	0.48	0.190	Air and soil lead are also relatively low in Kristensen *et al.* (2016) for the wine regions of McLaren Vale which is where we sampled this population.
5	Broken Hill	NSW	1500	(Kristensen & Taylor 2016)	5	8.262 (3)	Similar in Dong *et al.* (2015). Background levels ca. 100 mg/kg.
6	Sydney	NSW	210	(Laidlaw *et al.* 2017)	0.7	0.583	Similar in other studies that find about 50% of samples above threshold.
7	Armidale	NSW	46	(Rouillon *et al.* 2013)	0.15	-0.393	We used data from the regional centre of Lithgow which is similar to background levels across NSW (Olszowy *et al.* 1995; Rouillon *et al.* 2017).
8	Toowoomba	QLD	79	(Olszowy *et al.* 1995)	0.26	-0.196	Used QLD mean for old suburbs with low traffic.
9	Townsville	QLD	128	(Taylor 2015)	0.43	0.095	Mean from supplementary Table 4.
10	Longreach	QLD	79	(Olszowy *et al.* 1995)	0.26	-0.196	Used QLD mean for old suburbs with low traffic.
11	Mount Isa	QLD	638	(Taylor *et al.* 2010; Mackay *et al.* 2013)	2.13	3.131 (3)	Taylor (mean 346, n = 60, < 180 µm grain size) and Mackay (mean = 1560, n = 19) were averaged to get a mean. Mackay was focused on mining sites.

Table 2. Summary of significant SNPs from both models using BayeScEnv. From BayeScEnv significant outliers, due to ecological association with the environmental variable, can be assessed using q-values or the more conservative PEP (Posterior Error Probability) statistic. The exploratory model assumed all localities had average lead levels, score of 0, except for Broken Hill and Mount Isa that had high lead scores of 3. The estimated contamination model used the best estimates of lead from the literature (Table 1). We present both models because the estimates are similar and neither are likely to be 100% accurate.

Analysis	Total SNPs	No. SNPs linked to genes	No. of genes linked to SNPs	Percentage of SNPs linked to genes	Percentage of SNPs within genes	Percentage of SNPs within exons
Exploratory model – q-value	19	6	9	0.316	0.105	0
PEP	11	3	5	0.273	0.091	0
Estimated model – q-value	10	3	5	0.3	0	0
PEP	3	2	3	0.667	0	0

Figures

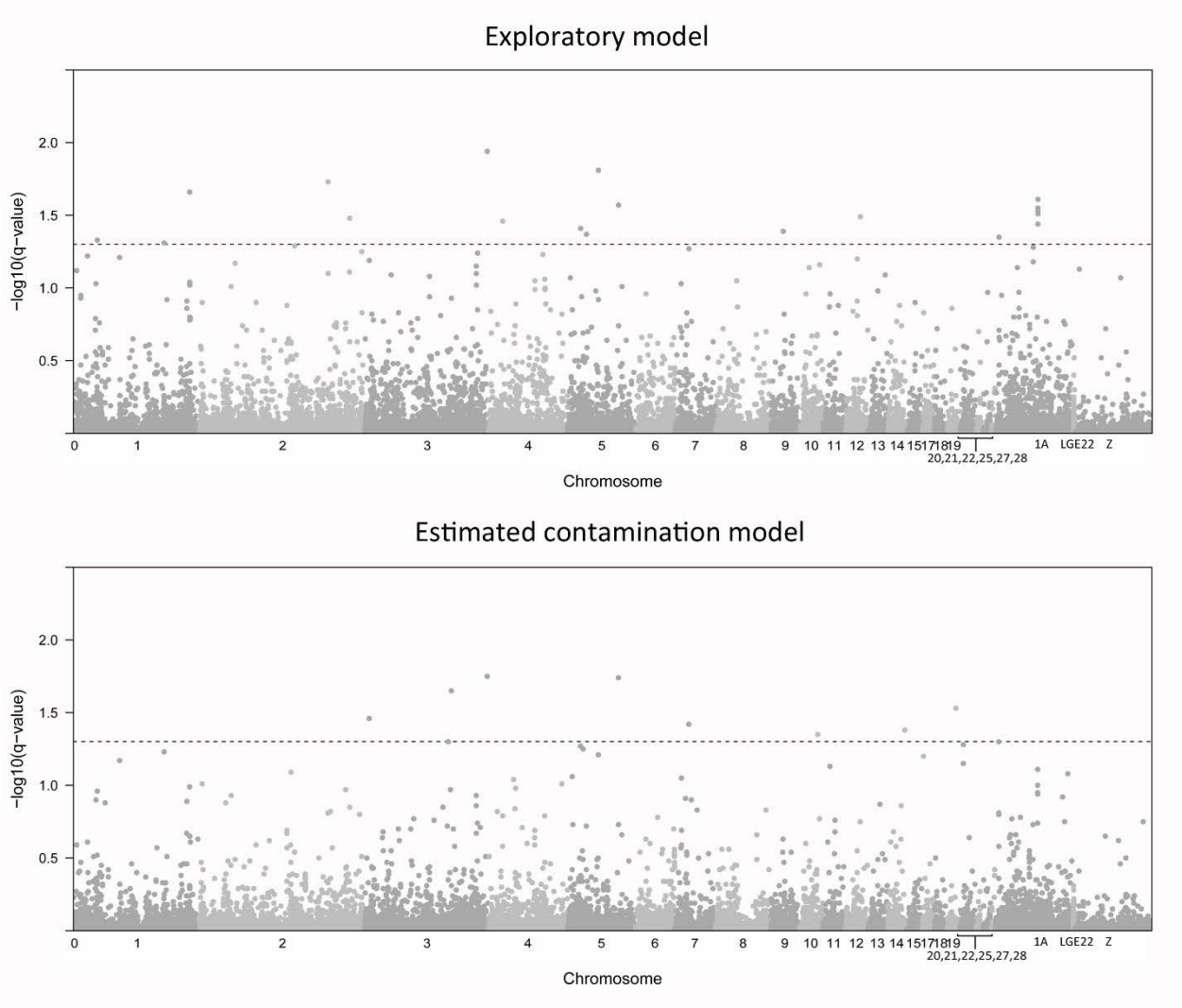


Figure 1. Manhattan plots for both BayeScEnv runs. The first exploratory model has an interesting peak on chromosome chr1A with 5 significant SNPs that are very close to each other. The exploratory model scored most of the locations as having normal levels of lead except for Mount Isa and Broken Hill which were allocated high environmental lead levels. The results for the estimated contamination model uses estimates of environmental lead from the literature (see Table 1). These plots display the log transformed q-values and SNPs above the dashed line are significant.

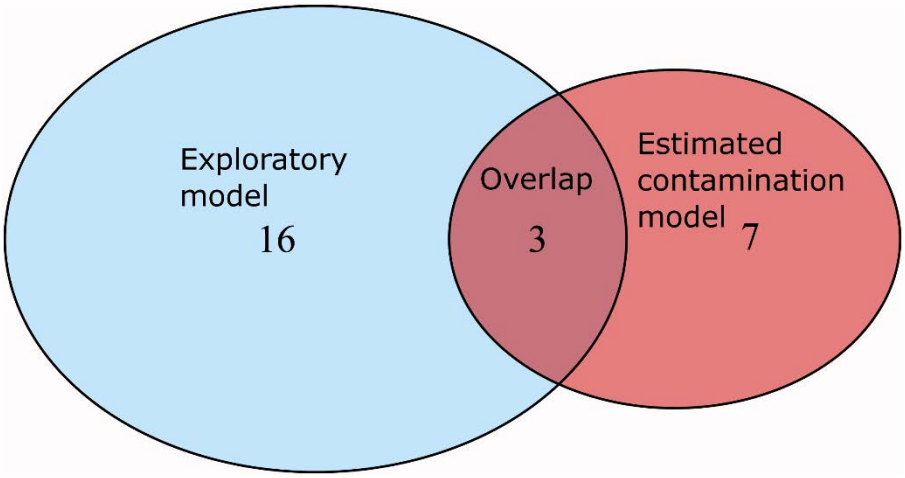


Figure 2. Venn Diagram for the overlap between the significant SNPs from the two models.

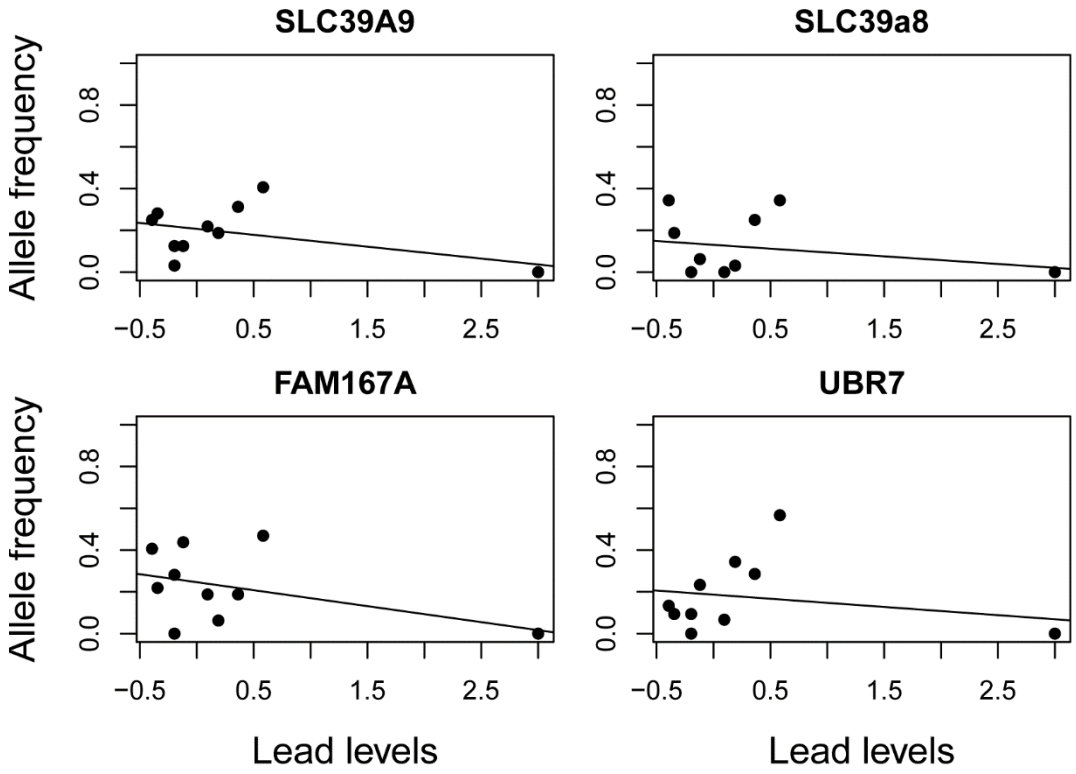
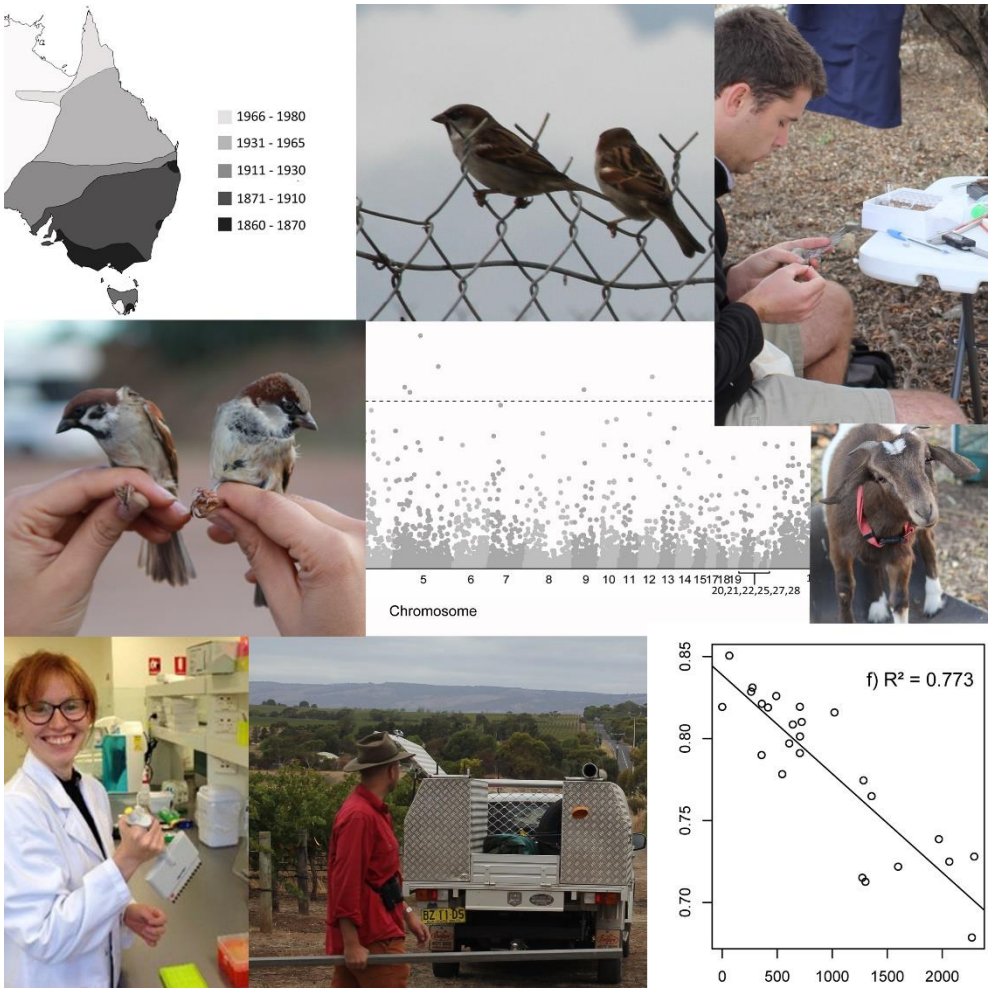


Figure 3. Allele frequencies for the SNPs linked to the four genes that are related to lead exposure. For all four genes the SNP frequency dropped to 0 in the Broken Hill and Mount Isa (lead level = 3). This is one pattern we might expect to see if there is selection taking place in the most differentiated environments but less so in the more moderate environments.

Chapter Eight

General Discussion



The globally distributed house sparrow (*Passer domesticus*) has historically been a model species for studying evolutionary ecology (Anderson 2006; Liebl *et al.* 2015). In Australia, the house sparrow has been introduced to a broad range of climates, creating a 'natural' experiment that can be used to observe how the species adapts to highly varied environments. In this thesis, we have: made a full description of the house sparrows' introduction to Australia using primary and secondary sources (Chapter Two); Observed clinal variation in morphology across the Australian and New Zealand distributions that is associated with average summer temperatures (Chapter Three); Described population structure using microsatellites, across the Australian and New Zealand distribution (Chapter Five); Identified signatures of local adaptation associated with climatic variation at a continental scale in the house sparrow using a high density SNP array (Chapter Six); Finally we have found convincing evidence of natural selection due to lead contamination in an avian species using a genome scan approach (Chapter Seven). Each of these studies connects to tell the story of how this species has become a ubiquitous member of urban environments across eastern Australia.

The house sparrow was successfully introduced to a number of locations in eastern Australia. The first introductions occurred in the 1860's and the current distribution across eastern Australia was established by the 1970's (Chapter Two, see Figure 1). This range expansion likely took multiple roots from several introduction sites near the capital cities where Acclimatisation Societies introduced a number of species to Australia (Rolls 1969; Courcy 2003; Lever 2005). The introductions to the five capital cities in eastern Australia are consistent with the five main genetic populations we see today across Australia (Chapter Five). The birds introduced to Sydney and Hobart were reported to have been sourced from the population imported to Melbourne (Chapter Two). Independent introductions of sparrows sent from England in the 1860s were made in Adelaide and Brisbane (supplemented with birds from Melbourne) (Chapter Two). A majority of the house sparrow shipments to

Australia and New Zealand contained birds from Great Britain and Western Europe. However, there are also reports of birds being imported from India to Melbourne (Chapter Two). With the collection of more genetic data from India it will be possible to resolve if there has been genetic admixture between the *P. d. domesticus* subspecies from Europe and the *P. d. indicus* subspecies from India. Admixture between these two subspecies could have resulted in increased genetic diversity and the addition of functional genetic diversity from the Indian sparrows that come from hotter and more tropical climates than the temperate European climates. This admixture with Indian sparrows could have provided genetic variation that helped the sparrows colonise Australia's hotter climates, and we are still attempting to obtain samples from India to test this hypothesis. Invasion history is a fundamental component of the biology of an invasive species and can be important for planning research on these species. It has become clear from our work on the sparrow that general sources of information on introduction history (Long 1981; Lever 2005; Higgins *et al.* 2006) are not always complete or consistent. The case of the house sparrow suggests this problem could be more widespread and these reporting inconsistences could affect future research on other systems.

Invasive species can be excellent systems for studying evolution and local adaptation. The potential for species to respond to 'overnight' changes in environment as they get translocated to new environments and climates can be realized by immediate adjustments through phenotypic plasticity as well as selection over multiple generations. The most studied ecogeographic rule that describes morphology differentiation across different climates, is Bergmann's rule (Blackburn *et al.* 1999). This rule predicts endotherms will have a larger body size in colder climates. This pattern can be observed within species and amongst closely related species with varied distributions (Salewski & Watt 2017). The house sparrow has been used in classic studies of Bergmann's rule that focused on the introduced populations of North America (Johnston & Selander 1964, 1971) and the native populations of Europe (Johnston

1969). When these studies were replicated in Australia and New Zealand, the best predictor of the clinal variation in body size (of larger body size in colder climates) was summer maximum temperature rather than the previously more studied winter minimum temperature (Chapter Three).

There is growing support for a hypothesis that predicts that the temperature experienced during development could be affecting avian body size, with higher temperatures resulting in a smaller body size attained (Adams 2008; Gardner *et al.* 2009, 2014a; b; Van Buskirk *et al.* 2010; Gienapp & Merilä 2014). To measure the effects of temperature on development we moved to the zebra finch (*Taeniopygia guttata*), a species more amenable to breeding experiments. We analysed observational data from wild zebra finch fledglings growing across the breeding season and experimental data from fledglings reared in temperature controlled rooms, to test if this negative relationship between temperature and body size, is explained by developmental plasticity that is regulated by temperature exposure during growth (Chapter Four). We found consistent results that showed higher temperatures during development result in a smaller body size. This phenotypic plasticity is in a direction that is consistent with Bergmann's rule and the thermoregulatory rule that a smaller body size increases the surface area to volume ratio, which increases heat dissipation. Therefore, the plasticity observed here is in the direction we would expect local adaptation to take so it should be considered as adaptive plasticity (Ghalambor *et al.* 2007). The results presented in Chapters Three and Four suggest Bergmann's rule needs to be reconsidered so researchers acknowledge that this clinal variation within species can be explained by a combination of phenotypic plasticity driven by developmental temperatures and genetic differentiation due to selection, working either independently or in combination. For the house sparrow populations, in addition to the response variable of summer temperature, approximately 10 % of the variation in body size was explained by the random factor of location (Chapter Three see Table 3). This random factor could be partitioning out variation explained by other

environmental variables not conflated with temperature or by genetic differentiation between populations.

Genetic population structure is relatively strong across the Australian house sparrow's distribution (Chapter Five). This population structure seems to fit with the introduction and range expansion of the species. To achieve their current distribution, this obligate commensal species must have had to make some significant jumps between the isolated towns in the rural areas of Australia away from the coast. These jumps between suitable habitat likely resulted in demographic bottlenecks and this is demonstrated by the decline in genetic diversity in more recently established populations (Chapter Five see Figure 4). Across Australia, we observe significant pair-wise F_{ST} and a pattern similar to Isolation by Distance (IBD), even though it is technically not one continuous distribution. These results suggest the populations in towns and cities across Australia form a meta-population spread across isolated habitat patches. This meta-population, which is spread across a broad range of climates, provides a good context for evolutionary research. Population genetics studies on invasive species are not just important for learning more about the population structure and introduction history, the information can also be useful for informing the study design of genome scans for local adaptation. The primary benefit of information about population structure and demographic history is that it can improve sampling design for genome scan projects and hence the reliability of results (De Mita *et al.* 2013; Rellstab *et al.* 2015; Francois *et al.* 2016). Therefore, broader sampling and genotyping with a few genetic markers can be advantageous for focusing in on a subset of populations for intensive genome scan studies.

Using genomic SNP data, we tested if signatures of selection were present across Australian house sparrow populations. We applied both Population Differentiation (PD) and Ecological Association (EA) genome scan methods and found hundreds of significant outliers (Chapter Six). Approximately 38% of these SNPs were physically linked to know protein coding genes on the house sparrow reference genome (Elgvin *et al.* 2017). It is not feasible to

do a detailed evaluation of all these outlier genes here but we did find that a number of them were previously identified in studies on local adaptation in other species and were related to traits such as thermal tolerance, immune function and morphology that are expected to be relevant to local adaptation (Chapter Six see Table 3). Interestingly, the only outlier gene that was significant in three out of four methods was a gene linked to body fat production. The NPY2R gene is linked to a pathway that regulates feeding behaviour and fat acquisition in a number of species (Kuo *et al.* 2007; Coccurello *et al.* 2009). This gene could be linked to the observed differentiation in body mass across Australian sparrow populations that is not explained by temperature (Chapter Three) and should be considered as an important target for future research. However, not all of the outlier loci are likely to be true signals of natural selection, recent colonization events and high genetic drift can make it hard to identify true signals of natural selection due to increased false discovery rates (de Villemereuil *et al.* 2014; de Villemereuil & Gaggiotti 2015; Francois *et al.* 2016). The research presented in Chapter Six asks similar questions to research on Norwegian house sparrows using the 200K SNP array, however, our sample sites are spread over a larger geographical area that covers a wider range of climates in comparison to a concentrated effort on a small part of the Norwegian coast (Lundregan *et al.* in prep; Jensen *et al.* 2013; Silva *et al.* 2017). This work is also novel because it is one of the first to have performed genome scan analyses to look for local adaptation in an introduced avian species at a continental scale.

Apart from living across varied climates, the house sparrow also faces other challenges in the urban environment including pollution and contamination from trace elements such as lead (Kristensen 2015; Kristensen *et al.* 2017). Lead (Pb) contamination is known to be toxic to avian species and other animals (Beyer *et al.* 1988, 2000; Lanphear *et al.* 2005; National Toxicology Program 2012; Lanphear 2015; González *et al.* 2017). Lead levels can be extremely elevated in urban environments due to human industry (Mielke *et al.* 2011; Callender 2014; Kristensen 2015; Kristensen *et al.* 2017). It would seem plausible that there

would be some selection pressure for adaptations that reduce the uptake of lead or the impact of lead on the body in sparrow populations from highly contaminated mining and smelting communities (Taylor *et al.* 2010; Dong *et al.* 2015; Kristensen & Taylor 2016). We tested if there were signs of selection as a result of lead contamination across Australia. An EA analyses using different estimates of lead contamination detected a small number of significant outlier loci (Chapter Seven). Interestingly, some of these outliers could be relevant to lead contamination. One outlier was linked to a metal transporter (SLC39A9) and another suggestive outlier (top 30 lowest q-values) was linked to a second metal ion transporter (SLC39A8) on a separate chromosome (Chapter Seven). These transporters primarily transport zinc but the second transporter has also been shown to transport lead and other heavy metals (Liu *et al.* 2008; Ding *et al.* 2014). Another outlier gene was associated with endoplasmic reticulum structure, and the endoplasmic reticulum is known to be stressed by increased lead exposure (Qian & Tiffany-Castiglioni 2003). This stress may result in misfolded proteins on the endoplasmic reticulum. Interestingly, another outlier and a suggestive outlier were linked to two ubiquitin proteins that break down misfolded proteins on the endoplasmic reticulum (Burr *et al.* 2011; Smith *et al.* 2011). These results show that there are outlier loci that show a significant association between genetic differentiation and environmental variation in average lead contamination (Chapter Seven). Some of these outlier loci were also linked to candidate genes that could be part of mechanisms involved in responding to lead contamination. If society wants to increase diversity in habitat heavily affected by human disturbance then understanding how species need to adapt to contamination from trace elements should be a critical focus for future research.

Changes in the allele frequencies of functional loci is the fundamental mechanism behind evolution. Our ability to observe the changes in allele frequency that are the signatures of selection, above the general genetic differentiation between populations due to drift, is possible due to current genome scan analyses methods. The ability to detect changes in allele

frequency that covary with climate across different genetic populations is important for identifying if natural selection is taking place. The demographic history of the house sparrow in Australia is relatively complex for studies of local adaptation due to multiple introduction events, successive founder events and demographic bottlenecks. Recently, improved genome scan analyses methods have been developed to start addressing these issues (de Villemereuil *et al.* 2014; Lotterhos & Whitlock 2014; Rellstab *et al.* 2015; Francois *et al.* 2016; Hoban *et al.* 2016). The characteristics of the house sparrow's introduction have created genetic population structure due to genetic drift, across a broad range of climates. This population structure and environmental variation creates an excellent 'natural' experiment to observe local adaptation. Across this broad range of climates, we see morphological differentiation that is partially explained by adaptive plasticity. In addition to plasticity, we find signatures of genetic adaptation that are observed above the existing genetic population structure at potentially ecologically relevant loci (Chapters Six and Seven). These patterns are exactly what we would expect to observe in an invasive species that has expanded its distribution to new environments and has had to adapt to the new conditions. There is still high potential for future research to use the house sparrow system to build our understanding of local adaptation. One important goal for future research relating to Bergmann's rule is to develop an understanding of the mechanisms behind the developmental plasticity in body size that is affected by temperatures during development. The mechanisms behind thermal tolerance in sparrows is another important trait for future research and the heat shock proteins that had linked SNPs that were significantly associated with summer maximum temperature provide targets for future research (Chapter Six). The consistent genotyping of thousands of SNPs across the house sparrow genome using the 200K SNP array, mean it should be viable to scale up these studies to collect genomic data across the global distribution of the species to explore these questions further and replicate the genomic results at a global scale.

The PD and EA genome scan methods discussed here consider changes in the allele frequency of loci independently across the genome. However, many of the traits which we expect to be important for local adaptation are likely to be polygenic. Within the gene network of a polygenic trait there may be many small-effect loci that do not show any significant differentiation or an association with environment, due to selection acting more strongly on combinations of large-effect loci. This problem is similar to the technical issues faced in Quantitative Trait Loci (QTL) and Genome Wide Association Studies (GWAS) analyses used to identify loci that explain phenotypic variation (Ernst & Steibel 2013; Van De Geer *et al.* 2014; Zhu *et al.* 2016; Li *et al.* 2017). When using these methods to identify the genetic architecture of polygenic traits, small-effect loci can be missed and these methods do not always accurately estimate the effect sizes of loci resulting in inflated estimates and false positives due to this random inflation (Kemppainen *et al.* 2016; Mcfarlane *et al.* 2016). Both GWAS and QTL methods have also been criticised for the long standing issue of 'missing heritability', because the combined variance explained by the significant QTL and GWAS loci is often much lower than the variance explained by traditional narrow sense heritability (h^2) (Brachi *et al.* 2011). One potential direction for improving our understanding of the genetic architecture of traits and for closing the 'missing heritability' gap is by combining genomic data (GWAS) with gene expression data (eQTL) (Zhu *et al.* 2016).

The most interesting direction I see for expanding our understanding of the house sparrow system in Australia is to collect phenotypic data on key traits such as bill morphology (Silva *et al.* 2017), immune response (Liebl & Martin 2009) and thermal tolerance (if possible), and combine it with high density SNP and gene expression data (Chutimanitsakun *et al.* 2011; Smith *et al.* 2013; Ekblom *et al.* 2014). The GWAS and gene expression (transcriptome sequencing) outliers significantly associated with phenotypic variation could then be used to estimate 'Outlier Loci Heritability', this new measure would be similar to SNP heritability (h^2_{SNP}) which is an estimate of the phenotypic variance explained by the

additive contributions of SNPs (Speed *et al.* 2017). The combination of genetic and gene expression outliers would estimate the similarity between individuals using the loci that are most strongly associated with phenotypic variation, and use this kind of 'relatedness' to estimate heritability. This strategy could help elucidate the networks of loci responsible for phenotypic variation and close the missing heritability gap.

Invasive species have always been excellent models for understanding adaptation to different environments and to rapidly changing environmental conditions through colonization (Lee 2002; Sax *et al.* 2007). The house sparrow has been one of these key model species with work undertaken on this species around the world (Anderson 2006; Liebl *et al.* 2015). In this thesis, I have expanded this research to the Australian continent, collecting data from populations across the broad range of the species in this region. However, there is still a long way to go before we can fully characterise the mechanisms behind the house sparrow's resilience and adaptability, that makes this species such a successful climate generalist. Common species such as the house sparrow represent a disproportionately large part of ecosystem biomass and have a large effect on spatial species richness (Gaston 2008) and are therefore enormously important to the overall composition and functioning of ecosystems (Gaston 2010). Understanding what makes species common and how to prevent their decline or localised extinction, is fundamental to conserving the diversity of ecosystems (Gaston 2010). Ultimately, the success of conservation will depend on an overall understanding of biology that includes the few common species and the many uncommon species. Understanding how species adapt to different environmental conditions through both phenotypic plasticity and selection will potentially make it possible to predict the response of wildlife and ecosystems to rapid changes in habitat and climate caused by humans. If humanity is going to act to preserve threatened species, we need to understand the mechanisms behind the traits that are limiting the potential of threatened species to adapt and respond to change.

References

Adams DC (2008) Phylogenetic Meta-Analysis. *Evolution*, **62**, 567–572.

Anderson TR (2006) Biology of the Ubiquitous House Sparrow: from Genes to Populations. In: *University Press New York.*, pp. 31–56. Oxford University Press,Inc, New York.

Beyer WN, Audet DJ, Heinz GH, Hoffman DJ, Day D (2000) Relation of waterfowl poisoning to sediment lead concentrations in the Coeur d'Alene River Basin. *Ecotoxicology*, **9**, 207–218.

Beyer WN, Spann JW, Siteo L, Franson JC (1988) Lead poisoning in 6 captive avian species. *Archives of Environmental Contamination and Toxicology*, **17**, 121–130.

Blackburn TM, Gaston KJ, Loder N, Jul N (1999) Geographic gradients in body size: A clarification of Bergmann's rule. *Biodiversity Research*, **5**, 165–174.

Brachi B, Morris GP, Borevitz JO (2011) Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology*, **12**, 232.

Burr ML, Cano F, Svobodova S *et al.* (2011) HRD1 and UBE2J1 target misfolded MHC class I heavy chains for endoplasmic reticulum-associated degradation. *PNAS*, **108**, 2034–2039.

Van Buskirk J, Mulvihill RS, Leberman RC (2010) Declining body sizes in North American birds associated with climate change. *Oikos*, **119**, 1047–1055.

Callender E (2014) Heavy Metals in the Environment – Historical Trends. In: *Treatise on Geochemistry*, pp. 67–105.

Chutimanitsakun Y, Nipper RW, Cuesta-Marcos A *et al.* (2011) Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. *BMC genomics*, **12**, 4.

Coccurello R, D'Amato FR, Moles A (2009) Chronic social stress, hedonism and vulnerability to obesity: Lessons from Rodents. *Neuroscience and Biobehavioral Reviews*, **33**, 537–550.

Courcy C (2003) Evolution of a Zoo: A History of the Melbourne Zoological Gardens. pp. 1–15. Quiddlers Press, Auburn, Australia.

Ding D, Salvi R, Roth JA (2014) Cellular localization and developmental changes of Zip8, Zip14 and transferrin receptor 1 in the inner ear of rats. *BioMetals*, **27**, 731–744.

Dong C, Taylor MP, Kristensen LJ, Zahran S (2015) Environmental contamination in an Australian mining community and potential influences on early childhood health and behavioural outcomes. *Environmental Pollution*, **207**, 345–356.

Ekblom R, Wennekes P, Horsburgh GJ, Burke T (2014) Characterization of the house sparrow (*Passer domesticus*) transcriptome: a resource for molecular ecology and immunogenetics. *Molecular ecology resources*, **14**, 636–646.

Elgvin TO, Trier CN, Tørresen OK *et al.* (2017) The genomic mosaicism of hybrid speciation. *Science Advances*, **3**, e1602996.

Ernst CW, Steibel JP (2013) Molecular advances in QTL discovery and application in pig breeding. *Trends in Genetics*, **29**, 215–224.

Francois O, Martins H, Caye K, Schoville S (2016) Controlling false discoveries in genome scans for selection. *Molecular Ecology*, **25**, 454–469.

Gardner JL, Amano T, Backwell PRY *et al.* (2014a) Temporal patterns of avian body size reflect linear size responses to broadscale environmental change over the last 50 years. *Journal of Avian Biology*, **45**, 529–535.

Gardner JL, Amano T, Mackey BG *et al.* (2014b) Dynamic size responses to climate change: prevailing effects of rising temperature drive long-term body size increases in a semi-arid passerine. *Global Change Biology*, **20**, 2062–2075.

Gardner JL, Heinsohn R, Joseph L (2009) Shifting latitudinal clines in avian body size correlate with global warming in Australian passerines. *Proceedings. Biological sciences / The Royal Society*, **276**, 3845–52.

Gaston KJ (2008) Biodiversity and extinction: the importance of being common. *Progress in Physical Geography*, **32**, 73–79.

Gaston KJ (2010) Valuing common species. *Science*, **327**, 154–5.

Van De Geer S, Bühlmann P, Ritov Y, Dezeure R (2014) On asymptotically optimal confidence regions and tests for high-dimensional models. *Annals of Statistics*, **42**, 1166–1202.

Ghalambor CK, McKay JK, Carroll SP, Reznick DN (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, **21**, 394–407.

Gienapp P, Merilä J (2014) Disentangling plastic and genetic changes in body mass of Siberian jays. *Journal of Evolutionary Biology*, **27**, 1849–1858.

González F, López I, Suarez L, Moraleda V, Rodríguez C (2017) Levels of blood lead in Griffon vultures from a Wildlife Rehabilitation Center in Spain. *Ecotoxicology and Environmental Safety*, **143**, 143–150.

Higgins PJ, Peter JM, Cowling SJ (2006) Handbook of Australian, New Zealand and Antarctic birds: Volume 7 Boatbill to Starlings, pp. 1102–26. Oxford University Press, Melbourne.

Hoban S, Kelley JL, Lotterhos KE *et al.* (2016) Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist*, **188**, 379–397.

Jensen H, Moe R, Hagen IJ *et al.* (2013) Genetic variation and structure of house sparrow populations: is there an island effect? *Molecular Ecology*, **22**, 1792–805.

Johnston RF (1969) Character Variation and Adaptation in European Sparrows. *Systematic Zoology*, **18**, 206.

Johnston RF, Selander RK (1964) House sparrows: rapid evolution of races in North America. *Science*, **144**, 548–50.

Johnston RF, Selander RK (1971) Evolution in the house sparrow. II. Adaptive differentiation in North American populations. *Evolution*, **25**, 1–28.

Kemppainen P, Rønning B, Kvalnes T *et al.* (2017) Controlling for p-value inflation in allele frequency change in experimental evolution and artificial selection experiments. *Molecular Ecology Resources*, **17,** 770-782.

Kristensen LJ (2015) Quantification of atmospheric lead emissions from 70 years of leaded petrol consumption in Australia. *Atmospheric Environment*, **111**, 195–201.

Kristensen LJ, Taylor MP (2016) Unravelling a "miner"s myth' that environmental contamination in mining towns is naturally occurring. *Environmental Geochemistry and Health*, **38**, 1015–1027.

Kristensen LJ, Taylor MP, Flegal AR (2017) An odyssey of environmental pollution: The rise, fall and remobilisation of industrial lead in Australia. *Applied Geochemistry*, **83**, 3–13.

Kuo LE, Kitlinska JB, Tilan JU *et al.* (2007) Neuropeptide Y acts directly in the periphery on fat tissue and mediates stress-induced obesity and metabolic syndrome. *Nature Medicine*, **13**, 803–811.

Lanphear BP (2015) The Impact of Toxins on the Developing Brain. *Annual Review of Public Health*, **36**, 211–230.

Lanphear BP, Hornung R, Khoury J *et al.* (2005) Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. *Environmental Health Perspectives*, **113**, 894–899.

Lee CEE (2002) Evolutionary genetics of invasive species. *Trends in Ecology and Evolution*, **17**, 386–391.

Lever C (2005) Naturalised Birds of the world. pp. 204–218. T & A D Poyser, London.

Li Z, Guo B, Yang J *et al.* (2017) Deciphering the genomic architecture of the stickleback brain with a novel multi-locus gene-mapping approach. *Molecular Ecology*, **26**, 1557–1575.

Liebl AL, Martin LB (2009) Simple quantification of blood and plasma antimicrobial capacity using spectrophotometry. *Functional Ecology*, **23**, 1091–1096.

Liebl AL, Schrey AW, Andrew SC, Sheldon EL, Griffith SC (2015) Invasion genetics: Lessons from a ubiquitous bird, the house sparrow *Passer domesticus*. *Current Zoology*, **61**, 465–476.

Liu Z, Li H, Soleimani M *et al.* (2008) Cd2+ versus Zn2+ uptake by the ZIP8 HCO3--dependent symporter: Kinetics, electrogenicity and trafficking. *Biochemical and Biophysical Research Communications*, **365**, 814–820.

Long JL (1981) Introduced Birds of the World: The Worldwide History, Distribution and Influence of Birds Introduced to New Environments. pp. 372–381. Universe Books, New South Wales.

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, **23**, 2178–2192.

Lundregan S, Jensen H, Hagen I, Husby A, Gohli J. Genetic architecture of bill morphology in a free-living house sparrow metapopulation. *in prep*.

Mcfarlane E, Husby A, Kawakami T *et al.* (2016) Increasing the power of genome wide association studies in natural populations using repeated measures evaluation and implementation. *Methods in Ecology and Evolution*, **7**, 792-799.

Mielke HW, Laidlaw MAS, Gonzales CR (2011) Estimation of leaded (Pb) gasoline's continuing material and health impacts on 90 US urbanized areas. *Environment International*, **37**, 248–257.

De Mita S, Thuillet A-C, Gay L *et al.* (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, **22**, 1383–1399.

National Toxicology Program (2012) National Toxicology Program: health effects of low-level lead. *NTP Monograph*, 1–148.

Qian Y, Tiffany-Castiglioni E (2003) Lead-induced endoplasmic reticulum (ER) stress responses in the nervous system. *Neurochemical Research*, **28**, 153–162.

Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015) A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, **24**, 4348–4370.

Rolls EC (1969) They all Ran Wild. pp. 225–250. Angus & Robertson LTD, Sydney.

Salewski V, Watt C (2017) Bergmann's rule: a biophysiological rule examined in birds. *Oikos*, **126**, 161–172.

Sax DF, Stachowicz JJ, Brown JH *et al.* (2007) Ecological and evolutionary insights from species invasions. *Trends in ecology & evolution*, **22**, 465–71.

Silva CNS, McFarlane SE, Hagen IJ *et al.* (2017) Insights into the genetic architecture of morphological traits in two passerine bird species. *Heredity*, 1–9.

Smith S, Bernatchez L, Beheregaray LB (2013) RNA-seq analysis reveals extensive transcriptional plasticity to temperature stress in a freshwater fish species. *BMC genomics*, **14**, 375.

Smith MH, Ploegh HL, Weissman JS (2011) Road to Ruin: Targeting Proteins for Degradation in the Endoplasmic Reticulum. *Science*, **334**, 1086–1090.

Speed D, Cai N, Johnson M, Nejentsev S, Balding D (2017) Re-evaluation of SNP heritability in complex human traits. *Nature Genetics*, **49**, 986–992.

Taylor MP, Mackay AK, Hudson-Edwards KA, Holz E (2010) Soil Cd, Cu, Pb and Zn contaminants around Mount Isa city, Queensland, Australia: Potential sources and risks to human health. *Applied Geochemistry*, **25**, 841–855.

de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.

de Villemereuil P, Gaggiotti OE (2015) A new FST-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*, **6**, 1248–1258.

Zhu Z, Zhang F, Hu H *et al.* (2016) Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nature genetics*, **48**, 481–7.

Appendix 1: Publications during candidature

Related to Ph.D. project

Liebl AL, Schrey AW, Andrew SC, et al. (2015) Invasion genetics: Lessons from a ubiquitous bird, the house sparrow *Passer domesticus*. *Current Zoology,* **61**, 465–476.

Andrew SC and Griffith SC (2016) Inaccuracies in the history of a well-known introduction: a case study of the Australian House Sparrow (*Passer domesticus*). *Avian Research*, **7**, 9. (Chapter Two)

Unrelated to Ph.D. project

Andrew SC, Perry CJ, Barron AB, et al. (2014) Peak shift in honey bee olfactory learning. *Animal Cognition*, **17,** 1177–1186.

Andrew SC and Kemp DJ (2016) Stress tolerance in a novel system: Genetic and environmental sources of (co)variation for cold tolerance in the butterfly *Eurema smilax*. *Austral Ecology*, **41**, 529–537.

Griffith SC, Crino OL, Andrew SC et al. (2017) Variation in Reproductive Success Across Captive Populations: Methodological Differences, Potential Biases and Opportunities. *Ethology*, **123**, 1–29.

Griffith SC, Crino OL and Andrew SC (2017) Commentary: The Challenge of Being Ecologically Relevant in Captivity. *Frontiers in Ecology and Evolution*, doi: 10.3389/fevo.2017.00021.

Appendix 2: List of Conference Presentations

Conference presentations made during my PhD candidature at Macquarie University.

Andrew SC, Perry CJ, Barron AB, Berthon K, Peralta V and Cheng K. 2014. Discrimination learning and peak shift in honey bees. Australasian Society for the Study of Animal Behaviour, Katoomba, Australia (talk)

Andrew SC, Rollins LA and Griffith SC. 2015. Drivers of population differentiation: A natural experiment with the Australian house sparrow. Genetics Society of Australasia, Adelaide, Australia (talk)

Andrew SC, Awasthy M, Jensen H, Lien S, Nakagawa S, Rollins LA and Griffith SC. 2016. Population differentiation and local adaptation in the Australian house sparrow. Society of Molecular Biology and Evolution, Gold Coast, Australia (poster)

Appendix 3: Animal Ethics Approvals

Australian house sparrow ethics approval letter

MACQUARIE UNIVERSITY

ANIMAL RESEARCH AUTHORITY (ARA)

AEC Reference No.: 2014/014 **Date of Expiry:** 31 March 2015

Full Approval Duration: *1 April 2014 to 31 March 2016 (24 Months)*

This ARA remains in force until the Date of Expiry (unless suspended, cancelled or surrendered) **and will only be renewed upon receipt of a satisfactory Progress Report before expiry (see Approval email for submission details).**

Principal Investigator:
A/Prof Simon Griffith
Biological Sciences
Macquarie University, NSW 2109
0425 746 674
Simon.Griffith@mq.edu.au

Associate Investigators:
Lee Rollins 0425 327 112
Samuel Andrew 0404 119 473

In case of emergency, please contact:
the Principal Investigator / Associate Investigator named above
Animal Welfare Officer - 9850 7758 / 0439 497 383, *or* **Manager, Fauna Park** - 9850 4109 / 0425 213 420

The above-named are authorised by MACQUARIE UNIVERSITY ANIMAL ETHICS COMMITTEE to conduct the following research:

Title of the project: **The population genetics of the house sparrow in Australia**

Purpose: **7 - Research: Environmental Study**

Aims: **To establish a long-term study to monitor reproduction across the whole range of the species in Australia to understand how climatic variability influences behavior and physiology**

Surgical Procedures category: 1 - Observation Involving Minor Interference

All procedures must be performed as per the AEC-approved protocol, unless stated otherwise by the AEC and/or AWO.

Maximum numbers approved (for the Full Approval Duration):

Species	Strain	Age/Sex	Total	Supplier/Source
17 – Exotic wild bird	House sparrow, Passer Domesticus	Adult / Any	700	Wild
		TOTAL	700	

Location of research:

Location	*Full street address*
Central Animal Facility	Building F9A, Research Park Drive, Macquarie University, NSW 2109
ASAM	Level 1, F10A, 2 Technology Place, Macquarie University, NSW 2109

Amendments approved by the AEC since initial approval: N/A

Conditions of Approval:
1. *The "Banders Diary Sheets" as required by the Australian Bat and Bird Banding Scheme are to be submitted to the AEC in the annual reports for this protocol.*

Being animal research carried out in accordance with the Code of Practice for a recognised research purpose and in connection with animals (other than exempt animals) that have been obtained from the holder of an animal suppliers licence.

Professor Mark Connor (Chair, Animal Ethics Committee) **Approval Date: 20 March 2014**

Adapted from Form C (issued under part IV of the Animal Research Act, 1985)

New Zealand house sparrow ethics approval letter

OTAGO UNIVERSITY ANIMAL ETHICS COMMITTEE

Hunter Centre
C/o Faculty Office
Medical School

3 November 2008

Dr S Nakagawa
Zoology
University of Otago

Dear Dr Nakagawa

Re: Application 87/08

This is to acknowledge receipt of your application for the use of live animals in the programme/project entitled: *The role of pathogen dynamics in maintaining vertebrate genetic variation - Avian malaria in house sparrows*.

Meetings of the Committee are held the second Wednesday of every month (except January). Your application will be considered for approval at the next meeting.

Yours sincerely

Barbara Lee
Secretary
Otago University Animal Ethics Committee

MACQUARIE University

ANIMAL RESEARCH AUTHORITY (ARA)

AEC Reference No.: 2013/029-8 **Date of Expiry:** 31 July 2015

Full Approval Duration: *20 September 2013 to 31 July 2015 (22 Months)*

This ARA remains in force until the Date of Expiry (unless suspended, cancelled or surrendered) **and will only be renewed upon receipt of a satisfactory Progress Report before expiry (see Approval email for submission details).**

Principal Investigator:
Dr Simon Griffith
Biological Sciences
Macquarie University, NSW 2109
simon.griffith@mq.edu.au
0425 746 674

Associate Investigators:
Kate L Buchanan 0429 398 460
Mark Mainwaring (02) 9850 1302
Andrea Crino 0406 398 817
Larissa Trompf 0419 371 403

Post-doctoral Fellow
Camille Duval 0475 209 328
Postgraduate Student:
Laura (Lori) Hurley 0419 978 077
Volunteer Research Assistant :
Tiare Broadhead 0431 570 967
Renee Borg 0468 622 372

In case of emergency, please contact:
the Principal Investigator / Associate Investigator named above
Animal Welfare Officer - 9850 7758 / 0439 497 383, *or* **Manager, Fauna Park** - 9850 4109 / 0425 213 420

The above-named are authorised by MACQUARIE UNIVERSITY ANIMAL ETHICS COMMITTEE to conduct the following research:
Title of the project: Physiological coordination of reproduction in a socially monogamous bird
Purpose: 4 - Research: Human or Animal Biology
Aims: To understand how male and female zebra finches are able to coordinate their physiology and how optimize their different reproductive attempts
Surgical Procedures category: 3 - Minor Conscious Intervention
All procedures must be performed as per the AEC-approved protocol, unless stated otherwise by the AEC and/or AWO.
Maximum numbers approved (for the Full Approval Duration):

Species	Strain	Sex	Weight	Age	Already Approved	New Request	New total	Supplier/ Source
18	Zebra finch (*Taeniopygia guttata*)	Both	12g	Adult	1080	112	1192	Current stock
					0	72	72	
18	Zebra finch (*Taeniopygia guttata*)	Both	17g	Juvenile	1080	840	1920	Current stock
18	Long Tailed Finch	Both	12g	Adult	0	80	80	Current stock
18	Long Tailed Finch	Both	12g	Juvenile	0	480	480	Current stock
							3744	

Location of research:

Location	*Full street address*
Fauna Park	209 Culloden Road, Marsfield, NSW 2122

Amendments approved by the AEC since initial approval:

1. ***Amendment #1****: Addition of Laura Hurley as Postgraduate Student* ***subject to the Research Assistant attending an Animal Ethics Course*** *(Executive Approved 7 November 2013, AEC ratified 5 December 2013).*
2. ***Amendment #2:*** *Addition of Camille Duval as Post-doctoral fellow* ***subject to attending an Animal Ethics Course*** *(Executive approved 28 May 2014, AEC ratified 19 June 2014)*
3. ***Amendment # 3.1 -*** *Amend current procedure, addition of additional animals, addition of 3 techniques. End point of animals remains the same as original AR. (Approved by AEC meeting on 14 August 2014)A*
4. ***Amendment # 3.2 -*** *Addition of animal type and number. (Approved by AEC meeting on 14 August 2014)*
5. ***Amendment # 3.3 -*** *Addition of number of animals, add a "more intensive" data collecting regime to compare data collected by the concurrently run already approve "less intensive" population in the Large Old Aviaries, already approved less intensive regime. Birds to be added and procedures outlined. Same animal endpoint as original ARA. (Approved by AEC meeting on 14 August 2014)*
6. ***Amendment # 3.4 -*** *Addition of new temperature manipulation study. (Approved by AEC meeting on 14 August 2014)*
7. ***Amendment # 3.5 -*** *Birds to be added, new procedures outlined, additional housing detailed. Change in animal endpoint (Approved by AEC meeting on 14 August 2014)*
8. ***Amendment #4****: Addition of Larissa Trompf as Associate Investigator (Approved by AEC meeting on 14 August 2014)*
9. ***Amendment #5****: Addition of Tiare Broadhead as Volunteer Research Assistant(Approved by Exec, to be ratified by AEC 19 March 2015)*
10. ***Amendment #6****: Addition of Renee Borg as Volunteer Research Assistant (Approved by Exec, to be ratified by AEC 19 March 2015)*

Conditions of Approval:

1. ***Amendment #1:*** *Laura (Lori) Hurley to attend the next WWRAW (Working with Research Animals Workshop) on 25 July 2014.*
2. ***Amendment #2****: Dr Camille Duval to attend the next WWRAW on 25 July 2014 in order to be familiar with NSW legislation and expectations.*
3. ***Amendment 5 & 6,*** *Tiare Broadhead and Renee Borg to be supervised by Principal Investigator only.*

Being animal research carried out in accordance with the Code of Practice for a recognised research purpose and in connection with animals (other than exempt animals) that have been obtained from the holder of an animal suppliers licence

Professor Mark Connor (Chair, Animal Ethics Committee) **Approval Date:** 03 March 2015

Adapted from Form C (issued under part IV of the Animal Research Act, 1985)

Fowlers Gap zebra finch ethics approval letter

MACQUARIE UNIVERSITY ~ SYDNEY

ANIMAL RESEARCH AUTHORITY

AEC Reference No.: 2007/038
Full Approval Duration: 01 January 2008 to 31 December 2010 (36 months)

To: A/Prof Simon Griffiths
CISAB
Macquarie University
Phone: (02) 9850 4186
Email: simon.griffith@mq.edu.au

Associate Investigator (s)
Dr Sarah Pryke Phone: 9850 4187
Ms Barbara Tschirren Phone: 9850 4187

Other people participating
Mr James Brazil-Boast Phone: 9850 4187
Ms Amanda Gilby Phone: 9850 4187

Is authorised by:

MACQUARIE UNIVERSITY to conduct the following research:

Title of the project: LIFE-HISTORY AND BIODIVERSITY IN FOUR ESTRILDID FINCHES: BEHAVIOUR, EVOLUTION AND CONSERVATION — IN THE WILD

Type of animal research and description of project:
Evolutionary and behavioural research - wildlife: The project will use wild finch species to investigate the evolutionary processes related to reproduction and life-history that underlie species diversity in birds. Research includes (1) mapping and characterizing variation of wild populations (individual phenotypes) (2) Breeding ecology and nest competition in the wild. Birds will be captured in the wild using mist nets. Experimental procedures include: banding (ABBBS); colour and morphological measurements; blood sampling; immunocompetence testing using PHA skin test; recording of parental care; sperm sampling; nest monitoring and observation; cross-fostering off-spring between nests. All experimental procedures to be conducted in accordance with details provided in the approved protocol.

Species of animal: Gouldian finch (*Erythrura gouldiae*); Long-tail finch (*Poephila acuticauda*); Black-throated finch (*Poephila cincta*); Zebra finch (*Taeniopygia guttata*)

Number: Gouldian finches: up to 400 per year (total 1200); Long-tail finches: up to 600 per year (total 1800); Black-throated finches: up to 600 per year (total 1800); Zebra finches: up to 1000 per year (total 3000)

Location/s: (1) Conservation land – Wyndham Shire, North Kimberley, WA 6740
(2) Fowler's Gap UNSW Arid Zone Research Station, via Broken Hill NSW 2880

Amendments considered by the AEC during last period: N/A

As approved by and in accordance with the establishment's Animal Ethics Committee.

MACQUARIE UNIVERSITY AEC

Approval was granted subject to compliance with the following conditions:

(This authority has been issued as the above condition (s) has been addressed to the satisfaction of the AEC)

Being animal research carried out in accordance with the Code of Practice for a recognised research purpose and in connection with animals (other than exempt animals) that have been obtained from the holder of an animal suppliers licence.

This authority remains in force from **01 January 2008** to **31 December 2008**, unless suspended, cancelled or surrendered, **and will only be renewed upon receipt of a PROGRESS report at the end of this period.**

Dr Darren Burke
Acting Chair of AEC, Macquarie University

Date: 12-12-07

Form C (issued under part IV of the Animal Research Act, 1985)