Mating system, sex-specific selection and the evolution of the avian sex chromosomes

Alison Elizabeth Wright
Linacre College
University of Oxford

Submitted for the degree of
Doctor of Philosophy, D.Phil
Trinity 2014

Abstract

MATING SYSTEM, SEX-SPECIFIC SELECTION AND THE EVOLUTION OF THE AVIAN SEX CHROMOSOMES

ALISON ELIZABETH WRIGHT

D.PHIL, LINACRE COLLEGE, UNIVERSITY OF OXFORD, TRINITY 2014

Sex chromosomes experience distinct evolutionary environments, due to their unusual pattern of inheritance, and studies of sex chromosome evolution can shed light on the fundamental evolutionary forces acting across the genome as a whole. Here, I combine genomic and transcriptomic data across a wide range of avian species to explore the evolutionary processes governing sex chromosome evolution. Birds are female heterogametic and therefore it is possible, via comparisons with male heterogametic species, to identify the fundamental factors driving sex chromosome evolution, versus those associated with sex.

In this thesis, I uncover a complex mosaic of recombination suppression between the Z and W chromosomes, characterized by repeated and independent divergence of gametologs, together with ongoing genetic exchange. Additionally, I highlight the role of mating system, and interplay between evolutionary forces, in driving coding and expression evolution on the Z and W chromosomes. My findings indicate that although the Z chromosome is masculinized for male-specific effects, the magnitude of genetic drift acting on Z-linked genes is elevated in promiscuous relative to monogamous mating systems. In contrast, evolution of the female-limited W chromosome is governed predominately by purifying selection. Together, my results suggest that the role of the Z chromosome in encoding sexual dimorphisms may be limited, but that W-linked genes play a significant role in female-specific fitness. In conclusion, my findings reveal the power of mating system in shaping broad patterns of genome evolution.

Acknowledgements

I am extremely grateful to my supervisor, Judith Mank, for her constant support and encouragement. Her unique insights and exceptional guidance have been invaluable, and she has created an extremely positive and creative environment to work in. I would like to thank all past and present members of the Mank group for helping to make my D.Phil such an enjoyable experience, and for their comments and contribution to this thesis; Natasha Bloch, Becky Dean, Peter Harrison, Hooman Moghadam, Stephen Montgomery, Jen Perry, Marie Pointer and Fabian Zimmer. I would also like to express my gratitude to Tom Pizzari for kindly stepping into an official advisory role when Judith and I moved to UCL, and Linacre College who provided a Writing-Up Bursary, and Travel Grant. Finally, I am very grateful to Chris Cooney for his support, thought provoking discussions and for sharing my fascination with the natural world. I would like to thank my parents and brother who always make time to listen and advise, and for their interest in, and enthusiasm for my work.

Contents

Figure List

Table List

Chapter 1

The scope and strength of sex-specific selection in genome evolution

1.1 ABSTRACT

Males and females share the vast majority of their genomes, and yet are often subject to different, even conflicting, selection. Genomic and transcriptomic developments have made it possible to assess sex-specific selection at the molecular level, and it is clear that sex-specific selection shapes the evolutionary properties of several genomic characteristics, including transcription, post-transcriptional regulation, imprinting, genome structure and gene sequence. Sex-specific selection is strongly influenced by mating system, which also causes neutral evolutionary changes that affect different regions of the genome in different ways. Here I synthesize theoretical and molecular work in order to provide a cohesive view of the role of sex-specific selection and mating system in genome evolution. I also highlight the need for a combined approach, incorporating both genomic data and experimental phenotypic studies, in order to understand precisely how sex-specific selection drives evolutionary change across the genome.

1.2 INTRODUCTION

The ability to attract mates and reproduce is a central component of Darwinian fitness. As well as primary sexual dimorphisms directly involved in reproduction, including numerous gametic proteins key to syngamy, there are many secondary sexual traits involved in mate acquisition (Andersson 1994; Swanson and Vacquier 2002; Nadeau et al. 2007). Selection related to sex can be a powerful force given that it is a crucial component in Darwinian fitness, and it can act in opposite directions for males and females due to their distinct reproductive roles and biology (Arnqvist and Rowe 2005). Aside from sex-limited regions, such as Y or W chromosomes, contradictory female- and male-specific selection regimes act on a genome that is shared between the sexes.

Sex-specific selection is strongly influenced by mating system (Fig. 1.1), which also affects neutral diversity and evolution. At one end of the spectrum, the potential for conflict between female- and male-specific selection is lowest in monogamous systems. In these species, sex-specific selection often acts primarily on reproductive biology, resulting in very few pronounced secondary sexual dimorphisms (Helfenstein et al. 2004). Promiscuous systems show much more potential for sexual conflict (Arnqvist and Rowe 2005; Lifjeld et al. 1993; Lindenfors et al. 2002). Polyandrous and polygynous species are often characterized by large differences between male and female parental effort and other aspects of life histories. This creates ample scope for sex-specific selection (Fig. 1.1) (Kokko et al 2012). Mating system also influences neutral diversity and rates of evolution. Large differences in the variance in reproductive success between males and females increase the rate of genetic drift, the strength of which varies across the genome depending on the degree and direction of

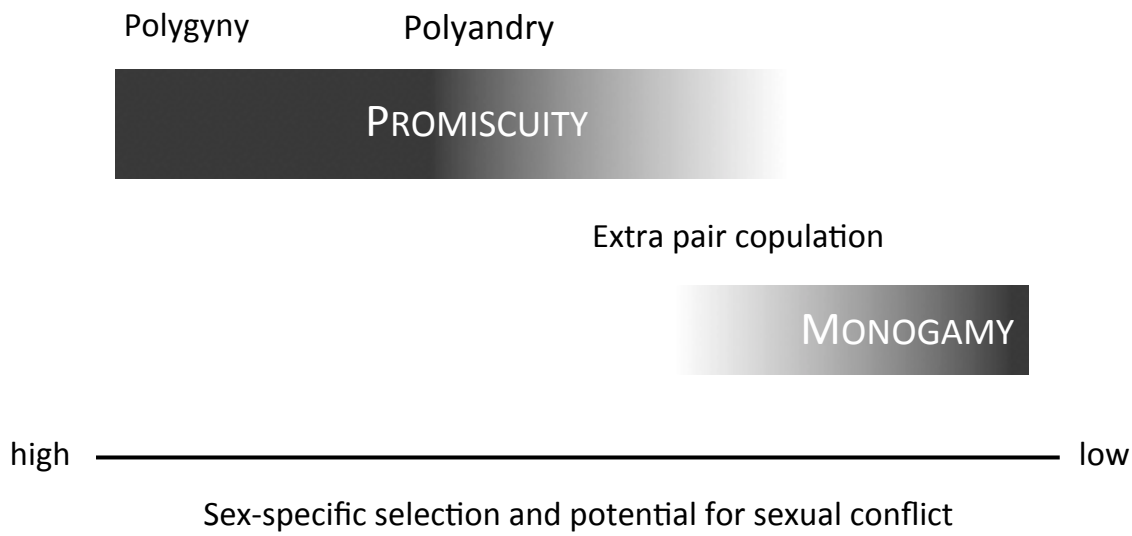

Fig. 1.1 Mating systems and sex-specific selection.

Sex-specific selection is strongly influenced by mating system. Sexual conflict arises when female- and male-specific selection act in opposing directions, and can result in significant fitness costs. The probability that sex-specific selection forces are contradictory is lowest in monogamous systems. However, in mating systems with large differences in reproductive potential, the divergence in male and female fitness optima is greater, as is the probability of contradictory sex-specific selective forces. As a result, the potential for sexual conflict is predicted to increase with the magnitude of difference in mating success. Additionally, polyandrous and polygynous species are often characterized by large differences between male and female parental effort and other aspects of life histories, which creates ample scope for sex-specific selection. Thus, sex-specific selection and sexual conflict may play a more significant role in polygnous and polyandrous mating systems than monogamous systems.

sexual asymmetry in inheritance (Vicoso and Charlesworth 2009; Mank et al. 2010a). Contrasting evolutionary rates and diversity among these regions, most often between the sex chromosomes and the autosomes, can be used to infer the power of neutral evolutionary forces at work in the genome, as well as estimate mating system (Corl and Ellegren 2012).

Conflicting sex-specific selection has the potential to create a significant evolutionary burden on a population (Foerster et al. 2007; Morrow et al. 2008; Connallon et al. 2010), and resolving this conflict, when it is possible at all, both relieves this burden and allows the sexes to reach separate fitness optima (Chapman et al. 2003). Phenotypic-level studies have revealed a significant level of conflict in many organisms (Chippindale et al. 2001; Arnqvist and Rowe 2002; Magurran and Seghers 1994, among many others), affecting many different phenotypes.

Recent genomic studies have made it possible to take studies to the molecular level, and assess the magnitude and locus of sex-specific selection, and to identify mechanisms of resolution for conflicting sex-specific selection (Innocenti and Morrow 2010; Moghadam et al. 2012). Signatures of selection and neutral evolution can be detected at the molecular level using a number of approaches. Firstly, the ratio of rate of nonsynonymous substitutions per nonsynonymous site (d_N) to synonymous substitutions per synonymous site (d_S) can be calculated from coding sequence alignments. $d_N/d_S > 1$ indicates positive selection, $d_N/d_S < 1$ indicates purifying selection and $d_N/d_S = 1$ is indicative of neutral evolution. Secondly, the McDonald-Kreitman test contrasts polymorphism and divergence estimates to distinguish contrasting selective regimes. Under relaxed purifying selection, we expect a relative excess of nonsynonymous polymorphisms ($(D_N/D_S) < (P_N/P_S)$), whereas under positive

selection we expect a relative excess of nonsynonymous divergence ((D_N/D_S) > (P_N/P_S)). However, it is important to note the distinction between positive selection and adaptive evolution, where adaptation refers to phenotypes that enhance the reproductive success of an organism in its natural environment.

Recent molecular studies indicate that sex-specific evolutionary forces, both adaptive and neutral, affect a large proportion of the genome (Connallon et al. 2010) and play a significant role in determining rates of both adaptive and neutral change for gene sequence, gene expression, and genome structure. This is reinforced by the disproportionately large effect that sexually antagonistic loci are predicted to have on genetic variation for fitness (Arnqvist 2011; Long et al. 2012). Additionally, it is clear from molecular studies that there are different routes to resolving conflict between female- and male- specific selection (Connallon 2007; Haig 2009; Gallach and Betran 2011).

Molecular approaches offer the promise of identifying sexually antagonistic genes and alleles, and creating an understanding of how discrete male and female phenotypes are encoded. However, there are a number of important assumptions made when genomic data are used in isolation to infer the nature of sexual conflict and sex-specific selection. One key assumption is that, for the phenotypic traits in question, intra-locus conflict between female- and male- specific selection has been resolved through the evolution of sex-specific gene regulation or some other mechanism to break down inter-sexual correlation. However, inter-locus conflict cannot be resolved by the same mechanisms, but instead may result in a cycling of adaptation and counter-adaptation. Incorporating experimental studies in addition to genomic approaches can shed light on this fundamental assumption, and provide

detailed information we lack on the fitness consequences of many genomic mechanisms thought to resolve sexual conflict (Tregenza et al. 2006).

Despite the potential of using a combined approach incorporating experimental or population-level phenotypic and reproductive studies with molecular data, aside from a few exceptions (Innocenti and Morrow 2010; Moghadam et al. 2012), these methods have proceeded largely independent of each other. My purpose in this chapter is to synthesize recent molecular genomic advances in order to create a cohesive picture of the importance of sex-specific selection in shaping the evolution of various genomic properties. Ultimately, my goal is to understand how sex-specific selection and mating system affect genome evolution through adaptive and neutral processes, and reinforce the need for a combined approach, incorporating both experimental phenotypic and molecular studies, to create a cohesive understanding of sex-specific selection, its fitness consequences, genomic targets and mechanisms of resolution.

1.3 SEX-SPECIFIC SELECTION AND ADAPTIVE CHANGE

Most molecular genetic analyses of conflicting sex-specific selection focus on intra-locus conflict (conflict where an allele at a given loci is beneficial to one sex but detrimental to the other), where opposite female- and male-selection acts on the same locus. Intra-locus conflict can be resolved by a number of genetic mechanisms, the most studied of which is sex-biased expression (Connallon and Knowles 2005; Gallach and Betran 2011), which represents the breakdown in inter-sexual correlation in gene regulation. Additionally, duplication of genes with sex-specific functions (Gallach et al. 2010) may also be important in resolving conflict. In either case, sex-

biased or sex-specific gene expression is used as a signature of at least partially resolved conflict, as the breakdown of correlation between male and female transcription allows for sex-specific fitness optima (Connallon and Knowles 2005; Mank 2009a). In contrast, it is more difficult to detect inter-locus conflict (conflict between alleles at different loci, where if one allele is beneficial to males and detrimental to females, the other allele displays a reverse fitness effect) based on molecular data alone, as there is no expected resolution through the breakdown in inter-sexual correlation. However, imprinting and parent-of-origin expression may indicate sites linked in inter-locus conflict (Haig 2009; Gregg et al. 2010a; Gregg et al. 2010b), as well as allelic imbalance in expression, although with caveats. The genetic mechanisms by which sexual conflict can be resolved are discussed in further detail below.

1.3.1 Types of genes influenced by sex-specific selection

Initial studies examining the influence of sex-specific selection at the genetic level focused on known reproductive genes (Swanson and Vacquier 2002), which evolve rapidly across a diverse range of taxa. Many of the genetic changes have been shown to be adaptive (Swanson et al. 2001), with one of the best examples perhaps being the evolution of male accessory gland proteins (Acps) related to sperm competition in *Drosophila*. These proteins are present in seminal fluid, and act to increase male mating success by promoting ovulation, reducing female receptivity to re-mating and promoting sperm storage (Wolfner 2002). Acp loci in *Drosophila* are 50% more divergent than non-reproductive proteins, with many exhibiting rapid turnover between species and high rates of functional change indicative of positive

selection (Swanson et al. 2001; Begun and Lindfors 2005). This is a broad pattern, as similar signatures of adaptive change have been found in a wider group of reproduction related genes, including seminal fluid proteins, in *Drosophila* (Haerty et al. 2007), primates (Clark and Swanson 2005) and rodents (Turner et al. 2008), as well as gamete recognition proteins in marine invertebrates (Metz and Palumbi 1996; Palumbi 1999). The rapid evolutionary rates of male reproduction related genes suggest post-copulatory selection is important in driving adaptive divergence. Indeed, sperm competitive ability in *Drosophila* has been directly linked to polymorphism in certain male reproductive genes (Fiumera et al. 2005). Presumably, increasing competition between males for mating increases the intensity of sperm competition, thereby resulting in higher rates of evolution for male reproduction related genes (Walters and Harrison 2011). Reproductive character displacement may also contribute to the rapid divergence of male reproductive genes (Matute 2010) as well as frequency dependent selection (Clark et al. 1999).

More recently, female reproductive proteins in *Drosophila* have been shown to undergo similarly high rates of functional change (Swanson et al. 2004; Panhuis and Swanson 2006), possibly due to the conflict between the sexes over polyspermy. Sperm competition generates selection on males to increase the speed of fertilization. Increased sperm fertilization rate can result in a cost to females through the elevated possibility of multiple sperm penetrating the ovum, which generally results in lethality. The major cost to females of egg inviability generates female-specific selection to slow down fertilisation. ZP3, a protein found within the egg coat, is responsible for binding to sperm and thus facilitating fertilisation, and there is evidence in mammals of positive selection within the gene region responsible for sperm binding (Swanson et al.

2001). Similar results have been shown in birds (Calkins et al. 2007; Berlin et al. 2008), but there is some debate as to whether selection against polyspermy is responsible for driving this adaptive change, as birds may be more tolerant of polyspermy than other animals (Birkhead and Fletcher 1998; Stepinska and Bakst 2007; Tarin and Cano 2000; Wishart 1987). Instead, cryptic female choice for specific male sperm types may be responsible for the high rates of functional change seen at sperm binding regions of some egg coat proteins (Calkins et al. 2007; Berlin et al. 2008).

Male and female fitness is not solely reliant on reproductive proteins, and secondary sexual characters can play an important role in mate acquisition. As expected, there is evidence of adaptive change as a result of sex-specific selection acting on these characters. Male plumage colour in Galliformes is highly diverse and shown to be involved in female mate choice and between male-to-male competition. The MC1R locus, which contributes to plumage pigmentation, has been shown to undergo high rates of functional change (Nadeau et al. 2007). Additionally, this rate correlates with the degree of sexual dichromatism exhibited across Galliform species, demonstrating that the MC1R locus is a target for sex-specific selection acting on plumage colouration. However, many somatic dimorphisms are complex aggregates of many loci, and this complexity requires fundamentally different approaches, explained below.

1.3.2 Sex-biased expression

More complex sexual phenotypes composed of dozens to hundreds of loci can be examined at the genomic scale with transcriptome data. The majority of polygenic sexual dimorphisms result from differences in gene expression between males and

females, and this sex-biased expression is the product of the breakdown in inter-sexual correlation in expression, and therefore represent loci where inter-locus conflict has been at least partially resolved (Connallon and Knowles 2005). Additionally, rates of evolution of sex-biased genes, measured in aggregate, may offer insight into the relative strength of male- versus female-specific selection.

In adults, the differences between males and females in expression are prevalent in the transcriptomes of many animals, including *Drosophila* (Jin et al. 2001; Ranz et al. 2003), mouse (Yang et al. 2006), *Anopheles* mosquitoes (Marinotti et al. 2006), birds (Mank et al. 2008a; Naurin et al. 2011), *C. elegans* (Cutter and Ward 2005) and ants (Ometto et al. 2010). Some of the pattern of sex-bias is condition-dependent (Wyman et al. 2010), consistent with predictions about some types of sexually-selected traits. Additionally, many of the differences in gene expression between the sexes are thought to arise from androgen- or oestrogen- mediated regulation (Zauner et al. 2003). Changes in regulation have produced large variation in the proportion of the transcriptome showing sex-bias among species (Zhang et al. 2007), as well as variation in sex-bias among populations within species (Muller et al. 2011; Moghadam et al. 2012).

Characterizing the degree of sex-biased gene expression at the tissue level can provide insight into the phenotypic traits subject to the greatest degree of sex-specific selection. As expected from studies on reproductive proteins, sex-biased expression is most evident in gonad transcriptomes (Parisi et al. 2004; Rinn et al. 2004; Mank et al. 2008a). However, sex-biased gene expression is observed across a large majority of somatic tissues in many animals (Yang et al. 2006; Mank et al. 2008a), particularly in the liver. Evidence from mammals and birds suggest that the degree of sex-biased

expression is lowest in the brain (Yang et al. 2006; Mank et al. 2008a; Reinius et al. 2008). However, these studies tend to examine the brain as a whole and thus a more detailed examination of specific areas of the brain may reveal a more obvious pattern of sex bias (Naurin et al. 2011).

It is possible to estimate the relative strength of sex-specific selection at the molecular level by estimating rates of divergence for sex-biased genes. Numerous studies have found evidence of accelerated rates of evolution, particularly due to positive selection, in adult male-biased genes across a range of species (Good and Nachman 2005; Khaitovich et al. 2005; Ellegren and Parsch 2007). In addition to divergence in coding sequence, expression divergence is more pronounced for male-biased genes in *Drosophila* than female-biased genes (Meiklejohn et al. 2003; Llopart 2012). However, there are exceptions to the higher rates of evolution seen in male-biased compared to female-biased genes. Some studies have found either no difference in divergence patterns between the sexes (Metta et al. 2006), potentially linked to a reduction in sexual selection, or the reverse pattern (Mank et al. 2010b). The latter study highlights the shifting nature of sex-specific selection throughout development, as female-biased genes were found to have higher rates of adaptive change when measured in the gonad during embryonic development compared to rates in adults.

Genetic architecture, the number and type of loci underlying a given trait, plays an important role in determining the genetic and phenotypic outcome of sex-specific selection and evolution of sex-biased expression. The involvement of a single gene in the development of multiple traits, pleiotropy, determines the extent to which intra-locus sexual conflict can be resolved and thus the ability of sex-specific selection to

facilitate evolutionary change (Harano et al. 2010). Less pleiotropic genes, measured as a function of tissue specificity, exhibit sex-biased gene expression and faster rates of evolutionary change compared to pleiotropic genes (Mank et al. 2008b; Meisel 2011). This may suggest that pleiotropy hinders the breakdown in inter-sexual correlation in expression, and therefore the capacity of the genome to respond to sex-specific selection, reinforcing the widely acknowledged role of pleiotropy as an evolutionary constraint (Fisher 1930; Orr 2000; Snell-Rood et al 2010). Thus, the effect of selection on gene expression will not be the same between genes under different architectural constraints. Interestingly, female-biased genes display greater pleiotropic effects than male-biased genes, potentially contributing to the different rates of evolution between the two classes of genes (Assis et al. 2012).

Surveys of species and population variation in expression, such as those cited above, provide a long-term evolutionary view of sex-specific selection and sexual conflict. Using sex-bias expression data to infer the targets and strength of sex-specific selection relies on a number of assumptions. First, the relationship assumes that sex-biased genes encode sex-specific phenotypes and often have sex-specific fitness effects. It is difficult to test this assumption with species and population transcriptome data alone, as these data cannot directly connect large aggregates of genes with differential expression to concrete phenotypes, although there is some empirical support (Connallon and Clark 2011a). Second, the relationship assumes that sexual conflict can be at least partially resolved via the breakdown of inter-sexual expression correlation, therefore pleiotropic constraints may mask loci that are subject to sexual conflict.

Studies that combine molecular and phenotypic approaches are just now appearing, to date only two have been published to my knowledge (Innocenti and Morrow 2010; Moghadam et al. 2012) and these have the added power of being able to measure sex-specific fitness effects and therefore estimate sex-specific phenotypic optima. In some cases, these sorts of studies fail to detect the predicted relationship between sex-biased expression and sex-specific fitness, possibly due to insufficient variation in sex-biased regulatory variation within study populations. If this is a general trend, it may suggest that short-term sex-specific regulatory changes are rare (Innocenti and Morrow 2010). Other work is somewhat contradictory, and shows that changes in sex-specific selection over short evolutionary histories causes substantial changes in sex-biased expression, and that this has sex-specific fitness consequences (Moghadam et al. 2012). Further experimental evolution studies are planned or in progress, and will no doubt provide exciting developments to this debate.

1.3.3 Sex-limited expression and imprinting

The negative fitness consequences of sexual conflict (Morrow et al. 2008) can be avoided if antagonistic genes are sex-limited, with expression completely restricted to either males or females. Sex-limited expression, where a gene is expressed in only one sex, is somewhat different from sex-biased expression, where a gene is expressed in both females and males, but at different levels. True sex-limited expression is relatively rare for genes not linked to the sex-limited Y or W chromosomes. This may be because selection against expression in one sex decreases as expression level wanes for most loci, creating a saturation point at which sex-specific selection is not strong enough to further reduce expression. However, the point at which sex-biased

expression becomes functionally sex-limited is somewhat arguable, as very low levels of expression in one sex are likely to have little phenotypic effect. This may suggest that most very sex-biased genes are functionally sex-limited.

Sex-limited genes avoid fitness penalties in the sex lacking expression (Rice 1984). Genomic imprinting is one potential mechanism to achieve sex-limited expression without the need to breakdown inter-sexual regulatory mechanisms. For an imprinted allele, expression depends upon the parent of origin, and studies have suggested that imprinting can both resolve intra-locus conflict and exacerbate inter-locus conflict (Day and Bonduriansky 2004; Swaney et al. 2007; Van Cleve and Feldman 2007; Haig 2009).

The resolution of intra-locus conflict by the evolution of imprinting has been modelled under a wide range of selection and dominance parameters (Day and Bonduriansky 2004; Van Cleve and Feldman 2007; Patten and Haig 2008). Negative mother-son fitness correlations can result in selection for invading modifier loci that silence sexually antagonistic maternally inherited alleles in males. In support of these predictions, sex-specific differences in imprinting effects of loci contributing to body size have been shown in mice (Hager et al. 2008).

For the majority of imprinted loci however, the phenotypic effect is poorly understood, and it is possible that modelling sexual antagonism at only one locus is unrealistic. Instead, sexual conflict between interacting genes may drive the evolution of imprinting. The insulin-like growth factor 2 (Igf2) and insulin-like growth factor 2 receptor (Igf2r) are imprinted in many mammals and heavily involved in the regulation of growth (Day and Bonduriansky 2004). This imprinting pattern of Igf2 and Igf2r is thought to be the product of conflict over maternal and paternal reproductive

interests as they play out through the offspring. The expression of Igf2 is paternally inherited and associated with increased growth rates, and a mutation in this gene causes the poor growth associated with Silver-Russell syndrome (Obermann et al. 2004). The paternal allele of Igf2 works in concert with Igf2r, which is maternally imprinted and is responsible for degrading Igf2 (Wilkins and Haig 2003) and limiting nutrient extraction from the mother. In addition to Igf2 and Igf2r, inter-locus conflict has been implicated in the evolution of imprinting for other developmental traits such as suckling intensity and age of sexual maturation (Haig 2009) and presumably as the phenotype of more imprinted genes is examined, the exact role of inter-locus conflict will become apparent.

Recent genome-wide assessments of genomic imprinting based on parent-of-origin expression imbalanced in the offspring (Gregg et al. 2010a; Gregg et al. 2010b) indicated that a large proportion of the genome exhibits some form of imprinting, and that partial imprinting may be common for many genes, or even parts of genes. However, for many loci, the phenotypic and evolutionary effects of imprinting are unclear and there are recent concerns regarding the accurate identification of imprinted loci (DeVeale et al. 2012). Further work characterising the fitness consequences of imprinting and ascertaining whether they exhibit a sexually antagonistic pattern is key for cementing the relationship between imprinting and sexual conflict.

Despite the controversy of genome-wide evidence of imprinting, evidence from the studies referenced above suggest that imprinted loci are beacons of sexual conflict, and careful examination of numbers and rates of evolution of imprinted genes will provide insight into the strength of sexual conflict arising from sex-specific

selection. However, as both inter and intra-locus sexual conflict increase, the number of imprinted loci is predicted to also increase, and further work on a larger number of sexually antagonistic imprinted loci over a range of mating systems can be used to explore this. Additionally, there is very little information about how parent-of-origin imprinting varies within and among populations, and how this regulatory mechanism responds to sex-specific selection in an experimental evolution framework.

1.3.4 Post-transcriptional regulation

In addition to transcriptional differences between the sexes, post-transcriptional mechanisms also differ. Alternative splicing is a gene regulatory mechanism that produces multiple distinct transcripts from one locus, thereby increasing transcriptomic complexity without unduly adding to genome size. Sex-specific splicing is a key component of sex-specific phenotypes, such as sex determination in *Drosophila* (Telonis-Scott et al. 2009), and therefore may be a general mechanism to resolve sexual conflict over gene function. Comprehensive attempts have recently been made to quantify the exact extent of sex-specific splicing throughout the genome, and have shown significant conservation of sex-specific splice variants (Blekhman et al. 2009; Telonis-Scott et al. 2009; Prince et al. 2010).

It is easy to imagine the potential for sex-specific splicing, both in response to sex-specific selection and a route to resolve sexual conflict. Studies of alternative splicing evolution have suggested that splice variants are adaptive, and that they allow for increasing phenotypic complexity without the need to increase genome size (Barbosa-Morais et al. 2012; Merkin et al. 2012). However, aside from the role of sex-specific splicing in sex determination in insects, little is known at this point about the

exact role of alternate splicing in sexually dimorphic phenotypes, or how splice variants evolve in response to sex-specific selection.

In addition to alternative splicing, sex-specific post-transcriptional regulation can be achieved via small non-coding mRNAs. Small non-coding mRNAs are thought to play an important role in development (Stefani and Slack 2008) via both transcriptional and post-transcriptional regulation (Engels and Hutvagner 2006). In a recent study of ten candidate small non-coding mRNAs in two *Drosophila* species, three were found to exhibit a highly male-biased expression pattern (Jiang et al. 2011). The extent of sex bias was large, with one candidate expressed 40 times more in males than females. It has yet to be verified how these differences translate to the level of the phenotype, but a role in sex-specific RNA regulation seems possible. Further work on this area is needed to explore how these small non-coding RNAs regulate gene networks and phenotypic change, before inferences about the strength of sex-specific selection can be made.

1.3.5 Gene duplication

Sexual conflict at a specific locus can also be resolved via gene duplication. The duplication of a sexually antagonistic locus generates paralogs, which provide the potential for sex-specific neo- or sub-functionalization. Sex-specific functionalization can act on both paralogs to theoretically generate male and female beneficial duplicates, which would eventually exhibit sex-biased or sex-limited expression (Connallon and Clark 2011; Gallach and Betran 2011). Alternatively, one paralog may maintain its original function, particularly when it exhibits pleiotropic functions, and

the resolution of sexual conflict can be achieved through sex-biased expression of the other copy.

Evidence indicates that selection to resolve sexual conflict acts to retain duplicates within the genome, and the larger the degree of conflict, the stronger the selection to resolve this conflict via the retention of a resolving duplicate. Nuclear-encoded mitochondrial genes are thought to be under sexually antagonistic selection for the rate of energy production. High rates increase the fitness of sperm but also result in a higher mutation rate, which is disadvantageous for female reproductive tissues. It has been shown amongst relocated duplicate genes of this class, that a large number exhibit testis-specific expression. This duplication and subsequent sex-limited expression decouples the sexes divergent fitness optima and allows resolution of the conflict over energy production (Gallach et al. 2010; Gallach and Betran 2011). Recent evidence suggests that duplications are associated primarily with the evolution of male-biased not female-biased expression (Wyman et al 2012). However, further experimental studies across organisms with different mating systems are necessary to determine whether this pattern is due to variation in the strength of sex-specific selection, or simply a function of the higher rate of male meiosis, and therefore origin of gene duplicates, associated with continuous sperm production.

1.3.6 Beyond the autosomes

Sex-biased or sex-limited inheritance, exhibited by the homogametic sex chromosomes, the heterogametic sex chromosomes, and extra-nuclear genomes carried by mitochondria and chloroplasts, influences the effect of sex-specific selection. Contrasting the evolutionary signatures on sex chromosomes in particular

with the autosomes is increasingly useful for uncovering the strength of sex-specific selection (Rice 1984; Dean et al. 2012) (Fig. 1.2).

Sexual conflict and the homogametic X and Z chromosomes, are linked in several ways. First, the unique sex-specific selection pressures foster the non-random patterns of gene traffic to and from sex chromosomes for loci with sex-specific benefits (Zhang et al. 2010a; Zhang et al. 2010b). In addition to gene movement, sex-biased inheritance of the homogametic sex chromosomes suggests that they play a disproportionately large role in the evolution of sexual dimorphism via intra-locus sexual conflict (Rice 1984). The theoretical prediction that the X chromosome is both feminized and demasculinized, and the Z chromosome is masculinized and defeminized, is supported by numerous genomic analyses in animals (reviewed in Mank 2009a), as well as plants (Spigler et al. 2011). Finally, sexual conflict may also foster the origin and/or turnover of sex chromosomes. In many clades, there is a high rate of sex chromosome turn-over, which has been linked to sexual conflict (van Doorn and Kirkpatrick 2007; van Doorn and Kirkpatrick 2010). The theoretical link has been supported by direct empirical evidence in cichlids, with the invasion of a novel sex determining locus (Roberts et al. 2009) as well as sticklebacks, with the origin of a neo-sex chromosome that contains loci for male courtship traits (Kitano et al. 2009). This suggests a complex relationship between conflict and sex chromosomes, involving gene movement, the origin of new sex chromosomes, or gene expression changes on existing sex chromosomes.

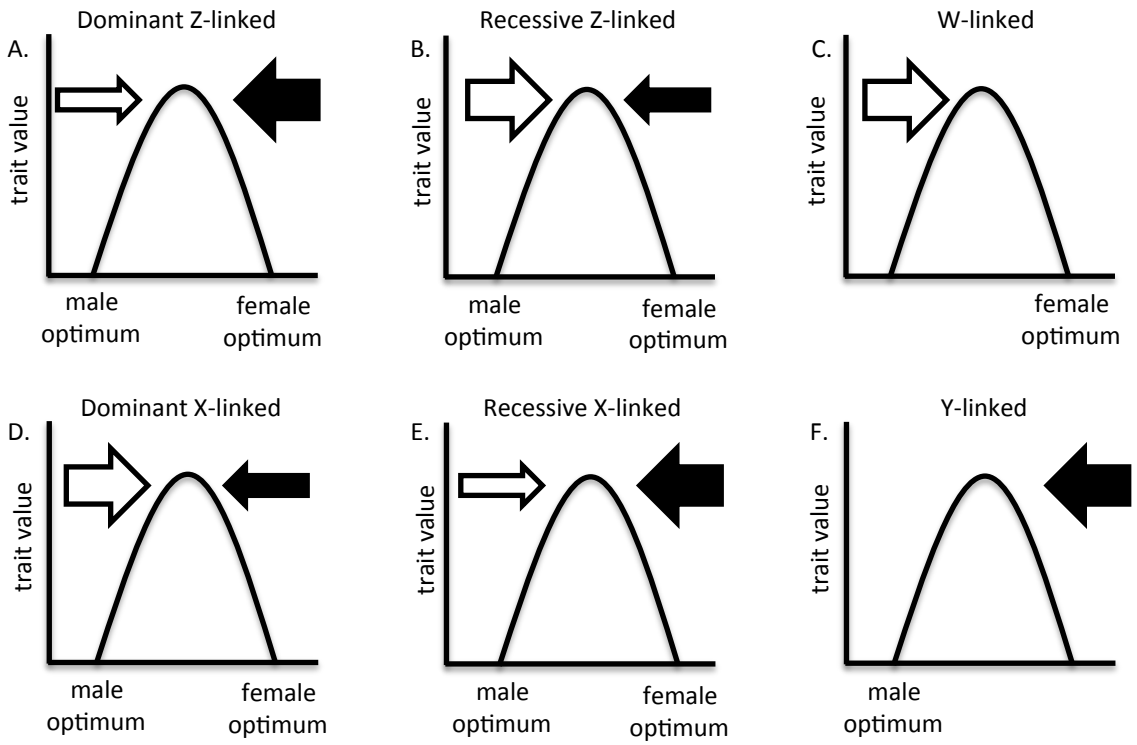

Fig. 1.2. Relative strength of sex-specific selection acting on different sex chromosomes.

Relative sex-specific selection is shown by arrow size, white arrows represent female-specific selection, black arrows represent selection towards male-specific optima. For dominant Z-linked alleles (panel A), male-specific selection is relatively stronger due to the fact that the Z is present more often in males than females. Recessive Z-linked alleles (panel B) are more often exposed in females to selection due to female hemizygosity, therefore female-specific selection is relatively stronger. W-linked genes (panel C) are only selected for female-specific effects. Dominant X-linked alleles (panel D) are more often selected for female-specific effects because the X is more often present in females, while recessive X-linked alleles (panel E) are more often exposed in males due to male hemizygosity. Y-linked genes (panel F) are only selected for male-specific effects.

Mating systems may also influence the degeneration rate of the heterogametic Y and W chromosomes. For sex chromosomes that evolve from existing autosomes, linkage between a sex determining locus and a nearby locus with sex-specific effects will result in selection to suppress recombination between the heterogametic and homogametic sex chromosome. As recombination suppression spreads across the heterogametic sex chromosomes, the Y and W chromosome coding content degrades by neutral processes (Charlesworth 2008). Additionally, the complete lack of recombination of the heterogametic sex chromosome promotes degradation by background selection, hitchhiking and Hill-Robertson effects (where linkage between alleles under selection, arising in different genetic backgrounds, reduces the efficacy of selection). However, degeneration may be halted by intra- or inter-chromosomal recombination. Homologous recombination can result in the unidirectional transfer of genetic material between homologous sequences, termed gene conversion (Chen et al. 2007), and has been documented between X and Y chromosomes in mammals (Iwase et al. 2010; Trombetta et al. 2014). For relatively stable genes with low levels of gene translocation, the likelihood that a new sex determining gene will be located proximate to a sexually antagonistic locus is roughly predicted by the proportion of loci in the genome that produce a sex-specific benefit. As mating system is one of the determinants controlling the proportion of the genome subject to sexually antagonistic selection, we might expect mating systems with high levels of conflict to result in faster spread of recombination suppression and therefore heterogametic sex chromosomes degeneration (Charlesworth and Mank 2010).

Sex chromosomes are also subject to selection for dosage compensation (Charlesworth 1996). The degeneration of sex-limited sequences results in an

imbalance in gene dose between males and females, and studies have shown that halving gene dose results in an expression fold-change of 1.5 (Pollack et al. 2002; Pessia et al. 2012). Correspondingly, the negative fitness effects of this expression imbalance for dosage sensitive genes are thought to select for the evolution of compensatory gene regulation mechanisms (Pessia et al. 2012; Ohno et al. 1967). These act to compensate gene dose in the heterogametic sex between the homogametic sex chromosome (X in XY males, Z in ZW females) and the autosomes. However, there is considerable variation in completeness of dosage compensation mechanisms across species (Mank and Ellegren 2009a; Deng et al. 2011; Vicoso and Bachtrog 2011; Julien et al. 2012; Lin et al. 2012; Muyle et al. 2012; Pessia et al. 2012; Vicoso et al. 2013a).

Similar to the sex chromosomes, extra-nuclear genomes show sex-biased transmission patterns that influence the pattern of sex-specific selection. Mitochondria, although present and essential to both males and females, are only transmitted through the matriline and therefore mitochondrial genomes are predicted to accumulate alleles beneficial to females but detrimental to males (Frank and Hurst 1996). A recent study supports this prediction, showing that the maternal transmission pattern results in a sieve that allows deleterious effects to persist if they are limited to males (Innocenti et al. 2011). This sex-specific sieve may act on other uniparentally inherited organelles, such as chloroplasts, though the sieve effect likely varies greatly between dioecious and monecious species.

1.4 NEUTRAL SEX-SPECIFIC PATTERNS

Identifying signatures of genetic drift at the genetic level provides insight into the strength of neutral evolution. Mating systems define the variance in reproductive success between the sexes and thus the transmission of genetic material to subsequent generations, and the effect of mating system on the direction and rate of transmissions differs among regions of the genome.

Effective population size (N_E) is the number of individuals in a population who contribute offspring to the next generation, and is calculated as the probability that two copies of a gene will be sampled and present in the next generation (Hartl and Clark 2007). The effective population size is therefore an important factor influencing the efficacy of selection, where the power of selection to purge mildly deleterious alleles is reduced in populations with lower N_E. Although males and females share the autosomal portion of their genome equally, there is a pronounced asymmetry in the inheritance of the X chromosome (more often present in females), the Y chromosome (male-limited), the Z chromosome (more often present in males) and the W chromosome (female-limited). These inheritance patterns mean that different regions of the genome differ from each other in both their absolute effective populations size (N_E), as well as their response in N_E to differences in male and female mating success (Fig. 1.3).

Drift for autosomal loci is minimized in monogamous species compared to other mating systems, and deviations from monogamy will lead to elevated rates of genetic drift and decrease the efficacy of selection across the genome as a whole (Hartl and Clark 2007). However, the relationship between drift and selection plays out differently on the sex chromosomes. The effective population size of both the X and Z

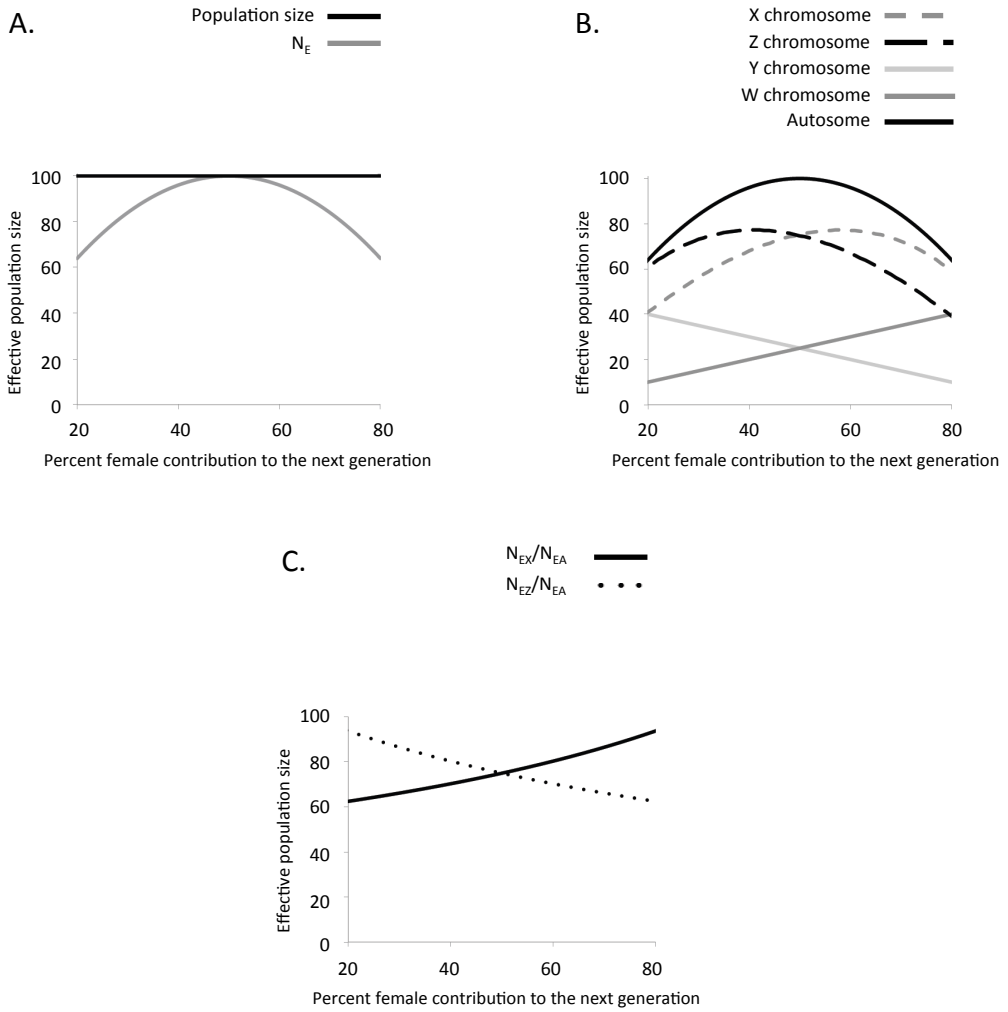

Fig. 1.3. Mating systems and effective population size (N_E).

Increasing differences between male and female reproductive success reduces N_E (panel A), despite a constant overall population size. This difference between the sexes in reproductive success influences the N_E of different portions of the genome in different ways (panel B). For autosomal genes, the largest effective population size (N_{EA}) is seen when the variance in male and female reproductive success is equal. However, N_{EX} and N_{EW} increase with a greater proportion of females than males contributing to the next generation. The opposite is seen for N_{EZ} and N_{EY}, which are maximized when there are more males than females in the reproductive pool. This difference in the effect of mating system on the effective population size of different chromosomes makes contrasts between sex chromosomes and autosomes revealing (panel C).

chromosomes (N_{EX} and N_{EZ}) = ¾ N_{EA}, creates a potential for genetic drift to act on homogametic sex chromosomes (Charlesworth et al. 1993; Vicoso and Charlesworth 2009). This is termed Faster-X and Faster-Z Evolution (Mank et al. 2010a). Increased variance in male reproductive success associated with most forms of sexual selection increases N_{EX} and decreases N_{EZ} relative to N_{EA} (Fig. 1.3B and 1.3C), therefore sexual selection on males is predicted to increase Faster-Z more than Faster-X. This is supported by some evidence from birds (Mank et al. 2007a; Mank et al. 2010c), mammals (Lau et al. 2009) and *Drosophila* (Connallon 2007; Singh et al. 2007; Baines et al. 2008). However, other data are discordant with the role of mating system and sex chromosome evolution (Haddrill et al. 2010).

These predictions for Faster-X and Faster-Z Evolution are slightly altered under female promiscuity, primarily because the degree of variation in female reproductive success seen in polyandry is predicted to be less than the variation in male reproductive success seen in polygyny (Liker et al. 2001). Additionally, although female promiscuity can increase male-specific selection, it can also erode variance in male mating success (Collet et al. 2012).

The exact relationship between mating system and the strength of Faster-X or Faster-Z Evolution is complicated by differences in the rate of recombination on the sex chromosomes, particularly the absence of recombination in *Drosophila* males, which raises the N_{EX} to near N_{EA}, thus reducing the strength of drift (Connallon 2007; Vicoso and Charlesworth 2009). It is important to make clear that variance in male and female reproductive success affects all portions of the genome. However, contrasting the rates of neutral evolution for different portions makes it possible to quantify the effect of mating system on genetic drift.

Mating system also affects the rates of evolution for the heterogametic W and Y chromosomes. N_{EW} and N_{EY} are both equal to ¼ N_{EA} under monogamy, but sexual selection acting on males will decrease N_{EY} and increase N_{EW} compared to N_{EA}. Although there are a host of factors affecting the evolution of heterogametic sex chromosomes, (reviewed in Charlesworth 2008), the difference in N_{EY} and N_{EW} under sexual selection may accelerate the rate of degeneration of Y chromosomes compared to W chromosomes (Bachtrog et al. 2011). In addition to decay, changes in mating system may also result in accelerated rates of change and structural rearrangement in heterogametic sex chromosomes. Consistent with this, the Y chromosome is highly conserved between human, gorilla (Goto et al. 2009) and rhesus macaque (Hughes et al. 2012) but has shown rapid change in the intermediary chimpanzee lineage (Wilson and Makova 2009a; Hughes et al. 2010), and potential explanations for this rest on the promiscuous mating system observed in chimps. However, although the primate example fits the theoretical predictions, it is also anecdotal, and it is not yet possible to robustly test the relationship between mating system and Y chromosome degeneration. Additionally, the exact relationship between heterogametic sex chromosomes and mating system is complicated by several factors, including differential rates of gene acquisition (Koerich et al. 2008) and intra-chromosomal recombination (Lange et al. 2009). Additionally, in some cases, recombination between the homogametic and heterogametic sex chromosome occurs in sex-reversed individuals, (Stöck et al. 2011) which may act to prevent Y chromosome decay (Perrin 2009).

1.5 CONCLUSIONS

Mating system can have profound effects on both neutral and adaptive genome evolution, and can also foster change in gene sequence, expression and post-translational modification for a large proportion of loci. At this point, the theoretical predictions linking mating system to genome evolution have been supported by many anecdotal species-specific studies. By utilising recent advances in molecular genomics, many of these studies provide a long-term overview of the nature of sex-specific selection acting across thousands of genes over millions of years. However, this approach relies on a number of assumptions, namely that conflict is resolvable, which may not be the case for many loci. We also lack detailed information on the sex-specific fitness effects of certain genomic mechanisms thought to be shaped by sex-specific selection. Phenotypic studies can provide much needed information on sex-specific fitness optima and the consequences of deviating from these optima. I therefore emphasize the need for a combined approach, incorporating phenotypic studies with genomic data in order to examine the nature of sexual conflict currently acting on the genome and the fitness consequences of the genomic mechanisms thought to resolve sexual conflict.

Furthermore, recent developments in next generation sequencing methods now facilitate studies utilising the variation in mating system across a wide range of species. Currently, we lack a holistic, robust and quantitative understanding of the relationship between mating system and genome evolution and a cohesive picture of how the myriad of sex-specific forces integrate with one another. This robust, quantitative and cohesive understanding requires comprehensive studies in clades with a range of mating systems. Although model organisms can be useful, the most

relevant behavioural ecologies for this type of analysis are often associated with non-model organisms. A comparative approach across a range of mating systems is much needed to shed further light on the nature of sex-specific selection acting on the genome.

1.6 THESIS OUTLINE

Here, I combine genomic and transcriptomic data across a wide range of avian species to explore the evolutionary processes governing sex chromosome evolution. Specifically, I examine the role of mating system in driving adaptive and neutral aspects of coding sequence and expression evolution on the avian sex chromosomes.

In Chapter 2, I firstly characterise the evolutionary history of the *Gallus gallus* Z chromosome using newly identified gametologs, to reveal a stepwise process of sex chromosome divergence, with four recombination suppression events over 130 million years. The four regions of the Z chromosome have been sex-linked, and therefore subject to masculinizing selection, for different lengths of time. This has resulted in nearly clocklike increase in male-biased expression, with older regions significantly more male-biased than younger regions of the Z chromosomes (Wright et al. 2012).

In Chapter 3, I build on the findings of Chapter 2 in order to characterise the short- and long-term dynamics of sex chromosome divergence. To do this, I combine transcriptomic and sequence based approaches to identify the largest catalog of W-linked genes assembled to date, spanning three species and 90 million years of avian evolution. Using these gametologs, I map the dynamics of recombination suppression across the avian Z chromosome. Divergence and gene convergence tests indicate that individual loci show evidence of residual inter-chromosomal genetic exchange. Finally,

I show that the female-limited W chromosome evolves primarily under purifying selection (Wright et al. In press).

In Chapter 4, I examined the role of mating system in driving Z chromosome divergence and genetic diversity, and measure the relative contribution of neutral and adaptive forces. To do so, I combined expression, sequence and polymorphism data with measures of sperm competition and promiscuity, across a clade of birds spanning 90 million years of evolutionary history. The results show that variance in male mating success can increase rates of neutral evolution on the Z chromosome, leading to elevated rates of genetic drift and the fixation of mildly deleterious alleles. This may limit the adaptive role of the Z chromosome in general, and in the evolution of sexually selected traits in particular (Wright et al. In review).

In Chapter 5, I discuss my findings in a wider context in order to explore the fundamental features of sex chromosome evolution. Of particular focus are the relative roles of selection versus genetic drift in driving Z and W chromosome evolution, and their respective contributions to sex-specific fitness.

1.7 GLOSSARY

Dosage compensation – Sex chromosome gene regulation mechanism in the heterogametic sex to correct for gene dose differences between the homogametic sex chromosome (X in XY males and Z in ZW females) and the autosomes. A side effect of this mechanism is that gene dose is equalised between males and females.

Effective population size – The number of individuals that contribute to the next generation.

Faster-X Effect (Faster-Z Effect) – Higher ratio of nonsynonymous to synonymous substitutions for X- and Z-linked genes compared to autosomal genes, due to a combination of hemizygous exposure of mutations in the heterogametic sex and lowered effective population size.

Gametologs - Orthologous genes on the Z and W (or X and Y) chromosomes which have diverged from the ancestral autosomal gene.

Gene conversion – A mechanism of homologous recombination, initiated by DNA double strand breaks, that results in the unidirectional transfer of genetic material between homologous sequences.

Hemizygosity – Having unpaired chromosomes in a diploid cell. The heterogametic sex (males in XY systems and females in ZW species) has only one copy of each sex chromosome, and is therefore hemizygous for the sex chromosomes.

Heterogametic sex chromosomes – The sex chromosomes that are sex-limited in the heterogametic sex, namely the Y in male heterogamety and the W in female heterogamety.

Homogametic sex chromosomes – The X chromosome in male heterogamety and the Z in female heterogamety.

Indel – Small insertion or deletion of nucleotides. An indel where the number of nucleotide is not a divisible by three and is in the coding region of a gene can result in a frameshift mutation.

Inter-locus sexual conflict – Type of antagonism resulting when sexual conflict plays out over two or more loci.

Intra-locus sexual conflict – Type of antagonism resulting when male and female optima differ for a single locus. Intra-locus sexual conflict can be resolved via a variety of genomic mechanisms.

Monogamy – Mating system where males and females pair and are sexually exclusive.

Orthologs - Genes that have diverged from the same common ancestral gene by speciation.

Paralogs – Genes that have diverged from the same common ancestral gene by gene duplication.

Polyandry – Mating system where females have more than one mate at any given time, resulting in greater variance in female compared to male reproductive success.

Polygamy – General term describing a deviation from monogamy, encompassing both polyandrous and polygynous mating systems.

Polygyny – Mating system where males have more than one mate at a time, resulting in greater variance in male compared to female reproductive success.

Polyspermy - A condition resulting when more than one sperm cell fertilizes an ovum, often resulting in inviability.

Reproductive character displacement – Natural selection acting on maladaptive hybrids can promote the rapid divergence of phenotypes crucial to reproduction between closely related taxa. This acts to reinforce prezygotic reproductive isolation between populations.

Single nucleotide polymorphism (SNP) – Variable site within a population where the nucleotide identify differs frequently. In contrast, variable nucleotide sites across non-recombining populations are called fixed differences.

Stratum – Region on the sex chromosomes where recombination has been suppressed, either by an inversion or large scale changes in heterochromatin patterning. Strata can be identified by clusters of gametologs with similar divergence.

1.8 STATEMENT OF CONTRIBUTION

I am the primary author of the text in this thesis, and corresponding author for all publications arising from this work.

For Chapter 2, I identified and characterized the divergence of Z-W gametologs, evolutionary history of the Z chromosome and gene expression evolution. Dr. Hooman Moghadam identified the W-linked genes used in this study, and obtained FPKM values from raw RNA-seq data. This is detailed in Moghadam et al. 2012 PNAS.

For Chapter 3, I identified novel W-linked genes, characterized the divergence of Z-W gametologs and assessed the evolutionary forces driving W chromosome evolution. Professor Judith Mank, Dr Marie Pointer and I collected tissue samples. Dr. Peter Harrison quality trimmed raw RNA-seq data.

For Chapter 4, I estimated the magnitude of Faster-Z Effect, the relationship between sexual selection and Z-linked divergence, calculated effective population sizes and established the relative role of drift versus positive selection in Z chromosome evolution. Professor Judith Mank, Dr Marie Pointer and I collected tissue samples. Dr Peter Harrison assembled *de novo* transcriptomes, calculated sex-biased expression, divergence and polymorphism estimates, and conducted site-model tests. Fabian Zimmer and I developed a pipeline to identify reciprocal orthologs.

Chapter 2

Trade-off between selection for dosage compensation and masculinization on the avian Z chromosome

2.1 ABSTRACT

Following the suppression of recombination, gene expression levels decline on the sex-limited chromosome, and this can lead to selection for dosage compensation in the heterogametic sex to balance average expression from the X or Z chromosome with average autosomal expression. At the same time, due to their unequal pattern of inheritance in males and females, the sex chromosomes are subject to unbalanced sex-specific selection, which contributes to a non-random distribution of sex-biased genes compared to the remainder of the genome. These two forces act against each other, and the relative importance of each is currently unclear. The *Gallus gallus* Z chromosome provides a useful opportunity to study the importance and trade-offs between sex-specific selection and dosage compensation in shaping the evolution of the genome as it shows incomplete dosage compensation and is also present twice as often in males than females, and therefore is predicted to be enriched for male-biased genes. Here, I refine our understanding of the evolution of the avian Z chromosome, and show that multiple strata formed across the chromosome over roughly 130 million

years. I then use this evolutionary history to examine the relative strength of selection for sex chromosome dosage compensation versus the cumulative effects of masculinizing selection on gene expression. I find that male-biased expression increases over time, indicating that selection for dosage compensation is relatively less important than masculinizing selection in shaping Z chromosome gene expression.

2.2 INTRODUCTION

Males and females of the same species are subject to distinct selective forces, often resulting in contradictory sex-specific selection pressures acting on a given locus. These selective forces have been shown to affect large portions of the genome (Connallon et al. 2010) and can place a significant genetic and evolutionary burden on a species (Chippindale et al. 2001; Arnqvist and Rowe 2005; Morrow et al. 2008). At the genetic level, sexually antagonistic selection pressures are thought to contribute to gene expression differences observed between the sexes in many species (Rice 1984; Connallon and Knowles 2005; Ellegren and Parsch 2007).

The effect of sex-specific selection is particularly evident on the sex chromosomes, where the unbalanced pattern of inheritance creates uneven sex-specific selection pressures (Rice 1984; Charlesworth et al. 1987; Connallon and Clark 2010). Consequentially, the sex chromosomes are a hotspot of intra-locus sexual conflict within the genome (Mank 2009a; Innocenti and Morrow 2010), and unequal sex-specific selection is predicted to have contributed to the complex pattern of sex-biased gene distribution observed across the sex chromosomes of many animals (Parisi et al. 2003; Khil et al. 2004; Sturgill et al. 2007; Zhang et al. 2010a; Zhang et al. 2010b; Chen et al. 2011). Indeed, the avian Z chromosome exhibits a pervasive pattern of male-biased gene expression, which could be interpreted in the light of masculinizing selection for dominant male-benefit genes (Ellegren et al. 2007; Itoh et al. 2007).

At the same time, sex chromosomes experience changes in gene dose. Due to lack of recombination with their homologs, Y and W chromosome gene activity slowly degenerates by neutral processes (Charlesworth 1996), and the buildup of non-synonymous and nonsense mutations as well as small indels (Zhou and Bachtrog

2012), where the rate of decay declines with the number of functionally constrained loci (Bachtrog 2008a). Gene expression loss shows a range of dominance (Agrawal and Whitlock 2011), and for some genes, loss-of-function mutations on the W or Y chromosome will cause negative fitness effects (Charlesworth 1978).

The resulting loss of gene activity in the heterogametic sex and thus unbalanced gene dose with the autosomes is predicted to be costly, especially for X or Z-linked genes that interact with autosomal genes in large protein complexes (Pessia et al. 2012), thereby selecting for the evolution of dosage compensation mechanisms (Ohno 1967; Charlesworth 1978). Selection for dosage compensation can be strong, leading in some organisms to complex regulatory machinery to equalize gene dose across the entirety of the chromosome (Muyle et al. 2012; Pessia et al. 2012). However, more often selection acts on dosage sensitive genes not entire chromosomes (Mank and Ellegren 2009a; Vicoso and Bachtrog 2011; Pessia et al. 2012). In *Drosophila*, males hyper-transcribe the single X chromosome (Lucchesi 1973), whereas in mammals a subset of X-linked genes appear to be up-regulated in both sexes, and then down-regulated in females via X chromosome inactivation to restore balanced gene dose with the autosomes (Deng et al. 2011; Julien et al. 2012; Lin et al. 2012; Pessia et al. 2012). A by-product of the evolution of dosage compensation is that gene expression levels are also balanced between the heterogametic and homogametic sexes (Bachtrog et al. 2011; Mank et al. 2011).

Birds lack complete Z chromosome dosage compensation, and the majority of Z-linked genes are expressed more in males than females due to the fact that males have two copies of Z-linked loci and females only one (Ellegren et al. 2007; Itoh et al. 2007). For dosage sensitive genes, the predicted cost of unbalanced gene dose with

the autosomes selects for the evolution of dosage compensation mechanisms, such as hyper-transcription of Z-linked dosage sensitive genes in females, to restore balanced expression levels, and consequentially to equalize expression between the sexes (Ohno 1967; Pessia et al. 2012). Therefore, selection for dosage compensation theoretically reduces male-biased expression on the Z chromosome, and this predicted outcome is opposite to what would be expected under a model of masculinizing selection for gene expression, where sex-specific differences in Z-linked expression are exacerbated.

The evolutionary history of the Z chromosome itself can be used to distinguish signatures of masculinizing selection from selection for dosage compensation. The avian Z chromosome resulted from at least two (Handley et al. 2004; Suh et al. 2011) recombination cessation events over the last 130 million years, resulting in strata of different ages. These strata are subject to unbalanced sex-specific selection and selection for dosage compensation, for different lengths of time (Charlesworth 1991).

Consequently, patterns of sex-biased expression between the different strata will reflect the selective regime to which the Z chromosome has been subject. Under selection for dosage compensation with no masculinization, Z-linked expression in the youngest stratum is predicted to reflect the ratio of expression from two copies in males to only one in females. Changes in gene dose show a range of expression level effects (Torres et al. 2008), and twofold change in human gene copy number produces an average fold-change in expression of 1.5 (Pollack et al. 2002), consistent with the sex-biased expression ratio observed under lack of dosage compensation (Deng et al. 2011; Pessia et al. 2012), potentially due to the robust nature of gene networks (Oliver 2007). Therefore, dose effects without masculinization are expected to have, on

average, 1.5-fold effects on expression. Z-linked expression is predicted to become increasingly less male-biased in older regions where effective compensation mechanisms have evolved. However, the opposite pattern is expected under masculinizing selection. In this scenario, the extent of male-biased expression should increase with the age of the stratum and thus length of exposure to cumulative masculinizing selection.

Here I expand and refine our understanding of the topology of the chicken Z chromosome through the addition of several newly-discovered Z-W orthologs (gametologs), and find evidence for up to four evolutionary strata. I then use this evolutionary history to test these two predictions for gene expression evolution.

2.3 MATERIALS AND METHODS

2.3.1 Identification of mis-assembled W-linked genes

Fertilised Red Jungle Fowl eggs were obtained (T. Pizzari, Oxford University) and kept under standard incubator conditions. After collecting the left gonad at embryonic day (ed) 19, tissue was stored in RNAlater until preparation. Previous work has shown this stage is just prior to meiotic sex chromosome inactivation (MSCI) in chickens, which affects both the W and Z chromosome (Schoenmakers et al. 2009).

Therefore, ed19 gonads were used in this study in order to maximize the number of active W genes and minimise the risk that Z expression estimates are confounded by MSCI. Detailed methods for the discovery of new W-linked genes are described elsewhere (Moghadam et al. 2012). Briefly, RNA was extracted from four samples of each sex using standard purification methods, which was then used for

Illumina RNA-seq at The Wellcome Trust Centre for Human Genetics Facility at Oxford University resulting in 16 million 50 bp paired-end mappable reads on average per sample. Sequences were mapped to the chicken reference genome (WUGSC 2.1/galGal3) using Bowtie v0.12.7 (Langmead et al. 2009) and Tophat v1.1.1. Transcript abundances for the Ensembl annotated genes were estimated using Cufflinks v0.9.3 (Trapnell et al. 2010) and putative W–linked genes identified through *in silico* analysis of the gene expression profiles between males and females. PCR genotyping of genomic DNA was used to validate that these genes had been mis-assembled in the current Ensembl annotation (WUGSC 2.1/galGal3) and were indeed W-linked (Moghadam et al. 2012).

2.3.2 Identification and divergence of Z-W orthologs

In total, I had sequence data on 14 known and 12 newly identified W-linked genes from Ensembl, some of which represented multiple orthologs of Z-linked loci. These sequences were BLASTed (Altschul et al. 1997) against the chicken genome sequence (WUGSC 2.1/galGal3) to identify orthologs on the Z chromosome. The physical position of these orthologs was taken from the genome assembly. There was one exception, ENSGALG00000004349, which does not have any annotated Z-linked orthologs in Ensembl.

Synonymous divergence estimates can be used to identify the number and boundaries of evolutionary strata, as they provide an appropriate measure of the length of time over which gametologs have differentiated from an original ancestral autosomal pair. I therefore calculated synonymous divergence (d_S) for each gametologous gene pair using whole coding sequences for chicken genes downloaded

from Biomart with the exception of SPINZ and SPINW. The SPIN gametologs are poorly annotated therefore I identified and excluded incorrect exon annotations from the analysis by BLASTing the coding sequences against the NCBI EST library. After translating coding sequences for Z-W orthologs to protein sequences I used PRANK (http://tinyurl.com/prank-msa) to align gametologs. I checked alignments by eye, and removed poorly aligned regions where at least 50% of the amino acids were mismatched. PRANK's method of distinguishing insertions from deletions avoids overmatching of aligned sequences, consequentially producing more evolutionarily accurate alignments for increasingly diverged sequences in comparison to ClustalW (Loytynoja and Goldman 2005; Tamura et al. 2011). However, in line with previous studies, to check for consistency I repeated alignments using ClustalW in MEGA5 (Tamura et al. 2011). I used CODEML in the PAML package v4.4 (Yang 2007) to calculate maximum likelihood estimates of synonymous (d_S) and nonsynonymous (d_N) divergence in pairwise comparisons and generate standard errors with the curvature method. In one case, I removed ENSGALG00000013670 from the analysis to avoid biasing my results, as d_S was found to be greater than two for the multiple Z-linked orthologs, making accurate divergence estimates in avian coding sequence impossible due to mutational saturation and double hits (Axelsson et al. 2008). As problems with the use of PAML to calculate d_S values have been documented (Bierne and Eyre-Walker 2003) I also used an additional method, K-estimator 6.1 (Comeron 1995, 1999), to verify my synonymous divergence estimates for the 18 Z-W orthologs.

Some Z-W orthologs had multiple W or Z paralogs. To identify the true gametolog, I calculated d_S individually for each potential orthologous gene pair, and used the gene pair with the lowest d_S estimate in the final results. The surplus genes

were excluded. This is different from previous methods which employed a consensus sequence (Nam and Ellegren 2008), but avoids potential problems of relaxed or diversifying selection acting on gene copies (Graur and Li 2000) likely due to gene duplications following recombination suppression, which can result in significant divergence in both gene sequence and expression (Busby et al. 2011). This approach and the subsequent exclusion of multi-copy genes meant that whilst I had sequence data for 26 W-linked genes, my conclusions regarding the number and boundaries of strata were based on d_S estimates from 18 true Z-W gametologs.

Gene conversion (Slattery et al. 2000; Ross et al. 2005) between differentiated Z-W orthologs has the potential to bias estimates of d_S by homogenizing the two independently evolving sequences. To ensure this did not influence my identification of true gametologs, I calculated the GC content at the third codon position of Z-linked genes (GC3) and their multiple candidate W orthologs using the R package seqinR (Charif and Lobry 2007). GC content and recombination rate are positively correlated in vertebrates (Galtier 2003; Meunier and Duret 2004), leading to the prediction that, in the absence of gene conversion, the GC3 content of W-linked genes should not exceed that of Z-linked gametologs. I obtained 95% confidence intervals with bootstrapping of 1000 replicates.

Additionally, I used GENECONV as an independent test of gene conversion (http://www.math.wustl.edu/~sawyer) between Z-W gametologs, using *Anolis carolinensis* as an outgroup. For the 18 Z-linked genes I identified 16 *Anolis* orthologs and downloaded the longest coding transcript from BioMart. I used PRANK (http://tinyurl.com/prank-msa) to generate alignments and GENECONV to identify aligned global fragments for which gametologs are sufficiently similar to be suggestive

of past gene conversion. GENECONV assigns p-values by two methods, the first based on 10,000 random permutations and the second by a method of Karlin and Altshul (Karlin and Altschul 1990, 1993). Both methods are corrected for multiple comparisons and sequence length when detecting global fragments.

I mapped d_S estimates to the Z chromosome and used four criteria to delineate stratum boundaries: 1) average and range of d_S in physically-proximate regions; 2) k-means clustering of d_S estimates into a defined number of groups in R (Forgy 1965; MacQueen 1967; Hartigan and Wong 1979; Lloyd 1982), (www.R-project.org); 3) Z-W ortholog density; 4) physical location on the Z chromosome. To verify the refined strata, I used one-tailed t-tests to test my explicit predictions regarding the increase in d_S estimates with strata age. Although this is a biased estimate, it is useful as a means of comparison with k-means clustering results.

I was also able to use d_S estimates to calculate divergence dates for Z-W orthologs with a molecular clock that accounts for sex-specific mutation rates in Z and W chromosome lineages based on male-mutation bias estimates in birds. In Galliformes, the W chromosome rate is approximately half the autosomal rate (Axelsson et al. 2004). Based on avian and mammalian divergence dates, the autosomal mutation rate for the chicken lineage is estimated to be 2.5×10^{-9}/site/year (Dimcheff et al. 2002; Hedges 2002; Webster et al. 2006; Nam and Ellegren 2008) and correspondingly 1.25×10^{-9}/site/year for the W chromosome. Previous work in Galliformes indicates the Z chromosome mutation rate is only slightly larger than the autosomal rate leading to an estimate of 2.6×10^{-9}site/year (Nam and Ellegren 2008). The Z and W mutation rates mentioned above are therefore consistent with previous studies which estimate male mutation bias in Galliformes as lying between two to four

(Harlid et al. 2003; Berlin et al. 2006). These estimates result in a chicken Z-W divergence rate of 3.8 x 10^{-9}/site/year (Z chromosome mutation rate = 2.6 x 10^{-9} + W chromosome mutation rate = 1.2 x 10^{-9}) and divergence time of d_S / 3.8 x 10^{-9} (Nam and Ellegren 2008) years, although this has a high variance.

To enhance our understanding of the formation of avian strata through time, I used synteny with the turkey which diverged from the chicken approximately 30 mya (Dimcheff et al. 2002). Genomic rearrangements are not frequent across the avian genome and birds therefore show highly conserved synteny across vast evolutionary distances (Backstrom et al. 2010; Warren et al. 2010). For the 18 Z-linked genes used to identify chicken strata, I identified 1-1 turkey orthologs and their position along the Z chromosome from Ensembl (Turkey_2.01/GCA_000146605.1), and plotted them against Z-linked positional information from the chicken.

2.3.3 Comparing expression across strata

After defining putative strata boundaries based on d_S estimates, I used FPKM (Fragments Per Kilobase of exon per Million mapped reads) values obtained from the left gonad RNA-Seq data to compare expression among strata. The assembly method is discussed above under Identification of mis-assembled W-linked genes. I filtered the expression dataset to exclude any genes not expressed in at least four individuals, and all genes within the male-hypermethylated region (MHM), located between 2.5 Mb to 3.5 Mb (Teranishi et al. 2001). The MHM was excluded to avoid biasing my comparison of expression between strata, as microarray analysis indicates that this region is strongly down regulated in males, a markedly different pattern from the rest of the Z chromosome (Melamed and Arnold 2007).

With the filtered dataset, I separately averaged female and male FPKM values for each stratum, and calculated male-biased expression as a function of $\log_2$ (male FPKM/female FPKM) = $\log_2$ (male FPKM) - $\log_2$ (female FPKM). I compared average male-biased expression among strata using permutation tests with 1000 replicates, and calculated 95% confidence intervals using bootstrapping with 1000 replicates.

In addition, as dosage compensation acts to equalize expression across the Z chromosome and autosomes, I calculated the Z:A expression ratio separately for males and females to test for signatures of this selective regime. I calculated average male and female expression separately for all Z-linked genes (minus the MHM) and autosomal genes, and finally a Z:A expression ratio for each sex. This was also repeated separately for each stratum. 95% confidence intervals were calculated by bootstrapping male and female Z-linked and autosomal averages with 1000 replicates.
I was also able to calculate and compare the rate of masculinization of the Z chromosome under both logarithmic and linear functions to describe the relationship between average stratum age, and male-biased expression. In addition, to verify that the relationship I observe is not due to chance, I randomized the ratio of male to female gene expression across the Z chromosome (1000 replicates). Each time, I recalculated average male-biased expression for each stratum to obtain a distribution of randomized r^2 values.

2.3.4 Gene function analysis

To determine whether non-random distributions of gene function were contributing to gene expression differences, I used GOrilla to perform a Gene Ontology

enrichment analysis between each stratum and the entire Z chromosome (minus the MHM), and between the Z chromosome (minus the MHM) and autosomes (Eden et al. 2007; Eden et al. 2009). I conducted each GOrilla analysis with 1-1 mouse orthologs, a threshold of $p < 10^{-2}$ using standard hyper-geometric statistics and the Benjamini and Hochberg correction for multiple tests (Benjamini and Hochberg 1995).

2.4 RESULTS

2.4.1 Resolving stratum boundaries

Synonymous estimates provide an appropriate measure of the length of time over which Z-W orthologs have differentiated from an original autosomal pair, and therefore have been used to establish the number and boundaries of evolutionary strata in several organisms (Lahn and Page 1999; Ellegren and Carmichael 2001; Handley et al. 2004; Bergero et al. 2007; Nam and Ellegren 2008). I therefore estimated d_S values for 18 Z-W gametologs (Table 2.1), 6 of which were previously unknown, representing a 50% increase in the known number of Z-W orthologs. With the exception of the HINTW gene which is known to have undergone adaptive molecular evolution in birds (Handley et al. 2004), d_N estimates ranged from 0.0002 to 0.0831, suggesting a high degree of conservation between Z-W gametologs. d_S estimates ranged from 0.128 to 0.513. Alignments repeated for consistency using ClustalW were extremely similar to those generated by PRANK, with the exception of four gametologs. This is because PRANK's treatment of insertions and deletions, in comparison to ClustalW, results in the productions of more evolutionarily accurate alignments for increasingly diverged sequences

Table 2.1. Sequence divergence between chicken Z-W orthologs.

Gene pair (Z/W)	Ensembl ID (Z/W)[a]	Position on Z (Mb)	d_S (SE)[b]	d_N (SE)[b]	Putative stratum	Divergence (MYA)
ST8SIA3Z/ ST8SIA3W	03049/ 22039	0.317	0.184 (0.0319)	0.007 (0.0035)	3	48
ZNF532Z/ ZNF532W	02852/ 14003	0.747	0.183 (0.0171)	0.043 (0.0045)	3	48
RPL17Z/ RPL17W	02696/ 22174	0.965	0.169 (0.0212)	0.000 (0.0000)	3	44
SMAD2Z/ SMAD2W	14697/ 10056	1.290	0.183 (0.0679)	0.018 (0.0107)	3	48
8030462N17RikZ/ 8030462N17RikW	01763/ 01585	1.874	0.172 (0.0311)	0.083 (0.0126)	3	45
ATP5A1Z/ ATP5A1W	14644/ 01756	1.938	0.203 (0.0284)	0.029 (0.0061)	3	54
UBAP2Z/ UBAP2W	13809/ 05785	6.876	0.206 (0.0176)	0.056 (0.0054)	3	54
VCPZ/ VCPW	01986/ 00386	7.932	0.189 (0.0231)	0.009 (0.0017)	3	50
GOLPH3Z/ GOLPH3W	03151/ 18586	9.013	0.128 (0.0340)	0.007 (0.0052)	3	34
ZFRZ/ ZFRW	03235/ 14545	9.104	0.164 (0.0160)	0.020 (0.0031)	3	43
NIPBLZ/ NIPBLW	03605/ 22678	10.720	0.147 (0.0524)	0.018 (0.0103)	3	39
MIER3Z/ MIER3W	14721/ 00140	16.837	0.219 (0.0619)	0.044 (0.0136)	2b	58
hnRNPKZ/ hnRNPKW	12591/ 14366	39.553	0.269 (0.0350)	0.004 (0.0021)	2b	71
SPINZ/ SPINW	14916/ 14641	42.572	0.224 (0.0534)	0.012 (0.0062)	2b	59
HINTZ/ HINTW	00428/ 22685	44.169	0.513 (0.1144)	0.314 (0.0459)	1	135
CHD1Z/ CHD1W	14642/ 15278	50.156	0.500 (0.0357)	0.047 (0.0044)	1	132
KCMF1Z/ KCMF1	15391/ 14441	52.658	0.286 (0.0385)	0.040 (0.0072)	2a	75
RASA1Z/ RASA1W	15639/ 22611	59.665	0.275 (0.0597)	0.034 (0.0105)	2a	72

[a] Ensgalg000000...
[b] Standard errors generated by taking the inverse of the second derivative of the log likelihood in Paml

(Loytynoja and Goldman 2005; Tamura et al. 2011). However, for previously identified orthologs my d_S estimates after alignment with PRANK were qualitatively identical (<0.03 different) to my d_S estimates, providing independent evidence for the multiple cessation of recombination across the avian Z chromosome.

My results do differ from previous studies in some instances where multiple W-linked genes share the same Z-linked orthologs and vice versa, likely due to gene duplications following recombination suppression. This is because I used the copy with the lowest d_S estimate in order to avoid problems with artificially high values due to relaxed selection acting on gene duplicates (Ridley 2003). As a consequence of this approach, my d_S estimates at three locations along the Z chromosome are somewhat lower than in previous studies (Nam and Ellegren 2008), reducing slightly my estimate of when recombination ceased to form Stratum 2. Additionally, newly identified Z-W orthologs extended my analysis across a wider region of the Z chromosome, allowing a reassessment of the number and boundaries of strata.

A comparison of GC content at the third codon position (GC3) between these Z-W orthologs revealed no significant difference in GC3, indicating that the lowest d_S estimates are not a result of biased gene conversion and can be used to infer true gametologs. I was also able to use GENECONV (http://www.math.wustl.edu/~sawyer) to identify aligned global fragments for which gametologs are sufficiently similar to be suggestive of past gene conversion. I used *Anolis carolinensis* as an outgroup, for which 16 of the 18 Z-W gametologs had *Anolis* orthologs. Of the 16 gametologs, 9 have not undergone gene conversion. However, there was limited evidence of gene conversion for 7 gametologs, with 1 significant inner fragment detected for Ensgalg00000014003

(permutation p_1 value = 0.002, K-A p_2 value = 0.003), Ensgalg00000001756 (p_1 value = 0.007, p_2 value = 0.012), Ensgalg00000005785 (p_1 value < 0.001, p_2 value < 0.001), Ensgalg00000014545 (p_1 value = 0.032, p_2 value = 0.040), Ensgalg00000022678 (p_1 value = < 0.001, p_2 value = 0.011) and Ensgalg00000022685 (p_1 value < 0.001, p_2 value < 0.001). Significant inner fragments are evidence of a possible gene conversion event between ancestors of the gametologs, but GENECONV also detects outer fragments that represent conversion events originating from outside the alignment, or within the sequence but where evidence has been destroyed by later mutation or gene conversion. Three significant outer fragments were detected for the Z-linked ortholog of Ensgalg00000014003 (p_1 value < 0.001, 0.003, 0.036, p_2 value < 0.001, 0.004, 0.038) and one for Ensgalg00000014545 (p_1 value = 0.006, p_2 value = 0.008) and Ensgalg00000000386 (p_1 value < 0.001, p_2 value < 0.001). Of the seven gametologs with evidence of gene conversion, six of these are located in the youngest stratum and one in the oldest. Genes within the youngest stratum have the largest number of multiple copies as the cumulative strength of degenerative forces are weakest, and W-linked degeneration is minimal compared to older stratum. Correspondingly, I expect a certain degree of intra-chromosomal recombination for these genes (Rozen et al. 2003; Backstrom et al. 2005; Davis et al. 2010), which might bias GENECONV's detection of conversion events between Z and W orthologs and explain the non-random distribution of genes with evidence of gene conversion across the strata. In line with this, four of the seven W-linked genes for which GENECONV detected significant fragments are present in multiple copies or have a W-linked paralog.

Overall, synonymous divergence estimates cluster into at least three groups (Handley et al. 2004; Nam and Ellegren 2008; Suh et al. 2011). However, I also find

evidence for a potential fourth stratum located within the previously identified Stratum 2 (Fig. 2.1). These strata correspond to 44-50 Mb (d_S = 0.500-0.513, Stratum 1), 50-60 Mb (d_S = 0.275-0.286, Stratum 2a), 17-44 Mb (d_S = 0.219-0.269, Stratum 2b) and 0.3-11 Mb (d_S = 0.128-0.206, Stratum 3). The physical separation of Stratum 2a and 2b by Stratum 1, suggests that recombination suppression occurred in these regions in quick succession.

Genes within Stratum 1 have a significantly higher d_S value than Stratum 2 ($p < 0.001$, t-test), supporting the notion that recombination ceased first in this region. My d_S estimates also differ significantly between putative Stratum 2a and 2b ($p = 0.050$, t-test), indicating recombination was suppressed independently twice in close succession. Finally, recombination ceased most recently in Stratum 3, where d_S values are significantly lower than Stratum 1 ($p < 0.001$, t-test), 2a ($p < 0.001$, t-test) and 2b ($p < 0.001$, t-test).

The results from the k-means analysis (Forgy 1965; MacQueen 1967; Hartigan and Wong 1979; Lloyd 1982), (www.R-project.org) provide further support for multiple recombination cessation events. I used k-means to partition d_S estimates into three clusters. The three groups defined by k-means are identical to the groups supported by the t-tests, providing an independent line of evidence to support the boundaries of Stratum 1, 2 and 3 corresponding to 44-50 Mb (Stratum 1), 17-60 Mb (Stratum 2), and 0.3-11 Mb (Stratum 3). Due to the very short divergence interval between Stratum 2a and 2b, the power of k-means to identify these as independent clusters is weakened.

In addition to significant clustering of d_S estimates, the density of genes within each region provides independent evidence to identify strata (Lahn and Page 1999). The suppression of recombination between Z-W orthologs promotes the degeneration

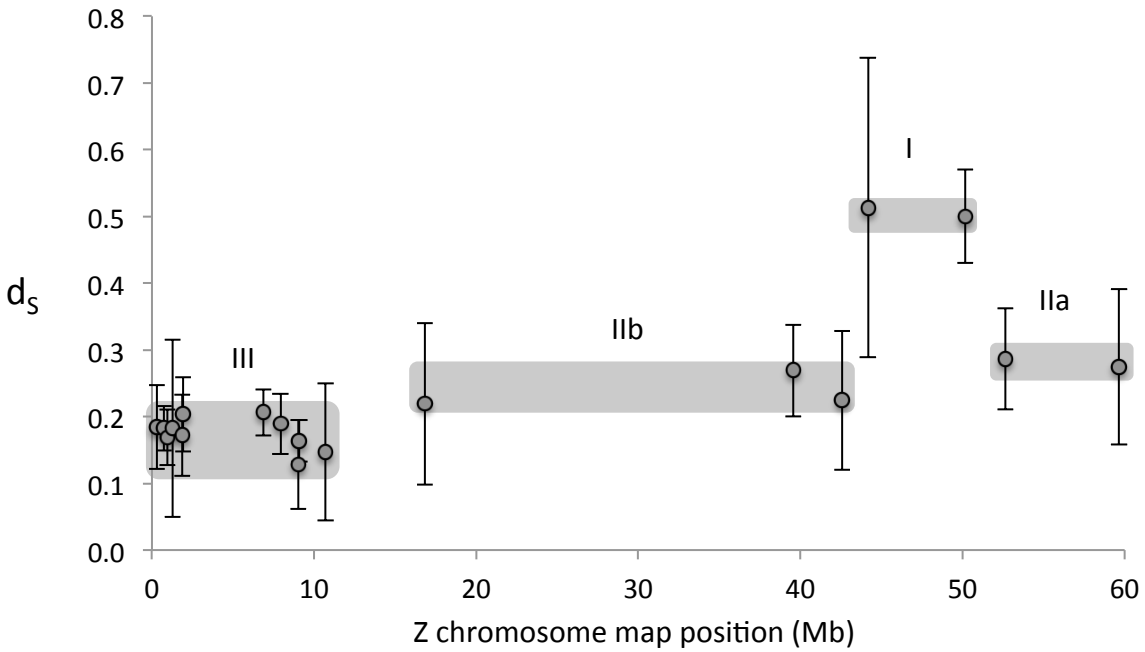

Fig. 2.1. Evolutionary history of the chicken Z chromosome.

The distribution of synonymous divergence estimates (d_S) for Z-W orthologs is shown. Physical position on the Z chromosome is based on the Ensembl genome assembly (WUGSC 2.1/galGal3). 95% confidence intervals are based on 1000 bootstrap replicates. Synonymous divergence estimates cluster into up to four groups, which differ significantly from each other, providing support for the existence of multiple Z chromosome strata.

of W-linked genes by neutral processes, the rate of which declines with the number of functionally constrained loci (Bachtrog 2008a). As recombination ceased most recently in the youngest stratum, the cumulative strength of degenerative forces are weakest, and I predict minimal W-linked degeneration compared to older stratum.

Correspondingly, the largest density of identifiable gametologs is expected in the youngest stratum. Gene density across the chicken Z chromosome is consistent with this prediction (Fig. 2.1) where the youngest stratum (1.06 genes per Mb) has a gene density of gametologs at least three times larger than the oldest (0.33 genes per Mb). The intermediate stratum has a gene density of 0.22 genes per Mb, excluding the male hyper-methylated region (MHM). Reasons for excluding the MHM are discussed below. Furthermore, the majority of the multi-gene copies excluded from the analysis were located within the youngest stratum.

As gene conversion has the potential to bias my identification of stratum boundaries, I repeated the one-tailed t-tests and k-means clustering after excluding the 7 Z-W orthologs with evidence of gene conversion. k-means clusters the 11 remaining gametologs into the same 3 groups corresponding to Stratum 1, 2 and 3. There is also still a significant difference between d_S estimates between Stratum 2 and 3 ($p < 0.001$, t-test), 2a and 3 ($p < 0.001$, t-test), 2b and 3 ($p = 0.001$, t-test) as with the previous analysis. However, there is now no significant difference between Stratum 2a and 2b ($p = 0.051$, t-test). As only 1 gene pair remains in Stratum 1 after the 7 gametologs are removed I could not conduct any tests of significance for this group. The consistency between strata boundaries defined from analyses excluding and including gametologs subject to gene conversion indicates my results are not biased by gene conversion.

Molecular clock estimates show that Stratum 1 arose between 132-135 MYA, Stratum 2a between 72-75 MYA, Stratum 2b between 58-71 MYA and Stratum 3 between 34-54 MYA. 16 of the 18 chicken Z-linked orthologs were identified on the turkey Z chromosome, of which 7 are 1-1 orthologs. Synteny of these Z-linked orthologs is conserved between both species, indicating a lack of lineage-specific genomic rearrangements.

2.4.2 Gene expression and the role of dosage compensation and masculinizing selection

I was able to calculate and compare average sex-biased expression for the 33 transcribed genes located within Stratum 1, 154 in Stratum 2 (41 genes in Stratum 2a, 113 genes in Stratum 2b) and the 101 genes in Stratum 3. The three established strata show clear differences in sex-biased expression (Table 2.2), with a progressive stepwise relationship where Stratum 1 has the greatest degree of male-biased expression, Stratum 2 is intermediate in male-biased expression and Stratum 3 shows the lowest level of average male-bias. Interestingly, the extent of male-bias in Stratum 3 (1.51 unlogged FPKM) is consistent with previous estimates of sex-bias associated with uncompensated genes on the mammalian X chromosome (Deng et al. 2011; Pessia et al. 2012). As Stratum 3 is the youngest and so the cumulative effects of degenerative forces are weakest, this sex-bias likely reflects the ancestral expression level associated with incomplete dosage compensation (Table 2.2). This pattern is strongly indicative of masculinizing selection acting on the avian Z chromosome in a cumulative manner, with increasing male-bias accruing over time. This is opposite to predictions under selection for dosage compensation, which would produce a

Table 2.2 Sex-biased gene expression across Z chromosome strata.

Stratum	Estimated age (MYA)	Number of expressed genes	Male-bias (log_2 male FPKM- log_2 female FPKM) (unlogged male FPKM/female FPKM)
1	133	33	0.86 (2.10)
2[1]	67	154	0.64 (1.80)
2a	73	41	0.69 (1.86)
2b[1]	63	113	0.63 (1.77)
3	46	101	0.33 (1.51)

[1] Genes from the male hyper-methylated region excluded.

decrease in male-biased expression as a function of stratum age. The differences in male-biased expression between Stratum 1 and Stratum 2 ($p = 0.021$, permutation test with 1000 replicates), Stratum 2 and 3 ($p < 0.001$, permutation test with 1000 replicates), and Stratum 1 and 3 ($p < 0.001$, permutation test with 1000 replicates) are all statistically significant. The division of Stratum 2 into two putative strata follows the same pattern of successive masculinization of the Z chromosome (Table 2.2), providing further support to the idea that recombination ceased independently four times. Stratum 2a has a higher average male-bias than Stratum 2b as predicted, however, due to the very short divergence interval I would not expect this difference to be significant ($p = 0.247$, permutation test with 1000 replicates). Despite this, the pattern of male-biased expression is still strongly indicative of masculinizing selection, with male-bias in both Stratum 2a and 2b greater than Stratum 3 ($p < 0.001$, $p < 0.001$, permutation tests with 1000 replicates) but smaller than Stratum 1 ($p = 0.03$, $p = 0.023$, permutation tests with 1000 replicates).

I excluded the MHM from Stratum 2b in the analyses described above in order to avoid biasing our comparison of expression between strata, as microarray analysis indicates that this region is strongly down regulated in males, a markedly different pattern from the rest of Z chromosome (Melamed and Arnold 2007). Although the expression of the MHM is not representative of the Z chromosome, when included in the analysis, average male-bias in Stratum 3 is still significantly lower than Stratum 2 ($p < 0.001$, permutation test with 1000 replicates), leaving the significant relationship between male-biased gene expression and age largely unchanged.

Results from a GOrilla analysis (Eden et al. 2007; Eden et al. 2009) suggest the observed difference in expression between the strata is not due to a non-random

distribution of genes based on function. A comparison of each stratum to the whole Z chromosome revealed no significant enrichment of GO terms after correction for multiple testing ($p<10^{-2}$). Even before correction, only Stratum 2b and Stratum 2 were enriched for any GO terms; (Stratum 2b = behaviour and negative regulation of multicellular organismal processes ($p = 2.25 \times 10^{-4}$, $p = 6.84 \times 10^{-4}$), Stratum 2 = behaviour ($p = 2.10 \times 10^{-4}$). However, none of these terms were significant after the Benjamini and Hochberg correction.

Additionally, when the Z chromosome is compared to the autosomes, there is no significant enrichment of GO terms ($p<10^{-2}$) after the Benjamini and Hochberg correction. Before correction for multiple testing, the Z chromosome is enriched for thrombin receptor signalling pathway, oncostatin-M-mediated signaling pathway, positive regulation of peptidyl-tyrosine phosphorylation, positive regulation of Rho protein signal transduction, protein-lipid complex subunit organization, plasma lipoprotein particle organization, positive regulation of Ras protein signal transduction, regulation of biological quality, positive regulation of cell migration, positive regulation of positive chemotaxis, thrombin receptor activity, cytokine receptor activity, signal transducer activity, signal transducer activity, transmembrane signaling receptor activity, intrinsic to plasma membrane ($p = 1.77 \times 10^{-5}$, $p = 8.42 \times 10^{-5}$, $p = 2.03 \times 10^{-4}$, $p = 3.26 \times 10^{-4}$, $p = 3.88 \times 10^{-4}$, $p = 3.88 \times 10^{-4}$, $p = 4.66 \times 10^{-4}$, $p = 5.43 \times 10^{-4}$, $p = 8.89 \times 10^{-4}$, $p = 9.47 \times 10^{-4}$, $p = 3.98 \times 10^{-5}$, $p = 8.42 \times 10^{-5}$, $p = 4.54 \times 10^{-4}$, $p = 4.54 \times 10^{-4}$, $p = 6.39 \times 10^{-4}$, $p = 8.60 \times 10^{-4}$). However, it is not apparent how these processes would explain the unique expression pattern observed.

As dosage compensation acts to equalize expression across the Z chromosome and autosomes, calculating the Z:A expression ratio separately for males and females

is essential for accurately assessing the selective forces driving expression evolution. I was able to calculate average male and female expression for 492 Z-linked genes (minus the MHM) and 10530 autosomal genes, and show that the Z:A expression ratio differs significantly between the sexes (Fig. 2.2). The female Z:A ratio is consistent with incomplete dosage compensation observed in previous studies (Deng et al. 2011; Pessia et al. 2012), whereas the male ratio reflects the ZZ:AA karyotype. There is no significant difference in female and male Z:A ratio across strata, and this may be due to the variation in overall expression level among individual genes, which would be expected to have a strong stochastic effect in strata with relatively few genes.

2.4.3 Rate of masculinization

The degree of male-bias increases across the Z chromosome as a function of time since divergence from the orthologous W chromosome (Figs. 2.3A and 2.3B). I found a stronger correlation between stratum age and degree of sex-bias under a logarithmic rather than a linear function for both the three stratum (r^2 = 0.930) and four stratum (r^2 = 0.888) models of Z chromosome evolution. Consequentially, this suggests an upper limit to the increase in masculinization of Z-linked gene expression over evolutionary time.

I randomized gene expression across the Z chromosome (1000 replicates) in order to assess the probability that the observed differences in sex-bias are due to chance alone. My observed r^2 values for the positive logarithmic relationship between stratum and sex-biased expression was significant (p = 0.04), but the linear relationship was not (p = 0.08), reinforcing my finding that the rate of masculinization declines over time.

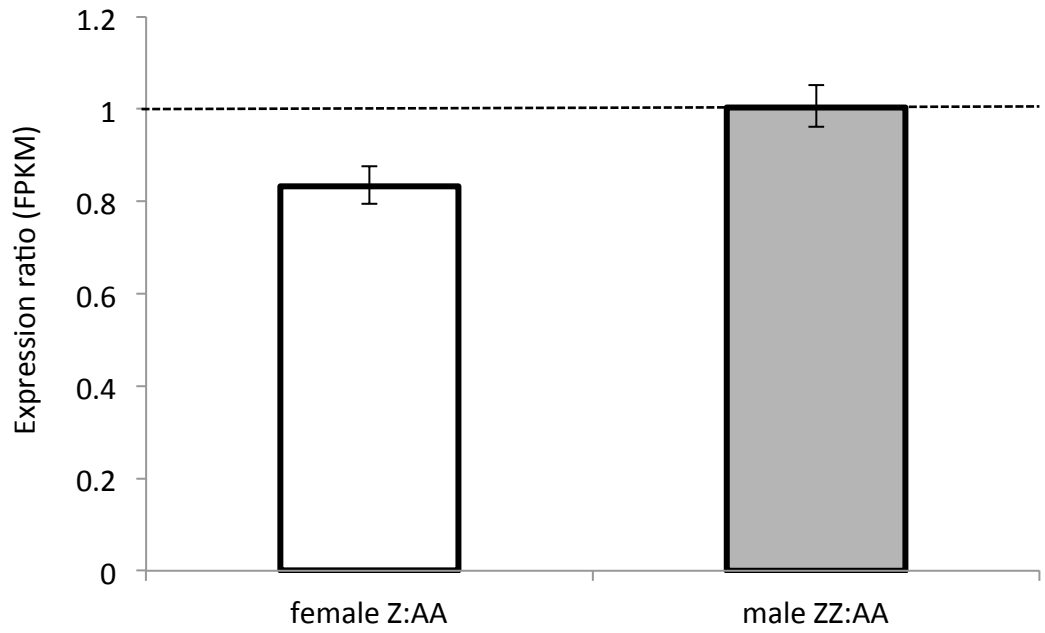

Fig. 2.2. Comparison of gene expression across the Z chromosome and autosomes.

The average Z:A expression ratio is shown for males and females separately. 95% confidence intervals (based on 1000 bootstrap replicates) are shown. Dotted line represents the expression ratio expected under equal gene dose.

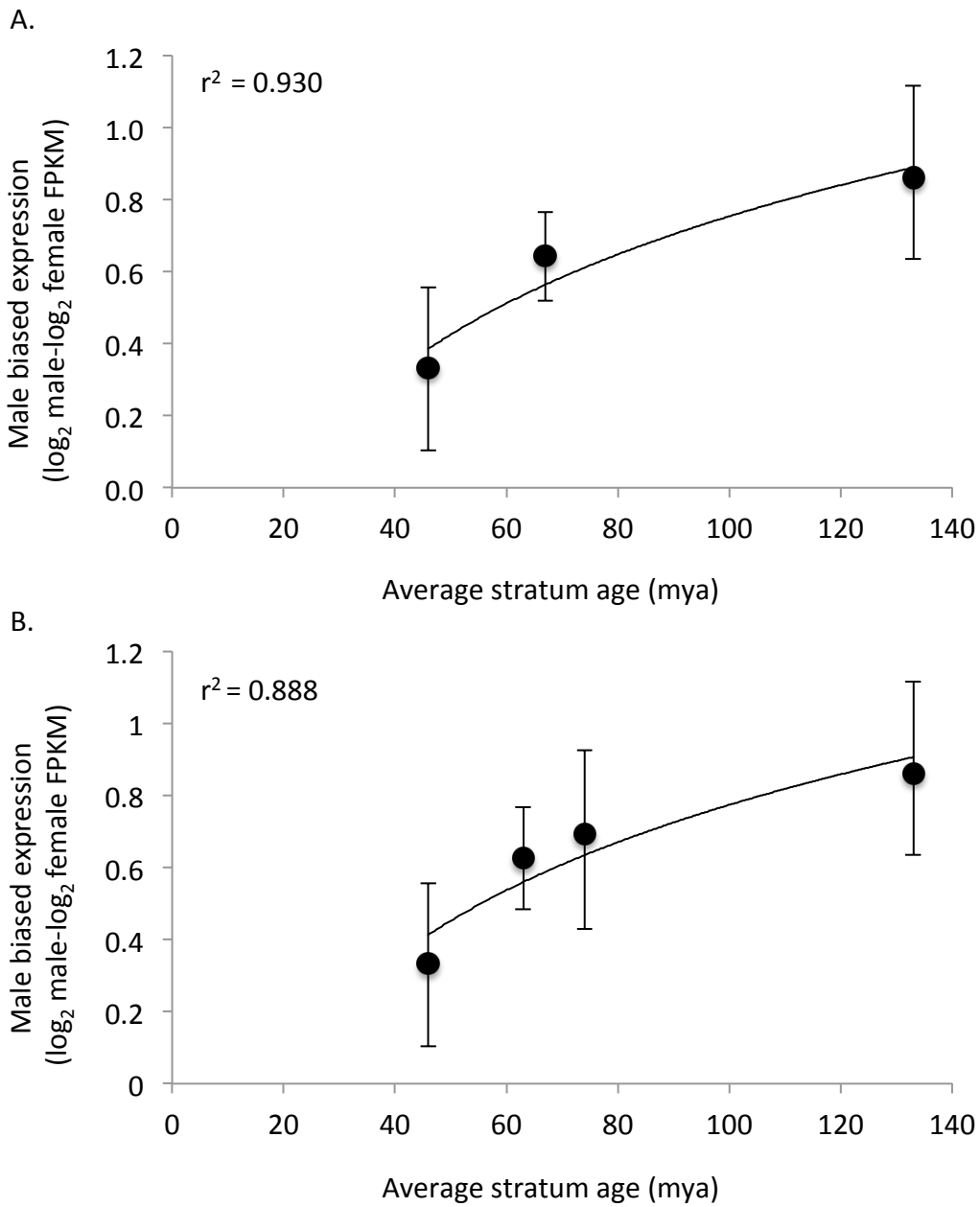

Fig. 2.3. Rate of masculinization of the Z chromosome.

The degree of male-biased expression for inferred Z chromosome strata is shown for both three strata (panel A) and four strata (panel B) models of Z chromosome evolution. 95% confidence intervals (based on 1000 bootstrap replicates) are shown. Masculinization of Z-linked gene expression increases as a function of age, but ultimately levels out, thereby limiting the role of the Z chromosome in the evolution of sexual dimorphism.

2.5 DISCUSSION

The Z chromosome could potentially be subject to two different sex-specific selection regimes. Due to the unequal inheritance pattern of the chromosome, male-specific selection for dominant alleles is predicted to drive the evolution of the Z chromosome coding content (Rice 1984), leading to the accumulation of male-benefit alleles. Correspondingly, masculinizing selection acting on gene expression (Connallon and Clark 2010) is expected to increase male-bias over time.

Loss of function mutations for W-linked genes are expected to cause negative fitness effects (Charlesworth 1978). The resulting loss of gene activity in females, and thus unbalanced gene dose with the autosomes is predicted to be costly, especially for dosage sensitive Z-linked genes that interact with autosomal genes in large protein complexes (Pessia et al. 2012). This is predicted to result in selection for hyper-transcription of dosage sensitive genes to equalise expression with the autosomes (Ohno 1967). If selection for dosage compensation were cumulative, I would expect it to result in a pattern opposite to that produced by masculinizing selection, as an indirect effect of selection for dosage compensation is to decrease sex-biased expression on the Z over time.

The Z chromosome has been previously observed to show a pervasive pattern of male-biased expression (Kaiser and Ellegren 2006; Storchova and Divina 2006; Mank and Ellegren 2009a). However, the relative role of masculinizing selection versus incomplete dosage compensation in driving this pattern has been difficult to untangle. Here I employ a novel method to detect signatures of sex-specific selection from selection for dosage compensation. My results indicate that masculinizing selection predominates over selection for dosage compensation, and I was able to calculate the

rate of masculinization using the evolutionary history of the chicken Z chromosome itself. My analysis includes 18 Z-W gametologs; a sizeable dataset in line with mammalian studies (Lahn and Page 1999) and to date the largest in birds, as well as 492 expressed genes along the Z chromosome.

2.5.1 Definition of Z chromosome strata

This study is reliant on accurate knowledge of the number and precise boundaries of strata, and using synonymous divergence estimates for newly identified Z-W orthologs together with gene density I was able to re-examine the chicken strata across a wider region of the Z chromosome than in previous studies. In doing so, I found support for the existence of up to four strata along the chicken Z chromosome. My data suggest that recombination ceased first in the 44-50 Mb region of the Z chromosome, between 132-135 MYA (Stratum 1). Then, in rapid succession, d_S estimates indicate recombination ceased in the 53-60 Mb region between 72-75 MYA (Stratum 2a) and the 17-43 Mb region roughly 58-71 MYA (Stratum 2b). The final suppression occurred in the 0.3-11 Mb region most recently, between 34-54 MYA (Stratum 3), as shown in Fig 2.1. The large reduction in gene density across the oldest and intermediate strata compared to Stratum 3 provides additional evidence for up to four chicken strata. In line with patterns of gene density observed across the mammalian X chromosome (Lahn and Page 1999), the number of gametologs is predicted to decline with age of the strata where the cumulative strength of degenerative forces is weakest and W-linked divergence is minimal. This decay is not predicted to follow a linear progression, but instead should decline with the number of functionally constrained loci (Bachtrog 2008a).

The identification of a possible fourth stratum may suggest that recombination suppression occurred on both sides of Stratum 1 in close succession, rather than requiring a suppression event for Stratum 2 followed by a subsequent structural rearrangement across the boundaries of Stratum 1 and 2 leading to a non-linear distribution of strata (Nam and Ellegren 2008; Bergero and Charlesworth 2009). The conservation of relative gene order between turkey and chicken Z-linked orthologs argues against a chicken-specific inversion, but does not rule out the possibility of this event occurring before the two lineages diverged. Therefore, my model to describe the successive formation of avian strata does not evoke additional cessation of recombination events but instead proposes an alternative arrangement of these events over evolutionary time.

2.5.2 Successive masculinization of the Z chromosome

By comparing sex-biased expression among the chicken strata, I uncover a pattern consistent with cumulative masculinizing selection acting on gene expression across the Z chromosome, where the magnitude of male-biased expression correlates closely with the age of the stratum. Masculinization may have occurred to some extent in the youngest stratum, which ceased to recombine with the W chromosome between 34-54 MYA. However, the effect is greater in the older strata. This logarithmic relationship holds when both three (r^2 = 0.930) and four strata (r^2 = 0.888) are considered, and is supported by randomization tests (p = 0.04). The significance of the randomization test, although marginal, is convergent with the correlation between stratum age and male-biased expression level. The cumulative masculinization of Z chromosome expression is likely due to the effect of male-specific selection, as the Z

chromosome is twice as often present in males than females, therefore dominant alleles are more often selected for their male-specific fitness effects (Rice 1984; Connallon and Clark 2010).

Unbalanced sex-specific selection on sex chromosomes has been suggested to explain the complex non-random distribution of sex-biased genes in mammals (Khil et al. 2004) and *Drosophila* (Parisi et al. 2003; Sturgill et al. 2007; Chen et al. 2011). However, it has not been clear from previous work in birds how much of the male-bias on the Z chromosome is due to masculinization (Ellegren 2011) and how much due to incomplete dosage compensation (Ellegren et al. 2007; Itoh et al. 2007). Previous attempts to circumvent the problems of gene dose compared embryonic and adult gene expression levels (Mank and Ellegren 2009b). However, this did not account for the fact that sex-specific selection and sex-biased gene expression shift rapidly throughout development and the adult life cycle (Mank et al. 2010b). My method allows male-biased selection to be detected without conducting a comparison between developmental time points, avoiding the problems encountered in previous studies of varying magnitudes of sex-specific selection. Consequentially, my results suggest that the increase in male-bias over time is consistent with a cumulative effect of masculinizing selection rather than selection for dosage compensation. The older strata have been subjected to longer periods of stronger selection for dominant male-benefit alleles, and hence show a greater degree of masculinization. This finding is consistent with theoretical predictions and a recent study showing that genes expressed in the testis are overrepresented among newly emerged Z-linked genes (Ellegren 2011). The results can also explain previous findings that the extent of sex-biased expression varies across the Z chromosome (Melamed and Arnold 2007).

Determining whether sex-specific selection is responsible for the analogous feminization of the X chromosome in *Drosophila* (Parisi et al. 2003; Sturgill et al. 2007; Chen et al. 2011) and mice (Khil et al. 2004), is complicated by meiotic sex chromosome inactivation, a genetic mechanism inactivating sex-linked genes during late spermatogenesis (Vibranovski et al. 2009a). MSCI has been proposed as a driving force behind the under representation of spermatogenesis genes on the X chromosome; however, this is surrounded by much debate (Khil et al. 2004; Vibranovski et al. 2009b; Meiklejohn et al. 2011). In birds, MSCI is ephemeral, lasting only from early pachytene to diplotene phases during oogenesis (Schoenmakers et al. 2009). As samples were taken before this stage, MSCI is unlikely to drive the majority of the non-random pattern of sex-biased genes on the Z chromosome I observe.

The pattern of successive masculinization across the Z chromosome is in contrast to that seen on the human X chromosome, whereby more genes escape X chromosome inactivation, and are therefore expressed more in females, in the younger regions of the X than the older regions (Carrel and Willard 2005). This hints at the important role of dosage compensation in shaping the pattern of sex-biased genes on sex chromosomes, as the dosage compensation mechanism is less effective in younger regions of the mammalian X chromosome (Lin et al. 2007; Deng et al. 2011). The *Drosophila* X chromosome does not exhibit strata, possibly because the lack of recombination in males means that the Y decayed as a single unit or because the Y and X are not orthologous (Hackstein et al. 1996). Birds lack global sex chromosome dosage compensation (Ellegren et al. 2007; Itoh et al. 2007), and we would expect that if there were significant selection for dosage compensation, the pattern of male-biased expression would decrease as a function of stratum age. I predict that these

differences should be strongly visible across the Z chromosome, as evidence in several animals suggests that dosage compensation mechanisms evolve relatively slowly. The sex chromosomes of monotremes and birds are ancient and yet still display incomplete dosage compensation (Itoh et al. 2007; Deakin et al. 2008) and whilst there is a current debate regarding the status of dosage compensation in eutherian mammals, evidence suggests that these chromosomes also lack complete dosage compensation (Deng et al. 2011; Pessia et al. 2012). However, limited evidence from plants indicates that in some cases selection for dosage compensation may be more rapid (Muyle et al. 2012). Despite this, my data are not consistent with selection for dosage compensation, as I observe male-bias increasing over time in a cumulative fashion. Additionally, comparing Z-linked and autosomal expression confirms that females are not under selection for dosage compensation, as the expression ratio is consistent with a lack of hyper-transcription of the single Z chromosome (Deng et al. 2011; Pessia et al. 2012). This suggests a trade-off between dosage compensation mechanisms and sex-specific selection on gene expression, with dosage compensation mechanisms acting against the effects of cumulative sex-specific selection on sex chromosome transcription.

Interestingly, the static architecture of the avian genome may also contribute to the pattern of male-biased expression on the Z chromosome. Recently, it has been shown that there is little gene traffic on and off the Z chromosome, potentially due to a lack of active transposons within the avian genome (Toups et al. 2011). The unequal inheritance pattern of the Z chromosome renders it unfavourable for dominant female-benefit alleles (Rice 1984). However, the potential for these genes to move to

the autosomes is severely limited. Subsequently, all Z-linked genes are subject to strong male-specific selection, favouring the evolution of male-biased expression.

2.6 CONCLUSIONS

A considerable body of theory focuses on the importance of sex chromosomes in facilitating the evolution of sexually dimorphic phenotypes (Rice 1984). Indeed, the Z chromosome is thought to be especially conducive to sexual selection, and predicted to play an important role in encoding sexually selected traits, due to the male biased inheritance pattern (Reeve and Pfennig 2003; Kirkpatrick and Hall 2004a; Albert and Otto 2005). Implicit in these predictions is that the importance of sex linkage in encoding sexually dimorphic phenotypes increases over evolutionary time with cumulative exposure to sex-specific selection. However, chromosome or stratum age is typically not considered when attempting to measure the importance of sex chromosomes in the evolution of sexually selected traits. Here I show that the cumulative effects of male-specific selection vary across the chromosome, indicating that older regions harbour more genes that may contribute to sexual dimorphic phenotypes via sex-biased expression than younger ones.

Chapter 3

Independent stratum formation on the avian sex chromosomes reveals inter-chromosomal gene conversion and predominance of purifying selection on the W chromosome

3.1 ABSTRACT

I used a comparative approach spanning three species and 90 million years to study the evolutionary history of the avian sex chromosomes. Using whole transcriptomes, I assembled the largest cross-species dataset of W-linked coding content to date. My results show that recombination suppression in large portions of the avian sex chromosomes has evolved independently, and that long-term sex chromosome divergence is consistent with repeated and independent inversions spreading progressively to restrict recombination. In contrast, over short-term periods I observe heterogeneous and locus specific divergence. I also uncover four instances of gene conversion between both highly diverged and recently evolved gametologs, suggesting a complex mosaic of recombination suppression across the sex chromosomes. Lastly, evidence from 16 gametologs reveal that the W chromosome is evolving with a significant contribution of purifying selection, consistent with previous findings that W-linked genes play an important role in encoding sex-specific fitness.

3.2 INTRODUCTION

Recombination suppression initiates the divergence of sex chromosomes from an ancestral autosomal pair, and promotes the degeneration of the W (in female heterogamety) and Y (in male heterogamety) chromosome coding content (Ohno 1967; Charlesworth 1991; Charlesworth and Charlesworth 2000; Charlesworth et al. 2005; Bachtrog et al. 2008b). In spite of the pivotal role recombination suppression plays in sex chromosome evolution, our understanding of the underlying mechanisms and dynamics is largely limited to either highly divergent sex chromosomes (e.g. mammals and *Drosophila*), or incipient sex chromosome systems, such as those observed in plants and fish (Chibalina and Filatov 2011; Muyle et al. 2012; Bergero et al. 2013; Natri et al. 2013; Qiu et al. 2013).

Older sex chromosomes support the classic model of sex chromosome evolution, where intra-chromosomal rearrangements are proposed to initiate the suppression of recombination across large regions (Charlesworth et al. 2005). Theoretically, these inversions result in instantaneous recombination suppression between the sex chromosomes. This leaves a spatial signature of clusters of gametologs, orthologous genes on the X and Y (or Z and W) chromosomes, of similar divergence, often referred to as strata. Strata are thought to form independently in a stepwise process over millions of years. At least four chromosomal rearrangements occurring over 300 million years have been documented on the human X and Y chromosomes (Lahn and Page 1999; Ross et al. 2005; Lemaitre et al. 2009; Wilson and Makova 2009b; Pandey et al. 2013; Bellott et al. 2014), and there is evidence of four strata on the avian sex chromosomes spanning 130 million years (Wright et al. 2012).

Newly evolved sex chromosomes show less concordant support for the strata model. The threespine stickleback (*Gasterosteus aculeatus*) and white campion (*Silene latifolia)* X and Y chromosomes originated around 30 million years ago and less than 10 million years ago respectively. Work on these systems indicates that recombination suppression evolves heterogeneously along the length of the chromosome (Chibalina and Filatov 2011; Muyle et al. 2012; Bergero et al. 2013; Natri et al. 2013; Qiu et al. 2013). This suggests that although selection to suppress recombination between the X and Y chromosomes occurs over large regions, within those regions, the effects are variable and genetic exchange between the sex chromosomes persists in some places.

We need to fully integrate short- and long-term views of sex chromosome evolution in order to provide a complete temporal overview of the dynamics of recombination suppression. Additionally, much of our understanding of sex chromosome evolution is based on studies of neo-sex chromosomes in *Drosophila* (Bachtrog et al. 2008b; Kaiser et al. 2011; Zhou and Bachtrog 2012). These have been highly informative. However, males are achiasmatic in *Drosophila* and therefore cannot be used to study progressive recombination suppression events. Instead, comparative analyses of sex chromosome divergence across a range of species with recombination in both sexes are required to provide an evolutionary framework to characterise the temporal dynamics of sex chromosome evolution.

The avian sex chromosomes are homologous across the entire clade. However recombination suppression between the Z and W has evolved repeatedly in different lineages across 130 million years (Tsuda et al. 2007; Suh et al. 2011). I use the repeated evolution of avian sex chromosome strata to create a cohesive view of short-, medium- and long-term dynamics in Z-W divergence patterns. To do this, I assembled the largest

cross-species dataset of W-linked coding content to date spanning the Galliformes (the landfowl) and the Anseriformes (the waterfowl). These two sister-orders, the Galloanserae, last shared a common ancestor approximately 90 million years ago (van Tuinen and Hedges 2001). My analysis indicates that the majority of the Z and W chromosomes formed independently in each of these orders. Additionally, my results suggest ongoing recombination between gametologs in the most recent region of the Anseriform sex chromosomes, and gene conversion throughout the sex chromosomes.

3.3 Materials and Methods

3.3.1 Identification of W-linked genes

Previous efforts to identify W-linked sequences in birds (Chen et al. 2012) have resulted primarily in non-coding sequence, which makes comparisons to the Z chromosome difficult. To expand the known coding content of both the mallard duck (*Anas platyrhynchos*) and wild turkey (*Meleagris gallopavo*) W chromosomes, I used a combined approach based on RNA-seq data and sequence similarity to known red jungle fowl (*Gallus gallus*) W-linked genes.

RNA-seq data was obtained from captive populations of *A. platyrhynchos* and *M. gallopavo*, taken at the start of their first breeding season (year one for *A. platyrhynchos*, year two for *M. gallopavo*). All samples were collected in accordance with national guidelines and with permission from institutional ethical review committees. The left gonad was collected separately from five male and five female *A. platyrhynchos*, and seven male and five female *M. gallopavo*. The spleen was collected from five male and five female *A. platyrhynchos* and four male and two female *M.*

gallopavo. The samples were homogenized and stored in RNAlater until preparation. RNA was extracted using the Animal Tissue RNA Kit (Qiagen) and The Wellcome Trust Centre for Human Genetics, University of Oxford, prepared and barcoded samples using standard methods. RNA was sequenced on an Illumina HiSeq 2000 resulting in on average 26 million 100 bp paired-end reads per sample.

The data was assessed for quality using FastQC v0.10.1 (www.bioinformatics.babraham.ac.uk/projects/fastqc) and filtered using Trimmomatic v0.22 (Lohse et al. 2012). Specifically, reads with residual adaptor sequences were removed and reads were trimmed if the leading or trailing bases had a Phred score <4 or the sliding window average Phred score over four bases was <15. Post filtering, reads were removed if either read pair was <25 bases in length. Filtered reads from each species were mapped to the respective reference genomes obtained from Ensembl v74 (Flicek et al. 2013) (*M. gallopavo* v2.01/GCA_000146605.1) (Dalloul et al. 2010; Dalloul et al. 2014) and *A. platyrhynchos* v1.0/GCA_000355885.1) (Huang et al. 2013)) using TopHat2 v2.09 (Trapnell et al. 2009; Kim et al. 2013) and Bowtie2 v2.1.1 (Langmead et al. 2009; Langmead and Salzberg 2012). For recently evolved gametologs with high sequence similarity, it is possible that Z-linked reads will spuriously map to W-linked exons. To avoid this and allow accurate identification of female-limited genes and therefore putative W-linkage, both reads of a pair had to map concordantly to the reference sequence and no mismatches in the alignment were permitted. This strategy, although essential to accurately differentiate gametologs, could fail to identify W-linked genes when the entire locus is not encompassed on a single scaffold in the genome assembly.

To identify substitutions between Z and W coding sequences the mapping criteria were relaxed. For each sample, I used Cufflinks v2.1.0 (Trapnell et al. 2010; Trapnell et al. 2013) to estimate transcript abundances for Ensembl annotated genes as fragments per kilobase per million mappable reads (FPKM). I identified strongly female-biased and female-limited genes as putative W-linked genes using gene expression profiles in males and females (Fig. S.3.1).

Putative W-linked genes were also identified by high sequence similarity to known *G. gallus* W-linked genes. *G. gallus* (Galgal4.0/GCA_000002315.2), *M. gallopavo* and *A. platyrhynchos* cDNA sequences were obtained from Ensembl v74 (Flicek et al. 2013) and for each species the longest transcript for each gene was identified. I determined putative orthology using BLAST v2.2.26 (Altschul et al. 1990) with an e-value cutoff of 1×10^{-10} and identified putative *A. platyrhynchos* and *M. gallopavo* W-linked genes that had the highest BLAST score to known *G. gallus* W-linked genes.

For both *A. platyrhynchos* and *M. gallopavo*, I validated putative W-linked genes by PCR using genomic DNA from three individuals of each sex.

3.3.2 Identification of gametologs

Previously, I used newly identified *G. gallus* gametologs to identify four strata on the Z chromosome (Wright et al. 2012). Subsequent to this work, the Ensembl annotation of the *G. gallus* W and Z chromosomes has been revised (Flicek et al. 2013) and additional W-linked genes have been identified (Ayers et al. 2013). To account for these changes, I updated gametologs and re-analyzed divergence estimates using coding sequences from Ensembl v74 (Flicek et al. 2013) and PAML v4.7 (Yang 2007). Revised divergence estimates were compared to previous estimates (Wright et al.

2012) with Spearman's correlation in the rcorr function (Hmisc package (Schemper 2003; Tian et al. 2007)) in R (R Core Team 2011). Across gametologs, revised d_S estimates were strongly correlated with previous estimates from Wright et al. 2012 (Spearman's rank correlation p-value<0.01, correlation coefficient 0.78). The d_S of three newly identified gametologs were also consistent with the previously identified strata, with overlapping 95% confidence intervals (Table S.3.1), in line with the previous findings that the *G. gallus* Z chromosome was formed by four independent recombination suppression events.

A. platyrhynchos, M. gallopavo, *G. gallus* and *T. guttata* (taeGut3.2.4) coding sequences were obtained from Ensembl v74 (Flicek et al. 2013), and additional *M. gallopavo* sequences were obtained from Genbank (Backstrom et al. 2005; Berlin and Ellegren 2006). For each species the longest transcript for each gene was identified. *G. gallus* sequences were BLASTed reciprocally against *A. platyrhynchos, M. gallopavo* and *T. guttata* coding sequences using an e-value cut-off of 1×10^{-10} and the highest BLAST score was used to identify the best BLAST hit.

For *A. platyrhynchos* and *M. gallopavo* W-linked genes, I identified orthologous Z-linked genes using the following steps. For each W-linked gene; 1) I identified the orthologous *G. gallus* W-linked gene using BLASTn and the *G. gallus* Z-linked ortholog from Wright et al. 2012; 2) *A. platyrhynchos* and *M. gallopavo* Z-linked orthologs were defined as the reciprocal best hit of the *G. gallus* Z-linked ortholog; 3) in a few instances there was no reciprocal best hit, often because of high sequence similarity between Z and W gametologs. For these genes, the Z-linked reciprocal ortholog was identified as the next best BLAST hit of the *G. gallus* Z-linked ortholog.

Avian genomes are particularly stable with few major genomic rearrangements (Stiglec et al. 2007) and little gene movement off and onto the avian Z chromosome, potentially due to a lack of active transposons (Toups et al. 2011). I therefore expect Z-linkage and synteny to be highly conserved across *A. platyrhynchos, M. gallopavo* and *G. gallus*. I verified Z-linkage of reciprocal orthologs using positional information in the *M. gallopavo* genome assembly. The *A. platyrhynchos* genome assembly lacks annotated chromosomes and as recently diverged Z- and W-linked genes share a high degree of sequence similarity, to avoid mis-identifying the Z-linked reciprocal ortholog as a W-linked paralog, I verified the reciprocal ortholog was not located on the W chromosome using PCR. However, this approach does not distinguish Z chromosome to autosomal gene duplication, but given the pronounced stability of avian genomes (Toups et al. 2011), this is unlikely to bias my results.

3.3.3 Divergence of gametologs

To identify the number and boundaries of evolutionary strata in *A. platyrhynchos* and *M. gallopavo*, I estimated rates of synonymous divergence (d_S) (Wright et al. 2012), as d_S provides a relative measure of the relative length of time gametologs have differentiated from an ancestral autosomal gene. For each species, I obtained Z and W coding sequences from Biomart, and aligned the translated protein sequences with MUSCLE (Edgar 2004b, a) in MEGA5 (Tamura et al. 2011). Alignments were visually inspected and poorly aligned regions were removed. The CODEML package in PAML v4.7 (Yang 2007) was used to calculate maximum likelihood estimates of synonymous (d_S) and nonsynonymous (d_N) divergence in pairwise comparisons and generate standard errors. I verified $d_S \leq 1$ for all gametologs, thereby

avoiding inaccurate divergence estimates due to mutational saturation and double hits (Axelsson et al. 2008).

Three *A. platyrhynchos* gametologs had $d_S < 0.02$. I verified that the annotated Z and W genes were distinct coding sequences by identifying fixed differences between males and females from RNA-Seq data and between Z- and W-linked Ensembl annotated sequences.

Some gametologs had multiple W paralogs, and to identify the true W-linked ortholog I used the method outlined in Wright et al (Wright et al. 2012). Briefly, I used the gametologous pair with the lowest d_S estimate to avoid potential problems of relaxed or diversifying selection acting on gene copies (Graur and Li 2000) likely due to gene duplications following the suppression of recombination (Busby et al. 2011). However, the *A. platyrhynchos* draft assembly contains short sequences that may represent partial fragments of longer W-linked genes. To avoid biasing divergence estimates on short stretches of sequence, I excluded W-linked genes if coverage of the *G. gallus* W-linked ortholog was <25% and the aligned length with the Z-linked ortholog was <150bp. This resulted in the removal of three *A. platyrhynchos* W-linked genes from the analysis. All *M. gallopavo* W-linked genes met these criteria, and therefore none were excluded.

For each species, I mapped d_S estimates to the Z chromosome, taking physical position from the orthologous *G. gallus* Z-linked ortholog. Divergence dates for gametologs were calculated using a molecular clock. Briefly, based on male-mutation bias estimates in birds, the molecular clock accounts for sex-specific mutation rates in Z- and W-linked genes. This results in a divergence time of $d_S / 3.8 \times 10^{-9}$ years (Dimcheff et al. 2002; Hedges 2002; Harlid et al. 2003; Axelsson et al. 2004; Berlin et al.

2006; Webster et al. 2006; Nam and Ellegren 2008).

I used *M. gallopavo* gametologs to independently confirm the four strata previously identified in *G. gallus* (Wright et al. 2012) and to refine the dates at which recombination was suppressed. I assessed longer-term dynamics of sex chromosome divergence using *A. platyrhynchos* gametologs. To do so, I built gene trees to determine whether recombination ceased independently between gametologs in each species, or before the last common ancestor of the Galloanserae. Specifically I; 1) obtained coding sequences for gametologs in the *G. gallus*, *A. platyrhynchos, M. gallopavo* and *T. guttata* from Biomart, and aligned the translated protein sequences with MUSCLE (Edgar 2004b, a) in MEGA5 (Tamura et al. 2011). I checked alignments by eye and removed poorly aligned regions. *T. guttata* Z-linked orthologs, identified from a reciprocal BLAST with *G. gallus*, were used as outgroups; 2) I constructed maximum-likelihood gene trees with 1000 bootstrap replicates in MEGA5 (Tamura et al. 2011). In the absence of overlapping sequences, gene trees were constructed between *G. gallus* and *A. platyrhynchos* or *M. gallopavo* gametologs separately. In all cases, the *T. guttata* Z-linked ortholog was designated as the outgroup.

In order to measure the completeness of recombination suppression between gametologs, I tested for evidence of gene conversion using GENECONV (Sawyer 1999). GENECONV identifies potential gene conversion sites as long fragments of identical sequence bounded by variable sites, assessing significance by 10,000 random permutations of variable sites across the alignment. P-values for global fragments are corrected for multiple comparisons and sequence length. I tested for gene conversion using the multiple sequence alignments described above (used to build gene trees) and specified the *–group* option to look only for conversion between gametologs in *G.*

gallus, *A. platyrhynchos* and *M. gallopavo*. I used the –Seqtype=SILENT option to specify coding sequence and that gene conversion fragments should only be detected using silent site amino acid polymorphisms. High false-positive rates have been documented when this option is not implemented, particularly if genes are subject to contrasting selective regimes (Ezawa et al. 2006; Ezawa et al. 2010). I used the strict default *-gscale* value (*gscale* = 0) where no mismatches are permitted between aligned putative conversion fragments and repeated the analysis with a relaxed mismatch penalty (*gscale* = 2) to ensure the number of mismatches did not bias the analysis. GENECONV detects inner fragments, which arise from gene conversion events between the ancestors of two aligned sequences and outer fragments, which represent a conversion events originating outside of the alignment or events whose signature has been weakened by subsequent mutation. I identified significant inner gene conversion fragments between gametologs as evidence of ongoing recombination between the sex chromosomes.

On closer inspection, gene conversion events identified in this study span several exons that in turn are separated by extremely dissimilar intronic sequences. As GENECONV identifies potential gene conversion sites as long fragments of sequence bounded by variable sites, it estimates the maximum sequence length over which recombination acts. To avoid falsely identifying gene conversion sites, I excluded exons where GENECONV fragments span less than 50 bp. I used informative divergent sites and maximum-likelihood exon trees to establish the direction of gene conversion.

3.3.4 Sequence evolution of Z- and W-linked genes

To assess the selective regime acting on Z and W coding sequences, I used the branch models and branch-site test in the CODEML package in PAML v4.7 (Yang 2007).

Branch models (model=2, nssites=0) allow ω (d_N/d_S) to vary across branches. By comparing a model where ω is estimated, to one where it is fixed to equal 1 (the expected d_N/d_S under neutral evolution) or 0 (strict purifying selection) it is possible to assess the contribution of neutral and purifying selection to coding sequence evolution. To do this, I used the multiple sequence alignments described above (used to build gene trees and test for gene conversion). If *A. platyrhynchos* and *M. gallopavo* gametologs ceased recombining before the last common ancestor of the Galloanserae the following gene tree was specified (*T. guttata Z,* (((*G. gallus* W, *M. gallopavo* W), *A. platyrhynchos* W), ((*G. gallus* Z, *M. gallopavo* Z), *A. platyrhynchos* Z))) and if *A. platyrhynchos* gametologs diverged independently the following gene tree was used (*T. guttata Z*, (((*G. gallus* W, *M. gallopavo* W), (*G. gallus* Z, *M. gallopavo* Z)), (*A. platyrhynchos Z, A. platyrhynchos W*))).

The branch-site model (model=2, nssites=2) allows ω to vary across branches and among sites and permits tests to detect positive selection acting on a subset of sites along specific lineages. Under this model, I estimated ω across W and Z branches separately and then tested whether ω was significantly different to 1 using the settings fix_omega=1,omega=1.

3.4 RESULTS

3.4.1 Characterising the coding content of the avian W chromosome across species

I used a combination of bioinformatics, molecular genetic and phylogenetic methods to validate putative W-linked genes and identify previously unknown W-linked coding content in the wild turkey (*Meleagris gallopavo*) and mallard duck (*Anas platyrhynchos*). My combined dataset comprised 14 W-linked genes in *M. gallopavo* and 21 in *A. platyrhynchos*, together with 27 previously identified W-linked genes in the red jungle fowl (*G. gallus*) (Table 3.1). Some W-linked genes are present in multiple copy number (Backstrom et al. 2005; Moghadam et al. 2012) and these W-linked genes correspond to 7 gametologs in *M. gallopavo*, 11 gametologs in *A. platyrhynchos and* 20 gametologs in *G. gallus*. To date, this is the largest cross-species dataset of W-linked coding sequences yet assembled.

3.4.2 Four Z chromosome strata are conserved in the Galliformes

Sex chromosome strata are identified from clusters of gametologs with similar synonymous (d_S) divergence estimates (Lahn and Page 1999; Skaletsky et al. 2003), which provide a measure of the relative length of time gametologs have been isolated from each other. Previously, I identified four evolutionary strata on the *G. gallus* Z chromosome, where recombination was suppressed progressively over 130 million years (Wright et al. 2012).

I used seven gametologs in *M. gallopavo* to independently verify the *G. gallus* strata. Z-linked physical position was conserved between *G. gallus* and *M. gallopavo* orthologs, and each stratum is represented by at least one *M. gallopavo* gametolog.

Table 3.1. Known and newly identified W-linked coding genes.

Gene name		Z-linked gene ID[a]	W-linked gene ID[a]
G. gallus	*HINT1*	00428	22674[1], 22683[2], 22690[2](Ayers et al. 2013), 22679[2]
	CHD1	14642	*CHDW*[2](Ayers et al. 2013)
	RASA1	17706	22611[1,2]
	KCMF1	15391	14441
	MIER3	14721	00140[1,2], 14584[1,2]
	SPIN	14916	*SPINW*[2]
	HNRNPK	12591	14366[1,2]
	CZH18ORF25	01763	01585[1,2]
	VCP	01986	00386[1](Moghadam et al. 2012)
	SIAT8C	03049	26991
	ZFR	03235	14545[1,2]
	NIPBL	03605	13312[1,2]
	UBAP2	13809	05785
	ATP5A1	14644	01756[1,2]
	MADH2	14697	14184
	BTF3	13512	00395[1,2]
	UBE2R2	01668	09227
	ZSWIM6	14734	27170
	ZNF532	02852	14003[2]
	RPL17L	02696	22174[1,2]
	Novel	-	29099
	Novel	-	25736
	Novel	-	25865
M. gallopavo	*HINT1*	06655	16898, 09917, *HINTW*(AY713488.1)[3]
	RASA1	08292	*RASA1W*(AH015047.1)[4]
	SPIN	02500	*SPINW*(AH015050.1)[4], 06743
	VCP	-*	09037
	NIPBL	01989	13606, *IDN3W*(AH015048.1)[4]
	UBAP2	01837	*UBAP2W*(AY188758.1)[4], 05595
	ATP5A1	01414	09963
	MADH2	*MADH2Z*[4]	*MADH2W*(AH015049.1)[4]
	Novel	-	16394
A. platyrhynchos	*CHD1*	09965	05191, 02506
	RASA1	05627	05611, 11371, 10611
	KCMF1	13426	03026, 03106
	MIER3	06634	10850
	SPIN	13922	02923
	HNRNPK	10856	10986
	VCP	04533	05806
	ZFR	15627	15519
	NIPBL	08473	10290, 05315, 10560, 02953, 03022
	UBE2R2	03800	16000
	ZSWIM6	06992	13555,14338
	UBAP2	06449	16155

[a]Ensgalg000000(*G.gallus*)/Ensmgag000000(*M.gallopavo*)/Ensaplg000000(*A.platyrhynchos*)
[1]Moghadam et al 2012, [2]Ayers et al 2013, [3]Backstrom et al 2005, [4]Berlin and Ellegren 2006
*Amalgamation of Z and W sequences in current genome assembly

For all gametologs, d_S estimates were consistent with orthologous *G. gallus* Z-W d_S estimates and 95% confidence intervals overlapped (Fig. 3.1A, Table S.3.2). Maximum likelihood gene trees of *G. gallus* and *M. gallopavo* gametologs, with the zebra finch (*Taeniopygia guttata*) Z-linked ortholog included as an outgroup, indicate that all four strata predate the *G. gallus-M. gallopavo* divergence estimated at roughly 30 million years ago (Dimcheff et al. 2002).

For all seven *M. gallopavo* gametologs, the gene trees clustered by sex chromosome rather than by species, and the phylogenetic grouping of *G. gallus* and *M. gallopavo* W-linked genes showed high bootstrap support (in all cases ≥95%, Fig. 3.1A, Fig. S.3.2), indicating that gametolog divergence predates the common ancestor of *G. gallus* and *M. gallopavo*. This provides independent verification that the four recombination suppression events previously identified in *G. gallus* are conserved across the Galliformes.

Divergence dates for recombination suppression in the youngest strata should predate the divergence between *G. gallus* and *M. gallopavo* approximately 30 million years ago (Dimcheff et al. 2002). Using a molecular clock that accounts for sex-specific mutation rates in Z- and W-linked genes (Wright et al. 2012), I found that Stratum III, corresponding to 17-43 Mb (d_S = 0.156-0.268), arose between 41 and 71 mya and Stratum IV, corresponding to 0-11 Mb (d_S = 0.137-0.257), between 36 and 68 mya (Fig. 3.1A). These dates are consistent with the finding that four strata are conserved in the Galliformes.

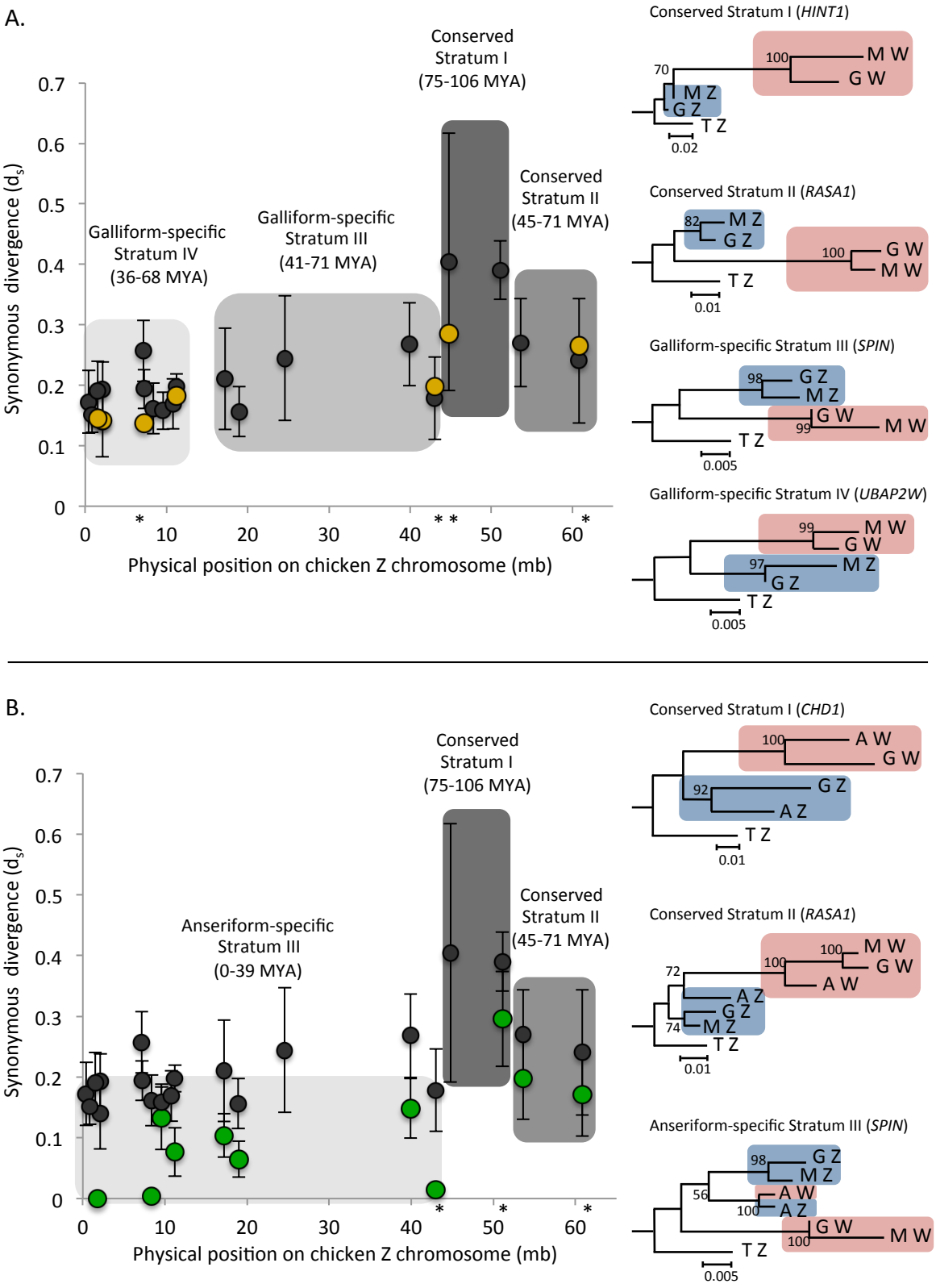
A.
Conserved Stratum I (75-106 MYA)
Conserved Stratum II (45-71 MYA)
Galliform-specific Stratum IV (36-68 MYA)
Galliform-specific Stratum III (41-71 MYA)
Synonymous divergence (ds)
Physical position on chicken Z chromosome (mb)
Conserved Stratum I (HINT1)
Conserved Stratum II (RASA1)
Galliform-specific Stratum III (SPIN)
Galliform-specific Stratum IV (UBAP2W)
B.
Conserved Stratum I (75-106 MYA)
Conserved Stratum II (45-71 MYA)
Anseriform-specific Stratum III (0-39 MYA)
Synonymous divergence (ds)
Physical position on chicken Z chromosome (mb)
Conserved Stratum I (CHD1)
Conserved Stratum II (RASA1)
Anseriform-specific Stratum III (SPIN)

Fig. 3.1: Dynamics of recombination suppression on the avian sex chromosomes.

The distribution of synonymous (d_S) divergence estimates between *G. gallus* (black), *M. gallopavo* (yellow, Panel A) and *A. platyrhynchos* (green, Panel B) gametologs, mapped against physical position of the *G. gallus* Z-linked ortholog. The 95% confidence intervals were calculated using standard errors. Bootstrap values were calculated using 1000 permutations. AZ, GZ, MZ, TZ corresponds to *A. platyrhynchos*, *G. gallus*, *M. gallopavo* and *T. guttata* Z-linked genes. The physical Z-linked position of the loci for which gene trees are shown are indicated (*). *M. gallopavo* d_S estimates map closely onto the *G. gallus* strata boundaries, and maximum-likelihood gene trees for all gametologs cluster by sex chromosome, not species, indicating that four strata are conserved across the Galliformes (Fig. S.3.2).

A. platyrhynchos d_S estimates map closely onto the two oldest *G. gallus* strata boundaries, and maximum-likelihood gene trees for gametologs within these regions cluster by sex chromosome indicating that these strata are conserved across the Galloanserae. For the remaining gametologs, maximum-likelihood gene trees reveal that recombination was suppressed independently in the *A. platyrhynchos* lineage (Fig. S.3.3). The d_S estimates are highly heterogeneous and locus specific, indicating gradual divergence of gametologs with residual recombination across large tracts of the sex chromosomes.

3.4.3 Recombination suppression evolved independently in the Anseriformes

Using 11 newly-identified *A. platyrhynchos* gametologs I calculated d_S estimates and maximum-likelihood gene trees to assess the dynamics of recombination suppression and sex chromosome divergence. Avian genomes are relatively stable, with few major genomic rearrangements, potentially due to a lack of active transposons (Toups et al. 2011). As synteny of the Z chromosome is highly conserved among all extant birds (Vicoso et al. 2013b), including the Galloanserae *A. platyrhynchos* and *G. gallus* (Skinner et al. 2009), I took physical position of the *A. platyrhynchos* gametologs from the location of the corresponding *G. gallus* Z-linked ortholog. To ensure W-linked paralogs were not mis-identified as Z-linked orthologs, I used PCR to confirm they were not physically female-limited.

For the three *A. platyrhynchos* gametologs located in the two oldest Galliform strata, d_S estimates were consistent with orthologous *G. gallus* Z-W d_S estimates, where 95% confidence intervals overlapped (Fig. 3.1B, Table S.3.3). Maximum-likelihood gene trees for all loci cluster by sex chromosome rather than by species, with 100% bootstrap support for the grouping of *G. gallus* and *A. platyrhynchos* W-linked orthologs distinct from Z-linked genes (Fig. 3.1B, Fig. S.3.3). For *RASA1Z* and *RASA1W*, I have coding sequence for *G. gallus*, *A. platyrhynchos* and *M. gallopavo* gametologs, and W-linked sequences cluster together with 100% bootstrap support. The three gametolog sets therefore convergently indicate that the two oldest strata are conserved across the Galloanserae. These strata correspond to 45-51 Mb (d_s = 0.285-0.404, Conserved Stratum I), 54-61 Mb (d_s = 0.171-0.271, Conserved Stratum II), and arose between 75 and 106 mya, and 45 and 71 mya respectively.

In contrast, of the eight *A. platyrhynchos* gametologs located in the youngest Galliform strata, d_S estimates for five gametologs are significantly lower than the corresponding *G. gallus* gametolog d_S estimates, where 95% confidence intervals are non-overlapping (Fig. 3.1B, Table S.3.3). Of these eight *A. platyrhynchos* gametologs, four cluster together in maximum-likelihood gene trees with bootstrap values ≥ 95, indicating that recombination suppression occurred relatively recently and independently in the Anseriformes in this region. However, owing to low bootstrap values I was unable to resolve the phylogenetic topology for the other four gametologs. Using a molecular clock, I calculated that recombination ceased independently in Anseriformes between 0 and 39 million years ago (Fig. 3.1B), well after the divergence of Galliformes and Anseriformes approximately 90 million years ago (van Tuinen and Hedges 2001).

3.4.4 Heterogeneous divergence of young Anseriform gametologs

There is marked heterogeneity in d_S across the youngest region of the Anseriform sex chromosomes, ranging from 0-0.148. This variation is greater than that observed across any other avian strata, and is not a result of a single outlier. Three gametologs have $d_S < 0.02$, indicating either very recent divergence or persistent recombination. I verified that Z- and W-linked coding sequences were distinct by identifying fixed differences between males and females from RNA-Seq data and between Z-linked and W-linked Ensembl annotated sequences. Therefore, the heterogeneity in divergence is not consistent with instantaneous suppression of recombination resulting from a large-scale intra chromosomal event such as an inversion. Instead, this may suggest more gradual divergence of gametologs with

residual recombination across large tracts of the sex chromosomes.

To further assess the possibility of ongoing recombination between gametologs, I identified signatures of gene conversion between Z- and W-linked coding sequences in *G. gallus*, *A. platyrhynchos* and *M. gallopavo* using GENECONV (Sawyer 1999). GENECONV identifies potential gene conversion sites as long fragments of identical sequence bounded by variable sites. Fragments that are significantly longer than expected given the length of the whole sequence, and therefore indicative of past conversion events, are identified by 10,000 random permutations of variable sites across the alignment.

Three gametologs in *A. platyrhynchos* showed evidence of significant inner gene conversion following a permutation test with 10,000 iterations (Table 3.2). I repeated the analysis allowing mismatches between gene conversion fragments (*gscale* = 2) and found no significant difference between the results. These gametologs are uniformly distributed across the Z chromosome, located in Conserved Strata I, II and Anseriform-specific Stratum III, indicating ongoing recombination along the whole length of the avian sex chromosomes. In total, I detected four Z- and four W-linked exons subject to gene conversion (Table 3.2). Of these, I could determine the direction of gene conversion for the exons in the Conserved Strata I and II. Informative variant sites indicate genetic material was transferred from *KCMF1W* to *KCMF1Z* and from *CHD1Z* to *CHD1W*. Exon tree topologies support the proposed direction of gene conversion in *KCMF1* and cluster by species for *CHD1*, revealing putative gene conversion across *G. gallus CHD1* gametologs, the signal of which was not detected with GENECONV.

Table 3.2. Gene conversion between gametologs in *A. platyrhynchos.*

Gene name	Gene ID (Z/W)[a]	Stratum	p-value [b/c]	Spanning Z-linked exon no.[d]	Spanning W-linked exon no.[d]	Length of gene conversion track (bp)
CHD1	09965/ 05191	Conserved I	0.0006/ 0.0027*	27	7	174
KCMF1	13426/ 03026	Conserved II	0.0056/ 0.0273	4	3	142
HNRNPK	10856/ 10986	Anseriform-specific III	0.0081/ 0.0294	10	8	255

[a]Ensaplg000000...
[b]Simulated p-values are based on 10,000 permutations.
[c]Karlin-Altshul p-values are Bonferroni-corrected KA (BLAST-like)
[d]Exons where gene conversion fragments span <50bp are not listed
*Significant after correcting for multiple comparisons

I used exonic data to identify signatures of gene conversion, and as a result cannot accurately distinguish the exact conversion boundaries. This may help explain my finding that the conversion events I detect are surrounded by divergent intronic sequence. It is possible that gene conversion extends into the introns, but I was unable to detect this. Sex-linked intronic sequences are notoriously difficult to assemble and Z- and W-linked sequences are most likely concatenated together in Ensembl. Therefore, without separately sequencing Z- and W-linked introns, my results are only indicative and not conclusive evidence for gene conversion.

Gene conversion between gametologs can lower d_S estimates and generate molecular gene trees where gametologs cluster by species. This may lead to false conclusions that recombination was suppressed independently in a given lineage. I detected no gene conversion between the four *A. platyrhynchos* gametologs where gene trees cluster by species, or the three gametologs with $d_S < 0.02$. However, previous studies have shown that GENECONV's power to detect gene conversion events between similar sequences is limited (McGrath et al. 2009; Ezawa et al. 2010). I recalculated d_S for three *A. platyrhynchos* gametologs after removing exons subject to gene conversion and as expected, the estimated divergence time increased. However, the range of divergence dates over which recombination was suppressed in Conserved Stratum I and II remained constant. The recalculated gene tree for the gametolog in Anseriform-specific Stratum III revealed a complex mosaic of recombination across specific exons. Before excluding the exons subject to gene conversion, the phylogenetic topology was unresolvable, however after removal of these exons, *A. platyrhynchos* and *G. gallus HNRNPKW* clustered together with high bootstrap support (≥95%).

3.4.5 Comparative analysis of W-linked coding content evolution

Studies examining the evolutionary forces acting after recombination suppression are mainly limited to male heterogametic systems (Chibalina and Filatov 2011; Hughes et al. 2012; Bachtrog 2013). As a result, the factors driving the differentiation and degeneration of Z and W sex chromosomes are partially unclear. Now equipped with the largest dataset of W-linked genes to date, I used the branch models and branch-site test in the CODEML package in PAML v4.7 (Yang 2007) to assess the selective regime driving Z and W coding sequence change over 90 million years of avian evolution.

For W-linked branches, ω (d_N/d_S) was significantly lower than 1 in 14 gametolog families, indicating these genes are evolving primarily due to purifying selection (Table 3.3, Table S.3.4). The exception was *HINT1W*, where ω was not significantly different to 1. However, *HINT1W* is a known ampliconic gene with documented gene conversion among W-linked copies (Backstrom et al. 2005). This likely violates the assumptions made in CODEML, where d_S is equivalent to the neutral mutation rate (Yang 2007). For Z-linked branches, ω was not significantly different to 0 across four loci, indicating these are extremely conserved and evolving under strict purifying selection. For the other 10 loci, ω was significantly different to both 1 and 0 for all but one locus, *UBAP2Z* (Table S.3.5).

I tested for positive selection using the branch-site model. This model allows ω to vary across both branches and among sites, and tests for adaptive evolution acting on a subset of codons along a branch(es) defined a priori (Zhang et al. 2005). I found no evidence of positive selection on any Z-linked gene families (Table S.3.5) but found

Table 3.3. Branch models and branch-site test for W-linked branches.

Gene name	Species included [a]	ω	Branch model test (ω =1) Likelihood Ratio	Branch model test (ω =0) Likelihood Ratio	Branch-site test (ω =1) Likelihood Ratio
HINT1W	G,M,T	1.10	0.05	**3697.05****	0.05
CHD1W	G,A,T	0.14	**94.45****	**7012.15****	0.00
RASA1W	G,A,M,T	0.09	**31.33****	**1386.95****	0.00
KCMF1W	G,A,T	0.22	**35.49****	**5674.71****	0.22
MIER3W	G,A,T	0.24	**17.97****	**167.25****	0.00
SPINW	G,A,M,T	0.09	**28.45****	**1749.98****	0.00
HNRNPKW	G,A,T	0.03	**108.75****	**84.21****	0.03
VCPW	G,A,T	0.03	**25.58****	**19.13****	0.16
ZFRW	G,A,T	0.18	**27.16****	**173.50****	0.00
NIPBLW	G,A,T	0.05	**44.08****	**63.91****	0.00
NIPBLW	G,M,T	0.03	**38.40****	**343.5****	0.00
UBAP2W	G,M,T	0.15	**6.19***	**354.82****	0.00
ATP5A1W	G,M,T	0.09	**16.54****	**650.83****	**3.58***
MADH2W	G,M,T	0.06	**22.37****	**333.50****	0.01
UBE2R2W	G,A,T	0.04	**15.03****	**17.39****	0.00
ZSWIM6W	G,A,T	0.25	**11.25****	**128.88****	0.00

[a]G=*G. gallus*, M=*M. gallopavo,* A=*A. platyrhynchos,* T=*T. guttata*
** p-value <0.01
* p-value <0.05
Likelihood values for each model and the gene IDs are listed in Table S.3.4.
Significant likelihood ratio values are in bold

significant evidence for positive selection acting on one W-linked locus, *ATP5A1W*, (proportion of sites where $\omega > 1$ was 0.01 and $\omega = 17.70$) (Table 3.3, Table S.3.4).

Gene conversion violates the assumption in PAML that d_S is constant across a gene and the GC bias in the mismatch repair process can inflate ω (d_N/d_S) (Berglund et al. 2009). Consequentially, previous studies have found that gene conversion can lead to false signals of positive selection (Casola and Hahn 2009; Ratnakumar et al. 2010). Although I found no evidence of gene conversion between *ATP5A1W* and *ATP5A1Z* in any species, *ATP5A1W* is present in two copies in the *G. gallus* genome (Moghadam et al. 2012) and may be subject to intra-chromosomal gene conversion. In contrast, my polymorphism data indicate there is only a single copy in the *M. gallopavo* genome, as I failed to identify polymorphisms in this locus within any single female, which would indicate multiple copy number. In order to exclude the possibility that gene conversion between W-linked *G. gallus* paralogs is driving a false signal of positive selection, I removed *G. gallus ATP5A1W* and repeated the analysis. After excluding this gene, the branch-site test for positive selection remained significant (proportion of sites where $\omega > 1$ was 0.01 and $\omega = 41.87$). *ATP5A1W* encodes a mitochondrial ATP synthase subunit. As the W chromosome and mitochondrial genome are both maternally inherited and effectively linked (Berlin et al. 2007), one plausible explanation for this signature of positive evolution may be that it reflects adaptive co-evolution.

3.5 DISCUSSION

The process of recombination suppression is a crucial component of sex chromosome evolution, facilitating the divergence and differentiation of gametologs (Charlesworth and Charlesworth 1978). Studies from highly differentiated sex

chromosomes suggest that recombination is suppressed instantaneously in a stepwise manner by intra-chromosomal rearrangements (Lahn and Page 1999; Pandey et al. 2013). However, my ability to determine whether inversions are a cause or consequence of this process is hampered by the limited numbers of sex-linked genes in these species. Conversely, in newly evolved sex chromosomes, recombination appears to be suppressed in a gradual and heterogeneous nature (Chibalina and Filatov 2011; Muyle et al. 2012; Bergero et al. 2013; Natri et al. 2013; Qiu et al. 2013).

I used a comparative approach to study the dynamics of sex chromosome recombination suppression on the largest collection of W-linked coding content yet assembled across 90 million years of avian evolutionary history. My results reveal that recombination between the Z and W chromosomes ceased independently multiple times across different avian lineages. The pattern I uncover is consistent with both the inversion model (Charlesworth et al. 2005) where evolutionary strata form in a stepwise process, and heterogeneous and locus specific recombination suppression over short-term periods. Additionally, I show that recombination suppression is not necessarily complete, as highly diverged gametologs recombine across specific exons via gene conversion.

3.5.1 Dynamics of sex chromosome divergence

Using both maximum-likelihood gene trees and synonymous divergence estimates, I confirm that four strata are conserved across the Galliformes (Fig. 3.2). These four strata were previously identified in *G. gallus* (Wright et al. 2012), and follow the same stepwise pattern observed across the mammalian X chromosome (Lahn and Page 1999; Sandstedt and Tucker 2004; Pearks Wilkerson et al. 2008; Van Laere et al.

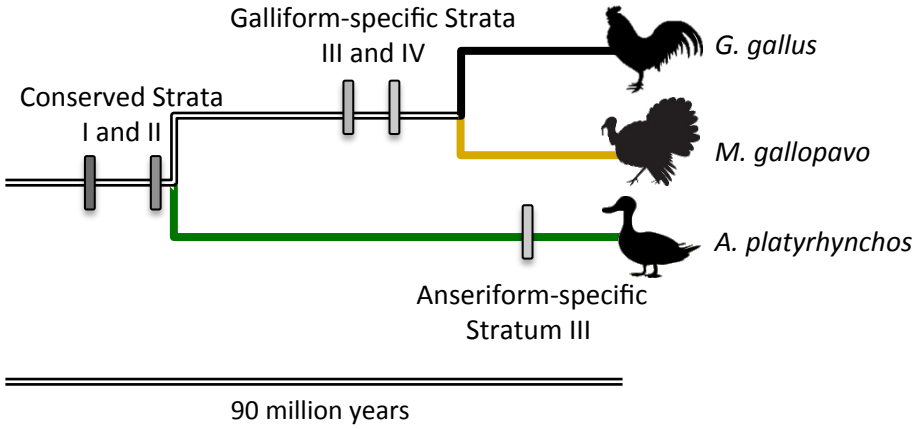

Fig. 3.2. Independent expansion of the non-recombining regions across the Galloanserae.

Phylogeny illustrating the independent cessation of recombination on the sex chromosomes in the Galliformes and Anseriformes. Estimated date of recombination suppression (millions of years), calculated using a molecular clock that accounts for sex-specific mutation rates in Z- and W-linked genes, is mapped onto the phylogeny.

2008). Using newly identified gametologs, I was able to refine the dates at which recombination was suppressed across the Galliform strata. I find that the oldest strata originated approximately 90 million years ago, and the youngest 46 million years ago, both of which are consistent with the divergence date between *G. gallus* and *M. gallopavo* (Dimcheff et al. 2002).

Similarly, maximum-likelihood gene trees reveal that the two oldest Galliform strata are conserved on the Anseriform Z chromosome. In contrast, recombination suppression is heterogeneous across the remainder of the *A. platyrhynchos* sex chromosomes, and the topology of gene trees revealed that gametologs diverged independently in this region (Fig. 3.1). Divergence estimates vary from 0 to 39 MYA, a range that is greater than that observed across any other avian strata, and not a result of a single outlier. I verified that Z- and W-linked sequences were in fact distinct by identifying fixed differences between males and females from RNA-Seq data and between Z-linked and W-inked Ensembl annotated sequences. This heterogeneity in divergence is not consistent with instantaneous suppression of recombination resulting from a large-scale intra-chromosomal event and is instead indicative of Z-W divergence combined with ongoing recombination. However, due to the limited number of gametologs, I lack sufficient power to formally test for differences in the variance of d_S across strata.

The lack of positional information for the *A. platyrhynchos* Z chromosome, means I cannot rule out the possibility that Anseriform-specific Stratum III is actually made up of two distinct strata formed by a lineage specific inversion. Under this scenario, the three gametologs with d_S estimates < 0.02 would comprise the youngest stratum with ongoing recombination between the W and Z orthologs. Positional

information may support this scenario, as two of the three gametologs are located between 0-10 Mb on the Z chromosome.

Some gametologs are subject to gene conversion, lending further support to the notion that recombination suppression is not complete. Three *A. platyrhynchos* gametologs showed evidence of significant gene conversion, and were distributed across the length of the Z chromosome. Therefore, even for highly differentiated gametologs, certain exons are prevented from diverging. This is consistent with previous studies documenting gene conversion between X- and Y-linked orthologs in mammals (Iwase et al. 2010; Trombetta et al. 2014). Gene conversion may therefore be an important mechanism halting the degeneration of sex-limited gametologs (Rosser et al. 2009) in addition to the intra-chromosomal gene conversion documented on the avian W and mammalian Y chromosomes (Backstrom et al. 2005; Marais et al. 2010; Hallast et al. 2013; Bellott et al. 2014). Interestingly, a previous study used whole chromosome-painting to show that the *A. platyrhynchos* W chromosome is highly divergent in comparison to closely related avian W chromosomes (Stiglec et al. 2007). I find a high degree of conservation among coding genes, indicating exon-specific gene conversion may facilitate large-scale turnover of intronic and repetitive regions whilst still conserving gene function.

This heterogeneous pattern of sex chromosome divergence and ongoing recombination is consistent with *G. aculeatus* and *S. latifolia* sex chromosomes. The *S. latifolia* XY system originated 10 million years ago, and a neo-sex chromosome formed 2 million years ago in *G. aculeatus*. In both these system, recombination suppression evolves heterogeneously (Chibalina and Filatov 2011; Muyle et al. 2012; Bergero et al. 2013; Natri et al. 2013; Qiu et al. 2013). The underlying mechanism of

recombination suppression is unclear, but changes in heterochromatin and methylation, which have been shown to alter patterns of recombination (Mirouze et al. 2012), or local changes in the binding sequence that recruits recombination machinery to the chromosome (Baudat et al. 2010; Myers et al. 2010), may be responsible.

My divergence estimates for recombination suppression are consistent with divergence dates for *G. gallus*, *A. platyrhynchos* and *M. gallopavo* with the exception of Conserved Stratum II. Using a molecular clock that accounts for sex-specific mutation rates in Z- and W-linked genes, the divergence of Conserved Stratum II was estimated to have occurred between 45 and 71 mya. Galliformes and Anseriformes diverged 90 million years ago (van Tuinen and Hedges 2001). Although these estimates are not vastly distant, they are inconsistent and may indicate that the molecular clock underestimates the time of gametolog divergence due to the unique evolutionary forces acting on sex chromosomes (Bachtrog et al. 2011).

3.5.2 Evolution of heterogametic sex chromosomes after recombination suppression

Studies examining the evolutionary forces acting after recombination has ceased have mainly been limited to male heterogametic systems, such as the Y chromosomes of mammals and *Drosophila* (Chibalina and Filatov 2011; Hughes et al. 2012; Bachtrog 2013). These studies have uncovered strong purifying selection acting on the Y chromosome (Kaiser et al. 2011; Hughes et al. 2012; Bellott et al. 2014; Wilson Sayres et al. 2014) together with signatures of positive selection (Gerrard and Filatov 2005; Li et al. 2013). However, the selective forces driving the evolution of the W chromosome are unclear. This is partly due to a previous paucity of known W-linked

genes, which has restricted the potential of inter-specific analyses (Berlin and Ellegren 2006).

Theoretical models predict key differences between W and Y chromosome evolution (Kirkpatrick and Hall 2004b; Mank 2012; Wright and Mank 2013). When sexual selection is acting primarily on males, the Y chromosome has a lower effective population size than W-linked sequences (Wright and Mank 2013), and is subject to male-mutation bias (Kirkpatrick and Hall 2004b). These factors have been hypothesized to result in rapid degeneration of Y-linked coding sequences, a pattern that has been observed on *Drosophila* and primate Y chromosomes (Hughes et al. 2010; Hughes et al. 2012; Bachtrog 2013).

Using the largest catalog of W-linked genes to date, I find that avian W-linked genes are evolving with a significant contribution of purifying selection. This is consistent with a previous study showing that sex-specific selection drives evolutionary changes in the expression of W-linked genes (Moghadam et al. 2012), and indicates that although the W chromosome does not determine sex directly (Smith et al. 2009), it has an important role in encoding sex-specific fitness. The large number of genes conserved across 90 million years of avian evolution is also potentially indicative of lower degeneration rates on the W compared to the Y chromosome.

I found evidence of positive selection in one W-linked gene that encodes a mitochondrial ATP synthase subunit. As the W chromosome and mitochondria are both maternally inherited and are effectively linked (Berlin et al. 2007), we might expect strong co-evolution between these portions of the genome, and this may be reflected by rapid divergence in sequence across different avian lineages.

My findings indicate that adaptive evolution together with conservation of

gene sequence is not restricted to male heterogametic systems and may be a fundamental feature of heterogametic sex chromosome evolution.

3.6 CONCLUSIONS

Recombination suppression is a key force driving sex chromosome evolution, and sex chromosomes are therefore important for the study of genome evolution, particularly the interplay between selective forces and recombination. Here, I use a comparative approach to examine short- and long-term dynamics of sex chromosome evolution. My results reveal that recombination has ceased multiple and independent times over 90 million years of avian evolution. I show that sex chromosome divergence over the long term is characterised by regions of similar divergence among gametologs, consistent with the inversion model of sex chromosome evolution (Lahn and Page 1999; Charlesworth et al. 2005; Vicoso et al. 2013a). In contrast, over short time periods I observe heterogeneous and locus specific divergence. Additionally, I uncover gene conversion between highly diverged gametologs across both long and short evolutionary trajectories. Moreover, if sex chromosomes do evolve via inversions, additional mechanisms may be required to completely suppress recombination. My data also show that the W chromosome is evolving primarily due to purifying selection, although I also detect the signature of positive selection at one locus. My findings indicate that the W chromosome plays an important role in encoding sex-specific fitness, and potentially exhibits a lower rate of degeneration compared to Y chromosome.

Chapter 4

Variation in promiscuity and sperm competition drives the evolution of the avian Z chromosome

4.1 ABSTRACT

Higher rates of coding sequence evolution have been observed on the Z chromosome relative to the autosomes across a wide range of species. However, despite a considerable body of theory, we lack empirical evidence explaining variation in the strength of Faster-Z Effect. To assess the magnitude and drivers of Faster-Z Evolution, I used six *de novo* transcriptomes, spanning 90 million years of avian evolution. My analysis combines expression, sequence and polymorphism data with measures of sperm competition and promiscuity. In doing so, I present the first empirical evidence demonstrating the positive relationship between Faster-Z Effect and measures of promiscuity, and therefore variance in male mating success. My results from multiple lines of evidence indicate that selection is less effective on the Z chromosome, particularly in promiscuous species, and that Faster-Z Evolution in birds is due primarily to genetic drift. In doing so, I reveal the power of mating system and sexual selection in shaping broad patterns in genome evolution.

4.2 INTRODUCTION

Sex chromosomes are subject to unique evolutionary forces as a result of their unusual pattern of inheritance (Charlesworth et al. 1987; Vicoso and Charlesworth 2009; Connallon et al. 2012). The magnitude of selection, genetic drift and recombination are all predicted to differ between the sex chromosomes and autosomes (Rice 1984; Kirkpatrick and Hall 2004b; Mank et al. 2010c; Meisel and Connallon 2013) and studies contrasting the evolution of sex-linked to autosomal genes can shed light on the fundamental evolutionary forces acting across the genome as a whole.

Faster rates of coding sequence evolution have been observed on the Z and X chromosomes relative to the autosomes across a wide range of species (recently reviewed by Meisel and Connallon 2013), and Faster-X and Faster-Z Effects appear to be a common feature of sex chromosome evolution. However, despite elevated rates of evolution for both X-linked and Z-linked genes, the underlying causes of Faster-X and Faster-Z Evolution are predicted to differ (Vicoso and Charlesworth 2009; Meisel and Connallon 2013). The effective population size of X and Z chromosomes (N_{EX} and N_{EZ}) is ¾ that of the autosomes (N_{EA}) when there is no difference in the variance of male and female reproductive success, such as in strictly monogamous breeding systems (Charlesworth et al. 1987). However, many forms of sexual selection cause elevated variance in male reproductive success (Andersson 1994) which reduces N_{EZ}/N_{EA}, and in extreme cases where a single male monopolises the reproductive output of many females, N_{EZ} approaches $½N_{EA}$ (Vicoso and Charlesworth 2009; Wright and Mank 2013) (Fig. 4.1). Correspondingly, genetic drift and fixation of weakly deleterious mutations is greater on the Z chromosome (Charlesworth 2009),

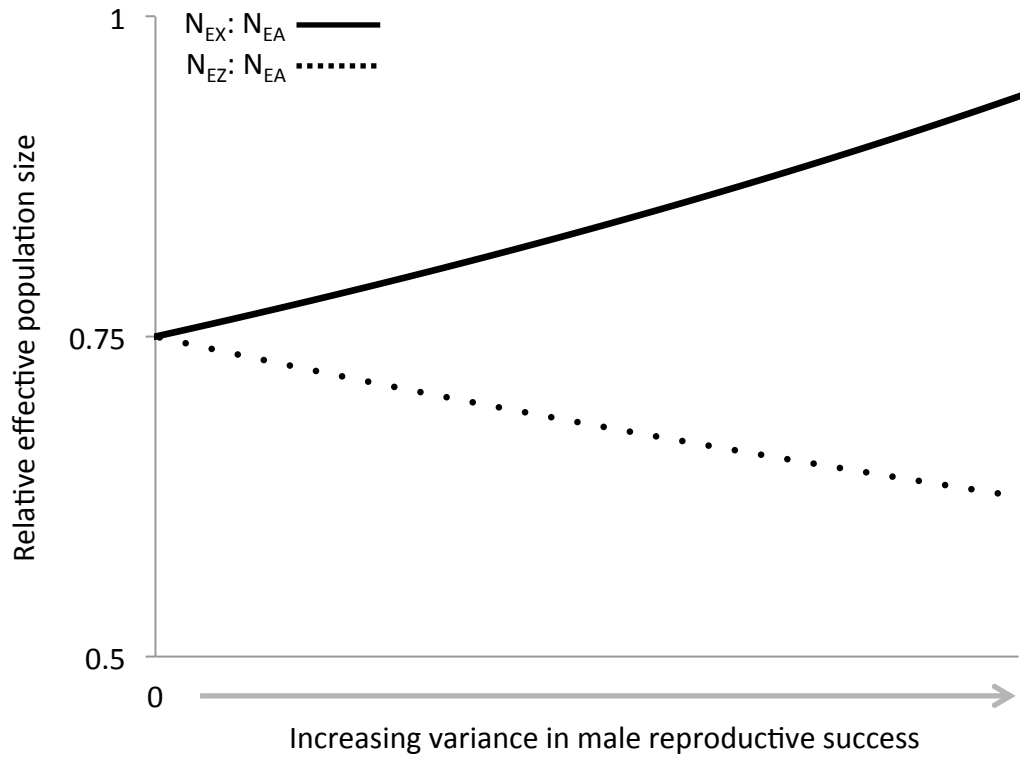

Fig. 4.1. Relationship between effective population size (N_E) and variance in male reproductive success.

Schematic outlining the predicted relationship between variance in male reproductive success and relative N_{EZ} and N_{EX} . When variance in reproductive success is the same in males and females, under monogamy, both N_{EZ} and N_{EX} = ¾ N_{EA}. As variance in male mating success increases, N_{EZ} < ¾ N_{EA} and N_{EX} > ¾ N_{EA}.

and we predict a Faster-Z Effect largely due to neutral, non-adaptive processes. Empirical evidence in birds and snakes is consistent with this non-adaptive and neutral explanation of Faster-Z (Mank et al. 2010c; Corl and Ellegren 2012; Vicoso et al. 2013a). However, silkmoths may present a recent exception (Sackton et al. 2014).

The opposite relationship is predicted for the X chromosome in male heterogametic systems (Laporte and Charlesworth 2002; Vicoso and Charlesworth 2009; Wright and Mank 2013), as increasing variance in male reproductive success results in $N_{EX}/N_{EA} > ¾$, and approaches 1 in extreme cases (Fig. 4.1). Conversely, the higher ratio of N_{EX}/N_{EA} is expected to decrease the effect of genetic drift in Faster-X Evolution. Instead, Faster-X Evolution is more often thought to be the product of increased efficacy of selection acting on recessive X-linked alleles in the heterogametic sex and increased rates of fixation of beneficial alleles relative to the autosomes. Signatures of positive selection have been uncovered on the X chromosome of mammals and *Drosophila*, consistent with adaptive Faster-X Evolution (Thornton and Long 2005; Baines et al. 2008; Hvilsom et al. 2012; Langley et al. 2012).

A key prediction of this theory is that the magnitude of Faster-Z Evolution can be explained by variation in the effective population size of the sex chromosomes relative to the autosomes driven by sexual selection (Vicoso and Charlesworth 2009). Here, I explicitly test this prediction in the Galloanserae, a clade of birds spanning 90 million years (Fig. 4.2) and for which there is extensive variation in the strength of sexual selection (Moller 1988, 1991; Birkhead and Petrie 1995). Using *de novo* transcriptomes for six Galloanserae species, I combined sequence divergence, polymorphism and expression with phenotypic measures of mating system to explicitly explore the nature Faster-Z Evolution. My results build on previous findings to reveal

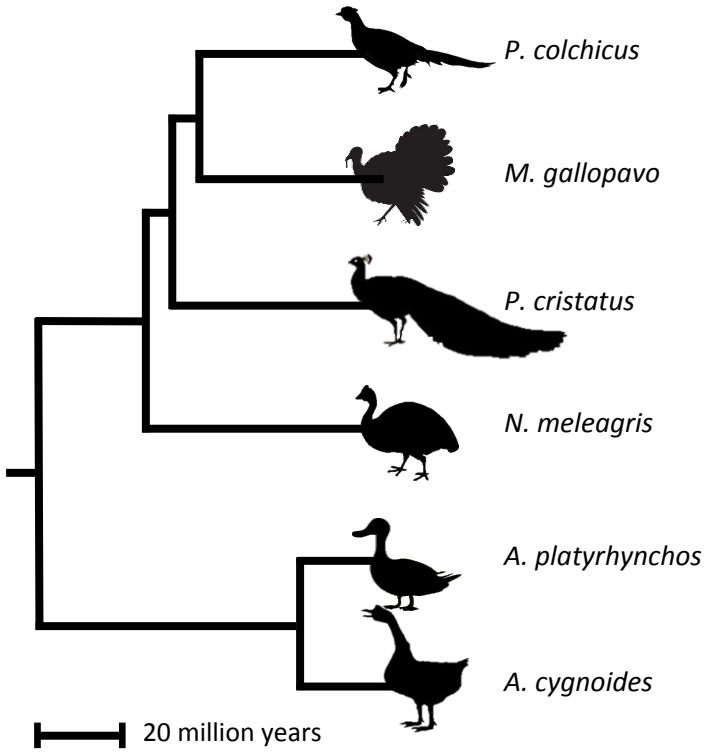

Fig. 4.2. Phylogenetic relationship of the Galloanserae species in this study.

Phylogeny of six Galloanserae species, separated by approximately 90 million years of avian evolution. There is considerable variation in the intensity of sperm competition and therefore variance in male reproductive success across this clade.

the dominant role genetic drift plays in Faster-Z. Furthermore, I uncover a positive correlation between Faster-Z and measures of sperm competition, a widely used indicator of the strength of sexual selection (Birkhead and Moller 1998), suggesting that variation in male mating success drives Z-linked divergence. In doing so, I present the first empirical evidence in support of the considerable body of theory (Charlesworth et al. 1987; Vicoso and Charlesworth 2009) outlining the relationship between sexual selection and sex chromosome evolution.

4.3 MATERIALS AND METHODS

4.3.1 De novo transcriptome assembly

RNA-Seq data was obtained from captive populations of the following Galloanserae species at the start of their first breeding season; *Anas platyrhynchos* (mallard duck), *Meleagris gallopavo* (wild turkey), *Phasianus colchicus* (common pheasant), *Numida meleagris* (helmeted guineafowl), *Pavo cristatus* (indian peafowl) *and Anser cygnoides* (swan goose) (Fig. 4.2). Samples were collected with permission from institutional ethical review committees and in accordance with national guidelines. The left gonad and spleen were dissected separately from five males and five females of each species. The exceptions were *P. colchicus*, where six male gonad and spleen samples were collected, and *M. gallopavo*, where four male and two female spleens were collected. Samples were homogenzied and stored in RNAlater until preparation. The Animal Tissue RNA Kit (Qiagen) was used to extract RNA and the samples were prepared and barcoded at The Wellcome Trust Centre for Human Genetics, University of Oxford using standard methods. RNA was sequenced on an

Illumina HiSeq 2000 resulting in on average 26 million 100 bp paired-end reads per sample.

The data was quality assessed using FastQC v0.10.1 (www.bioinformatics.babraham.ac.uk/projects/fastqc) and filtered using Trimmomatic v0.22 (Lohse et al. 2012). Specifically, reads containing adaptor sequences were removed and reads were trimmed if the sliding window average Phred score over four bases was <15 or if the leading/trailing bases had a Phred score <4. Reads were removed post filtering if either read pair was <25 bases in length. *De novo* transcriptome assemblies were constructed for each species using Trinity (Grabherr et al. 2011). All of the reads from each sample were separately mapped back to the Trinity contigs using RSEM v1.1.21 (Li and Dewey 2011) to obtain expression levels. A minimum expression filter of 2 reads per kilobase per million mapped reads (RPKM) required that each contig have expression above 2 RPKM in at least half of any of the tissues from either sex. For each Trinity contig cluster, the isoform with the highest expression level was selected for further analysis. rRNA transcripts were removed using *G. gallus* known sequences. This generated 37,453 contigs for *A. platyrhynchos*, 50,817 for *M. gallopavo,* 56,090 for *P. colchicus,* 45,535 for *N. meleagris,* 56,604 for *P. cristatus* and 44,144 for *A. cygnoides*.

4.3.2 Identification of Galloanserae orthogroups

G. gallus (Galgal4/GCA_000002315.2) cDNA sequences were obtained from Ensembl v73 (Flicek et al. 2013) and the longest transcript for each gene was identified. Orthology was determined using reciprocal BLASTn v2.2.27+ (Altschul et al. 1990) with an E-value cutoff of 1×10^{-10} and minimum percentage identity of 30%.

3116 groups of reciprocal 1-1 orthologs across all seven species (orthogroups) were identified using the highest BLAST score.

The Galloanserae last shared a common ancestor approximately 90 million years ago (van Tuinen and Hedges 2001; Hedges et al. 2006). Avian chromosome structure is stable potentially due to a lack of active transposons (Toups et al. 2011) and major genomic rearrangements are infrequent (Stiglec et al. 2007). Synteny of the Z chromosome has previously been shown to be highly conserved across both extant birds (Vicoso et al. 2013b), as well as within the Galloanserae (Skinner et al. 2009). Chromosomal location was therefore assigned from *G. gallus* reciprocal orthologs.

4.3.3 Estimating sequence divergence across orthogroups

In order to extract Galloanserae protein-coding sequences, *G. gallus* (Galgal4/GCA_000002315.2) protein sequences were obtained from Ensembl v73 (Flicek et al. 2013). For each orthogroup, each contig was translated into all potential reading frames and BLASTed against the orthologous *G. gallus* protein sequence using BLASTx. BLASTx outputs were used to determine coding frame, and protein-coding sequences for each species were extracted. Protein-coding sequences were defined as sequences starting with the amino acid M and terminating with a stop codon or end of the contig. Orthogroups with no BLASTx hits or valid protein-coding sequence were excluded.

Orthogroups were aligned with PRANK v121218 using the orthologous *Taeniopygia guttata* cDNA (taeGut3.2.4.75) as an outgroup and specifying the following guidetree; (((*A .cygnoides*, *A. platyrhynchos*), (*N .meleagris*, (*P. cristatus*, (*M. gallopavo*, *P. colchicus*)))), *T. guttata*). Retrotransposons were removed with

RepeatMasker (vopen-4.0.3) and sequences with internal stop codons were also filtered. SWAMP v0.9 (Harrison et al. In review) with a cut-off of 4 and window size of 15, and a minimum length of 75bp was used to preprocess the data.

The branch model (model=2, nssites=0) in the CODEML package in PAML v4.7a (Yang 2007) was used to obtain divergence estimates for each orthogroup, using the specified phylogeny; ((*A. cygnoides*, *A. platyrhynchos*), (*N .meleagris*, (*P. cristatus*, (*M .gallopavo*, *P. colchicus*))), *T. guttata*). Firstly, the branch model was used to calculate average d_N/d_S across all Galloanserae branches, excluding *T. guttata* outgroup. I will refer to this model as the Galloanserae model. Secondly, average d_N/d_S from the terminal tip to the Galloanserae common ancestor was calculated for each of the six Galloanserae species separately. I will refer to this model as the species-specific model. This approach ensures that the branch length over which d_N/d_S is calculated is identical for each species and therefore prevents inter-specific variation in branch length biasing my conclusions (Montgomery et al. 2011). As mutational saturation and double hits can lead to inaccurate divergence estimates (Axelsson et al. 2008), orthogroups were excluded if tree length $d_S > 2$ in the Galloanserae branch model.

4.3.4 Using sequence divergence to estimate the Faster-Z Effect

After filtering, there were divergence estimates for 160 Z-linked and 2431 autosomal orthogroups. There is considerable variation in chromosome size across the avian genomes, with many small microchromosomes (<19.5 Mb). Microchromosomes display distinct properties in contrast to the rest of the genome; specifically, they are gene dense, enriched in GC and have an elevated recombination rate (Burt 2002; Ellegren 2013). These characteristics can impact the nature of molecular evolution,

and therefore the most accurate way to measure the Faster-Z Effect is to compare divergence between the Z chromosome and similarly sized autosomes 1-10 (Mank et al. 2007a; Mank et al. 2010c).

I calculated mean d_N, d_S and d_N/d_S separately for each genomic category. Specifically, mean divergence was calculated as the sum of the number of substitutions (SdS/NdN) divided by the number of sites (S/N). This approach avoids the problems of infinitely high d_N/d_S estimates arising from genes with extremely low d_S (Mank et al. 2007a; Mank et al. 2010c), and prevents disproportionate weighting of shorter genes.

I used bootstrapping with 1000 repetitions to generate 95% confidence intervals and significance differences between genomic categories were determined from 1000 permutation tests. One-tailed p-values are reported because I specifically test whether d_N, d_S and d_N/d_S are significantly higher for Z-linked genes versus autosomal genes.

Mean Z-linked and autosomal d_N, d_S and d_N/d_S values were calculated for the whole Galloanserae (Galloanserae model) and for each of the six species (species-specific model). Faster-Z Effect was calculated as $d_{NZ}/d_{SZ} : d_{NA}/d_{SA}$.

4.3.5 Testing the relationship between sexual selection and Faster-Z Effect

To test the hypothesis that the magnitude of Faster-Z increases with increased variance in male reproductive success, I performed phylogenetically controlled regression analyses between Faster-Z ($d_{NZ}/d_{SZ} : d_{NA}/d_{SA}$) and relative N_{EZ} for each Galloanserae species, and two measures of female promiscuity. The intensity of sperm competition, a widely used proxy for the magnitude of postcopulatory sexual selection and therefore variance in male reproductive success, is strongly predicted by relative

testes weight and sperm number (Moller 1991; Moller and Briskie 1995; Birkhead and Moller 1998). Residual testes weight was calculated using the following equation describing the linear relationship between log testes mass and body mass in birds: log(testes mass(g))) = -1.56 + 0.61 log (body mass(g)) (Moller 1988, 1991; Birkhead and Petrie 1995; Pitcher et al. 2005). Log sperm number (10^6) has been measured in previous studies (Moller 1988, 1991; Birkhead and Petrie 1995). Estimates for body weight and sperm number were not available for *A. cygnoides* and therefore *A. anser* estimates were used instead, because these species are closely related (Ruokonen et al. 2000) and both exhibit strictly monogamous mating systems.

These analyses were performed using phylogenetic generalized least squares models (PGLS) in BayesTraits v2-beta (Pagel 1999; Pagel et al. 2004) with maximum likelihood and 1000 runs for each analysis. Phylogenies were obtained from birdtree.org using the Ericson dataset. I calculated mean r^2 for each regression analysis and average t-values (average correlation coefficient of the slope/average standard error correlation coefficient of the slope). A one-tailed t-test with four degrees of freedom was used to determine whether the slope was significantly greater than 0.

Differences in the rate of male-biased mutation across the six species could contribute to variation in Faster-Z Effect because the Z chromosome is more often present in males than the autosomes (Kirkpatrick and Hall 2004b). I explicitly tested for significant differences in average Z-linked d_S across the six species using permutation tests with 1000 replicates in order to verify that there were no underlying differences in mutation rate.

4.3.6 Tests of positive selection using sequence data

The site models in the CODEML package in PAML v4.7a (Yang 2007) were used to test for signatures of positive selection acting at a subset of sites. These models allow d_N/d_S to vary among sites but not across lineages. To test for positive selection I compared likelihoods from two models; M1a (Nearly neutral, model=0, nssites=1) and M2a (Positive selection, model=0, nssites=2). Under model M1a, sites can fall into one of two categories (purifying selection $d_N/d_S < 1$ and neutral evolution $d_N/d_S = 1$), whereas there is an additional category under model M2a (positive selection $d_N/d_S > 1$). The following phylogeny was specified ((*A. cygnoides*, *A. platyrhynchos*), (*N. meleagris*, (*P. cristatus*, (*M. gallopavo*, *P . colchicus*))), *T. guttata*).

4.3.7 Tests of positive selection using polymorphism data

I tested for deviations from neutrality using polymorphism data. Polymorphism data was obtained by first mapping Trinity reads to orthogroups using the two pass alignment method of the STAR aligner (Dobin et al. 2013). SNPs were called using varscan v2.3.6 (Koboldt et al. 2009; Koboldt et al. 2012) and Samtools (Li et al. 2009) following the recommendations of Quinn et al 2013 (Quinn et al. 2013). SNPs were required to have a minor allele frequency >0.15 and to be from regions where at least 4 samples had a read depth >20 and have a Phred quality >20. SNP clusters of more than 10 SNPs in 100 bases were excluded (Stevenson et al. 2013) . Valid SNPs were matched to the reading frame to determine whether they were synonymous or non-synonymous.

For each species, I calculated mean nonsynonymous polymorphism (p_N), synonymous polymorphism (p_S) and p_N/p_S separately for Z-linked and autosomal 1-10

orthogroups. Specifically, mean polymorphism was calculated as the sum of the number of polymorphic sites (SpS/NpN) divided by the number of sites (S/N). Faster-Z was calculated as $p_{NZ}/p_{SZ} : p_{NA}/p_{SA}$. Bootstrapping with 1000 repetitions was used to generate 95% confidence intervals and significance differences between genomic categories were determined from 1000 permutation tests.

For each species, I used the McDonald Kreitman test (McDonald and Kreitman 1991) to estimate the number of contigs evolving under adaptive and neutral evolution. The McDonald Kreitman test compares the number of nonsynonymous and synonymous substitutions (D_N and D_S) with polymorphisms (P_N and P_S). An excess of nonsynonymous substitutions relative to polymorphism is indicative of positive selection ($(D_N/D_S) > (P_N/P_S)$), and underrepresentation of nonsynonymous substitutions relative to polymorphism is indicative of relaxed purifying selection ($(D_N/D_S) < (P_N/P_S)$). I tested for departures from neutrality using a 2x2 contingency table and Pearson's Chi-squared test (Hope 1968; Patefield 1981) in R version 3.1.0 (R Core Team 2014). I used the qvalue function in R with a false discovery rate = 0.05 and lambda = 0 to correct for multiple testing. Contigs were only included in the analysis if the sum of each marginal row and column of the 2x2 contingency table was greater or equal than 6 (Begun et al. 2007; Andolfatto 2008).

The McDonald Kreitman test is restricted to genes with sufficient numbers of polymorphisms or substitutions (Begun et al. 2007; Andolfatto 2008). Therefore, to maximise the power of my dataset, for each species I separately summed P_N and P_S and used Pearson's Chi-squared test to test for deviations from neutrality (Mank et al. 2007a). Specifically, I tested whether the Z chromosome, relative to the autosomes, had 1) an excess of nonsynonymous substitutions predicted under Faster-Z Evolution;

2) an excess/underrepresentation of nonsynonymous polymorphisms indicative of elevated drift/selection respectively.

4.3.8 Calculating relative effective population size of the Z chromosome

I calculated the effective population size (N_E) of the Z chromosome and autosomes 1-10 for each species. Nucleotide diversity per base pair (π) was calculated for each genomic category by summing P_N, P_S, N and S across genes. $\pi = (P_N+P_S)/(N+S)$. The mutation rate per site per year (U) was taken from previous Galliform estimates and corrected for male-biased mutation rate and generation time (Dimcheff et al. 2002; Axelsson et al. 2004; van Tuinen and Dyke 2004; Mank et al. 2010a). $N_{EA} = \pi/(4*(U*\text{generation time}))$ and $N_{EZ} = \pi/(3*(U*\text{generation time}))$ (Hartl and Clark 2007). Bootstrapping with 1000 repetitions was used to generate 95% confidence intervals.

4.3.9 Tests of positive selection using gene expression

The relative role of selection versus drift in driving Faster-Z Evolution can be disentangled using gene expression (Baines et al. 2008; Mank et al. 2010c; Sackton et al. 2014). If Faster-Z Evolution is driven predominantly by increased efficiency of selection acting on recessive mutations in the heterozygous sex, I predict the effect to be largest for female-biased, followed by unbiased and then male-biased genes. Conversely, if Faster-Z is driven by elevated drift arising from a reduction in Z-linked effective population size relative to the autosomes, I predict no relationship between the magnitude of Faster-Z and expression category.

Gene expression was quantified using adult gonad samples, because this tissue exhibits the greatest magnitude of sex-biased transcription (Mank et al. 2007b; Pointer

et al. 2013) and therefore maximises the number of female-biased genes used in the analysis. Expression was estimated as reads per kilobase per million mappable reads (RPKM). Log_2 RPKM was normalized across species to control for differences in sequencing depth using cross-species normalisation (Brawand et al. 2011).

Mean male and female RPKM of each orthogroup were calculated separately for each species, together with fold change (a measure of sex-bias: log_2(male RPKM)-log_2(female RPKM)). A t-test was used to identify significantly sex-biased genes, and the Benjamini Hochberg method (FDR of 5%) (Benjamini and Hochberg 1995) used to correct for multiple testing. Female-biased and male-biased contigs were classified as significantly sex-biased ($p<0.05$) or sex-limited with a log_2 fold change of <-1 and >1 respectively. Unbiased genes had a log_2 fold change between <1 and >-1.

In order to test the predictions of the selection and drift hypotheses outlined above I used three approaches detailed below.

First, I calculated Faster-Z for orthogroups where expression category was conserved across all six species. This was to avoid diluting significant signals of selection or drift by including orthogroups where exposure to the dominant evolutionary force has not been consistent over time due to rapid expression turnover. Mean d_N, d_S and d_N/d_S were calculated separately for each expression category for Z-linked and autosomal genes using divergence estimates from the Galloanserae model in CODEML (Yang 2007). Bootstrapping with 1000 repetitions was used to generate 95% confidence intervals. Significant differences between genomic categories were determined using permutation tests with 1000 repetitions.

I then repeated this analysis with relaxed criteria in order to maximise the number of orthogroups in each expression category. Specifically, I compared the

Faster-Z Effect between putatively female-biased genes (defined as $\log_2$ fold change <0 across all species) and male-biased genes (defined as $\log_2$ fold change >0 across all species).

Finally, I assessed the relationship between species-specific Faster-Z Evolution and gene expression. For each species, I calculated d_{NZ}/d_{SZ} : d_{NA}/d_{SA} for female-, male- and unbiased contigs for each species as defined with t-tests and fold change thresholds. Significance was assessed using permutation tests with 1000 repetitions.

4.4 RESULTS

4.4.1 Faster-Z Evolution

De novo transcriptomes were assembled for six Galloanserae species, separated by approximately 90 million years of avian evolution (Fig. 4.2). After quality checks and filtering, 160 Z-linked and 2431 autosomal orthogroups were identified using BLASTn (Altschul et al. 1990). Divergence estimates for each orthogroup were calculated using the Galloanserae model in the CODEML package in PAML (Yang 2007), which averages d_N/d_S across all Galloanserae branches and sites.

Across the Galloanserae, mean d_N/d_S of the Z chromosome is significantly higher than that of the autosomes, due to significantly elevated d_{NZ} (Table 4.1, Fig. 4.3). There is no difference in d_S between the Z chromosomes and all autosomes ($p = 0.859$). 741 autosomal orthogroups are located on microchromosomes in the chicken genome. Microchromosomes exhibit an elevated recombination rate, greater gene density and GC content and these factors have been shown to impact the nature

Table 4.1. d_N, d_S and d_N/d_S for Z-linked and autosomal genes across Galloanserae clade.

	Z chromosome (160 contigs)	**Autosomes 1-10 (1690 contigs)**	**Microchromosomes (741 contigs)**	**All Autosomes (2431 contigs)**
d_N/d_S	0.130	0.110	0.078	0.100
95% CI	(0.112-0.150)	(0.104-0.116)	(0.072-0.085)	(0.095-0.105)
		p=0.019	**p < 0.001**	**p=0.002**
d_S	0.432	0.424	0.510	0.447
95% CI	(0.413-0.452)	(0.417-0.431)	(0.495-0.527)	(0.440-0.454)
		p=0.230	p=1.000	p=0.859
d_N	0.056	0.047	0.040	0.045
95% CI	(0.048-0.065)	(0.044-0.049)	(0.037-0.043)	(0.043-0.047)
		p=0.013	**p < 0.001**	**p=0.005**

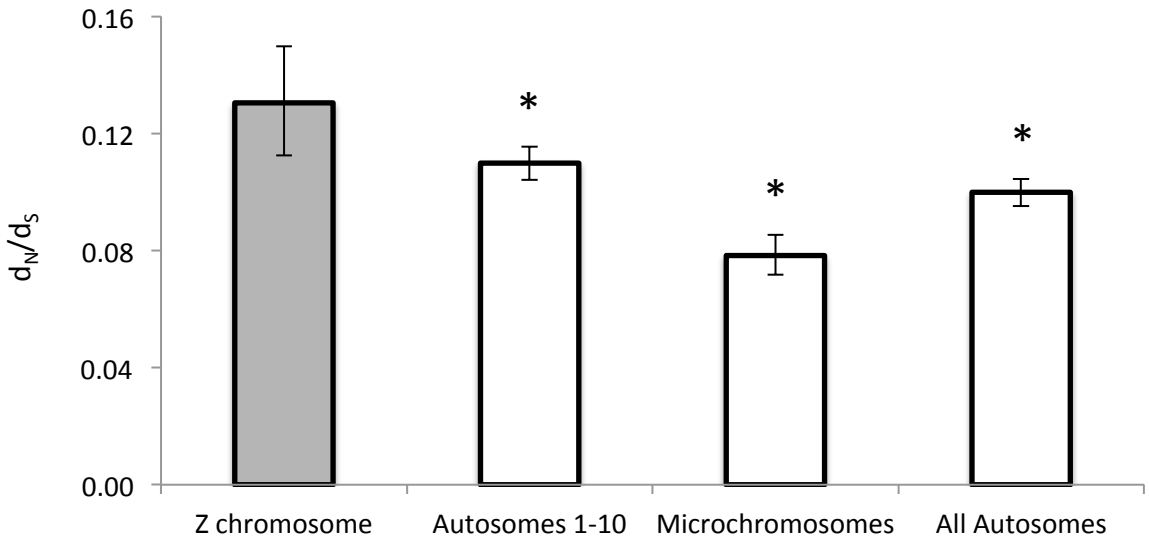

Fig. 4.3. Estimates of mean d_N/d_S for loci on autosomes and the Z chromosome across the Galloanserae.

Synonymous and nonsynonymous divergence estimates were calculated using the Galloanserae branch model in PAML. 95% confidence intervals were calculated by bootstrapping with 1000 replicates, and significant differences in d_N/d_S (permutation test, 1000 replicates) are indicated (*).

and efficacy of selection (Burt 2002; Ellegren 2013). Therefore, the fairest measure of Faster-Z Evolution is to contrast divergence between the Z chromosome and similarly sized autosomes 1-10 (Mank et al. 2010c). Mean d_{NZ}/d_{SZ} and d_{NZ} are both significantly higher than mean d_N/d_S and d_N of autosomal 1-10 orthogroups (Table 4.1, Fig. 4.3). This pattern is consistent with the results of the previous analysis using all autosomes, and with previous estimates of Faster-Z Evolution in birds (Mank et al. 2007a; Dalloul et al. 2010; Mank et al. 2010c; Ellegren et al. 2012; Wang et al. 2014). For the rest of the manuscript, autosomal will refer to autosomal 1-10 orthogroups and Faster-Z will refer to the comparison between Z-linked and autosomal 1-10 orthogroups d_{NZ}/d_{SZ} : d_{NA}/d_{SA}.

In each of the six Galloanserae species d_{NZ}/d_{SZ} is higher than d_{NA}/d_{SA} based on the species-specific model, and there is inter-specific variation in the magnitude of this difference (Table 4.2). I found no significant difference in d_S between the Z chromosome and autosomes for any species, consistent with previous findings that male-biased mutation rate is weak across the Galloanserae (Bartosch-Harlid et al. 2003; Axelsson et al. 2004), and suggesting that Z-linked mutation rate does not vary significantly across the six species.

4.4.2 Variation in male mating success drives Faster-Z Evolution

I conducted phylogenetically controlled regression analyses between d_{NZ}/d_{SZ} : d_{NA}/d_{SA} and log sperm number as well as residual testes weight (Moller 1988, 1991; Birkhead and Petrie 1995). Numerous studies have demonstrated that the intensity of sperm competition, a widely used indicator of postcopulatory sexual selection and therefore one measure of variance in male mating success, is strongly predicted by

Table 4.2. d_N, d_S and d_N/d_S for Z-linked and autosomal genes across Galloanserae species.

Species	Z chromosome			Autosomes 1:10			Faster Z
	d_N (95% CI)	d_S (95% CI)	d_N/d_S (95% CI)	d_N (95% CI)	d_S (95% CI)	d_N/d_S (95% CI)	d_{NZ}/d_{SZ}: d_{NA}/d_{SA} (95% CI)
Meleagris gallopavo	0.023 (0.020-0.027)	0.163 (0.155-0.171)	0.144 (0.123-0.167)	**0.019 (0.018-0.020) p=0.005**	0.158 (0.155-0.161) p=0.172	**0.120 (0.113-0.126) p=0.017**	1.205 (1.031-1.410)
Phasianus colchicus	0.021 (0.018-0.025)	0.161 (0.153-0.169)	0.134 (0.114-0.153)	**0.018 (0.017-0.200) p=0.047**	0.157 (0.153-0.160) p=0.183	0.118 (0.111-0.124) p=0.086	1.137 (0.973-1.314)
Numida meleagris	0.019 (0.016-0.022)	0.133 (0.127-0.140)	0.140 (0.118-0.161)	0.016 (0.015-0.017) p=0.057	0.132 (0.129-0.135) p=0.444	0.123 (0.116-0.130) p=0.057	1.140 (0.964-1.342)
Anas platyrhnchos	0.015 (0.013-0.018)	0.116 (0.106-0.125)	0.131 (0.108-0.155)	**0.013 (0.012-0.014) p=0.040**	0.116 (0.103-0.116) p=0.533	**0.109 (0.103-0.116) p=0.029**	1.200 (0.981-1.445)
Anser cygnoides	0.012 (0.011-0.015)	0.100 (0.093-0.108)	0.125 (0.104-0.147)	0.011 (0.011-0.012) p=0.06	0.099 (0.097-0.010) p=0.361	0.111 (0.104-0.117) p=0.075	1.129 (0.934-1.346)
Pavo cristatus	0.020 (0.017-0.023)	0.147 (0.140-0.154)	0.134 (0.114-0.156)	0.017 (0.016-0.018) p=0.082	0.147 (0.144-0.150) p=0.480	0.118 (0.111-0.125) p=0.073	1.133 (0.960-1.323)

Significance values were determined from 1000 permutation tests and bootstrapping with 1000 repetitions was used to generate 95% confidence intervals. Significant values are shown in bold.

relative testes weight and sperm number in birds (Moller 1991; Birkhead and Moller 1998; Pitcher et al. 2005).

I recovered a significant positive correlation between magnitude of Faster-Z Evolution and both log sperm number (r^2 = 0.675, p = 0.012, t = 3.550) and residual testes weight (r^2 = 0.581, p = 0.022, t = 2.911) (Fig. 4.4). To test the strength of these correlations, I sequentially removed each species, and repeated the analyses (Table S.4.1). Despite the reduction in sample size and therefore statistical power, there was no change to the significance or direction of the slope for log sperm number. For residual testes weight, there was no change to the direction of the slope but when either *A. cynoides* or *A. platyrhynchos* was excluded the relationship was marginally non-significant (Table S.4.1).

There are two plausible explanations for my finding that the magnitude of Z-linked divergence increases with increasing female promiscuity. A recent study in silkmoths has shown that Faster-Z Evolution is adaptive, and results from increased efficacy of selection acting on recessive mutations in the hemizygous sex (Sackton et al. 2014). Conversely, a study in birds suggested that avian Faster-Z Evolution is a neutral process, driven by elevated genetic drift as a consequence of relative differences in N_{EZ}/N_{EA} (Mank et al. 2010c). Under the latter hypothesis, variation in male reproductive success, associated with sexual selection, is predicted to alter the relationship between N_{EZ} and N_{EA}, and therefore the relative magnitude of drift acting on the Z chromosome (Charlesworth et al. 1993; Vicoso and Charlesworth 2009). Specifically, with increasing variance in male reproductive success, relative N_{EZ} decreases, resulting in greater magnitude of drift and therefore Faster-Z Effect (Wright and Mank 2013).

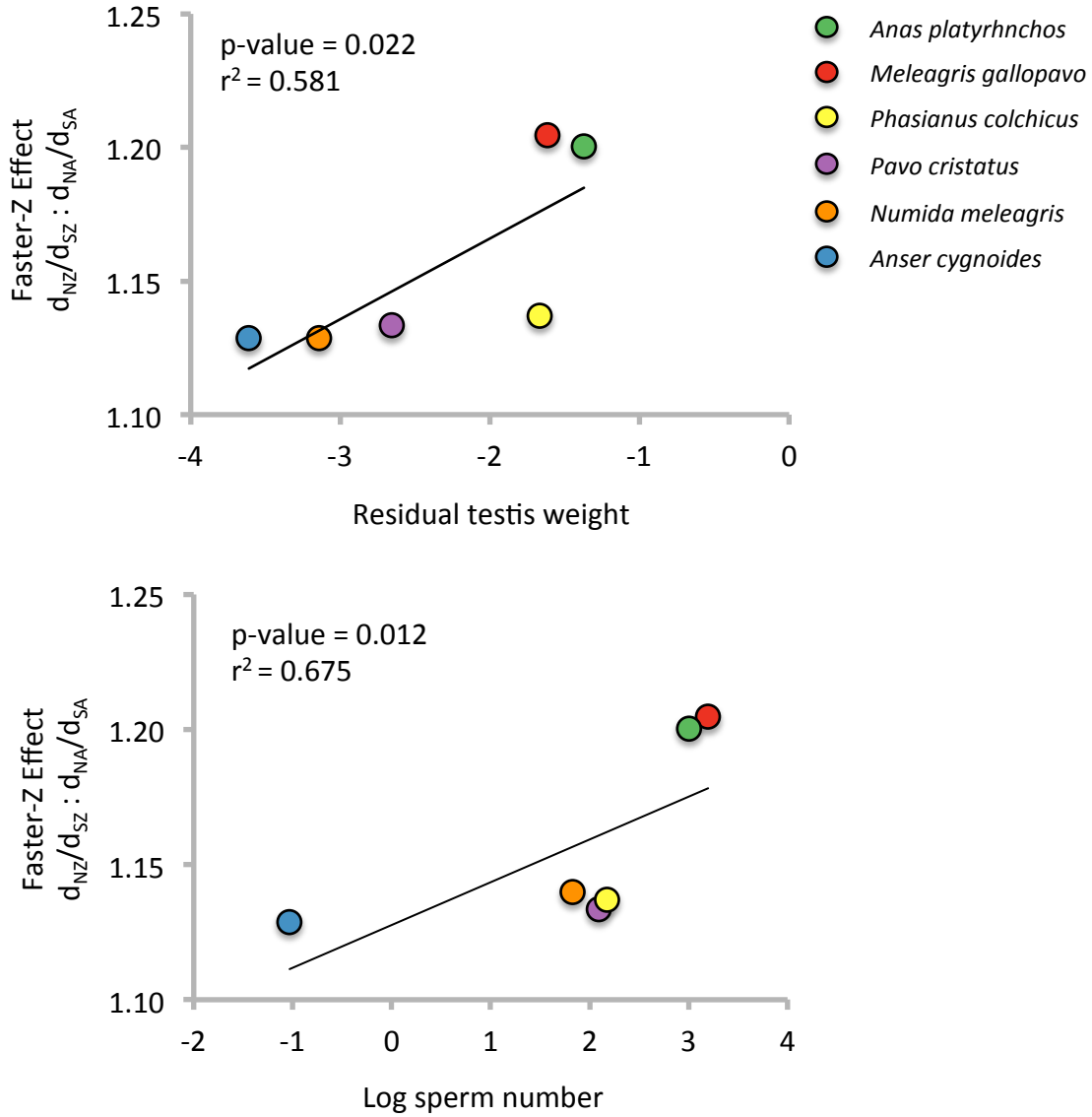

Fig. 4.4. Phylogenetically controlled regression between proxies of sperm competition and Faster-Z Effect.

Data points are raw species values but p-values and r^2 estimates were calculated using phylogenetic generalized least squares models with maximum likelihood and 1000 runs for each analysis.

I use sequence divergence, polymorphism and expression data to test whether the relationship between female promiscuity and Faster-Z evolution is adaptive or neutral.

4.4.3 Estimates of relative N_{EZ}

After filtering for quality and read depth, across Z-linked and autosomal 1-10 contigs 12061 SNPs in *A. platyrhynchos,* 4508 in *M. gallopavo,* 6759 in *P. colchicus,* 5109 in *N. meleagris,* 1917 in *P. cristatus* and 7943 in *A. cygnoides* were identified (Table S.4.2).

For each species, I calculated the effective population size of the Z chromosome (N_{EZ}) and autosomes 1-10 (N_{EA}). I corrected for male-biased mutation rate and generation time using previous Galliform estimates (Dimcheff et al. 2002; Axelsson et al. 2004; van Tuinen and Dyke 2004; Mank et al. 2010a). For each species, N_{EZ}/N_{EA} falls into the neutral range of Faster-Z Evolution (Table 4.3) (Vicoso and Charlesworth 2009). Additionally, N_{EZ} is significantly less than the ¾ N_{EA} predicted under strict monogamy for all species, with the exception of *P. cristatus,* although the 95% CI for this species was unusually wide.

The relationship between N_{EZ}/N_{EA} and sperm number or residual testes weight was not statistically significant (sperm number: $r^2 = 0.070$, $p = 0.269$, $t = 0.673$; residual testes weight: $r^2 = 0.061$, $p = 0.285$, $t = 0.622$). Additionally, the autosomal effective population size of *P. cristatus* is significantly smaller than the other six species, indicating either a very recent bottleneck or variation in family structure across the individuals sampled in this study. This finding hints at the sensitivity of N_E calculations to many factors (Hartl and Clark 2007), including past demographic perturbations, and

Table 4.3. Effective population size estimates of the Z chromosome and autosomes.

Species	N_{EZ} (E+05) (95% CI)	N_{EA1-10} (E+05) (95% CI)	N_{EZ}/N_{EA1-10} (95% CI)
Meleagris gallopavo	0.787 (0.512-1.140)	1.950 (1.832-2.079)	0.404 (0.254-0.593)
Phasianus colchicus	1.427 (1.056-1.888)	2.904 (2.769-3.057)	0.491 (0.358-0.641)
Numida meleagris	0.757 (0.336-1.362)	2.220 (2.048-2.414)	0.341 (0.153-0.614)
Anas platyrhynchos	2.560 (1.701-3.575)	5.181 (4.935-5.453)	0.494 (0.327-0.692)
Anser cygnoides	1.826 (1.203-2.557)	3.401 (3.208-3.590)	0.537 (0.358-0.744)
Pavo cristatus	0.552 (0.136-1.201)	0.812 (0.726-0.911)	0.679 (0.191-1.493)

N_E was calculated using the same method as Mank et al 2010a. Mutation rate estimates are from Axelsson et al 2004, Dimcheff et al 2002 and van Tuinen and Dyke 2004.

this may explain both the unusually low N_E estimates in *P. cristatus* as well as the lack of significant association between N_{EZ}/N_{EA} and measures of sperm competition (addressed further in the Discussion).

4.4.4 Tests of positive selection

I used sequence and polymorphism data from my six species to test whether selection is more effective for Z-linked versus autosomal loci.

First, I used site model test results in the CODEML package in PAML (Yang 2007) to test for signatures of positive selection acting on a subset of sites. Of the 160 Z-linked orthogroups, I found significant evidence for positive selection acting on 5 loci, of which 1 remained significant after sequential Bonferroni correction (Table 4.4, Table S.4.3). Of the 1690 orthogroups located on autosomes 1-10, I found significant evidence for positive selection acting on 51 loci, 5 of which remained significant after sequential Bonferroni correction (Table 4.4, Table S.4.3). I also detected 13 microchromosomal loci that were evolving under positive selection, of which 3 remained significant after sequential Bonferroni correction (Table S.4.3).

There was no significant difference in the proportion of positively selected loci on the Z chromosome or autosomes 1-10 either before or after multiple testing correction (χ^2, d.f. = 1, $p > 0.400$ in both comparisons). This finding indicates that selection is not more effective on the Z chromosome.

I next used polymorphism data to test for deviations from neutrality. With the exception of *N. meleagris* and *P. cristatus*, p_{NZ}/p_{SZ} is significantly greater than p_{NA}/p_{SA} (Table 4.5). This finding of excess nonsynonymous polymorphism on the Z chromosome relative to the autosomes suggests that selection is less effective at

Table 4.4. Site model test results for contigs under positive selection.

Gene name [a]	Chr	*ω*	Proportion of sites	M1a LR	M2a LR	LRT	p-value	p-fdr value[b]
22552	1	2.897	0.122	-6535.857	-6522.227	27.26	< 0.001	0.003
21101	1	4.155	0.033	-14063.297	-14050.29	26.02	< 0.001	0.006
31776	3	4.608	0.130	-1270.098	-1256.430	27.34	< 0.001	0.003
39919	6	4.226	0.310	-1630.735	-1611.278	38.92	< 0.001	< 0.001
03831	8	4.817	0.080	-9607.226	-9560.287	93.88	< 0.001	< 0.001
10504	15	3.343	0.072	-5389.616	-5375.473	28.29	< 0.001	0.002
01868	20	9.422	0.013	-4192.958	-4179.195	27.53	< 0.001	0.003
02022	28	4.914	0.068	-2768.690	-2753.634	30.11	< 0.001	0.001
29351	Z	6.802	0.104	-2974.892	-2930.430	88.93	< 0.001	< 0.001

[a] *G.gallus* ortholog Ensembl ID ENSGALT000000…..
[b] Sequential Bonferroni correction (Holm 1979)

Table 4.5. p_N, p_S and p_N/p_S for Z-linked and autosomal genes across Galloanserae species.

Species	Z chromosome			Autosomes 1:10			Faster Z
	p_N (95% CI)	p_S (95% CI)	p_N/p_S (95% CI)	p_N (95% CI)	p_S (95% CI)	p_N/p_S (95% CI)	p_{NZ}/p_{SZ}: p_{NA}/p_{SA} (95% CI)
Meleagris gallopavo	0.000 (0.000-0.000)	0.001 (0.001-0.002)	0.229 (0.147-0.336)	0.001 (0.001-0.001) p = 1.000	0.004 (0.004-0.005) p = 1.000	**0.130 (0.119-0.143) p<0.001**	1.757 (1.095-2.665)
Phasianus colchicus	0.000 (0.000-0.001)	0.002 (0.002-0.003)	0.212 (0.138-0.309)	0.001 (0.001-0.001) p = 1.000	0.006 (0.006-0.007) p = 1.000	**0.121 (0.112-0.130) p<0.001**	1.748 (1.162-2.603)
Numida meleagris	0.000 (0.000-0.000)	0.001 (0.001-0.003)	0.108 (0.066-0.192)	0.001 (0.001-0.001) p = 1.000	0.005 (0.004-0.001) p = 1.000	0.129 (0.117-0.141) p = 0.876	0.836 (0.516-1.492)
Anas platyrhnchos	0.001 (0.000-0.001)	0.005 (0.003-0.007)	0.121 (0.074-0.177)	0.001 (0.001-0.001) p = 1.000	0.0124 (0.012-0.013) p = 1.000	**0.089 (0.081-0.096) p=0.009**	1.367 (0.856-2.013)
Anser cygnoides	0.001 (0.000-0.001)	0.003 (0.002-0.004)	0.207 (0.129-0.301)	0.001 (0.000-0.001) p = 1.000	0.007 (0.007-0.008) p = 1.000	**0.135 (0.125-0.147) p<0.001**	1.527 (0.942-2.285)
Pavo cristatus	0.000 (0.000-0.000)	0.001 (0.000-0.002)	0.221 (0.113-0.691)	0.000 (0.000-0.000) p = 1.000	0.002 (0.001-0.002) p = 0.947	0.168 (0.144-0.194**)** p = 0.119	1.317 (0.657-4.241)

Significance values were determined from 1000 permutation tests and bootstrapping with 1000 repetitions was used to generate 95% confidence intervals. Significant values are shown in bold.

removing mildly deleterious mutations from the Z chromosome. Interestingly, *N. meleagris* Z chromosome exhibits a non-significant deficit of p_N, potentially as a consequence of monogamy.

For each species, I estimated the number of contigs evolving under adaptive and neutral evolution using the McDonald Kreitman test (McDonald and Kreitman 1991), which contrasts the number of nonsynonymous and synonymous substitutions (D_N and D_S) with polymorphisms (P_N and P_S). An excess of nonsynonymous substitutions relative to polymorphism is indicative of positive selection ((D_N/D_S) > (P_N/P_S)). I detected no Z-linked contigs with signatures of positive selection, and there was no difference between the Z and autosomes in the proportion of loci under positive selection in any species (χ^2, d.f. = 1, $p > 0.600$ in all cases) (Table S.4.4). However, only contigs with sufficient numbers of substitutions and polymorphisms were included in the analysis (Begun et al. 2007; Andolfatto 2008) and therefore my ability to draw species-specific conclusions is limited by low sample sizes. I therefore summed across all species. Again, across all six species, there was no statistical difference in the proportion of Z-linked and autosomal loci showing signatures of positive selection (χ^2, d.f. = 1, $p > 0.500$).

In order to further maximise the power of my dataset, for each species I summed the total number of P_N and P_S across all Z-linked and all autosomal 1-10 contigs separately and tested for deviations from neutrality (Table 4.6). For each species, there is a significant excess of Z-linked nonsynonymous polymorphism relative to the autosomes for all species with the exceptions of *P. cristatus* and *N. meleagris*. The excess of non-synonymous polymorphism on the Z chromosome is again

Table 4.6. Significant differences between Z-linked and autosomal P_N and P_S.

Species	**Z chromosome**		**Autosomes 1:10**		$P_{NZ}/P_{SZ} : P_{NA}/P_{SA}$
	P_N	P_S	P_N	P_S	p-value
Meleagris gallopavo	51	83	1148	3226	1.727 **p=0.003**
Phasianus colchicus	88	155	1618	4898	1.719 **p<0.001**
Numida meleagris	29	100	1299	3681	0.822 p=0.412
Anas platyrhnchos	107	329	2265	9360	1.344 **p=0.011**
Anser cygnoides	111	200	2060	5572	1.501 **p=0.001**
Pavo cristatus	35	59	573	1250	1.294 p=0.287

Significance values were determined from 1000 permutation tests and bootstrapping with 1000 repetitions was used to generate 95% confidence intervals. Significant values are shown in bold.

consistent with a reduction in the power of selection to remove mildly deleterious alleles from this chromosome.

The lack of difference in Z-linked and autosomal nonsynonymous polymorphism in *P. cristatus* and *N. meleagris* could be attributed to a number of factors. It could reflect biological differences in sexual selection and therefore the magnitude of drift acting on the Z chromosome. However, although this explanation is consistent with the monogamous mating system of *N. meleagris*, it is not consistent with the polygnous *P. cristatus*. More likely, this pattern reflects the limitations of polymorphism data and the difficulty in controlling for family structure and demographic effects. For example, the number of SNPs in *P. cristatus* is much lower than the other five species, and therefore the power of this analysis is limited (Table 4.6).

4.4.5 Gene expression

I used gene expression data from gonads of our six avian species to identify the dominant force driving Faster-Z Evolution across the Galloanserae clade.

If Faster-Z Evolution is adaptive and driven by increased efficacy of selection acting on recessive mutations in the hemizygous sex, I predict the Faster-Z Effect to be largest for female-biased, followed by unbiased and then male-biased genes. If it is due to neutral causes, there will be no difference in the rate of Faster-Z Evolution among expression classes (Baines et al. 2008; Mank et al. 2010c; Sackton et al. 2014). I tested this prediction at three levels in my data.

First, I identified orthogroups with consistent male-, female- and unbiased expression across all six species, thereby excluding any orthogroups where the nature

of sex-bias, and therefore exposure to the dominant evolutionary force, has varied over Galloanserae evolutionary history. The rapid change in sex bias across this clade (Harrison et al. in prep) means that relatively few genes are consistently sex-biased in my dataset, resulting in 17 male-biased, 9 female-biased and 7 unbiased Z-linked orthogroups alongside 103 male-biased, 111 female-biased and 205 unbiased autosomal orthogroups. Among these gene sets, there was no significant difference in Faster-Z Effect (male-biased vs female-biased p = 1.00, female-biased vs unbiased p = 1.00, male-biased vs unbiased p= 1.00, all two-tailed pairwise permutation tests with 1000 repetitions), shown in Fig. 4.5.

To exclude the possibility that I lack statistical power to distinguish between drift and positive selection due to low sample sizes, I next repeated the analysis and relaxed the definition of sex-bias. In this approach, orthogroups were classified into different expression categories based on fold change alone (see Materials and Methods). In doing so, I doubled the number of orthogroups in each expression category; identifying 54 male-biased and 18 female-biased orthogroups, together with 349 male-biased and 335 female-biased autosomal orthogroups. Again, there was no significant difference in Faster-Z Effect between these gene sets (p = 1.000, permutation test, 1000 repetitions), with female-biased $d_{NZ}/d_{SZ} : d_{NA}/d_{SA}$ = 1.362 (95% CI = 0.990 -1.888) and male-biased $d_{NZ}/d_{SZ} : d_{NA}/d_{SA}$ = 1.461 (95% CI = 1.104 -1.859).

Finally, I assessed whether there was any species-specific pattern in Faster-Z Evolution across expression categories. I identified male-, female- and unbiased genes in each species separately, and then tested for differences in $d_{NZ}/d_{SZ} : d_{NA}/d_{SA}$ across expression categories (Table S.4.5). There is no significant difference between Faster-Z of any expression category in any species, with the exception of *N. meleagris* where I

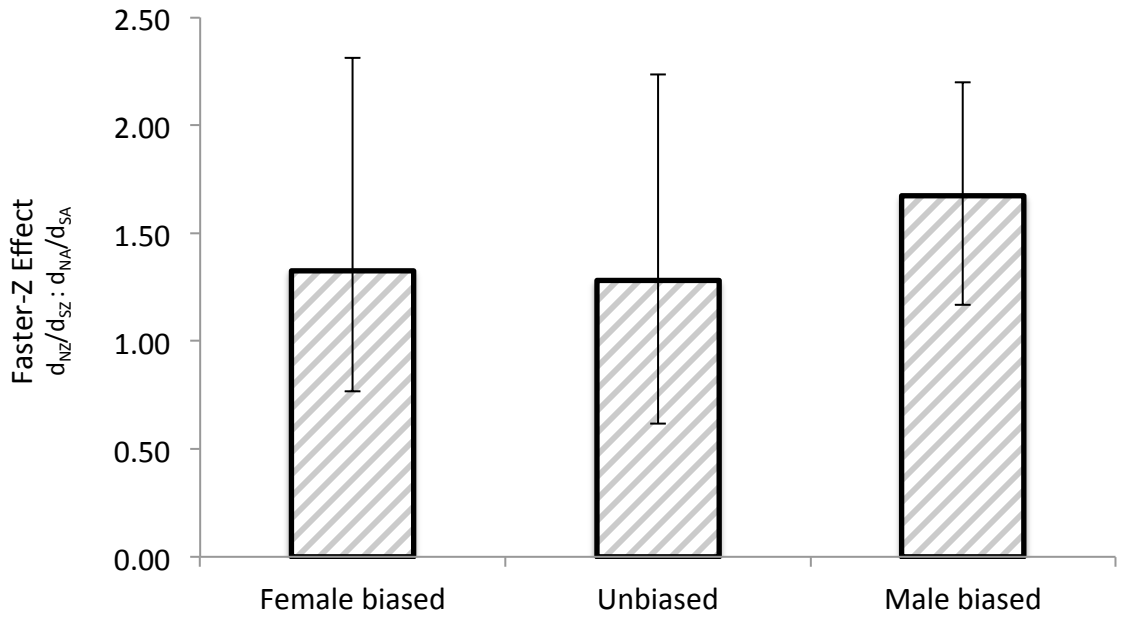

Fig. 4.5. Estimates of mean Faster-Z across sex-biased gene expression categories.

Sex-bias was defined using fold change thresholds and t-tests. 95% confidence intervals were calculated by bootstrapping with 1000 replicates.

found a significant difference between male-biased and unbiased Faster-Z Effect (Table S.4.6).

At all three levels of analysis, my gene expression data are consistent with Faster-Z Evolution resulting predominantly from neutral forces.

4.5 DISCUSSION

Faster rates of coding sequence evolution on the Z chromosome relative to the autosomes have been observed across a wide range of species (Mank et al. 2007a; Dalloul et al. 2010; Mank et al. 2010c; Ellegren et al. 2012; Sackton et al. 2014; Wang et al. 2014), however the underlying cause is unclear. Indirect evidence from an expression based approach suggests that avian Faster-Z Evolution is driven by genetic drift (Mank et al. 2010c), but a recent study in silkmoths postulated an adaptive explanation (Sackton et al. 2014). In order to determine the cause of Faster-Z Evolution, I used *de novo* transcriptomes for six Galloanserae species, spanning 90 million years of avian evolution and combined expression, sequence and polymorphism data with measures of sperm competition and promiscuity. I present the first empirical evidence demonstrating the positive relationship between the Faster-Z Effect and measures of sperm competition and variance in male mating success.

This pattern is consistent with a considerable body of theory predicting that Faster-Z evolution is driven by changes in the relative strength of genetic drift as a result of increased variance in male mating success (Vicoso and Charlesworth 2009). In support of the predominant role of genetic drift in shaping rates of Z chromosome evolution, I used multiple sequence, polymorphism and expression based approaches.

My expression analysis is consistent with previous work which found no difference in Faster-Z Evolution among sex-biased expression categories (Mank et al. 2010c). However, my analysis significantly extends this previous work by incorporating tests of positive selection based on divergence and polymorphism. The results from these multiple lines of evidence are broadly convergent, indicating that selection is not more effective on the Z chromosome. I conclude that Faster-Z Evolution in birds is due primarily to genetic drift and that the magnitude of this effect is dependent on the nature of sexual selection.

4.5.1 Variation in male mating success drives Faster-Z Evolution

Changes in the skew of male reproductive success are commonly associated with promiscuity and the intensity of sexual selection (Andersson 1994), both of which would decrease the N_{EZ}/N_{EA} ratio. If Faster-Z is neutral and non-adaptive, we predict that the magnitude of Faster-Z Evolution should increase as N_{EZ}/N_{EA} decreases (Vicoso and Charlesworth 2009), and therefore we should expect both lower N_{EZ}/N_{EA} and increased rate of Faster-Z Evolution in promiscuous compared to monogamous populations (Fig. 4.1) .

After correcting for phylogeny, I uncovered a significant and positive correlation between the magnitude of Faster-Z and relative testes weight and sperm number, both reliable predictors of the intensity of sperm competition in birds (Fig. 4.4) (Moller 1991; Birkhead and Moller 1998; Pitcher et al. 2005). Sperm competition is a widely used indicator of the strength of postcopulatory sexual selection and therefore a good proxy for variance in male mating success and the magnitude of drift acting on the Z chromosome (Moller 1991; Birkhead and Moller 1998).

However, although the relationship between N_{EZ}/N_{EA} and sperm number or residual testes size was not significant, N_{EZ}/N_{EA} across the Galloanserae does fall into the neutral range for Faster-Z Evolution (Vicoso and Charlesworth 2009) and is significantly less than the 0.75 predicted under strict monogamy, with the exception of *P. cristatus* (Table 4.3). Estimates of N_E are highly sensitive to many factors, particularly short fluctuations in demography (Hartl and Clark 2007). Additionally, I calculated effective population size using parameters estimated from previous Galliform studies (Dimcheff et al. 2002; Axelsson et al. 2004; van Tuinen and Dyke 2004; Mank et al. 2010a), which may reflect the biology of certain species more accurately than others. Although mutation rate, male-biased mutation and generation time are not expected to vary substantially across the Galloanserae, slight differences might be expected.

More importantly, polymorphism estimates are sensitive to past demographic perturbations and recent bottlenecks (Hartl and Clark 2007), and this may explain both the unusually low N_E estimates in *P. cristatus*, as well as the lack of a significant relationship between N_{EZ}/N_{EA} and measures of promiscuity and sperm competition. Previous attempts to estimate N_{EZ}/N_{EA} in birds (Corl and Ellegren 2012) showed sizable variation from what would be predicted by mating system, suggesting that N_{EZ}/N_{EA} estimates may simply be too inaccurate for the types of analyses used here. Because divergence data are not as sensitive to recent demographic perturbations, it can be argued that it is a fairer test for the role of male mating success and sperm competition in Faster-Z Evolution.

4.5.2 Tests of positive selection

I used sequence and polymorphism data to test the relative strength of selection on the Z chromosome versus autosomes. In both the site-model tests in PAML as well as species-specific McDonald Kreitman tests, there was no difference in the proportion of positively selected loci on the Z chromosome compared to the autosomes (Table 4.4, Table S.4.3, Table S.4.4). The McDonald Kreitman test is limited to sequences with sufficient numbers of substitutions and polymorphisms (McDonald and Kreitman 1991; Andolfatto 2008), and this restricted my analysis to a handful of Z-linked contigs. Therefore, to maximise the power of my dataset, I summed polymorphism data across all Z-linked and autosomal contigs (Mank et al. 2007a). For the majority of species, an excess of Z-linked nonsynonymous polymorphism was observed (Table 4.5 and Table 4.6), suggesting that selection is less able to purge mildly deleterious alleles from the Z chromosome. This pattern is consistent with the theoretical expectations of elevated levels of genetic drift. I would expect the opposite pattern, a deficit of Z-linked nonsynonymous polymorphism, under both positive and purifying selection.

Overall, I failed to detect any indication that selection is more effective for Z-linked loci, consistent with the non-adaptive explanations for Faster-Z Evolution. Given Faster-Z Evolution is neutral, theory predicts that the magnitude of Faster-Z Effect should increase as N_{EZ}/N_{EA} decreases (Vicoso and Charlesworth 2009), and therefore I would expect increased rates of Faster-Z Evolution in promiscuous compared to monogamous populations. This prediction is consistent with my finding that Faster-Z is positively correlated with variance in male reproductive success.

4.5.3 Faster-Z versus Faster-X Evolution

Faster rates of coding sequence divergence have repeatedly been documented on the X and Z chromosomes relative to the autosomes, and there is considerable variation in the magnitude of this difference across species (Meisel and Connallon 2013). Moreover, there is a stark contrast between my results and those of Faster-X Evolution in *Drosophila* and mammals, where X-linked male-biased genes evolve more rapidly than unbiased and female-biased genes (Khaitovich et al. 2005; Baines et al. 2008; Grath and Parsch 2012). This pattern is consistent with an adaptive explanation of Faster-X Evolution driven by increased efficacy of selection acting on recessive mutations. In addition, there is considerable evidence for signatures of adaptation on the X chromosome across many species (Thornton and Long 2005; Baines et al. 2008; Hvilsom et al. 2012; Langley et al. 2012).

The empirical evidence for neutral versus adaptive explanations of Faster-Z and Faster-X Evolution, respectively, are supported by theoretical predictions (Vicoso and Charlesworth 2009). As variance in male reproductive fitness increases, $N_{EZ} < ¾ N_{EA}$, reducing the ability of selection to purge mildly deleterious alleles. In contrast, $N_{EX} > ¾ N_{EA}$ under increased variance in male reproductive success, indicating that Faster-X is more often due to positive selection acting on recessive mutations exposed in the heterogametic sex. However, a recent study in silkmoths (Sackton et al. 2014) indicates that this prediction may not hold for all species and is dependent on numerous other factors, including overall population size and sex-specific recombination rates (Connallon et al. 2012).

4.5.4 Male-biased mutation

The relative rate of Z-linked divergence is thought to be influenced by multiple factors, not only variance in male reproductive success (Kirkpatrick and Hall 2004b Connallon et al. 2012). The number of cell divisions, and therefore potential for mutations, is inherently higher in spermatogenesis compared to oogenesis. This male-biased mutation has been documented across a number of species (Bartosch-Harlid et al. 2003; Axelsson et al. 2004; Xu et al. 2012), and as the Z chromosome is present more often in males than females, it could contribute to the observed differences in relative Z-linked divergence (Kirkpatrick and Hall 2004b; Xu et al. 2012). However, previous estimates indicate the magnitude of male-biased mutation may be relatively weak across the Galloanserae (Bartosch-Harlid et al. 2003; Axelsson et al. 2004). Moreover, my analysis showed that there was no significant difference between d_{SZ} and d_{SA} across species, indicating that male-mutation bias does not vary significantly across this clade.

4.5.5 Sexual selection and the Z chromosome

The sex chromosomes are predicted to play a disproportionate role in encoding sex-specific fitness due to their unequal inheritance pattern (Rice 1984). The Z chromosome in particular is thought to foster tight linkage between female preference genes and flashy male traits, and promote rapid evolution of some types of sexually selected traits (Rice 1984; Reeve and Pfennig 2003; Kirkpatrick and Hall 2004a). However, evidence that the Z chromosome harbours genes encoding sexually dimorphic phenotypes is mixed (Dean and Mank 2014). Z-linked male plumage genes have been documented in flycatchers (Saetre et al. 2003; Saether et al. 2007), but

other studies have failed to find an association between sexual dimorphisms and sex-linkage (Knief et al. 2012; Schielzeth et al. 2012; Pointer et al. 2013). My findings may help explain this discrepancy between theoretical and empirical data. The low effective population size of the Z chromosome relative to the autosomes may weaken the efficacy of sex-specific selection, particularly in the species under the strongest sexual selection regimes. This may limit the adaptive role of the Z chromosome in general, and in particular its role in encoding sexually selected traits. Given this, it is important to note that my results do not exclude the potential for selection acting on the Z chromosome, rather that genetic drift is the dominant force.

4.6 CONCLUSIONS

I assessed the magnitude and drivers of Faster-Z Evolution across a clade of birds spanning 90 million years of evolution. My analysis combines expression, sequence and polymorphism data with measures of sperm competition and promiscuity. The results from these multiple lines of evidence are broadly convergent, indicating that selection is less effective on the Z chromosome, and suggesting that Faster-Z Evolution in birds is due primarily to genetic drift. Moreover, I present the first empirical evidence demonstrating the positive relationship between the Faster-Z Effect and measures of promiscuity and therefore variance in male mating success. My results reveal the power of mating system in shaping broad patterns in genome evolution.

Chapter 5

Conclusions

Sex chromosomes are subject to unique evolutionary environments as a consequence of their unusual pattern of inheritance, and are predicted to follow very different evolutionary trajectories compared to the autosomes (Rice 1984; Charlesworth et al. 1987; Mank 2009b; Wright and Mank 2013). In this thesis, I combine genomic and transcriptomic data across a wide range of avian species to explore the evolutionary processes governing sex chromosome evolution. In doing so, I shed light on the dynamics of sex chromosome divergence, and highlight the role of mating system, and interplay between evolutionary forces, in driving coding sequence and expression evolution on the avian Z and W chromosomes.

5.1 Complex dynamics of recombination suppression across the avian sex chromosomes

Recombination suppression initiates the divergence and differentiation of sex chromosomes from an ancestral pair of autosomes (Charlesworth 1991). However, despite its critical role in sex chromosome evolution, our understanding of the evolutionary dynamics of the cessation of recombination is limited. Older sex chromosomes support the classic model of sex chromosome evolution where recombination is suppressed instantaneously in a stepwise manner consistent with inversions (Skaletsky et al. 2003; Ross et al. 2005; Bachtrog 2013; Vicoso et al. 2013a). However, data from newly evolved sex chromosomes reveals that recombination

suppression is far more heterogeneous (Chibalina and Filatov 2011; Muyle et al. 2012; Natri et al. 2013). By identifying the largest catalog of W-linked genes to date, across three species spanning 90 million years of avian evolution, I was able to assess the short- and long- term dynamics of sex chromosome divergence (Chapter 2 and 3).

I first show that the *Gallus gallus* Z chromosome was formed by up to four separate recombination suppression events, resulting in physical regions of similar divergence, called strata, over the course of 130 million years (Chapter 2, Wright et al. 2102). Subsequently, I demonstrated that these strata are conserved across Galliformes (Chapter 3, Wright et al. In press). This stepwise pattern of recombination cessation is consistent with the classic inversion model of sex chromosome evolution. Together, my results suggest that long-term sex chromosome divergence is characterized by repeated inversions spreading progressively to restrict recombination.

In contrast, over shorter evolutionary time scales I observe recent and independent evolution of recombination suppression across sex chromosomes of the sister clade, Anseriformes, together with heterogeneous and locus specific divergence. In addition, I uncover evidence of persistent genetic exchange between both highly diverged and recently evolved gametologs, indicating that recombination suppression is not instantaneous (Chapter 3, Wright et al. In press).

Taken together, my findings illustrate a complex mosaic of recombination suppression, with repeated and independent divergence of Z-W gametologs together with ongoing genetic exchange.

5.1.1 Future directions

My findings reveal the complex dynamics of sex chromosome divergence over long evolutionary periods. However, the factors promoting recombination suppression in the initial stages of sex chromosome differentiation are still unclear. Theoretical models predict that beneficial linkage between a sex determining gene and a nearby locus with sex-specific effects is the initial driver of sex chromosome evolution, as recombination is selected against in this region (Charlesworth and Charlesworth 1978; Charlesworth et al. 2005). However, empirical evidence in support of this model is limited and indirect; one instance of sex chromosome origination has been linked to sexual conflict over colouration (Roberts et al. 2009), and the formation of a neo-sex chromosome may be due to mate choice (Kitano et al. 2009). It is still therefore unclear whether inversions are the cause or neutral consequence of recombination suppression.

5.2 PURIFYING SELECTION DRIVES THE EVOLUTION OF THE W CHROMOSOME

Because W and Y chromosomes are non-recombining, they are predicted to follow very different evolutionary trajectories compared to the other regions of the genome (Kirkpatrick and Hall 2004b; Mank 2012). The lack of recombination is predicted to result in rapid gene loss and decay as a result of linkage and Hill-Robertson effects soon after recombination is halted (Charlesworth and Charlesworth 2000). Furthermore, due to their unique sex-limited inheritance patterns, the W and Y chromosomes are also predicted to be important hotspots for sex-specific fitness effects (Mank 2012). Despite these empirical predictions, the selective forces driving the evolution of the W chromosome are unclear, partly due to a previous paucity of

known W-linked genes. Additionally, the avian W chromosome does not determine sex (Smith et al. 2009), and therefore its role in encoding sex-specific fitness is unclear.

Using the largest catalog of W-linked genes to date, I was able to examine the relative roles of genetic drift and selection in driving W chromosome evolution (Chapter 3, Wright et al. In press). I uncovered a clear signal that W-linked genes are evolving with a significant contribution of purifying selection and found evidence of positive selection in one W-linked gene.

These findings are consistent with studies of male heterogametic species, which have revealed strong purifying selection acting on the Y chromosome (Kaiser et al. 2011; Hughes et al. 2012; Bellott et al. 2014; Wilson Sayres et al. 2014) and signatures of positive selection (Gerrard and Filatov 2005; Li et al. 2013). Additionally, a recent study in *G. gallus* examined changes in the expression of W-linked genes (Moghadam et al. 2012) to suggest that the W chromosome has an important role in encoding female fitness traits, consistent with my results. However, given the strong selective pressures presumably associated with domestication in *G. gallus*, until now it was unclear how widely applicable these results were to W chromosomes as a whole.

Interestingly, my results have important implications for our understanding of sex chromosome degeneration and the pervasive role of genetic drift across different evolutionary environments. I find that a large number of avian W-linked genes are conserved across 90 million years, in contrast to the rapid rate of gene loss documented in mammals (Hughes et al. 2010; Hughes et al. 2012). The *Drosophila* Y chromosome has also experienced rapid degeneration, however the relative roles of male heterogamety versus lack of recombination in males are difficult to disentangle (Bachtrog 2013). Theoretical models predict key differences in the rate of coding

sequence degeneration between W and Y chromosome evolution (Kirkpatrick and Hall 2004b Mank 2012; Wright and Mank 2013). The Y chromosome has a lower effective population size than the W chromosome when sexual selection is acting primarily on males (Wright and Mank 2013). As a result, the magnitude of genetic drift acting on Y-linked coding sequences is stronger, and this effect is compounded by male-mutation bias (Kirkpatrick and Hall 2004b). Correspondingly, we predict more rapid degeneration of Y- versus W-linked genes, consistent with my findings.

My results indicate that adaptive evolution together with a significant contribution of purifying selection is not restricted to male heterogametic systems and represents a fundamental feature of heterogametic sex chromosome evolution. However, there are key differences in the evolutionary trajectories of W versus Y chromosomes, predominantly driven by their distinct inheritance patterns.

5.2.1 Future directions

There is considerable potential to build on my findings on W chromosome evolution with the incorporation of additional species, spanning longer evolutionary time periods and across diverse mating systems. Expanding this approach would shed light on two key aspects of heterogametic sex chromosome evolution that remain enigmatic; 1) Dynamics of W chromosome expression and regulation, 2) Role of mating system in driving degeneration of the W chromosome.

Firstly, previous work has indicated that W chromosome expression is dynamic throughout development (Mank et al. 2010b), and likely regulated via a complex mixture of *trans* and *cis* mechanisms (Moghadam et al. 2012). However, we lack a detailed understanding of the relative role of each mechanism, the extent and

phenotypic implications of copy number variation and the developmental profile of W-linked gene expression across a diverse range of species.

Secondly, the role of mating system and W chromosome evolution has yet to be explored empirically (Bachtrog et al. 2011). Specifically, under increasing sexual selection and variance in male reproductive fitness, the effective population size of the W chromosome increases and therefore we expect weaker genetic drift and larger conservation of W-linkage. Exploring the extent of W chromosome turnover under varying levels of sexual selection would have important implications for our understanding of heterogametic sex chromosome evolution.

5.3 COMPLEX INTERPLAY BETWEEN GENETIC DRIFT AND SEX-SPECIFIC SELECTION ON THE Z CHROMOSOME

Sex chromosomes are subject to a multitude of distinct evolutionary forces (Wright and Mank 2013), and given their unequal pattern of inheritance, we expect the nature and efficacy of selection and genetic drift acting on Z- and X-linked genes to differ relative to the autosomes.

Transcriptomic data reveal that the Z chromosome is shaped by cumulative masculinizing selection, where older regions harbour more male-biased genes than younger regions (Chapter 2, Wright et al. 2012). This finding indicates that the avian Z chromosome is implicated in sex-specific fitness, and is consistent with theoretical predictions that Z-linked dominant male-benefit alleles are subject to strong male-specific selection as a result of the male-biased inheritance pattern (Rice 1984; Charlesworth et al. 1987). Interestingly, the Z chromosome is also theoretically under

selection for the evolution of dosage compensation mechanisms to equalize gene dose between the Z chromosome and autosomes (Charlesworth 1996). However, my findings indicate that there is a trade-off between selection for dosage compensation and sex-specific selection on the Z chromosome.

I also uncovered consistently faster rates of avian Z-linked sequence divergence relative to the autosomes. By combining expression, sequence and polymorphism data with measures of sperm competition and promiscuity, I present the first empirical evidence demonstrating the positive relationship between Faster-Z Effect and measures of promiscuity. However, this rapid divergence is not adaptive, as my earlier work on sex-specific selection might suggest. Instead I found that selection is less effective on the Z chromosome, particularly in promiscuous species, and that Faster-Z Evolution in birds is due primarily to genetic drift (Chapter 4, Wright et al. In review). These findings are consistent with theoretical models (Charlesworth et al. 1987; Vicoso and Charlesworth 2009). Under monogamy when male and female mating success is equal, genetic drift and fixation of weakly deleterious mutations is greater on the Z chromosome as the effective population size of Z-linked genes is ¾ that of autosomal genes (Charlesworth et al. 1987; Vicoso and Charlesworth 2009). The relative magnitude of drift acting on the Z chromosome is predicted to increase with increasing male promiscuity and sexual selection (Vicoso and Charlesworth 2009), which decreases the effective population size of the Z chromosome relative to the autosomes.

Together, my results suggest a complex interplay between mating system, sex-specific processes and genetic drift in driving coding and sequence evolution on the Z chromosome. This finding has important implications for our understanding of the role

of sex chromosomes in sexual selection. Considerable theory predicts that the sex chromosomes, and particularly the Z chromosome, promote the evolution of some types of sexual dimorphism and the resolution of sexual conflict (Rice 1984; Reeve and Pfennig 2003; Kirkpatrick and Hall 2004a). This is consistent with my results showing that the Z chromosome has been masculinized over long evolutionary time scales. However, empirical data in support of these predictions is generally mixed (Dean and Mank 2014), and many studies have revealed that the X and Z chromosomes across many species are not enriched for sex-specific fitness effects (Knief et al. 2012; Schielzeth et al. 2012; Pointer et al. 2013). My finding that short-term Z chromosome evolution is dominated by genetic drift may help explain this discrepancy, as drift would act to weaken the efficacy of sex-specific selection, particularly in promiscuity species. Consistent with this conclusion, I find that the rate of Z chromosome masculinization is relatively slow and that younger regions are not significantly enriched with male-biased genes. Consequently, genetic drift may limit the adaptive role of the Z chromosome in general, and in particular its role in encoding sexually selected traits.

Together, my findings reveal that the Z chromosome is subject to a multitude of selective agents, and the evolution of gene expression and coding sequence reflects a trade-off between these evolutionary forces.

5.4 CONTRASTING MALE AND FEMALE HETEROGAMETIC SYSTEMS TO IDENTIFY THE FUNDAMENTAL FORCES IN SEX CHROMOSOME EVOLUTION

Contrasting studies of male and female heterogametic systems are vital in order to disentangle the fundamental forces governing sex chromosome evolution, from the evolutionary processes associated primarily with sex. Here, I discuss key themes emerging from my findings.

Dosage compensation until recently was thought to evolve alongside the gradual divergence of sex chromosomes and degeneration of sex-limited sequences (Charlesworth 1996). However, recently it has been revealed that there is large variation in the degree and mechanism of compensation (Mank 2009b). A striking pattern to have emerged is that complete dosage compensation is restricted to male heterogametic species (Mank and Ellegren 2009a; Deng et al. 2011; Vicoso and Bachtrog 2011; Julien et al. 2012; Lin et al. 2012; Muyle et al. 2012; Pessia et al. 2012; Vicoso et al. 2013a), with the recent exception of the moth (Smith et al. 2014). My finding that a trade-off between sex-specific selection and dosage compensation exists on the Z chromosome may help explain this apparent divide in compensation status between XY and ZW systems (Chapter 2, Wright et al. 2012). If the Z chromosome is dominated by male-specific selection, this may over-ride selection in females for dosage compensation. In XY systems, selection for dosage compensation acts occurs in males on the X chromosome, and this may explain the higher proportion of male heterogametic species with some form of sex chromosome dosage compensation.

The relative exposure to male-specific processes between male and female heterogametic sex chromosomes has important consequences for our understanding

of many aspects of sex chromosome evolution. I have already discussed the impact of male-biased mutation rate, together with differences in effective population size, in driving relative rates of W- versus Y-linked degeneration rates (Chapter 3, Wright et al. In press). However, we might also expect differences in the nature of the Faster-Z and Faster-X Evolution. I revealed that Z chromosome evolution is governed primarily by drift, which may limits its adaptive role in encoding some types of sexually dimorphic phenotypes drift (Chapter 4, Wright et al. In review). However, numerous studies have uncovered signatures of positive selection on the X chromosome across different species, consistent with theoretical models that Faster X evolution is adaptive (Thornton and Long 2005; Baines et al. 2008; Grath et al. 2009; Hvilsom et al. 2012; Langley et al. 2012). The neutral versus adaptive explanations of Faster-Z and Faster-X Evolution respectively are supported by theoretical predictions (Vicoso and Charlesworth 2009), arising from differences in relative effective population size under increasing promiscuity. These disparities may have important consequences for the relative role of the Z and X chromosome in encoding sexual dimorphisms, and effect of mating system in driving sex chromosome evolution.

5.5 CONCLUSIONS

Due to their unique inheritance pattern, the sex chromosomes experience distinct evolutionary environments relative to the autosomes, and contrasts between the sex chromosomes and the autosomes can be used to shed light on the fundamental evolutionary forces acting across the genome as a whole. In particular, female heterogametic species such as birds provide an essential comparison to male heterogamety to identify those factors that affect sex chromosomes in general from

those that are unique to specific types of sex chromosome systems. In this thesis, I combine genomic and transcriptomic data across a wide range of avian species to explore the evolutionary processes governing sex chromosome evolution.

In doing so, I reveal a complex mosaic of recombination suppression between the avian Z and W chromosomes, characterized by repeated and independent divergence of gametologs, together with ongoing genetic exchange. Additionally, I uncover a trade-off between sex-specific processes and genetic drift on both the Z and W chromosomes, with very different consequences. Specifically, my findings indicate that although the Z chromosome is masculinized for male-specific effects, rapid divergence of Z-linked coding sequences is a consequence of elevated genetic drift, the magnitude of which varies across mating systems. Together, my results suggest that the role of the Z chromosome in encoding adaptive sex-specific phenotypes may be limited, particularly in species with high variance in male reproductive fitness. In contrast, evolution of the female-limited W chromosome is governed predominately by purifying selection, and despite considerable degeneration of W-linked coding sequences, my results suggest the remaining genes play a significant role in female-specific fitness. In aggregate, my findings reveal the power of mating system in shaping broad patterns of genome evolution.

Appendix

Table S.3.1. Sequence divergence between *G. gallus* gametologs.

Gene pair (Z/W)	Gene ID[a] (Z/W)	Position on *G .gallus* Z (mb)	*G. gallus* d_S (SE)[b]	*G.gallus* d_N (SE)[b]	Putative Stratum	Divergence (MYA)
CHD1Z/	14642/		0.390	0.0322		
CHD1W	*CHDW*[1]	51.10	(0.0246)	(0.0029)	Conserved I	103
HINT1Z/	00428/		0.404	0.275		
HINT1W	22690	44.76	(0.1087)	(0.0503)	Conserved I	106
RASA1Z/	17706/		0.241	0.028		
RASA1W	22611	60.83	(0.0526)	(0.0094)	Conserved II	63
KCMF1Z/	15391/		0.271	0.036		
KCMF1W	14441	53.60	(0.0371)	(0.0068)	Conserved II	71
SPINZ/	14916/		0.178	0.007		
SPINW	*SPINW*[1]	43.03	(0.0347)	(0.0035)	Galliform-specific III	47
HNRNPKZ/	12591/		0.268	0.003		
HNRNPKW	14366	39.94	(0.0351)	(0.0019)	Galliform-specific III	71
BTF3Z/	13512/		0.245	0.024		
BTF3W	00395	24.60	(0.0526)	(0.0087)	Galliform-specific III	64
ZSWIM6Z/	14734/		0.156	0.024		
ZSWIM6W	27170	18.95	(0.0210)	(0.0048)	Galliform-specific III	41
MIER3Z/	14721/		0.210	0.025		
MIER3W	00140	17.22	(0.0426)	(0.0077)	Galliform-specific III	55
NIPBLZ/	03605/		0.198	0.023		
NIPBLW	13312	11.23	(0.0111)	(0.0018)	Galliform-specific IV	52
RPL17LZ/	02696/		0.169	0.000		
RPL17LW	22174	10.80	(0.0212)	(0.0000)	Galliform-specific IV	44
ZFRZ/	03235/		0.158	0.018		
ZFRW	14545	9.61	(0.0157)	(0.0029)	Galliform-specific IV	42
VCPZ/	01986/		0.162	0.000		
VCPW	00386	8.40	(0.0214)	(0.0009)	Galliform-specific IV	43
UBAP2Z/	13809/		0.194	0.049		
UBAP2W	05785	7.22	(0.0165)	(0.0049)	Galliform-specific IV	51
UBE2R2Z/	01668/		0.257	0.000		
UBE2R2W	09227	7.16	(0.0258)	(0.0000)	Galliform-specific IV	68
ATP5A1Z/	14644/		0.193	0.019		
ATP5A1W	01756	2.15	(0.0231)	(0.0042)	Galliform-specific IV	51
CZH18ORF25Z/	01763/		0.139	0.072		
CZH18ORF25W	01585	2.10	(0.0294)	(0.0126)	Galliform-specific IV	37
MADH2Z/	14697/		0.191	0.005		
MADH2W	10056	1.52	(0.0251)	(0.0023)	Galliform-specific IV	50
ZNF532Z/	02852/		0.152	0.032		
ZNF532W	14003	0.86	(0.0151)	(0.0039)	Galliform-specific IV	40
SIAT8CZ/	03049/		0.173	0.021		
SIAT8CW	26991	0.44	(0.0266)	(0.0053)	Galliform-specific IV	45

[a]Ensembl gene ID Ensgalg000000...
[b]Standard errors generated in Paml.
[1]Ayers et al. 2013

Table S.3.2. Sequence divergence between *M. gallopavo* gametologs.

Gene pair (Z/W)	Gene ID[a] (Z/W)	Position on *G gallus*/ *M.gallopavo* Z (mb)	*M.gallopavo* d_S (SE)[b]	*M.gallopavo* d_N (SE)[b]	Putative Stratum	Divergence (MYA)
HINT1Z/	06655/		0.285	0.182	Conserved	
HINT1W	16898	44.76/47.46	(0.0961)	(0.0434)	I	75
RASA1Z/	08292/		0.266	0.0312	Conserved	
RASA1W	AH015047.1[1]	60.83/64.54	(0.0600)	(0.0106)	II	70
SPINZ/	02500/		0.198	0.0129	Galliform-	
SPINW	AH015050.1[1]	43.03/45.85	(0.0423)	(0.0054)	specific III	52
NIPBLZ/	01989/		0.182	0.0087	Galliform-	
NIPBLW	13606	11.23/12.01	(0.0453)	(0.0051)	specific IV	48
UBAP2Z/	01837/		0.137	0.0273	Galliform-	
UBAP2W	AY188758.1[1]	7.22/75.78	(0.0487)	(0.0139)	specific IV	36
ATP5A1Z/	09963/		0.141	0.0076	Galliform-	
ATP5A1W	01414	2.15/2.23	(0.0326)	(0.0044)	specific IV	37
MADH2Z/	01351/		0.146	0.0159	Galliform-	
MADH2W	AH015049.1[1]	1.52/1.51	(0.0485)	(0.0093)	specific IV	38

[a]Ensembl gene ID Ensmgag000000...
[b]Standard errors generated in Paml
[1]Berlin and Ellegren 2006

Table S.3.3. Sequence divergence between *A. platyrhynchos* gametologs.

Gene pair (Z/W)	Gene ID[a] (Z/W)	Position on *G. gallus* Z (mb)	*A. platyrhynchos* d_S (SE)[b]	*A. platyrhynchos* d_N (SE)[b]	Putative Stratum	Divergence (MYA)
CHD1Z/	09965/		0.296	0.0317		
CHD1W	05191	51.10	(0.0399)	(0.0057)	Conserved I	78
RASA1Z/	05627/		0.171	0.008	Conserved	
RASA1W	05611	60.83	(0.0346)	(0.0039)	II	45
KCMF1Z/	13426/		0.198	0.0371	Conserved	
KCMF1W	03026	53.60	(0.0343)	(0.0081)	II	52
SPINZ/	13922/		0.014*	0.000	Anseriform-	
SPINW	02923	43.03	(0.0000)	(0.0000)	specific III	4
HNRNPKZ/	10856/		0.148	0.005	Anseriform-	
HNRNPKW	10986	39.94	(0.0249)	(0.0023)	specific III	39
ZSWIM6Z/	06992/		0.065*	0.0131	Anseriform-	
ZSWIM6W	13555	19.00	(0.0153)	(0.0045)	specific III	17
MIER3Z/	06634/		0.104	0.0278	Anseriform-	
MIER3W	010850	17.22	(0.0182)	(0.0055)	specific III	27
NIPBLZ/	08473/		0.077*	0.0017	Anseriform-	
NIPBLW	10290	11.23	(0.0204)	(0.0017)	specific III	20
ZFRZ/	15627/		0.133	0.0256	Anseriform-	
ZFRW	15519	9.61	(0.0264)	(0.0062)	specific III	35
VCPZ/	04533/		0.005*	0.000	Anseriform-	
VCPW	05806	8.40	(0.0000)	(0.0000)	specific III	1
UBE2R2Z/	03800/		0.000*	0.004	Anseriform-	
UBE2R2W	16000	1.76	(0.0000)	(0.0042)	specific III	0

[a]Ensembl gene ID Ensaplg000000...
[b]Standard errors generated in Paml
*Significantly different from *G. gallus* d_S estimates (non-overlapping 95% confidence intervals)

Table S.3.4. Branch models and branch-site test across W-linked branches.

Gene name	Chicken (Z/W)[a]	Duck (Z/W)[b]	Turkey (Z/W)[c]	Zebra finch (Z)[d]	Branch model (estimate ω) -lnL	ω	Branch model test (ω =1) –lnL	LR	Branch model test (ω =0) –lnL	LR	Branch-site test (ω =1) –lnL (estimate ω)	–lnL (ω =1)	LR
HINT1	00428/ 22690	-	06655/ 16898	00395	-365.13	1.10	-365.15	0.05	-2213.65	**3697.05****	-362.45	-362.47	0.05
CHD1	14642/ *CHDW*[1]	09965/ 05191	-	01185	-3027.39	0.14	-3074.62	**94.45****	-6533.47	**7012.15****	-3009.74	-3009.74	0.00
RASA1	17706/ 22611	05611/ 05627	08292/ AH015047.1[2]	01169	-697.16	0.09	-712.82	**31.33****	-1390.63	**1386.95****	-697.08	-697.08	0.00
KCMF1	15391/ 14441	13426/ 03026	-	00265	-1746.53	0.22	-1764.27	**35.49****	-4583.88	**5674.71****	-1736.81	-1736.92	0.22
MIER3	14721/ 00140	06634/ 10850	-	02601	-1292.61	0.24	-1301.60	**17.97****	-1376.24	**167.25****	-1293.55	-1293.55	0.00
SPIN	14916/ *SPINW*[1]	13922/ 02923	02500/ AH015050.1[2]	00468	-1153.90	0.09	-1168.12	**28.45****	-2027.39	**1749.98****	-1154.67	-1154.67	0.00
HNRNPK	12591/ 14366	10856/ 10986	-	03138	-2340.51	0.03	-2394.89	**108.75****	-2382.62	**84.21****	-2332.91	-2332.92	0.03
VCP	01986/ 00386	04533/ 05806	-	01760	-1081.34	0.03	-1094.13	**25.58****	-1090.90	**19.13****	-1081.23	-1081.31	0.16
ZFR	03235/ 14545	15627/ 15519	-	01909	-1973.85	0.18	-1987.42	**27.16****	-2060.59	**173.50****	-1963.95	-1963.95	0.00
NIPBL	03605/ 13312	08473/ 10290	-	02047	-1487.07	0.05	-1509.10	**44.08****	-1519.02	**63.91****	-1487.31	-1487.31	0.00

NIPBL	03605/ 13312	-	13606/ 01989	02047	-838.30	0.03	-857.50	**38.40****	-1010.05	**343.5****	-838.37	-838.37	0.00
UBAP2	13809/ 05785	-	01837/ AY188758.1[2]	01689	-406.73	0.15	-409.83	**6.19***	-584.14	**354.82****	-406.74	-406.74	0.00
ATP5A1	14644/ 01756	-	09963/ 01414	01582	-1101.40	0.09	-1109.67	**16.54****	-1426.82	**650.83****	-1096.96	-1098.75	**3.58***
MADH2	14697/ 14184	-	01351/ AH015049.1[2]	00016	-1151.96	0.06	-1163.14	**22.37****	-1318.71	**33.50****	-1149.58	-1149.58	0.01
UBE2R2	01668/ 09227	03800/ 16000	-	01683	-514.68	0.04	-522.19	**15.03****	-523.37	**17.39****	-514.68	-514.68	0.00
ZSWIM6	06992/ 13555	14734/ 27170	-	02716	-1014.36	0.25	-1019.98	**11.25****	-1078.80	**128.88****	-1014.15	-1014.15	0.00

[a]Ensgalg000000...
[b]Ensaplg000000...
[c]Ensmgag000000...
[d]Enstgug000000...
** p-value <0.01
* p-value <0.05
[1]Ayers et al. 2013
[2]Berlin and Ellegren 2006

Table S.3.5. Branch models and branch-site test across Z-linked branches.

Gene name	Chicken (Z/W)[a]	Duck (Z/W)[b]	Turkey (Z/W)[c]	Zebra finch (Z)[d]	Branch model (estimate ω)		Branch model test (ω =1)		Branch model test (ω =0)		Branch-site test (ω =1)		
					–lnL	ω	–lnL	LR	–lnL	LR	–lnL (estimate ω)	–lnL (ω =1)	LR
HINT1	00428/ 22690	-	06655/ 16898	00395	-367.32	0.00	-370.99	**7.34****	-367.32	0.00	-370.17	-370.17	0.00
CHD1	14642/ *CHDW*[1]	09965/ 05191	-	01185	-3035.79	0.03	-3122.01	**172.44****	-4005.27	**1938.96****	-3026.57	-3026.57	0.00
RASA1	17706/ 22611	05611/ 05627	08292/ AH015047.1[2]	01169	-696.13	0.03	-709.51	**26.77****	-868.30	**344.34****	-697.08	-697.08	0.00
KCMF1	15391/ 14441	13426/ 03026	-	00265	-1749.50	0.12	-1764.45	**29.89****	-3019.36	**259.71****	-1740.93	-1740.93	0.00
MIER3	14721/ 00140	06634/ 10850	-	02601	-1293.87	0.11	-1306.99	**26.23****	-1324.28	**60.82****	-1294.25	-1294.25	0.00
SPIN	14916/ *SPINW*[1]	13922/ 02923	02500/ AH015050.1[2]	00468	-1156.96	0.02	-1172.79	**31.65****	-1326.73	**339.55****	-1157.20	-1157.20	0.00
HNRNPK	12591/ 14366	10856/ 10986	-	03138	-2339.38	0.01	-2393.62	**108.49****	-2343.99	**9.22****	-2333.75	-2333.75	0.00
VCP	01986/ 00386	04533/ 05806	-	01760	-1081.20	0.00	-1095.52	**28.65****	-1081.20	0.00	-1081.67	-1081.67	0.00
ZFR	03235/ 14545	15627/ 15519	-	01909	-1975.63	0.12	-2003.04	**54.84****	-2055.82	**160.38****	-1968.12	-1968.16	0.09

NIPBL	03605/ 13312	08473/ 10290	-	02047	-1487.64	0.02	-1529.38	**83.48****	-1503.87	**32.46****	-1488.06	-1488.06	0.00
NIPBL	03605/ 13312	-	13606/ 01989	02047	-838.46	0.04	-845.81	**14.70****	-1008.86	**340.80****	-838.47	-838.47	0.00
UBAP2	13809/ 05785	-	01837/ AY188758.1[2]	01689	-406.59	0.25	-407.82	2.47	-751.58	**689.98****	-406.61	-406.61	0.00
ATP5A1	14644/ 01756	-	09963/ 01414	01582	-1105.90	0.00	-1132.66	**53.53****	-1105.89	-0.01	-1103.10	-1103.10	0.00
MADH2	14697/ 14184	-	01351/ AH015049.1[2]	00016	-1152.46	0.02	-1182.98	**61.03****	-1485.78	**666.65****	-1150.21	-1150.21	0.00
UBE2R2	01668/ 09227	03800/ 16000	-	01683	-515.67	0.00	-520.67	**10.01****	-515.66	0.00	-515.82	-515.82	0.00
ZSWIM6	06992/ 13555	14734/ 27170	-	02716	-1014.55	0.05	-1021.05	**12.99****	-1020.57	**12.05****	-1015.06	-1015.06	0.00

[a]Ensgalg000000...
[b]Ensaplg000000...
[c]Ensmgag000000...
[d]Enstgug000000...
** p-value <0.01
* p-value <0.05
[1]Ayers et al. 2013
[2]Berlin and Ellegren 2006

Table S.4.1. Phylogenetically controlled regression analyses between Faster-Z and log sperm number, and residual testes weight.

Species excluded	Residual testis weight				Log sperm number			
	t	p-value	r^2	Beta	t	p-value	r^2	Beta
Meleagris gallopavo	4.256	**0.007**	0.775	0.025	5.619	**0.003**	0.855	0.017
Phasianus colchicus	6.614	**0.001**	0.897	0.039	3.526	**0.013**	0.710	0.021
Numida meleagris	2.542	**0.032**	0.559	0.033	3.301	**0.016**	0.683	0.021
Anas platyrhnchos	1.503	0.104	0.311	0.031	2.240	**0.046**	0.499	0.028
Anser cygnoides	1.732	0.079	0.375	0.036	17.934	**<0.010**	0.985	0.063
Pavo cristatus	2.491	**0.034**	0.548	0.031	3.410	**0.014**	0.696	0.021

Parameters are shown for the regression analysis when the given species is excluded
Significant values are shown in bold.

Table S.4.2. Number of nonsynonymous and synonymous polymorphic and fixed sites.

Species	Z chromosome				Autosomes 1-10			
	D_N	D_S	P_N	P_S	D_N	D_S	P_N	P_S
Meleagris gallopavo	4472	11563	51	83	38835	118922	1148	3226
Phasianus colchicus	4094	11419	88	155	37842	117996	1618	4898
Numida meleagris	3565	9472	29	100	33399	99449	1299	3681
Anas platyrhnchos	2901	8233	107	329	26162	87629	2265	9360
Anser cygnoides	2379	7093	111	200	22486	74388	2060	5572
Pavo cristatus	3750	10420	35	59	35767	110749	573	1250

D_N, D_S, P_N and P_S are rounded to the nearest integer.

Table S.4.3. Site model test results for contigs with signatures of positive selection.

Gene name [a]	Chr	*ω*	Proportion of sites	M1a LR	M2a LR	LRT	p-value	p-fdr value[b]
22552	1	2.897	0.122	-6535.857	-6522.227	27.259	< 0.001	0.003
21101	1	4.155	0.033	-14063.297	-14050.286	26.023	< 0.001	0.006
37104	1	4.926	0.036	-4435.291	-4424.814	20.954	< 0.001	0.073
36845	1	4.415	0.075	-3419.672	-3409.907	19.531	< 0.001	0.148
16502	1	3.237	0.144	-2264.272	-2255.709	17.126	< 0.001	0.492
21418	1	3.079	0.055	-5553.341	-5545.776	15.130	0.001	1.334
21457	1	1.954	0.069	-12752.825	-12745.364	14.922	0.001	1.479
24837	1	3.717	0.093	-1309.612	-1303.078	13.069	0.001	3.731
23319	1	5.404	0.008	-11327.088	-11320.745	12.686	0.002	4.515
14357	1	2.471	0.174	-3251.079	-3244.950	12.258	0.002	5.588
26993	1	20.333	0.008	-1343.505	-1338.433	10.145	0.006	16.008
23629	1	3.562	0.044	-3033.798	-3029.348	8.900	0.012	29.785
27590	1	5.082	0.013	-6679.670	-6675.904	7.531	0.023	58.805
26389	1	20.985	0.002	-2662.400	-2659.000	6.800	0.033	84.670
27463	1	4.641	0.021	-2788.178	-2784.963	6.429	0.040	101.587
44252	2	2.831	0.165	-1432.129	-1427.436	9.386	0.009	23.370
45482	2	10.290	0.002	-4950.364	-4946.989	6.749	0.034	86.841
26112	2	1.687	0.167	-5851.565	-5848.226	6.677	0.035	89.889
46085	2	1.705	0.326	-2701.980	-2698.729	6.502	0.039	98.009
20347	2	1.705	0.099	-9450.060	-9446.902	6.315	0.043	107.477
31776	3	4.608	0.130	-1270.098	-1256.430	27.337	< 0.001	0.003
14587	3	4.393	0.040	-2761.834	-2753.484	16.700	< 0.001	0.609
16201	3	3.925	0.027	-8837.436	-8829.793	15.286	< 0.001	1.235
24568	3	5.232	0.015	-4629.778	-4625.429	8.697	0.013	32.942
17328	3	2.585	0.112	-1886.600	-1883.057	7.085	0.029	73.470
37447	3	2.523	0.078	-1740.906	-1737.739	6.333	0.042	106.550
16068	3	6.676	0.006	-2283.082	-2280.007	6.150	0.046	116.661
17992	4	3.812	0.033	-10248.079	-10239.377	17.406	< 0.001	0.428
23095	4	4.149	0.023	-10276.768	-10269.900	13.737	0.001	2.674
22955	4	4.580	0.013	-9871.493	-9864.857	13.274	0.001	3.368
14387	4	6.044	0.014	-5452.766	-5447.180	11.173	0.004	9.600
32094	4	4.363	0.038	-2032.598	-2028.562	8.072	0.018	44.952
22858	4	21.691	0.003	-3252.065	-3248.231	7.668	0.022	54.956
25298	4	3.750	0.055	-2108.140	-2104.827	6.626	0.036	92.137

12746	5	3.833	0.042	-3266.303	-3260.081	12.445	0.002	5.090
12190	5	7.557	0.007	-7413.128	-7407.132	11.992	0.002	6.382
17502	5	4.310	0.012	-9231.539	-9225.994	11.091	0.004	9.996
18863	5	1.521	0.177	-10833.724	-10829.701	8.046	0.018	45.523
14067	5	4.615	0.004	-5474.941	-5471.907	6.067	0.048	121.549
39919	6	4.226	0.310	-1630.735	-1611.278	38.915	< 0.001	0.000
10311	6	6.691	0.043	-1636.382	-1626.707	19.351	< 0.001	0.162
13711	6	2.190	0.129	-7523.106	-7517.618	10.975	0.004	10.580
06741	7	5.639	0.020	-4454.582	-4447.832	13.499	0.001	3.011
18021	7	2.755	0.046	-8843.415	-8837.885	11.060	0.004	10.147
21611	7	3.740	0.024	-7320.455	-7316.242	8.426	0.015	37.692
14983	7	26.827	0.002	-3520.059	-3515.955	8.207	0.017	42.024
03831	8	4.817	0.080	-9607.226	-9560.287	93.878	< 0.001	0.000
09004	8	8.990	0.007	-4487.428	-4483.521	7.813	0.020	51.111
00626	8	2.660	0.029	-7503.872	-7500.522	6.700	0.035	88.907
013645	9	6.164	0.005	-4614.796	-4610.506	8.580	0.014	34.924
11804	10	4.057	0.006	-3908.867	-3903.212	11.309	0.004	8.972
39713	11	2.279	0.254	-2696.490	-2691.001	10.980	0.004	10.559
41379	11	2.222	0.330	-974.822	-971.467	6.711	0.035	88.436
01593	11	3.991	0.061	-1953.134	-1949.807	6.654	0.036	90.900
44152	11	8.060	0.020	-1087.697	-1084.699	5.997	0.050	125.827
34137	12	1.757	0.408	-3027.514	-3022.796	9.435	0.009	22.810
10504	15	3.343	0.072	-5389.616	-5375.473	28.287	< 0.001	0.002
11333	17	3.164	0.041	-7288.959	-7283.483	10.952	0.004	10.699
05557	18	3.172	0.050	-6000.642	-5995.829	9.626	0.008	20.744
02792	19	4.005	0.038	-1981.892	-1971.930	19.923	< 0.001	0.122
01367	19	25.280	0.001	-5913.302	-5909.030	8.544	0.014	35.534
01868	20	9.422	0.013	-4192.958	-4179.195	27.526	< 0.001	0.003
07089	21	2.003	0.294	-2347.262	-2343.741	7.042	0.030	75.049
02022	28	4.914	0.068	-2768.690	-2753.634	30.110	< 0.001	0.001
29351	Z	6.802	0.104	-2974.892	-2930.430	88.925	< 0.001	0.000
24394	Z	4.087	0.029	-6594.462	-6587.094	14.736	0.001	1.624
23931	Z	9.163	0.007	-3248.608	-3242.892	11.431	0.003	8.443
08032	Z	13.117	0.001	-8849.748	-8844.406	10.685	0.005	12.226
24368	Z	5.886	0.011	-7464.550	-7461.493	6.113	0.047	118.812

[a]ENSGALT000000…..
[b]Sequential Bonferroni correction (Holm 1979)

Table S.4.4. McDonald Kreitman test results.

Species	Residual testis weight		Log sperm number	
	No. of genes[a]	Positive selection[b/c]	No. of genes	Positive selection
Meleagris gallopavo	3	0/0	250	1/0
Phasianus colchicus	12	0/0	374	2/0
Numida meleagris	4	0/0	253	1/1
Anas platyrhnchos	15	0/0	572	8/0
Anser cygnoides	15	0/0	423	3/0
Pavo cristatus	3	0/0	78	1/0
Total	**52**	**0/0**	**1950**	**16/1**

[a]Contigs was excluded if the sum of each marginal row and column of the 2x2 contingency table was less than 6.
[b]P values were corrected with a false discovery rate of 0.05 and lambda 0 to obtain q values. The number of contigs with a significant uncorrected p value is shown first, followed by the number after correction for multiple testing.
[c]Excess of nonsynonymous divergence relative to polymorphism indicates positive selection.

Table S.4.5. Faster-Z Effect across expression classes.

Species	Female-biased[a] $d_{NZ}/d_{SZ} : d_{NA}/d_{SA}$ 95% CI	Unbiased $d_{NZ}/d_{SZ} : d_{NA}/d_{SA}$ 95% CI	Male-biased $d_{NZ}/d_{SZ} : d_{NA}/d_{SA}$ 95% CI
Meleagris gallopavo	1.309 (0.773-2.143) Z = 22, A = 387	1.104 (0.873-1.338) Z = 73, A = 853	1.217 (0.905-1.543) Z = 63, A = 437
Phasianus colchicus	1.09 (0.662-1.600) Z = 21, A = 417	1.015 (0.806-1.26) Z = 77, A = 863	1.199 (0.887-1.556) Z = 60, A = 403
Numida meleagris	1.69 (0.730-1.500) Z = 25, A = 397	0.827 (0.651-1.024) Z = 75, A = 897	1.491 (1.125-1.888) Z = 58, A = 380
Anas platyrhnchos	1.041 (0.709-1.540) Z = 44, A = 450	1.141 (0.828-1.490) Z = 65, A = 828	1.354 (0.949-1.855) Z = 50, A = 399
Anser cygnoides	1.146 (0.731-1.657) Z = 24, A = 413	0.869 (0.656-1.103) Z = 74, A = 835	1.434 (1.087-1.875) Z = 61, A = 426
Pavo cristatus	1.315 (0.943-1.763) Z = 26, A = 409	1.151 (0.709-1.151) Z = 64, A =803	1.248 (0.929-1.613) Z = 69, A = 458

[a]Expression category is assigned using both fold change and t-tests separately for each species.

Table S.4.6. Differences in Faster-Z Effect between expression classes.

Species	Female- and male-biased[a] p-value	Female- and unbiased p-value	Male- and unbiased p-value
Meleagris gallopavo	1.000	0.944	1.000
Phasianus colchicus	1.000	1.000	0.812
Numida meleagris	0.400	0.484	**0.004**
Anas platyrhnchos	0.692	1.000	1.000
Anser cygnoides	0.940	0.560	0.056
Pavo cristatus	1.000	0.124	0.240

[a]Expression category is assigned using both fold change and t-tests separately for each species
Significant values are shown in bold and calculated using 1000 permutation tests.

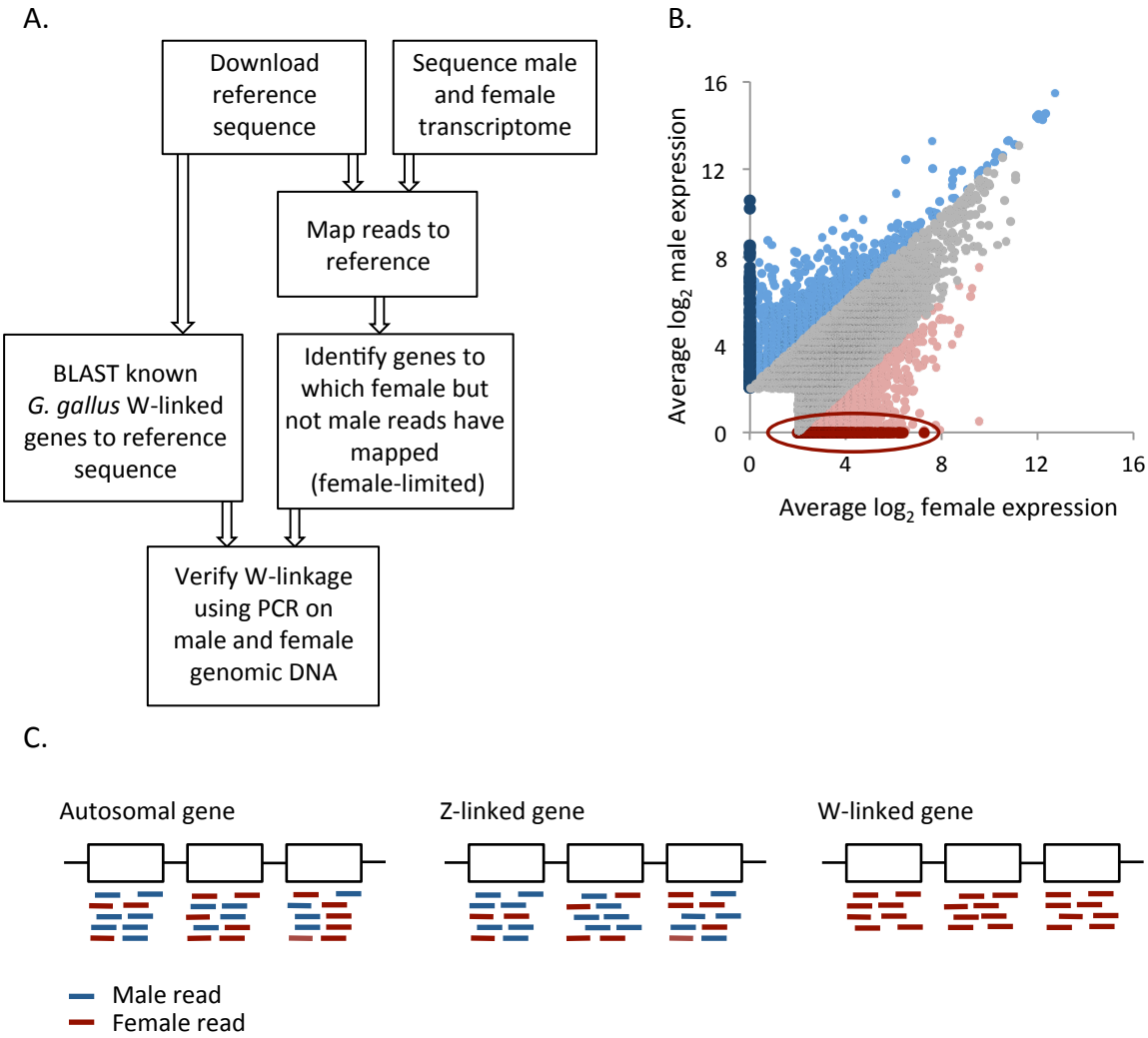

Fig. S.3.1. Identification of W-linked genes.

Panel A. Pipeline of expression and BLAST based approaches used to identify W-linked genes. Panel B. Schematic illustrating the distribution of sex-biased genes after Illumina reads are mapped to the reference genome. Female-limited genes and strongly female biased genes are highlighted as putative W-linked genes. Panel C. Diagram outlining the distribution of male and female Illumina reads after mapping to the reference genome. Female but not male reads will map to putative W-linked genes.

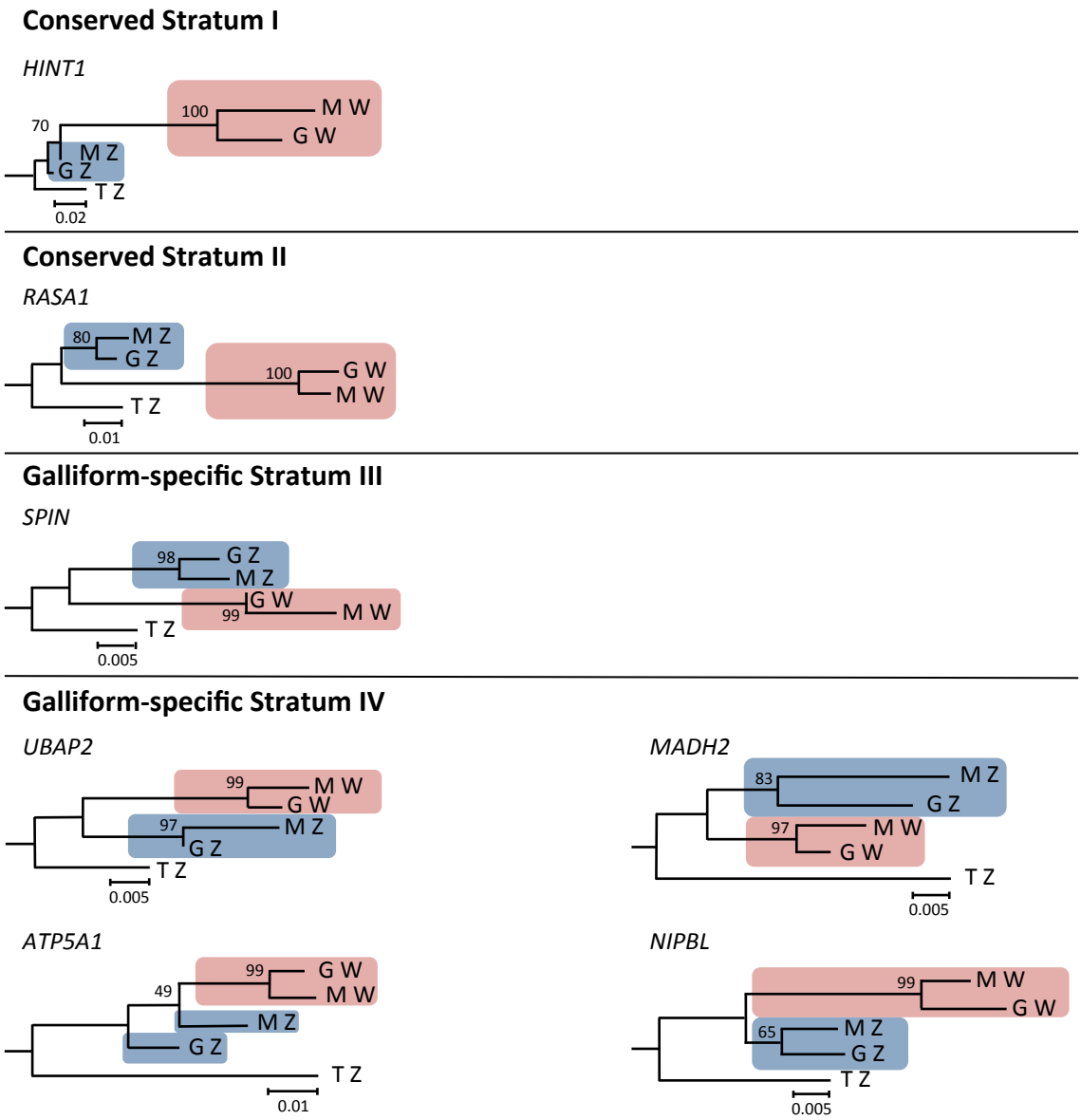

Fig. S.3.2. Gene trees for *M. gallopavo* gametologs.

Maximum-likelihood gene phylogenies for *M. gallopavo* gametologs are shown. Bootstrap values were calculated using 1000 permutations. AZ, GZ, MZ, TZ corresponds to *A. platyrhynchos*, *G. gallus*, *M. gallopavo* and *T. guttata* Z-linked genes.

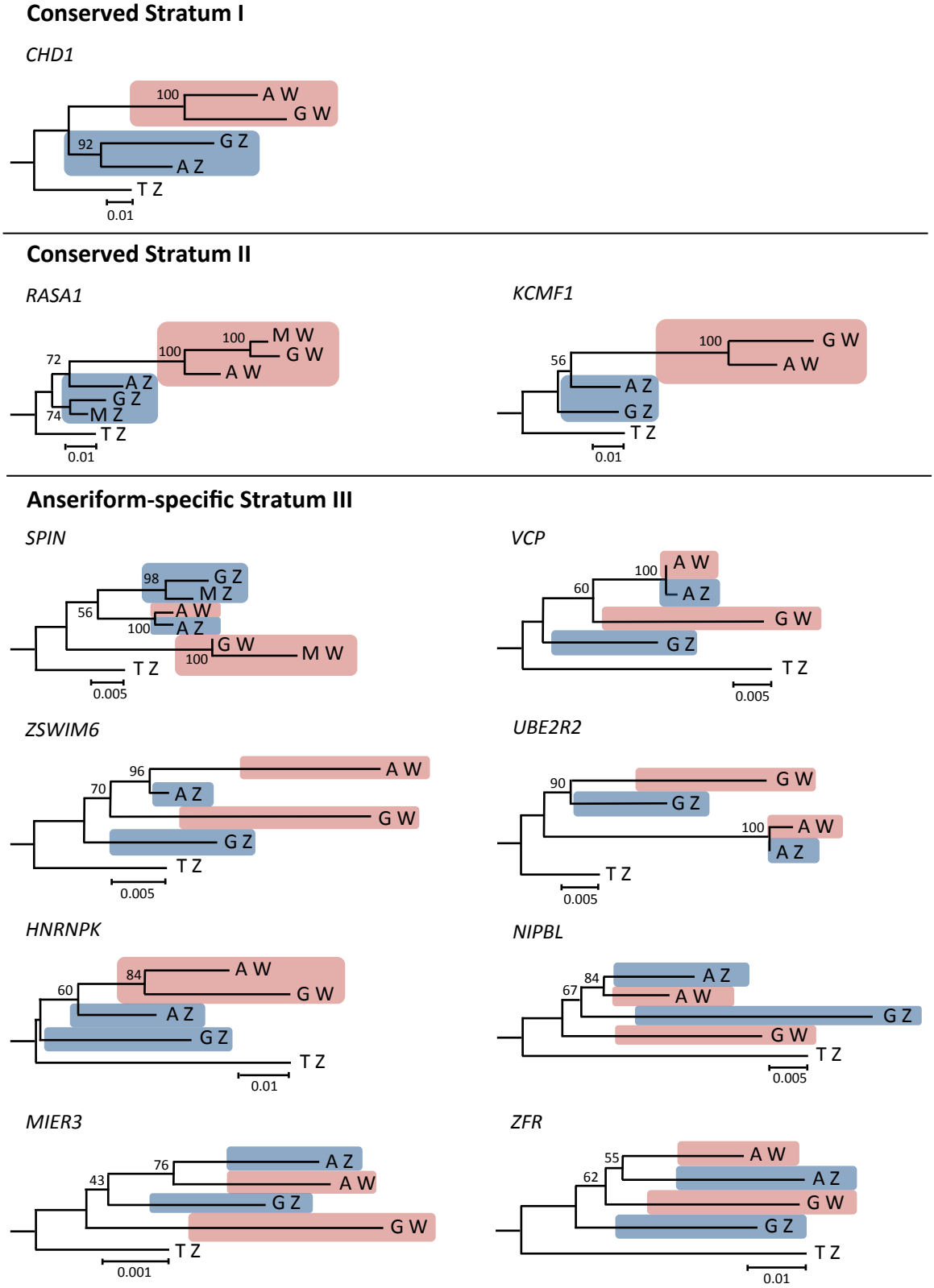

Fig. S.3.3. Gene trees for *A. platyrhynchos* gametologs.

Maximum-likelihood gene phylogenies for *A. platyrhynchos* gametologs are shown. Bootstrap values were calculated using 1000 permutations. AZ, GZ, MZ, TZ corresponds to *A. platyrhynchos*, *G. gallus*, *M. gallopavo* and *T. guttata* Z-linked genes.

Key publications

Wright AE, HK Moghadam, and JE Mank. (2012) Trade-off between selection for dosage compensation and masculinization on the avian Z chromosome. Genetics. 192:1433-1445.

Moghadam HK, MA Pointer, **AE Wright**, S Berlin and JE Mank. (2012) W chromosome expression responds to female-specific selection. Proceedings of the National Academy of Sciences, USA. 109:8207-8211.

Wright AE and JE Mank. (2012) Battle of the sexes: Conflict over dosage-sensitive genes and the origin of X chromosome inactivation. Proceedings of the National Academy of Sciences, USA. 109:5144-5145.

Harrison PW*, **AE Wright*,** and JE Mank. (2012) The evolution of gene expression and the transcriptome-phenotype relationship. Seminars in Cell & Developmental Biology. 23:222-229. (* Joint first authors)

Wright AE and JE Mank (2013). The scope and strength of sex-specific selection in genome evolution. Journal of Evolutionary Biology. 26:1841-53.

Pointer MA, PW Harrison, **AE Wright** and JE Mank (2013) Masculinization of gene expression is associated with exaggeration of male sexual dimorphism. PLoS Genetics. 9:1-9.

Wright AE, PW Harrison, SH Montgomery, MA Pointer and JE Mank. (In press.) Independent stratum formation on the avian sex chromosomes reveals inter-chromosomal gene conversion and predominance of purifying selection on the W chromosome. Evolution.

Mullon C, **AE Wright**, M Reuter, A Pomiankowski and JE Mank. (In review). The effects of sexual selection on the evolution of dosage compensation differ between the X and Z chromosomes.

Wright AE, PW Harrison, F Zimmer, SH Montgomery, MA Pointer, and JE Mank. (In review). Variation in promiscuity and sperm competition drives the evolution of the avian Z chromosome.

JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12201

REVIEW

The scope and strength of sex-specific selection in genome evolution

A. E. WRIGHT*† & J. E. MANK†

*Department of Zoology, University of Oxford, Edward Grey Institute, Oxford, UK
†Department of Genetics, Evolution and Environment, University College London, London, UK

Keywords:

sex chromosomes;
sex-biased gene expression;
sex-specific selection;
sexual dimorphism.

Abstract

Males and females share the vast majority of their genomes and yet are often subject to different, even conflicting, selection. Genomic and transcriptomic developments have made it possible to assess sex-specific selection at the molecular level, and it is clear that sex-specific selection shapes the evolutionary properties of several genomic characteristics, including transcription, post-transcriptional regulation, imprinting, genome structure and gene sequence. Sex-specific selection is strongly influenced by mating system, which also causes neutral evolutionary changes that affect different regions of the genome in different ways. Here, we synthesize theoretical and molecular work in order to provide a cohesive view of the role of sex-specific selection and mating system in genome evolution. We also highlight the need for a combined approach, incorporating both genomic data and experimental phenotypic studies, in order to understand precisely how sex-specific selection drives evolutionary change across the genome.

The ability to attract mates and reproduce is a central component of Darwinian fitness. As well as primary sexual dimorphisms directly involved in reproduction, including numerous gametic proteins key to syngamy, there are many secondary sexual traits involved in mate acquisition (Andersson, 1994; Swanson & Vacquier, 2002; Nadeau *et al.*, 2007). Selection related to sex can be a powerful force given that it is a crucial component in Darwinian fitness, and it can act in opposite directions for males and females due to their distinct reproductive roles and biology (Arnqvist & Rowe, 2005). Aside from sex-limited regions, such as Y or W chromosomes, contradictory female- and male-specific selection regimes act on a genome that is shared between the sexes.

Sex-specific selection is strongly influenced by mating system (Fig. 1), which also affects neutral diversity and evolution. At one end of the spectrum, the potential for conflict between female- and male-specific selection is lowest in monogamous systems. In these species, sex-specific selection often acts primarily on reproductive biology, resulting in very few pronounced secondary sexual dimorphisms (Helfenstein *et al.*, 2004). Promiscuous systems show much more potential for sexual conflict (Lifjeld *et al.*, 1993; Lindenfors *et al.*, 2002; Arnqvist & Rowe, 2005). Polyandrous and polygynous species are often characterized by large differences between male and female parental effort and other aspects of life histories. This creates ample scope for sex-specific selection (Fig. 1) (Kokko *et al.*, 2012). Mating system also influences neutral diversity and rates of evolution. Large differences in the variance in reproductive success between males and females increase the rate of genetic drift, the strength of which varies across the genome depending on the degree and direction of sexual asymmetry in inheritance (Vicoso & Charlesworth, 2009; Mank *et al.*, 2010c). Contrasting evolutionary rates and diversity among these regions, most often between the sex chromosomes and the autosomes, can be used to infer the power of neutral evolutionary forces at work in the genome, as well as estimate mating system (Corl & Ellegren, 2012).

Conflicting sex-specific selection has the potential to create a significant evolutionary burden on a population (Foerster *et al.*, 2007; Morrow *et al.*, 2008; Connallon *et al.*, 2010), and resolving this conflict,

Correspondence: Alison E. Wright, Department of Genetics Evolution and Environment, University College London, The Darwin Building, Gower Street, London WC1E 6BT, United Kingdom.
Tel.: +44 207 679 2170; e-mail: alison.e.wright@ucl.ac.uk

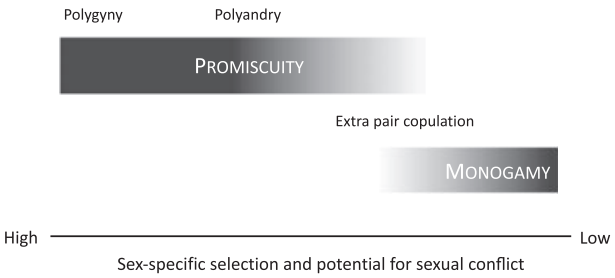

Fig. 1 Mating systems and sex-specific selection. Sex-specific selection is strongly influenced by mating system. The potential for conflict between female- and male-specific selection is lowest in monogamous systems. In mating systems with large differences in reproductive potential, the divergence in male and female fitness optima is greater, as is the probability that sex-specific selection forces are contradictory. As a result, the potential for sexual conflict is predicted to increase with the magnitude of difference in mating success. Additionally, polyandrous and polygynous species are often characterized by large differences between male and female parental effort and other aspects of life histories, which creates ample scope for sex-specific selection. Thus, sex-specific selection and sexual conflict may play a more significant role in polygnous and polyandrous mating systems than monogamous systems.

when it is possible at all, both relieves this burden and allows the sexes to reach separate fitness optima (Chapman *et al.*, 2003). Phenotypic-level studies have revealed a significant level of conflict in many organisms (Chippindale *et al.*, 2001; Arnqvist & Rowe, 2002; Magurran & Seghers, 1994, among many others), affecting many different phenotypes.

Recent genomic and transcriptomic studies have made it possible to take studies to the molecular level, and assess the magnitude, locus, and resolution of conflicting sex-specific selection (Innocenti & Morrow, 2010; Moghadam *et al.*, 2012). These studies indicate that sex-specific evolutionary forces, both adaptive and neutral, affect a large proportion of the genome (Connallon *et al.*, 2010) and play a significant role in determining rates of both adaptive and neutral change for gene sequence, gene expression and genome structure. This is reinforced by the disproportionately large effect that sexually antagonistic loci are predicted to have on genetic variation for fitness (Arnqvist, 2011; Long *et al.*, 2012). Additionally, it is clear from molecular studies that there are different routes to resolving conflict between female- and male- specific selection (Connallon, 2007; Haig, 2009; Gallach & Betran, 2011).

Molecular approaches offer the promise of identifying sexually antagonistic genes and alleles, as well as creating an understanding of how discrete male and female phenotypes are encoded. However, there are a number of important assumptions made when genomic data is used in isolation to infer the nature of sexual conflict and sex-specific selection. One key assumption is that, for the phenotypic traits in question, conflict between female- and male- specific selection has been resolved through the evolution of sex-specific gene regulation or some other mechanism to break down intersexual correlation. Incorporating experimental studies in addition to genomic approaches can shed light on this fundamental assumption and provide detailed information we lack on the fitness consequences of many genomic mechanisms thought to resolve sexual conflict (Tregenza *et al.*, 2006).

Despite the potential of combining phenotypic and molecular data, aside from a few exceptions (Innocenti & Morrow, 2010; Moghadam *et al.*, 2012), these methods have proceeded largely independent of each other. Our purpose here is to synthesize recent molecular genomic advances in order to create a cohesive picture of the importance of sex-specific selection in shaping the evolution of various genomic properties. Ultimately, our goal is to understand how sex-specific selection and mating system affects genome evolution through adaptive and neutral processes, and reinforce the need for a combined approach, incorporating both experimental phenotypic and molecular studies, to create a cohesive understanding of sex-specific selection, its fitness consequences, genomic targets and mechanisms of resolution.

Sex-specific selection and adaptive change

Most molecular genetic analysis of conflicting sex-specific selection focuses on intralocus conflict (conflict where an allele at a given loci is beneficial to one sex but detrimental to the other), where opposite female- and male-selection acts on the same locus. Intralocus conflict can be resolved by a number of genetic mechanisms, the most studied of which is sex-biased expression (Connallon & Knowles, 2005; Gallach & Betran, 2011), which represents the breakdown in intersexual correlation in gene regulation. Additionally, duplication of genes with sex-specific functions (Gallach *et al.*, 2010) may also be important in resolving conflict. In either case, sex-biased or sex-specific gene expression is used as a signature of at least partially resolved conflict, as the breakdown of correlation between male and female transcription allows for sex-specific fitness optima (Connallon & Knowles, 2005; Mank, 2009). In contrast, it is more difficult to detect interlocus conflict (conflict between alleles at different loci, where if one allele is beneficial to males and detrimental to females, the other allele displays a reverse fitness effect) based on molecular data alone, as there is no expected resolution through the breakdown in intersexual correlation. However, imprinting and parent-of-origin expression may indicate sites linked in interlocus conflict (Haig, 2009; Gregg *et al.*, 2010a,b), as well as allelic imbalance in expression, although with caveats. The genetic mechanisms by which sexual conflict can be resolved are discussed in further detail below.

Types of genes influenced by sex-specific selection

Initial studies examining the influence of sex-specific selection at the genetic level focused on known reproductive genes (Swanson & Vacquier, 2002), which evolve rapidly across a diverse range of taxa. Many of the genetic changes have been shown to be adaptive (Swanson *et al.*, 2001), with one of the best examples perhaps being the evolution of male accessory gland proteins (Acps) related to sperm competition in *Drosophila*. These proteins are present in seminal fluid, and act to increase male mating success by promoting ovulation, reducing female receptivity to remating and promoting sperm storage (Wolfner, 2002). Acp loci in *Drosophila* are 50% more divergent than nonreproductive proteins, with many exhibiting rapid turnover between species and high rates of functional change indicative of positive selection (Swanson *et al.*, 2001; Begun & Lindfors, 2005). This is a broad pattern, as similar signatures of adaptive change have been found in a wider group of reproduction-related genes, including seminal fluid proteins, in *Drosophila* (Haerty *et al.*, 2007), primates (Clark & Swanson, 2005) and rodents (Turner *et al.*, 2008), as well as gamete recognition proteins in marine invertebrates (Metz & Palumbi, 1996; Palumbi, 1999). The rapid evolutionary rates of male reproduction-related genes suggest post-copulatory selection is important in driving adaptive divergence. Indeed, sperm competitive ability in *Drosophila* has been directly linked to polymorphism in certain male reproductive genes (Fiumera *et al.*, 2005). Presumably, increasing competition between males for mating increases the intensity of sperm competition, thereby resulting in higher rates of evolution for male reproduction-related genes (Walters & Harrison, 2011). Reproductive character displacement may also contribute to the rapid divergence of male reproductive genes (Matute, 2010) as well as frequency-dependent selection (Clark *et al.*, 1999).

More recently, female reproductive proteins in *Drosophila* have been shown to undergo similarly high rates of functional change (Swanson *et al.*, 2004; Panhuis & Swanson, 2006), possibly due to the conflict between the sexes over polyspermy. Sperm competition generates selection on males to increase the speed of fertilization. Increased sperm fertilization rate can result in a cost to females through the elevated possibility of multiple sperm penetrating the ovum, which generally results in lethality. The major cost to females of egg inviability generates female-specific selection to slow down fertilization. ZP3, a protein found within the egg coat, is responsible for binding to sperm and thus facilitating fertilization, and there is evidence in mammals of positive selection within the gene region responsible for sperm binding (Swanson *et al.*, 2001). Similar results have been shown in birds (Calkins *et al.*, 2007; Berlin *et al.*, 2008); however, there is some debate as to whether selection against polyspermy is responsible for driving this adaptive change, as birds may be more tolerant of polyspermy than other animals (Wishart, 1987; Birkhead & Fletcher, 1998; Tarin & Caro, 2000; Stepinska & Bakst, 2007). Instead, cryptic female choice for specific male sperm type may be responsible for the high rates of functional change seen at sperm binding regions of some egg coat proteins (Calkins *et al.*, 2007; Berlin *et al.*, 2008).

Male and female fitness is not solely reliant on reproductive proteins, and secondary sexual characters can play an important role in mate acquisition. As expected, there is evidence of adaptive change as a result of sex-specific selection acting on these characters. Male plumage colour in galliforms is highly diverse and shown to be involved in female mate choice and between male-to-male competition. The MC1R locus, which contributes to plumage pigmentation, has been shown to undergo high rates of functional change (Nadeau *et al.*, 2007). Additionally, this rate correlates with the degree of sexual dichromatism exhibited across galliform species, demonstrating that the MC1R locus is a target for sex-specific selection acting on plumage coloration. However, many somatic dimorphisms are complex aggregates of many loci, and this complexity requires fundamentally different approaches, explained below.

Sex-biased expression

More complex sexual phenotypes composed of dozens to hundreds of loci can be examined at the genomic scale with transcriptome data. The majority of polygenic sexual dimorphisms result from differences in gene expression between males and females, and this sex-biased expression is the product of the breakdown in intersexual correlation in expression, and therefore represent loci where intralocus conflict has been at least partially resolved (Connallon & Knowles, 2005). Additionally, rates of evolution of sex-biased genes, measured in aggregate, may offer insight into the relative strength of male- vs. female-specific selection.

In adults, the differences between males and females in expression are prevalent in the transcriptomes of many animals, including *Drosophila* (Jin *et al.*, 2001; Ranz *et al.*, 2003), mouse (Yang *et al.*, 2006), *Anopheles* mosquitoes (Marinotti *et al.*, 2006), birds (Mank *et al.*, 2008a; Naurin *et al.*, 2011), *C. elegans* (Cutter & Ward, 2005) and ants (Ometto *et al.*, 2010). Some of the pattern of sex-bias is condition-dependent (Wyman *et al.*, 2010), consistent with predictions about some types of sexually selected traits. Additionally, many of the differences in gene expression between the sexes are thought to arise from androgen- or oestrogen-mediated regulation (Zauner *et al.*, 2003), and changes in regulation have produced large variation in the proportion of the transcriptome showing sex-bias among species (Zhang *et al.*, 2007), as well as variation in sex-bias among

populations within species (Muller *et al.*, 2011; Moghadam *et al.*, 2012).

Characterizing the degree of sex-biased gene expression at the tissue level can provide insight into the phenotypic traits subject to the greatest degree of sex-specific selection. As expected from studies on reproductive proteins, sex-biased expression is most evident in gonad transcriptomes (Parisi *et al.*, 2004; Rinn *et al.*, 2004; Mank *et al.*, 2008a). However, sex-biased gene expression is observed across a large majority of somatic tissues in many animals (Yang *et al.*, 2006; Mank *et al.*, 2008a), particularly in the liver. Evidence from mammals and birds suggests that the degree of sex-biased expression is lowest in the brain (Yang *et al.*, 2006; Mank *et al.*, 2008a; Reinius *et al.*, 2008). However, these studies tend to examine the brain as a whole and thus, a more detailed examination of specific areas of the brain may reveal a more obvious pattern of sex-bias (Naurin *et al.*, 2011).

It is possible to estimate the relative strength of sex-specific selection at the molecular level by estimating rates of divergence for sex-biased genes. Numerous studies have found evidence of accelerated rates of evolution, particularly due to positive selection, in adult male-biased genes across a range of species (Good & Nachman, 2005; Khaitovich *et al.*, 2005; Ellegren & Parsch, 2013). In addition to divergence in coding sequence, expression divergence is more pronounced for male-biased genes in *Drosophila* than female-biased genes (Meiklejohn *et al.*, 2003; Llopart, 2012). However, there are exceptions to the higher rates of evolution seen in male-biased compared to female-biased genes. Some studies have found either no difference in divergence patterns between the sexes (Metta *et al.*, 2006), potentially linked to a reduction in sexual selection, or the reverse pattern (Mank *et al.*, 2010a). The latter study highlights the shifting nature of sex-specific selection throughout development, as female-biased genes were found to have higher rates of adaptive change when measured in the gonad during embryonic development compared to rates in adults.

Genetic architecture, the number and type of loci underlying a given trait, plays an important role in determining the genetic and phenotypic outcome of sex-specific selection and evolution of sex-biased expression. The involvement of a single gene in the development of multiple traits, pleiotropy, determines the extent to which intralocus sexual conflict can be resolved and thus, the ability of sex-specific selection to facilitate evolutionary change (Harano *et al.*, 2010). Less pleiotropic genes, measured as a function of tissue specificity, exhibit sex-biased gene expression and faster rates of evolutionary change compared to pleiotropic genes (Mank *et al.*, 2008b; Meisel, 2011). This may suggest that pleiotropy hinders the breakdown in intersexual correlation in expression, and therefore the capacity of the genome to respond to sex-specific selection, reinforcing the widely acknowledged role of pleiotropy as an evolutionary constraint (Fisher, 1930; Orr, 2000; Snell-Rood *et al.*, 2010). Thus, the effect of selection on gene expression will not be the same between genes under different architectural constraints. Interestingly, female-biased genes display greater pleiotropic effects than male-biased genes, potentially contributing to the different rates of evolution between the two classes of genes (Assis *et al.*, 2012).

Surveys of species and population variation in expression, such as those cited above, provide a long-term evolutionary view of sex-specific selection and sexual conflict. Using sex-bias expression data to infer the targets and strength of sex-specific selection relies on a number of assumptions. First, the relationship assumes that sex-biased genes encode sex-specific phenotypes and often have sex-specific fitness effects. It is difficult to test this assumption with species and population transcriptome data alone, as these data cannot directly connect large aggregates of genes with differential expression to concrete phenotypes, although there is some empirical support (Connallon & Clark, 2011a). Second, the relationship assumes that sexual conflict can be at least partially resolved via the breakdown of intersexual expression correlation; therefore, pleiotropic constraints may mask loci that are subject to sexual conflict.

Studies that combine molecular and phenotypic approaches are just now appearing, to date only two have been published to our knowledge (Innocenti & Morrow, 2010; Moghadam *et al.*, 2012) and these have the added power of being able to measure sex-specific fitness effects and therefore estimate sex-specific phenotypic optima. In one case the predicted relationship between sex-biased expression and sex-specific fitness was not evident, possibly due to insufficient variation in sex-biased regulatory variation within study populations. If this is a general trend, it may suggest that short-term sex-specific regulatory changes are rare (Innocenti & Morrow, 2010). However, the other study was somewhat contradictory, and showed that changes in sex-specific selection over short evolutionary histories cause substantial changes in sex-biased expression, and that this has sex-specific fitness consequences (Moghadam *et al.*, 2012). Further experimental evolution studies are planned or in progress and will no doubt provide exciting developments to this debate.

Sex-limited expression and imprinting

The negative fitness consequences of sexual conflict (Morrow *et al.*, 2008) can be avoided if antagonistic genes are sex-limited, with expression completely restricted to either males or females. Sex-limited expression, where a gene is expressed in only one sex, is somewhat different from sex-biased expression, where a gene is expressed in both females and males, but at

different levels. True sex-limited expression is relatively rare for genes not linked to the sex-limited Y or W chromosomes. This may be because selection against expression in one sex decreases as expression level wanes for most loci, creating a saturation point at which sex-specific selection is not strong enough to further reduce expression. However, the point at which sex-biased expression becomes functionally sex-limited is somewhat arguable, as very low levels of expression in one sex are likely to have little phenotypic effect. This may suggest that most very sex-biased genes are functionally sex-limited.

Sex-limited genes avoid fitness penalties in the sex lacking expression (Rice, 1984). Genomic imprinting is one potential mechanism to achieve sex-limited expression without the need to breakdown intersexual regulatory mechanisms. For an imprinted allele, expression depends upon the parent of origin, and studies have suggested that imprinting can both resolve intralocus conflict and exacerbate interlocus conflict (Day & Bonduriansky, 2004; Swaney *et al.*, 2007; Van Cleve & Feldman, 2007; Hager *et al.*, 2008; Haig, 2009).

The resolution of intralocus conflict by the evolution of imprinting has been modelled under a wide range of selection and dominance parameters (Day & Bonduriansky, 2004; Van Cleve & Feldman, 2007; Patten & Haig, 2008). Negative mother–son fitness correlations can result in selection for invading modifier loci that silence sexually antagonistic maternally inherited alleles in males. In support of these predictions, sex-specific differences in imprinting effects of loci contributing to body size have been shown in mice (Hager *et al.*, 2008).

For the majority of imprinted loci however, phenotypic effect is poorly understood, and it is possible that modelling sexual antagonism at only one locus is unrealistic. Instead, sexual conflict between interacting genes may drive the evolution of imprinting. The insulin-like growth factor 2 (Igf2) and insulin-like growth factor 2 receptor (Igf2r) are imprinted in many mammals and heavily involved in the regulation of growth (Day & Bonduriansky, 2004). This imprinting pattern of Igf2 and Igf2r is thought to be the product of conflict over maternal and paternal reproductive interests as they play out through the offspring. The expression of Igf2 is paternally inherited and associated with increased growth rates, and a mutation in this gene causes the poor growth associated with Silver–Russell syndrome (Obermann *et al.*, 2004). The paternal allele of Igf2 works in concert with Igf2r, which is maternally imprinted and is responsible for degrading Igf2 (Wilkins & Haig, 2003) and limiting nutrient extraction from the mother. In addition to Igf2 and Igf2r, interlocus conflict has been implicated in the evolution of imprinting for other developmental traits such as suckling intensity and age of sexual maturation (Haig, 2009) and presumably as the phenotype of more imprinted genes is examined, the exact role of interlocus conflict will become apparent.

Recent genome-wide assessments of genomic imprinting based on parent-of-origin expression imbalanced in the offspring (Gregg *et al.*, 2010a,b) indicated that a large proportion of the genome exhibits some form of imprinting, and that partial imprinting may be common for many genes, or even parts of genes. However, for many loci, the phenotypic and evolutionary effects of imprinting are unclear, and there are recent concerns regarding the accurate identification of imprinted loci (DeVeale *et al.*, 2012). Further work characterizing the fitness consequences of imprinting and ascertaining whether they exhibit a sexually antagonistic pattern is key for cementing the relationship between imprinting and sexual conflict.

Despite the controversy of genome-wide evidence of imprinting, evidence from the studies referenced above suggests that imprinted loci are beacons of sexual conflict, and careful examination of numbers and rates of evolution of imprinted genes will provide insight into the strength of sexual conflict arising from sex-specific selection. However, as both inter- and intralocus sexual conflict increase, the number of imprinted loci is predicted to also increase, and further work on a larger number of sexually antagonistic imprinted loci over a range of mating systems can be used to explore this. Additionally, there is very little information about how parent-of-origin imprinting varies within and among populations, and how this regulatory mechanism responds to sex-specific selection in an experimental evolution framework.

Post-transcriptional regulation

In addition to transcriptional differences between the sexes, post-transcriptional mechanisms also differ. Alternative splicing is a gene regulatory mechanism that produces multiple distinct transcripts from one locus, thereby increasing transcriptomic complexity without unduly adding to genome size. Sex-specific splicing is a key component of sex-specific phenotypes, such as sex determination in *Drosophila* (Telonis-Scott *et al.*, 2009), and therefore, may be a general mechanism to resolve sexual conflict over gene function. Comprehensive attempts have recently been made to quantify the exact extent of sex-specific splicing throughout the genome and have shown significant conservation of sex-specific splice variants (Blekhman *et al.*, 2009; Telonis-Scott *et al.*, 2009; Prince *et al.*, 2010).

It is easy to imagine the potential for sex-specific splicing, both in response to sex-specific selection and a route to resolve sexual conflict. Studies of alternative splicing evolution have suggested that splice variants are adaptive, and that they allow for increasing phenotypic complexity without the need to increase genome size (Barbosa-Morais *et al.*, 2012; Merkin *et al.*, 2012).

However, aside from the role of sex-specific splicing in sex determination in insects, little is known at this point about the exact role of alternate splicing in sexually dimorphic phenotypes, or how splice variants evolve in response to sex-specific selection.

In addition to alternative splicing, sex-specific post-transcriptional regulation can be achieved via small noncoding mRNAs. Small noncoding mRNAs are thought to play an important role in development (Stefani & Slack, 2008) via both transcriptional and post-transcriptional regulation (Engels & Hutvagner, 2006). In a recent study of ten candidate small noncoding mRNAs in two *Drosophila* species, three were found to exhibit a highly male-biased expression pattern (Jiang *et al.*, 2011). The extent of sex-bias was large, with one candidate expressed 40 times more in males than females. It has yet to be verified how these differences translate to the level of the phenotype, but a role in sex-specific RNA regulation seems possible. Further work on this area is needed to explore how these small noncoding RNAs regulate gene networks and phenotypic change, before inferences about the strength of sex-specific selection can be made.

Gene duplication

In some cases, sexual conflict at a specific locus can be resolved via gene duplication. The duplication of a sexually antagonistic locus generates paralogs, which provide the potential for sex-specific neo- or subfunctionalization. Sex-specific functionalization can act on both paralogs to theoretically generate male and female beneficial duplicates, which would eventually exhibit sex-biased or sex-limited expression (Connallon & Clark, 2011b; Gallach & Betran, 2011). Alternatively, one paralog may maintain its original function, particularly when it is pleiotropic, and the resolution of sexual conflict can be achieved through sex-biased expression of the other copy.

Evidence indicates that selection to resolve sexual conflict acts to retain duplicates within the genome, and the larger the degree of conflict, the stronger the selection to resolve this conflict via the retention of a resolving duplicate. Nuclear-encoded mitochondrial genes are thought to be under sexually antagonistic selection for the rate of energy production. High rates increase the fitness of sperm but also result in a higher mutation rate, which is disadvantageous for female reproductive tissues. It has been shown among relocated duplicate genes of this class, that a large number exhibit testis-specific expression. This duplication and subsequent sex-limited expression decouples the sexes divergent fitness optima and allows resolution of the conflict over energy production (Gallach *et al.*, 2010; Gallach & Betran, 2011). Recent evidence suggests that duplications are associated primarily with the evolution of male-biased not female-biased expression (Wyman *et al.*, 2012); however, further experimental studies across organisms with different mating systems are necessary to determine whether this pattern is due to variation in the strength of sex-specific selection, or simply a function of the higher rate of male meiosis, and therefore origin of gene duplicates, associated with continuous sperm production.

Beyond the autosomes

Sex-biased or sex-limited inheritance, exhibited by the homogametic sex chromosomes, the heterogametic sex chromosomes and extra-nuclear genomes carried by mitochondria and chloroplasts, influences the effect of sex-specific selection. Contrasting the evolutionary signatures on sex chromosomes in particular with the autosomes is increasingly useful for uncovering the strength of sex-specific selection (Rice, 1984; Dean *et al.*, 2012) (Fig. 2).

Sexual conflict and the homogametic X and Z chromosomes are linked in several ways. First, the unique sex-specific selection pressures foster the nonrandom patterns of gene traffic to and from sex chromosomes for loci with sex-specific benefits (Zhang *et al.*, 2010a, b). In addition to gene movement, sex-biased inheritance of the homogametic sex chromosomes suggests that they play a disproportionately large role in the evolution of sexual dimorphism via intralocus sexual conflict (Rice, 1984). The theoretical prediction that the X chromosome is both feminized and demasculinized, and the Z chromosome is masculinized and defeminized, is supported by numerous genomic analyses in animals (reviewed in Mank, 2009; Wright *et al.*, 2012), as well as plants (Spigler *et al.*, 2011). Finally, sexual conflict may also foster the origin and/or turnover of sex chromosomes. In many clades, there is a high rate of sex chromosome turnover, which has been linked to sexual conflict (van Doorn & Kirkpatrick, 2007, 2010). The theoretical link has been supported by direct empirical evidence in cichlids, with the invasion of a novel sex determining locus (Roberts *et al.*, 2009) as well as sticklebacks, with the origin of a neosex chromosome that contains loci for male courtship traits (Kitano *et al.*, 2009; although see Natri *et al.*, 2013). This suggests a complex relationship between conflict and sex chromosomes, involving gene movement, the origin of new sex chromosomes, or gene expression changes on existing sex chromosomes.

Mating systems may also influence the degeneration rate of the heterogametic Y and W chromosomes. For sex chromosomes that evolve from existing autosomes, linkage between a sex determining locus and a nearby locus with sex-specific effects will result in selection to suppress recombination between the heterogametic and homogametic sex chromosome. As recombination suppression spreads across the heterogametic sex chromosomes, the Y and W chromosome–coding content

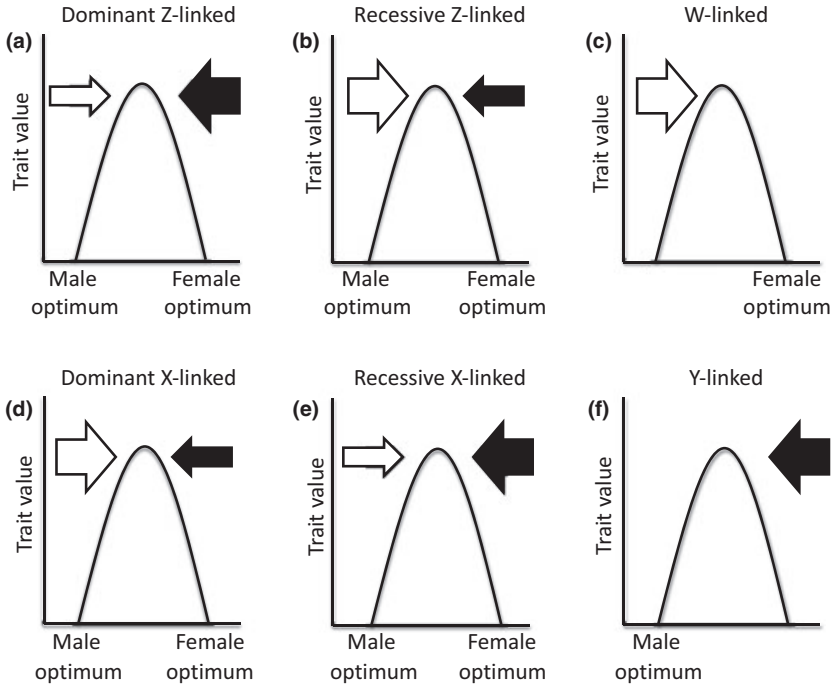

Fig. 2 Relative strength of sex-specific selection acting on different sex chromosomes. Relative sex-specific selection is shown by arrow size, white arrows represent female-specific selection, black arrows represent selection towards male-specific optima. For dominant Z-linked alleles (panel a), male-specific selection is relatively stronger due to the fact that the Z is present more often in males than females. Recessive Z-linked alleles (panel b) are more often exposed in females to selection due to female hemizygosity; therefore, female-specific selection is relatively stronger. W-linked genes (panel c) are only selected for female-specific effects. Dominant X-linked alleles (panel d) are more often selected for female-specific effect because the X is more often present in females, whereas recessive X-linked alleles (panel e) are more often exposed in males due to male hemizygosity. Y-linked genes (panel f) are only selected for male-specific effects.

degrades by neutral processes (Charlesworth, 2008). Additionally, the lowered effective population size of the heterogametic sex chromosome compared to the homogametic sex chromosome and the autosomes promotes degradation by background selection and hitchhiking. For relatively stable genes with low levels of gene translocation, the likelihood that a new sex determining gene will be located proximate to a sexually antagonistic locus may be roughly predicted by the proportion of loci in the genome that produce a sex-specific benefit. As mating system is one of the determinants controlling the proportion of the genome subject to sexually antagonistic selection, we might expect mating systems with high levels of conflict to result in faster spread of recombination suppression and therefore heterogametic sex chromosomes degeneration (Charlesworth & Mank, 2010).

Similar to the sex chromosomes, extra-nuclear genomes show sex-biased transmission patterns that influence the pattern of sex-specific selection. Mitochondria, although present and essential to both males and females, are only transmitted through the matriline and therefore, mitochondrial genomes are predicted to accumulate alleles beneficial to females but detrimental to males (Cosmides & Tooby, 1981; Frank & Hurst, 1996). A recent study supports this prediction, showing that the maternal transmission pattern results in a sieve that allows deleterious effects to persist if they are limited to males (Innocenti *et al.*, 2011). This sex-specific sieve may act on other uniparentally inherited organelles, such as chloroplasts, thought the sieve effect likely varies greatly between dioecious and monecious species.

Neutral sex-specific patterns

Identifying signatures of genetic drift at the genetic level provides insight into the strength of neutral evolution. Mating systems define the variance in reproductive success between the sexes and thus the transmission of genetic material to subsequent generations, and the effect of mating system on the direction and rate of transmissions differs among regions of the genome. Although males and females share the autosomal portion of their genome equally, there is a pronounced asymmetry in the inheritance of the X chromosome (more often present in females), the Y chromosome (male-limited), the Z chromosome (more often present

in males) and the W chromosome (female-limited). These inheritance patterns mean that different regions of the genome differ from each other in both their absolute effective populations size (N_E), as well as their response in N_E to differences in male and female mating success (Fig. 3). Drift for autosomal loci is minimized in monogamous species compared to other mating systems, and deviations from monogamy will lead to elevated rates of genetic drift and decrease the efficacy of selection across the genome as a whole (Hartl & Clark, 2006). However, the relationship between drift and selection plays out differently on the sex chromosomes. The effective population size of both the X and Z chromosomes (N_{EX} and N_{EZ}) = ¾ that of the autosomes (N_{EA}) in monogamous mating systems, creating a potential for increased genetic drift to act on homogametic sex chromosomes (Charlesworth *et al.*, 1993; Vicoso & Charlesworth, 2009). Increased variance in male reproductive success associated with most forms of sexual selection increases N_{EX} and decreases N_{EZ} relative to N_{EA} (Fig. 3b–c); therefore, sexual selection on males is predicted to increase rates of neutral evolution for Z chromosomes more than X chromosomes, termed Faster-Z and Faster-X evolution (Mank *et al.*, 2010c). This is supported by some evidence from birds (Mank *et al.*, 2007, 2010b), mammals (Lau *et al.*, 2009) and *Drosophila* (Connallon, 2007; Singh *et al.*, 2007; Baines *et al.*, 2008); however, other data are discordant with the role of mating sytem and sex chromosome evolution (Haddrill *et al.*, 2010).

These predictions for Faster-X and Faster-Z evolution are slightly altered under female promiscuity, primarily because the degree of variation in female reproductive success seen in polyandry is predicted to be less than the variation in male reproductive success seen in polygyny (Liker *et al.*, 2001). Additionally, although female promiscuity can increase male-specific selection, it can also erode variance in male mating success (Collet *et al.*, 2012).

The exact relationship between mating system and the strength of Faster-X or Faster-Z evolution is complicated by differences in the rate of recombination on the sex chromosomes, particularly the absence of recombination in *Drosophila* males, which raises the N_{EX} to near N_{EA}, thus reducing the strength of drift (Connallon, 2007; Vicoso & Charlesworth, 2009). It is important to make clear that variance in male and female reproductive success affects all portions of the genome, however contrasting the rates of neutral evolution for different portions makes it possible to quantify the effect of mating system on genetic drift.

Mating system also affects the rates of evolution for the heterogametic W and Y chromosomes. N_{EW} and N_{EY} are both equal to 1/4 N_{EA} under monogamy, but sexual selection acting on males will decrease N_{EY} and increase N_{EW} compared to N_{EA}. Although there are a host of factors affecting the evolution of heterogametic sex chromosomes, (reviewed in Charlesworth, 2008), the difference in N_{EY} and N_{EW} under sexual selection may accelerate the rate of degeneration of Y chromosomes compared to W chromosomes (Bachtrog *et al.*, 2011). In addition to decay, changes in mating system may also result in accelerated rates of change and structural rearrangement in heterogametic sex chromosomes. Consistent with this, the Y chromosome is highly conserved between human, gorilla (Goto *et al.*, 2009) and rhesus macaque (Hughes *et al.*, 2012) but has shown rapid change in the intermediary chimpanzee lineage (Wilson & Makova, 2009; Hughes *et al.*, 2010), and potential explanations for this rest on the

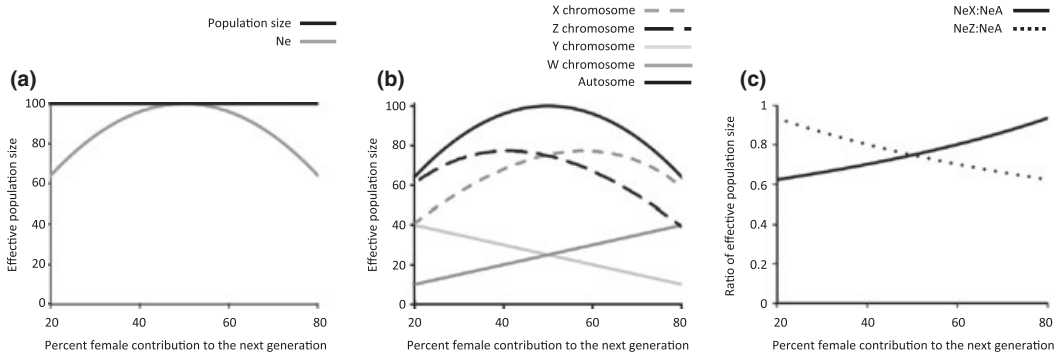

Fig. 3 Mating systems and effective population size (N_E). Increasing differences between male and female reproductive success reduces N_E (panel a), despite a constant overall population size. This difference between the sexes in reproductive success influences the N_E of different portions of the genome in different ways (panel b). For autosomal genes, the largest effective population size (N_{EA}) is seen when the variance in male and female reproductive success is equal; however, N_{EX} and N_{EW} increase with a greater proportion of females than males contributing to the next generation. The opposite is seen for N_{EZ} and N_{EY}, which are maximized when there are more males than females in the reproductive pool. This difference in the effect of mating system on the effective population size of different chromosomes makes contrasts between sex chromosomes and autosomes revealing (panel c).

promiscuous mating system observed in chimps. However, although the primate example fits the theoretical predictions, it is also anecdotal, and it is not yet possible to robustly test the relationship between mating system and Y chromosome degeneration. Additionally, the exact relationship between heterogametic sex chromosomes and mating system is complicated by several factors, including differential rates of gene acquisition (Koerich *et al.*, 2008) and intrachromosomal recombination (Lange *et al.*, 2009). Additionally, in some cases, recombination between the homogametic and heterogametic sex chromosome occurs in sex-reversed individuals (Stöck *et al.*, 2011), which may act to prevent Y chromosome decay (Perrin, 2009).

Conclusions and synthesis

Mating system can have profound effects on both neutral and adaptive genome evolution, and can also foster change in gene sequence, expression and post-translational modification for a large proportion of loci. At this point, the theoretical predictions linking mating system to genome evolution have been supported by many anecdotal species-specific studies. By utilizing recent advances in molecular genomics, many of these studies provide a long-term overview of the nature of sex-specific selection acting across thousands of genes over millions of years. However, this approach relies on a number of assumptions, namely that conflict is resolvable, which may not be the case for many loci. We also lack detailed information on the sex-specific fitness effects of certain genomic mechanisms thought to be shaped by sex-specific selection. Phenotypic studies can provide much needed information on sex-specific fitness optima and the consequences of deviating from these optima. We therefore emphasize the need for a combined approach, incorporating phenotypic studies with genomic data in order to examine the nature of sexual conflict currently acting on the genome and the fitness consequences of the genomic mechanisms thought to resolve sexual conflict.

Furthermore, recent developments in next generation sequencing methods now facilitate studies utilizing the variation in mating system across a wide range of species. Currently, we lack a holistic, robust and quantitative understanding of the relationship between mating system and genome evolution and a cohesive picture of how the myriad of sex-specific forces integrate with one another. This robust, quantitative and cohesive understanding requires comprehensive studies in clades with a range of mating systems. Although model organisms can be useful, the most relevant behavioural ecologies for this type of analysis are often associated with nonmodel organisms. A comparative approach across a range of mating systems is much needed to shed further light on the nature of sex-specific selection acting on the genome.

Acknowledgments

We thank J. Perry, M. Pointer and H. Moghadam for helpful comments and discussions. AW is funded by a NERC Post-Graduate Studentship, and additional support comes from the BBSRC and the ERC 7th Framework Programme (grant agreement no 260233).

References

Andersson, M. 1994. *Sexual Selection*. Princeton University Press, Princeton, NJ.

Arnqvist, G. 2011. Assortative mating by fitness and sexually antagonistic genetic variation. *Evolution* **65**: 2111–2116.

Arnqvist, G. & Rowe, L. 2002. Antagonistic coevolution between the sexes in a group of insects. *Nature* **415**: 787–789.

Arnqvist, G. & Rowe, L. 2005. *Sexual Conflict*. Princeton University Press, Princeton, NJ.

Assis, R., Zhou, Q. & Bachtrog, D. 2012. Sex-biased transcriptome evolution in Drosophila. *Genome Biol. Evol.* **4**: 1189–1200.

Bachtrog, D., Kirkpatrick, M., Mank, J.E., McDaniel, S.F., Pires, J.C., Rice, W. *et al.* 2011. Are all sex chromosomes created equal? *Trends Genet.* **27**: 350–357.

Baines, J.F., Sawyer, S.A., Hartl, D.L. & Parsch, J. 2008. Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in *Drosophila*. *Mol. Biol. Evol.* **25**: 1639–1650.

Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J. *et al.* 2012. The evolutionary landscape of alternative splicing in vertebrate species. *Science* **338**: 1587–1593.

Begun, D.J. & Lindfors, H.A. 2005. Rapid evolution of genomic Acp complement in the melanogaster subgroup of *Drosophila*. *Mol. Biol. Evol.* **22**: 2010–2021.

Berlin, S., Qu, L.J. & Ellegren, H. 2008. Adaptive evolution of gamete-recognition proteins in birds. *J. Mol. Evol.* **67**: 488–496.

Birkhead, T.R. & Fletcher, F. 1998. Sperm transport in the reproductive tract of female zebra finches (*Taeniopygia guttata*). *J. Reprod. Fertil.* **114**: 141–145.

Blekhman, R., Marioni, J.C., Zumbo, P., Stephens, M. & Gilad, Y. 2009. Sex-specific and lineage-specific alternative splicing in primates. *Genome Res.* **20**: 180–189.

Calkins, J.D., El-Hinn, D. & Swanson, W.J. 2007. Adaptive evolution in an avian reproductive protein: ZP3. *J. Mol. Evol.* **65**: 555–563.

Chapman, T., Arnqvist, G., Bangham, J. & Rowe, L. 2003. Sexual conflict. *Trends Ecol. Evol.* **9**: 289–293.

Charlesworth, D. 2008. Sex chromosome origins and evolution. In: *Evolutionary Genomics and Proteomics*. (M. Pagel & A. Pamiankowski, eds), pp. 207–239. Sinauer Associates Inc., Sunderland, MA.

Charlesworth, D. & Mank, J.E. 2010. The birds and the bees and the flowers and the trees: lessons from genetic mapping of sex determination in plants and animals. *Genetics* **186**: 9–31.

Charlesworth, B., Coyne, J.A. & Orr, H.A. 1993. Meiotic drive and unisexual hybrid sterility - a comment. *Genetics* **133**: 421–424.

Chippindale, A.K., Gibson, J.R. & Rice, W.R. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **98**: 1671–1675.

Clark, N.L. & Swanson, W.J. 2005. Pervasive adaptive evolution in primate seminal proteins. *PLoS Genet.* **1**: 335–342.

Clark, A.G., Begun, D.J. & Prout, T. 1999. Female x male interactions in *Drosophila* sperm competition. *Science* **283**: 217–220.

Collet, J., Richardson, D.S., Worley, K. & Pizzari, T. 2012. Sexual selection and the differential effect of polyandry. *Proc. Natl. Acad. Sci. USA* **109**: 8641–8645.

Connallon, T. 2007. Adaptive protein evolution of X-linked and autosomal genes in *Drosophila*: implications for faster-X hypotheses. *Mol. Biol. Evol.* **24**: 2566–2572.

Connallon, T. & Clark, A.G. 2011a. Association between sex-biased gene expression and mutations with sex-specific phenotypic consequences in Drosophila. *Genome Biol. Evol.* **3**: 151–155.

Connallon, T. & Clark, A.G. 2011b. The resolution of sexual antagonism by gene duplication. *Genetics* **187**: 919–937.

Connallon, T. & Knowles, L.L. 2005. Intergenomic conflict revealed by patterns of sex-biased gene expression. *Trends Genet.* **21**: 495–499.

Connallon, T., Cox, R.M. & Calsbeek, R. 2010. Fitness consequences of sex-specific selection. *Evolution* **64**: 1671–1682.

Corl, A. & Ellegren, H. 2012. The genomic signature of sexual selection in the genetic diversity of the sex chromosomes and autosomes. *Evolution* **66**: 2138–2149.

Cosmides, L.M. & Tooby, J. 1981. Cytoplasmic inheritance and intragenomic conflict. *J. Theor. Biol.* **89**: 83–129.

Cutter, A.D. & Ward, S. 2005. Sexual and temporal dynamics of molecular evolution in *C. elegans* development. *Mol. Biol. Evol.* **22**: 178–188.

Day, T. & Bonduriansky, R. 2004. Intralocus sexual conflict can drive the evolution of genomic imprinting. *Genetics* **167**: 1537–1546.

Dean, R., Perry, J.C., Pizzari, T., Mank, J.E. & Wigby, S. 2012. Experimental evolution of a novel sexually antagonistic allele. *PLoS Genet.* **8**: e1002917.

DeVeale, B., van der Kooy, D. & Babak, T. 2012. Critical evaluation of imprinted gene expression by RNA-Seq: a new perspective. *PLoS Genet.* **8**: e1002600.

van Doorn, G.S. & Kirkpatrick, M. 2007. Turnover of sex chromosomes induced by sexual conflict. *Nature* **449**: 909–912.

van Doorn, G.S. & Kirkpatrick, M. 2010. Transitions between male and female heterogamety caused by sex-antagonistic selection. *Genetics* **186**: 629–645.

Ellegren, H. & Parsch, J. 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nat. Rev. Genet.* **14**: 83–87.

Engels, B.M. & Hutvagner, G. 2006. Principles and effects of microRNA-mediated post-transcriptional gene regulation. *Oncogene* **25**: 6163–6169.

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

Fiumera, A.C., Dumont, B.L. & Clark, A.G. 2005. Sperm competitive ability in *Drosophila melanogaster* associated with variation in male reproductive proteins. *Genetics* **169**: 243–257.

Foerster, K., Coulson, T., Sheldon, B.C., Pemberton, J.M., Clutton-Brock, T.H. & Kruuk, L.E.B. 2007. Sexually antagonistic genetic variation for fitness in red deer. *Nature* **447**: 1107–1110.

Frank, S.A. & Hurst, L.D. 1996. Mitochondria and male disease. *Nature* **383**: 224.

Gallach, M. & Betran, E. 2011. Intralocus sexual conflict resolved through gene duplication. *Trends Ecol. Evol.* **26**: 222–228.

Gallach, M., Chandrasekaran, C. & Betrán, E. 2010. Analyses of nuclearly encoded mitochondrial genes suggest gene duplication as a mechanism for resolving intralocus sexually antagonistic conflict in *Drosophila*. *Genome Biol. Evol.* **2**: 835–850.

Good, J.M. & Nachman, M.W. 2005. Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis. *Mol. Biol. Evol.* **22**: 1044–1052.

Goto, H., Peng, L. & Makova, K. 2009. Evolution of X-degenerate Y chromosome genes in greater apes: conservation of gene content in human and gorilla, but not chimpanzee. *J. Mol. Evol.* **68**: 134–144.

Gregg, C., Zhang, J.W., Butler, J.E., Haig, D. & Dulac, C. 2010a. Sex-specific parent-of-origin allelic expression in the mouse brain. *Science* **329**: 682–685.

Gregg, C., Zhang, J.W., Weissbourd, B., Luo, S.J., Schroth, G.P., Haig, D. *et al.* 2010b. High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science* **329**: 643–648.

Haddrill, P.R., Loewe, L. & Charlesworth, B. 2010. Estimating the parameters of selection on nonsynonymous mutations in *Drosophila pseudoobscura* and *D. miranda*. *Genetics* **185**: 1381–1396.

Haerty, W., Jagadeeshan, S., Kulathinal, R.J., Wong, A., Ravi Ram, K., Sirot, L.K. *et al.* 2007. Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*. *Genetics* **177**: 1321–1335.

Hager, R., Cheverud, J.M., Leamy, L.J. & Wolf, J.B. 2008. Sex dependent imprinting effects on complex traits in mice. *BMC Evol. Biol.* **8**: 303.

Haig, D. 2009. Transfers and transitions: parent-offspring conflict, genomic imprinting, and the evolution of human life history. *Proc. Natl. Acad. Sci. USA* **107**: 1731–1735.

Harano, T., Okada, K., Nakayama, S., Miyatake, T. & Hosken, D.J. 2010. Intralocus sexual conflict unresolved by sex-limited trait expression. *Curr. Biol.* **20**: 2036–2039.

Hartl, D.L. & Clark, A.G. 2006. *Principles of Population Genetics*. Sinauer Associates Inc., Sunderland, MA.

Helfenstein, F., Tirard, C., Danchin, E. & Wagner, R.H. 2004. Low frequency of extra-pair paternity and high frequency of adoption in black-legged kittiwakes. *The Condor* **106**: 149–155.

Hughes, J.F., Skaletsky, H., Pyntikova, T., Graves, T.A., van Daalen, S.K.M., Minx, P.J. *et al.* 2010. Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. *Nature* **463**: 536–539.

Hughes, J.F., Skaletsky, H., Brown, L.G., Pyntikova, T., Graves, T., Fulton, R.S. *et al.* 2012. Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature* **483**: 82–86.

Innocenti, P. & Morrow, E.H. 2010. The sexually antagonistic genes of *Drosophila melanogaster*. *PLoS Biol.* **8**: e1000335.

Innocenti, P., Morrow, E.H. & Dowling, D.K. 2011. Experimental evidence supports a sex-specific selective sieve in mitochondrial genome evolution. *Science* **332**: 845–848.

Jiang, Z.F., Croshaw, D.A., Wang, Y., Hey, J. & Machado, C.A. 2011. Enrichment of mRNA-like noncoding RNAs in the divergence of *Drosophila* males. *Mol. Biol. Evol.* **28**: 1339–1348.

Jin, W., Riley, R.M., Wolfinger, R.D., White, K.P., Passador-Gurgel, G. & Gibson, G. 2001. The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nat. Genet.* **29**: 389–395.

Khaitovich, P., Hellmann, I., Enard, W., Nowick, K., Leinweber, M., Franz, H. *et al.* 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science* **309**: 1850–1854.

Kitano, J., Ross, J.A., Mori, S., Kume, M., Jones, F.C., Chan, Y.F. *et al.* 2009. A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**: 1079–1083.

Koerich, L.B., Wang, X.Y., Clark, A.G. & Carvalho, A.B. 2008. Low conservation of gene content in the *Drosophila* Y chromosome. *Nature* **456**: 949–951.

Kokko, H., Klug, H. & Jennions, M.D. 2012. Unifying cornerstones of sexual selection: operational sex ratio, Bateman gradient and the scope for competitive investment. *Ecol. Lett.* **15**: 1340–1351.

Lange, J., Skaletsky, H., van Daalen, S.K.M., Embry, S.L., Korver, C.M., Brown, L.G. *et al.* 2009. Isodicentric Y chromosomes and sex disorders as byproducts of homologous recombination that maintains palindromes. *Cell* **138**: 855–869.

Lau, A.N., Peng, L., Goto, H., Chemnick, L., Ryder, O.A. & Makova, K.D. 2009. Horse domestication and conservation genetics of Przewalski's horse inferred from sex chromosomal and autosomal sequences. *Mol. Biol. Evol.* **26**: 199–208.

Lifjeld, J.T., Dunn, P.O., Robertson, R.J. & Boag, P.T. 1993. Extra-pair paternity in monogamous tree swallows. *Anim. Behav.* **45**: 213–229.

Liker, A., Reynolds, J.D. & Székely, T. 2001. The evolution of egg size in socially polyandrous shorebirds. *Oikos* **95**: 3–14.

Lindenfors, P., Tullberg, B.S. & Biuw, M. 2002. Phylogenetic analyses of sexual selection and sexual size dimorphism in pinnipeds. *Behav. Ecol. Sociobiol.* **52**: 188–193.

Llopart, A. 2012. The rapid evolution of X-linked male-biased gene expression and the large-X effect in *Drosophila yakuba, D. santomea* and their hybrids. *Mol. Biol. Evol.* **29**: 3873–3886.

Long, T.A.F., Agrawal, A.F. & Rowe, L. 2012. The effect of sexual selection on offspring fitness depends on the nature of genetic variation. *Curr. Biol.* **22**: 204–208.

Magurran, A.E. & Seghers, B.H. 1994. Sexual conflict as a consequence of ecology - Evidence from guppy, *Poecilia-reticulata*, populations in Trinidad. *Proc. Biol. Sci.* **255**: 31–36.

Mank, J.E. 2009. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. *Am. Nat.* **173**: 141–150.

Mank, J.E., Axelsson, E. & Ellegren, H. 2007. Fast-X on the Z: rapid evolution of sex-linked genes in birds. *Genome Res.* **17**: 618–624.

Mank, J.E., Hultin-Rosenberg, L., Webster, M.T. & Ellegren, H. 2008a. The unique genomic properties of sex-biased genes: insights from avian microarray data. *BMC Genomics* **9**: 148.

Mank, J.E., Hultin-Rosenberg, L., Zwahlen, M. & Ellegren, H. 2008b. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. *Am. Nat.* **171**: 35–43.

Mank, J.E., Nam, K., Brunstrom, B. & Ellegren, H. 2010a. Ontogenetic complexity of sexual dimorphism and sex-specific selection. *Mol. Biol. Evol.* **27**: 1570–1578.

Mank, J.E., Nam, K. & Ellegren, H. 2010b. Faster-Z evolution is predominantly due to genetic drift. *Mol. Biol. Evol.* **27**: 661–670.

Mank, J.E., Vicoso, B., Berlin, S. & Charlesworth, B. 2010c. Effective population size and the faster-X effect: empirical results and their interpretation. *Evolution* **64**: 663–674.

Marinotti, O., Calvo, E., Nguyen, Q.K., Dissanayake, S., Ribeiro, J.M.C. & James, A.A. 2006. Genome-wide analysis of gene expression in adult *Anopheles gambiae*. *Insect Mol. Biol.* **15**: 1–12.

Matute, D.R. 2010. Reinforcement can overcome gene flow during speciation in *Drosophila*. *Curr. Biol.* **20**: 2229–2233.

Meiklejohn, C.D., Parsch, J., Ranz, J.M. & Hartl, D.L. 2003. Rapid evolution of male-biased gene expression in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **100**: 9894–9899.

Meisel, R. 2011. Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein coding sequence evolution. *Mol. Biol. Evol.* **28**: 1893–1900.

Merkin, J., Russell, C., Chen, P. & Burge, C.B. 2012. Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science* **338**: 1593–1599.

Metta, M., Gudavalli, R., Gibert, J.M. & Schlotterer, C. 2006. No accelerated rate of protein evolution in male-biased *Drosophila pseudoobscura* genes. *Genetics* **174**: 411–420.

Metz, E.C. & Palumbi, S.R. 1996. Positive selection and sequence rearrangements generate extensive polymorphisms in the gamete recognition protein binding. *Mol. Biol. Evol.* **13**: 397–406.

Moghadam, H.K., Pointer, M.A., Wright, A.E., Berlin, S. & Mank, J.E. 2012. W chromosome expression responds to female-specific selection. *Proc. Natl. Acad. Sci. USA* **109**: 8207–8211.

Morrow, E.H., Stewart, A.D. & Rice, W.R. 2008. Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. *J. Evol. Biol.* **21**: 1046–1054.

Muller, L., Hutter, S., Stamboliyska, R., Saminadin-Peter, S.S., Stephan, W. & Parsch, J. 2011. Population transcriptomics of *Drosophila melanogaster* females. *BMC Genomics* **12**: 81.

Nadeau, N.J., Burke, T. & Mundy, N.I. 2007. Evolution of an avian pigmentation gene correlates with a measure of sexual selection. *Proc. Biol. Sci.* **274**: 1807–1813.

Natri, H.M., Shikano, T. & Merila, J. 2013. Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. *Mol. Biol. Evol.* **30**: 1131–1144.

Naurin, S., Hansson, B., Hasselquist, D., Kim, Y. & Bensch, S. 2011. The sex-biased brain: sexual dimorphism in gene expression in two species of songbirds. *BMC Genomics* **12**: 37.

Obermann, C., Meyer, E., Prager, S., Tomiuk, J.R., Wollmann, H.A. & Eggermann, T. 2004. Searching for genomic variants in IGF2 and CDKN1C in Silver-Russell syndrome patients. *Mol. Genet. Metab.* **82**: 246–250.

Ometto, L., Shoemaker, D., Ross, K.G. & Keller, L. 2010. Evolution of gene expression in fire ants: the effects of developmental stage, caste, and species. *Mol. Biol. Evol.* **28**: 1381–1392.

Orr, H.A. 2000. Adaptation and the cost of complexity. *Evolution* **54**: 13–20.

Palumbi, S.R. 1999. All males are not created equal: fertility differences depend on gamete recognition polymorphisms in sea urchins. *Proc. Natl. Acad. Sci. USA* **96**: 12632–12637.

Panhuis, T.M. & Swanson, W.J. 2006. Molecular evolution and population genetic analysis of candidate female reproductive genes in *Drosophila*. *Genetics* **173**: 2039–2047.

Parisi, M., Nuttall, R., Edwards, P., Minor, J., Naiman, D., Lu, J.N. *et al.* 2004. A survey of ovary-, testis-, and soma-biased gene expression in *Drosophila melanogaster* adults. *Genome Biol.* **5**: R40.

Patten, M.M. & Haig, D. 2008. Reciprocally imprinted genes and the response to selection on one sex. *Genetics* **179**: 1389–1394.

Perrin, N. 2009. Sex reversal: a fountain of youth for sex chromosomes. *Evolution* **63**: 3043–3049.

Prince, E.G., Kirkland, D. & Demuth, J.P. 2010. Hyper-expression of the X chromosome in both sexes results in extensive female-bias of X-linked gene in the flour beetle. *Genome Biol. Evol.* **2**: 336–346.

Ranz, J.M., Castillo-Davis, C.I., Meiklejohn, C.D. & Hartl, D.L. 2003. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* **300**: 1742–1745.

Reinius, B., Saetre, P., Leonard, J.A., Blekhman, R., Merino-Martinez, R., Gilad, Y. *et al.* 2008. An evolutionarily conserved sexual signature in the primate brain. *PLoS Genet.* **4**: e1000100.

Rice, W.R. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38**: 735–742.

Rinn, J.L., Rozowsky, J.S., Laurenzi, I.J., Petersen, P.H., Zou, K., Zhong, W. *et al.* 2004. Major molecular differences between mammalian sexes are involved in drug metabolism and renal function. *Dev. Cell* **6**: 791–800.

Roberts, R.B., Ser, J.R. & Kocher, T.D. 2009. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. *Science* **326**: 998–1001.

Singh, N.D., Macpherson, J.M., Jensen, J.D. & Petrov, D.A. 2007. Similar levels of X-linked and autosomal nucleotide variation in African and non-African populations of *Drosophila melanogaster*. *BMC Evol. Biol.* **7**: 202.

Snell-Rood, E.C., Van Dyken, J.D., Cruickshank, T., Wade, M.J. & Moczek, A.P. 2010. Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. *BioEssays* **32**: 71–81.

Spigler, R.B., Lewers, K.S. & Ashman, T.-L. 2011. Genetic architecture of sexual dimorphism in a subdioecious plant with a proto-sex chromosome. *Evolution* **65**: 1114–1126.

Stefani, G. & Slack, F.J. 2008. Small non-coding RNAs in animal development. *Nat. Rev. Mol. Cell Biol.* **9**: 219–230.

Stepinska, U. & Bakst, M.R. 2007. *Fertilisation. Reproductive Biology and Phylogeny of Birds: Phylogeny, Morphology, Hormones, Fertilization*. Science Publishers, Enfield, NH.

Stöck, M., Horn, A., Grossen, C., Lindtke, D., Sermier, R., Betto-Colliard, C. *et al.* 2011. Ever-young sex chromosomes in European tree frogs. *PLoS Biol.* **9**: e1001062.

Swaney, W.T., Curley, J.P., Champagne, F.A. & Keverne, E.B. 2007. Genomic imprinting mediates sexual experience-dependent olfactory learning in male mice. *Proc. Natl. Acad. Sci. USA* **104**: 6084–6089.

Swanson, W.J. & Vacquier, V.D. 2002. The rapid evolution of reproductive proteins. *Nat. Rev. Genet.* **3**: 137–144.

Swanson, W.J., Clark, A.G., Waldrip-Dail, H.M., Wolfner, M.F. & Aquadro, C.F. 2001. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **98**: 7375–7379.

Swanson, W.J., Wong, A., Wolfner, M.F. & Aquadro, C.F. 2004. Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection. *Genetics* **168**: 1457–1465.

Tarin, J.J. & Caro, A. 2000. *Fertilization in Protozoa and Metazoan Animals*. Springer, Berlin.

Telonis-Scott, M., Kopp, A., Wayne, M.L., Nuzhdin, S.V. & McIntyre, L.M. 2009. Sex-specific splicing in *Drosophila*: widespread occurrence, tissue specificity and evolutionary conservation. *Genetics* **181**: 421–434.

Tregenza, T., Wedell, N. & Chapman, T. 2006. Introduction. Sexual conflict: a new paradigm? *Philos. Trans. R. Soc. Lond.* **361**: 229–234.

Turner, L.M., Chuong, E.B. & Hoekstra, H.E. 2008. Comparative analysis of testis protein evolution in rodents. *Genetics* **179**: 2075–2089.

Van Cleve, J. & Feldman, M.W. 2007. Sex-specific viability, sex linkage and dominance in genomic imprinting. *Genetics* **176**: 1101–1118.

Vicoso, B. & Charlesworth, B. 2009. Effective population size and the faster-X effect: an extended model. *Evolution* **63**: 2413–2426.

Walters, J.R. & Harrison, R.G. 2011. Decoupling of rapid and adaptive evolution among seminal fluid proteins in *Heliconius* butterflies with divergent mating systems. *Evolution* **65**: 2855–2871.

Wilkins, J.F. & Haig, D. 2003. What good is genomic imprinting: the function of parent-specific gene expression. *Nat. Rev. Genet.* **4**: 359–368.

Wilson, M.A. & Makova, K.D. 2009. Genomic analyses of sex chromosome evolution. *Annu. Rev. Genomics Hum. Genet.* **10**: 333–354.

Wishart, G.J. 1987. Regulation of the length of the fertile period in the domestic fowl by numbers of oviductal spermatozoa. *J. Reprod. Fertil.* **80**: 493–498.

Wolfner, M.F. 2002. The gifts that keep on giving: physiological functions and evolutionary dynamics of male seminal proteins in *Drosophila*. *Heredity* **88**: 85–93.

Wright, A.E., Moghadam, H.K. & Mank, J.E. 2012. Trade-off between selection for dosage compensation and masculinization of the avian Z chromosome. *Genetics* **192**: 1433–1445.

Wyman, M.J., Agrawal, A.F. & Rowe, L. 2010. Condition-dependence of the sexually dimorphic transcriptome in *Drosophila melanogaster*. *Evolution* **64**: 1836–1848.

Wyman, M.J., Cutter, A.D. & Rowe, L. 2012. Gene duplication in the evolution of sexual dimorphism. *Evolution* **66**: 1556–1566.

Yang, X., Schadt, E.E., Wang, S., Wang, H., Arnold, A.P., Ingram-Drake, L. *et al.* 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res.* **16**: 995–1004.

Zauner, H., Begemann, G., Mari-Beffa, M. & Meyer, A. 2003. Differential regulation of msx genes in the development of the gonopodium, an intromittent organ, and of the 'sword', a sexually selected trait of swordtail fishes (Xiphophorus). *Evol Dev* **5**: 466–477.

Zhang, Y., Sturgill, D., Parisi, M., Kumar, S. & Oliver, B. 2007. Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. *Nature* **450**: 233–238.

Zhang, Y.E., Vibranovski, M.D., Krinsky, B.H. & Long, M. 2010a. Age-dependent chromosomal distribution of male-biased genes in *Drosophila*. *Genome Res.* **20**: 1526–1533.

Zhang, Y.E., Vibranovski, M.D., Landback, P., Marais, G.A.B. & Long, M. 2010b. Chromosomal redistribution of male-biased genes in Mammalian evolution with two bursts of gene gain on the X chromosome. *PLoS Biol.* **8**: e1000494.

Glossary

Fast -X Effect (Fast-Z Effect): Higher ratio of nonsynonymous to synonymous substitutions for X- and Z-linked genes compared to autosomal genes, due to a combination of hemizygous exposure of mutations in the heterogametic sex and lowered effective population size.

Hemizygosity: Having unpaired chromosomes in a diploid cell. The heterogametic sex (males in XY systems and females in ZW species) has only one copy of each sex chromosome and is therefore hemizygous for the sex chromosomes.

Heterogametic sex chromosomes: The sex chromosomes that are sex-limited in the heterogametic sex, namely the Y in male heterogamety and the W in female heterogamety.

Homogametic sex chromosomes: The X chromosome in male heterogamety and the Z in female heterogamety.

Interlocus sexual conflict: Type of antagonism resulting when sexual conflict plays out over two or more loci. This often results in an unresolvable arms race between females and males.

Intralocus sexual conflict: Type of antagonism resulting when male and female optima differ for a single locus. Intralocus sexual conflict can be resolved via a variety of genomic mechanisms.

Monogamy: Mating system where males and females pair and are sexually exclusive.

Polyandry: Mating system where females have more than one mate at any given time, resulting in greater variance in female compared to male reproductive success.

Polygamy: General term describing a deviation from monogamy, encompassing both polyandrous and polygynous mating systems.

Polygyny: Mating system where males have more than one mate at a time, resulting in greater variance in male compared to female reproductive success.

Polyspermy: A condition resulting when more than one sperm cell fertilizes an ovum, often resulting in inviability.

Received 7 March 2013; revised 14 May 2013; accepted 15 May 2013

Trade-off Between Selection for Dosage Compensation and Masculinization on the Avian Z Chromosome

Alison E. Wright,*,†,[1] Hooman K. Moghadam,* and Judith E. Mank†
*Department of Zoology, Edward Grey Institute, University of Oxford, Oxford OX1 3PS, United Kingdom, and †Department of Genetics, Evolution and the Environment, University College London, WC1E 6BT, United Kingdom

ABSTRACT Following the suppression of recombination, gene expression levels decline on the sex-limited chromosome, and this can lead to selection for dosage compensation in the heterogametic sex to rebalance average expression from the X or Z chromosome with average autosomal expression. At the same time, due to their unequal pattern of inheritance in males and females, the sex chromosomes are subject to unbalanced sex-specific selection, which contributes to a nonrandom distribution of sex-biased genes compared to the remainder of the genome. These two forces act against each other, and the relative importance of each is currently unclear. The *Gallus gallus* Z chromosome provides a useful opportunity to study the importance and trade-offs between sex-specific selection and dosage compensation in shaping the evolution of the genome as it shows incomplete dosage compensation and is also present twice as often in males than females, and therefore predicted to be enriched for male-biased genes. Here, we refine our understanding of the evolution of the avian Z chromosome, and show that multiple strata formed across the chromosome over ~130 million years. We then use this evolutionary history to examine the relative strength of selection for sex chromosome dosage compensation *vs.* the cumulative effects of masculinizing selection on gene expression. We find that male-biased expression increases over time, indicating that selection for dosage compensation is relatively less important than masculinizing selection in shaping Z chromosome gene expression.

MALES and females of the same species are subject to distinct selective forces, often resulting in contradictory sex-specific selection pressures acting on a given locus. These selective forces have been shown to affect large portions of the genome (Connallon *et al.* 2010) and can place a significant genetic and evolutionary burden on a species (Chippindale *et al.* 2001; Arnqvist and Rowe 2005; Morrow *et al.* 2008). At the genetic level, sexually antagonistic selection pressures are thought to contribute to gene expression differences observed between the sexes in many species (Rice 1984; Connallon and Knowles 2005; Ellegren and Parsch 2007).

The effect of sex-specific selection is particularly evident on the sex chromosomes, where the unbalanced pattern of inheritance creates uneven sex-specific selection pressures (Rice 1984; Charlesworth *et al.* 1987; Connallon and Clark 2010). Consequentially, the sex chromosomes are a hotspot of intralocus sexual conflict within the genome (Mank 2009; Innocenti and Morrow 2010), and unequal sex-specific selection is predicted to have contributed to the complex pattern of sex-biased gene distribution observed across the sex chromosomes of many animals (Parisi *et al.* 2003; Khil *et al.* 2004; Vicoso and Charlesworth 2006; Sturgill *et al.* 2007; Zhang *et al.* 2010a,b; Chen *et al.* 2011). Indeed, the avian Z chromosome exhibits a pervasive pattern of male-biased gene expression, which could be interpreted in the light of masculinizing selection for dominant male-benefit genes (Rice 1984; Ellegren *et al.* 2007; Itoh *et al.* 2007).

At the same time, sex chromosomes experience changes in gene dose. Due to lack of recombination with their homologs, Y and W chromosome gene activity slowly degenerates by neutral processes (Charlesworth 1996) and the buildup of nonsynonymous and nonsense mutations as well as small indels (Zhou and Bachtrog 2012), where the rate of decay declines with the number of functionally

doi: 10.1534/genetics.112.145102
Manuscript received August 21, 2012; accepted for publication September 12, 2012
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.112.145102/-/DC1.
[1]Corresponding author: Department of Genetics, Evolution and Environment, UCL, Darwin Bldg., Gower St., London, WC1E 6BT, United Kingdom. E-mail: Alison.e.wright@ucl.ac.uk

constrained loci (Bachtrog 2008). Gene expression loss shows a range of dominance (Agrawal and Whitlock 2011), and for some genes, loss-of-function mutations on the W or Y chromosome will cause negative fitness effects (Charlesworth 1978).

The resulting loss of gene activity in the heterogametic sex and thus unbalanced gene dose with the autosomes is predicted to be costly, especially for X- or Z-linked genes that interact with autosomal genes in large protein complexes (Pessia *et al.* 2012), thereby selecting for the evolution of dosage compensation mechanisms (Ohno 1967; Charlesworth 1978). Selection for dosage compensation can be strong, leading in some organisms to complex regulatory machinery to equalize gene dose across the entirety of the chromosome (Muyle *et al.* 2012; Pessia *et al.* 2012); however, more often selection acts on dosage sensitive genes rather than entire chromosomes (Mank and Ellegren 2009a; Vicoso and Bachtrog 2011; Pessia *et al.* 2012). In *Drosophila*, males hypertranscribe the single X chromosome (Lucchesi 1973), whereas in mammals a subset of X-linked genes appears to be up-regulated in both sexes and then down-regulated in females via X chromosome inactivation to restore balanced gene dose with the autosomes (Deng *et al.* 2011; Julien *et al.* 2012; Lin *et al.* 2012; Pessia *et al.* 2012). A by-product of the evolution of dosage compensation is that gene expression levels are also balanced between the heterogametic and homogametic sexes (Bachtrog *et al.* 2011; Mank *et al.* 2011).

Birds lack complete Z chromosome dosage compensation, and the majority of Z-linked genes are expressed more in males than females due to the fact that males have two copies of Z-linked loci and females only one (Ellegren *et al.* 2007; Itoh *et al.* 2007). For dosage-sensitive genes, the predicted cost of unbalanced gene dose with the autosomes selects for the evolution of dosage compensation mechanisms, such as hypertranscription of Z-linked dosage-sensitive genes in females, to restore balanced expression levels, and consequentially equalize expression between the sexes (Ohno 1967; Pessia *et al.* 2012). Therefore, selection for dosage compensation theoretically reduces male-biased expression on the Z chromosome, and this predicted outcome is opposite to what would be expected under a model of masculinizing selection for gene expression, where sex-specific differences in Z-linked expression are exacerbated.

The evolutionary history of the Z chromosome itself can be used to distinguish signatures of masculinizing selection from selection for dosage compensation. The avian Z chromosome resulted from at least two (Handley *et al.* 2004; Suh *et al.* 2011) recombination cessation events over the last 130 million years, resulting in strata of different ages. These strata are subject to unbalanced sex-specific selection and selection for dosage compensation, for different lengths of time (Charlesworth *et al.* 2005).

Consequently, patterns of sex-biased expression between the different strata will reflect the selective regime to which the Z chromosome has been subject. Under selection for dosage compensation with no masculinization, Z-linked expression in the youngest stratum is predicted to reflect the ratio of expression from two copies in males to only one in females. Changes in gene dose show a range of expression level effects (Torres *et al.* 2008), and twofold change in human gene copy number produces an average fold change in expression of 1.5 (Pollack *et al.* 2002), consistent with the sex-biased expression ratio observed under lack of dosage compensation (Deng *et al.* 2011; Pessia *et al.* 2012), potentially due to the robust nature of gene networks (Oliver 2007). Therefore, dose effects without masculinization are expected to have, on average, 1.5-fold effects on expression. Z-linked expression is predicted to become increasingly less male biased in older regions where effective compensation mechanisms have evolved. However, the opposite pattern is expected under masculinizing selection. In this scenario, the extent of male-biased expression should increase with the age of the stratum and thus length of exposure to cumulative masculinizing selection.

Here we expand and refine our understanding of the topology of the chicken Z chromosome through the addition of several newly discovered Z-W orthologs (gametologs), and find evidence for up to four evolutionary strata. We then use this evolutionary history to test these two predictions for gene expression evolution.

Materials and Methods

Identification of misassembled W-linked genes

To expand the known coding content of the chicken W chromosome, we used RNA-seq data obtained from fertilized Red Jungle Fowl eggs (T. Pizzari, Oxford University), kept under standard incubator conditions. After collecting the left gonad at embryonic day (ed) 19, we stored the tissue in RNAlater until preparation. Previous work has shown this stage is just prior to meiotic sex chromosome inactivation (MSCI) in chickens, which affects both the W and Z chromosomes (Schoenmakers *et al.* 2009). Therefore, we used ed19 gonads to maximize the number of active W genes and minimize the risk that our Z expression estimates are confounded by MSCI. Detailed methods for the discovery of new W-linked genes are described elsewhere (Moghadam *et al.* 2012). Briefly, we extracted RNA from four samples of each sex using standard purification methods, which was then used for Illumina RNA-seq at the Wellcome Trust Centre for Human Genetics facility at Oxford University, resulting in 16 million 50-bp paired-end mappable reads on average per sample. We mapped sequences to the chicken reference genome (Washington University Genome Sequencing Center, WUGSC 2.1/galGal3) using Bowtie (v.0.12.7) (Langmead *et al.* 2009) and Tophat (v.1.1.1). To estimate transcript abundances for the Ensembl annotated genes we used Cufflinks (v.0.9.3) (Trapnell *et al.* 2010) and identified putative W-linked genes through *in silico* analysis of the gene expression profiles between

males and females. We validated that these genes had been misassembled in the current Ensembl annotation (WUGSC 2.1/galGal3) and were indeed W-linked by PCR genotyping of genomic DNA (Moghadam *et al.* 2012).

Identification and divergence of Z-W orthologs

In total, we had sequence data on 14 known and 12 newly identified W-linked genes from Ensembl, some of which represented multiple orthologs of Z-linked loci. We BLASTed (Altschul *et al.* 1997) sequences against the chicken genome sequence (WUGSC 2.1/galGal3) to identify orthologs on the Z chromosome. We took the physical positions of these orthologs from the genome assembly. There was one exception, ENSGALG00000004349, which does not have any annotated Z-linked orthologs in Ensembl.

Synonymous divergence estimates can be used to identify the number and boundaries of evolutionary strata, as they provide an appropriate measure of the length of time over which gametologs have differentiated from an original ancestral autosomal pair. We therefore calculated synonymous divergence (d_S) for each gametologous gene pair using whole coding sequences for chicken genes downloaded from Biomart with the exception of SPINZ and SPINW. The SPIN gametologs are poorly annotated; therefore, we identified and excluded incorrect exon annotations from the analysis by BLASTing the coding sequences against the National Center for Biotechnology Information Expressed Sequence Tag library. After translating coding sequences for Z-W orthologs to protein sequences, we used PRANK (http://tinyurl.com/prank-msa) to align gametologs. We checked alignments by eye, and removed poorly aligned regions where at least 50% of the amino acids were mismatched. PRANK's method of distinguishing insertions from deletions avoids overmatching of aligned sequences, consequentially producing more evolutionary accurate alignments for increasingly diverged sequences in comparison to ClustalW (Loytynoja and Goldman 2005; Tamura *et al.* 2011). However, in line with previous studies, to check for consistency we repeated alignments using ClustalW in MEGA5 (Tamura *et al.* 2011). We used CODEML in the PAML package version 4.4 (Yang 2007) to calculate maximum likelihood estimates of synonymous (d_S) and nonsynonymous (d_N) divergence in pairwise comparisons and generate standard errors with the curvature method. In one case, we removed ENSGALG00000013670 from the analysis to avoid biasing our results, as d_S was found to be greater than two for the multiple Z-linked orthologs, making accurate divergence estimates in avian coding sequence impossible due to mutational saturation and double hits (Axelsson *et al.* 2008). As problems with the use of PAML to calculate d_S values have been documented (Bierne and Eyre-Walker 2003) we also used an additional method, *K*-estimator 6.1 (Comeron 1995, 1999), to verify our synonymous divergence estimates for the 18 Z-W orthologs.

Some Z-W orthologs had multiple W or Z paralogs. To identify the true gametolog, we calculated d_S individually for each potential orthologous gene pair and used the gene pair with the lowest d_S estimate in the final results. The surplus genes were excluded. This is different from previous methods that employed a consensus sequence (Nam and Ellegren 2008), but avoids potential problems of relaxed or diversifying selection acting on gene copies (Graur and Li 2000) likely due to gene duplications following recombination suppression, which can result in significant divergence in both gene sequence and expression (Busby *et al.* 2011). This approach and the subsequent exclusion of multicopy genes meant that while we had sequence data for 26 W-linked genes, our conclusions regarding the number and boundaries of strata were based on d_S estimates from 18 true Z-W gametologs.

Gene conversion (Slattery *et al.* 2000; Ross *et al.* 2005) between differentiated Z-W orthologs has the potential to bias our estimates of d_S by homogenizing the two independently evolving sequences. To ensure this did not influence our identification of true gametologs, we calculated the GC content at the third codon position of Z-linked genes (GC3) and their multiple candidate W orthologs using the R package seqinR (Charif and Lobry 2007). GC content and recombination rate are positively correlated in vertebrates (Galtier 2003; Meunier and Duret 2004), leading to the prediction that, in the absence of gene conversion, the GC3 content of W-linked genes should not exceed that of Z-linked gametologs. We obtained 95% confidence intervals with bootstrapping of 1000 replicates.

Additionally, we used GENECONV as an independent test of gene conversion (http://www.math.wustl.edu/~sawyer) between Z-W gametologs, using *Anolis carolinensis* as an outgroup. For the 18 Z-linked genes, we identified 16 *Anolis* orthologs and downloaded the longest coding transcript from BioMart. We used PRANK (http://tinyurl.com/prank-msa) to generate alignments and GENECONV to identify aligned global fragments for which gametologs are sufficiently similar to be suggestive of past gene conversion. GENECONV assigns *P*-values by two methods, the first based on 10,000 random permutations and the second by a method of Karlin and Altschul (1990, 1993). Both methods are corrected for multiple comparisons and sequence length when detecting global fragments.

We mapped d_S estimates to the Z chromosome and used four criteria to delineate stratum boundaries: (1) average and range of d_S in physically proximate regions, (2) *k*-means clustering of d_S estimates into a defined number of groups in R (Forgy 1965; MacQueen 1967; Hartigan and Wong 1979; Lloyd 1982; www.R-project.org), (3) Z-W ortholog density, and (4) physical location on the Z chromosome. To verify the refined strata, we used one-tailed *t*-tests to test our explicit predictions regarding the increase in d_S estimates with strata age. Although this is a biased estimate, it is useful as a means of comparison with *k*-means clustering results.

We were also able to use d_S estimates to calculate divergence dates for Z-W orthologs with a molecular clock that accounts for sex-specific mutation rates in Z and W

chromosome lineages based on male-mutation bias estimates in birds. In galliforms, the W chromosome rate is approximately half the autosomal rate (Axelsson *et al.* 2004). Based on avian and mammalian divergence dates, the autosomal mutation rate for the chicken lineage is estimated to be 2.5×10^{-9}/site/year (Dimcheff *et al.* 2002; Hedges 2002; Webster *et al.* 2006; Nam and Ellegren 2008) and correspondingly 1.25×10^{-9}/site/year for the W chromosome. Previous work in galliforms indicates the Z chromosome mutation rate is only slightly larger than the autosomal rate leading to an estimate of 2.6×10^{-9}site/year (Nam and Ellegren 2008). The Z and W mutation rates mentioned above are therefore consistent with previous studies that estimate male mutation bias in galliforms as lying between 2 and 4 (Harlid *et al.* 2003; Berlin *et al.* 2006). These estimates result in a chicken Z-W divergence rate of 3.8×10^{-9}/site/year (Z chromosome mutation rate = 2.6×10^{-9} + W chromosome mutation rate = 1.2×10^{-9}) and divergence time of $d_S \times 3.8 \times 10^{-9}$ (Nam and Ellegren 2008) years, although this has a high variance.

To enhance our understanding of the formation of avian strata through time, we used syteny with the turkey, which diverged from the chicken ~30 MYA (Dimcheff *et al.* 2002). Genomic rearrangements are not frequent across the avian genome and birds therefore show highly conserved synteny across vast evolutionary distances (Backstrom *et al.* 2010; Warren *et al.* 2010). For the 18 Z-linked genes used to identify chicken strata, we identified 1-1 turkey orthologs and their position along the Z chromosome from Ensembl (Turkey_2.01/ GCA_000146605.1), and plotted them against Z-linked positional information from the chicken.

Comparing expression across strata

After defining putative strata boundaries based on d_S estimates, we used fragments per kilobase of exon per million mapped reads (FPKM) values obtained from the left gonad RNA-Seq data to compare expression among strata. The assembly method is discussed above under *Identification of misassembled W-linked genes*. We filtered the expression dataset to exclude any genes not expressed in at least four individuals, and all genes within the male-hypermethylated region (MHM), located between 2.5 Mb and 3.5 Mb (Teranishi *et al.* 2001). The MHM was excluded to avoid biasing our comparison of expression between strata, as microarray analysis indicates that this region is strongly downregulated in males, a markedly different pattern from the rest of the Z chromosome (Melamed and Arnold 2007).

With the filtered dataset, we separately averaged female and male FPKM values for each stratum and calculated male-biased expression as a function of $\log_2$ (male FPKM/ female FPKM) = $\log_2$ (male FPKM) − $\log_2$ (female FPKM). We compared average male-biased expression among strata using permutation tests with 1000 replicates and calculated 95% confidence intervals using bootstrapping with 1000 replicates.

In addition, as dosage compensation acts to equalize expression across the Z chromosome and autosomes, we calculated the Z:A expression ratio separately for males and females to test for signatures of this selective regime. We calculated average male and female expression separately for all Z-linked genes (minus the MHM) and autosomal genes and finally a Z:A expression ratio for each sex. This was also repeated separately for each stratum. The 95% confidence intervals were calculated by bootstrapping male and female Z-linked and autosomal averages with 1000 replicates.

We were also able to calculate and compare the rate of masculinization of the Z chromosome under both logarithmic and linear functions to describe the relationship between average stratum age and male-biased expression. In addition, to verify that the relationship we observe is not due to chance, we randomized the ratio of male to female gene expression across the Z chromosome (1000 replicates). Each time, we recalculated average male-biased expression for each stratum to obtain a distribution of randomized r^2 values.

Gene function analysis

To determine whether nonrandom distributions of gene function were contributing to gene expression differences, we used GOrilla to perform a gene ontology (GO) enrichment analysis between each stratum and the entire Z chromosome (minus the MHM), and between the Z chromosome (minus the MHM) and autosomes (Eden *et al.* 2007, 2009). We conducted each GOrilla analysis with 1-1 mouse orthologs, a threshold of $P < 10^{-2}$ using standard hypergeometric statistics and the Benjamini and Hochberg correction for multiple tests (Benjamini and Hochberg 1995).

Results

Resolving stratum boundaries

As synonymous estimates provide an appropriate measure of the length of time over which Z-W orthologs have differentiated from an original autosomal pair, and this method has been used in several organisms to establish the number and boundaries of evolutionary strata (Lahn and Page 1999; Ellegren and Carmichael 2001; Handley *et al.* 2004; Bergero *et al.* 2007; Nam and Ellegren 2008). We therefore estimated d_S values for 18 Z-W gametologs (Table 1), 6 of which were previously unknown, representing a 50% increase in the known number of Z-W orthologs. With the exception of the HINTW gene, which is known to have undergone adaptive molecular evolution in birds (Handley *et al.* 2004), d_N estimates ranged from 0.0002 to 0.0831, suggesting a high degree of conservation between Z-W gametologs. Our d_S estimates ranged from 0.128 to 0.513. Alignments repeated for consistency using ClustalW were extremely similar to those generated by PRANK, with the

Table 1 Sequence divergence between chicken Z-W orthologs

Gene pair (Z/W)	Ensembl ID (Z/W)[a]	Position on Z (Mb)	d_S (SE)[b]	d_N (SE)[b]	Putative stratum	Divergence (MYA)
ST8SIA3Z/	03049/	0.317	0.184	0.007	3	48
ST8SIA3W	22039		(0.0319)	(0.0035)		
ZNF532Z/	02852/	0.747	0.183	0.043	3	48
ZNF532W	14003		(0.0171)	(0.0045)		
RPL17Z/	02696/	0.965	0.169	0.000	3	44
RPL17W	22174		(0.0212)	(0.0000)		
SMAD2Z/	14697/	1.290	0.183	0.018	3	48
SMAD2W	10056		(0.0679)	(0.0107)		
8030462N17RikZ/	01763/	1.874	0.172	0.083	3	45
8030462N17RikW	01585		(0.0311)	(0.0126)		
ATP5A1Z/	14644/	1.938	0.203	0.029	3	54
ATP5A1W	01756		(0.0284)	(0.0061)		
UBAP2Z/	13809/	6.876	0.206	0.056	3	54
UBAP2W	05785		(0.0176)	(0.0054)		
VCPZ/	01986/	7.932	0.189	0.009	3	50
VCPW	00386		(0.0231)	(0.0017)		
GOLPH3Z/	03151/	9.013	0.128	0.007	3	34
GOLPH3W	18586		(0.0340)	(0.0052)		
ZFRZ/	03235/	9.104	0.164	0.020	3	43
ZFRW	14545		(0.0160)	(0.0031)		
NIPBLZ/	03605/	10.720	0.147	0.018	3	39
NIPBLW	22678		(0.0524)	(0.0103)		
MIER3Z/	14721/	16.837	0.219	0.044	2b	58
MIER3W	00140		(0.0619)	(0.0136)		
hnRNPKZ/	12591/	39.553	0.269	0.004	2b	71
hnRNPKW	14366		(0.0350)	(0.0021)		
SPINZ/	14916/	42.572	0.224	0.012	2b	59
SPINW	14641		(0.0534)	(0.0062)		
HINTZ/	00428/	44.169	0.513	0.314	1	135
HINTW	22685		(0.1144)	(0.0459)		
CHD1Z/	14642/	50.156	0.500	0.047	1	132
CHD1W	15278		(0.0357)	(0.0044)		
KCMF1Z/	15391/	52.658	0.286	0.040	2a	75
KCMF1	14441		(0.0385)	(0.0072)		
RASA1Z/	15639/	59.665	0.275	0.034	2a	72
RASA1W	22611		(0.0597)	(0.0105)		

[a] Ensgalg000000...
[b] Standard errors generated by taking the inverse of the second derivative of the log likelihood in Paml.

exception of 4 gametologs. This is because PRANK's treatment of insertions and deletions, in comparison to ClustalW, results in the productions of more evolutionarily accurate alignments for increasingly diverged sequences (Loytynoja and Goldman 2005; Tamura *et al.* 2011). However, for previously identified orthologs, our d_S estimates after alignment with PRANK were qualitatively identical to prior work (Nam and Ellegren 2008). To verify our synonymous divergence estimates, we used *K*-estimator 6.1 (Comeron 1995, 1999) to calculate K_s estimates for 18 Z-W orthologs. With the exception of four gene pairs, K_s was qualitatively identical (<0.03 different) to our d_S estimates, providing independent evidence for the multiple cessation of recombination across the avian Z chromosome.

Our results do differ from previous studies in some instances where multiple W-linked genes share the same Z-linked orthologs and vice versa, likely due to gene duplications following recombination suppression. This is because we used the copy with the lowest d_S estimate to avoid problems with artificially high values due to relaxed selection acting on gene duplicates (Ridley 2003). As a consequence of this approach, our d_S estimates at three locations along the Z chromosome are somewhat lower than in previous studies (Nam and Ellegren 2008), reducing slightly our estimate of when recombination ceased to form stratum 2. Additionally, newly identified Z-W orthologs extended our analysis across a wider region of the Z chromosome, allowing a reassessment of the number and boundaries of strata.

A comparison of GC content at the third codon position (GC3) between these Z-W orthologs revealed no significant difference in GC3, indicating that the lowest d_S estimates are not a result of biased gene conversion and can be used to infer true gametologs. We were also able to use GENECONV (http://www.math.wustl.edu/~sawyer) to identify aligned global fragments for which gametologs are sufficiently similar to be suggestive of past gene conversion. We used *A. carolinensis* as an outgroup, for which 16 of the 18 Z-W gametologs had *Anolis* orthologs. Of the 16 gametologs, 9

have not undergone gene conversion. However, there was limited evidence of gene conversion for 7 gametologs, with 1 significant inner fragment detected for Ensgalg00000014003 (permutation P_1 value = 0.002, Karlin and Altschul P_2 value = 0.003), Ensgalg00000001756 (P_1 value = 0.007 and P_2 value = 0.012), Ensgalg00000005785 (P_1 value <0.001 and P_2 value <0.001), Ensgalg00000014545 (P_1 value = 0.032 and P_2 value = 0.040), Ensgalg00000022678 (P_1 value <0.001 and P_2 value = 0.011), and Ensgalg00000022685 (P_1 value <0.001 and P_2 value <0.001). Significant inner fragments are evidence of a possible gene conversion event between ancestors of the gametologs, but GENECONV also detects outer fragments that represent conversion events originating from outside the alignment or within the sequence but where evidence has been destroyed by later mutation or gene conversion. Three significant outer fragments were detected for the Z-linked ortholog of Ensgalg00000014003 (P_1 value <0.001, 0.003, and 0.036; P_2 value <0.001, 0.004, and 0.038) and one for Ensgalg00000014545 (P_1 value = 0.006 and P_2 value = 0.008) and Ensgalg00000000386 (P_1 value <0.001 and P_2 value <0.001). Of the 7 gametologs with evidence of gene conversion, 6 of these are located in the youngest stratum and 1 in the oldest. Genes within the youngest stratum have the largest number of multiple copies as the cumulative strength of degenerative forces are weakest, and W-linked degeneration in minimal compared to older strata. Correspondingly, we expect a certain degree of intrachromosomal recombination for these genes (Rozen *et al.* 2003; Backstrom *et al.* 2005; Davis *et al.* 2010), which might bias GENECONV's detection of conversion events between Z and W orthologs and explain the nonrandom distribution of genes with evidence of gene conversion across the strata. In line with this, four of the seven W-linked genes for which GENECONV detected significant fragments are present in multiple copies or have a W-linked paralog.

Overall, synonymous divergence estimates cluster into at least three groups, (Handley *et al.* 2004; Nam and Ellegren 2008; Suh *et al.* 2011); however, we also find evidence for a potential fourth stratum located within the previously identified stratum 2 (Figure 1). These strata correspond to 44–50 Mb (d_S = 0.500–0.513, stratum 1), 50–60 Mb (d_S = 0.275–0.286, stratum 2a), 17–44 Mb (d_S = 0.219–0.269, stratum 2b), and 0.3–11 Mb (d_S = 0.128–0.206, stratum 3). The physical separation of stratum 2a and 2b by stratum 1 suggests that recombination suppression occurred in these regions in quick succession.

Genes within stratum 1 have a significantly higher d_S value than stratum 2 ($P < 0.001$, *t*-test), supporting the notion that recombination ceased first in this region. Our d_S estimates also differ significantly between putative strata 2a and 2b ($P = 0.050$, *t*-test), indicating recombination was suppressed independently twice in close succession. Finally, recombination ceased most recently in stratum 3, where d_S values are significantly lower than strata 1 ($P < 0.001$, *t*-test), 2a ($P < 0.001$, *t*-test), and 2b ($P < 0.001$, *t*-test).

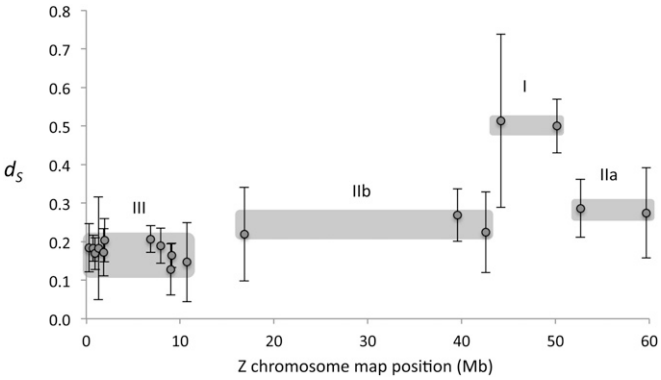

Figure 1 Evolutionary history of the chicken Z chromosome. The distribution of synonymous divergence estimates (d_S) for Z-W orthologs is shown. Physical position on the Z chromosome is based on current Ensembl genome assembly (WUGSC 2.1/galGal3). The 95% confidence intervals are based on 1000 bootstrap replicates. Synonymous divergence estimates cluster into up to four groups, which differ significantly from each other, providing support for the existence of multiple Z chromosome strata.

The results from the *k*-means analysis (Forgy 1965; MacQueen 1967; Hartigan and Wong 1979; Lloyd 1982; www.R-project.org) provide further support for multiple recombination cessation events. We used *k*-means to partition our d_S estimates into three clusters. The three groups defined by *k*-means are identical to the groups supported by the *t*-tests, providing an independent line of evidence to support the boundaries of strata 1, 2, and 3 corresponding to 44–50 Mb (stratum 1), 17–60 Mb (stratum 2), and 0.3–11 Mb (stratum 3). Due to the very short divergence interval between strata 2a and 2b, the power of *k*-means to identify these as independent clusters is weakened.

In addition to significant clustering of d_S estimates, the density of genes within each region provides independent evidence to identify strata (Lahn and Page 1999). The suppression of recombination between Z-W orthologs promotes the degeneration of W-linked genes by neutral processes, the rate of which declines with the number of functionally constrained loci (Bachtrog 2008). As recombination ceased most recently in the youngest stratum, the cumulative strength of degenerative forces are weakest, and we predict minimal W-linked degeneration compared to older stratum. Correspondingly, the largest density of identifiable gametologs is expected in the youngest stratum. Gene density across the chicken Z chromosome is consistent with this prediction (Figure 2) where the youngest stratum (1.06 genes per megabase) has a gene density of gametologs at least three times larger than the oldest (0.33 genes per megabase). The intermediate stratum has a gene density of 0.22 genes per megabase, excluding the MHM. Reasons for excluding the MHM are discussed below. Furthermore, the majority of the multigene copies excluded from the analysis were located within the youngest stratum.

As gene conversion has the potential to bias our identification of stratum boundaries, we repeated the one-tailed

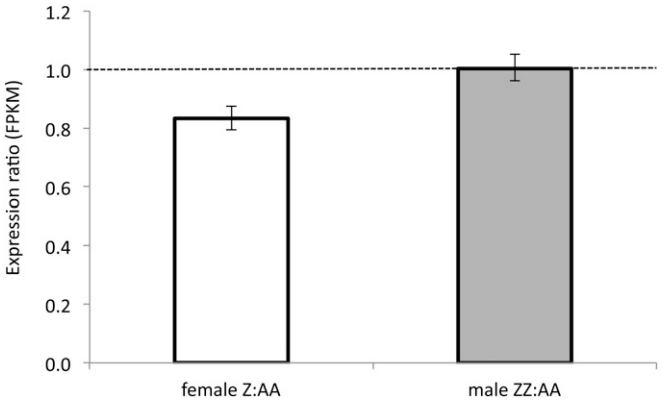

Figure 2 Comparison of gene expression across the Z chromosome and autosomes. The average Z:A expression ratio is shown for males and females separately. The 95% confidence intervals (based on 1000 bootstrap replicates) are shown. Dotted line represents the expression ratio expected under equal gene dose.

t-tests and *k*-means clustering after excluding the 7 Z-W orthologs with evidence of gene conversion. *K*-means clusters the 11 remaining gametologs into the same three groups corresponding to strata 1, 2, and 3. There is also still a significant difference between d_S estimates between strata 2 and 3 ($P < 0.001$, *t*-test), 2a and 3 ($P < 0.001$, *t*-test), 2b and 3 ($P = 0.001$, *t*-test), as with the previous analysis; however, there is now no significant difference between strata 2a and 2b ($P = 0.051$, *t*-test). As only one gene pair remains in stratum 1 after the 7 gametologs are removed, we could not conduct any tests of significance for this group. The consistency between strata boundaries defined from analyses excluding and including gametologs subject to gene conversion indicates our results are not biased by gene conversion.

Molecular clock estimates show that stratum 1 arose between 132 and 135 MYA, stratum 2a between 72 and 75 MYA, stratum 2b between 58 and 71 MYA, and stratum 3 between 34 and 54 MYA.

Sixteen of the 18 chicken Z-linked orthologs were identified on the turkey Z chromosome, of which 7 are 1-1 orthologs. Synteny of these Z-linked orthologs is conserved between both species, indicating a lack of lineage-specific genomic rearrangements.

Gene expression and the role of dosage compensation and masculinizing selection

We were able to calculate and compare average sex-biased expression for the 33 transcribed genes located within stratum 1, 154 in stratum 2 (41 genes in stratum 2a and 113 genes in stratum 2b), and the 101 genes in stratum 3 (Table S1). The three established strata show clear differences in sex-biased expression (Table 2), with a progressive stepwise relationship where stratum 1 has the greatest degree of male-biased expression, stratum 2 is intermediate in male-biased expression, and stratum 3 shows the lowest level of average male bias. Interestingly, the extent of male bias in stratum 3 (1.51 unlogged FPKM) is consistent with previous estimates of sex-bias associated with uncompensated genes on the mammalian X chromosome (Deng *et al.* 2011; Pessia *et al.* 2012). As stratum 3 is the youngest and so the cumulative effects of degenerative forces are weakest, this sex bias likely reflects the ancestral expression level associated with incomplete dosage compensation (Table 2). This pattern is strongly indicative of masculinizing selection acting on the avian Z chromosome in a cumulative manner, with increasing male bias accruing over time. This is opposite to predictions under selection for dosage compensation, which would produce a decrease in male-biased expression as a function of stratum age.

The differences in male-biased expression between strata 1 and 2 ($P = 0.021$, permutation test with 1000 replicates), strata 2 and 3 ($P < 0.001$, permutation test with 1000 replicates), and strata 1 and 3 ($P < 0.001$, permutation test with 1000 replicates) are all statistically significant. The division of stratum 2 into two putative strata follows the same pattern of successive masculinization of the Z chromosome (Table 2), providing further support to the idea that recombination ceased independently four times. Stratum 2a has a higher average male bias than stratum 2b as predicted; however, due to the very short divergence interval we would not expect this difference to be significant ($P = 0.247$, permutation test with 1000 replicates). Despite this, the pattern of male-biased expression is still strongly indicative of masculinizing selection, with male bias in both strata 2a and 2b greater than stratum 3 ($P < 0.001$ and $P < 0.001$, permutation tests with 1000 replicates) but smaller than stratum 1 ($P = 0.03$ and $P = 0.023$, permutation tests with 1000 replicates).

We excluded the MHM from stratum 2b in the analyses described above to avoid biasing our comparison of expression between strata, as microarray analysis indicates that this region is strongly downregulated in males, a markedly different pattern from the rest of Z chromosome (Melamed and Arnold 2007). Although the expression of the MHM is not representative of the Z chromosome, when included in the analysis, average male bias in stratum 3 is still significantly lower than stratum 2 ($P < 0.001$, permutation test with 1000 replicates), leaving the significant relationship between male-biased gene expression and age largely unchanged.

Results from a GOrilla analysis (Eden *et al.* 2007, 2009) suggest the observed difference in expression between the strata is not due to a nonrandom distribution of genes based on function. A comparison of each stratum to the whole Z chromosome revealed no significant enrichment of GO terms after correction for multiple testing ($P < 10^{-2}$). Even before correction, only strata 2b and 2 were enriched for any GO terms; [stratum 2b = behavior and negative regulation of multicellular organismal processes ($P = 2.25 \times 10^{-4}$ and $P = 6.84 \times 10^{-4}$), stratum 2 = behavior ($P = 2.10 \times 10^{-4}$)]; however, none of these terms were significant after the Benjamini and Hochberg correction.

Table 2 Sex-biased gene expression across Z chromosome strata

Stratum	Estimated age (MYA)	No. of expressed genes	Male-bias (log_2 male FPKM−female FPKM) (unlogged male FPKM−female FPKM)
1	133	33	0.86 (2.10)
2[a]	67	154	0.64 (1.80)
2a	73	41	0.69 (1.86)
2b[a]	63	113	0.63 (1.77)
3	46	101	0.33 (1.51)

[a] Genes from the male hypermethylated region excluded.

Additionally, when the Z chromosome is compared to the autosomes, there is no significant enrichment of GO terms ($P < 10^{-2}$) after the Benjamini and Hochberg correction. Before correction for multiple testing, the Z chromosome is enriched for thrombin receptor signaling pathway, oncostatin-M–mediated signaling pathway, positive regulation of peptidyl-tyrosine phosphorylation, positive regulation of Rho protein signal transduction, protein–lipid complex subunit organization, plasma lipoprotein particle organization, positive regulation of Ras protein signal transduction, regulation of biological quality, positive regulation of cell migration, positive regulation of positive chemotaxis, thrombin receptor activity, cytokine receptor activity, signal transducer activity, and transmembrane signaling receptor activity, intrinsic to plasma membrane ($P = 1.77 \times 10^{-5}$, $P = 8.42 \times 10^{-5}$, $P = 2.03 \times 10^{-4}$, $P = 3.26 \times 10^{-4}$, $P = 3.88 \times 10^{-4}$, $P = 3.88 \times 10^{-4}$, $P = 4.66 \times 10^{-4}$, $P = 5.43 \times 10^{-4}$, $P = 8.89 \times 10^{-4}$, $P = 9.47 \times 10^{-4}$, $P = 3.98 \times 10^{-5}$, $P = 8.42 \times 10^{-5}$, $P = 4.54 \times 10^{-4}$, $P = 4.54 \times 10^{-4}$, $P = 6.39 \times 10^{-4}$, and $P = 8.60 \times 10^{-4}$). However, it is not apparent how these processes would explain the unique expression pattern observed.

As dosage compensation acts to equalize expression across the Z chromosome and autosomes, calculating the Z:A expression ratio separately for males and females is essential for accurately assessing the selective forces driving expression evolution. We were able to calculate average male and female expression for 492 Z-linked genes (minus the MHM) and 10,530 autosomal genes and show that the Z:A expression ratio differs significantly between the sexes (Figure 2). The female Z:A ratio is consistent with incomplete dosage compensation observed in previous studies (Ellegren *et al.* 2007; Itoh *et al.* 2007, 2010), whereas the male ratio reflects the ZZ:AA karyotype. There is no significant difference in female and male Z:A ratio across strata, and this may be due to the variation in overall expression level among individual genes, which would be expected to have a strong stochastic effect in strata with relatively few genes.

Rate of masculinization

The degree of male bias increases across the Z chromosome as a function of time since divergence from the orthologous W chromosome (Figure 3, A and B). We found a stronger correlation between stratum age and degree of sex bias under a logarithmic rather than a linear function for both the three strata ($r^2 = 0.93$) and four strata ($r^2 = 0.89$) models of Z chromosome evolution. Consequentially, this suggests an upper limit to the increase in masculinization of Z-linked gene expression over evolutionary time.

We randomized gene expression across the Z chromosome (1000 replicates) to assess the probability that the observed differences in sex bias are due to chance alone. Our observed r^2 values for the positive logarithmic relationship between stratum and sex-biased expression was significant ($P = 0.04$), but the linear relationship was not ($P = 0.08$), reinforcing our results that the rate of masculinization declines over time.

Discussion

The Z chromosome could potentially be subject to two different sex-specific selection regimes. Due to the unequal inheritance pattern of the chromosome, male-specific selection for dominant alleles is predicted to drive the evolution of the Z chromosome coding content (Rice 1984), leading to the accumulation of male-benefit alleles. Correspondingly, masculinizing selection acting on gene expression (Connallon and Clark 2010) is expected to increase male bias over time.

Loss-of-function mutations for W-linked genes is expected to cause negative fitness effects (Charlesworth 1978). The resulting loss of gene activity in females, and thus unbalanced gene dose with the autosomes, is predicted to be costly, especially for dosage-sensitive Z-linked genes that interact with autosomal genes in large protein complexes (Pessia *et al.* 2012). This is predicted to result in selection for hypertranscription of dosage-sensitive genes to equalize expression with the autosomes (Ohno 1967). If selection for dosage compensation were cumulative, we would expect it to result in a pattern opposite to that produced by masculinizing selection, as an indirect effect of selection for dosage compensation is to decrease sex-biased expression on the Z over time.

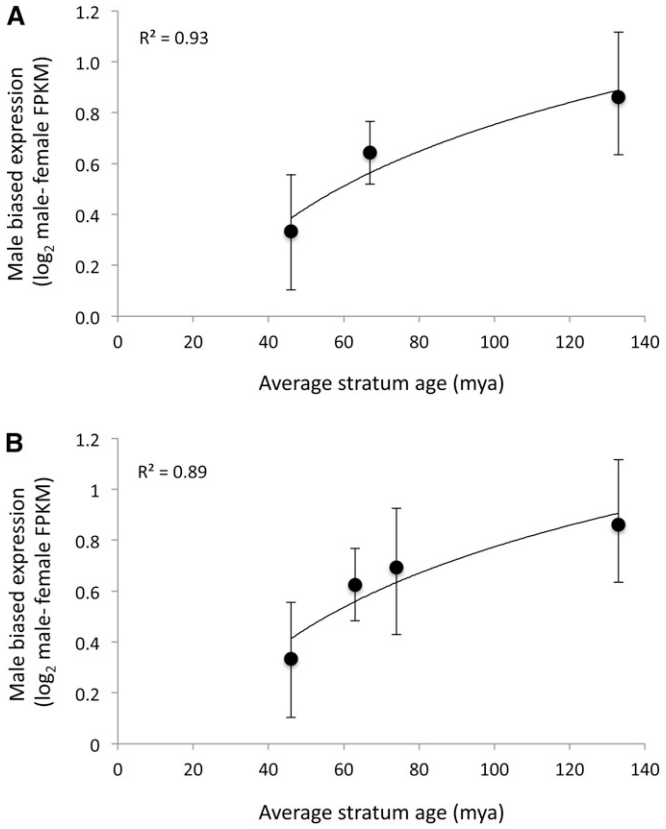

Figure 3 Rate of masculinization of the Z chromosome. The degree of male-biased expression for inferred Z chromosome strata is shown for both three strata (A) and four strata (B) models of Z chromosome evolution. The 95% confidence intervals (based on 1000 bootstrap replicates) are shown. Masculinization of Z-linked gene expression increases as a function of age, but ultimately levels out, thereby limiting the role of the Z chromosome in the evolution of sexual dimorphism.

The Z chromosome has been previously observed to show a pervasive pattern of male-biased expression (Kaiser and Ellegren 2006; Storchova and Divina 2006; Mank and Ellegren 2009a); however, the relative role of masculinizing selection *vs.* incomplete dosage compensation in driving this pattern has been difficult to untangle. Here we employ a novel method to detect signatures of sex-specific selection from selection for dosage compensation. Our results indicate that masculinizing selection predominates over selection for dosage compensation, and we were able to calculate the rate of masculinization using the evolutionary history of the chicken Z chromosome itself. Our analysis includes 18 Z-W gametologs; a sizable dataset in line with mammalian studies (Lahn and Page 1999) and to date the largest in birds, as well as 492 expressed genes along the Z chromosome.

Definition of Z chromosome strata

This study is reliant on accurate knowledge of the number and precise boundaries of strata, and using synonymous divergence estimates for newly identified Z-W orthologs together with gene density we were able to re-examine the chicken strata across a wider region of the Z chromosome than in previous studies. In doing so, we found support for the existence of up to four strata along the chicken Z chromosome. Our data suggest that recombination ceased first in the 44- to 50-Mb region of the Z chromosome, between 132 and 135 MYA (stratum 1). Then, in rapid succession, d_S estimates indicate recombination ceased in the 53- to 60-Mb region between 72 and 75 MYA (stratum 2a) and the 17- to 43-Mb region ~58–71 MYA (stratum 2b). The final suppression occurred in the 0.3- to 11-Mb region most recently, between 34 and 54 MYA (stratum 3), as shown in Figure 1. The large reduction in gene density across the oldest and intermediate strata compared to stratum 3 provides additional evidence for up to four chicken strata. In line with patterns of gene density observed across the mammalian X chromosome (Lahn and Page 1999), the number of gametologs is predicted to decline with age of the strata where the cumulative strength of degenerative forces is weakest and W-linked divergence is minimal. This decay is not predicted to follow a linear progression, but instead should decline with the number of functionally constrained loci (Bachtrog 2008).

The identification of a possible fourth stratum may suggest that recombination suppression occurred on both sides of stratum 1 in close succession, rather than requiring a suppression event for stratum 2 followed by a subsequent structural rearrangement across the boundaries of strata 1 and 2 leading to a nonlinear distribution of strata (Nam and Ellegren 2008; Bergero and Charlesworth 2009). The conservation of relative gene order between turkey and chicken Z-linked orthologs argues against a chicken-specific inversion, but does not rule out the possibility of this event occurring before the two lineages diverged. Therefore, our model to describe the successive formation of avian strata does not evoke additional cessation of recombination events but instead proposes an alternative arrangement of these events over evolutionary time.

Successive masculinization of the Z chromosome

By comparing sex-biased expression among the chicken strata, we uncover a pattern consistent with cumulative masculinizing selection acting on gene expression across the Z chromosome, where the magnitude of male-biased expression correlates closely with stratum age. Masculinization may have occurred to some extent in the youngest stratum, which ceased to recombine with the W chromosome between 34 and 54 MYA; however, the effect is greater in the older strata. This logarithmic relationship holds when both three ($r^2 = 0.93$) and four strata ($r^2 = 0.89$) are considered and is supported by randomization tests ($P = 0.04$). The significance of the randomization test, although marginal, is convergent with the correlation between stratum age and male-biased expression level. The cumulative masculinization of Z chromosome expression is likely due to the effect of male-specific selection, as the Z chromosome is twice as often present in males than females; therefore,

dominant alleles are more often selected for their male-specific fitness effects (Rice 1984; Connallon and Clark 2010).

Unbalanced sex-specific selection on sex chromosomes has been suggested to explain the complex nonrandom distribution of sex-biased genes in mammals (Khil *et al.* 2004) and *Drosophila* (Parisi *et al.* 2003; Sturgill *et al.* 2007; Chen *et al.* 2011). However, it has not been clear from previous work in birds how much of the male bias on the Z chromosome is due to masculinization (Ellegren 2011) and how much to incomplete dosage compensation (Ellegren *et al.* 2007; Itoh *et al.* 2007). Previous attempts to circumvent the problems of gene dose compared embryonic and adult gene expression levels (Mank and Ellegren 2009b); however, this did not account for the fact that sex-specific selection and sex-biased gene expression shift rapidly throughout development and the adult life cycle (Mank *et al.* 2010). Our method allows male-biased selection to be detected without conducting a comparison between developmental time points, avoiding the problems encountered in previous studies of varying magnitudes of sex-specific selection. Consequentially, our results suggest that the increase in male bias over time is consistent with a cumulative effect of masculinizing selection rather than selection for dosage compensation. The older strata have been subjected to longer periods of stronger selection for dominant male-benefit alleles, and hence show a greater degree of masculinization. This finding is consistent with theoretical predictions and a recent study showing that genes expressed in the testis are overrepresented among newly emerged Z-linked genes (Ellegren 2011). The results can also explain previous findings that the extent of sex-biased expression varies across the Z chromosome (Melamed and Arnold 2007).

Determining whether sex-specific selection is responsible for the analogous feminization of the X chromosome in *Drosophila* (Parisi *et al.* 2003; Sturgill *et al.* 2007; Chen *et al.* 2011) and mice (Khil *et al.* 2004), is complicated by meiotic sex chromosome inactivation, a genetic mechanism inactivating sex-linked genes during late spermatogenesis (Vibranovski *et al.* 2009a). MSCI has been proposed as a driving force behind the underrepresentation of spermatogenesis genes on the X chromosome; however, this is surrounded by much debate (Khil *et al.* 2004; Vibranovski *et al.* 2009b; Meiklejohn *et al.* 2011). In birds, MSCI is ephemeral, lasting only from early pachytene to diplotene phases during oogenesis (Schoenmakers *et al.* 2009). As our samples were taken before this stage, MSCI is unlikely to drive the majority of the nonrandom pattern of sex-biased genes on the Z chromosome we observe.

The pattern of successive masculinization across the Z chromosome is in contrast to that seen on the human X chromosome, whereby more genes escape X chromosome inactivation and are therefore expressed more in females, in the younger regions of the X than the older regions (Carrel and Willard 2005). This hints at the important role of dosage compensation in shaping the pattern of sex-biased genes on sex chromosomes, as the dosage compensation mechanism is less effective in younger regions of the mammalian X chromosome (Lin *et al.* 2007; Deng *et al.* 2011). The *Drosophila* X chromosome does not exhibit strata, possibly because the lack of recombination in males means that the Y decayed as a single unit or because the Y and X are not orthologous (Hackstein *et al.* 1996). Birds lack global sex chromosome dosage compensation (Ellegren *et al.* 2007; Itoh *et al.* 2007), and we would expect that if there were significant selection for dosage compensation, the pattern of male-biased expression would decrease as a function of stratum age. We predict that these differences should be strongly visible across the Z chromosome, as evidence in several animals suggests that dosage compensation mechanisms evolve relatively slowly. The sex chromosomes of monotremes and birds are ancient and yet still display incomplete dosage compensation (Itoh *et al.* 2007; Deakin *et al.* 2008) and while there is a current debate regarding the status of dosage compensation in eutherian mammals, evidence suggests that these chromosomes also lack complete dosage compensation (Deng *et al.* 2011; Pessia *et al.* 2012). However, limited evidence from plants indicates that in some cases selection for dosage compensation may be more rapid (Muyle *et al.* 2012). Despite this, our data are not consistent with selection for dosage compensation, as we observe male bias increasing over time in a cumulative fashion. Additionally, comparing Z-linked and autosomal expression confirms that females are not under selection for dosage compensation, as the expression ratio is consistent with a lack of hypertranscription of the single Z chromosome (Ellegren *et al.* 2007; Itoh *et al.* 2007, 2010). This suggests a trade-off between dosage compensation mechanisms and sex-specific selection on gene expression, with dosage compensation mechanisms acting against the effects of cumulative sex-specific selection on sex chromosome transcription.

Interestingly, the static architecture of the avian genome may also contribute to the pattern of male-biased expression we observe on the Z chromosome. Recently, it has been shown that there is little gene traffic on and off the Z chromosome, potentially due to a lack of active transposons within the avian genome (Toups *et al.* 2011). The unequal inheritance pattern of the Z chromosome renders it unfavorable for dominant female-benefit alleles (Rice 1984); however, the potential for these genes to move to the autosomes is severely limited. Subsequently, all Z-linked genes are subject to strong male-specific selection, favoring the evolution of male-biased expression.

Concluding remarks

A considerable body of theory focuses on the importance of sex chromosomes in facilitating the evolution of sexually dimorphic phenotypes (Rice 1984). Indeed, the Z chromosome is thought to be especially conducive to sexual selection and predicted to play an important role in encoding sexually selected traits, due to the male biased inheritance pattern (Reeve and Pfennig 2003; Kirkpatrick and Hall

2004; Albert and Otto 2005). Implicit in these predictions is that the importance of sex linkage in encoding sexually dimorphic phenotypes increases over evolutionary time with cumulative exposure to sex-specific selection; however, chromosome or stratum age is typically not considered when attempting to measure the importance of sex chromosomes in the evolution of sexually selected traits. Here we show that the cumulative effects of male-specific selection vary across the chromosome, indicating that older regions harbor more genes that may contribute to sexual dimorphic phenotypes via sex-biased expression than younger ones.

Acknowledgments

We thank Tommaso Pizzari (Oxford University), who provided Red Jungle Fowl samples and the use of incubation facilities, as well as Peter Harrison, Marie Pointer, Stephen Montgomery, Laura Ross, and Jarrod Hadfield for helpful comments on the manuscript. We thank Lorna Gregory and the Welcome Trust Centre for Human Genetics (funded by Welcome Trust grant reference 090532/Z09/Z and Medical Research Council hub grant G0900747 91070) for the generation of the sequencing data. The work described here was funded by the Biotechnology and Biological Sciences Research Council and the European Research Council under the Framework 7 agreement (grant agreement 260233). A.E.W. is supported by a National Environment Research Council Ph.D. studentship.

Literature Cited

Agrawal, A. F., and M. C. Whitlock, 2011 Inferences about the distribution of dominance drawn from yeast gene knockout data. Genetics 187: 553–566.

Albert, A. Y. K., and S. P. Otto, 2005 Sexual selection can resolve sex-linked sexual antagonism. Science 310: 119–121.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. H. Zhang, Z. Zhang *et al.*, 1997 Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389–3402.

Arnqvist, G., and L. Rowe, 2005 *Sexual Conflict*. Princeton University Press, Princeton, NJ.

Axelsson, E., N. G. C. Smith, H. Sundström, S. Berlin, and H. Ellegren, 2004 Male-biased mutation rate and divergence in autosomal, Z-Linked and W-Linked introns of chicken and turkey. Mol. Biol. Evol. 21: 1538–1547.

Axelsson, E., L. Hultin-Rosenberg, M. Brandström, M. Zwahlen, D. F. Clayton *et al.*, 2008 Natural selection in avian protein-coding genes expressed in brain. Mol. Ecol. 17: 3008–3017.

Bachtrog, D., 2008 The temporal dynamics of processes underlying Y chromosome degeneration. Genetics 179: 1513–1525.

Bachtrog, D., M. Kirkpatrick, J. E. Mank, S. F. McDaniel, J. C. Pires *et al.*, 2011 Are all sex chromosomes created equal? Trends Genet. 27: 350–357.

Backstrom, N., H. Ceplitis, S. Berlin, and H. Ellegren 2005 Gene conversion drives the evolution of HINTW, an ampliconic gene on the female-specific avian W chromosome. Mol. Biol. Evol. 22: 1992–1999.

Backstrom, N., E. Palkopoulou, A. Qvarnstrom, and H. Ellegren, 2010 No evidence for Z-chromosome rearrangements between the pied flycatcher and the collared flycatcher as judged by gene-based comparative genetic maps. Mol. Ecol. 19: 3394–3405.

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57: 289–300.

Bergero, R., and D. Charlesworth, 2009 The evolution of restricted recombination in sex chromosomes. Trends Ecol. Evol. 24: 94–102.

Bergero, R., A. Forrest, E. Kamau, and D. Charlesworth, 2007 Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. Genetics 175: 1945–1954.

Berlin, S., M. Brandstrom, N. Backstrom, E. Axelsson, N. Smith *et al.*, 2006 Substitution rate heterogeneity and the male mutation bias. J. Mol. Evol. 62: 226–233.

Bierne, N., and A. Eyre-Walker, 2003 The problem of counting sites in the estimation of the synonymous and nonsynonymous substitution rates: implications for the correlation between the synonymous substitution rate and codon usage bias. Genetics 165: 1587–1597.

Busby, M., J. Gray, A. Costa, C. Stewart, M. Stromberg *et al.*, 2011 Expression divergence measured by transcriptome sequencing of four yeast species. BMC Genomics 12: 635.

Carrel, L., and H. F. Willard, 2005 X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434: 400–404.

Charif, D., and J. R. Lobry, 2007 Seqin{R} 1.0–2: a contributed package to the {R} project for statistical computing devoted to biological sequences retrieval and analysis, pp. 207–232 in *Structural Approaches to Sequence Evolution: Molecules*, *Networks*, *Populations*. Springer-Verlag, Berlin.

Charlesworth, B., 1978 Model for evolution of Y chromosomes and dosage compensation. Proc. Natl. Acad. Sci. USA 75: 5618–5622.

Charlesworth, B., 1996 The evolution of chromosomal sex determination and dosage compensation. Curr. Biol. 6: 149–162.

Charlesworth, B., J. A. Coyne, and N. H. Barton, 1987 The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. 130: 113–146.

Charlesworth, D., B. Charlesworth, and G. Marais, 2005 Steps in the evolution of heteromorphic sex chromosomes. Heredity 95: 118–128.

Chen, Z.-X., Y. E. Zhang, M. Vibranovski, J. Luo, G. Gao *et al.*, 2011 Deficiency of X-linked inverted duplicates with male-biased expression and the underlying evolutionary mechanisms in the Drosophila genome. Mol. Biol. Evol. 28: 2823–2832.

Chippindale, A. K., J. R. Gibson, and W. R. Rice, 2001 Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. Proc. Natl. Acad. Sci. USA 98: 1671–1675.

Comeron, J. M., 1995 A method for estimating the numbers of synonymous and nonsynonymous substitutions per site. J. Mol. Evol. 41: 1152–1159.

Comeron, J. M., 1999 K-estimator: calculation of the number of nucleotide substitutions per site and the confidence intervals. Bioinformatics 15: 763–764.

Connallon, T., and L. L. Knowles, 2005 Intergenomic conflict revealed by patterns of sex-biased gene expression. Trends Genet. 21: 495–499.

Connallon, T., and A. G. Clark, 2010 Sex linkage, sex-specific selection and the role of recombination in the evolution of sexually dimorphic gene expression. Evolution 64: 3417–3442.

Connallon, T., R. M. Cox, and R. Calsbeek, 2010 Fitness consequences of sex-specific selection. Evolution 64: 1671–1682.

Davis, J. K., P. J. Thomas, and N. C. S. Progra, 2010 A W-linked palindrome and gene conversion in New World sparrows and blackbirds. Chromosome Res. 18: 543–553.

Deakin, J. E., T. A. Hore, E. Koina, and J. A. M. Graves, 2008 The status of dosage compensation in the multiple X chromosomes of the platypus. PLoS Genet. 4: e1000140.

Deng, X., J. B. Hiatt, N. Di Kim, S. Ercan, D. Sturgill *et al.*, 2011 Evidence for compensatory upregulation of expressed X-linked genes in mammals, *Caenorhabditis elegans* and *Drosophila melanogaster*. Nat. Genet. 43: 1179–1185.

Dimcheff, D. E., S. V. Drovetski, and D. P. Mindell, 2002 Phylogeny of *Tetraoninae* and other galliform birds using mitochondrial 12S and ND2 genes. Mol. Phylogenet. Evol. 24: 203–215.

Eden, E., D. Lipson, S. Yogev, and Z. Yakhini, 2007 Discovering motifs in ranked lists of DNA sequences. PLOS Comput. Biol. 3: 508–522.

Eden, E., R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini, 2009 GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10: 48.

Ellegren, H., 2011 Emergence of male-biased genes on the chicken Z-chromosome: contrasts between male and female heterogametic systems. Genome Res. 21: 2082–2086.

Ellegren, H., and A. Carmichael, 2001 Multiple and independent cessation of recombination between avian sex chromosomes. Genetics 158: 325–331.

Ellegren, H., and J. Parsch, 2007 The evolution of sex-biased genes and sex-biased gene expression. Nat. Rev. Genet. 8: 689–698.

Ellegren, H., L. Hultin-Rosenberg, B. Brunstrom, L. Dencker, K. Kultima *et al.*, 2007 Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. BMC Biol. 5: 40.

Forgy, E. W., 1965 Cluster analysis of multivariate data: efficiency *vs.* interpretability of classifications. Biometrics 21: 768–769.

Galtier, N., 2003 Gene conversion drives GC content evolution in mammalian histones. Trends Genet. 19: 65–68.

Graur, D., and W.-H. Li, 2000 *Fundamentals of Molecular Evolution*. Sinauer Associates, Sunderland, MA.

Hackstein, J. H. P., R. Hockstenbach, E. Hacksteinjungen, and L. W. Beukeboom, 1996 Is the Y chromosome of *Drosophila* an evolved supernumerary chromosome? Bioessays 18: 317–323.

Handley, L. L., H. Ceplitis, and H. Ellegren, 2004 Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. Genetics 167: 367–376.

Harlid, A. B., S. Berlin, N. G. C. Smith, A. P. Mosller, and H. Ellegren, 2003 Life history and the male mutation bias. Evolution 57: 2398–2406.

Hartigan, J. A., and M. A. Wong, 1979 Algorithm AS 136: a k-means clustering algorithm. J. R. Stat. Soc. Ser. C Appl. Stat. 28: 100–108.

Hedges, S. B., 2002 The origin and evolution of model organisms. Nat. Rev. Genet. 3: 838–849.

Innocenti, P., and E. H. Morrow, 2010 The sexually antagonistic genes of Drosophila melanogaster. PLoS Biol. 8: e1000336.

Itoh, Y., E. Melamed, X. Yang, K. Kampf, S. Wang *et al.*, 2007 Dosage compensation is less effective in birds than in mammals. J. Biol. 6: 2.

Itoh, Y., E. Replogle, Y. H. Kim, J. Wade, D. F. Clayton *et al.*, 2010 Sex bias and dosage compensation in the zebra finch *vs.* chicken genomes: general and specialized patterns among birds. Genome Res. 20: 512–518.

Julien, P., D. Brawand, M. Soumillon, A. Necsulea, A. Ljuechti *et al.*, 2012 Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. PLoS Biol. 10: e1001328.

Kaiser, V. B., and H. Ellegren, 2006 Nonrandom distribution of genes with sex-biased expression in the chicken genome. Evolution 60: 1945–1951.

Karlin, S., and S. F. Altschul, 1990 Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proc. Natl. Acad. Sci. USA 87: 2264–2268.

Karlin, S., and S. F. Altschul, 1993 Applications and statistics for multiple high-scoring segments in molecular sequences. Proc. Natl. Acad. Sci. USA 90: 5873–5877.

Khil, P. P., N. A. Smirnova, P. J. Romanienko, and R. D. Camerini-Otero, 2004 The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. Nat. Genet. 36: 642–646.

Kirkpatrick, M., and D. W. Hall, 2004 Sexual selection and sex linkage. Evolution 58: 683–691.

Lahn, B. T., and D. C. Page, 1999 Four evolutionary strata on the human X chromosome. Science 286: 964–967.

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10: R25.

Lin, F., K. Xing, J. Zhang,, and X. He, 2012 Expression reduction in mammalian X chromosome evolution refutes Ohno's hypothesis of dosage compensation. Proc. Natl. Acad. Sci. USA 109: 11752–11757.

Lin, H., V. Gupta, M. D. VerMilyea, F. Falciani, J. T. Lee *et al.*, 2007 Dosage compensation in the mouse balances up-regulation and silencing of X-linked genes. PLoS Biol. 5: 2809–2820.

Lloyd, S. P., 1982 Least-squares quantization in PCM. IEEE Trans. Inf. Theory 28: 129–137.

Loytynoja, A., and N. Goldman, 2005 An algorithm for progressive multiple alignment of sequences with insertions. Proc. Natl. Acad. Sci. USA 102: 10557–10562.

Lucchesi, J. C., 1973 Dosage compensation in *Drosophila*. Annu. Rev. Genet. 7: 225–237.

MacQueen, J., 1967 Some methods for classification and analysis of multivariate observations, pp. 281–297 in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, Berkley, CA.

Mank, J. E., 2009 Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. Am. Nat. 173: 141–150.

Mank, J. E., and H. Ellegren, 2009a All dosage compensation is local: gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. Heredity 102: 312–320.

Mank, J. E., and H. Ellegren, 2009b Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. Evolution 63: 1464–1472.

Mank, J. E., K. Nam, B. Brunstrom, and H. Ellegren, 2010 Ontogenetic complexity of sexual dimorphism and sex-specific selection. Mol. Biol. Evol. 27: 1570–1578.

Mank, J. E., D. J. Hosken, and N. Wedell, 2011 Some inconvenient truths about sex chromosome dosage compensation and the potential role of sexual conflict. Evolution 65: 2133–2144.

Meiklejohn, C. D., E. L. Landeen, J. M. Cook, S. B. Kingan, and D. C. Presgraves, 2011 Sex chromosome-specific regulation in the *Drosophila* male germline but little evidence for chromosomal dosage compensation or meiotic inactivation. PLoS Biol. 9: e1001126.

Melamed, E., and A. P. Arnold, 2007 Regional differences in dosage compensation on the chicken Z chromosome. Genome Biol. 8: R202.

Meunier, J., and L. Duret, 2004 Recombination drives the evolution of GC-content in the human genome. Mol. Biol. Evol. 21: 984–990.

Moghadam, H. K., M. A. Pointer, A. E. Wright, S. Berlin, and J. E Mank, 2012 W chromosome expression responds to female-specific selection. Proc. Natl. Acad. Sci. USA 109: 8207–8211.

Morrow, E. H., A. D. Stewart, and W. R. Rice, 2008 Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. J. Evol. Biol. 21: 1046–1054.

Muyle, A., N. Zemp, C. Deschamps, S. Mousset, A. Widmer *et al.*, 2012 Rapid de novo evolution of X chromosome dosage

compensation in *Silene latifolia*, a plant with young sex chromosomes. PLoS Biol. 10: e1001308.
Nam, K., and H. Ellegren, 2008 The chicken (*Gallus gallus*) Z chromosome contains at least three nonlinear evolutionary strata. Genetics 180: 1131–1136.
Ohno, S., 1967 *Sex Chromosomes and Sex Linked Genes*. Springer-Verlag, Berlin.
Oliver, B., 2007 Sex, dose, and equality. PLoS Biol. 5: e340.
Parisi, M., R. Nuttall, D. Naiman, G. Bouffard, J. Malley *et al.*, 2003 Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. Science 299: 697–700.
Pessia, E., T. Makino, M. Bailly-Bechet, A. McLysaght, and G. A. B. Marais, 2012 Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. Proc. Natl. Acad. Sci. USA 109: 5346–5351.
Pollack, J. R., T. Sorlie, C. M. Perou, C. A. Ress, S. S. Jeffrey *et al.*, 2002 Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proc. Natl. Acad. Sci. USA 99: 12963–12968.
Reeve, H. K., and D. W. Pfennig, 2003 Genetic biases for showy males: Are some genetic systems especially conducive to sexual selection? Proc. Natl. Acad. Sci. USA 100: 1089–1094.
Rice, W. R., 1984 Sex chromosomes and the evolution of sexual dimorphism. Evolution 38: 735–742.
Ridley, M., 2003 *Evolution*. John Wiley & Sons, New York.
Ross, M. T., D. V. Grafham, A. J. Coffey, S. Scherer, K. McLay *et al.*, 2005 The DNA sequence of the human X chromosome. Nature 434: 325–337.
Rozen, S., H. Skaletsky, J. D. Marszalek, P. J. Minx, and H. S. Cordum, 2003 Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature 423: 873–876.
Schoenmakers, S., E. Wassenaar, J. W. Hoogerbrugge, J. S. E. Laven, J. A. Grootegoed *et al.*, 2009 Female meiotic sex chromosome inactivation in chicken. PLoS Genet. 5: e1000466.
Slattery, J. P., L. Sanner-Wachter, and S. J. O'Brien, 2000 Novel gene conversion between X-Y homologues located in the nonrecombining region of the Y chromosome in *Felidae* (Mammalia). Proc. Natl. Acad. Sci. USA 97: 5307–5312.
Storchova, R., and P. Divina, 2006 Nonrandom representation of sex-biased genes on chicken Z chromosome. J. Mol. Evol. 63: 676–681.
Sturgill, D., Y. Zhang, M. Parisi, and B. Oliver, 2007 Demasculinization of X chromosomes in the *Drosophila* genus. Nature 450: 238–241.
Suh, A., J. O. Kriegs, J. Brosius, and J. Schmitz, 2011 Retroposon insertions and the chronology of avian sex chromosome evolution. Mol. Biol. Evol. 28: 2993–2997.
Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei *et al.*, 2011 MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28: 2731–2739.
Teranishi, M., Y. Shimada, T. Hori, O. Nakabayashi, T. Kikuchi *et al.*, 2001 Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. Chromosome Res. 9: 147–165.
Torres, E. M., B. R. Williams, and A. Amon, 2008 Aneuploidy: cells losing their balance. Genetics 179: 737–746.
Toups, M. A., J. B. Pease, and M. W. Hahn, 2011 No excess gene movement is detected off the avian or lepidopteran Z chromosome. Genome Biol. Evol. 3: 1381–1390.
Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan *et al.*, 2010 Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28: 511–515.
Vibranovski, M. D., H. F. Lopes, T. L. Karr, and M. Long, 2009a Stage-specific expression profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. PLoS Genet. 5: e1000731.
Vibranovski, M. D., Y. Zhang, and M. Long, 2009b General gene movement off the X chromosome in the *Drosophila* genus. Genome Res. 19: 897–903.
Vicoso, B., and B. Charlesworth, 2006 Evolution on the X chromosome: unusual patterns and processes. Nat. Rev. Genet. 7: 645–653.
Vicoso, B., and D. Bachtrog, 2011 Lack of global dosage compensation in Schistosoma mansoni, a female-heterogametic parasite. Genome Biol. Evol. 3: 230–235.
Warren, W. C., D. F. Clayton, H. Ellegren, A. P. Arnold, L. W. Hillier *et al.*, 2010 The genome of a songbird. Nature 464: 757–762.
Webster, M. T., E. Axelsson, and H. Ellegren, 2006 Strong regional biases in nucleotide substitution in the chicken genome. Mol. Biol. Evol. 23: 1203–1216.
Yang, Z., 2007 PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24: 1586–1591.
Zhang, Y. E., M. D. Vibranovski, B. H. Krinsky, and M. Long, 2010a Age-dependent chromosomal distribution of male-biased genes in *Drosophila*. Genome Res. 20: 1526–1533.
Zhang, Y. E., M. D. Vibranovski, P. Landback, G. A. B. Marais, and M. Long, 2010b Chromosomal redistribution of male-biased genes in mammalian evolution with two bursts of gene gain on the X Chromosome. PLoS Biol. 8: e1000494.
Zhou, Q., and D. Bachtrog, 2012 Sex-specific adaptation drives early sex chromosome evolution in Drosophila. Science 337: 341–345.

Communicating editor: B. Payseur

doi:10.1111/evo.12493

INDEPENDENT STRATUM FORMATION ON THE AVIAN SEX CHROMOSOMES REVEALS INTER-CHROMOSOMAL GENE CONVERSION AND PREDOMINANCE OF PURIFYING SELECTION ON THE W CHROMOSOME

Alison E. Wright,[1,2,3] Peter W. Harrison,[2] Stephen H. Montgomery,[2] Marie A. Pointer,[1] and Judith E. Mank[2]

[1]*Department of Zoology, Edward Grey Institute, University of Oxford, Oxford OX1 3PS, United Kingdom*

[2]*Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, United Kingdom*

[3]*E-mail: alison.e.wright@ucl.ac.uk*

Received May 4, 2014
Accepted July 15, 2014

We used a comparative approach spanning three species and 90 million years to study the evolutionary history of the avian sex chromosomes. Using whole transcriptomes, we assembled the largest cross-species dataset of W-linked coding content to date. Our results show that recombination suppression in large portions of the avian sex chromosomes has evolved independently, and that long-term sex chromosome divergence is consistent with repeated and independent inversions spreading progressively to restrict recombination. In contrast, over short-term periods we observe heterogeneous and locus-specific divergence. We also uncover four instances of gene conversion between both highly diverged and recently evolved gametologs, suggesting a complex mosaic of recombination suppression across the sex chromosomes. Lastly, evidence from 16 gametologs reveal that the W chromosome is evolving with a significant contribution of purifying selection, consistent with previous findings that W-linked genes play an important role in encoding sex-specific fitness.

KEY WORDS: **Chromosome inversion, female heterogamety, recombination suppression.**

Recombination suppression initiates the divergence of sex chromosomes from an ancestral autosomal pair, and promotes the degeneration of the W (in female heterogamety) and Y (in male heterogamety) chromosome coding content (Ohno 1967; Charlesworth 1991; Charlesworth and Charlesworth 2000; Charlesworth et al. 2005; Bachtrog et al. 2008). In spite of the pivotal role recombination suppression plays in sex chromosome evolution, our understanding of the underlying mechanisms and dynamics is largely limited to either highly divergent sex chromosomes (e.g., mammals and *Drosophila*), or incipient sex chromosome systems, such as those observed in plants and fish (Chibalina and Filatov 2011; Muyle et al. 2012; Bergero et al. 2013; Natri et al. 2013; Qiu et al. 2013).

Older sex chromosomes support the classic model of sex chromosome evolution, where intra-chromosomal rearrangements are proposed to initiate the suppression of recombination across large regions (Charlesworth et al. 2005). Theoretically, these inversions result in instantaneous recombination suppression between the sex chromosomes. This leaves a spatial signature of clusters of gametologs, orthologous genes on the X and Y (or Z and W) chromosomes, of similar divergence, often referred to as strata. Strata are thought to form independently in a

This article was published online on 29 August 2014. Subsequently, it was determined that a typographical error had been introduced in Table 2 caption, the correction was published on 24 September 2014.

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

stepwise process over millions of years. At least four chromosomal rearrangements occurring over 300 million years have been documented on the human X and Y chromosomes (Lahn and Page 1999; Ross et al. 2005; Lemaitre et al. 2009; Wilson and Makova 2009; Pandey et al. 2013; Bellott et al. 2014), and there is evidence of four strata on the avian sex chromosomes spanning 130 million years (Wright et al. 2012).

Newly evolved sex chromosomes show less concordant support for the strata model. The threespine stickleback (*Gasterosteus aculeatus*) and white campion (*Silene latifolia*) X and Y chromosomes originated around 30 million years ago (Mya) and less than 10 Mya, respectively. Work on these systems indicates that recombination suppression evolves heterogeneously along the length of the chromosome (Chibalina and Filatov 2011; Muyle et al. 2012; Bergero et al. 2013; Natri et al. 2013; Qiu et al. 2013; Roesti et al. 2013). This suggests that although selection to suppress recombination between the X and Y chromosomes occurs over large regions, within those regions, the effects are variable and genetic exchange between the sex chromosomes persists in some places.

We need to fully integrate short- and long-term views of sex chromosome evolution to provide a complete temporal overview of the dynamics of recombination suppression. Additionally, much of our understanding of sex chromosome evolution is based on studies of neo-sex chromosomes in *Drosophila* (Bachtrog et al. 2008; Kaiser et al. 2011; Zhou and Bachtrog 2012). These have been highly informative, however, males are achiasmatic in *Drosophila* and therefore cannot be used to study progressive recombination suppression events. Instead, comparative analyses of sex chromosome divergence across a range of species with recombination in both sexes are required to provide an evolutionary framework to characterize the temporal dynamics of sex chromosome evolution.

The avian sex chromosomes are homologous across the entire clade, however recombination suppression between the Z and W has evolved repeatedly in different lineages across 130 million years (Tsuda et al. 2007; Suh et al. 2011). We use the repeated evolution of avian sex chromosome strata to create a cohesive view of short-, medium-, and long-term dynamics in Z–W divergence patterns. To do this, we assembled the largest cross-species dataset of W-linked coding content to date spanning the Galliformes (the landfowl) and the Anseriformes (the waterfowl). These two sister orders, the Galloanserae, last shared a common ancestor approximately 90 Mya (van Tuinen and Hedges 2001). Our analysis indicates that the majority of the Z and W chromosomes formed independently in each of these orders. Additionally, our results suggest ongoing recombination between gametologs in the most recent region of the Anseriform sex chromosomes, and gene conversion throughout the sex chromosomes.

Materials and Methods

IDENTIFICATION OF W-LINKED GENES

Previous efforts to identify W-linked sequences in birds (Chen et al. 2012) have resulted primarily in noncoding sequence, which makes comparisons to the Z chromosome difficult. To expand the known coding content of both the mallard duck (*Anas platyrhynchos*) and wild turkey (*Meleagris gallopavo*) W chromosomes, we used a combined approach based on RNA-seq data and sequence similarity to known red jungle fowl (*Gallus gallus*) W-linked genes.

RNA-seq data were obtained from captive populations of *A. platyrhynchos* and *M. gallopavo*, taken at the start of their first breeding season (year 1 for *A. platyrhynchos*, year 2 for *M. gallopavo*). All samples were collected in accordance with national guidelines and with permission from institutional ethical review committees. The left gonad was collected separately from five male and five female *A. platyrhynchos*, and seven male and five female *M. gallopavo*. The spleen was collected from five male and five female *A. platyrhynchos* and four male and two female *M. gallopavo*. The samples were homogenized and stored in RNAlater until preparation. We extracted RNA using the Animal Tissue RNA Kit (Qiagen) and The Wellcome Trust Centre for Human Genetics, University of Oxford, prepared and barcoded samples using standard methods. RNA was sequenced on an Illumina HiSeq 2000 resulting in on average 26 million 100 bp paired-end reads per sample.

The data were assessed for quality using FastQC version 0.10.1 (www.bioinformatics.babraham.ac.uk/projects/fastqc) and filtered using Trimmomatic version 0.22 (Lohse et al. 2012). Specifically, reads with residual adaptor sequences were removed and reads were trimmed if the leading or trailing bases had a Phred score <4 or the sliding window average Phred score over four bases was <15. Post filtering, reads were removed if either read pair was <25 bases in length. Filtered reads from each species were mapped to the respective reference genomes obtained from Ensembl version 74 (Flicek et al. 2013; *M. gallopavo* version 2.01/GCA_000146605.1; Dalloul et al. 2010; Dalloul et al. 2014) and *A. platyrhynchos* version 1.0/GCA_000355885.1; Huang et al. 2013) using TopHat2 version 2.09 (Trapnell et al. 2009; Kim et al. 2013) and Bowtie2 version 2.1.1 (Langmead et al. 2009; Langmead and Salzberg 2012). For recently evolved gametologs with high sequence similarity, it is possible that Z-linked reads will spuriously map to W-linked exons. To avoid this and allow accurate identification of female-limited genes and therefore putative W-linkage, both reads of a pair had to map concordantly to the reference sequence and no mismatches in the alignment were permitted. This strategy, although essential to accurately differentiate gametologs, could fail to identify W-linked genes

when the entire locus is not encompassed on a single scaffold in the genome assembly.

To identify polymorphisms between Z and W coding sequences the mapping criteria were relaxed. For each sample, we used Cufflinks version 2.1.0 (Trapnell et al. 2010, 2013) to estimate transcript abundances for Ensembl annotated genes as fragments per kilobase per million mappable reads. We identified strongly female-biased and female-limited genes as putative W-linked genes using gene expression profiles in males and females (Fig. S1).

Putative W-linked genes were also identified by high sequence similarity to known G. gallus W-linked genes. *Gallus gallus* (Galgal4.0/GCA_000002315.2), *M. gallopavo*, and *A. platyrhynchos* cDNA sequences were obtained from Ensembl version 74 (Flicek et al. 2013) and for each species the longest transcript for each gene was identified. We determined putative orthology using BLAST version 2.2.26 (Altschul et al. 1990) with an e-value cutoff of 1×10^{-10} and identified putative *A. platyrhynchos* and *M. gallopavo* W-linked genes that had the highest BLAST score to known *G. gallus* W-linked genes.

For both *A. platyrhynchos* and *M. gallopavo*, we validated putative W-linked genes by PCR using genomic DNA from three individuals of each sex.

IDENTIFICATION OF GAMETOLOGS

Previously, we used newly identified *G. gallus* gametologs to identify four strata on the Z chromosome (Wright et al. 2012). Subsequent to this work, the Ensembl annotation of the *G. gallus* W and Z chromosomes has been revised (Flicek et al. 2013) and additional W-linked genes have been identified (Ayers et al. 2013). To account for these changes, we updated gametologs and reanalyzed divergence estimates using coding sequences from Ensembl version 74 (Flicek et al. 2013) and PAML version 4.7 (Yang 2007). Revised divergence estimates were compared to previous estimates (Wright et al. 2012) with Spearman's correlation in the rcorr function (Hmisc package [Schemper 2003; Tian et al. 2007] in R [R Core Team 2011]). Across gametologs, revised d_S estimates were strongly correlated with previous estimates from Wright et al. 2012 (Spearman's rank correlation P-value < 0.01, correlation coefficient 0.78). The d_S of three newly identified gametologs were also consistent with the previously identified strata, with overlapping 95% confidence intervals (Fig. 1, Table S1), in line with the previous findings that the *G. gallus* Z chromosome was formed by four independent recombination suppression events.

Anas platyrhynchos, M. gallopavo, *G. gallus,* and *T. guttata* (taeGut3.2.4) coding sequences were obtained from Ensembl version 74 (Flicek et al. 2013), and additional *M. gallopavo* sequences were obtained from Genbank (Backstrom et al. 2005; Berlin and Ellegren 2006). For each species the longest transcript for each gene was identified. *Gallus gallus* sequences were BLASTed reciprocally against *A. platyrhynchos, M. gallopavo,* and *T. guttata* coding sequences using an e-value cutoff of 1×10^{-10} and the highest BLAST score was used to identify the best BLAST hit.

For *A. platyrhynchos* and *M. gallopavo* W-linked genes, we identified orthologous Z-linked genes using the following steps. For each W-linked gene (1) we identified the orthologous *G. gallus* W-linked gene using BLAST and the *G. gallus* Z-linked ortholog from Wright et al. (2012); (2) *A. platyrhynchos* and *M. gallopavo* Z-linked orthologs were defined as the reciprocal best hit of the *G. gallus* Z-linked ortholog; (3) in a few instances there was no reciprocal best hit, often because of high sequence similarity between Z and W gametologs. For these genes, the Z-linked reciprocal ortholog was identified as the next best BLAST hit of the *G. gallus* Z-linked ortholog.

Avian genomes are particularly stable with few major genomic rearrangements (Stiglec et al. 2007) and little gene movement off and onto the avian Z chromosome, potentially due to a lack of active transposons (Toups et al. 2011). We therefore expect Z-linkage and synteny to be highly conserved across *A. platyrhynchos, M. gallopavo,* and *G. gallus*. We verified Z-linkage of reciprocal orthologs using positional information in the *M. gallopavo* genome assembly. The *A. platyrhynchos* genome assembly lacks annotated chromosomes and as recently diverged Z- and W-linked genes share a high degree of sequence similarity, to avoid misidentifying the Z-linked reciprocal ortholog as a W-linked paralog, we verified the reciprocal ortholog was not located on the W chromosome using PCR. However, this approach does not distinguish Z chromosome to autosomal gene duplication, but given the pronounced stability of avian genomes (Toups et al. 2011), this is unlikely to bias our results.

DIVERGENCE OF GAMETOLOGS

To identify the number and boundaries of evolutionary strata in *A. platyrhynchos* and *M. gallopavo*, we estimated rates of synonymous divergence (d_S) (Wright et al. 2012), as d_S provides a relative measure of the relative length of time gametologs have differentiated from an ancestral autosomal gene. For each species, we obtained Z and W coding sequences from Biomart, and aligned the translated protein sequences with MUSCLE (Edgar 2004a,b) in MEGA5 (Tamura et al. 2011). Alignments were visually inspected and poorly aligned regions were removed. The CODEML package in PAML version 4.7 (Yang 2007) was used to calculate maximum-likelihood estimates of synonymous (d_S) and nonsynonymous (d_N) divergence in pairwise comparisons and generate standard errors. We verified $d_S \leq 1$ for all gametologs, thereby avoiding inaccurate divergence estimates due to mutational saturation and double hits (Axelsson et al. 2008).

Three *A. platyrhynchos* gametologs had $d_S < 0.02$. We verified that the annotated Z and W genes were distinct coding

sequences by identifying female-limited SNPs using RNA-seq data and W-linked SNPs from Ensembl annotated sequences.

Some gametologs had multiple W paralogs, and to identify the true W-linked ortholog we used the method outlined in Wright et al. (2012). Briefly, we used the gametologous pair with the lowest d_S estimate to avoid potential problems of relaxed or diversifying selection acting on gene copies (Graur and Li 2000) likely due to gene duplications following the suppression of recombination (Busby et al. 2011). However, the *A. platyrhynchos* draft assembly contains short sequences that may represent partial fragments of longer W-linked genes. To avoid biasing divergence estimates on short stretches of sequence, we excluded W-linked genes if coverage of the *G. gallus* W-linked ortholog was <25% and the aligned length with the Z-linked ortholog was <150 bp. This resulted in the removal of three *A. platyrhynchos* W-linked genes from the analysis. All *M. gallopavo* W-linked genes met these criteria, and therefore none were excluded.

For each species, we mapped d_S estimates to the Z chromosome, taking physical position from the orthologous *G. gallus* Z-linked ortholog. Divergence dates for gametologs were calculated using a molecular clock. Briefly, based on male-mutation bias estimates in birds, the molecular clock accounts for sex-specific mutation rates in Z- and W-linked genes. This results in a divergence time of $d_S/(3.8 \times 10^{-9})$ years (Dimcheff et al. 2002b; Hedges 2002; Harlid et al. 2003; Axelsson et al. 2004; Berlin et al. 2006; Webster et al. 2006; Nam and Ellegren 2008).

We used *M. gallopavo* gametologs to independently confirm the four strata previously identified in *G. gallus* (Wright et al. 2012) and to refine the dates at which recombination was suppressed. We assessed longer term dynamics of sex chromosome divergence using *A. platyrhynchos* gametologs. To do so, we built gene trees to determine whether recombination ceased independently between gametologs in each species, or before the last common ancestor of the Galloanserae. Specifically we (1) obtained coding sequences for gametologs in the *G. gallus*, *A. platyrhynchos, M. gallopavo,* and *T. guttata* from Biomart, and aligned the translated protein sequences with MUSCLE (Edgar 2004a, b) in MEGA5 (Tamura et al. 2011). We checked alignments by eye and removed poorly aligned regions. *Taeniopygia guttata* Z-linked orthologs, identified from a reciprocal BLAST with *G. gallus*, were used as outgroups; (2) we constructed maximum-likelihood gene trees with 1000 bootstrap replicates in MEGA5 (Tamura et al. 2011). In the absence of overlapping sequences, gene trees were constructed between *G. gallus* and *A. platyrhynchos* or *M. gallopavo* gametologs separately. In all cases, the *T. guttata* Z-linked ortholog was designated as the outgroup.

To measure the completeness of recombination suppression between gametologs, we tested for evidence of gene conversion using GENECONV (Sawyer 1999). GENECONV identifies potential gene conversion sites as long fragments of identical sequence bounded by variable sites, assessing significance by 10,000 random permutations of variable sites across the alignment. *P*-values for global fragments are corrected for multiple comparisons and sequence length. We tested for gene conversion using the multiple sequence alignments described above (used to build gene trees) and specified the *–group* option to look only for conversion between gametologs in *G. gallus*, *A. platyrhynchos,* and *M. gallopavo*. We used the –Seqtype = SILENT option to specify coding sequence and that gene conversion fragments should only be detected using silent site amino acid polymorphisms. High false-positive rates have been documented when this option is not implemented, particularly if genes are subject to contrasting selective regimes (Ezawa et al. 2006, 2010). We used the strict default *-gscale* value (*gscale* = 0) where no mismatches are permitted between aligned putative conversion fragments and repeated the analysis with a relaxed mismatch penalty (*gscale* = 2) to ensure the number of mismatches did not bias our analysis. GENECONV detects inner fragments, which arise from gene conversion events between the ancestors of two aligned sequences and outer fragments, which represent a conversion events originating outside of the alignment or events whose signature has been weakened by subsequent mutation. We identified significant inner gene conversion fragments between gametologs as evidence of ongoing recombination between the sex chromosomes.

On closer inspection, gene conversion events identified in this study span several exons that in turn are separated by extremely dissimilar intronic sequences. As GENECONV identifies potential gene conversion sites as long fragments of sequence bounded by variable sites, it estimates the maximum sequence length over which recombination acts. To avoid falsely identifying gene conversion sites, we excluded exons where GENECONV fragments span less than 50 bp. We used informative polymorphic sites and maximum-likelihood exon trees to establish the direction of gene conversion.

SEQUENCE EVOLUTION OF Z- AND W-LINKED GENES

To assess the selective regime acting on Z and W coding sequences, we used the branch models and branch-site test in the CODEML package in PAML version 4.7 (Yang 2007).

Branch models (model = 2, nssites = 0) allow ω (d_N/d_S) to vary across branches. By comparing a model where ω is estimated to one where it is fixed to equal 1 (the expected d_N/d_S under neutral evolution) or 0 (strict purifying selection) it is possible to assess the contribution of neutral and purifying selection to coding sequence evolution. To do this, we used the multiple sequence alignments described above (used to build gene trees and test for gene conversion). If *A. platyrhynchos* and *M. gallopavo* gametologs ceased recombining before the last common ancestor of the Galloanserae the following gene tree was specified (*T. guttata Z,* (((*G. gallus* W, *M. gallopavo* W), *A. platyrhynchos* W),

((*G. gallus* Z, *M. gallopavo* Z), *A. platyrhynchos* Z))) and if *A. platyrhynchos* gametologs diverged independently the following gene tree was used (*T. guttata Z*, (((*G. gallus* W, *M. gallopavo* W), (*G. gallus* Z, *M. gallopavo* Z)), (*A. platyrhynchos Z, A. platyrhynchos W*))).

The branch-site model (model = 2, nssites = 2) allows ω to vary across branches and among sites and permits tests to detect positive selection acting on a subset of sites along specific lineages. Under this model, we estimated ω across W and Z branches separately and then tested whether ω was significantly different to 1 using the settings fix_omega = 1, omega = 1.

Results

CHARACTERISING THE CODING CONTENT OF THE AVIAN W CHROMOSOME ACROSS SPECIES

We used a combination of bioinformatics, molecular genetic, and phylogenetic methods to validate putative W-linked genes and identify previously unknown W-linked coding content in the wild turkey (*M. gallopavo*) and mallard duck (*A. platyrhynchos*). Our combined dataset comprised 14 W-linked genes in *M. gallopavo* and 21 in *A. platyrhynchos*, together with 27 previously identified W-linked genes in the red jungle fowl (*G. gallus*) (Table 1). Some W-linked genes are present in multiple copy number (Backstrom et al. 2005; Moghadam et al. 2012) and these W-linked genes correspond to seven gametologs in *M. gallopavo*, 11 gametologs in *A. platyrhynchos*, and 20 gametologs in *G. gallus*. To date, this is the largest cross-species dataset of W-linked coding sequences yet assembled.

FOUR Z CHROMOSOME STRATA ARE CONSERVED IN THE GALLIFORMES

Sex chromosome strata are identified from clusters of gametologs with similar synonymous (d_S) divergence estimates (Lahn and Page 1999; Skaletsky et al. 2003), which provide a measure of the relative length of time gametologs have been isolated from each other. Previously, we identified four evolutionary strata on the *G. gallus* Z chromosome, where recombination was suppressed progressively over 130 million years (Wright et al. 2012).

We used seven gametologs in *M. gallopavo* to independently verify the *G. gallus* strata. Z-linked physical position was conserved between *G. gallus* and *M. gallopavo* orthologs, and each stratum is represented by at least one *M. gallopavo* gametolog. For all gametologs, d_S estimates were consistent with orthologous *G. gallus* Z–W d_S estimates and 95% confidence intervals overlapped (Fig. 1A, Table S2). Maximum-likelihood gene trees of *G. gallus* and *M. gallopavo* gametologs, with the zebra finch (*T. guttata*) Z-linked ortholog included as an outgroup, indicate that all four strata predate the *G. gallus–M. gallopavo* divergence estimated at roughly 30 MYA (Dimcheff et al. 2002a). For all seven gametologs, the gene trees clustered by sex chromosome rather than by species, and the phylogenetic grouping of *G. gallus* and *M. gallopavo* W-linked genes showed high bootstrap support (in all cases ≥95%, Figs. 1A, S2), indicating that gametolog divergence predates the common ancestor of *G. gallus* and *M. gallopavo*. This provides independent verification that the four recombination suppression events previously identified in *G. gallus* are conserved across the Galliformes.

Divergence dates for recombination suppression in the youngest strata should predate the divergence between *G. gallus* and *M. gallopavo* approximately 30 Mya (Dimcheff et al. 2002a). Using a molecular clock that accounts for sex-specific mutation rates in Z- and W-linked genes (Wright et al. 2012), we found that Stratum III, corresponding to 17–43 Mb ($d_S = 0.156–0.268$), arose between 41 and 71 Mya and Stratum IV, corresponding to 0–11 Mb ($d_S = 0.137–0.257$), between 36 and 68 Mya (Fig. 1A). These dates are consistent with the finding that four strata are conserved in the Galliformes.

RECOMBINATION SUPPRESSION EVOLVED INDEPENDENTLY IN THE ANSERIFORMES

Using 11 newly identified *A. platyrhynchos* gametologs, we calculated d_S estimates and maximum-likelihood gene trees to assess the dynamics of recombination suppression and sex chromosome divergence. Avian genomes are relatively stable, with few major genomic rearrangements, potentially due to a lack of active transposons (Toups et al. 2011). As synteny of the Z chromosome is highly conserved among all extant birds (Vicoso et al. 2013b), including the Galloanserae *A. platyrhynchos* and *G. gallus* (Skinner et al. 2009), we took physical position of the *A. platyrhynchos* gametologs from the location of the corresponding *G. gallus* Z-linked ortholog. To ensure W-linked paralogs were not misidentified as Z-linked orthologs, we used PCR to confirm they were not physically female-limited.

For the three *A. platyrhynchos* gametologs located in the two oldest Galliform strata, d_S estimates were consistent with orthologous *G. gallus* Z–W d_S estimates, where 95% confidence intervals overlapped (Fig. 1B, Table S3). Maximum-likelihood gene trees for all loci cluster by sex chromosome rather than by species, with 100% bootstrap support for the grouping of *G. gallus* and *A. platyrhynchos* W-linked orthologs distinct from Z-linked genes (Figs. 1B, S3). For *RASA1Z* and *RASA1W*, we have coding sequence for *G. gallus*, *A. platyrhynchos,* and *M. gallopavo* gametologs, and W-linked sequences cluster together with 100% bootstrap support. The three gametolog sets therefore convergently indicate that the two oldest strata are conserved across the Galloanserae. These strata correspond to 45–51 Mb (d_S = 0.285–0.404, Conserved Stratum I), 54–61 Mb

Table 1. Known and newly identified *G. gallus*, *A. platyrhynchos,* and *M. gallopavo* W-linked coding genes.

Name	Z-linked gene ID[1]	W-linked gene ID[1]
G. gallus		
HINT1	00428	22674,[2] 22683,[3] 22690,[3] 22679,[3]
CHD1	14642	*CHDW*[3]
RASA1	17706	22611[2,3]
KCMF1	15391	14441
MIER3	14721	00140,[2,3] 14584[2,3]
SPIN	14916	*SPINW*[3]
HNRNPK	12591	14366[2,3]
CZH18ORF25	01763	01585[2,3]
VCP	01986	00386[2]
SIAT8C	03049	26991
ZFR	03235	14545[2,3]
NIPBL	03605	13312[2,3]
UBAP2	13809	05785
ATP5A1	14644	01756[2,3]
MADH2	14697	14184
BTF3	13512	00395[2,3]
UBE2R2	01668	09227
ZSWIM6	14734	27170
ZNF532	02852	14003[3]
RPL17L	02696	22174[2,3]
Novel	-	29099
Novel	-	25736
Novel	-	25865
M. gallopavo		
HINT1	06655	16898, 09917, *HINTW*(AY713488.1)[4]
RASA1	08292	*RASA1W*(AH015047.1)[5]
SPIN	02500	*SPINW*(AH015050.1),[5] 06743
VCP	-[6]	09037
NIPBL	01989	13606, *IDN3W*(AH015048.1)[5]
UBAP2	01837	*UBAP2W*(AY188758.1),[5] 05595
ATP5A1	01414	09963
MADH2	*MADH2Z*[5]	*MADH2W*(AH015049.1)[5]
Novel	-	16394
A. platyrhynchos		
CHD1	09965	05191, 02506
RASA1	05627	05611, 11371, 10611
KCMF1	13426	03026, 03106
MIER3	06634	10850
SPIN	13922	02923
HNRNPK	10856	10986
VCP	06634	05806
ZFR	15627	15519
NIPBL	08473	10290, 05315, 10560, 02953, 03022
UBE2R2	03800	16000
ZSWIM6	06992	13555, 14338
UBAP2	06449	16155

[1]Ensgalg000000(*G. gallus*)/Ensmgag000000(*M. gallopavo*)/Ensaplg000000(*A. platyrhynchos*).

[2]Moghadam et al. (2012).

[3]Ayers et al. (2013).

[4]Backstrom et al. (2005).

[5]Berlin and Ellegren (2006).

[6]Amalgamation of Z and W sequences in current genome assembly.

(d_S = 0.171–0.271, Conserved Stratum II), and arose between 75 and 106 Mya, and 45 and 71 Mya, respectively.

In contrast, of the eight *A. platyrhynchos* gametologs located in the youngest Galliform strata, d_S estimates for five gametologs are significantly lower than the corresponding *G. gallus* gametolog d_S estimates, where 95% confidence intervals are nonoverlapping (Fig. 1B, Table S3). Of these eight *A. platyrhynchos* gametologs, four cluster together in maximum-likelihood gene trees with bootstrap values ≥95, indicating that recombination suppression occurred relatively recently and independently in the Anseriformes in this region. However, owing to low bootstrap values we were unable to resolve the phylogenetic topology for the other four gametologs. Using a molecular clock, we calculated that recombination ceased independently in Anseriformes between 0 and 39 Mya (Fig. 1B), well after the divergence of Galliformes and Anseriformes approximately 90 Mya (van Tuinen and Hedges 2001).

HETEROGENEOUS DIVERGENCE OF YOUNG ANSERIFORM GAMETOLOGS

There is marked heterogeneity in d_S across the youngest region of the Anseriform sex chromosomes, ranging from 0 to 0.148. This variation is greater than that observed across any other avian strata, and is not a result of a single outlier. Three gametologs have $d_S < 0.02$, indicating either very recent divergence or persistent recombination. We verified that Z- and W-linked coding sequences were distinct using female-limited SNPs identified from RNA-seq data and W-linked SNPs from Ensembl annotated sequences. Therefore, the heterogeneity in divergence is not consistent with instantaneous suppression of recombination resulting from a large-scale intra-chromosomal event such as an inversion. Instead, this may suggest more gradual divergence of gametologs with residual recombination across large tracts of the sex chromosomes.

To further assess the possibility of ongoing recombination between gametologs, we identified signatures of gene conversion between Z- and W-linked coding sequences in *G. gallus*, *A. platyrhynchos,* and *M. gallopavo* using GENECONV (Sawyer 1999). Three gametologs in *A. platyrhynchos* showed evidence of significant inner gene conversion following a permutation test with 10,000 iterations (Table 2). We repeated the analysis allowing mismatches between gene conversion fragments (*gscale* = 2) and found no significant difference between the results. These gametologs are uniformly distributed across the Z chromosome, located in Conserved Strata I, II, and Anseriform-specific Stratum III, indicating ongoing recombination along the whole length of the avian sex chromosomes. In total, we detected four Z- and four W-linked exons subject to gene conversion (Table 2). Of these, we could determine the direction of gene conversion for the exons in the Conserved Strata I and II. Informative variant sites indicate genetic material was transferred from *KCMF1W* to *KCMF1Z* and from *CHD1Z* to *CHD1W*. Exon tree topologies support the proposed direction of gene conversion in *KCMF1* and cluster by species for *CHD1*, revealing putative gene conversion across *G. gallus CHD1* gametologs, the signal of which was not detected with GENECONV.

Gene conversion between gametologs can lower d_S estimates and generate molecular gene trees where gametologs cluster by species. This may lead to false conclusions that recombination was suppressed independently in a given lineage. We detected no gene conversion between the four *A. platyrhynchos* gametologs where gene trees cluster by species, or the three gametologs with $d_S <$ 0.02. However, previous studies have shown that GENECONV's power to detect gene conversion events between similar sequences is limited (McGrath et al. 2009; Ezawa et al. 2010).

We recalculated d_S for three *A. platyrhynchos* gametologs after removing exons subject to gene conversion and as expected, the estimated divergence time increased. However, the range of divergence dates over which recombination was suppressed in Conserved Stratum I and II remained constant. The recalculated gene tree for the gametolog in Anseriform-specific Stratum III revealed a complex mosaic of recombination across specific exons. Before excluding the exons subject to gene conversion, the phylogenetic topology was unresolvable, however after removal of these exons, *A. platyrhynchos* and *G. gallus HNRNPKW* clustered together with high bootstrap support (≥95%).

COMPARATIVE ANALYSIS OF W-LINKED CODING CONTENT EVOLUTION

Studies examining the evolutionary forces acting after recombination suppression are mainly limited to male heterogametic systems (Chibalina and Filatov 2011; Hughes et al. 2012; Bachtrog 2013). As a result, the factors driving the differentiation and degeneration of Z and W sex chromosomes are partially unclear. Now equipped with the largest dataset of W-linked genes to date, we used the branch models and branch-site test in the CODEML package in PAML version 4.7 (Yang 2007) to assess the selective regime driving Z and W coding sequence change over 90 million years of avian evolution.

For W-linked branches, ω (d_N/d_S) was significantly lower than 1 in 14 gametolog families, indicating these genes are evolving primarily due to purifying selection (Tables 3, S4). The exception was *HINT1W*, where ω was not significantly different to 1. However, *HINT1W* is a known ampliconic gene with documented gene conversion among W-linked copies (Backstrom et al. 2005). This likely violates the assumptions made in CODEML, where d_S is equivalent to the neutral mutation rate (Yang 2007). For Z-linked branches, ω was not significantly different to 0 across four loci, indicating these are extremely conserved and evolving under strict purifying selection. For the other 10 loci, ω was

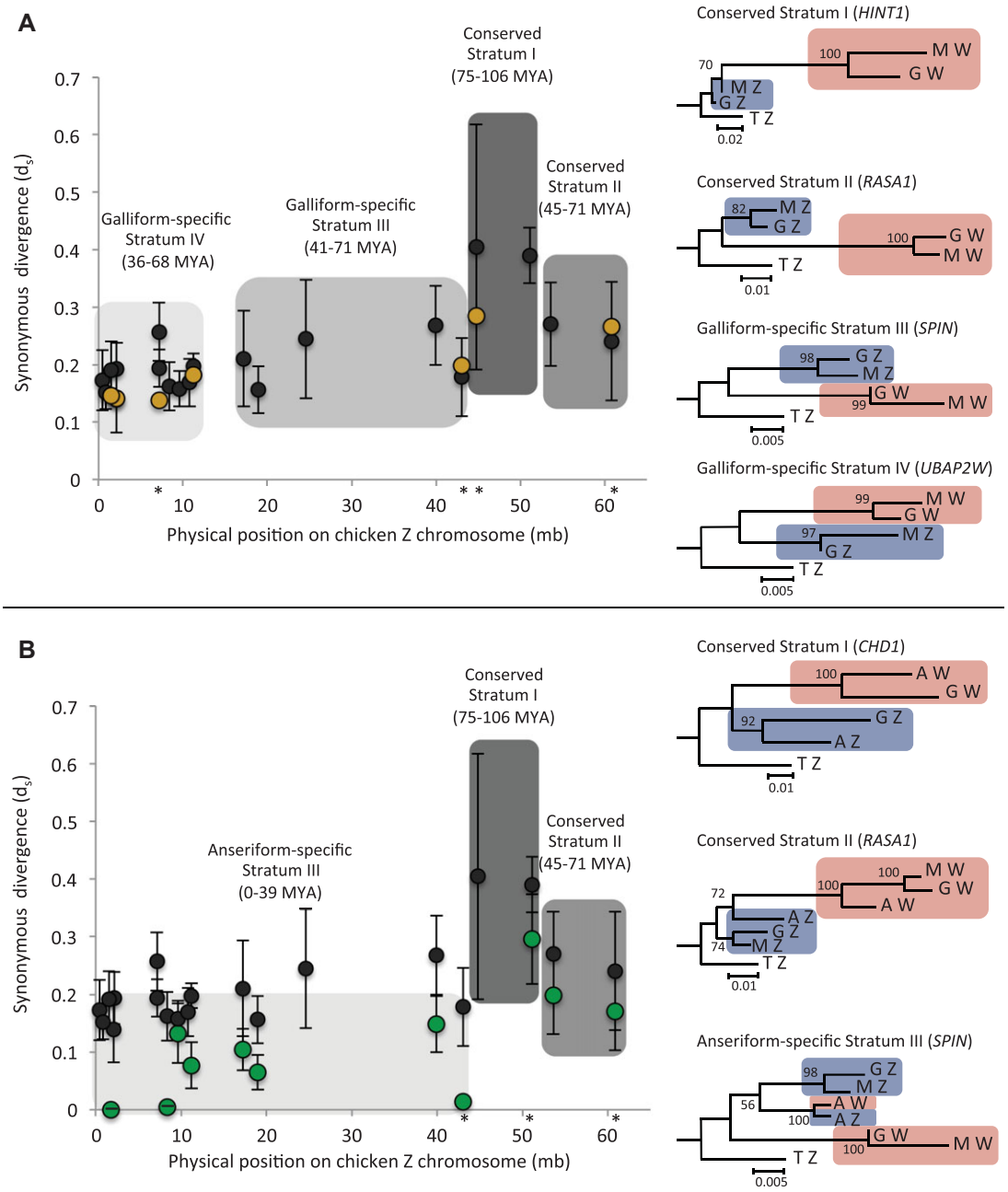

Figure 1. Dynamics of recombination suppression on the avian sex chromosomes. The distribution of synonymous (d_S) divergence estimates between *G. gallus* (black), *M. gallopavo* (yellow, panel A) and *A. platyrhynchos* (green, panel B) gametologs, mapped against physical position of the *G. gallus* Z-linked ortholog. The 95% confidence intervals were calculated using standard errors. Bootstrap values were calculated using 1000 permutations. AZ, GZ, MZ, TZ corresponds to *A. platyrhynchos*, *G. gallus*, *M. gallopavo,* and *T. guttata* Z-linked genes. The physical Z-linked position of the loci for which gene trees are shown are indicated (*). *Meleagris gallopavo* d_S estimates map closely onto the *G. gallus* strata boundaries, and maximum-likelihood gene trees for all gametologs cluster by sex chromosome, not species, indicating that four strata are conserved across the Galliformes (Fig. S2). *Anas platyrhynchos* d_S estimates map closely onto the two oldest *G. gallus* strata boundaries, and maximum-likelihood gene trees for gametologs within these regions cluster by sex chromosome indicating that these strata are conserved across the Galloanserae. For the remaining gametologs, maximum-likelihood gene trees reveal that recombination was suppressed independently in the *A. platyrhynchos* lineage (Fig. S3). The d_S estimates are highly heterogeneous and locus specific, indicating gradual divergence of gametologs with residual recombination across large tracts of the sex chromosomes.

Table 2. Gene conversion between gametologs in *A. platyrhynchos*.

Gene name	Gene ID (Z/W)[a]	Stratum	*P*-value [b/c]	Spanning Z-linked exon no.[d]	Spanning W-linked exon no.[d]	Length of gene conversion track (bp)
CHD1	09965/ 05191	Conserved I	0.0006/ 0.0027*	27	7	174
KCMF1	13426/ 03026	Conserved II	0.0056/ 0.0273	4	3	142
HNRNPK	10856/ 10986	Anseriform-specific III	0.0081/ 0.0294	10	8	255
HNRNPK	10856/ 10986	Anseriform-specific III	0.0314/ 0.0775	18	13	159

[a] Ensaplg000000.
[b] Simulated *P*-values are based on 10,000 permutations.
[c] Karlin–Altshul *P*-values are Bonferroni-corrected KA (BLAST-like).
[d] Exons where gene conversion fragments span <50 bp are not listed.
* Significant after correcting for multiple comparisons.

significantly different to both 1 and 0 for all but one locus, *UBAP2Z* (Table S5).

We tested for positive selection using the branch-site model. This model allows ω to vary across both branches and among sites, and tests for adaptive evolution acting on a subset of codons along a branch(es) defined a priori (Zhang et al. 2005). We found no evidence of positive selection on any Z-linked gene families (Table S5) but found significant evidence for positive selection acting on one W-linked loci, *ATP5A1W* (proportion of sites where $\omega > 1$ was 0.012 and $\omega = 17.70$) (Tables 3, S4).

Gene conversion violates the assumption in PAML that d_S is constant across a gene and the GC bias in the mismatch repair process can inflate ω (d_N/d_S) (Berglund et al. 2009). Consequentially, previous studies have found that gene conversion can lead to false signals of positive selection (Casola and Hahn 2009; Ratnakumar et al. 2010). Although we found no evidence of gene conversion between *ATP5A1W* and *ATP5A1Z* in any species, *ATP5A1W* is present in two copies in the *G. gallus* genome (Moghadam et al. 2012) and may be subject to intra-chromosomal gene conversion. In contrast, our polymorphism data indicate there is only a single copy in the *M. gallopavo* genome, as we failed to identify polymorphisms in this locus within any single female, which would indicate multiple copy number. To exclude the possibility that gene conversion between W-linked *G. gallus* paralogs is driving a false signal of positive selection, we removed *G. gallus ATP5A1W* and repeated the analysis. After excluding this gene, the branch-site test for positive selection remained significant (proportion of sites where $\omega > 1$ was 0.006 and $\omega = 41.87$). *ATP5A1W* encodes a mitochondrial ATP synthase subunit. As the W chromosome and mitochondrial genome are both maternally inherited and effectively linked (Berlin et al. 2007), one plausible explanation for this signature of positive evolution may be that it reflects adaptive coevolution.

Discussion

The process of recombination suppression is a crucial component of sex chromosome evolution, facilitating the divergence and differentiation of gametologs (Charlesworth and Charlesworth 1978). Studies from highly differentiated sex chromosomes suggest that recombination is suppressed instantaneously in a stepwise manner by intra-chromosomal rearrangements (Lahn and Page 1999; Pandey et al. 2013). However, our ability to determine whether inversions are a cause or consequence of this process is hampered by the limited numbers of sex-linked genes in these species. Conversely, in newly evolved sex chromosomes, recombination appears to be suppressed in a gradual and heterogeneous nature (Chibalina and Filatov 2011; Muyle et al. 2012; Bergero et al. 2013; Natri et al. 2013; Qiu et al. 2013).

We used a comparative approach to study the dynamics of sex chromosome recombination suppression on the largest collection of W-linked coding content yet assembled across 90 million years of avian evolutionary history. Our results reveal that recombination between the Z and W chromosomes ceased independently multiple times across different avian lineages. The pattern we uncover is consistent with both the inversion model (Charlesworth et al. 2005) where evolutionary strata form in a stepwise process, and heterogeneous and locus specific recombination suppression over short-term periods. Additionally, we show that recombination suppression is not necessarily complete, as highly diverged gametologs recombine across specific exons via gene conversion.

DYNAMICS OF SEX CHROMOSOME DIVERGENCE

Using both maximum-likelihood gene trees and synonymous divergence estimates, we confirm that four strata are conserved across the Galliformes (Fig. 2). These four strata were previously identified in *G. gallus* (Wright et al. 2012), and follow the

Table 3. Branch models and branch-site test for W-linked branches.

Gene name	Species included[1]	ω	Branch model test (ω=1) likelihood ratio	Branch model test (ω=0) likelihood ratio	Branch-site test (ω =1) likelihood ratio
HINT1W	G, M, T	1.10	0.05	**3697.05****	0.05
CHD1W	G, A, T	0.14	**94.45****	**7012.15****	0.00
RASA1W	G, A, M, T	0.09	**31.33****	**1386.95****	0.00
KCMF1W	G, A, T	0.22	**35.49****	**5674.71****	0.22
MIER3W	G, A, T	0.24	**17.97****	**167.25****	0.00
SPINW	G, A, M, T	0.09	**28.45****	**1749.98****	0.00
HNRNPKW	G, A, T	0.03	**108.75****	**84.21****	0.03
VCPW	G, A, T	0.03	**25.58****	**19.13****	0.16
ZFRW	G, A, T	0.18	**27.16****	**173.50****	0.00
NIPBLW	G, A, T	0.05	**44.08****	**63.91****	0.00
NIPBLW	G, M, T	0.03	**38.40****	**343.5****	0.00
UBAP2W	G, M, T	0.15	**6.19***	**354.82****	0.00
ATP5A1W	G, M, T	0.09	**16.54****	**650.83****	**3.58***
MADH2W	G, M, T	0.06	**22.37****	**333.50****	0.01
UBE2R2W	G, A, T	0.04	**15.03****	**17.39****	0.00
ZSWIM6W	G, A, T	0.25	**11.25****	**128.88****	0.00

[1]G = *G. gallus*, M = *M. gallopavo*, A = *A. platyrhynchos*, T = *T. guttata.*

***P*-value < 0.01.

**P*-value < 0.05.

Likelihood values for each model and the gene IDs are listed in Table S4. Significant likelihood ratios are in bold.

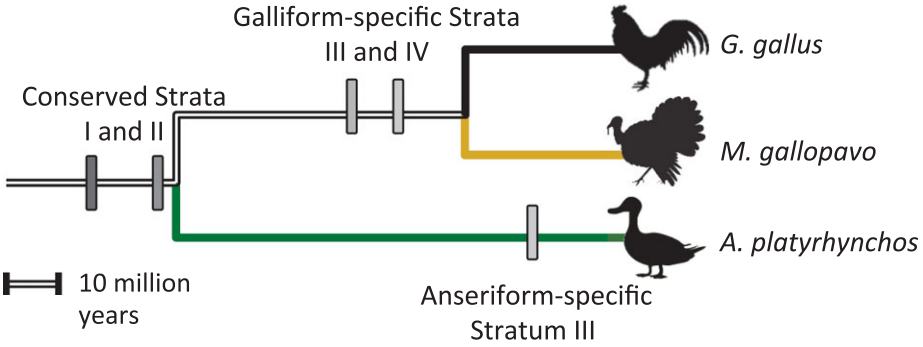

Figure 2. **Independent expansion of the nonrecombining regions across the Galloanserae. Phylogeny illustrating the independent cessation of recombination on the sex chromosomes in the Galliformes and Anseriformes. Estimated date of recombination suppression (millions of years), calculated using a molecular clock that accounts for sex-specific mutation rates in Z- and W-linked genes, is mapped onto the phylogeny.**

same stepwise pattern observed across the mammalian X chromosome (Lahn and Page 1999; Sandstedt and Tucker 2004; Pearks Wilkerson et al. 2008; Van Laere et al. 2008). Using newly identified gametologs, we were able to refine the dates at which recombination was suppressed across the Galliform strata. We find that the oldest strata originated approximately 90 Mya, and the youngest 46 Mya, both of which are consistent with the divergence date between *G. gallus* and *M. gallopavo* (Dimcheff et al. 2002a).

Similarly, maximum-likelihood gene trees reveal that the two oldest Galliform strata are conserved on the Anseriform Z chromosome. In contrast, recombination suppression is heterogeneous across the remainder of the *A. platyrhynchos* sex chromosomes, and the topology of gene trees revealed that gametologs diverged independently in this region (Fig. 2). Divergence estimates vary from 0 to 39 Mya, a range that is greater than that observed across any other avian strata, and not a result of a single outlier. We verified that Z- and W-linked sequences were in fact distinct using polymorphism data and Ensembl coding sequences. This heterogeneity in divergence is not consistent with instantaneous suppression of recombination resulting from a large-scale intrachromosomal event and is instead indicative of Z–W divergence combined with ongoing recombination. However, due to the limited number of gametologs, we lack sufficient power to formally test for differences in the variance of d_S across strata.

The lack of positional information for the *A. platyrhynchos* Z chromosome means we cannot rule out the possibility that Anseriform-specific Stratum III is actually made up of two distinct strata formed by a lineage specific inversion. Under this scenario, the three gametologs with d_S estimates <0.02 would comprise the youngest stratum with ongoing recombination between the W and Z orthologs. Positional information may support this scenario, as two of the three gametologs are located between 0 and 10 Mb on the Z chromosome.

Some gametologs are subject to gene conversion, lending further support to the notion that recombination suppression is not complete. Three *A. platyrhynchos* gametologs showed evidence of significant gene conversion, and were distributed across the length of the Z chromosome. Therefore, even for highly differentiated gametologs, certain exons are prevented from diverging. This is consistent with previous studies documenting gene conversion between X- and Y-linked orthologs in mammals (Iwase et al. 2010; Trombetta et al. 2014). Gene conversion may therefore be an important mechanism halting the degeneration of sex-limited gametologs (Rosser et al. 2009) in addition to the intra-chromosomal gene conversion documented on the avian W and mammalian Y chromosomes (Backstrom et al. 2005; Marais et al. 2010; Hallast et al. 2013; Bellott et al. 2014). Interestingly, a previous study used whole chromosome painting to show that the *A. platyrhynchos* W chromosome is highly divergent in comparison to closely related avian W chromosomes (Stiglec et al. 2007). We find a high degree of conservation among coding genes, indicating exon-specific gene conversion may facilitate large-scale turnover of intronic and repetitive regions while still conserving gene function.

This heterogeneous pattern of sex chromosome divergence and ongoing recombination is consistent with *G. aculeatus* and *S. latifolia* sex chromosomes. The *S. latifolia* XY system originated 10 Mya, and a neo-sex chromosome formed 2 Mya in *G. aculeatus*. In both these system, recombination suppression evolves heterogeneously (Chibalina and Filatov 2011; Muyle et al. 2012; Bergero et al. 2013; Natri et al. 2013; Qiu et al. 2013; Roesti et al. 2013). The underlying mechanism of recombination suppression is unclear, but changes in heterochromatin and methylation, which have been shown to alter patterns of recombination (Mirouze et al. 2012), or local changes in the binding sequence that recruits recombination machinery to the chromosome (Baudat et al. 2010; Myers et al. 2010), may be responsible.

Our divergence estimates for recombination suppression are consistent with divergence dates for *G. gallus*, *A. platyrhynchos,* and *M. gallopavo* with the exception of Conserved Stratum II. Using a molecular clock that accounts for sex-specific mutation rates in Z- and W-linked genes, the divergence of Conserved Stratum II was estimated to have occurred between 45 and 71 Mya. Galliformes and Anseriformes diverged 90 Mya (van Tuinen and Hedges 2001). Although these estimates are not vastly distant, they are inconsistent and may indicate that the molecular clock underestimates the time of gametolog divergence due to the unique evolutionary forces acting on sex chromosomes (Bachtrog et al. 2011).

EVOLUTION OF HETEROGAMETIC SEX CHROMOSOMES AFTER RECOMBINATION SUPPRESSION

Studies examining the evolutionary forces acting after recombination has ceased have mainly been limited to male heterogametic systems, such as the Y chromosomes of mammals and *Drosophila* (Chibalina and Filatov 2011; Hughes et al. 2012; Bachtrog 2013). These studies have uncovered strong purifying selection acting on the Y chromosome (Kaiser et al. 2011; Hughes et al. 2012; Bellott et al. 2014; Wilson Sayres et al. 2014) together with signatures of positive selection (Gerrard and Filatov 2005; Li et al. 2013). However, the selective forces driving the evolution of the W chromosome are unclear. This is partly due to a previous paucity of known W-linked genes, which has restricted the potential of interspecific analyses (Berlin and Ellegren 2006).

Theoretical models predict key differences between W and Y chromosome evolution (Kirkpatrick and Hall 2004; Mank 2012; Wright and Mank 2013). When sexual selection is acting primarily on males, the Y chromosome has a lower effective population size than W-linked sequences (Wright and Mank 2013), and is subject to male-mutation bias (Kirkpatrick and Hall 2004). These factors have been hypothesized to result in rapid degeneration of Y-linked coding sequences, a pattern that has been observed on *Drosophila* and primate Y chromosomes (Hughes et al. 2010, 2012; Bachtrog 2013).

Using the largest catalog of W-linked genes to date, we find that avian W-linked genes are evolving with a significant contribution of purifying selection. This is consistent with a previous study showing that sex-specific selection drives evolutionary changes in the expression of W-linked genes (Moghadam et al. 2012), and indicates that although the W chromosome does not determine sex directly (Smith et al. 2009), it has an important role in encoding sex-specific fitness. The large number of genes conserved across 90 million years of avian evolution is also potentially indicative of lower degeneration rates on the W compared to the Y chromosome.

We found evidence of positive selection in one W-linked gene that encodes a mitochondrial ATP synthase subunit. As the W chromosome and mitochondria are both maternally inherited and are effectively linked (Berlin et al. 2007), we might expect strong coevolution between these portions of the genome, and this may be reflected by rapid divergence in sequence across different avian lineages.

Our findings indicate that adaptive evolution together with conservation of gene sequence is not restricted to male heterogametic systems and may be a fundamental feature of heterogametic sex chromosome evolution.

CONCLUDING REMARKS

Recombination suppression is a key force driving sex chromosome evolution, and sex chromosomes are therefore important for the study of genome evolution, particularly the interplay between selective forces and recombination. Here, we use a comparative approach to examine short- and long-term dynamics of sex chromosome evolution. Our results reveal that recombination has ceased multiple and independent times over 90 million years of avian evolution. We show that sex chromosome divergence over the long term is characterized by regions of similar divergence among gametologs, consistent with the inversion model of sex chromosome evolution (Lahn and Page 1999; Charlesworth et al. 2005; Vicoso et al. 2013a). In contrast, over short time periods we observe heterogeneous and locus specific divergence. Additionally, we uncover gene conversion between highly diverged gametologs across both long and short evolutionary trajectories. Moreover, if sex chromosomes do evolve via inversions, additional mechanisms may be required to completely suppress recombination. Our data also show that the W chromosome is evolving primarily due to purifying selection, although we also detect the signature of positive selection at one locus. Our findings indicate that the W chromosome plays an important role in encoding sex-specific fitness, and potentially exhibits a lower rate of degeneration compared to Y chromosome.

ACKNOWLEDGMENTS

We thank V. Roberts and J. Hall for collection assistance, G. Parker for assistance in securing samples, Z. Yang for helpful comments on the analyses conducted in PAML, and D. Burt for providing *A. platyrhynchos* coding sequences in advance of the release of the full genome annotation. We also thank J. Perry, F. Zimmer, V. Oostra, and three anonymous reviewers for helpful suggestions and comments on the manuscript. Sequencing was performed by the Wellcome Trust Centre for Human Genetics Sequencing Hub (funded by Wellcome Trust Grant reference 090532/Z09/Z and MRC Hub Grant G0900747 91070). The authors acknowledge the use of the UCL *Legion* High Performance Computing Facility (Legion@UCL) and UCL *Unity* SMP Facility, and associated support services, in the completion of this work. The work described here was funded by the European Research Council under the Framework 7 agreement (grant agreement 260233). AEW is funded by a National Environment Research Council

Ph.D. studentship. SHM is funded by a research fellowship from the Royal Commission for the Exhibition of 1851.

LITERATURE CITED

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.

Axelsson, E., N. G. C. Smith, H. Sundström, S. Berlin, and H. Ellegren. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. Mol. Biol. Evol. 21:1538–1547.

Axelsson, E., L. Hultin-Rosenberg, M. Brandström, M. Zwahlen, D. F. Clayton, and H. Ellegren. 2008. Natural selection in avian protein-coding genes expressed in brain. Mol. Ecol. 17:3008–3017.

Ayers, K. L., N. M. Davidson, D. Demiyah, K. N. Roeszler, F. Gruetzner, A. H. Sinclair, A. Oshlack, and C. A. Smith. 2013. RNA sequencing reveals sexually dimorphic gene expression before gonadal differentiation in chicken and allows comprehensive annotation of the W-chromosome. Genome Biol. 14:R24.

Bachtrog, D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat. Rev. Genet. 14:113–124.

Bachtrog, D., E. Hom, K. M. Wong, X. Maside, and P. de Jong. 2008. Genomic degradation of a young Y chromosome in *Drosophila miranda*. Genome Biol. 9:R30.

Bachtrog, D., M. Kirkpatrick, J. E. Mank, S. F. McDaniel, J. C. Pires, W. R. Rice, and N. Valenzuela. 2011. Are all sex chromosomes created equal? Trends Genet. 27:350–357.

Backstrom, N., H. Ceplitis, S. Berlin, and H. Ellegren. 2005. Gene conversion drives the evolution of HINTW, an ampliconic gene on the female-specific avian W chromosome. Mol. Biol. Evol. 22:1992–1999.

Baudat, F., J. Buard, C. Grey, A. Fledel-Alon, C. Ober, M. Przeworski, G. Coop, and B. de Massy. 2010. PRDM9 Is a major determinant of meiotic recombination hotspots in humans and mice. Science 327:836–840.

Bellott, D. W., J. F. Hughes, H. Skaletsky, L. G. Brown, T. Pyntikova, T.-J. Cho, N. Koutseva, S. Zaghlul, T. Graves, S. Rock, et al. 2014. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. Nature 508:494–499.

Bergero, R., S. Qiu, A. Forrest, H. Borthwick, and D. Charlesworth. 2013. Expansion of the pseudo-autosomal region and ongoing recombination suppression in the *Silene latifolia* sex chromosomes. Genetics 194:673–686.

Berglund, J., K. S. Pollard, and M. T. Webster. 2009. Hotspots of biased nucleotide substitutions in human genes. PLoS Biol. 7:45–62.

Berlin, S., and H. Ellegren. 2006. Fast accumulation of nonsynonymous mutations on the female-specific W chromosome in birds. J. Mol. Evol. 62:66–72.

Berlin, S., M. Brandstrom, N. Backstrom, E. Axelsson, N. Smith, and H. Ellegren. 2006. Substitution rate heterogeneity and the male mutation bias. J. Mol. Evol. 62:226–233.

Berlin, S., D. Tomaras, and B. Charlesworth. 2007. Low mitochondrial variability in birds may indicate Hill-Robertson effects on the W chromosome. Heredity 99:389–396.

Busby, M., J. Gray, A. Costa, C. Stewart, M. Stromberg, D. Barnett, J. Chuang, M. Springer, and G. Marth. 2011. Expression divergence measured by transcriptome sequencing of four yeast species. BMC Genomics 12:635.

Casola, C., and M. W. Hahn. 2009. Gene conversion among paralogs results in moderate false detection of positive selection using likelihood methods. J. Mol. Evol. 68:679–687.

Charlesworth, B. 1991. The evolution of sex chromosomes. Science 251: 1030–1033.

Charlesworth, B., and D. Charlesworth. 1978. A model for the evolution of dioecy and gynodioecy. Am. Nat. 112:975–997.

———. 2000. The degeneration of Y chromosomes. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 355:1563–1572.

Charlesworth, D., B. Charlesworth, and G. Marais. 2005. Steps in the evolution of heteromorphic sex chromosomes. Heredity 95:118–128.

Chen, N., D. W. Bellott, D. C. Page, and A. G. Clark. 2012. Identification of avian W-linked contigs by short-read sequencing. BMC Genomics 13:183. doi:10.1186/1471-2164-13-183.

Chibalina, M. V., and D. A. Filatov. 2011. Plant Y chromosome degeneration is retarded by haploid purifying selection. Curr. Biol. 21:1475–1479.

Dalloul, R. A., J. A. Long, A. V. Zimin, L. Aslam, K. Beal, L. A. Blomberg, P. Bouffard, D. W. Burt, O. Crasta, R. P. M. A. Crooijmans, et al. 2010. Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis. PLoS Biol. 8:e1000475.

Dalloul, R. A., A. V. Zimin, R. E. Settlage, S. Kim, and K. M. Reed. 2014. Next-generation sequencing strategies for characterizing the turkey genome. Poult. Sci. 93:479–484.

Dimcheff, D. E., S. V. Drovetski, and D. P. Mindell. 2002a. Phylogeny of *Tetraoninae* and other galliform birds using mitochondrial 12S and ND2 genes. Mol. Phylogenet. Evol. 24:203–215.

———. 2002b. Phylogeny of *Tetraoninae* and other galliform birds using mitochondrial 12S and ND2 genes. Mol. Phylogenet. Evol. 24:203–215.

Edgar, R. C. 2004a. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:1–19.

———. 2004b. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Ezawa, K., S. Oota, and N. Saitou. 2006. Genome-wide search of gene conversions in duplicated genes of mouse and rat. Mol. Biol. Evol. 23:927–940.

Ezawa, K., K. Ikeo, T. Gojobori, and N. Saitou. 2010. Evolutionary pattern of gene homogenization between primate-specific paralogs after human and macaque speciation using the 4–2–4 method. Mol. Biol. Evol. 27:2152–2171.

Flicek, P., I. Ahmed, M. R. Amode, D. Barrell, K. Beal, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fairley, et al. 2013. Ensembl 2013. Nucleic Acids Res. 41:D48–D55.

Gerrard, D. T., and D. A. Filatov. 2005. Positive and negative selection on mammalian Y chromosomes. Mol. Biol. Evol. 22:1423–1432.

Graur, D., and W.-H. Li. 2000. Fundamentals of molecular evolution. Sinauer Associates, Sunderland, MA.

Hallast, P., P. Balaresque, G. R. Bowden, S. Ballereau, and M. A. Jobling. 2013. Recombination dynamics of a human Y-chromosomal palindrome: rapid GC-biased gene conversion, multi-kilobase conversion tracts, and rare inversions. PLoS Genet. 9:e1003666.

Harlid, A. B., S. Berlin, N. G. C. Smith, A. P. Mosller, and H. Ellegren. 2003. Life history and the male mutation bias. Evolution 57:2398–2406.

Hedges, S. B. 2002. The origin and evolution of model organisms. Nat. Rev. Genet. 3:838–849.

Huang, Y., Y. Li, D. W. Burt, H. Chen, Y. Zhang, W. Qian, H. Kim, S. Gan, Y. Zhao, J. Li, et al. 2013. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. Nat. Genet. 45:776–783.

Hughes, J. F., H. Skaletsky, T. Pyntikova, T. A. Graves, S. K. M. van Daalen, P. J. Minx, R. S. Fulton, S. D. McGrath, D. P. Locke, C. Friedman, et al. 2010. Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. Nature 463:536–539.

Hughes, J. F., H. Skaletsky, L. G. Brown, T. Pyntikova, T. Graves, R. S. Fulton, S. Dugan, Y. Ding, C. J. Buhay, C. Kremitzki, et al. 2012. Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. Nature 483:82–86.

Iwase, M., Y. Satta, H. Hirai, Y. Hirai, and N. Takahata. 2010. Frequent gene conversion events between the X and Y homologous chromosomal regions in primates. BMC Evol. Biol. 10:225. doi:10.1186/1471-2148-10-225.

Kaiser, V. B., Q. Zhou, and D. Bachtrog. 2011. Nonrandom gene loss from the *Drosophila miranda* neo-Y chromosome. Genome Biol. Evol. 3:1329–1337.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S. L. Salzberg. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14:R36.

Kirkpatrick, M., and D. W. Hall. 2004. Male-biased mutation, sex linkage, and the rate of adaptive evolution. Evolution 58:437–440.

Lahn, B. T., and D. C. Page. 1999. Four evolutionary strata on the human X chromosome. Science 286:964–967.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9:357–359.

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10:R25.

Lemaitre, C., M. D. V. Braga, C. Gautier, M.-F. Sagot, E. Tannier, and G. A. B. Marais. 2009. Footprints of inversions at present and past pseudoautosomal boundaries in human sex chromosomes. Genome Biol. Evol. 1:56–66.

Li, G., B. W. Davis, T. Raudsepp, A. J. P. Wilkerson, V. C. Mason, M. Ferguson-Smith, P. C. O'Brien, P. D. Waters, and W. J. Murphy. 2013. Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. Genome Res. 23:1486–1495.

Lohse, M., A. M. Bolger, A. Nagel, A. R. Fernie, J. E. Lunn, M. Stitt, and B. Usadel. 2012. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res. 40:W622–W627.

Mank, J. E. 2012. Small but mighty: the evolutionary dynamics of W and Y sex chromosomes. Chromosome Res. 20:21–33.

Marais, G. A. B., P. R. A. Campos, and I. Gordo. 2010. Can intra-Y gene conversion oppose the degeneration of the human Y chromosome? A simulation study. Genome Biol. Evol. 2:347–357.

McGrath, C. L., C. Casola, and M. W. Hahn. 2009. Minimal effect of ectopic gene conversion among recent duplicates in four mammalian genomes. Genetics 182:615–622.

Mirouze, M., M. Lieberman-Lazarovich, R. Aversano, E. Bucher, J. Nicolet, J. Reinders, and J. Paszkowski. 2012. Loss of DNA methylation affects the recombination landscape in *Arabidopsis*. Proc. Natl. Acad. Sci. USA 109:5880–5885.

Moghadam, H. K., M. A. Pointer, A. E. Wright, S. Berlin, and J. E. Mank. 2012. W chromosome expression responds to female-specific selection. Proc. Natl. Acad. Sci. USA 109:8207–8211.

Muyle, A., N. Zemp, C. Deschamps, S. Mousset, A. Widmer, and G. A. B. Marais. 2012. Rapid de novo evolution of X chromosome dosage compensation in *Silene latifolia*, a plant with young sex chromosomes. PLoS Biol. 10:e1001308.

Myers, S., R. Bowden, A. Tumian, R. E. Bontrop, C. Freeman, T. S. MacFie, G. McVean, and P. Donnelly. 2010. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science 327:876–879.

Nam, K., and H. Ellegren. 2008. The chicken (*Gallus gallus*) Z chromosome contains at least three nonlinear evolutionary strata. Genetics 180:1131–1136.

Natri, H. M., T. Shikano, and J. Merila. 2013. Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. Mol. Biol. Evol. 30:1131–1144.

Ohno, S. 1967. Sex chromosomes and sex linked genes. Springer-Verlag, Berlin.

Pandey, R. S., M. A. Wilson Sayres, and R. K. Azad. 2013. Detecting evolutionary strata on the human X chromosome in the absence of gametologous Y-linked sequences. Genome Biol. Evol. 5:1863–1871.

Pearks Wilkerson, A. J., T. Raudsepp, T. Graves, D. Albracht, W. Warren, B. P. Chowdhary, L. C. Skow, and W. J. Murphy. 2008. Gene discovery and comparative analysis of X-degenerate genes from the domestic cat Y chromosome. Genomics 92:329–338.

Qiu, S., R. Bergero, and D. Charlesworth. 2013. Testing for the footprint of sexually antagonistic polymorphisms in the pseudoautosomal region of a plant sex chromosome pair. Genetics 194:663–672.

Ratnakumar, A., S. Mousset, S. Glémin, J. Berglund, N. Galtier, L. Duret, and M. T. Webster. 2010. Detecting positive selection within genomes: the problem of biased gene conversion. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 365:2571–2580.

R Core Team. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at http://www.R-project.org.

Roesti, M., D. Moser, and D. Berner. 2013. Recombination in the threespine stickleback genome–patterns and consequences. Mol. Ecol. 22: 3014–3027.

Ross, M. T., D. V. Grafham, A. J. Coffey, S. Scherer, K. McLay, D. Muzny, M. Platzer, G. R. Howell, C. Burrows, C. P. Bird, et al. 2005. The DNA sequence of the human X chromosome. Nature 434:325–337.

Rosser, Z. H., P. Balaresque, and M. A. Jobling. 2009. Gene conversion between the X chromosome and the male-specific region of the Y chromosome at a translocation hotspot. Am. J. Med. Genet. 85:130–134.

Sandstedt, S. A., and P. K. Tucker. 2004. Evolutionary strata on the mouse X chromosome correspond to strata on the human X chromosome. Genome Res. 14:267–272.

Sawyer, S. A. 1999. GENECONV: a computer package for the statistical detection of gene conversion. Distributed by the author www.math.wustl.edu/∼sawyer/geneconv/.

Schemper, M. 2003. Predictive accuracy and explained variation. Stat. Med. 22:2299–2308.

Skaletsky, H., T. Kuroda-Kawaguchi, P. J. Minx, H. S. Cordum, L. Hillier, L. G. Brown, S. Repping, T. Pyntikova, J. Ali, T. Bieri, et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825–837.

Skinner, B. M., L. B. W. Robertson, H. G. Tempest, E. J. Langley, D. Ioannou, K. E. Fowler, R. P. M. A. Crooijmans, A. D. Hall, D. K. Griffin, and M. Voelker. 2009. Comparative genomics in chicken and Pekin duck using FISH mapping and microarray analysis. BMC Genomics 10:357. doi:10.1186/1471-2164-10-357.

Smith, C. A., K. N. Roeszler, T. Ohnesorg, D. M. Cummins, P. G. Farlie, T. J. Doran, and A. H. Sinclair. 2009. The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature 461:267–271.

Stiglec, R., T. Ezaz, and J. A. M. Graves. 2007. A new look at the evolution of avian sex chromosomes. Cytogenet. Genome Res. 117:103–109.

Suh, A., J. O. Kriegs, J. Brosius, and J. Schmitz. 2011. Retroposon insertions and the chronology of avian sex chromosome evolution. Mol. Biol. Evol. 28:2993–2997.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731–2739.

Tian, L., T. Cai, E. Goetghebeur, and L. J. Wei. 2007. Model evaluation based on the sampling distribution of estimated absolute prediction error. Biometrika 94:297–311.

Toups, M. A., J. B. Pease, and M. W. Hahn. 2011. No excess gene movement is detected off the avian or lepidopteran Z chromosome. Genome Biol. Evol. 3:1381–1139.

Trapnell, C., L. Pachter, and S. L. Salzberg. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105–1111.

Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, and L. Pachter. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28:511–515.

Trapnell, C., D. G. Hendrickson, M. Sauvageau, L. Goff, J. L. Rinn, and L. Pachter. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat. Biotechnol. 31:46–53.

Trombetta, B., D. Sellitto, R. Scozzari, and F. Cruciani. 2014. Inter- and intra-species phylogenetic analyses reveal extensive X-Y gene conversion in the evolution of gametologous sequences of human sex chromosomes. Mol. Biol. Evol. 31:2108–2123.

Tsuda, Y., C. Nishida-Umehara, J. Ishijima, K. Yamada, and Y. Matsuda. 2007. Comparison of the Z and W sex chromosomal architectures in elegant crested tinamou (*Eudromia elegans*) and ostrich (*Struthio camelus*) and the process of sex chromosome differentiation in palaeognathous birds. Chromosoma 116:159–173.

Van Laere, A.-S., W. Coppieters, and M. Georges. 2008. Characterization of the bovine pseudoautosomal boundary: documenting the evolutionary history of mammalian sex chromosomes. Genome Res. 18:1884–1895.

van Tuinen, M., and S. B. Hedges. 2001. Calibration of avian molecular clocks. Mol. Biol. Evol. 18:206–213.

Vicoso, B., J. J. Emerson, Y. Zektser, S. Mahajan, and D. Bachtrog. 2013a. Comparative sex chromosome genomics in snakes: differentiation, evolutionarystrata, and lack of global dosage compensation. PLoS Biol. 11:e1001643.

Vicoso, B., V. B. Kaiser, and D. Bachtrog. 2013b. Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. Proc. Natl. Acad. Sci. USA 110:6453–6458.

Webster, M. T., E. Axelsson, and H. Ellegren. 2006. Strong regional biases in nucleotide substitution in the chicken genome. Mol. Biol. Evol. 23:1203–1216.

Wilson, M. A., and K. D. Makova. 2009. Evolution and survival on eutherian sex chromosomes. PLoS Genet. 5: e1000568.

Wilson Sayres, M. A., K. E. Lohmueller, and R. Nielsen. 2014. Natural selection reduced diversity on Human Y chromosomes. PLoS Genet. 10:e1004064.

Wright, A. E., and J. E. Mank. 2013. The scope and strength of sex-specific selection in genome evolution. J. Evol. Biol. 26:1841–1853.

Wright, A. E., H. K. Moghadam, and J. E. Mank. 2012. Trade-off between selection for dosage compensation and masculinization on the avian Z chromosome. Genetics 192:1433–1445.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.

Zhang, J., R. Nielsen, and Z. Yang. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol. Biol. Evol. 22:2472–2479.

Zhou, Q., and D. Bachtrog. 2012. Chromosome-wide gene silencing initiates Y degeneration in *Drosophila*. Curr. Biol. 22:522–525.

Associate Editor: M. Hahn

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. Identification of W-linked genes.
Figure S2. Gene trees for *M. gallopavo* gametologs.
Figure S3. Gene trees for *A. platyrhynchos* gametologs.
Table S1. Sequence divergence between *G. gallus* gametologs.
Table S2. Sequence divergence between *M. gallopavo* gametologs.
Table S3. Sequence divergence between *A. platyrhynchos* gametologs.
Table S4. Branch models and branch-site test across W-linked branches.
Table S5. Branch models and branch-site test across Z-linked branches.

References

Agrawal, A. F., and M. C. Whitlock. 2011. Inferences about the distribution of dominance drawn from yeast gene knockout data. Genetics 187:553-566.

Albert, A. Y. K., and S. P. Otto. 2005. Sexual selection can resolve sex-linked sexual antagonism. Science 310:119-121.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. Journal of Molecular Biology 215:403-410.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. H. Zhang, Z. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25:3389-3402.

Andersson, M. 1994. Sexual selection. Princeton University Press.

Andolfatto, P. 2008. Controlling type-I error of the McDonald-Kreitman Test in genomewide scans for selection on noncoding DNA. Genetics 180:1767-1771.

Arnqvist, G., and L. Rowe. 2005. Sexual conflict. Princeton University Press.

Arnqvist, G. 2011. Assortative mating by fitness and sexually antagonistic genetic variation. Evolution 65: 2111–2116.

Assis, R., Zhou, Q. & Bachtrog, D. 2012. Sex-biased transcriptome evolution in *Drosophila*. Genome Biology and Evolution 4: 1189–1200.

Axelsson, E., N. G. C. Smith, H. Sundstrom, S. Berlin, and H. Ellegren. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. Molecular Biology and Evolution 21:1538-1547.

Axelsson, E., L. Hultin-Rosenberg, M. Brandström, M. Zwahlen, D. F. Clayton, and H. Ellegren. 2008. Natural selection in avian protein-coding genes expressed in brain. Molecular Ecology 17:3008-3017.

Ayers, K. L., N. M. Davidson, D. Demiyah, K. N. Roeszler, F. Gruetzner, A. H. Sinclair, A. Oshlack, and C. A. Smith. 2013. RNA sequencing reveals sexually dimorphic gene expression before gonadal differentiation in chicken and allows comprehensive annotation of the W-chromosome. Genome Biology 14:R26.

Bachtrog, D. 2008a. The temporal dynamics of processes underlying Y chromosome degeneration. Genetics 179:1513-1525.

Bachtrog, D., E. Hom, K. M. Wong, X. Maside, and P. de Jong. 2008b. Genomic degradation of a young Y chromosome in *Drosophila miranda*. Genome Biology 9:R30.

Bachtrog, D., M. Kirkpatrick, J. E. Mank, S. F. McDaniel, J. C. Pires, W. Rice, and N. Valenzuela. 2011. Are all sex chromosomes created equal? Trends in Genetics 27:350-357.

Bachtrog, D. 2013. Y-chromosome evolution: Emerging insights into processes of Y-chromosome degeneration. Nature Review Genetics 14:113-124.

Backstrom, N., H. Ceplitis, S. Berlin, and H. Ellegren. 2005. Gene conversion drives the evolution of HINTW, an ampliconic gene on the female-specific avian W chromosome. Molecular Biology and Evolution 22:1992-1999.

Backstrom, N., E. Palkopoulou, A. Qvarnstrom, and H. Ellegren. 2010. No evidence for Z-chromosome rearrangements between the pied flycatcher and the collared flycatcher as judged by gene-based comparative genetic maps. Molecular Ecology 19:3394-3405.

Baines, J. F., S. A. Sawyer, D. L. Hartl, and J. Parsch. 2008. Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in *Drosophila*. Molecular Biology and Evolution 25:1639-1650.

Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov,S., Lee, L.J. et al. 2012. The evolutionary landscape of alternative splicing in vertebrate species. Science 338:1587–1593.

Bartosch-Harlid, A., S. Berlin, N. G. C. Smith, A. P. Moller, and H. Ellegren. 2003. Life history and the male mutation bias. Evolution 57:2398-2406.

Baudat, F., J. Buard, C. Grey, A. Fledel-Alon, C. Ober, M. Przeworski, G. Coop, and B. de Massy. 2010. PRDM9 Is a major determinant of meiotic recombination hotspots in humans and mice. Science 327:836-840.

Begun, D. J., and H. A. Lindfors. 2005. Rapid evolution of genomic Acp complement in the melanogaster subgroup of *Drosophila*. Molecular Biology and Evolution 22:2010-2021.

Begun, D. J., A. K. Holloway, K. Stevens, L. W. Hillier, Y-P. Poh, M. W. Hahn, P. M. Nista, C. D. Jones, A. D. Kern, C. N. Dewey, et al. 2007. Population genomics: Whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. PLoS Biology 5:2534-2559.

Bellott, D. W., J. F. Hughes, H. Skaletsky, L. G. Brown, T. Pyntikova, T-J. Cho, N. Koutseva, S. Zaghlul, T. Graves, S. Rock, et al. 2014. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. Nature 508:494-499.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate- A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B-Methodological 57:289-300.

Bergero, R., A. Forrest, E. Kamau, and D. Charlesworth. 2007. Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: Evidence from new sex-linked genes. Genetics 175:1945-1954.

Bergero, R., and D. Charlesworth. 2009. The evolution of restricted recombination in sex chromosomes. Trends in Ecology and Evolution 24:94-102.

Bergero, R., S. Qiu, A. Forrest, H. Borthwick, and D. Charlesworth. 2013. Expansion of the pseudo-autosomal region and ongoing recombination suppression in the *Silene latifolia* sex chromosomes. Genetics 194:673-686.

Berglund, J., K. S. Pollard, and M. T. Webster. 2009. Hotspots of biased nucleotide substitutions in human genes. PLoS Biology 7:45-62.

Berlin, S., and H. Ellegren. 2006. Fast accumulation of nonsynonymous mutations on the female-specific W chromosome in birds. Journal of Molecular Evolution 62:66-72.

Berlin, S., M. Brandstrom, N. Backstrom, E. Axelsson, N. Smith, and H. Ellegren. 2006. Substitution rate heterogeneity and the male mutation bias. Journal of Molecular Evolution 62:226-233.

Berlin, S., D. Tomaras, and B. Charlesworth. 2007. Low mitochondrial variability in birds may indicate Hill-Robertson effects on the W chromosome. Heredity 99:389-396.

Berlin, S., L. J. Qu, and H. Ellegren. 2008. Adaptive evolution of gamete-recognition proteins in birds. Journal of Molecular Evolution 67:488-496.

Bierne, N., and A. Eyre-Walker. 2003. The problem of counting sites in the estimation of the synonymous and nonsynonymous substitution rates: Implications for the correlation between the synonymous substitution rate and codon usage bias. Genetics 165:1587-1597.

Birkhead, T. R., and M. Petrie. 1995. Ejaculate features and sperm utilization in Peafowl *Pave Cristatus*. Proceedings of the Royal Society B-Biological Sciences 261:153-158.

Birkhead, T.R. & Fletcher, F. 1998. Sperm transport in the reproductive tract of female zebra finches *(Taeniopygia guttata*). Journal of Reproduction and Fertility 114: 141–145.

Birkhead, T., and A. P. Moller. 1998. Sperm competition and sexual selection. Academic Press Inc.

Blekhman, R., J. C. Marioni, P. Zumbo, M. Stephens, and Y. Gilad. 2009. Sex-specific and lineage-specific alternative splicing in primates. Genome Research 20:180-189.

Brawand, D., M. Soumillon, A. Necsulea, P. Julien, G. Csardi, P. Harrigan, M. Weier, A. Liechti, A. Aximu-Petri, M. Kircher, et al. 2011. The evolution of gene expression levels in mammalian organs. Nature 478:343-348.

Burt, D. W. 2002. Origin and evolution of avian microchromosomes. Cytogenetic and Genome Research 96:97-112.

Busby, M., J. Gray, A. Costa, C. Stewart, M. Stromberg, D. Barnett, J. Chuang, M. Springer, and G. Marth. 2011. Expression divergence measured by transcriptome sequencing of four yeast species. BMC Genomics 12:635.

Calkins, J. D., D. El-Hinn, and W. J. Swanson. 2007. Adaptive evolution in an avian reproductive protein: ZP3. Journal of Molecular Evolution 65:555-563.

Carrel, L., and H. F. Willard. 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature 434:400-404.

Casola, C., and M. W. Hahn. 2009. Gene conversion among paralogs results in moderate false detection of positive selection using likelihood methods. Journal of Molecular Evolution 68:679-687.

Chapman, T., Arnqvist, G., Bangham, J. & Rowe, L. 2003. Sexual conflict. Trends in Ecology and Evolution 9: 289–293.

Charif, D., and J. R. Lobry. 2007. SeqinR 1.0-2: A contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. Pp. 207-232. Structural approaches to sequence evolution: Molecules, networks, populations. Springer Verlag.

Charlesworth, B. 1978. Model for evolution of Y chromosomes and dosage compensation. Proceedings of the National Academy of Sciences, USA 75:5618-5622.

Charlesworth, B., and D. Charlesworth. 1978. A model for the evolution of dioecy and gynodioecy. American Naturalist 112:975-997.

Charlesworth, B., J. A. Coyne, and N. H. Barton. 1987. The relative rates of evolution of sex chromosomes and autosomes. American Naturalist 130:113-146.

Charlesworth, B. 1991. The evolution of sex chromosomes. Science 251:1030-1033.

Charlesworth, B., J. A. Coyne, and H. A. Orr. 1993. Meiotic drive and unisexual hybrid sterility - A comment Genetics 133:421-424.

Charlesworth, B., M. T. Morgan, and D. Charlesworth. 1993. The effect of deleterious mutations on neutral molecular variation. Genetics 134:1289-1303.

Charlesworth, B. 1996. The evolution of chromosomal sex determination and dosage compensation. Current Biology 6:149-162.

Charlesworth, B., and D. Charlesworth. 2000. The degeneration of Y chromosomes. Philosophical Transactions of the Royal Society B-Biological Sciences 355:1563-1572.

Charlesworth, D., B. Charlesworth, and G. Marais. 2005. Steps in the evolution of heteromorphic sex chromosomes. Heredity 95:118-128.

Charlesworth, D. 2008. Sex chromosome origins and evolution. Pp. 207-239 in M. Pagel, and A. Pomiankowski, eds. Evolutionary genomics and proteomics. Sinauer Associates, Inc.

Charlesworth, B. 2009. Effective population size and patterns of molecular evolution and variation. Nature Reviews Genetics 10:195-205.

Charlesworth, D., and J. E. Mank. 2010. The birds and the bees and the flowers and the trees: Lessons from genetic mapping of sex determination in plants and animals. Genetics 186:9-31.

Chen, N., D. W. Bellott, D. C. Page, and A. G. Clark. 2012. Identification of avian W-linked contigs by short-read sequencing. BMC Genomics 13:183.

Chen, J.-M., D. N. Cooper, N. Chuzhanova, C. Ferec, and G. P. Patrinos. 2007. Gene conversion: mechanisms, evolution and human disease. Nature Review Genetics 8:762-775.

Chen, Z-X., Y. E. Zhang, M. Vibranovski, J. Luo, G. Gao, and M. Long. 2011. Deficiency of X-linked inverted duplicates with male-biased expression and the underlying evolutionary mechanisms in the Drosophila genome. Molecular Biology and Evolution 28:2823-2832.

Chibalina, M. V., and D. A. Filatov. 2011. Plant Y chromosome degeneration is retarded by haploid purifying selection. Current. Biology 21:1475-1479.

Chippindale, A. K., J. R. Gibson, and W. R. Rice. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in Drosophila. Proceedings of the National Academy of Sciences, USA 98:1671-1675.

Clark, A.G., Begun, D.J. & Prout, T. 1999. Female x male interactions in *Drosophila* sperm competition. Science 283:217–220.

Clark, N. L., and W. J. Swanson. 2005. Pervasive adaptive evolution in primate seminal proteins. PLoS Genetics 1:335-342.

Collet, J., D. S. Richardson, K. Worley, and T. Pizzari. 2012. Sexual selection and the differential effect of polyandry. Proceedings of the National Academy of Sciences, USA. 109: 8641-8645.

Comeron, J. M. 1995. A method for estimating the numbers of synonymous and nonsynonymous substitutions per site. Journal of Molecular Evolution 41:1152-1159.

Comeron, J. M. 1999. K-estimator: Calculation of the number of nucleotide substitutions per site and the confidence intervals. Bioinformatics 15:763-764.

Connallon, T., and L. L. Knowles. 2005. Intergenomic conflict revealed by patterns of sex-biased gene expression. Trends in Genetics 21:495-499.

Connallon, T. 2007. Adaptive protein evolution of X-linked and autosomal genes in *Drosophila*: Implications for Faster-X hypotheses. Molecular Biology and Evolution 24:2566-2572.

Connallon, T., R. M. Cox, and R. Calsbeek. 2010. Fitness consequences of sex-specific selection. Evolution 64:1671-1682.

Connallon, T., and A. G. Clark. 2010. Sex linkage, sex-specific selection and the role of recombination in the evolution of sexually dimorphic gene expression. Evolution 64:3417-3442.

Connallon, T., and A. G. Clark. 2011. Association between sex-biased gene expression and mutations with sex-specific phenotypic consequences in Drosophila. Genome Biology and Evolution 3:151-155.

Connallon, T., and A. G. Clark. 2011. The resolution of sexual antagonism by gene duplication. Genetics. 187:919-937.

Connallon, T., N. D. Singh, and A. G. Clark. 2012. Impact of genetic architecture on the relative rates of X versus autosomal adaptive substitution. Molecular Biology and Evolution 29:1933-1942.

Corl, A., and H. Ellegren. 2012. The genomic signature of sexual selection in the genetic diversity of the sex chromosomes and autosomes. Evolution 66:2138-2149.

Cutter, A. D., and S. Ward. 2005. Sexual and temporal dynamics of molecular evolution in *C. elegans* development. Molecular Biology and Evolution 22:178-188.

Dalloul, R. A., J. A. Long, A. V. Zimin, L. Aslam, K. Beal, L. A. Blomberg, P. Bouffard, D. W. Burt, O. Crasta, R. P. M. A. Crooijmans, et al. 2010. Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis. PLoS Biology 8:e1000475.

Dalloul, R. A., A. V. Zimin, R. E. Settlage, S. Kim, and K. M. Reed. 2014. Next-generation sequencing strategies for characterizing the turkey genome. Poultry Science 93:479-484.

Davis, J., P. Thomas, N. C. S. Program, and J. Thomas. 2010. A W-linked palindrome and gene conversion in New World sparrows and blackbirds. Chromosome Research 18:543-553.

Day, T., and R. Bonduriansky. 2004. Intralocus sexual conflict can drive the evolution of genomic imprinting. Genetics 167:1537-1546.

Deakin, J. E., T. A. Hore, E. Koina, and J. A. M. Graves. 2008. The Status of Dosage Compensation in the Multiple X Chromosomes of the Platypus. PLoS Genetics 4: e1000140.

Dean, R., J. C. Perry, T. Pizzari, J. E. Mank, and S. Wigby. 2012. Experimental evolution of a novel sexually antagonistic allele. PLoS Genetics 8:e1002917.

Dean, R., and J. E. Mank. 2014. The role of sex chromosomes in sexual dimorphism: discordance between molecular and phenotypic data. Journal of Evolutionary Biology 27:1443-1453.

Deng, X., J. B. Hiatt, N. Di Kim, S. Ercan, D. Sturgill, L. W. Hillier, F. Schlesinger, C. A. Davis, V. J. Reinke, T. R. Gingeras, et al. 2011. Evidence for compensatory upregulation of expressed X-linked genes in mammals, *Caenorhabditis elegans* and *Drosophila melanogaster*. Nature Genetics 43:1179-1185.

DeVeale, B., D. van der Kooy, and T. Babak. 2012. Critical evaluation of imprinted gene expression by RNA-Seq: a new perspective. PLoS Genetics 8:e1002600.

Dimcheff, D. E., S. V. Drovetski, and D. P. Mindell. 2002. Phylogeny of *Tetraoninae* and other galliform birds using mitochondrial 12S and ND2 genes. Molecular Phylogenetics and Evolution 24:203-215.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T. R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15-21.

Eden, E., D. Lipson, S. Yogev, and Z. Yakhini. 2007. Discovering motifs in ranked lists of DNA sequences. PLoS Computational Biology 3:508-522.

Eden, E., R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini. 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10:48.

Edgar, R. C. 2004a. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:1-19.

Edgar, R. C. 2004b. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Researcg 32:1792-1797.

Ellegren, H., and A. Carmichael. 2001. Multiple and independent cessation of recombination between avian sex chromosomes. Genetics 158:325-331.

Ellegren, H., and J. Parsch. 2007. The evolution of sex-biased genes and sex-biased gene expression. Nature Reviews Genetics 8:689-698.

Ellegren, H., L. Hultin-Rosenberg, B. Brunstrom, L. Dencker, K. Kultima, and B. Scholz. 2007. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. BMC Biology 5:40.

Ellegren, H. 2011. Emergence of male-biased genes on the chicken Z-chromosome: Contrasts between male and female heterogametic systems. Genome Research 21:2082-2086.

Ellegren, H., L. Smeds, R. Burri, P. I. Olason, N. Backstrom, T. Kawakami, A. Kunstner, H. Makinen, K. Nadachowska-Brzyska, A. Qvarnstrom, et al. 2012. The genomic landscape of species divergence in Ficedula flycatchers. Nature 491:756-760.

Ellegren, H. 2013. The evolutionary genomics of birds. Annual Review of Ecology, Evolution, and Systematics, 44:239-259.

Engels, B. M., and G. Hutvagner. 2006. Principles and effects of microRNA-mediated post-transcriptional gene regulation. Oncogene 25:6163-6169.

Ezawa, K., S. Oota, and N. Saitou. 2006. Genome-wide search of gene conversions in duplicated genes of mouse and rat. Molecular Biology and Evolution 23:927-940.

Ezawa, K., K. Ikeo, T. Gojobori, and N. Saitou. 2010. Evolutionary pattern of gene homogenization between primate-specific paralogs after human and macaque speciation using the 4-2-4 method. Molecular Biology and Evolution 27:2152-2171.

Fisher, R.A. 1930. The Genetical theory of Natural Selection. The Clarendon Press, Oxford.

Fiumera, A.C., Dumont, B.L. & Clark, A.G. 2005. Sperm competitive ability in *Drosophila* melanogaster associated with variation in male reproductive proteins. Genetics 169:243–257.

Flicek, P., I. Ahmed, M. R. Amode, D. Barrell, K. Beal, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fairley, et al. 2013. Ensembl 2013. Nucleic Acids Research 41:D48-D55.

Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. B. Kruuk. 2007. Sexually antagonistic genetic variation for fitness in red deer. Nature 447:1107-1109.

Forgy, E. W. 1965. Cluster analysis of multivariate data: Efficiency vs interpretability of classifications. Biometrics 21:768-769.

Frank, S. A., and L. D. Hurst. 1996. Mitochondria and male disease. Nature. 383:224.

Gallach, M., C. Chandrasekaran, and E. Betrán. 2010. Analyses of nuclearly encoded mitochondrial genes suggest gene duplication as a mechanism for resolving intralocus sexually antagonistic conflict in *Drosophila*. Genome Biology and Evolution 2:835-850.

Gallach, M., and E. Betran. 2011. Intralocus sexual conflict resolved through gene duplication. Trends in Ecology & Evolution 26:222-228.

Galtier, N. 2003. Gene conversion drives GC content evolution in mammalian histones. Trends in Genetics 19:65-68.

Gerrard, D. T., and D. A. Filatov. 2005. Positive and negative selection on Mammalian Y chromosomes. Molecular Biology and Evolution 22:1423-1432.

Good, J. M., and M. W. Nachman. 2005. Rates of protein evolution are positively correlated with developmental timing of expression during mouse spermatogenesis. Molecular Biology and Evolution 22:1044-1052.

Goto, H., L. Peng, and K. Makova. 2009. Evolution of X-degenerate Y chromosome genes in Greater Apes: Conservation of gene content in Human and Gorilla, but not Chimpanzee. Journal of Molecular Evolution 68:134-144.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29:644-652.

Grath, S., J. F. Baines, and J. Parsch. 2009. Molecular evolution of sex-biased genes in the *Drosophila ananassae* subgroup. BMC Evolutionary Biology 9:291.

Grath, S., and J. Parsch. 2012. Rate of amino acid substitution is influenced by the degree and conservation of male-biased transcription over 50 Myr of *Drosophila* evolution. Genome Biology and Evolution 4:346-359.

Graur, D., and W-H. Li. 2000. Fundamentals of molecular evolution. Sinauer Associates.

Gregg, C., J. W. Zhang, B. Weissbourd, S. J. Luo, G. P. Schroth, D. Haig, and C. Dulac. 2010b. High-resolution analysis of parent-of-origin allelic expression in the mouse brain. Science 329:643-648.

Gregg, C., J. W. Zhang, J. E. Butler, D. Haig, and C. Dulac. 2010a. Sex-specific parent-of-origin allelic expression in the mouse brain. Science 329:682-685.

Hackstein, J. H. P., R. Hochstenbach, E. HauschteckJungen, and L. W. Beukeboom. 1996. Is the Y chromosome of *Drosophila* an evolved supernumerary chromosome? Bioessays 18:317-323.

Haddrill, P. R., L. Loewe, and B. Charlesworth. 2010. Estimating the parameters of selection on nonsynonymous mutations in *Drosophila pseudoobscura* and *D. miranda*. Genetics. 185:1381-1396.

Haerty, W., S. Jagadeeshan, R. J. Kulathinal, A. Wong, K. Ravi Ram, L. K. Sirot, L. Levesque, C. G. Artieri, M. F. Wolfner, A. Civetta, et al. 2007. Evolution in the fast lane: Rapidly evolving sex-related genes in *Drosophila*. Genetics 177:1321-1335.

Hager, R., J. M. Cheverud, L. J. Leamy, and J. B. Wolf. 2008. Sex dependent imprinting effects on complex traits in mice. BMC Evolutionary Biology 8:303.

Haig, D. 2009. Transfers and transitions: Parent-offspring conflict, genomic imprinting, and the evolution of human life history. Proceedings of the National Academy of Sciences, USA.107:1731-1735.

Hallast, P., P. Balaresque, G. R. Bowden, S. Ballereau, and M. A. Jobling. 2013. Recombination dynamics of a human Y-chromosomal palindrome: Rapid GC-biased gene conversion, multi-kilobase conversion tracts, and rare inversions. PLoS Genetics 9:e1003666.

Handley, L. L., H. Ceplitis, and H. Ellegren. 2004. Evolutionary strata on the chicken Z chromosome: Implications for sex chromosome evolution. Genetics 167:367-376.

Harano, T., K. Okada, S. Nakayama, T. Miyatake, and D. J. Hosken. 2010. Intralocus sexual conflict unresolved by sex-limited trait expression. Current Biology 20:2036-2039.

Harlid, A. B., S. Berlin, N. G. C. Smith, A. P. Mosller, and H. Ellegren. 2003. Life history and the male mutation bias. Evolution 57:2398-2406.

Harrison, P. W., J. GE, and S. H. Montgomery. In review. Sliding Window Alignment Masker for PAML. Evolutionary Bioinformatics.

Harrison, P. W., A. E. Wright, F. Zimmer, R. F. Dean, S. H. Montgomery, M. A. Pointer, and J. E. Mank. In prep. Rapid turnover of genes under male-specific selection and high levels of genetic variation in male-biased genes.

Hartigan, J. A., and M. A. Wong. 1979. Algorithm AS 136: A k-means clustering algorithm. Journal of the Royal Statistical Society. Series C. 28:100-108.

Hartl, D. L., and A. G. Clark. 2007. Principles of population genetics. Sinauer Associates, Inc.

Hedges, S. B. 2002. The origin and evolution of model organisms. Nature Reviews Genetics 3:838-849.

Hedges, S. B., J. Dudley, and S. Kumar. 2006. TimeTree: a public knowledge-base of divergence times among organisms. Bioinformatics 22:2971-2972.

Helfenstein, F., C. Tirard, E. Danchin, and R. H. Wagner. 2004. Low frequency of extra-pair paternity and high frequency of adoption in black-legged kittiwakes The Condor 106:149-155.

Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6:65-70.

Hope, A. C. A. 1968. A simplified Monte Carlo significance test procedure. Journal of the Royal Statistical Society. Series B. 30:582-598.

Huang, Y., Y. Li, D. W. Burt, H. Chen, Y. Zhang, W. Qian, H. Kim, S. Gan, Y. Zhao, J. Li, et al. 2013. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. Nature Genetics 45:776-783.

Hughes, J. F., H. Skaletsky, T. Pyntikova, T. A. Graves, S. K. M. van Daalen, P. J. Minx, R. S. Fulton, S. D. McGrath, D. P. Locke, C. Friedman, et al. 2010. Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. Nature 463:536-539.

Hughes, J. F., H. Skaletsky, L. G. Brown, T. Pyntikova, T. Graves, R. S. Fulton, S. Dugan, Y. Ding, C. J. Buhay, C. Kremitzki, et al. 2012. Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. Nature 483:82-86.

Hvilsom, C., Y. Qian, T. Bataillon, Y. Li, T. Mailund, B. Salle, F. Carlsen, R. Li, H. Zheng, T. Jiang, et al. 2012. Extensive X-linked adaptive evolution in central chimpanzees. Proceedings of the National Academy of Sciences, USA 109:2054-2059.

Innocenti, P., and E. H. Morrow. 2010. The sexually antagonistic genes of *Drosophila melanogaster*. PLoS Biology 8:e1000335.

Innocenti, P., E. H. Morrow, and D. K. Dowling. 2011. Experimental evidence supports a sex-specific selective sieve in mitochondrial genome evolution. Science 332:845-848.

Itoh, Y., E. Melamed, X. Yang, K. Kampf, S. Wang, N. Yehya, A. Van Nas, K. Replogle, M. Band, D. Clayton, et al. 2007. Dosage compensation is less effective in birds than in mammals. Journal of Biology 6:2.

Iwase, M., Y. Satta, H. Hirai, Y. Hirai, and N. Takahata. 2010. Frequent gene conversion events between the X and Y homologous chromosomal regions in primates. BMC Evolutionary Biology 10:225.

Jiang, Z. F., D. A. Croshaw, Y. Wang, J. Hey, and C. A. Machado. 2011. Enrichment of mRNA-like noncoding RNAs in the divergence of *Drosophila* males. Molecular Biology and Evolution 28:1339-1348.

Jin, W., R. M. Riley, R. D. Wolfinger, K. P. White, G. Passador-Gurgel, and G. Gibson. 2001. The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. Nature Genetics 29:389-395.

Julien, P., D. Brawand, M. Soumillon, A. Necsulea, A. Liechti, F. Schütz, T. Daish, F. Grützner, and H. Kaessmann. 2012. Mechanisms and evolutionary patterns of Mammalian and Avian dosage compensation. PLoS Biology 10:e1001328.

Kaiser, V. B., and H. Ellegren. 2006. Nonrandom distribution of genes with sex-biased expression in the chicken genome. Evolution 60:1945-1951.

Kaiser, V. B., Q. Zhou, and D. Bachtrog. 2011. Nonrandom gene loss from the *Drosophila miranda* neo-Y chromosome. Genome Biology and Evolution 3:1329-1337.

Karlin, S., and S. F. Altschul. 1990. Methods for assessing the statistical significance of molecular sequence features by using general scoring schemes. Proceedings of the National Academy of Sciences, USA 87:2264-2268.

Karlin, S., and S. F. Altschul. 1993. Applications and statistics for multiple high-scoring segments in molecular sequences. Proceedings of the National Academy of Sciences, USA 90:5873-5877.

Khaitovich, P., I. Hellmann, W. Enard, K. Nowick, M. Leinweber, H. Franz, G. Weiss, M. Lachmann, and S. Pääbo. 2005. Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. Science 309:1850-1854.

Khil, P. P., N. A. Smirnova, P. J. Romanienko, and R. D. Camerini-Otero. 2004. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. Nature Genetics 36:642-646.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S. L. Salzberg. 2013. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biology 14:R36.

Kirkpatrick, M., and D. W. Hall. 2004a. Sexual selection and sex linkage. Evolution 58:683-691.

Kirkpatrick, M., and D. W. Hall. 2004b. Male-biased mutation, sex linkage, and the rate of adaptive evolution. Evolution 58:437-440.

Kitano, J., J. A. Ross, S. Mori, M. Kume, F. C. Jones, Y. F. Chan, D. M. Absher, J. Grimwood, J. Schmutz, R. M. Myers, et al. 2009. A role for a neo-sex chromosome in stickleback speciation. Nature 461:1079-1083.

Knief, U., H. Schielzeth, B. Kempenaers, H. Ellegren, and W. Forstmeier. 2012. QTL and quantitative genetic analysis of beak morphology reveals patterns of standing genetic variation in an Estrildid finch. Molecular Ecology 21:3704-3717.

Koboldt, D. C., K. Chen, T. Wylie, D. E. Larson, M. D. McLellan, E. R. Mardis, G. M. Weinstock, R. K. Wilson, and L. Ding. 2009. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25:2283-2285.

Koboldt, D. C., Q. Zhang, D. E. Larson, D. Shen, M. D. McLellan, L. Lin, C. A. Miller, E. R. Mardis, L. Ding, and R. K. Wilson. 2012. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Research 22:568-576.

Koerich, L. B., X. Y. Wang, A. G. Clark, and A. B. Carvalho. 2008. Low conservation of gene content in the *Drosophila* Y chromosome. Nature 456:949-951.

Kokko, H., H. Klug, and M. D. Jennions. 2012. Unifying cornerstones of sexual selection: operational sex ratio, Bateman gradient and the scope for competitive investment. Ecology Letters 15:1340-1351.

Lahn, B. T., and D. C. Page. 1999. Four evolutionary strata on the human X chromosome. Science 286:964-967.

Lange, J., H. Skaletsky, S. K. M. van Daalen, S. L. Embry, C. M. Korver, L. G. Brown, R. D. Oates, S. Silber, S. Repping, and D. C. Page. 2009. Isodicentric Y chromosomes and sex disorders as byproducts of homologous recombination that maintains palindromes. Cell 138:855-869.

Langley, C. H., K. Stevens, C. Cardeno, Y. C. G. Lee, D. R. Schrider, J. E. Pool, S. A. Langley, C. Suarez, R. B. Corbett-Detig, B. Kolaczkowski, et al. 2012. Genomic variation in natural populations of *Drosophila melanogaster*. Genetics 192:533-598.

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology 10:R25.

Langmead, B., and S. L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods 9:357-359.

Laporte, V., and B. Charlesworth. 2002. Effective population size and population subdivision in demographically structured populations. Genetics 162:501-519.

Lau, A. N., L. Peng, H. Goto, L. Chemnick, O. A. Ryder, and K. D. Makova. 2009. Horse domestication and conservation genetics of Przewalski's horse inferred from sex chromosomal and autosomal sequences. Molecular Biology and Evolution 26:199-208.

Lemaitre, C., M. D. V. Braga, C. Gautier, M-F. Sagot, E. Tannier, and G. A. B. Marais. 2009. Footprints of inversions at present and past pseudoautosomal boundaries in human sex chromosomes. Genome Biology and Evolution 1:56-66.

Li, B., and C. N. Dewey. 2011. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and P. Genome Project Data. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079.

Li, G., B. W. Davis, T. Raudsepp, A. J. P. Wilkerson, V. C. Mason, M. Ferguson-Smith, P. C. O'Brien, P. D. Waters, and W. J. Murphy. 2013. Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. Genome Research 23:1486-1495.

Lifjeld, J.T., Dunn, P.O., Robertson, R.J. & Boag, P.T. 1993. Extra-pair paternity in monogamous tree swallows. Animal Behaviour 45: 213–229.

Liker, A., J. D. Reynolds, and T. Székely. 2001. The evolution of egg size in socially polyandrous shorebirds. Oikos 95:3-14.

Lin, F., K. Xing, J. Zhang, and X. He. 2012. Expression reduction in mammalian X chromosome evolution refutes Ohno's hypothesis of dosage compensation. Proceedings of the National Academy of Sciences, USA. 109:11752-11757.

Lin, H., V. Gupta, M. D. VerMilyea, F. Falciani, J. T. Lee, L. P. O'Neill, and B. M. Turner. 2007. Dosage compensation in the mouse balances up-regulation and silencing of X-linked genes. PLoS Biology 5:2809-2820.

Lindenfors, P., Tullberg, B.S. & Biuw, M. 2002. Phylogenetic analyses of sexual selection and sexual size dimorphism in pinnipeds. Behavioural Ecology and Sociobiology. 52: 188–193.

Llopart, A. 2012. The rapid evolution of X-linked male-biased gene expression and the large-X effect in *Drosophila yakuba*, *D. santomea* and their hybrids. Molecular Biology and Evolution 29:3873-3886.

Lloyd, S. P. 1982. Least-squares qunatization in PCM. IEEE Transactions on Information Theory 28:129-137.

Lohse, M., A. M. Bolger, A. Nagel, A. R. Fernie, J. E. Lunn, M. Stitt, and B. Usadel. 2012. RobiNA: A user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Research 40:W622-W627.

Long, T.A.F., Agrawal, A.F. & Rowe, L. 2012. The effect of sexual selection on offspring fitness depends on the nature of genetic variation. Current Biology 22: 204–208.

Loytynoja, A., and N. Goldman. 2005. An algorithm for progressive multiple alignment of sequences with insertions. Proceedings of the National Academy of Sciences, USA 102:10557-10562.

Lucchesi, J. C. 1973. Dosage compensation in *Drosophila*. Annual Review of Genetics 7:225-237.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. Pp. 281–297. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. University of California Press.

Magurran, A.E. & Seghers, B.H. 1994. Sexual conflict as a consequence of ecology - Evidence from guppy, *Poecilia-reticulata*, populations in Trinidad. Proceedings of the Royal Society B-Biological Sciences. 255: 31–36.

Mank, J. E., E. Axelsson, and H. Ellegren. 2007a. Fast-X on the Z: Rapid evolution of sex-linked genes in birds. Genome Research 17:618-624.

Mank, J. E., L. Hultin-Rosenberg, E. Axelsson, and H. Ellegren. 2007b. Rapid evolution of female-biased, but not male-biased, genes expressed in the avian brain. Molecular Biology and Evolution 24:2698-2706.

Mank, J. E., L. Hultin-Rosenberg, M. T. Webster, and H. Ellegren. 2008a. The unique genomic properties of sex-biased genes: Insights from avian microarray data. BMC Genomics 9:148

Mank, J. E., L. Hultin-Rosenberg, M. Zwahlen, and H. Ellegren. 2008b. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. American Naturalist 171:35-43.

Mank, J. E. 2009a. Sex chromosomes and the evolution of sexual dimorphism: Lessons from the genome. American Naturalist 173:141-150.

Mank, J. E. 2009b. The W, X, Y and Z of sex-chromosome dosage compensation. Trends in Genetics 25:226-233.

Mank, J. E., and H. Ellegren. 2009a. All dosage compensation is local: Gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. Heredity 102:312-320.

Mank, J. E., and H. Ellegren. 2009b. Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. Evolution 63:1464-1472.

Mank, J. E., B. Vicoso, S. Berlin, and B. Charlesworth. 2010a. Effective population size and the Faster-X effect: Empirical results and their interpretation. Evolution 64:663-674.

Mank, J. E., K. Nam, B. Brunstrom, and H. Ellegren. 2010b. Ontogenetic complexity of sexual dimorphism and sex-specific selection. Molecular Biology and Evolution 27:1570-1578.

Mank, J. E., K. Nam, and H. Ellegren. 2010c. Faster-Z Evolution is predominantly due to genetic drift. Molecular Biology and Evolution 27:661-670.

Mank, J. E., D. J. Hosken, and N. Wedell. 2011. Some inconvenient truths about sex chromosome dosage compensation and the potential role of sexual conflict. Evolution 65:2133-2144.

Mank, J. E. 2012. Small but mighty: The evolutionary dynamics of W and Y sex chromosomes. Chromosome Research 20:21-33.

Marais, G. A. B., P. R. A. Campos, and I. Gordo. 2010. Can intra-Y gene conversion oppose the degeneration of the human Y chromosome? A simulation study. Genome Biology and Evolution 2:347-357.

Marinotti, O., E. Calvo, Q. K. Nguyen, S. Dissanayake, J. M. C. Ribeiro, and A. A. James. 2006. Genome-wide analysis of gene expression in adult *Anopheles gambiae*. Insect Molecular Biology 15:1-12.

Matute, D.R. 2010. Reinforcement can overcome gene flow during speciation in *Drosophila*. Current Biology 20: 2229–2233.

McDonald, J. H., and M. Kreitman. 1991. Adatpive protein evolution at the ADH locus in *Drosophila*. Nature 351:652-654.

McGrath, C. L., C. Casola, and M. W. Hahn. 2009. Minimal effect of ectopic gene conversion among recent duplicates in four Mammalian genomes. Genetics 182:615-622.

Meiklejohn, C. D., J. Parsch, J. M. Ranz, and D. L. Hartl. 2003. Rapid evolution of male-biased gene expression in *Drosophila*. Proceedings of the National Academy of Sciences, USA 100:9894-9899.

Meiklejohn, C. D., E. L. Landeen, J. M. Cook, S. B. Kingan, and D. C. Presgraves. 2011. Sex chromosome-specific regulation in the *Drosophila* male germline but little evidence for chromosomal dosage compensation or meiotic inactivation. PLoS Biology 9:e1001126.

Meisel, R. 2011. Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein coding sequence evolution. Molecular Biology and Evolution 28:1893-1900.

Meisel, R. P., and T. Connallon. 2013. The Faster-X effect: integrating theory and data. Trends in Genetics 29:537-544.

Melamed, E., and A. P. Arnold. 2007. Regional differences in dosage compensation on the chicken Z chromosome. Genome Biology 8: R202.

Merkin, J., Russell, C., Chen, P. & Burge, C.B. 2012. Evolutionary dynamics of gene and isoform regulation in mammalian tissues. Science 338: 1593–1599.

Metta, M., Gudavalli, R., Gibert, J.M. & Schlotterer, C. 2006. No accelerated rate of protein evolution in male-biased *Drosophila pseudoobscura* genes. Genetics 174: 411–420.

Metz, E.C. & Palumbi, S.R. 1996. Positive selection and sequence rearrangements generate extensive polymorphisms in the gamete recognition protein binding. Molecular Biology and Evolution 13: 397–406.

Meunier, J., and L. Duret. 2004. Recombination drives the evolution of GC-content in the human genome. Molecular Biology and Evolution 21:984-990.

Mirouze, M., M. Lieberman-Lazarovich, R. Aversano, E. Bucher, J. Nicolet, J. Reinders, and J. Paszkowski. 2012. Loss of DNA methylation affects the recombination landscape in *Arabidopsis*. Proceedings of the National Academy of Sciences, USA 109:5880-5885.

Moghadam, H. K., M. A. Pointer, A. E. Wright, S. Berlin, and J. E. Mank. 2012. W chromosome expression responds to female-specific selection. Proceedings of the National Academy of Sciences, USA 109:8207-8211.

Moller, A. P. 1988. Testes size, ejaculate quality and sperm competition in birds. Biological Journal of the Linnean Society 33:273-283.

Moller, A. P. 1991. Sperm competition, sperm depletion, paternal care, and relative testis size in birds. American Naturalist 137:882-906.

Moller, A. P., and J. V. Briskie. 1995. Extra-pair paternity, sperm competition and the evolution of testis size in birds. Behavioral Ecology and Sociobiology 36:357-365.

Montgomery, S.H, I. Capellini, C. Venditti, R.A, Barton and N. I. Mundy 2011. Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. Molecular Biology and Evolution. 28:625-638.

Morrow, E. H., A. D. Stewart, and W. R. Rice. 2008. Assessing the extent of genome-wide intralocus sexual conflict via experimentally enforced gender-limited selection. Journal of Evolutionary Biology 21:1046-1054.

Muyle, A., N. Zemp, C. Deschamps, S. Mousset, A. Widmer, and G. A. B. Marais. 2012. Rapid de novo evolution of X chromosome dosage compensation in *Silene latifolia,* a plant with young sex chromosomes. PLoS Biology 10:e1001308.

Myers, S., R. Bowden, A. Tumian, R. E. Bontrop, C. Freeman, T. S. MacFie, G. McVean, and P. Donnelly. 2010. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science 327:876-879.

Nadeau, N. J., T. Burke, and N. I. Mundy. 2007. Evolution of an avian pigmentation gene correlates with a measure of sexual selection. Proceedings of the Royal Society B-Biological Sciences 274:1807-1813.

Nam, K., and H. Ellegren. 2008. The chicken (*Gallus gallus*) Z chromosome contains at least three nonlinear evolutionary strata. Genetics 180:1131-1136.

Natri, H. M., T. Shikano, and J. Merila. 2013. Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. Molecular Biology and Evolution 30:1131-1144.

Naurin, S., B. Hansson, D. Hasselquist, Y. Kim, and S. Bensch. 2011. The sex-biased brain: sexual dimorphism in gene expression in two species of songbirds. BMC Genomics 12:37.

Obermann, C., E. Meyer, S. Prager, J. R. Tomiuk, H. A. Wollmann, and T. Eggermann. 2004. Searching for genomic variants in IGF2 and CDKN1C in Silver-Russell syndrome patients. Molecular Genetics and Metabolism 82:246-250.

Ohno, S. 1967. Sex Chromosomes and Sex Linked Genes. Springer-Verlag, Berlin.

Oliver, B. 2007. Sex, Dose, and Equality. PLoS Biology 5:e340.

Ometto, L., D. Shoemaker, K. G. Ross, and L. Keller. 2010. Evolution of gene expression in fire ants: the effects of developmental stage, caste, and species. Molecular Biology and Evolution. 28:1381-1392.

Orr, H.A. 2000. Adaptation and the cost of complexity. Evolution 54: 13–20.

Pagel, M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877-884.

Pagel, M., A. Meade, and D. Barker. 2004. Bayesian estimation of ancestral character states on phylogenies. Systematic Biology 53:673-684.

Palumbi, S.R. 1999. All males are not created equal: Fertility differences depend on gamete recognition polymorphisms in sea urchins. Proceedings of the National Academy of Sciences, USA 96: 12632–12637

Pandey, R. S., M. A. Wilson Sayres, and R. K. Azad. 2013. Detecting evolutionary strata on the human X chromosome in the absence of gametologous Y-linked sequences. Genome Biology and Evolution 5:1863-1871.

Panhuis, T. M., and W. J. Swanson. 2006. Molecular evolution and population genetic analysis of candidate female reproductive genes in *Drosophila*. Genetics 173:2039-2047.

Parisi, M., R. Nuttall, D. Naiman, G. Bouffard, J. Malley, J. Andrews, S. Eastman, and B. Oliver. 2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. Science 299:697-700.

Parisi, M., R. Nuttall, P. Edwards, J. Minor, D. Naiman, J. N. Lu, M. Doctolero, M. Vainer, C. Chan, J. Malley, et al. 2004. A survey of ovary-, testis-, and soma-biased gene expression in Drosophila melanogaster adults. Genome Biology 5:R40.

Patefield, W. M. 1981. Algorithm AS159. An efficient method of generating r x c tables with given row and column totals. Applied Statistics 30:91–97.

Patten, M. M., and D. Haig. 2008. Reciprocally imprinted genes and the response to selection on one sex. Genetics 179:1389-1394.

Pearks Wilkerson, A. J., T. Raudsepp, T. Graves, D. Albracht, W. Warren, B. P. Chowdhary, L. C. Skow, and W. J. Murphy. 2008. Gene discovery and comparative analysis of X-degenerate genes from the domestic cat Y chromosome. Genomics 92:329-338.

Perrin, N. 2009. Sex reversal: a fountain of youth for sex chromosomes. Evolution 63:3043-3049.

Pessia, E., T. Makino, M. Bailly-Bechet, A. McLysaght, and G. A. B. Marais. 2012. Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. Proceedings of the National Academy of Sciences, USA 109:5346-5351.

Pitcher, T. E., P. O. Dunn, and L. A. Whittingham. 2005. Sperm competition and the evolution of testes size in birds. Journal of Evolutionary Biology 18:557-567.

Pointer, M. A., P. W. Harrison, A. E. Wright, and J. E. Mank. 2013. Masculinization of gene expression Is associated with exaggeration of male sexual dimorphism. PLoS Genetics 9:1-9.

Pollack, J. R., T. Sorlie, C. M. Perou, C. A. Rees, S. S. Jeffrey, P. E. Lonning, R. Tibshirani, D. Botstein, A. L. Borresen-Dale, and P. O. Brown. 2002. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proceedings of the National Academy of Sciences, USA 99:12963-12968.

Prince, E. G., D. Kirkland, and J. P. Demuth. 2010. Hyper-expression of the X chromosome in both sexes results in extensive female-bias of X-linked gene in the flour beetle. Genome Biology and Evolution 2:336-346.

Qiu, S., R. Bergero, and D. Charlesworth. 2013. Testing for the footprint of sexually antagonistic polymorphisms in the pseudoautosomal region of a plant sex chromosome pair. Genetics 194:663-672.

Quinn, E. M., P. Cormican, E. M. Kenny, M. Hill, R. Anney, M. Gill, A. P. Corvin, and D. W. Morris. 2013. Development of strategies for SNP detection in RNA-Seq data: Application to lymphoblastoid cell lines and evaluation using 1000 genomes data. PLoS ONE 8:e58815.

R Core Team 2011. R: A language and environment for statistical computing. http://www.R-project.org/

R Core Team 2014. R: A language and environment for statistical computing. http://www.R-project.org/

Ranz, J. M., C. I. Castillo-Davis, C. D. Meiklejohn, and D. L. Hartl. 2003. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. Science 300:1742-1745.

Ratnakumar, A., S. Mousset, S. Glémin, J. Berglund, N. Galtier, L. Duret, and M. T. Webster. 2010. Detecting positive selection within genomes: The problem of biased gene conversion. Philosophical Transactions of the Royal Society B-Biological Sciences. 365:2571-2580.

Reeve, H. K., and D. W. Pfennig. 2003. Genetic biases for showy males: Are some genetic systems especially conducive to sexual selection? Proceedings of the National Academy of Sciences, USA 100:1089-1094.

Reinius, B., P. Saetre, J. A. Leonard, R. Blekhman, R. Merino-Martinez, Y. Gilad, and E. Jazin. 2008. An evolutionarily conserved sexual signature in the primate brain. PLoS Genetics 4: e1000100.

Rice, W. R. 1984. Sex chromosomes and the evolution of sexual dimorphism. Evolution 38:735-742.

Ridley, M. 2003. Evolution. John Wiley & Sons.

Rinn, J. L., J. S. Rozowsky, I. J. Laurenzi, P. H. Petersen, K. Zou, W. Zhong, M. Gerstein, and M. Snyder. 2004. Major molecular differences between Mammalian sexes are involved in drug metabolism and renal function. Developmental Cell 6:791-800.

Roberts, R. B., J. R. Ser, and T. D. Kocher. 2009. Sexual conflict resolved by invasion of a novel sex determiner in Lake Malawi cichlid fishes. Science 326:998-10001.

Ross, M. T., D. V. Grafham, A. J. Coffey, S. Scherer, K. McLay, D. Muzny, M. Platzer, G. R. Howell, C. Burrows, C. P. Bird, et al. 2005. The DNA sequence of the human X chromosome. Nature 434:325-337.

Rosser, Z. H., P. Balaresque, and M. A. Jobling. 2009. Gene conversion between the X chromosome and the male-specific region of the Y chromosome at a translocation hotspot. American Journal of Human Genetics. 85:130-134.

Rozen, S., H. Skaletsky, J. D. Marszalek, P. J. Minx, H. S. Cordum, R. H. Waterston, R. K. Wilson, and D. C. Page. 2003. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature 423:873-876.

Ruokonen, M., L. Kvist, and J. Lumme. 2000. Close relatedness between mitochondrial DNA from seven Anser goose species. Journal of Evolutionary Biology 13:532-540.

Sackton, T. B., R. B. Corbett-Detig, J. Nagaraju, L. Vaishna, K. P. Arunkumar, and D. L. Hartl. 2014. Positive selection drives Faster-Z evolution in silkmoths. Evolution 68:2331-2342.

Saether, S. A., G.-P. Saetre, T. Borge, C. Wiley, N. Svedin, G. Andersson, T. Veen, J. Haavie, M. R. Servedio, S. Bures, et al. 2007. Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. Science 318:95-97.

Saetre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon, C. Primmer, and A. C. Syvanen. 2003. Sex chromosome evolution and speciation in Ficedula flycatchers. Proceedings of the Royal Society B-Biological Sciences 270:53-59.

Sandstedt, S. A., and P. K. Tucker. 2004. Evolutionary strata on the mouse X chromosome correspond to strata on the human X chromosome. Genome Research 14:267-272.

Sawyer, S. A. 1999. A computer package for the statistical detection of gene conversion. Distributed by the author.

Schemper, M. 2003. Predictive accuracy and explained variation. Statistics in Medicine 22:2299-2308.

Schielzeth, H., B. Kempenaers, H. Ellegren, and W. Forstmeier. 2012. QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. Evolution 66:18-30.

Schoenmakers, S., E. Wassenaar, J. W. Hoogerbrugge, J. S. E. Laven, J. A. Grootegoed, and W. M. Baarends. 2009. Female meiotic sex chromosome inactivation in chicken. PLoS Genetics 5: e1000466.

Singh, N. D., J. M. Macpherson, J. D. Jensen, and D. A. Petrov. 2007. Similar levels of X-linked and autosomal nucleotide variation in African and non-African populations of Drosophila melanogaster. BMC Evolutionary Biology 7:202.

Skaletsky, H., T. Kuroda-Kawaguchi, P. J. Minx, H. S. Cordum, L. Hillier, L. G. Brown, S. Repping, T. Pyntikova, J. Ali, T. Bieri, et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825-837.

Skinner, B. M., L. B. W. Robertson, H. G. Tempest, E. J. Langley, D. Ioannou, K. E. Fowler, R. P. M. A. Crooijmans, A. D. Hall, D. K. Griffin, and M. Voelker. 2009. Comparative genomics in chicken and Pekin duck using FISH mapping and microarray analysis. BMC Genomics 10:357.

Slattery, J. P., L. Sanner-Wachter, and S. J. O'Brien. 2000. Novel gene conversion between X-Y homologues located in the nonrecombining region of the Y chromosome in *Felidae* (Mammalia). Proceedings of the National Academy of Sciences, USA 97:5307-5312.

Smith, C. A., K. N. Roeszler, T. Ohnesorg, D. M. Cummins, P. G. Farlie, T. J. Doran, and A. H. Sinclair. 2009. The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. Nature 461:267-271.

Smith, G., Y-R. Chen, G. W. Blissard, and A. D. Briscoe. 2014. Complete dosage compensation and sex-biased gene expression in the moth *Manduca sexta*. Genome Biology and Evolution. 6:526-537.

Snell-Rood, E.C., Van Dyken, J.D., Cruickshank, T., Wade, M.J. & Moczek, A.P. 2010. Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. BioEssays 32: 71–81.

Spigler, R. B., K. S. Lewers, and T-L. Ashman. 2011. Genetic architecture of sexual dimorphism in a subdioecious plant with a proto-sex chromosome Evolution 65:1114-1126.

Stefani, G., and F. J. Slack. 2008. Small non-coding RNAs in animal development. Nature Reviews Molecular Cell Biology 9:219-230.

Stepinska, U. & Bakst, M.R. 2007. Fertilisation. Reproductive Biology and Phylogeny of Birds: Phylogeny, Morphology, Hormones, Fertilization. Science Publishers, Enfield, NH.

Stevenson, K. R., J. D. Coolon, and P. J. Wittkopp. 2013. Sources of bias in measures of allele-specific expression derived from RNA-Seq data aligned to a single reference genome. BMC Genomics 14:536.

Stiglec, R., T. Ezaz, and J. A. M. Graves. 2007. A new look at the evolution of avian sex chromosomes. Cytogenetic Genome Research 117:103-109.

Stöck, M., A. Horn, C. Grossen, D. Lindtke, R. Sermier, C. Betto-Colliard, C. Dufresnes, E. Bonjour, Z. Dumas, E. Luquet, et al. 2011. Ever-young sex chromosomes in European tree frogs. PLoS Biology 9:e1001062.

Storchova, R., and P. Divina. 2006. Nonrandom representation of sex-biased genes on chicken Z chromosome. Journal of Molecular Evolution 63:676-681.

Sturgill, D., Y. Zhang, M. Parisi, and B. Oliver. 2007. Demasculinization of X chromosomes in the *Drosophila* genus. Nature 450:238-241.

Suh, A., J. O. Kriegs, J. Brosius, and J. Schmitz. 2011. Retroposon insertions and the chronology of avian sex chromosome evolution. Molecular Biology and Evolution 28:2993-2997.

Swaney, W. T., J. P. Curley, F. A. Champagne, and E. B. Keverne. 2007. Genomic imprinting mediates sexual experience-dependent olfactory learning in male mice. Proceedings of the National Academy of Sciences, USA 104:6084-6089.

Swanson, W. J., A. G. Clark, H. M. Waldrip-Dail, M. F. Wolfner, and C. F. Aquadro. 2001. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. Proceedings of the National Academy of Sciences, USA 98:7375-7379.

Swanson, W. J., and V. D. Vacquier. 2002. The rapid evolution of reproductive proteins. Nature Reviews Genetics 3:137-144.

Swanson, W. J., A. Wong, M. F. Wolfner, and C. F. Aquadro. 2004. Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection. Genetics 168:1457-1465.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28:2731-2739.

Tarin, J.J. & Caro, A. 2000. Fertilization in Protozoa and Metazoan Animals. Springer, Berlin.

Telonis-Scott, M., A. Kopp, M. L. Wayne, S. V. Nuzhdin, and L. M. McIntyre. 2009. Sex-specific splicing in *Drosophila*: Widespread occurrence, tissue specificity and evolutionary conservation. Genetics 181:421-434.

Teranishi, M., Y. Shimada, T. Hori, O. Nakabayashi, T. Kikuchi, T. Macleod, R. Pym, B. Sheldon, I. Solovei, H. Macgregor, et al. 2001. Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. Chromosome Research 9:147-165.

Tregenza, T., Wedell, N. & Chapman, T. 2006. Introduction. Sexual conflict: a new paradigm? Philosophical Transactions of the Royal Society B-Biological Sciences. 361: 229–234.

Thornton, K., and M. Long. 2005. Excess of amino acid substitutions relative to polymorphism between X-linked duplications in Drosophila melanogaster. Molecular Biology and Evolution 22:273-284.

Tian, L., T. Cai, E. Goetghebeur, and L. J. Wei. 2007. Model evaluation based on the sampling distribution of estimated absolute prediction error. Biometrika 94:297-311.

Torres, E. M., B. R. Williams, and A. Amon. 2008. Aneuploidy: Cells losing their balance. Genetics 179:737-746.

Toups, M. A., J. B. Pease, and M. W. Hahn. 2011. No excess gene movement is detected off the avian or lepidopteran Z chromosome. Genome Biology and Evolution 3:1381-1139.

Trapnell, C., L. Pachter, and S. L. Salzberg. 2009. TopHat: Discovering splice junctions with RNA-Seq. Bioinformatics 25:1105-1111.

Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, and L. Pachter. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nature Biotechnology 28:511-515.

Trapnell, C., D. G. Hendrickson, M. Sauvageau, L. Goff, J. L. Rinn, and L. Pachter. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nature Biotechnology 31:46-53.

Trombetta, B., D. Sellitto, R. Scozzari, and F. Cruciani. 2014. Inter- and intra-species phylogenetic analyses reveal extensive X-Y gene conversion in the evolution of gametologous sequences of human sex chromosomes. Molecular Biology and Evolution. 31:2108-2123.

Tsuda, Y., C. Nishida-Umehara, J. Ishijima, K. Yamada, and Y. Matsuda. 2007. Comparison of the Z and W sex chromosomal architectures in elegant crested tinamou (*Eudromia elegans*) and ostrich (*Struthio camelus*) and the process of sex chromosome differentiation in palaeognathous birds. Chromosoma 116:159-173.

Turner, L. M., E. B. Chuong, and H. E. Hoekstra. 2008. Comparative analysis of testis protein evolution in rodents. Genetics 179:2075-2089.

Van Cleve, J., and M. W. Feldman. 2007. Sex-specific viability, sex linkage and dominance in genomic imprinting. Genetics 176:1101-1118.

van Doorn, G. S., and M. Kirkpatrick. 2007. Turnover of sex chromosomes induced by sexual conflict. Nature 449:909-912.

van Doorn, G. S., and M. Kirkpatrick. 2010. Transitions between male and female heterogamety caused by sex-antagonistic selection. Genetics 186:629-645.

Van Laere, A-S., W. Coppieters, and M. Georges. 2008. Characterization of the bovine pseudoautosomal boundary: Documenting the evolutionary history of mammalian sex chromosomes. Genome Research 18:1884-1895.

van Tuinen, M., and S. B. Hedges. 2001. Calibration of avian molecular clocks. Molecular Biology and Evolution 18:206-213.

van Tuinen, M., and G. J. Dyke. 2004. Calibration of galliform molecular clocks using multiple fossils and genetic partitions. Molecular Phylogenetics and Evolution 30:74-86.

Vibranovski, M. D., H. F. Lopes, T. L. Karr, and M. Long. 2009a. Stage-specific expression profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. PLoS Genetics 5: e1000731.

Vibranovski, M. D., Y. Zhang, and M. Long. 2009b. General gene movement off the X chromosome in the *Drosophila* genus. Genome Research 19:897-903.

Vicoso, B., and B. Charlesworth. 2009. Effective population size and the Faster-X effect: An extended model. Evolution 63:2413-2426.

Vicoso, B., and D. Bachtrog. 2011. Lack of global dosage compensation in *Schistosoma mansoni*, a female-heterogametic parasite. Genome Biology and Evolution 3:230-235.

Vicoso, B., J. J. Emerson, Y. Zektser, S. Mahajan, and D. Bachtrog. 2013a. Comparative sex chromosome genomics in snakes: Differentiation, evolutionary strata, and lack of global dosage compensation. PLoS Biology 11:e1001643.

Vicoso, B., V. B. Kaiser, and D. Bachtrog. 2013b. Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. Proceedings of the National Academy of Sciences, USA 110:6453-6458.

Walters, J. R., and R. G. Harrison. 2011. Decoupling of rapid and adaptive evolution among seminal fluid proteins in *Heliconius* bufferflies with divergent mating systems. Evolution 65:2855-2871.

Wang, B., R. Ekblom, I. Bunikis, H. Siitari, and J. Hoglund. 2014. Whole genome sequencing of the black grouse (*Tetrao tetrix*): reference guided assembly suggests Faster-Z and MHC evolution. BMC Genomics 15:180.

Warren, W. C., D. F. Clayton, H. Ellegren, A. P. Arnold, L. W. Hillier, A. Kunstner, S. Searle, S. White, A. J. Vilella, S. Fairley, et al. 2010. The genome of a songbird. Nature 464:757-762.

Webster, M. T., E. Axelsson, and H. Ellegren. 2006. Strong regional biases in nucleotide substitution in the chicken genome. Molecular Biology and Evolution 23:1203-1216.

Wilkins, J. F., and D. Haig. 2003. What good is genomic imprinting: The function of parent-specific gene expression. Nature Reviews Genetics 4:359-368.

Wilson Sayres, M. A., K. E. Lohmueller, and R. Nielsen. 2014. Natural selection reduced diversity on Human Y chromosomes. PLoS Genetics 10:e1004064.

Wilson, M. A., and K. D. Makova. 2009a. Genomic analyses of sex chromosome evolution. Annual Review of Genomics and Human Genetics. 10:333-354.

Wilson, M. A., and K. D. Makova. 2009b. Evolution and survival on eutherian sex chromosomes. PLoS Genetics 5: e1000568.

Wishart, G.J. 1987. Regulation of the length of the fertile period in the domestic fowl by numbers of oviductal spermatozoa. Journal of Reproduction and Fertility 80: 493–498.

Wolfner, M. F. 2002. The gifts that keep on giving: physiological functions and evolutionary dynamics of male seminal proteins in *Drosophila*. Heredity 88:85-93.

Wright, A. E., H. K. Moghadam, and J. E. Mank. 2012. Trade-off between selection for dosage compensation and masculinization on the avian Z chromosome. Genetics 192:1433-1445.

Wright, A. E., and J. E. Mank. 2013. The scope and strength of sex-specific selection in genome evolution. Journal of Evolutionary Biology. 26:1841-1853.

Wright, A. E., P. W. Harrison, S. H. Montgomery, M. A. Pointer, and J. E. Mank. In press. Independent stratum formation on the avian sex chromosomes reveals inter-chromosomal gene conversion and predominance of purifying selection on the W chromosome. Evolution.

Wright, A. E., P. W. Harrison, F. Zimmer, S. H. Montgomery, M. A. Pointer, and J. E. Mank. In review. Variation in promiscuity and sperm competition drives the evolution of the avian Z chromosome.

Wyman, M. J., A. F. Agrawal, and L. Rowe. 2010. Condition-dependence of the sexually dimorphic transcriptome in *Drosophila melanogaster* Evolution 64:1836-1848.

Xu, K., S. Oh, T. Park, D. C. Presgraves, and S. V. Yi. 2012. Lineage-specific variation in slow- and fast-X evolution in primates. Evolution 66:1751-1761.

Yang, X., E. E. Schadt, S. Wang, H. Wang, A. P. Arnold, L. Ingram-Drake, T. A. Drake, and A. J. Lusis. 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome Research 16:995-1004.

Yang, Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution 24:1586-1591.

Zauner, H., Begemann, G., Mari-Beffa, M. & Meyer, A. 2003. Differential regulation of msx genes in the development of the gonopodium, an intromittent organ, and of the 'sword', a sexually selected trait of swordtail fishes (*Xiphophorus*). Evolution and Development 5: 466–477.

Zhang, J., R. Nielsen, and Z. Yang. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Molecular Biology and Evolution. 22:2472-2479.

Zhang, Y., D. Sturgill, M. Parisi, S. Kumar, and B. Oliver. 2007. Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. Nature 450:233-238.

Zhang, Y. E., M. D. Vibranovski, B. H. Krinsky, and M. Long. 2010a. Age-dependent chromosomal distribution of male-biased genes in *Drosophila*. Genome Research 20:1526-1533.

Zhang, Y. E., M. D. Vibranovski, P. Landback, G. A. B. Marais, and M. Long. 2010b. Chromosomal redistribution of male-biased genes in mammalian evolution with two bursts of gene gain on the X Chromosome. PLoS Biology 8.

Zhou, Q., and D. Bachtrog. 2012. Chromosome-wide gene silencing initiates Y degeneration in *Drosophila*. Current Biology 22:522-525.