UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1788

The evolution of sex chromosomes and sex-linked sequences in birds

HOMA PAPOLI YAZDI

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2019

ISSN 1651-6214
ISBN 978-91-513-0612-4
urn:nbn:se:uu:diva-379665

Dissertation presented at Uppsala University to be publicly examined in Lindahlsalen, Evolutionsbiologiskt centrum, Norbyvägen 18A, Uppsala, Monday, 20 May 2019 at 13:00 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Bengt Hansson (Department of Biology, Lund University).

Abstract
Yazdi, H. P. 2019. The evolution of sex chromosomes and sex-linked sequences in birds. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 1788. 51 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-513-0612-4.

Identifying the processes involved in the evolution of suppressed recombination between sex chromosomes and understanding their consequences for the evolutionary dynamics of sex-linked loci have been major topics of research during the last century. In this thesis, I used the avian ZW system, where females are the heterogametic sex, to investigate the underlying processes in sex chromosome evolution in birds. I identified the gametologous genes between the largely recombining Z and W chromosomes of ostrich and dated the timing of the cessation of recombination to prior to the split of modern birds. I then constructed a genetic map of the ostrich Z chromosome and corrected its assembly in order to obtain the ancestral organization of the Z chromosome in a basal clade of birds. By analyzing the inversion events across the avian phylogeny, I concluded that a combination of Z- and possibly W-linked inversions might have been responsible for the evolution of suppressed recombination in avian sex chromosomes. To understand the determinants of levels of genetic diversity on Z chromosome compared to autosomes, I calculated Z to autosome (Z:A) genetic diversity across 32 avian species. This revealed a broad range of Z:A genetic diversity, between 0.278 – 1.27. Lineage-specific estimates of the nonsynonymous to synonymous substitution rate ratio (d_N:d_S) for autosomal and Z-linked genes further revealed a Fast-Z effect in the majority of birds. The lack of a significant correlation between Z:A d_N:d_S and Z:A genetic diversity indicated that genetic drift might not be sufficient to explain faster evolution of Z-linked genes, suggesting that positive selection might also contribute to the observed values. Finally, I calculated genetic diversity and linkage disequilibrium (LD) along the pseudoautosomal region (PAR) of the Z chromosome using population genomics data of ostrich. In contrast to theoretical expectation, levels of diversity on the PAR were not significantly higher close to the sex-determining region (SDR) compared to autosomal values. Additionally, I observed a lower level of LD on the PAR compared to the average for the Z chromosome and no significant level of LD across the PAR boundary was detected, indicating recombination allows the boundary-proximal region of PAR to behave independently of SDR. Considered together with a higher level of recombination rate in females in the proximity of the SDR, this observation might help explain the maintenance of a long PAR in ostriches and other ratites. Altogether, the results of this thesis make a modest contribution to our understanding of sex chromosome evolution in birds.

Keywords: sex chromosomes, female heterogamety, recombination suppression, genetic map, inversions, genetic diversity, pseudoautosomal region

Homa Papoli Yazdi, Department of Ecology and Genetics, Norbyvägen 18 D, Uppsala University, SE-752 36 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-513-0612-4
urn:nbn:se:uu:diva-379665 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-379665)

To the memory of my best friend,

Akram

تقدیم به یاد و خاطره ی بهترین دوستم،

اکرم

Cover illustration named “Two best friends, Z and W” by Roy M. Francis depicts the degeneration of the W chromosome in the process of recombination cessation. Image credit Roy M. Francis.

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I **Yazdi, H. P.**, Ellegren, H. (2014) Old but Not (So) Degenerated-Slow Evolution of Largely Homomorphic Sex Chromosomes in Ratites. *Molecular Biology and Evolution*, 31(6): 1444 –1453

II **Yazdi, H. P.**, Ellegren, H. (2018) A Genetic Map of Ostrich Z Chromosome and the Role of Inversions in Avian Sex Chromosome Evolution. *Genome Biology and Evolution*, 10(8): 2049-2060

III **Yazdi, H. P.**[*], Bolivar, P.[*], Mugal C. F., Ellegren, H. (-) Variation in the Z Chromosome to Autosomes Ratio of Genetic Diversity across Birds and its Relationship to the Fast-Z Effect. *Manuscript*

IV **Yazdi, H. P.**, Kawakami, T., Unneberg, P., Schou, M. F., Cloete, S. W. P., Cornwallis, C. K. (-) Patterns of Nucleotide Diversity and Linkage Disequilibrium along the Ostrich Pseudoautosomal region. *Manuscript*

[*] These authors contributed equally to this work.

Additional papers

The following papers were published during the course of my doctoral studies but are not part of this thesis.

Dutoit, L., Mugal, C. F., Bolivar, P., Wang, M., Nadachowska-Brzyska K., Smeds, L., **Yazdi, H. P.**, Gustafsson, L., Ellegren, H. (2018) Sex-biased gene expression, sexual antagonism and levels of genetic diversity in the collared flycatcher (*Ficedula albicolli*s) genome. *Molecular Ecology*, 27(18): 3572-3581

Connallon, T.[*], Olito, C.[*], Dutoit, L., **Papoli, H.**, Ruzicka, F., Yong, L. (2018) Local adaptation and the evolution of inversions on sex chromosomes and autosomes. *Philos Trans R Soc Lond B Biol Sci,* 373: 20170423

[*] These authors contributed equally to this work.

Contents

Introduction

Sex chromosomes are an important part of the genome that carry the sex-determining region. They have particular characteristics such as different ploidy, sex-specific mode of inheritance and recombination pattern that distinguish them from autosomes. In the following sections, I first briefly introduce the sex determination mechanism in vertebrates to provide a context for the function of sex chromosomes. I then provide an introduction to the biological features of sex chromosomes with a focus on their level of gene expression and sex-specific recombination pattern. In the next section, I discuss the major steps in the evolution of heteromorphic sex chromosomes, including the causes and mechanisms of recombination suppression between sex chromosomes in their early stage and the subsequent degeneration of the non-recombining chromosome. I then give a background on sex chromosomes within a population genetics context and explain how evolutionary forces such as genetic drift and selection can influence sex chromosomes differently than autosomes due to their special mode of ploidy and inheritance. In the final section, I focus on the evolutionary dynamics of loci in the recombining part of the sex chromosomes, the pseudoautosomal region.

Sexual reproduction and sex determination in vertebrates

Sexual reproduction is a ubiquitous phenomenon across the animal kingdom and entails the fusion of an egg and a sperm in the process of fertilization to form a zygote. The zygote will develop to either a male or a female during the process of sex determination. Sex determination is a developmental process that directs the fate of a sexually reproducing organism towards developing into a male or female (Beukeboom and Perrin 2014). There are a wide variety of sex determination mechanisms throughout the animal kingdom. In vertebrates, the initial triggers of the molecular cascade that determine sex can be environmental, genetic or a mixture of both. In environmental sex determination, at a specific stage during embryonic development, cues from the environment such as temperature or population density trigger the differentiation of the embryo into male or female. In genetic sex determination, however, sex is determined by a gene or chromosomal differences between sexes (Capel 2017). The main players in genetic sex determination are the sex chromo-

somes which carry the sex-determining region (SDR). The SDR may be carried on largely differentiated sex chromosomes as in the case of human and chicken or as in pufferfish, it can be limited to just a single nucleotide polymorphism (SNP) (Kamiya, et al. 2012). Among the genes involved in sex determination cascade, DM-domain transcription factors contribute to sex differentiation across all animal kingdom (Matson and Zarkower 2012) with the *dmrt1* (double sex and mab-3 related transcription factor 1) gene playing a central switch role by switching on the male developmental program and switching off the female program (Beukeboom and Perrin 2014).

Mammals have an XX/XY sex determination system with male heterogamety (XY). In mammals, sex is determined by the dominant male determining gene *Sry* (Sex-determining region Y) located on the Y chromosome (Sinclair, et al. 1990). *Sry* plays a central role in the upstream part of the sex determination cascade which then, through a series of factors, upregulates *dmrt1*, leading to the initiation of male developmental program and inhibition of female development (Matson, et al. 2011). In XX females and in the absence of *Sry*, *dmrt1* is suppressed, leading to the activation of a pathway with the expression of genes that lead to ovary development. In birds, on other hand, *dmrt1* has moved to the top level of hierarchy in the molecular cascade of sex determination. The mechanism of sex determination is not fully understood in birds; however, evidence suggest a dosage dependent mechanism of sex determination (Smith, et al. 2009). Birds have a ZZ/ZW sex determination system with female heterogamety (ZW). Since *dmrt1* is located on the Z chromosome and is absent from the W, males receive a double dosage of DMRT1 and females a single dosage. Similar to mammals, DMRT1 promotes testis development by activating a pathway that leads to suppressing ovary developmental genes (Smith, et al. 2009; Chue and Smith 2011).

Sex chromosomes: A historical and biological perspective

The sex chromosomes are a pair of chromosomes carrying the SDR and differing from autosomes in shape, size, inheritance pattern and various genomic features. The X chromosome in firebug (*Pyrrhocoris apterus*) was the first to be discovered when Hermann Henking observed that one element in male meiosis was only transmitted to half of the sperms (Henking 1891). The nomination of the chromosome 'X' was done to signify its unknown function. Later, by analyzing *Drosophila* individuals with abnormal sex chromosomes, Bridges found evidence for the role of X and Y chromosomes in sex determination (Bridges 1916, 1925). Sex chromosomes in most vertebrates show a strong morphological differentiation. The homogametic sex carries two X chromosomes (females in male heterogamety) or Z chromosomes (males in female heterogamety). In contrast, the heterogametic sex carries one X and Y

chromosome (males in male heterogamety) or one Z and W chromosome (females in female heterogamety).

The heteromorphism of sex chromosomes has consequences on gene dosage in the heterogametic sex. Dosage of a chromosome refers to its copy number in the genome and is proportional to the level of gene expression as measured by the number of mRNAs or proteins (Ercan 2015). Maintenance of correct dosage is important for the organism's fitness because deviation from the diploid dosage of genes can be detrimental (Torres, et al. 2008; Tang and Amon 2013). Genes located on the X and Z chromosomes are found only in one copy in males and females, respectively. The single copy number of genes on these chromosomes means that there is a reduced level of expression of X and Z linked genes in the heterogametic sex. To alleviate the deleterious effect of dosage imbalance, various mechanisms of dosage compensation have evolved.

These mechanisms can range from a chromosome-wide compensation in mammals as well as systems in which compensation seems to act on individual genes (Graves 2016). In mammals, to compensate for the reduced dosage of X chromosomes in males, one of the X chromosomes in females is epigenetically silenced throughout development, maintaining equal dosage between males and females (Graves 2016). For example, the female to male expression ratio in different species of primates was shown to be nearly 1.0 (Julien, et al. 2012). In birds, on the other hand, the mechanism of dosage compensation is not chromosome-wide (Ellegren, et al. 2007; Julien, et al. 2012). In chicken, Z genes were shown to express 30-40% more strongly in ZZ males than in ZW females (Ellegren, et al. 2007; Julien, et al. 2012). Lack of a global mechanism for dosage compensation has also been shown in flycatcher (Uebbing, et al. 2013) and the W-degenerate segment of the ostrich Z chromosome (Adolfsson and Ellegren 2013).

Another specific feature of sex chromosomes is their sex-specific pattern of recombination. Heteromorphic sex chromosomes do not recombine along most of their length in the heterogametic sex. Recombination occurs along the X chromosome in females in male heterogamety (Z chromosome in female heterogamety) but it is restricted to a segment called the pseudoautosomal region (PAR) between X and Y (or Z and W) chromosomes. Diploid PARs share some features with autosomes but they are still partly sex-linked in the region close to the SDR (Otto, et al. 2011). Recombination rate is high in the small PAR and is even higher in the heterogametic sex due to the requirement of at least one cross-over for proper segregation of chromosomes during meiosis (Perry, et al. 2001; Hinch, et al. 2014; Smeds, et al. 2014). The PAR can also

consist of the majority of the sex chromosome in recently evolved sex chromosome systems (Bewick, et al. 2013) or in certain ancient systems such as in ratites (Otto, et al. 2011; Vicoso, Emerson, et al. 2013; Vicoso, Kaiser, et al. 2013; Yazdi and Ellegren 2014). The suppression of recombination is the cornerstone in the evolution of heteromorphic sex chromosomes. It is therefore important to understand the why and how of recombination suppression during the evolution of sex chromosomes. This is the topic for the next section.

The evolution of heteromorphic sex chromosomes through recombination suppression

Evolution of sex chromosomes has been a major topic of research during the last century. The earliest theory concerning the evolution of heteromorphic sex chromosomes was put forward by Muller (Muller 1914, 1918) with credit given to A. H. Sturtevant (Clark 1988). All the research has led to a consensus that sex chromosomes evolved from a pair of autosomes that gained a sex-determining function which led to the suppressed recombination between them. The sex chromosome in mammals and birds evolved independently from a separate pair of autosomes (Fridolfsson, et al. 1998), however, in both cases, the common feature is the cessation of recombination around the SDR. Several models have been suggested for the evolution of recombination suppression (Nei 1969; Charlesworth and Charlesworth 1980; Rice 1987a). At the heart of all these models is selection for reduced recombination between the SDR and a mutation with sexually antagonistic fitness effect (i.e., beneficial to one sex but detrimental to the other). The process of recombination suppression can be considered to take place in two steps: First is the initial recombination cessation around the SDR and second is the expansion of the segment with suppressed recombination through the gain of sexually antagonistic mutations (Charlesworth, et al. 2005).

Cessation of recombination in the proximity of the SDR can be studied together with the evolution of two sexes from a co-sexual population (Charlesworth & Charlesworth 1978). In the evolution of SDR coupled with the evolution of two sexes, at least two mutations are needed. The first mutation is a recessive male-sterility mutation which renders the individuals to become female. At this moment, the population will consist of females and hermaphrodites. In the next step, a dominant female sterility gene will create males. Recombination between the recessive male sterility and dominant female sterility mutations must decrease in order to avoid the production of intersex individuals with lower fitness.

After the initial suppression of recombination in the proximity of the SDR, in most sex chromosome systems, recombination continues to cease in a larger

segment. The main theory concerning the extension of recombination cessation is the accumulation of sexually antagonistic alleles (Rice 1987a). The main factors in selection for recombination suppression is the genetic distance between the sexually antagonistic locus and the SDR and the selective advantage of the sexually antagonistic mutation. In the case of XY system, the occurrence of a male beneficial mutation in the proximity of the *Sry* on Y chromosome could select for the cessation of recombination between *Sry* and the male beneficial allele. Given high selective advantage in males, the male beneficial mutation can spread into population even if detrimental to females, selecting for the extension of the non-recombining region (Rice 1987a). The ZW system is often considered to be equivalent to the XY system in this aspect by exchanging W for Y and male beneficial alleles occurring on Y for female beneficial alleles occurring on W. However, recent empirical work has failed to identify any female specific or female-beneficial genes on the W chromosome (Smeds, et al. 2015; Bellott, et al. 2017). As shown in the previous section, sex is determined by the dosage of DMRT1 on the Z chromosome in birds and there has been identification of male-beneficial alleles on the Z chromosome (Bellott, et al. 2010; Ellegren 2011).

Suppression of recombination between sex chromosomes leads to a gradual sequence divergence of the homologous X and Y (Z and W) linked loci from one another. The term, evolutionary strata, was coined to describe the stepwise pattern of recombination cessation between the sex chromosomes with each stratum dating back to a distinctive amount of time in the past (Lahn and Page 1999). Evolutionary strata are inferred from levels of sequence divergence between homologous genes on the non-recombining part of the sex chromosomes (i.e., gametologous genes). Examples of evolutionary strata have been reported in sex chromosomes of a variety of species, both in systems of male and female heterogamety, including mammals (Lahn and Page 1999; Sandstedt and Tucker 2004), birds (Handley, et al. 2004; Nam and Ellegren 2008; Wright, et al. 2014; Yazdi and Ellegren 2014), reptiles (Vicoso, Emerson, et al. 2013), fish (Roesti, et al. 2013; White, et al. 2015), and plants (Nicolas, et al. 2005; Bergero, et al. 2007; Wang, et al. 2012), as well as in mating-type chromosomes of fungi (Branco, et al. 2017). Structural rearrangements or molecular mechanisms can lead to cessation of recombination and create a stepwise or gradual pattern of sequence divergence between gametologs. These mechanisms are the topic for the next section.

Mechanisms of recombination suppression between sex chromosomes

Several mechanisms can prevent the pairing of chromosomes and hence lead to recombination cessation. Structural rearrangements such as chromosomal

inversions, through the reduction in fitness of the heterokaryotype individual (Kirkpatrick 2010) and inhibition of the normal pairing of chromosomes during meiosis (Griffiths 2000), can lead to recombination suppression over their length. This idea dates back to the original suggestion by Ohno (Ohno 1967) that recombination arrest between sex chromosomes can be established by a pericentric inversion (i.e., inversion containing the centromere). In humans, it was suggested that recombination between X and Y chromosomes stopped in a step-wise manner through a series of inversions on the Y chromosome (Lahn and Page 1999; Ross, et al. 2005).

The role of inversions was also highlighted in the case of avian sex chromosomes with inversions on Z suggested to be responsible for recombination arrest prior to the split of modern birds (Wright, et al. 2014; Zhou, et al. 2014). However, inversions, along with other types of structural rearrangements such as accumulation of transposable elements and other repetitive sequences can be as well the consequence of recombination cessation due to a decrease in effective population size (N_e) and hence the reduction in the efficacy of selection (Beukeboom and Perrin 2014). Moreover, in ancient sex chromosome systems, connecting inversions to the events of recombination cessation is quite challenging (Lemaitre, et al. 2009; Yazdi and Ellegren 2018). A prediction of cessation of recombination by chromosomal inversions is that the genomic loci within the inverted segments should show similar levels of divergence and in a phylogenetic analysis, gametologous genes resulting from recombination suppression due to an inversion event should cluster by chromosome rather than by species.

In contrast, genetic modifiers of recombination would potentially cause a more gradual spread of cessation of recombination (Chibalina and Filatov 2011; Bergero, et al. 2013; Natri, et al. 2013). For example, heterochromatinization could occur in a gene-by-gene basis. In papaya, the sex-determining gene is close to the centromere with an already reduced rate of recombination (Wang, et al. 2012). In that case, the heterochromatin state of the centromere provides a state of an already reduced recombination for the SDR.

Causes of the degeneration of the non-recombining sex chromosome

Once recombination has stopped, the non-recombining region on Y or W chromosomes starts to degenerate (Charlesworth and Charlesworth 2000). The processes involved in degeneration of the non-recombining chromosome have been a major subject of investigation for decades but the exact process or the relative role of each has been proven to be very difficult to determine empirically. Based on theoretical work, at the center of all these processes is the

consequence of reduction in N_e of the non-recombining chromosome. The main mechanisms suggested for the degeneration of DNA sequence of the non-recombining chromosome include Muller's ratchet (Muller 1964; Felsenstein 1974), hitchhiking effect of mutations with deleterious (background selection) or advantageous (selective sweep) fitness effects (Maynard-Smith and Haigh 1974; Charlesworth, et al. 1993; Charlesworth 1994) and Hill-Robertson interference (Hill and Robertson 1966).

Muller's ratchet is the process of stochastic loss of the class of chromosomes carrying the least number of deleterious mutations in a finite population. In the absence of recombination or back mutation (Charlesworth and Charlesworth 2000), no class of chromosomes with fewer mutations can occur, hence the population gets trapped in a situation with all individuals harboring a number of deleterious mutations. Therefore, every time an individual with the least number of mutations dies (e.g., one mutation), population will consist of individuals with at least two deleterious mutations and so forth. In this way, deleterious mutations such as deletions accumulate on the non-recombining chromosome, leading to its gradual loss of DNA sequence.

Hitchhiking effect of mutations with a fitness effect on the linked neutral loci can further reduce the N_e of the non-recombining chromosome, leading to further fixation of deleterious mutations. In a non-recombining population, any neutral mutation can only survive if it appears on the background of a chromosome free from deleterious mutations (Charlesworth 1994). Background selection leads to a reduction in the mean fitness of the Y (W) chromosomes relative to X (Z) by reducing the N_e of the non-recombining chromosome. This will further accelerate the fixation of deleterious mutations (Charlesworth and Charlesworth 2000). Similar effect can also occur through the effect of an advantageous allele on the neighboring neutral loci. Selective sweep can erode all genetic variation on the non-recombining sex chromosomes and lead to the fixation of deleterious mutant alleles along with the advantageous mutations, contributing to its further degeneration (Rice 1987b). Finally, in Hill-Robertson interference, loci that are both under selection influence the efficacy of selection through the action of one on another. This leads to the inhibition of removal of deleterious mutations and fixation of beneficial ones (Hill and Robertson 1966) which in the long term reduces the fitness of an evolving Y chromosome and contributes to its eventual loss of genetic material.

Population genetics of sex chromosomes

In this section, I only consider the female heterogametic system, the same arguments apply to the male heterogametic system by exchanging males for females, Z for X and W for Y. Sex chromosomes spend unequal amount of their

evolutionary time in males and females. The Z chromosome spends two-thirds of the time in males and one-thirds of the time in females and W chromosome is only inherited through female line. Moreover, for every male and female in the population, there exists three Z chromosomes and one W chromosome for every four autosomes. The sex-linked inheritance pattern and reduced N_e of sex chromosomes can have consequences on the action of evolutionary forces on these chromosomes and their expected levels of genetic diversity. The expected level of genetic diversity is a product of N_e and mutation rate (μ) and is expressed as ($\theta = 4N_e\mu$). Under the assumption of random variance of offspring number for males and females, the Z chromosome to autosomes (Z:A) ratio of N_e is expected to be 0.75 (Caballero 1995). Hence, assuming an equal mutation rate between Z chromosome and autosomes, the Z:A genetic diversity is also expected to be 0.75. However, several factors that can affect N_e and mutation rate differently on Z chromosomes compared to autosomes can cause deviation of genetic diversity from the expected 0.75. Over longer times, these differences will translate into differences in sequence divergence between sex chromosomes and autosomes. In what follows, I introduce the factors that can affect the N_e of sex chromosomes, explain the impact of male bias in mutation rate on levels of genetic diversity, briefly discuss the sex-specific recombination rate of sex chromosomes and finally introduce the consequences of hemizygosity of Z-linked loci for the action of selection.

Genetic drift

Genetic drift is the random loss of genetic diversity due to binomial sampling in a small or finite population size (Hedrick 2007) and is an important factor in the evolution of neutral loci in small populations. The extent of genetic drift is inversely related to N_e. The expected N_e for a Z locus to an autosomal gene in a population consisting of males and females can be calculated from the following equation:

$$\frac{N_{e_Z}}{N_{e_A}} = \frac{\frac{9N_{ef}N_{em}}{4N_{ef} + 2N_{em}}}{\frac{4N_{ef}N_{em}}{N_{ef} + N_{em}}} = \frac{3}{4}$$

where N_{ef} is the N_e in females and N_{em} is the N_e in males. The Z:A N_e of 0.75 is only maintained if $N_{ef} = N_{em}$ (Figure 1). The Z:A N_e can be influenced by several factors including life history traits, fluctuations in population size and the effect of linked selection. Variance in reproductive success causes a deviation from the randomly distributed number of offspring for males and females. For example, in polygynous mating system where one male mates with

multiple females, the N_e of Z chromosomes compared to autosomes is reduced. In addition to variance in reproductive success, N_e of sex chromosomes can be affected differently than autosomes in the face of demographic changes. For example, in the case of population size reduction or bottleneck, the N_e of Z chromosome is reduced more drastically than that of autosomes which leads to even faster rate of coalescence for Z-linked loci, reducing Z:A N_e (Pool and Nielsen 2007).

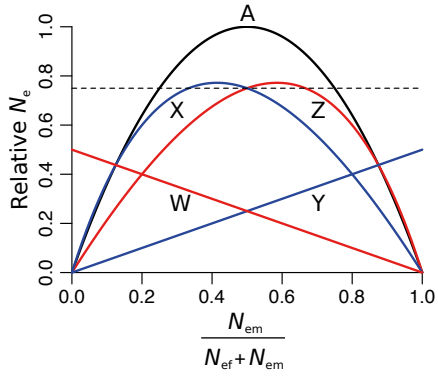

Figure 1. Relative N_e of sex chromosomes in relation to relative N_e of males and females.

Finally, linked selection, the diversity-reducing effect of mutations with fitness effect on linked neutral loci (Maynard-Smith and Haigh 1974; Charlesworth, et al. 1993) can reduce N_e at the proximity of sites under selection. Since the Z chromosome recombines only in the homogametic sex, it has a lower rate of sex-averaged recombination rate, which might increase the impact of selection on linked neutral sites leading to a reduction in genetic diversity.

Mutation rate

Mutation rate is the probability of occurrence of an error in DNA replication per unit of time, most commonly measured in generation. Germ line mutations get inherited from one generation to the next and therefore provide the raw material for genetic variation. Male gametes typically undergo many more rounds of cell division compared to female gametes during the process of gametogenesis, a consequence of the higher production of sperm than egg cells. This has led to the suggestion that the mutation rate should be higher in males compared to females, a phenomenon known as male-biased mutation (Miyata, et al. 1987; Ellegren 2007; Gao, et al. 2016). Male-biased mutation has been subsequently shown in a variety of species (Ellegren 2007, 2011; Wilson Sayres and Makova 2011).

Since Z chromosomes spend two-thirds of the time in males, they can be influenced by male-biased mutation, leading to an increase in male to female (male:female) mutation ratio. The relationship of male:female mutation ratio with Z:A genetic diversity is shown in the equation below:

$$\frac{\theta_Z}{\theta_A} = \frac{9(N_{ef} + N_{em})}{8(2N_{ef} + N_{em})} * \frac{2(2\alpha + 1)}{3(\alpha + 1)}$$

where α is the male:female mutation ratio, θ_Z and θ_A are the genetic diversity on Z and autosomes, respectively. N_{ef} and N_{em} are the effective population sizes in females and males, respectively. Based on the above equation, it is expected to have an increase in Z chromosome genetic diversity compared to autosomes with an increase in male:female mutation ratio (Figure 2A and B).

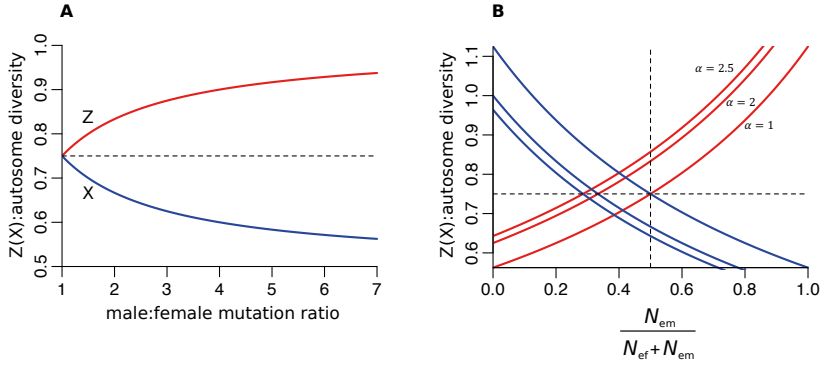

Figure 2. Effect of male-biased mutation in Z(X): autosome level of genetic diversity. A) Z(X): autosome genetic diversity and its relation to male:female mutation ratio for $N_{ef} = N_{em}$. B) Z(X): autosome genetic diversity and its relation to the relative N_e of males and females plotted for different male:female mutation ratio. Red and blue represent Z and X chromosome, respectively

Recombination rate

Recombination is the process of exchange of DNA sequence between two homologous chromosomes. The rate of recombination is measured as the probability of cross-over event that involves the exchange of DNA sequence per unit of physical distance for example million base pairs (Mb). There is a heterogeneous pattern of recombination across the genome and between sexes in different organisms (Petkov, et al. 2007; Groenen, et al. 2009; Kawakami, et al. 2014; Bherer, et al. 2017). The most striking sex-difference in recombination rate occurs between sex chromosomes. In the heterogametic sex, recombination between Z and W chromosomes is confined to the PAR. On the other hand, in the homogametic sex, recombination occurs along the Z chromosomes. The sex-averaged recombination rate for autosomal genes is the arithmetic mean of recombination rate between males and females as despite sex

difference in recombination, autosomal genes spend equal amount of times in males and females on average. However, since there are three Z chromosomes in the population and two undergo recombination in males, the average recombination frequency for the Z chromosome is:

$$\bar{r} = \frac{2r_m}{3}$$

where $\bar{r}$ is the sex-averaged recombination frequency and r_m is the recombination frequency in males.

Selection

Genes on Z chromosome with degenerate W chromosome might exhibit different patterns of evolution in comparison with autosomal genes due to their hemizygosity in the heterogametic sex (Charlesworth, et al. 2018). The Z:A ratio of substitution rate depends on the dominance coefficient and the sex-specific effects of mutations on fitness (Avery 1984; Rice 1984; Charlesworth, et al. 1987). For example, the Z:A ratio of substitution rate is larger than one for partially recessive mutations that are expressed in both sexes or in females. In this case, a Fast-Z effect is expected in which genes on the Z chromosome show faster rate of evolution due to the expression of recessive mutations in the heterogametic sex. Moreover, if mutations are sexually antagonistic, a dominant mutation on the Z chromosome will be expressed twice in males than in females, hence, it will increase in frequency if male beneficial and decrease in frequency if detrimental to males. In contrast, due to the hemizygosity of Z-linked genes, a recessive mutation will be initially expressed in females. Such mutation will be disfavored if it is detrimental to females or increase in frequency if it is beneficial in females (Rice 1984). In addition to Z-linked sequences with degenerate W which are suitable targets for the accumulation of sexually antagonistic genes and polymorphisms, the pseudoautosomal region of the Z chromosome, particularly in the segment close to the SDR is also an important target for sexually antagonistic mutations and is discussed in the last section of the introduction.

Evolutionary dynamics of loci on the pseudoautosomal region

While most attention in studies of sex chromosome evolution has been concentrated on the sex-specific region of sex chromosomes, the PAR can contain information about the evolutionary dynamics of sex-linked sequences. It has been assumed that the evolutionary dynamics of loci on the PAR should resemble that of autosomes as both are diploid and undergo recombination in both sexes. However, PAR loci that are in the proximity of the SDR, depending on the recombination rate in the heterogametic sex, may be sex-linked by

spending unequal amount of times in males and females. The sex-linked inheritance pattern of these loci increases the expected coalescence time as a function of genetic distance to the SDR (Kirkpatrick, et al. 2010) with the highest expected coalescence time for loci closest to the SDR. Moreover, the sex-linkage of the PAR loci close to the SDR make them a suitable substrate for the accumulation of sexually antagonistic mutations (Otto, et al. 2011; Kirkpatrick and Guerrero 2014) which leads to further increase in coalescence time between the two sex chromosomes and might manifest itself in higher levels of genetic differentiation between males and females.

Methods

In this section, I describe the key methods that I have used to address different questions in sex chromosome evolution.

Dating the timing of recombination suppression between gametologous genes

Identification of W-linked gametologs

Once the Z and W chromosomes stop to recombine, genes on the non-recombining region of the W chromosome will independently accumulate mutations. In the genotyping of sequenced individuals, substitutions in a still existing W-linked gametolog will be recognized as female-specific SNPs when reads from such a gene are mapped to the corresponding Z-linked gene given enough sequence conservation for the reliable mapping. Since females have one copy of the Z chromosome and one copy of the W chromosome, Z-linked genes outside the PAR for which the W copy has been degenerated are hemizygous in females and their genotypes are detected as homozygous in sequencing. W chromosome gametologs were identified as genes with heterozygous SNP calls in individual females for which they consistently showed the Z chromosome reference allele and a unique allele not previously identified as segregating sites in sequencing of male individuals. The sequence of the gametologous genes was then built by tilling reads originating from the W-linked copy identified by containing the non-reference allele at heterozygous sites.

Estimating divergence time between Z-linked and W-linked genes and the strength of selection

With the sequence of Z and W gametologs at hand, we can use the concept of molecular clock to calculate the timing of the divergence of Z and W chromosomes (i.e., the timing of the cessation of recombination) using the amount of sequence divergence between the two chromosomes. Molecular clock is the concept of a steady rate of change in DNA sequences over time which provides a basis for dating the time of divergence of lineages if the rate of change can be estimated (Futuyma and Kirkpatrick 2018). The time of divergence can then be calculated using the following equation:

$$t = \frac{d}{2\mu}$$

where t is the amount of time measured in generations or years since the two lineages have diverged, d is the amount of sequence divergence and μ is the mutation rate.

Using the principle above, in order to get an estimate of divergence time, we calculated the pairwise synonymous substitution rate (d_S) for the gametologous genes. Next, we needed to obtain a mutation rate specific to the lineage we are interested in, in this case, the ratites. Under neutrality, substitution rate is equal to the rate of mutation (Kimura 1971). To obtain a ratite-specific mutation rate, we used the emu-ostrich alignment of orthologous genes and calculated d_S. We next used the fossil calibrated divergence time between the two species of about 97 MY (Haddrath and Baker 2012). Using the equation above, we estimated a ratite-specific substitution rate (a proxy for mutation rate) as 1.1×10^{-9}. We next used this estimate of mutation rate and the calculated d_S to obtain the time of divergence between the Z and W gametologs.

In order to infer how selection has been acting on Z and W gametologs, we estimated the lineage-specific d_S and d_N for Z and W chromosomes using chicken orthologs as the outgroup. In lineage-specific calculation of sequence divergence, the outgroup can inform us about the direction of change in the in-group, in this case, the ostrich Z or W sequences. The ratio of the two, d_N:d_S can be used to infer selection in a given lineage with values above one indicating positive selection or relaxation of purifying selection, equal one indicating neutral evolution and below one, indicating purifying selection.

Genetic mapping

Recombination between genes on a chromosome

Homologous chromosomes can undergo an exchange of their segments during meiosis. The exchange of DNA segments (crossing over) between homologous chromosomes can result in recombination, yielding daughter chromosomes that carry combinations of alleles that are not present in the parents. Frequency of recombination is the proportion of gametes carrying combination of alleles that are not present in the parental chromosome (Hartl 2005). In order to know whether a combination of alleles on a chromosome is different from the parental type, it is important to know the configuration of the alleles along the chromosome (i.e., the phase). For example, for the genotype *AaBb*, two configurations are shown in Figure 3.

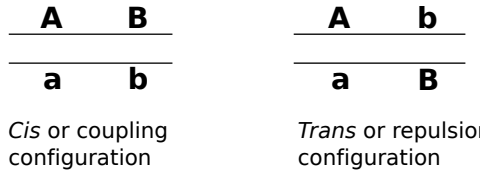

Figure 3. Configuration of the alleles *a* and *b* on the chromosome. The *cis* or coupling configuration has the alleles on the same chromosome while the *trans* or repulsion configuration has the two alleles on different chromosomes.

Basics of genetic mapping

The linkage of genes on a chromosome can be represented in the form of a genetic map, which shows the linear order of genes along the chromosome with the distances between adjacent genes proportional to the frequency of recombination between them (Hartl 2005). The unit of distance in a genetic map is called a map unit. Over short intervals, 1 map unit equals 1 percent recombination which is also called a centimorgan (cM). Since the maximum frequency of recombination is 50 percent, the maximum length of a chromosome is expected to not exceed 50 cM. However, the genetic map length often exceeds 50 cM. For example, the ostrich sex-averaged map of Z chromosome is about 94.2 cM. The reason for a longer length of chromosome than the maximum frequency of recombination is that more than one cross-over can occur on a chromosome.

Double cross-overs in a chromosomal segment between genes can remain undetected since it may not result in recombinant offspring. This is important since map units measure how much crossing-over takes place between genes while the recombination frequency reflects how much recombination is actually observed (Hartl 2005). Thus, double recombinants that cannot be observed do not contribute to recombination frequency but they do contribute to the map distance. If the distance between two genes or genetic markers is so short that the probability of double cross-over is very low, map unit and recombination frequency are the same.

The occurrence of one cross-over event reduces the probability of the cross-over event in the proximity of the first one. This phenomenon is called chromosome interference. In order to integrate the effect of interference in connecting recombination frequency with genetic distance, different mapping functions are used. Mapping functions convert a map distance between genetic markers into their recombination frequency. One of the most used mapping functions is the Kosambi mapping function (Kosambi 1944) which assumes

that interference decreases as a linear function of distance, therefore, larger map distances than 50 cM are possible.

Maximum likelihood and LOD scores in genetic mapping

In organisms such as ostrich that there are few offspring per mating and often, the phase of the alleles on a chromosome cannot be determined, information from multiple pedigrees are combined to have enough power to test the hypothesis of linkage between a pair of markers against the hypothesis of free recombination. In order to obtain a genetic map, a pedigree or mapping population of a minimum of two generations (Parents and F1) is needed. The principle of obtaining the genetic map from multiple pedigrees is that for each pedigree, the probability of pedigree, that is the number of recombinant and non-recombinant offspring is calculated given a recombination frequency (r). Next, the probability of pedigree given free recombination ($r = 0.5$) is calculated. Finally, for each pedigree, the ratio of the probability of the pedigree given an arbitrary value of recombination frequency to that with $r = 0.5$ is calculated which is called the likelihood ratio. By taking the logarithm of the likelihood ratio, the logarithm of odds or LOD score is calculated (Hartl 2005).

The frequency of recombination is estimated as that value of r that maximizes the LOD score for all the pedigrees. This is the principle of genetic mapping when using LOD score across many independent pedigrees each with a few number of offspring. When many loci are used, the calculations above must be done for every pair of loci which requires specific algorithms to increase efficiency. Such algorithms are implemented in various genetic mapping software such as the one used in this thesis, LEPMAP (Rastas, et al. 2015).

Analysis of chromosomal inversions

Structural rearrangements are one of the major algorithmic challenges in genome evolution. The problem of chromosomal inversions is often described as the problem of sorting a signed permutation (Bourque and Pevzner 2002; Lemaitre, et al. 2009). Each section of genome that is syntenic to the genome of another organism can be given a signed index. The goal is then to find the parsimonious solution for sorting this signed permutation. Here, the parsimonious solution is the one with the least number of inversion steps necessary to sort the signed permutation, that is to obtain the ancestral genome form. One of the algorithm used is called GRIMM (Tesler 2002).

In order to infer the inversions between two species, we must first infer synteny blocks between the species of interest. Synteny blocks are segments that can be converted into regions with conserved gene order without disrup-

tion by micro-rearrangements (Nadeau and Taylor 1984). In Figure 4, the principle of inversion analysis by sorting the signed permutation is shown. The source genome can be considered an ancestral genome and is represented with the number of synteny blocks (1, 2, 3, 4) and the destination genome, a lineage in which we want to infer the inversion events since the ancestor contains a series of inverted blocks (-3, -1, 2, 4). As shown in Figure 4, two blocks have a reversed orientation in the destination genome (-3, -1) compared to the source genome, however, three inversions are necessary to obtain the gene order and orientation in the destination genome. The set of parsimonious solution may involve different set of inversions but only one is reported by GRIMM (Braga 2009). Therefore, it is not possible to know the accuracy of the set of breakpoints obtained from a given inversion scenario.

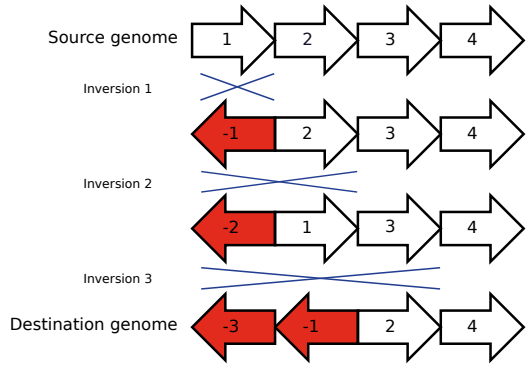

Figure 4. Principle of inversion analysis by sorting a signed permutation. Each arrow indicates a synteny block. Red indicates an inverted block. Blue crosses indicate the inversion event.

Genetic diversity: Heterozygosity

The neutral genetic diversity depends on the mutation rate and the effective population size (N_e). Assuming a small rate of mutation per site, the average heterozygosity (H) over the sites of a sequence is expected to be the product of N_e and mutation rate:

$$H = 4N_e\mu$$

A general way to quantify the amount of nucleotide variation in a population is to determine the proportion of nucleotide differences between pairs of sequences and then weigh these differences by the frequencies of the sequences and sum over all of the possible pairs of sequences (Nei and Li 1979). The following equation is used to calculate heterozygosity:

$$\pi = \sum_{ij} p_i p_j \pi_{ij}$$

where p_i and p_j denote the respective frequencies of the ith and jth sequences, π_{ij} denotes the number of nucleotide differences per nucleotide site between the ith and jth sequences.

Genetic differentiation: F_{ST}

The amount of differentiation between populations can inform us about several features of populations including the amount of time that they have last shared a common ancestor. One of the statistics used to measure differentiation between populations is F_{ST} and one of the most commonly used estimators of F_{ST} is Nei's F_{ST} (Nei 1973) which is computed as follows:

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

where H_T is the heterozygosity of the total population and H_S is the heterozygosity of the subpopulation. In addition to the use of F_{ST} as a measure of population differentiation, one can also use F_{ST} to measure the amount of differentiation between different groups for example between males and females. An increase in differentiation between males and females might indicate that a process such as sexually antagonistic selection is going on in that specific region of the genome which increases the frequency of allele under selection in one sex compared to the other (Cheng and Kirkpatrick 2016). However, it is important to know that elevated values of F_{ST} merely represent a certain level of differentiation among populations or groups of individuals and do not convey any information about the underlying processes. Increased F_{ST} at a genomic location compared to the background can be due to population structure as a result of limited migration or it can be a signal of positive selection.

Linkage disequilibrium: r^2

Linkage disequilibrium (LD), the non-random association of alleles (Slatkin 2008), means that there is a tendency for a certain allele to segregate with another more often than is expected entirely by chance. The definition for LD between a pair of loci is called *D* and is measured as the following quantity:

$$D = p_{AB} - p_A p_B$$

which is the difference between the frequency of gamete carrying A and B alleles, p_{AB}, and the expected frequency if alleles associated randomly on chromosomes, $p_A p_B$ (Gillespie 2004). A commonly used measure of LD which is less sensitive to allele frequencies is r^2. r^2 is the square of correlation coefficient between pairs of alleles at two loci that segregate together (Gillespie 2004) and is calculated as:

$$r^2 = \frac{D^2}{p_A(1-p_A)p_B(1-p_B)}$$

Several factors can influence levels of LD in population. Natural selection might favor a certain combination of alleles, leading to an increase in the frequency of specific haplotypes. Moreover, genetic drift can create non-random associations between alleles due to random sampling (Gillespie 2004). For loci on the same chromosome, LD decreases with an increase in recombination rate, therefore, in the context of sex chromosomes, loci on the pseudo-autosomal region are expected to experience lower levels of LD than loci within the non-recombining part of the chromosome.

Research Aims

The main objective of this thesis was to study the processes involved in the evolution of suppressed recombination in avian sex chromosomes and levels and patterns of genetic diversity on Z chromosomes across avian phylogeny. The specific aims of each of the presented papers of this thesis are listed below:

I. Identifying homologous genes in the non-recombining part of the ostrich sex chromosomes (i.e., gametologous genes) and estimating the timing of the cessation of recombination between Z and W chromosomes.

II. Constructing the genetic map of ostrich Z chromosome, inferring the number of inversion events across major avian lineages and investigating the role of Z-linked inversions in the evolution of recombination suppression in avian sex chromosomes.

III. Investigating the relative levels of Z chromosome to autosomes (Z:A) genetic diversity across avian phylogeny, exploring the determinants of Z:A levels of genetic diversity and its relationship to the faster evolution of Z-linked genes.

IV. Studying patterns of polymorphism across the pseudoautosomal region of ostrich Z chromosome in order to understand the effect of sex-linkage on the evolutionary dynamics of PAR loci.

Summary of the papers

Paper I

Old but Not (So) Degenerated – Slow Evolution of Largely Homomorphic Sex Chromosomes in Ratites

The main feature of sex chromosome evolution is the cessation of recombination along the proto-sex chromosomes. Recombination suppression leads to the degeneration of the non-recombining chromosome which is seen in a variety of taxa with highly differentiated sex chromosomes. The non-recombining chromosome (Y in male heterogametic system such as in mammals, W in female heterogametic system such as in birds) is usually in heterochromatic state and has an accumulation of repeat elements. However, in ratites (order Palaeognathae, including ostrich), the Z and W chromosomes have similar size and have maintained recombination along the majority of their length despite sex chromosome evolution was initiated more than 130 MYA.

In order to investigate the timing of the cessation of recombination and patterns of divergence between Z and W chromosomes in ostrich, we used transcriptomic data from six female ostriches to identify homologous genes that have stopped recombining between the sex chromosomes but have not yet been degenerated. I identified the W-linked gametologs by identifying genes that consistently showed a non-reference allele not previously found in male sequencing. We identified fourteen W gametologs and by aligning them with the Z gametologs and orthologous chicken genes, we could date the timing of the cessation of recombination and the amount of selection on Z and W gametologs.

The synonymous sequence divergence in the pairwise comparison of gametologous Z and W chromosome copies varied between 0.027 and 0.177. Using ratite-specific substitution rate of 1.1×10^{-9} per year, we dated the timing of cessation of recombination between the gametologous genes around 24-157 MYA. Based on the position of genes on Z chromosome in chicken, we found a clear correlation between chromosomal position and divergence.

We next sought to determine the strength of selection. The ratio of nonsynonymous to synonymous substitutions (d_N:d_S) ranged between 0.033 and 0.491.

One gene, *SMC5* contained three premature stop-codons and had the highest d_N:d_S. The presence of stop codons in the coding sequence of the W-linked gametolog is indicative of degeneration and relaxed constraint. To test whether Z and W genes have been subject to different selection pressures, we used lineage-specific rate of sequence divergence using chicken as the outgroup. All genes had a higher d_N:d_S in the W-linked than in the Z-linked gametologs. d_N:d_S was two times larger in the W-linked copy than in the Z-linked copy for seven genes, with *SMC5* showing the most pronounced difference, d_N:d_S = 1.425 for *SMC5W* versus d_N:d_S = 0.064 for *SMC5Z*.

Recombination rate can affect the base composition. Consistent with this, we observed that the ratio of the number of AT→GC to the number of GC→AT substitutions was higher in the Z-linked than in the W-linked gametologs which suggests that Z-linked gametologs evolve towards a higher GC content than their W-linked copies. Finally, we observed that the male to female expression ratio (M:F expression ratio) was close to one in some of the gametologs indicating an active expression of genes in females while in others, expression in females had become severely reduced with M:F expression ratios close to 2.

The observations in this study, including a gradual increase in the level of divergence along the Z and W chromosomes that have stopped recombining, the higher level of d_N:d_S on the W chromosome indicative of relaxed selective constraint and the difference in M:F expression ratio across the gametologous genes showed that sex chromosome evolution in ostriches follows similar path as of other sex chromosome systems. However, the rate of degeneration of the W chromosome in ostriches has been quite slow. We suggested that lack of a global mechanism for dosage compensation could be a factor preventing the degeneration of W chromosome in this lineage of birds.

Paper II

A Genetic Map of Ostrich Z chromosome and the Role of Inversions in Avian Sex Chromosome Evolution

Recombination arrest is a necessary step for the evolution of distinct sex chromosomes. One of the mechanisms that can cause recombination suppression is the occurrence of structural changes such as inversions. Inversions are structural rearrangements in which a chromosomal segment is cut at two locations and is replaced in the same position with a reversal of 180 degrees. In order to reveal the organization of sex chromosomes in a basal avian lineage, identify the position of gametologous genes along the ostrich Z chromosome and study

the role of inversions in avian sex chromosome evolution, we developed a sex-specific linkage map of ostrich Z chromosome.

We genotyped 384 SNPs from the putatively Z-linked scaffolds in 22 full-sib families. We found two linkage groups. The first linkage group had a male genetic map length of 92.3 cM and a female genetic map length of 81.2 cM with a sex-averaged map length of 94.2 cM. The second linkage group had a sex-averaged genetic length of 6.6 cM. The identification of two linkage groups was surprising since the markers were selected from putatively Z-linked scaffolds. We found that the first linkage group was syntenic to chicken chromosome Z and the second linkage group was syntenic to chicken chromosomes 3 and 12. Based on the formation of a separate linkage group and synteny to chicken autosomes, we decided that the second linkage group, corresponding to a 12 Mb region of ostrich Z assembly, represented an assembly error.

In sex chromosome systems with a small PAR, due to an obligate cross-over in the heterogametic sex, recombination rate in the PAR is very high. Although the PAR in ostrich is much larger compared to other sex chromosomes (around 60% of the ZW chromosomes), there was a higher rate of recombination in females than in males, particularly before the PAR boundary. We next sought to compare the organization of the Z chromosome in ostrich with that of six other avian and three reptile species. We identified 18 synteny blocks and found 25 inversions in the avian lineage. Six inversions were identified on the long unrooted branch between Palaeognathae and Neognathae node and lizard. We found a high degree of conservation between lizard, turtle and alligator in the genomic regions that correspond to the avian Z chromosome, therefore, we inferred that the six inversions occurred in an early avian ancestor. Four of these inversions occurred in the currently non-recombining segment of ostrich sex chromosome. In the Neognathae lineage, subsequent to the split from Palaeognathae, a large inversion occurred in an early ancestor and is thus shared among all Neognathae species included in the study. This large inversion maps to 25.5-62.1 Mb of the chicken Z chromosome. The remaining inversions in the Neognathae occurred within the boundaries of this large inversion which further shuffled the gene order and orientation in different avian lineages.

The results of this study first highlight the importance of integrating genome assembly and linkage map information to make reliable inferences for chromosomal evolution. Moreover, the inversions identified in this study suggest that recombination might have been suppressed through a series of Z-linked and/or W-linked inversions in the avian lineage.

Paper III

Variation in the Z Chromosome to Autosomes Ratio of Genetic Diversity across Birds and its Relationship to the Fast-Z Effect

Sex chromosomes have specific mode of inheritance and ploidy that distinguish them from autosomes. These differences have consequences for the levels and characteristics of genetic diversity on sex chromosomes. Levels of genetic diversity is proportional to the effective population size (N_e) and mutation rate. Under the assumption of random variance in offspring number for males and females and equal mutation rate in sexes, the genetic diversity on Z compared to autosomes is expected to be 0.75. Differences in N_e and mutation rate between sexes can influence levels of genetic diversity and divergence on Z chromosomes.

In this study, in order to obtain a range of Z to autosome (Z:A) diversity across avian species, understand the underlying causes of the deviation of Z:A from the expected value of 0.75 and the relationship of Z:A diversity with the faster rate of evolution of Z-linked genes, we used the whole genome sequence of 32 avian species for which a male genome assembly was available. We calculated genetic diversity as heterozygosity per site across three functional categories, intergenic, intronic and 4-fold degenerate sites. There was a strong correlation between diversity on Z, autosome and Z:A diversity between pairs of functional categories. We also obtained lineage-specific estimates of synonymous (d_S) and nonsynonymous (d_N) divergence for autosomal and Z-linked genes. We observed that the mean of the distribution of Z:A diversity across avian species was not significantly different from 0.75. However, there was a variation among species in Z:A diversity that ranged between 0.278 and 1.27.

We next sought to identify determinants of Z:A levels of genetic diversity. Since most studied species were (at least socially) monogamous, variance in reproductive success did not explain the range of observed values of Z:A diversity. Levels of Z:A diversity showed a positive correlation with male to female mutation ratio estimated from ratio of synonymous substitutions for the Z chromosome to the synonymous substitutions for the autosome (d_{SZ}:d_{SA}). This was expected since Z chromosome spends two-thirds of its evolutionary time in males and one-third in females, hence, it is more influenced by a higher mutation rate in males.

To understand the impact of long-term N_e on Z:A levels of genetic diversity, we used body mass as a proxy for N_e . While autosomal diversity showed a negative correlation with body mass, no such correlation was observed for the Z chromosome. The Z:A genetic diversity showed a positive correlation with

body mass however the correlation was driven by the lack of correlation with Z diversity. We investigated the Z:A level of diversity in regions of the genome with different densities of sites under selection represented by the length of intergenic sequences. The Z:A diversity was lower (though not significantly different from 0.75) in regions with higher gene density than in regions with lower density.

We did not find a significant correlation between d_N:d_S of Z chromosome and autosome and levels of Z:A genetic diversity. This indicates that genetic drift might not be the only driver of faster evolution of Z-linked genes. Moreover, similar to genetic diversity, autosomal d_N:d_S had a significant correlation with body mass but Z chromosome d_N:d_S did not show a significant correlation with body mass. The results of this study suggest that levels of diversity and divergence on Z chromosome is affected by specific features that distinguish this chromosome from autosomes such as its lower level of sex-averaged recombination rate which can lead to an enhanced action of linked selection on Z chromosome.

Paper IV

Patterns of Nucleotide Diversity and Linkage Disequilibrium along the Ostrich Pseudoautosomal Region

The pseudoautosomal region (PAR) of sex chromosomes shares some features such as diploidy and recombination with autosomes. However, loci on PAR are affected to various extent by their linkage to the sex determining region (SDR). The linkage of PAR loci to the SDR may lead to their sex-linked inheritance pattern. Theoretical studies have shown that sex-linkage leads to an increase in the expected coalescence time for a pair of loci in the proximity of the SDR. The effect of sex-linkage on two different sex chromosomes resembles the situation of population structure with low rate of migration between two demes. Low rate of recombination between the sampled locus and the SDR can lead to the accumulation of genetic differences between chromosomes. Moreover, the sex-linkage of these loci provide a suitable condition for the action of sexually antagonistic mutations. In most sex chromosome systems, the PAR is very small and has a very high rate of recombination particularly in the heterogametic sex. In contrast, the long PAR in ostrich provides a system where we can investigate patterns of neutral genetic diversity and levels of differentiation between males and females on the PAR and in particular in the proximity of the SDR.

To this end, we used population genomics data of three subspecies of ostrich, the Southern African ostrich, *Struthio camelus australis,* the Eastern African ostrich, *Struthio camelus massiacus,* and *Struthio camelus domesticus,* which is the result of interbreeding ostriches from all over Africa during the start of the 19th century. For each subspecies, 10 individuals, 5 males and 5 females were sequenced. We measured levels of genetic diversity as heterozygosity per site, linkage disequilibrium and levels of differentiation between males and females across the Z chromosome.

We observed a heterogeneous pattern of genetic diversity across the Z chromosome shared between the three subspecies. Levels of genetic diversity were closest to autosomal values in three regions on the Z chromosome: at both ends of the chromosome and in the PAR region close to the PAR boundary. Heterozygosity was higher on the PAR (mean = 0.000811) than the average of the Z chromosome (mean = 0.000769). LD measured as r^2 was lower on the PAR (mean = 0.179) compared to the mean for Z chromosome (mean = 0.202). These two patterns were expected since the PAR undergoes recombination in both sexes. We used genomic features of Z chromosome, including the sex-averaged recombination rate, GC content and gene density to explain the variation observed in diversity across the chromosome. We found that genetic diversity in both the PAR and the whole Z chromosome was lower in regions with higher gene density. However, gene density was correlated with recombination rate and GC density. This could indicate the influence of linked selection as a determining factor of genetic diversity on Z chromosome. We did not find a significant difference in levels of genetic differentiation between males and females in PAR and non-PAR as measured by F_{ST}.

The nearly autosomal levels of heterozygosity close to the SDR and an absence of significant level of genetic differentiation between males and females could suggest that sexual antagonism is not a particularly strong force in ostriches. Moreover, the increased rate of recombination close to the PAR boundary in females might have implications for the maintenance of long PAR in ostriches and other ratites.

Conclusions and Future Prospects

Sex chromosome evolution and the processes involved in the evolution of degenerated sex chromosomes have been a major topic of research during the last century. In this thesis, I investigated the steps in the evolution of avian sex chromosomes by providing further information about the ancestral state of sex chromosomes in ostrich, a species of the basal clade of birds, Palaeognathae. By analyzing a wide number of species across avian phylogeny, I could show that previous estimates of severe reduction in Z:A genetic diversity in birds might have been due to sparse sampling of genomic loci and that several factors including the male-biased mutation rate and the prevalence of linked selection might shape levels and patterns of diversity on Z chromosome differently compared to autosomes. Using the long PAR in ostriches, I showed that levels of genetic diversity close to the SDR are not increased compared to the autosomal value and concluded that the higher rate of recombination in females in this region might have implication both for the observed diversity pattern in the PAR and for the maintenance of the long PAR in ostriches and other ratites.

There are various potential paths that can be taken for future research. I think the role of inversions together with the accumulation of sexually antagonistic (SA) alleles in avian sex chromosome evolution requires further theoretical and empirical investigation. The prevalent theory of recombination cessation in sex chromosome evolution is that the accumulation of SA mutations in the proximity of the SDR leads to selection for a modifier of recombination such as an inversion (Rice 1987a). The equilibrium frequency of the chromosome carrying the SA allele depends on the recombination frequency between the SDR and the SA allele and the selection coefficient of the SA allele. With an increasing recombination frequency between the SDR and the SA allele, a higher selective advantage is needed for the SA allele to increase in frequency. Inversions that cause recombination cessation by capturing the SDR and the SA allele are expected to increase in frequency in the population. However, the size of a given inversion is an important determinant of its fate in the population (Charlesworth and Barton 2018). Small inversions might not increase in frequency since they can only suppress a small amount of recombination. On the other hand, larger inversions can suppress a higher amount of recom-

bination but they might cause heterozygote disadvantage. It is therefore important to integrate the size and the fitness effect of the inversion itself in studies of the evolutionary dynamics of Z and W linked inversions in birds. Moreover, an important empirical work is to determine the distribution of fitness effect of SA mutations and the plausibility of their spread at a given genetic distance to the SDR.

Another exciting area of research is a joint study of the evolution of mechanisms of dosage compensation and suppression of recombination. For example, it could be speculated that if recombination would cease in a gene-by-gene basis, a more gradual or gene-specific mechanism of dosage compensation could be in place. To further understand the co-evolution of dosage compensation mechanisms along with recombination cessation and the degeneration of the non-recombining chromosome, a comparative study of dosage compensation across taxa with varying degree of sex chromosome degeneration would be particularly interesting.

The largely recombining state of ancient ratite sex chromosomes is another puzzling phenomenon and has several potentials for further research. Several hypotheses have been proposed to explain the non-degenerated state of ratite sex chromosomes including a lack of sexual antagonism. An important step would be to investigate the presence of SA alleles on the ratite PAR, particularly, in the proximity of the SDR since loci close to the SDR are expected to be targets for future SA mutations. It has been suggested that an increased rate of recombination in the proximity of the SDR might help to maintain a stable PAR during evolution. Future empirical and theoretical work in combining genetic diversity data with recombination, particularly in the PAR boundary, will be particularly interesting. Finally, cessation of recombination of sex chromosomes can be considered a two-step process with an initial recombination suppression in the proximity of the SDR and the further extension of the region with suppressed recombination (Charlesworth, et al. 2005). I very much look forward to further empirical research to better understand the mechanisms involved in the expansion of the region with suppressed recombination to provide a more comprehensive view of the underlying processes in the evolution of dimorphic sex chromosomes.

Svenks Sammanfattning

Könskromosomerna är den del av genomet som bär de könsbestämmande generna och som har specifika särdrag som skiljer dem från autosomerna. Könskromosomerna har ett könsspecifikt nedärvnings- och rekombinationsmönster, och visar ofta på en viss grad av heteromorfism mellan de homologa paren. Evolutionen av könskromosomer kännetecknas av undertryckt rekombination mellan proto-könskromosomer, vilket leder till en gradvis minskning av genetiskt material från den kromosom som inte genomgår rekombination. I däggdjur är hanar det heterogametiska könet, där könsbestämning sker via ett XX/XY-system. I fåglar är däremot honor det heterogametiska könet (ZW) och hanar det homogametiska (ZZ). Könskromosomerna som är specifika för det heterogametiska könet (W och Y i fåglar respektive däggdjur) är vanligtvis degenererade och rekombination med Z respektive X är begränsad till en mindre del av respektive kromosom. Processerna som påverkar evolutionen av degenererade könskromosomer samt den evolutionära dynamiken hos loci kopplade till könskromosomer har under det senaste århundradet åtnjutit stort forskningsintresse inom evolutionsbiologin.

I den här doktorsavhandlingen undersökte jag det honligt heterogametiska systemet som återfinns hos fåglar för att generera förståelse på en evolutionär nivå kring hur rekombination upphör mellan könskromosomer. Dessutom undersökte jag vilka evolutionära krafter som påverkar nivån av genetisk diversitet i könsspecifika DNA-sekvenser. Fåglar är uppdelade i två fylogenetiska huvudgrupper, Paleognathae (ratiter och tinamofåglar) och Neognathae (alla andra fåglar, mer är 99% av nu levande arter). I kontrast mot Neognathae genomgår Z- och W-kromosomerna hos ratiterna rekombination över mer än hälften av kromosomlängden. Att studera evolutionen av könskromosomer i Palaeognathae är därmed intressant om vi vill förstå de olika steg som är involverade i evolutionen av fåglars könskromosomer. Detta eftersom Paleognathae är den ursprungliga fylogenetiska gruppen av fåglar och kan ge insikt i hur de ursprungliga könskromosomerna såg ut. Dessutom är den utbredda rekombinationen mellan könskromosomer i Paleognathae i sig självt ett fascinerande fenomen som behöver undersökas vidare.

De två första kapitlen i min doktorsavhandling handlar om evolutionen av könskromosomer i fåglar över långa evolutionära tidsavstånd, med ett fokus

på strutsens könskromosomer. I den första studien identifierade jag icke-rekombinerande homologa gener på Z- och W-kromosomerna i struts, vilka ännu inte har degenererats på W-kromosomen (s.k. gametologa gener). Från graden av sekvensdivergens i de här generna kunde jag se att upphörandet av rekombination skedde innan fåglarna separerades Paleognathae och Neognathae. Från positionen och sekvensvariationen hos gametologa gener kan vi få förståelse kring hur upphörandet av rekombination gick igenom olika steg. Detta beror på att gener som upphörde att genomgå rekombination längre tillbaks i tiden förväntas ha högre sekvensdivergens än gener som upphörde att rekombinera senare. Dock kunde inte positionen av de gametologa generna längs med kromosomen bestämmas eftersom strutsgenomet saknar information på kromosomnivå. I den andra studien konstruerade jag en kopplingskarta över strutsens Z-kromosom och använde denna för att korrigera Z-kromosomens sekvenssammansättning. I ett nästa steg identifierade jag 25 inversioner på Z-kromosomen genom att jämföra ordningen och positionen av Z-kromosomkopplade gener i struts med sex arter av Neognathae och en reptilart. Upptäckten av Z-kopplade inversioner ledde till slutsatsen att evolutionen kring upphörandet av rekombination mellan könskromosomer i fåglar kan ha skett genom en serie av inversioner av Z- och/eller W-kromosomen, eller andra processer som lett till ett mer gradvis upphörande av rekombination.

När könskromosomerna väl utvecklats till den nuvarande formen av heteromorphism som ses hos en majoritet av taxa, kommer olika evolutionära processer att ha en annan påverkan på dem jämfört med autosomerna. För varje hona och hane går det tre Z-kromosomer på fyra autosomer, det medför att den genetiska variationen hos Z-kromosomen kan förväntas uppgå till 0.75 av den autosomala variationen, förutsatt att mutationshastigheten och den effektiva populationsstorleken är densamma för båda könen. Faktorer som kan påverka den effektiva populationsstorleken, som till exempel polygama parningssystem eller könsbias i mutationshastighet, kan medföra att Z:A-variationen avviker från det förväntade värdet på 0.75. I avhandlingens tredje kapitel undersöker jag de evolutionära processer som kan ge avvikelser från den förväntade Z:A-variationen. Jag använde helgenomdata från hanar från 32 fågelarter för att analysera distributionen av Z:A-variationen hos fåglar. Vi fann att medelvärdet för Z:A-variationen generellt inte var signifikant skild från 0.75, men däremot var det stor spridning i Z:A-variationen mellan olika arter (0.278-1.27). De flesta undersökta arter är (åtminstone socialt) monogama, så spridningen i Z:A kan inte förklaras av variansen i reproduktiv framgång. Å andra sidan visade Z:A-variationen ett positivt samband med hanlig mutationsbias och en negativ korrelation med effektiv populationsstorlek, vilket kan vara en indikation på förekomst av kopplad selektion på Z-kromosomen.

Hos de flesta könskromosomsystem kvarstår rekombination i en liten region hos det heterogameta könet för att säkerställa korrekt kromosomparning och segregering under meios, regionen benämns pseudoautosomal region (PAR). Gener lokaliserade på PAR har vissa egenskaper gemensamt med gener på autosomer, som också har två kopior var hos honor och hanar, men genregioner som ligger nära den könsbestämmande regionen ("sex-determining region", SDR) uppvisar egenskaper med mer könsbundna särdrag. I fjärde kapitlet använde jag populationsgenetiska data från Z-kromosomen hos struts för att undersöka den evolutionära dynamiken hos PAR med särskilt fokus på dess avstånd till SDR. I motsats till teoretiska förutsägelser var den genetiska variationen nära SDR på Z-kromosomen inte högre än nivån hos autosomer. Lokus nära SDR uppvisade inte någon kopplingsojämnvikt med SDR och ingen signatur från sexuellt antagonistisk selektion kunde detekteras vid analys av könsskillnader i allelfrekvenser. Med en statistisk modell kunde vi visa att gentäthet verkar vara en god prediktor av diversitetsmönster längs Z-kromosomen. Resultatet kan bero på en lindrigt förhöjd rekombinationsfrekvens nära PAR hos honor. Detta är särskilt intressant med tanke på att det har föreslagits att PAR med en högre rekombinationsfrekvens nära SDR kan förhindra ansamling av sexuellt antagonistiska loci och därmed behålla en evolutionärt stabil PAR.

Sammanfattningsvis har jag undersökt de olika stegen i evolutionen av fåglars könskromosomer genom att ta fram ytterligare information om könskromosomernas ursprungliga tillstånd hos struts. Genom att analysera ett stort antal fåglar har jag visat att tidigare fynd av kraftig nedgång i Z:A-variationen hos fågel kan ha orsakats av för litet analysurval av genomiska loci och att Z:A-variationen verkar minska vid ökad effektiv populationsstorlek. Slutligen visade jag att variationen nära SDR längs PAR hos struts ligger nära värdet hos autosomer, en intressant observation med betydelse för hur ratiters långa PAR kan ha bevarats. Samtliga kapitel i min avhandling kan ligga till grund för flera potentiella fortsatta studier. Vidare undersökningar av inversioner på W-kromosomen hos struts kan ge ytterligare belägg för hur inversioner påverkar evolutionen av könskromosomer hos fåglar och för att undersöka förekomsten av kopplad selektion på könskromosomer behövs mer ingående analys med hjälp av populationsgenetiska data. Slutligen, den huvudsakliga frågan varför ratiter har behållit rekombination på könskromosomerna trots deras urgamla ursprung är fortfarande en gåta och fordrar vidare teoretiska och empiriska studier.

Translated to Swedish by Erik Gudmunds and Karin Näsvall

Acknowledgements

I would like to thank my supervisor, Hans Ellegren, for welcoming me to his lab and introducing me to the exciting area of sex chromosome evolution. I am very grateful that you let me think and discover. I strongly believe that a great supervisor is not the one who takes the student's hand and accompany them step by step, but it's the one who shows the path and lets the student discover on their own while being there as a support. I would like to thank my co-supervisor, Takeshi Kawakami. You have been a major support during my PhD, talking to you was always amazingly helpful. Thank you so much for believing in me and helping me develop my ideas. Above all, thanks a lot for all your comments on my work and for pushing me to work harder to finish my PhD during the last months. Carina Mugal, I would like to especially thank you for your patience and guidance of my project on levels of genetic diversity on sex chromosomes and autosomes. You came in at the time I was totally lost with this project and was so much in need of help. You are one of the brightest researchers I know. I admire so much your passion, the depth of your knowledge and the precision you have in your research.

I have often bothered almost all of the professors in the program with my questions. Alex Suh, thank you so much for all the discussion and support and explaining me for a hundred times about phylogeny, hopefully, one day, I will understand it properly and thank you for taking the time to comment on my Kappa despite your busy schedule. Niclas Backström, thank you so much for letting me in your office at any time and answering my questions with patience. Simon Whelan, thanks a lot for lending me your books and explaining about probability theory to me. Douglas Scofield, thank you very much for the time you spent on explaining me every single step of any question I had. It is always so exciting for me to talk to you because you always go beyond the question asked and explain concepts in their broader picture. Jochen, Arild, Elina, Robert and Anna, it has been and it is such a great pleasure to work with you in the program and learn from you.

Outside of our evolutionary biology program, I would like to specially thank Sylvain Glémin and Martin Lascoux. Sylvain, you really are a role model for young researchers. Your breath and depth of understanding is amazing. I hope I will be able to explain population genetics concepts as well as you do one

day. Martin, I cannot thank you enough for answering so many of my questions on coalescent theory and your support during my PhD. I would like to especially thank Charlie Cornwallis for giving me the opportunity to work with population genomics data of ostrich subspecies and for the amazing support to analyze and write the last chapter of my thesis. Tim Connallon and Colin Olito, it was a great pleasure working with you and learning from your vast amount of knowledge and experience, I hope I can develop my theoretical skills further and one day, work in another project with you. I would also like to thank my first supervisor during my master studies in Montpellier, Patrick Lemaire and his team for such an amazing internship in their lab. Two very special persons in my scientific education are the supervisors of my master's thesis in Montpellier, Renaud Vitalis and Mathieu Gautier. I decided to apply for a project with you because I understood absolutely nothing of your lecture on genomic signatures of positive selection. I never forget the first day when you asked me about my programming skills and if I knew how to code in AWK and I thought you are talking about the oak tree. I want to thank you extremely much for giving me the opportunity to work with you and learn from you and being patient enough to guide me through every single question. I am in this field because you decided to give me a chance.

I would like to thank two special persons, Annette Berger and Frida Svedberg. Annette, thanks a lot for providing me with an endless number of certificates for submission to the migration office and always being friendly and helpful. Frida, thank you so much for keeping track of our funding and money and maintaining such a discipline and efficiency and welcoming me in Swedish lunches. I would also like to thank the department board of EBC. Being part of the board was very educational to me. I learnt a lot about financing a research department, words such as overhead cost, "överskott" and "underskott" became, little by little, part of my vocabulary. I think it is fantastic that PhD students of every nationality have the right to be part of the board. It teaches greatly how the economics of a research department works and the effort that is required to keep our work environment and our research quality in high standard and our economic balance safe.

The PhD journey is exciting and difficult and on the way, there are friends and colleagues who lift you up when you are down or in doubt. I had the luck to start my PhD in an amazing office, with the three most wonderful office mates back then and friends now. For cheering me up and encouraging me all along the way and for the fantastic discussions, I would like to thank my dear friend, Matthias Weissenteiner. Willian da Silva, I cannot only express how much you helped me during the hard times but how much I enjoyed discussing with you and sharing our excitement for learning new concepts together. I am so lucky to have you as one of my best friends and colleagues. Roy Francis, thank

you very much for listening to our complaints and doubts and for helping us with your amazing graphic skills. It was such a pleasure to have you in the office. And thank you so much for designing such a beautiful cover for my thesis, I just love it!

Paulina, Ludo and Mi, we grew up together in this PhD, I am so happy to have had you as colleagues and friends. I learnt a lot from each of you, about molecular evolution, about Python and about statistics. This PhD would not have been possible without your support and friendship along the way. Linnéa Smeds, I feel so lucky to have become your friend, to have got to know you. I learnt so many things from you, for most, keeping a healthy balance between work and life, being efficient and organized and extremely honest in work. Agnes Rettelbach, thank you so much for believing in my idea and spending time in developing the theory. I plan to work on our project further to develop our work, that was the very first time someone took me serious and I am not letting it go that easy. Madeline Chase, you were a breath of fresh air at the most stressful time of my PhD. It has been amazing working with you, I learnt a lot from you and I hope that we continue working on our balancing selection project to get a great publication out of it.

Shadi and Ghazal, it is amazing to have you, to talk our mother tongue from time to time and to laugh a lot. I cannot express how happy I am to have both you at EBC. Sergio, Foteini, Venkat, Merce, Philipp and Valentina, one needs a great support system during the PhD and you all have been that. It has been so much fun to share the PhD experience with you, the happy and more challenging moments. Erik and Karin, thank you so much for the time and effort you spent in translating my thesis summary to Swedish. I would also like to thank all the past and present members of Ellegren lab, Reto, Marty, Krysha, Jente, Agnese for all the discussion and great comments. I would like to thank all the amazing people I have had the privilege to work with at the same department, the pressure of the last months of thesis writing has crippled my memory for names but I have not forgotten your kindness and support.

Life is not all within university. Elham and Shirin, you have been wonderful support during my PhD. Elham, your home was my safe haven to run to whenever I felt sad and the place to share the happy moments. Shirin, your wisdom and positive mind helped me so much. I cannot thank you enough for your support. Eline and Jasper, how wonderful it has been to have both of you during these years. Eline, thank you so much for being such a great friend, for all the beautiful postcards and for all your support of me and Thierry. Ellen, Johan and Michiel, I have not forgotten the first day I arrived in my student room in Groningen, seeing how you had prepared it with all the beautiful gifts and sweets, bringing a bike for me all the way from Belgium. I am so lucky to

have you in my life. My lovely Afsaneh, thank you so much for having been such a sweet wonderful friend to me since we were so little. Delaram, I am so happy to have had you as my friend and have shared so many memories with you. My dear Kosar, I cannot thank you enough for taking care of my mother whenever she was in need of help. Parivash, thanks a lot for visiting me in Leipzig and for all your kindness. Mansoureh and the Heravi family, you have been the sunshine in our life since we can remember. My dear Soudabeh, Salman, Mandis and Mojtaba, I am so lucky to have all of you in my life. Mandis, thank you so much for the lovely weekend and delicious Gheymeh in London, I enjoyed it so much.

There is one who's always been my rock, and that's you, Thierry. How I burdened you all the time with the pressure of my work? How many times I thought I could not make it and you made me feel strong again? You stood all these times by my side, you supported me, helped me, cheered me up and gave me the strength to go on while you had a pain of your own, tolerating it in silence just so I could continue working towards my dreams. Thierry, like you always say, we are a team, we really are and look, we finally did it! I would like to give a special thought to our two cats, Dropytje and Irmaatje! You two cute creatures accompanied me many times till late at night so I could work. I am so happy to have the three of you in my life. I would also like to thank Eric, Caroline and Febe for their great help in Marke. Annick and Matthias, I am so grateful for all your help and support during these years, it would not be possible without you.

I have a last word of acknowledgment for the PhD itself. May it be helpful to other students when facing difficulty or despair. I found PhD more than learning about science. You have to read a lot, a lot of books and papers and you must improve your skills in science. But also, you must develop a character. A strong one and at times, a stubborn one. You must know what you want to do and why you do what you do, to take initiatives and responsibility for the right and wrong of your work. To not seek confidence in the compliment of others but build confidence by your knowledge and skills. To criticize yourself and be honest. To dare ask questions, as many as you want and to not stop until you really are convinced. You are never too ambitious, believe in yourself and have confidence but work really hard and dream a lot, you'll make it!

اشرف خانوم، به خاطر مهربونیتون در لحظاتی که همه من و مامان رو در زندگی ترک کرده بودن، ازت ممنونم. لحظات غم و تنهایی با تو همیشه شاد میشن. غم برزگیو تجربه کردیم ولی اگر بهشتی باشی، لایق‌ترین ساکن براش اکرم ماست.

اکرم نازم، هیچ چیز رو جز صدای خنده هات نمیتونم به یاد بیارم. هنوز، هر روز، در خاطراتم زنده هستی. چیزی که همیشه بهم میگفتی رو نوشتم همیشه بهش نگاه می کنم. میگفتی مشکلات اگر میمش رو برداری میشه شکلات. برای همیشه دوستت دارم.

خاله نسرین و حامد آقا، از تمام مهربونیاتون ممنونم، خیلی دوستون دارم.

لیلا، عمران و کیان نازک، از همتون به خاطره محبتتون ممنونم. خیلی خوشحالم که شماها رو تو زندگی دارم.

مامان و بابا، از شما به خاطر تمام زحمتاتون تشکر می‌کنم. بابا، از اینکه به من کمک کردی که بتونم بیام اروپا و درس بخونم، ازت خیلی ممنونم. مامان، ازت ممنونم به خاطر تمام وقتایی که من و دوستام رو پذیرایی کردی و تمام شبایی که با صدای صامت تلویزیون نگاه کردی که من درس بخونم. تو، تو دنیا یکی هست. شما به من بال پرواز دادین. از هر دو تون به خاطره فکر بازتون و تربیت خوبتون خیلی ممنونم.

References

Adolfsson S, Ellegren H. 2013. Lack of Dosage Compensation Accompanies the Arrested Stage of Sex Chromosome Evolution in Ostriches. *Molecular Biology and Evolution* 30:806-810.

Avery PJ. 1984. The Population Genetics of Haplo-Diploids and X-Linked Genes. *Genetics Research* 44:321-341.

Bellott DW, Skaletsky H, Cho TJ, Brown L, Locke D, Chen N, Galkina S, Pyntikova T, Koutseva N, Graves T, et al. 2017. Avian W and mammalian Y chromosomes convergently retained dosage-sensitive regulators. *Nature Genetics* 49:387-394.

Bellott DW, Skaletsky H, Pyntikova T, Mardis ER, Graves T, Kremitzki C, Brown LG, Rozen S, Warren WC, Wilson RK, et al. 2010. Convergent evolution of chicken Z and human X chromosomes by expansion and gene acquisition. *Nature* 466:612-616.

Bergero R, Forrest A, Kamau E, Charlesworth D. 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, Qiu S, Forrest A, Borthwick H, Charlesworth D. 2013. Expansion of the Pseudo-autosomal Region and Ongoing Recombination Suppression in the *Silene latifolia* Sex Chromosomes. *Genetics* 194:673-686.

Beukeboom L, Perrin N. 2014. The evolution of sex determination: *Oxford University Press*.

Bewick AJ, Chain FJ, Zimmerman LB, Sesay A, Gilchrist MJ, Owens ND, Seifertova E, Krylov V, Macha J, Tlapakova T, et al. 2013. A large pseudoautosomal region on the sex chromosomes of the frog *Silurana tropicalis*. *Genome Biology and Evolution* 5:1087-1098.

Bherer C, Campbell CL, Auton A. 2017. Refined genetic maps reveal sexual dimorphism in human meiotic recombination at multiple scales. *Nature Communications* 8:14994.

Bourque G, Pevzner PA. 2002. Genome-scale evolution: Reconstructing gene orders in the ancestral species. *Genome Research* 12:26-36.

Braga MDV. 2009. baobabLUNA: the solution space of sorting by reversals. *Bioinformatics* 25:1833-1835.

Branco S, Badouin H, Rodriguez de la Vega RC, Gouzy J, Carpentier F, Aguileta G, Siguenza S, Brandenburg JT, Coelho MA, Hood ME, et al. 2017. Evolutionary strata on young mating-type chromosomes despite the

lack of sexual antagonism. *Proceedings of the National Academy of Sciences USA* 114:7067-7072.
Bridges C. 1916. Non-disjunction as proof of the chromosome theory of heredity (concluded). *Genetics* 1:107-163.
Bridges C. 1925. Sex in relation to chromosomes and genes. *American Naturalist* 59:127-137.
Caballero A. 1995. On the effective size of populations with separate sexes, with particular reference to sex-linked genes. *Genetics* 139:1007-1011.
Capel B. 2017. Vertebrate sex determination: evolutionary plasticity of a fundamental switch. *Nature Reviews Genetics* 18:675-689.
Charlesworth B. 1994. The effect of background selection against deleterious mutations on weakly selected, linked variants. *Genetics Research* 63:213-227.
Charlesworth B, Barton NH. 2018. The Spread of an Inversion with Migration and Selection. *Genetics* 208:377-382.
Charlesworth B, Campos JL, Jackson BC. 2018. Faster-X evolution: Theory and evidence from Drosophila. *Molecular Ecology* 27:3753-3771.
Charlesworth B, Charlesworth D. 2000. The degeneration of Y chromosomes. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 355:1563-1572.
Charlesworth B, Coyne JA, Barton NH. 1987. The Relative Rates of Evolution of Sex-Chromosomes and Autosomes. *American Naturalist* 130:113-146.
Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134:1289-1303.
Charlesworth D, Charlesworth B. 1980. Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genetics Research* 35:205-214.
Charlesworth D, Charlesworth B, Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* 95:118-128.
Cheng C, Kirkpatrick M. 2016. Sex-Specific Selection and Sex-Biased Gene Expression in Humans and Flies. *PloS Genetics* 12:e1006170.
Chibalina MV, Filatov DA. 2011. Plant Y Chromosome Degeneration Is Retarded by Haploid Purifying Selection. *Current Biology* 21:1475-1479.
Chue J, Smith CA. 2011. Sex determination and sexual differentiation in the avian model. *Febs Journal* 278:1027-1034.
Clark AG. 1988. The evolution of the Y chromosome with X-Y recombination. *Genetics* 119:711-720.
Ellegren H. 2007. Characteristics, causes and evolutionary consequences of male-biased mutation. *Proceedings of the Royal Society B* 274:1-10.
Ellegren H. 2011. Emergence of male-biased genes on the chicken Z-chromosome: sex-chromosome contrasts between male and female heterogametic systems. *Genome Research* 21:2082-2086.
Ellegren H, Hultin-Rosenberg L, Brunstrom B, Dencker L, Kultima K, Scholz B. 2007. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. *BMC Biology* 5:40.

Ercan S. 2015. Mechanisms of X chromosome dosage compensation. *J Genomics* 3:1-19.
Felsenstein J. 1974. The evolutionary advantage of recombination. *Genetics* 78:737-756.
Fridolfsson AK, Cheng H, Copeland NG, Jenkins NA, Liu HC, Raudsepp T, Woodage T, Chowdhary B, Halverson J, Ellegren H. 1998. Evolution of the avian sex chromosomes from an ancestral pair of autosomes. *Proceedings of the National Academy of Sciences USA* 95:8147-8152.
Futuyma DJ, Kirkpatrick M. 2018. Evolution. *New York: Sinauer Associates*.
Gao Z, Wyman MJ, Sella G, Przeworski M. 2016. Interpreting the Dependence of Mutation Rates on Age and Time. *PLoS Biology* 14:e1002355.
Gillespie JH. 2004. Population genetics : a concise guide. Baltimore ; London: *Johns Hopkins University Press*.
Graves JAM. 2016. Evolution of vertebrate sex chromosomes and dosage compensation. *Nature Reviews Genetics* 17:33-46.
Griffiths A, Miller, JH., Suzuki, DT., et al. 2000. An Introduction to Genetic Analysis. 7th edition: New York: *W. H. Freeman*.
Groenen MA, Wahlberg P, Foglio M, Cheng HH, Megens HJ, Crooijmans RP, Besnier F, Lathrop M, Muir WM, Wong GK, et al. 2009. A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Research* 19:510-519.
Haddrath O, Baker AJ. 2012. Multiple nuclear genes and retroposons support vicariance and dispersal of the palaeognaths, and an Early Cretaceous origin of modern birds. *Proceedings of the Royal Society B* 279:4617-4625.
Handley LL, Ceplitis H, Ellegren H. 2004. Evolutionary strata on the chicken Z chromosome: Implications for sex chromosome evolution. *Genetics* 167:367-376.
Hartl D, Jones, E. 2005. Genetics: Analysis of Genes & Genomes: *Jones and Bartlett Publishers*.
Hedrick PW. 2007. Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. *Evolution* 61:2750-2771.
Henking H. 1891. Untersuchungen uber die ersten Entwicklungsvorigoenge in den Eiern der Insektern II Uber spermatogenese und Beziehung zur Eientwicklung bei *Pyrrhocoris apterus*. Z Wiss Zool 51:51.
Hill WG, Robertson A. 1966. Effect of Linkage on Limits to Artificial Selection. *Genetics Research* 89: 311-336.
Hinch AG, Altemose N, Noor N, Donnelly P, Myers SR. 2014. Recombination in the human Pseudoautosomal region PAR1. *PloS Genetics* 10:e1004503.
Julien P, Brawand D, Soumillon M, Necsulea A, Liechti A, Schutz F, Daish T, Grutzner F, Kaessmann H. 2012. Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. *PLoS Biology* 10:e1001328.

Kamiya T, Kai W, Tasumi S, Oka A, Matsunaga T, Mizuno N, Fujita M, Suetake H, Suzuki S, Hosoya S, et al. 2012. A Trans-Species Missense SNP in *Amhr2* Is Associated with Sex Determination in the Tiger Pufferfish, *Takifugu rubripes* (Fugu). *PloS Genetics* 8:e1002798.

Kawakami T, Smeds L, Backstrom N, Husby A, Qvarnstrom A, Mugal CF, Olason P, Ellegren H. 2014. A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. *Molecular Ecology* 23:4035-4058.

Kimura M. 1971. Theoretical foundation of population genetics at the molecular level. *Theoretical Population Biology* 2:174–208.

Kirkpatrick M. 2010. How and why chromosome inversions evolve. *PLoS Biology* 8:e1000501.

Kirkpatrick M, Guerrero RF. 2014. Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics* 197:531-541.

Kirkpatrick M, Guerrero RF, Scarpino SV. 2010. Patterns of neutral genetic variation on recombining sex chromosomes. *Genetics* 184:1141-1152.

Kosambi D. 1944. The estimation of the map distance from recombination values. *Annals of Eugenics* 12:172-175.

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

Lemaitre C, Braga MDV, Gautier C, Sagot MF, Tannier E, Marais GAB. 2009. Footprints of Inversions at Present and Past Pseudoautosomal Boundaries in Human Sex Chromosomes. *Genome Biology and Evolution* 1:56-66.

Matson CK, Murphy MW, Sarver AL, Griswold MD, Bardwell VJ, Zarkower D. 2011. DMRT1 prevents female reprogramming in the postnatal mammalian testis. *Nature* 476:101-104.

Matson CK, Zarkower D. 2012. Sex and the singular DM domain: insights into sexual regulation, evolution and plasticity. *Nature Reviews Genetics* 13:163-174.

Maynard-Smith J, Haigh J. 1974. Hitch-Hiking Effect of a Favorable Gene. *Genetics Research* 23:23-35.

Miyata T, Hayashida H, Kuma K, Mitsuyasu K, Yasunaga T. 1987. Male-driven molecular evolution: a model and nucleotide sequence analysis. *Cold Spring Harbor Symposia on Quantitative Biology* 52:863-867.

Muller HJ. 1914. A Factor for the Fourth Chromosome of Drosophila. *Science* 39:906.

Muller HJ. 1918. Genetic Variability, Twin Hybrids and Constant Hybrids, in a Case of Balanced Lethal Factors. *Genetics* 3:422-499.

Muller HJ. 1964. The Relation of Recombination to Mutational Advance. *Mutation Research* 106:2-9.

Nadeau JH, Taylor BA. 1984. Lengths of Chromosomal Segments Conserved since Divergence of Man and Mouse. *Proceedings of the National Academy of Sciences USA* 81:814-818.

Nam K, Ellegren H. 2008. The Chicken (*Gallus gallus*) Z Chromosome Contains at Least Three Nonlinear Evolutionary Strata. *Genetics* 180:1131-1136.

Natri HM, Shikano T, Merila J. 2013. Progressive Recombination Suppression and Differentiation in Recently Evolved Neo-sex Chromosomes. *Molecular Biology and Evolution* 30:1131-1144.

Nei M. 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences USA* 70:3321-3323.

Nei M. 1969. Linkage modifications and sex difference in recombination. *Genetics* 63:681-699.

Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Science USA* 76:5269-5273.

Nicolas M, Marais G, Hykelova V, Janousek B, Laporte V, Vyskot B, Mouchiroud D, Negrutiu I, Charlesworth D, Moneger F. 2005. A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants. *PLoS Biology* 3:e4.

Ohno S. 1967. Sex chromosomes and sex-linked genes. Berlin, New York etc.: *Springer-Verlag*.

Otto SP, Pannell JR, Peichel CL, Ashman TL, Charlesworth D, Chippindale AK, Delph LF, Guerrero RF, Scarpino SV, McAllister BF. 2011. About PAR: The distinct evolutionary dynamics of the pseudoautosomal region. *Trends in Genetics* 27:358-367.

Perry J, Palmer S, Gabriel A, Ashworth A. 2001. A short pseudoautosomal region in laboratory mice. *Genome Research* 11:1826-1832.

Petkov PM, Broman KW, Szatkiewicz JP, Paigen K. 2007. Crossover interference underlies sex differences in recombination rates. *Trends in Genetics* 23:539-542.

Pool JE, Nielsen R. 2007. Population size changes reshape genomic patterns of diversity. *Evolution* 61:3001-3006.

Rastas P, Calboli FC, Guo B, Shikano T, Merila J. 2015. Construction of Ultradense Linkage Maps with Lep-MAP2: Stickleback F2 Recombinant Crosses as an Example. *Genome Biology and Evolution* 8:78-93.

Rice WR. 1987a. The Accumulation of Sexually Antagonistic Genes as a Selective Agent Promoting the Evolution of Reduced Recombination between Primitive Sex Chromosomes. *Evolution* 41:911-914.

Rice WR. 1987b. Genetic Hitchhiking and the Evolution of Reduced Genetic-Activity of the Y Sex Chromosome. *Genetics* 116:161-167.

Rice WR. 1984. Sex Chromosomes and the Evolution of Sexual Dimorphism. *Evolution* 38:735-742.

Roesti M, Moser D, Berner D. 2013. Recombination in the threespine stickleback genome--patterns and consequences. *Molecular Ecology* 22:3014-3027.

Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Platzer M, Howell GR, Burrows C, Bird CP, et al. 2005. The DNA sequence of the human X chromosome. *Nature* 434:325-337.

Sandstedt SA, Tucker PK. 2004. Evolutionary strata on the mouse X chromosome correspond to strata on the human X chromosome. *Genome Research* 14:267-272.

Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, Smith MJ, Foster JW, Frischauf AM, Lovell-Badge R, Goodfellow PN. 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 346:240-244.

Slatkin M. 2008. Linkage disequilibrium--understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics* 9:477-485.

Smeds L, Kawakami T, Burri R, Bolivar P, Husby A, Qvarnstrom A, Uebbing S, Ellegren H. 2014. Genomic identification and characterization of the pseudoautosomal region in highly differentiated avian sex chromosomes. *Nature Communications* 5:5448.

Smeds L, Warmuth V, Bolivar P, Uebbing S, Burri R, Suh A, Nater A, Bures S, Garamszegi LZ, Hogner S, et al. 2015. Evolutionary analysis of the female-specific avian W chromosome. *Nature Communications* 6:7330.

Smith CA, Roeszler KN, Ohnesorg T, Cummins DM, Farlie PG, Doran TJ, Sinclair AH. 2009. The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature* 461:267-271.

Tang YC, Amon A. 2013. Gene Copy-Number Alterations: A Cost-Benefit Analysis. *Cell* 152:394-405.

Tesler G. 2002. GRIMM: genome rearrangements web server. *Bioinformatics* 18:492-493.

Torres EM, Williams BR, Amon A. 2008. Aneuploidy: Cells losing their balance. *Genetics* 179:737-746.

Uebbing S, Kunstner A, Makinen H, Ellegren H. 2013. Transcriptome sequencing reveals the character of incomplete dosage compensation across multiple tissues in flycatchers. *Genome Biology and Evolution* 5:1555-1566.

Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D. 2013. Comparative Sex Chromosome Genomics in Snakes: Differentiation, Evolutionary Strata, and Lack of Global Dosage Compensation. *PloS Biology* 11:e1001643.

Vicoso B, Kaiser VB, Bachtrog D. 2013. 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.

Wang JP, Na JK, Yu QY, Gschwend AR, Han J, Zeng FC, Aryal R, VanBuren R, Murray JE, Zhang WL, et al. 2012. Sequencing papaya X and Yh chromosomes reveals molecular basis of incipient sex chromosome evolution. *Proceedings of the National Academy of Sciences USA* 109:13710-13715.

White MA, Kitano J, Peichel CL. 2015. Purifying Selection Maintains Dosage-Sensitive Genes during Degeneration of the Threespine Stickleback Y Chromosome. *Molecular Biology and Evolution* 32:1981-1995.

Wilson Sayres MA, Makova KD. 2011. Genome analyses substantiate male mutation bias in many species. *Bioessays* 33:938-945.

Wright AE, Harrison PW, Montgomery SH, Pointer MA, Mank JE. 2014. Independent Stratum Formation on the Avian Sex Chromosomes Reveals Inter-Chromosomal Gene Conversion and Predominance of Purifying Selection on the W Chromosome. *Evolution* 68:3281-3295.

Yazdi HP, Ellegren H. 2018. A Genetic Map of Ostrich Z Chromosome and the Role of Inversions in Avian Sex Chromosome Evolution. *Genome Biology and Evolution* 10:2049-2060.

Yazdi HP, Ellegren H. 2014. Old but not (so) degenerated--slow evolution of largely homomorphic sex chromosomes in ratites. *Molecular Biology and Evolution* 31:1444-1453.

Zhou Q, Zhang JL, Bachtrog D, An N, Huang QF, Jarvis ED, Gilbert MTP, Zhang GJ. 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* 346:1246338.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1788

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-379665

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2019