Evolution of sex chromosomes in Sylvioidea songbirds

Sigeman, Hanna

2021

Document Version:
Publisher's PDF, also known as Version of record

Evolution of sex chromosomes in Sylvioidea songbirds

HANNA SIGEMAN
DEPARTMENT OF BIOLOGY | FACULTY OF SCIENCE | LUND UNIVERSITY

Evolution of sex chromosomes in Sylvioidea songbirds

Hanna Sigeman

Evolution of sex chromosomes in Sylvioidea songbirds

Hanna Sigeman

LUND UNIVERSITY

DOCTORAL DISSERTATION

by due permission of the Faculty of Science, Lund University, Sweden.
To be defended in the Blue Hall, Ecology Building, Sölvegatan 37, Lund, Sweden
on the 9th of April 2021 at 9:00.

Faculty opponent
Prof. Paul Sunnucks

Monash University

Organization LUND UNIVERSITY Department of Biology	**Document name** DOCTORAL DISSERTATION
	Date of issue 9th of April 2021
Author: Hanna Sigeman	Sponsoring organization

Title and subtitle: Evolution of sex chromosomes in Sylvioidea songbirds

Abstract

Sex chromosomes were discovered more than 100 years ago. They have been studied intensely from a theoretical perspective since then, giving rise to a large body of testable predictions about their evolution from autosomes. A common feature of sex chromosomes is recombination suppression between the sex chromosome copies (X and Y in male heterogametic systems, or Z and W in female heterogametic systems). Without recombination, the sex-limited chromosome (Y or W) is expected to degenerate through the accumulation of deleterious mutations and repeat elements. Over long evolutionary time scales, this degeneration may leave the Y/W chromosomes short and almost completely devoid of functional genes.

Using genome sequencing technology, we can now study the full DNA sequence of sex chromosomes. The analysis of sequence data from a wide range of taxonomic groups has revealed that sex chromosomes are more dynamic and variable than previously believed. Several aspects of sex chromosome evolution, however, remain poorly understood, especially relating to the early stages of evolution from autosomes. This is partly because some hypotheses are challenging to test, but also because most well-studied sex chromosome systems are evolutionarily old and biased towards XY systems.

In this thesis, I study the evolution of sex chromosomes across Sylvioidea songbirds using genomic data and bioinformatic methodology. All members of this superfamily have a shared “neo-sex chromosome”: a fusion between an autosome (chromosome 4A) and the existing sex chromosomes. The sex chromosomes of birds (ZW) formed in a common ancestor more than 100 million years ago. Since then, the W chromosome has undergone severe degradation and shortening, obscuring almost all traces of their early evolution. Additions of new genetic material through autosome-sex chromosome fusions, however, allow us to study the early stages of sex chromosome evolution.

I developed a computational pipeline aimed at discovering and visualizing sex chromosomes. I applied this pipeline to genomic data from species belonging to 13 different Sylvioidea families, and found that four additional autosome-sex chromosome fusions have occurred in different lineages within the group (involving chromosomes 3, 4, 5 and 8). These different fused regions have intermediate to extremely low W degeneration levels, with dosage sensitive and evolutionarily constrained genes being retained to a higher degree than other genes. I also studied the structure of these neo-sex chromosomes, how female gene expression changes in response to W degeneration and how recombination suppression extends along newly added sex chromosome regions. The work in this thesis shows that Sylvioidea songbirds are an ideal system for testing theory relating to sex chromosome evolution, and that bird sex chromosomes are more variable than previously believed.

Key words: Sex chromosome, birds, aves, Sylvioidea, genomics, neo-sex chromosome, chromosome fusion

Classification system and/or index terms (if any)

Supplementary bibliographical information		**Language** English
ISSN and key title		**ISBN** 978-91-7895-781-1 (pdf) 978-91-7895-782-8 (print)
Recipient’s notes	**Number of pages** 228	Price
	Security classification	

Signature  Date 2021-03-04

Evolution of sex chromosomes in Sylvioidea songbirds

Hanna Sigeman

LUND
UNIVERSITY

Cover illustration by Martin Sigeman and Lego by Philip Downing

Faculty of Science
Department of Biology

ISBN 978-91-7895-781-1 (pdf)
ISBN 978-91-7895-782-8 (print)

Printed in Sweden by Media-Tryck, Lund University
Lund 2021

Sex chromosomes wonderfully illustrate what I call 'dumb design': systems that make no functional sense but that can be understood in terms of evolution.

Jennifer A. Marshall Graves

Table of contents

List of papers

I. **Hanna Sigeman,** Bella Sinclair, Bengt Hansson. 2021. XYZWfinder: a snakemake pipeline for detecting and visualising sex chromosomes using short read data. Manuscript.

II. **Hanna Sigeman,** Suvi Ponnikas, Pallavi Chauhan, Elisa Dierickx, M de L Brooke, Bengt Hansson. 2019. Repeated sex chromosome evolution in vertebrates supported by expanded avian sex chromosomes. *Proceedings of the Royal Society B*, 286, 20192051 (doi: 10.1098/rspb.2019.2051).

III. **Hanna Sigeman,** Suvi Ponnikas, Bengt Hansson. 2020. Whole-genome analysis across 10 songbird families within Sylvioidea reveals a novel autosome-sex chromosome fusion. *Biology Letters*, 16, 20200082 (doi: 10.1098/rsbl.2020.0082).

IV. **Hanna Sigeman**, Hongkai Zhang, Salwan Ali Adeb, Bengt Hansson. 2021. A novel neo-sex chromosome in *Sylvietta brachyura* (Macrosphenidae) adds to the unprecedented avian sex chromosome diversity among Sylvioidea songbirds. Manuscript.

V. **Hanna Sigeman**, Maria Strandh, Estelle Proux-Wéra, Verena E. Kutschera, Suvi Ponnikas, Hongkai Zhang, Max Lundberg, Lucile Soler, Ignas Bunikis, Maja Tarka, Dennis Hasselquist, Björn Nystedt, Helena Westerdahl, Bengt Hansson. 2020. Avian neo-sex chromosomes reveal dynamics of recombination suppression and W degeneration. *bioRxiv*, 2020.09.25.314088 (doi: 10.1101/2020.09.25.314088). Submitted manuscript.

VI. **Hanna Sigeman,** Bengt Hansson. 2021. Evolutionary dynamics of enlarged sex chromosomes and novel pseudoautosomal regions in Sylvioidea songbirds. Manuscript.

VII. **Hanna Sigeman**, Suvi Ponnikas, Elin Videvall, Hongkai Zhang, Pallavi Chauhan, Sara Naurin, Bengt Hansson. 2018. Insights into avian incomplete dosage compensation: sex-biased gene expression coevolves with sex chromosome degeneration in the common whitethroat. *Genes*, 9, 373 (doi: 10.3390/genes9080373).

VIII. **Hanna Sigeman**, Bengt Hansson. 2021. The rate of W chromosome degeneration across multiple autosome-sex chromosome fusions in birds. Manuscript.

Author contributions

I. H.S. and B.H. planned and designed the study. H.S. and B.S. developed the software. H.S. performed the bioinformatic analyses presented in the manuscript. H.S. wrote the paper with assistance of B.H. All authors reviewed and approved the final manuscript.

II. H.S., S.P. and B.H. planned and designed the study. H.S. and P.C. performed bioinformatic analyses. E.D. and M.d.L.B. provided samples. H.S. and B.H. wrote the paper with input from all other co-authors. All authors reviewed and approved the final manuscript.

III. H.S., S.P. and B.H. conceived the study. H.S. performed analyses with input from B.H. H.S. and B.H. wrote the paper with input from S.P. All authors reviewed and approved the final manuscript.

IV. H.S. and B.H. planned and designed the study. H.Z. planned and performed laboratory work. H.S. performed bioinformatic analyses. S.A.A. provided samples. H.S. and B.H. wrote the paper with input from all other co-authors. All authors reviewed and approved the final manuscript.

V. H.S., M.S., H.W. and B.H planned and designed the study. H.S., M.S., E.P.-W., V.E.K, S.P., H.Z., M.L., L.S. and I.B performed bioinformatic analyses. H.S. and B.H. wrote the paper with input from all authors. All authors reviewed and approved the final manuscript.

VI. H.S. and B.H. planned and designed the study. H.S. performed bioinformatic analyses. H.S. and B.H. wrote the paper. All authors reviewed and approved the final manuscript.

VII. H.S., S.P. and B.H. conceived and designed the study; S.N. collected samples and extracted RNA; H.S., H.Z., E.V. and P.C. performed analyses with input from B.H.; H.S. and B.H. wrote the manuscript with input from coauthors. All authors reviewed and approved the final manuscript.

VIII. H.S. and B.H. planned and designed the study. H.S. performed bioinformatic analyses. H.S. and B.H. wrote the paper. All authors reviewed and approved the final manuscript.

Additional papers published during PhD

The following papers, on which I am a co-author, were published during my PhD. These papers are not included in this thesis.

IX. Suvi Ponnikas, **Hanna Sigeman**, Jessica K Abbott, Bengt Hansson. 2018. Why do sex chromosomes stop recombining? *Trends in Genetics*, 34(7), 492–503.

X. Bengt Hansson, **Hanna Sigeman**, Martin Stervander, Maja Tarka, Suvi Ponnikas, Maria Strandh, et al. 2018. Contrasting results from GWAS and QTL mapping on wing length in great reed warblers. *Molecular Ecology Resources*, 18(4), 867-876.

XI. Nicolas Dussex, David W. G. Stanton, **Hanna Sigeman,** Per G. P. Ericson, Jacquelyn Gill, et al. 2020. Biomolecular analyses reveal the age, sex and species identity of a near-intact Pleistocene bird carcass. *Communications Biology*, 3(1), 84.

XII. Shaohong Feng, Josefin Stiller, Yuan Deng, Joel Armstrong, […**Hanna Sigeman**…], et al. 2020. Dense sampling of bird diversity increases power of comparative genomics. *Nature*, 587, 252–257.

Abstract

Sex chromosomes were discovered more than 100 years ago. They have been studied intensely from a theoretical perspective since then, giving rise to a large body of testable predictions about their evolution from autosomes. A common feature of sex chromosomes is recombination suppression between the sex chromosome copies (X and Y in male heterogametic systems, or Z and W in female heterogametic systems). Without recombination, the sex-limited chromosome (Y or W) is expected to degenerate through the accumulation of deleterious mutations and repeat elements. Over long evolutionary time scales, this degeneration may leave the Y/W chromosomes short and almost completely devoid of functional genes.

Using genome sequencing technology, we can now study the full DNA sequence of sex chromosomes. The analysis of sequence data from a wide range of taxonomic groups has revealed that sex chromosomes are more dynamic and variable than previously believed. Several aspects of sex chromosome evolution, however, remain poorly understood, especially relating to the early stages of evolution from autosomes. This is partly because some hypotheses are challenging to test, but also because most well-studied sex chromosome systems are evolutionarily old and biased towards XY systems.

In this thesis, I study the evolution of sex chromosomes across Sylvioidea songbirds using genomic data and bioinformatic methodology. All members of this superfamily have a shared “neo-sex chromosome”: a fusion between an autosome (chromosome 4A) and the existing sex chromosomes. The sex chromosomes of birds (ZW) formed in a common ancestor more than 100 million years ago. Since then, the W chromosome has undergone severe degradation and shortening, obscuring almost all traces of their early evolution. Additions of new genetic material through autosome-sex chromosome fusions, however, allow us to study the early stages of sex chromosome evolution.

I developed a computational pipeline aimed at discovering and visualizing sex chromosomes. I applied this pipeline to genomic data from species belonging to 13 different Sylvioidea families, and found that four additional autosome-sex chromosome fusions have occurred in different lineages within the group (involving chromosomes 3, 4, 5 and 8). These different fused regions have intermediate to extremely low W degeneration levels, with dosage sensitive and evolutionarily constrained genes being retained to a higher degree than other genes. I also studied

the structure of these neo-sex chromosomes, how female gene expression changes in response to W degeneration and how recombination suppression extends along newly added sex chromosome regions. The work in this thesis shows that Sylvioidea songbirds are an ideal system for testing theory relating to sex chromosome evolution, and that bird sex chromosomes are more variable than previously believed.

Popular science summary

The human genome consists of 23 pairs of chromosomes. Chromosomes are where your genes live. Of these 23 pairs of chromosomes, 22 looks identical. The pair that determines sex, however, are weird. The X chromosome, of which females have two and males one, is a large chromosome that contains over 1000 different genes. But the tiny Y chromosome, which is only found in males, contains almost no genes at all. It hasn't always been like this though. Around 150 million years ago the X and the Y looked the same too. So, what happened?

This is a question that evolutionary theoreticians have been trying to answer for over 100 years. The basic idea from theory is that the sex chromosomes evolved from an ordinary pair of chromosomes, by acquiring a sex determining gene. On normal chromosomes, genes are exchanged between chromosome copies during meiosis, through a process called genetic recombination. This reshuffling of the genes between chromosome pairs is one reason why siblings from the same parents are genetically similar, but not identical.

However, sex chromosome copies (X and Y) stop recombining, at least over parts of their length. The idea is that if a gene that is really good for males but bad for females moves to the Y chromosome, then recombination suppression would evolve as it prevents this gene from leaving the Y chromosome. Once this happens, the Y chromosome will start to degenerate because bad mutations, which arise randomly all the time on chromosomes, are difficult to get rid of without recombination. Over long evolutionary time scales, this leads to the sorry state of the Y chromosome that we see today. This is what theoreticians think.

To test this theory, we need to be able to read and study the genetic code of sex chromosomes. This wasn't easy to do until recently. Just twenty years ago, it cost $100 million to sequence a human genome, but today it only costs $1000. This development in DNA sequencing technology allows researchers, like me, to study the genetic code of sex chromosomes of any species whose DNA we can get hold of. At last, we can test if our theories about why recombination stops, and why the Y chromosome shrinks, are right.

During my PhD, I used DNA sequencing data from different species of songbirds which all belong to a superfamily called Sylvioidea. This group of birds include some of the most common ones in Sweden, such as warblers, swallows and larks. Just like in humans, the sex chromosomes in birds evolved many million years ago.

But unlike in humans, in birds it is the females that have the weird sex chromosome copy. Male birds have two Z chromosomes, while females have one Z chromosome and one tiny W chromosome. What's special about Sylvioidea birds is that an ordinary chromosome attached itself to the Z chromosome, forming what is called a "neo-sex chromosome". This new part of the sex chromosomes allows us to test the old theory because it should stop recombining and start to shrink too.

Amazingly, I found that four more fusions between ordinary chromosomes and the sex chromosomes have happened in the Sylvioidea. These fusions formed the biggest bird sex chromosomes known to date! Because these fusions happened at different times in the past, this let me measure the rate at which the W chromosome shrinks, and identify which genes disappeared and when they did so. By studying the sex chromosomes of more than twenty different Sylvioidea species, I learned that W chromosomes in these birds keep the more important genes and let the less important genes go. I also learned that the W chromosome becomes bigger, by accumulating a lot of junk DNA, before it gets smaller. So, does the old theory stand up to modern data? My results suggest so.

Populärvetenskaplig sammanfattning

Den mänskliga arvsmassan består av 23 par kromosomer. Kromosomerna är där våra gener bor. Av de 23 kromosomparen ser 22 helt identiska ut. Men det sista paret, könskromosomerna, är annorlunda. X-kromosomen, som finns i två kopior hos kvinnor och en i män, är en stor kromosom med mer än 1000 olika gener. Y kromosomen, som bara män har, är liten och där finns nästan inga gener alls. Men så här har det inte alltid varit. För ungefär 150 miljoner år sedan såg X och Y också likadana ut. Så, vad var det som hände?

Det är en fråga som evolutionsteoretiker har försökt svara på i över 100 år. Grundidén är att könskromosomer bildas från vanliga kromosomer genom att en könsbestämmande gen flyttar dit, eller skapas genom mutation. På vanliga kromosomer byter delar av kromosomparen plats med varandra genom en process som kallas genetisk rekombination. Denna process skapar unika kromosomer i varje generation, och är en anledning till att syskon till samma föräldrar är genetiskt lika, men inte identiska.

Men på könskromosomer upphör rekombinationen över delar eller hela kromosomerna. Idén bakom detta är att om gener som är bra för män men dåliga för kvinnor flyttar över till Y-kromosomen, så kommer mutationer som minskar rekombinationstakten att vara gynnsamma, eftersom det förhindrar att genen flyttar över till X. Men utan rekombination är det svårt för Y-kromosomen att bli av med skadliga mutationer, som uppstår slumpmässigt på kromosomer hela tiden. Detta leder till att Y-kromosomen börjar brytas ner, bli kortare och förlora gener. Resultatet 150 miljoner år senare ser vi när vi jämför den ståtliga X-kromosomen med den stackars Y-kromosomen. Detta är vad teoretikerna kom fram till.

För att testa denna teori behöver vi läsa av och analysera könskromosomernas DNA, vilket inte var möjligt fram tills nyligen. För 20 år sedan kostade det nästan en miljard kronor att ta fram den kompletta DNA-sekvensen från en människa. Idag kostar det mindre än 10,000 kronor. Denna teknologiska utveckling gör det möjligt för forskare, som jag, att studera könskromosomernas genetiska kod, i vilken art som helst som vi kan få tag på DNA från. Till sist är det alltså möjligt att testa teorierna om varför rekombinationen upphör, och varför Y krymper, faktiskt stämmer.

I mitt avhandlingsarbete har jag analyserat sekvenseringsdata från olika sångfågelarter som alla tillhör en superfamilj som heter Sylvioidea. I den här

gruppen ingår många av våra vanligaste fåglar i Sverige, som olika slags sångare, svalor och lärkor. Precis som hos människor bildades fåglarnas könskromosomer för många miljoner år sedan. Men till skillnad från i människor är det fågelhonorna som har den konstiga könskromosomen. Fågelhanar har två Z-kromosomer, medan honorna har en Z-kromosom och en liten W-kromosom. Det speciella med fåglarna inom Sylvioidea är att de har en så kallad ”neo-könskromosom”, en utökad könskromosom som bildades när en vanlig kromosom sattes ihop med den gamla Z-kromosomen. Eftersom den här nya delen av könskromosomerna också slutade rekombinera, fast för inte lika länge sedan, kan vi använda den för att testa teorier om hur och varför W-kromosomen börjar krympa.

Jag upptäckte att det hade skett inte mindre än fem sammanslagningar (fusioner) mellan vanliga kromosomer och könskromosomerna inom Sylvioidea. Detta har skapat, vad vi vet, fågelvärldens största könskromosom! Eftersom fusionerna skett vid olika tidpunkter kunde jag använda dem för att undersöka hur snabbt W-kromosomen krymper, vilka gener som försvinner och när de försvann. Genom att studera könskromosomerna från fler än 20 Sylvioidea-arter lärde jag mig att W-kromosomerna lyckas hålla fast vid de viktigaste generna men låter de mindre viktiga försvinna. Jag lärde mig också att W-kromosomen faktiskt blir större, genom att samla på sig ”skräp-DNA”, innan den krymper. Så, stämmer de gamla teorierna när vi testar dem med modern sekvenseringsdata? Mina resultat tyder på det.

Background

Nearly all vertebrates have separate sexes, with females that produce large gametes (eggs) and males that produce small gametes (sperm) (Adolfi et al. 2018). There are several variations on this theme. For example, sequential hermaphrodites change sex during their lifespan and parthenogens can reproduce without males. As a general rule, however, reproduction in vertebrates requires two individuals of the opposite sex, which are determined at the embryonic stage (Capel 2017; Adolfi et al. 2018).

Whether an embryo differentiates into the female or male phenotype is governed by essentially the same set of genes in all vertebrates, although their relative position in the sex determination pathways may differ between species (Graves & Peichel 2010; Capel 2017). These pathways are directed towards either the female or male developmental program, depending on the presence or absence of one or several cues (or "sex determining mechanisms"). These cues are regulated in two broad ways. Environmental sex determination (ESD) systems are triggered by external cues such as temperature or light (Bull 1983). This type of sex determination is found in all crocodiles, many turtles and some fish (Bachtrog et al. 2014). In genetic sex determination (GSD) systems, which is the most common form among vertebrates (Bull 1983; Beukeboom & Perrin 2014; Bachtrog et al. 2014), the sex determination pathway is triggered by one or several genes. (Graves & Peichel 2010). Sex determination genes may be conserved across large taxonomic groups, such as in eutherian mammals (*Sry*) and birds (*Dmrt1*), but can vary between closely related species and even between different populations of the same species (Ogata et al. 2008).

My thesis focuses on GSD, specifically, the evolution of sex chromosomes. These are chromosomes that carry one, or several, sex determining gene(s). Sex chromosomes have been studied intensively from a theoretical perspective for more than a century, giving rise to a rich body of testable predictions on their expected evolutionary trajectory. Thanks to advances in genome sequencing technology, we can now test this theory with data.

What are sex chromosomes?

Many vertebrates have genetic sex determination systems with sex chromosomes that carry sex determining genes. The presence of a sex determining gene separates the sex chromosomes from the other chromosomes in the genome, which are called autosomes. Sex chromosomes typically evolve recombination suppression around the sex determining gene, which over time results in a build-up of genetic differentiation between the chromosome copies (Bachtrog 2006). Structurally, sex chromosomes can be divided into two primary components: the fully sex-linked region containing the sex determining gene (the "sex-determining region", or SDR), and the partially sex-linked and still recombining pseudoautosomal region (or PAR). The relative size of these two components can vary; some sex chromosomes still recombine throughout most of their length, while others lack recombination completely.

There are two main types of genetic sex determination systems: XY (male heterogametic systems) and ZW (female heterogametic systems). These terms are not indicative of shared origin. Both XY systems and ZW systems have evolved many times independently (see e.g., Matsubara et al. 2006; Tree of Sex Consortium 2014). Instead, they are defined by chromosome configuration differences between the sexes. In XY systems, which is the prevailing system in eutherian mammals such as humans, females have two virtually identical sex chromosome copies (two X chromosomes) while males have two genetically differentiated copies (one X chromosome and one Y chromosome). The human X and Y chromosome were once identical, as they originated from the same chromosome pair. However, due to a lack of recombination for over 100 million years they have evolved extreme differences in both size and gene content. The X chromosome has retained most of its original size and gene content while the Y chromosome is small and contains almost no functional genes (Skaletsky et al. 2003; Bellott et al. 2010; Mueller et al. 2013).

ZW systems are a mirror image of XY systems: males have two identical chromosome copies (ZZ), and females have two distinct copies (ZW). There are further variations of sex chromosome systems in vertebrates, for example where the heterogametic sex has lost their unique (Y or W) chromosome copy completely ("XO systems" or "ZO systems"), or where the sex chromosomes consist of many different genetic elements (for example in the platypus; Grützner et al. 2004).

How do we find sex chromosomes?

The first sex chromosomes were discovered more than one hundred years ago using **cytogenetic** methods, which involved examining the contents of cells using

microscopes. Sex chromosomes were identified through sex-specific differences in either the size or number of chromosomes (Henking 1891; Wilson 1905; Stevens 1906). At this time, sex chromosome systems with large size differences between the X and Y (or Z and W) chromosomes were naturally more likely to be found. Cytogenetics is still an important and widely used method for identifying sex chromosomes to this day, using more sophisticated methods such as C-banding and fluorescent *in situ* hybridization (FISH) (see e.g., Kawai et al. 2007; Iannucci et al. 2019). Sex chromosome systems of low differentiation, however, remain challenging to identify using these methods.

Sex chromosomes can also be identified with genome sequencing data, by searching for genomic regions that differ between females and males. Broad comparative genomics studies have revealed two important insights. Firstly, that sex chromosome differentiation is a highly continuous trait, with sex chromosome copies often exhibiting only minor genetic differences. And secondly, that sex chromosomes have evolved independently numerous times across the tree of life (Bachtrog et al. 2014).

The discovery and characterisation of sex chromosome systems is still a major research area in the field of sex chromosome evolution (**Thesis aims 1-4**). The current high rate of discovery of novel systems (see e.g., Gamble et al. 2015; Gammerdinger & Kocher 2018; Jeffries et al. 2018) suggests that we are nowhere near a complete characterization of sex chromosome diversity in vertebrates, let alone in other taxonomic groups. Genome sequencing data is now being produced at an unprecedented rate; a telling fact is that data storage and the analysis of genomic data in energy efficient ways that reduce carbon emissions are major problems (Papageorgiou et al. 2018). To promote discovery of new sex chromosome systems, and to do so in an efficient, economic as well as sustainable way, development of computational methods aimed at utilizing different types of genomic data is imperative.

In addition to revolutionizing the discovery and characterisation of sex chromosome systems, genome sequencing technology allows us to test and expand theory on sex chromosome evolution that was developed in the pre-genomic era (Bachtrog 2006; Bachtrog et al. 2011; Abbott et al. 2017) (**Thesis aim 5 and 6**). In the following section, I outline the standard (and still prevailing) model of sex chromosome evolution, which was largely developed before and during the 1980's. Due to the available knowledge at that time, it was mainly developed with highly differentiated XY systems in mind. However, most aspects apply to ZW systems as well.

How do sex chromosomes evolve?

Formation from autosomes and recombination suppression

Sex chromosomes evolve from autosomes through the acquisition of a sex determining gene; either through mutation, translocation or duplication (Muller 1914; Muller 1918; Ohno 1967; Beukeboom & Perrin 2014) (Figure 1). Sex chromosomes can evolve both in hermaphrodites and in species with separate sexes (determined environmentally or genetically). Once a sex determining gene is acquired, this is typically followed by recombination suppression around this gene. An early and still prevailing theory is that mutations that halt recombination will be favoured by selection in regions containing both a sex determining gene and one or several sexually antagonistic genes (genes in which the alleles have opposing fitness effects in each sex). This is because genetic linkage with the sex determining gene allows these sexually antagonistic genes to be predominantly inherited through the sex in which they have a beneficial fitness effect (Fisher 1931; Bull 1983; Rice 1987; Bergero & Charlesworth 2009). It is hard to overstate the importance of recombination suppression for the evolution of sex chromosomes; its centrality is reflected in the vast theoretical as well as empirical work devoted to understanding this process (recently reviewed in Wright et al. 2016; Charlesworth 2021; and by us in Ponnikas et al. 2018). The lack of recombination allows genetic differentiation to build up between the sex chromosome copies, which over time may vanquish almost all original resemblance.

Inversions have long been recognized as a possible mechanism for ceased recombination on sex chromosomes (Bowen 1965; Charlesworth et al. 2005). An inversion capturing a sex determining gene and a sexually antagonistic gene would create instant linkage despite considerable physical distance. Additional inversions, that capture other sexually antagonistic genes, could then create recombination suppression over an increasingly large part of the sex chromosomes. In line with this hypothesis, “evolutionary strata” of genetic divergence caused by successive stepwise recombination suppression events have been found in many species, e.g., in mammals (Lahn & Page 1999), birds (Handley et al. 2004) and plants (Nicolas et al. 2004).

The sex chromosome copies spend unequal amounts of time in each sex. For instance, X chromosomes are predominantly inherited through females and Y chromosomes exclusively through males. This bias is expected to lead to sexualization of the gene content of sex chromosomes. X chromosomes are predicted to become enriched for recessive male-beneficial mutations, even when they come at a cost to females, as these are always exposed to selection in males. As X chromosomes spend more time in females, they are predicted to become enriched for dominant female-beneficial mutations (Rice 1984). Likewise, Z

chromosomes will accumulate dominant male-beneficial and recessive female-beneficial mutations (Bachtrog et al. 2011).

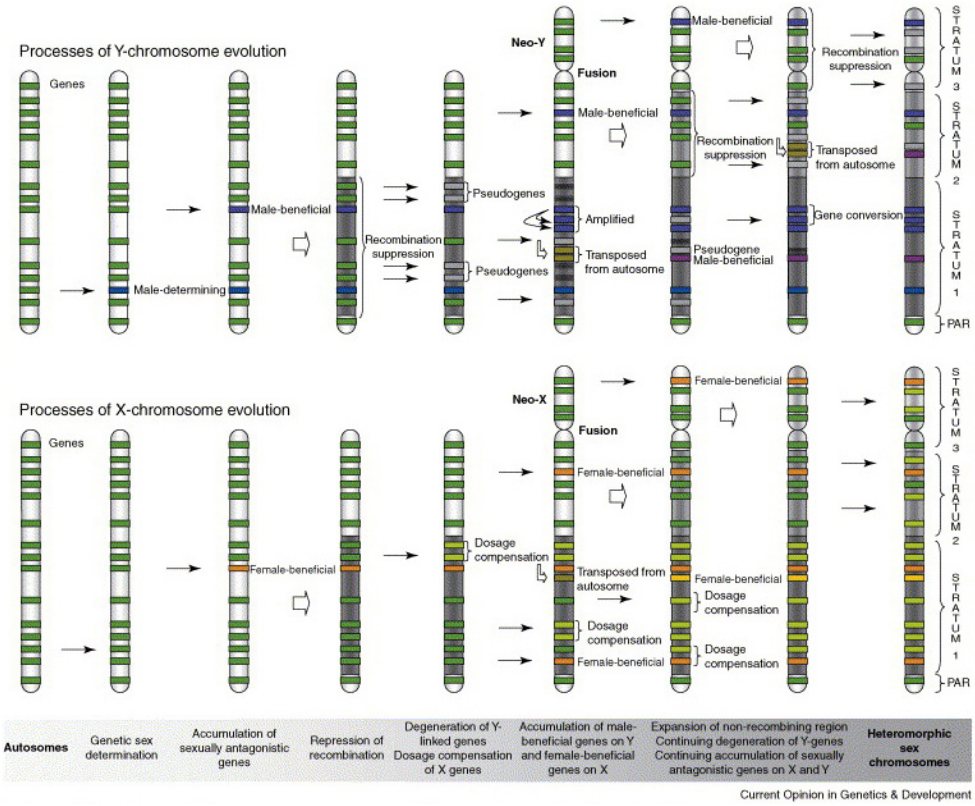


Figure 1. The standard model of sex chromosome evolution. Sex chromosomes evolve from autosomes as they acquire a sex determining gene. Recombination suppression is favored by selection as it increases linkage between the sex determining factor and sexually antagonistic genes. The lack of recombination allows the sex chromosome copies (X and Y) to genetically differentiate, while also resulting in Y chromosome degeneration. Dosage compensation may restore gene expression levels in the heterogametic sex to compensate for the loss of functional Y-linked gametologs. Fusions between sex chromosomes and autosomes may form “neo-sex chromosomes”. Reprinted from Bachtrog (2006) with permission from Elsevier.

Fusions between sex chromosomes and autosomes create “neo-sex chromosomes” (Figure 1). One explanation for why such rearrangements become fixed is that they may bring together favourable combinations of genes (Lewis & John 1968), for example genes with sexually antagonistic effects (Charlesworth & Charlesworth 1980). Neo-sex chromosomes could also become established though heterozygote advantage (Charlesworth & Wall 1999), as a result of genetic drift (Lande 1979) or through meiotic drive (Yoshida & Kitano 2012). Additional recombination suppression events following such fusions can then create full sex-linkage of this new region, which then starts to evolve in the same way as the original sex chromosome region (Charlesworth & Charlesworth 2000).

Consequences of recombination suppression

Detailed studies, primarily from highly differentiated XY systems such as eutherian mammals (Skaletsky et al. 2003; Soh et al. 2014) and fruit flies (Bachtrog et al. 2003; Bachtrog et al. 2013), have given us a nuanced understanding of the long-term consequences of recombination suppression (Bachtrog et al. 2014). Evolutionarily old sex-limited chromosomes (Y or W) are often highly heterochromatic and gene poor. The core idea behind Y (or W) chromosome degeneration is that the absence of recombination weakens the efficiency of selection against deleterious mutations and reduces the rate of fixation of beneficial mutations (reviewed in Bachtrog 2013). This is believed to occur due to evolutionary processes such as background selection (Charlesworth 1994), Muller's ratchet (Charlesworth & Charlesworth 2000) and genetic hitchhiking (Rice 1987). The rate of degeneration may depend on many factors, including the number of genes in the non-recombining sex chromosome region and the effective population size of Y and W chromosomes (Bachtrog 2008).

The accumulation of deleterious mutations in genic or regulatory regions is followed by adaptive silencing of gene expression. This change in gene expression, brought about by copy number differences, can lead to disruption of crucial gene networks (Livernois et al. 2012; Graves 2016). This silencing may therefore be followed by dosage compensation to restore equal expression levels between the sex (Figure 1). These steps – mutations, silencing and dosage compensation – may evolve in a sequential manner. However, they may also evolve simultaneously and form a feed-back loop ("degeneration by regulatory evolution"; Lenormand et al. 2020).

Mutations with only minor negative fitness effects are predicted to be especially important for Y degeneration, as highly deleterious mutations are likely to be selected against (Charlesworth & Charlesworth 2020). Genes surviving on Y chromosomes over long evolutionary timescales are therefore expected to have genetic signatures of strong purifying selection, with little genetic variation, as only strong selection can maintain them in the absence of recombination (Bachtrog 2013). In line with this reasoning, heavily degenerated Y chromosomes are often enriched for broadly expressed and dosage sensitive genes under strong purifying selection (Bellott et al. 2014; Bellott et al. 2017; Xu et al. 2019).

In conclusion, sex chromosome evolution can lead to extreme degeneration of the sex-limited chromosome copy, essentially leaving the heterogametic sex monoallelic for all genes while the homogametic sex retains two functional gene copies. This gene copy imbalance is expected to cause problems in gene expression optimization in one or both sexes, as well as between sex chromosomes and autosomes. Moreover, pronounced differences between X and Y (or Z and W) chromosomes may lead to difficulties in pairing up during meiosis. It is to these problems that the 'dumb design' quote in the start of this thesis refers.

Gaps and problems with the Standard Model

The standard theory of sex chromosome evolution was largely developed before and during the 1980's, with highly heteromorphic XY systems in mind (Abbott et al. 2017). This was partly due to the methodological bias (cytogenetics) which meant that empirical studies until this point had mostly been conducted on a few model organisms with highly heteromorphic XY sex chromosomes, such as eutherian mammals and fruit flies. Today, thanks to advances in DNA sequencing technology, empirical studies can move beyond these model systems and test the different steps in the standard model of sex chromosome evolution empirically in more detail than ever before.

Many sex chromosomes have been shown to evolve in concordance with the standard theory (Bachtrog et al. 2011; Beukeboom & Perrin 2014; Bachtrog et al., 2014; Bergero & Charlesworth 2009), but not in every aspect and not in all systems. For example, some sex chromosomes appear to remain homomorphic despite considerable age (Perrin 2009; Rodrigues et al. 2018). Furthermore, dosage compensation across the entire sex chromosomes was once thought to be universal (Ohno 1967) but recent studies have shown that the mode and degree of dosage compensation is in fact highly variable across vertebrates (reviewed in Graves 2016).

Many aspects of the standard model remain poorly explored to this day. In particular, we currently know much less about the early stages of sex chromosome evolution than we know about the long-term effects of sex-linkage. This is partly because of the study bias towards old and heteromorphic XY systems, which can tell us very little about the origins of sex chromosomes, and partly because many of the hypothesized drivers of sex chromosome evolution are difficult to characterize and may be ephemeral (for example sexual antagonism and meiotic drive). Below I highlight key areas of sex chromosome evolution that require further study.

Sex chromosome formation from autosomes

Differences between taxonomic groups

The rate of sex chromosome formation varies greatly across vertebrates. The sex chromosomes in eutherian mammals (XY) and birds (ZW) each originated over 100 million years ago (Cortez et al. 2014; Zhou et al. 2014) and have remained stable since then. In both taxonomic groups, the sex-limited chromosome (Y and W) has undergone massive degeneration and shortening. However, sex chromosome systems in other vertebrate groups, such as fish (Kitano & Peichel, 2012; Myosho et al. 2015) and amphibians (Jeffries et al. 2018), are less stable, with frequent shifts between sex chromosome systems, so called turnovers. In the latter groups, the sex

chromosomes are often homomorphic because frequent turnovers do not give Y/W chromosomes time to degenerate. Reptiles show a mix of stability and turnovers: some groups have highly stable sex chromosomes (Rovatsos et al. 2014), while others have undergone frequent turnovers (Gamble et al. 2015) (Figure 2).

Figure 2. Phylogeny of vertebrates with information about their sex determination system, or lack thereof. Some vertebrate groups, such as mammals and birds, have highly stable sex chromosomes (XY/ZW). Other vertebrates, such as amphibians and fish, have highly variable sex chromosome systems, and also variation between environmental (ESD) and genetic sex determination systems. Reprinted from Sandell & Otto (2016) with permission from the Genetics Society of America. The figure is an adaptation from The Tree of Sex Consortium (2014).

The "evolutionary trap" hypotheses states that turnovers are less likely in evolutionarily old and heteromorphic sex chromosome systems, as turnover events typically involve the fixation of either chromosome copy (Vicoso 2019). Fixation of the sex-limited (Y or W) chromosome in species with high Y/W chromosome degeneration is likely to be lethal, as many genes present on the other sex chromosome (X or Z) would disappear. On the other hand, sex-limited chromosomes may accumulate genes essential for sex-specific survival and fertility over time, which may prevent fixation of the X or Z. This hypothesis could therefore explain the observed difference in sex chromosome stability across vertebrate groups.

However, recent evidence of turnovers (Vicoso & Bachtrog 2013) or loss of the Y chromosome (Kuroiwa et al. 2010) in ancient sex chromosomes systems suggests

that sex chromosome heteromorphy does not constitute an unescapable evolutionary trap. Furthermore, a recent study (utilizing sex chromosome size differences identified through karyotype data among fish species) did not observe differences in the rate of turnovers between homomorphic and heteromorphic sex chromosomes (Pennell et al. 2018). However, as the authors point out, size differentiation is not necessarily representative of genetic differentiation and more studies on this question are needed. Continued efforts to characterize sex chromosome diversity will aid in our understanding of what evolutionary processes and ecological conditions underlie turnovers (Pennell et al. 2018).

Formation of neo-sex chromosomes

Another source of sex chromosome diversity are neo-sex chromosomes, which form through autosome-sex chromosome fusions (White 1973; Charlesworth & Charlesworth 1980). Similar to complete turnovers, these rearrangements seem to be more common in some groups than others (Tree of Sex Consortium 2014; Vicoso and Bachtrog 2015; Blackmon et al. 2016; Pennell et al. 2015). Autosome-sex chromosome fusions are likely to be less constrained by sex chromosome differentiation compared to complete turnovers as they are not as tightly associated with the loss of a chromosome copy (but see e.g., Bracewell et al. 2017). Broad scale comparative work across vertebrates has shown that the type of heterogamety may affect fixation rates of neo-sex chromosomes; they appear to occur more frequently in XY systems compared to ZW systems (Pennell et al. 2015). Specifically, Y-autosome fusions are significantly more common than other types, which has been suggested to be a result of genetic drift and high mutation rates in males (Pennell et al. 2015).

Which autosomes become sex-chromosomes?

Sex chromosomes have evolved independently numerous times across vertebrates. Comparative genomic studies, especially among amniotes, have revealed that many of these origins involve the same autosomes and sex determining genes (Graves & Peichel 2010; O'Meally et al. 2012; Ezaz et al. 2016). This suggests that certain genes and autosomes may be especially suited for a sex determining role (Graves & Peichel 2010; O'Meally et al. 2012; Ezaz et al. 2016). This may be because they carry genes that are sexually antagonistic or genes that are part of the vertebrate sex determination pathway. In some clades with recurrent turnovers, such as frogs and cichlid fish, the same chromosome has been involved multiple times (Jeffries et al. 2018; Gammerdinger & Kocher 2018). In the case of frogs, some chromosomes were significantly overrepresented (Jeffries et al. 2018). Likewise, it has been suggested that certain autosomes are more likely to fuse with existing sex chromosomes than others (Ross et al. 2009; Graves & Peichel 2010; Pala et al. 2012a). Additional data from independent formations of sex chromosomes from autosomes, either from complete turnovers or autosome-sex chromosome fusions, would give insights to these observed patterns (**Thesis aim 2**).

Recombination suppression

Ultimate explanations

The causal role of sexually antagonistic genes in bringing about recombination suppression ("the SA hypothesis") on sex chromosomes is a long-standing and generally accepted theory. However, there is surprisingly little empirical evidence (either experimental or observational) to support this hypothesis, which has been described as '*a widely accepted hypothesis ... due partly to the absence of an alternative*' (Charlesworth 2018). One problem is that it is challenging to identify genomic regions under sexual antagonism. Furthermore, even if such regions were identified on sex chromosomes, they may be there as a consequence of recombination suppression, rather than having caused it. This is because sex chromosomes may also accumulate sexually antagonistic genes after recombination suppression. A classic paper by Rice (1984) suggested such accumulation is more likely on sex chromosomes than autosomes, leading to sex chromosomes becoming "hot spots" for sexually antagonistic genes (but see Fry 2010, who argue that the acceptance of this hypothesis is premature). This naturally complicates the separation of cause and consequence of recombination suppression, especially in old sex chromosome systems (Ironside 2010).

Importantly, as non-recombining regions on mating-type chromosomes have been found in species without sexes, which by definition lack sexual antagonism (e.g., Branco et al. 2017), other explanations are needed to fully explain the evolution of recombination suppression (Ironside 2010). Non-selective processes, such as meiotic drive (Úbeda et al. 2014), heterozygote advantage (Charlesworth & Wall 1999) and genetic drift (Charlesworth et al. 1987) have also been proposed as possible explanations for recombination suppression between nascent sex chromosomes. However, we also lack conclusive evidence that these processes cause recombination suppression (Wright et al. 2016). Furthermore, heavy degeneration of the sex-limited (Y or W) chromosomes may eliminate all traces of what caused the initial recombination suppression. Studying young sex chromosome systems is, therefore, key to learning more about the causes of recombination suppression (**Thesis aim 5**).

Proximate explanations

The study-bias towards evolutionarily old sex chromosome systems has also affected our understanding of the proximate factors behind recombination suppression. For example, inversions are thought to play an important role in mechanistically achieving recombination suppression. However, because inversions tend to form in non-recombining regions, it may be impossible to distinguish cause from consequence in evolutionary old sex chromosome systems. Without recombination there is no longer selection to maintain gene order, and Y and W chromosomes are often heavily rearranged (Bergero & Charlesworth 2009). It is

also possible for sex chromosomes to establish in regions where an inversion polymorphism has already caused recombination suppression (Natri et al. 2019), which further complicates the distinction between cause and consequence.

The gradual spread of recombination suppression, for example through the accumulation of transposable elements (Iwase et al. 2003; Xu et al. 2019) or via DNA methylation (Gorelick et al. 2003), has been proposed as an alternative mechanism. This gradual spread is supported by evidence from some recently formed sex chromosome systems (Natri et al. 2013). To understand the relative importance of the proximate mechanisms underlying recombination suppression, it is important to study sex-linked regions in diverse taxonomic systems of more recent origin (**Thesis aim 5**).

The absence of recombination suppression

Many evolutionarily young sex chromosomes have evolved recombination suppression only around a small area (e.g., Liu et al. 2004; Kondo et al. 2009). In contrast, among taxa with evolutionarily old sex chromosomes, recombination suppression has progressed over almost the entire sex chromosomes (e.g., Matsubara et al. 2006; Vicoso et al. 2013; Zhou et al. 2014; Cortez et al. 2014; Bellott et al. 2014; Bellott et al. 2017). Although the correlation between the age and level of heteromorphism of sex chromosomes is generally strong, it is not absolute (Bachtrog et al. 2014). In ratite birds, for example, a large part of the sex chromosomes is still recombining despite having formed over 100 million years ago (Vicoso et al. 2013; Zhou et al. 2014). A possible explanation for the maintenance of this large pseudoautosomal region (PAR) is that there may be few sexually antagonistic genes outside of the sex determining region (SDR) and thus weak selection for recombination suppression to evolve (Rice 1987). Alternatively, sex-specific expression of sexually antagonistic genes could resolve sexual conflict, thus removing selection for recombination suppression (Jordan and Charlesworth 2012).

Where to look for answers

As already mentioned, one way to improve our understanding of the evolution of recombination suppression, both in terms of selective (or non-selective) forces and how it is achieved mechanistically, is to study sex-linked regions of recent origin (Charlesworth et al. 2005; Bergero & Charlesworth 2009; Furman et al. 2020). This naturally includes novel sex chromosome systems that formed *de novo*, but also genomic regions on older sex chromosome systems where recombination suppression has occurred recently (**Thesis aim 5**). The boundary between the PAR and the fully sex-linked SDR may be of special interest, as sexually antagonistic genes are predicted to have a higher likelihood of surviving in this region (Otto et al. 2011). Studying PARs with size differences between closely related species (or even intraspecific differences) may also be informative, as they may contain clues as to why some species evolve recombination suppression while others do not.

Furthermore, female and male heterogametic systems are expected to vary in selection pressures and mutation rates. Contrasting data from XY and ZW systems, which may reveal differential rates of recombination suppression, is however hindered by a lack of well-studied PARs from ZW systems. Detailed studies of large PARs in ZW systems, such as in birds, are therefore greatly needed (Bachtrog et al. 2011; **Thesis aim 4**).

Consequences of recombination suppression

Degeneration rates and the influence of heterogamety differences

The different evolutionary processes underlying Y/W chromosome degeneration are well studied from a theoretical perspective. There are clearly expressed predictions regarding their relative influence over time, as well as under varying Y/W chromosome effective population sizes and in relation to the number of sex-linked genes (Bachtrog 2008). However, the empirical bias towards studying old XY systems has limited our understanding of these evolutionary processes. This is not only because evolutionarily old sex chromosomes are poor model systems for exploring the evolutionary forces and molecular mechanisms driving degeneration (Bachtrog 2013), but also because Y and W chromosomes are expected to degenerate at different rates. Males typically have faster mutation rates (Kirkpatrick & Hall 2004), lower effective population sizes and experience stronger sexual selection pressures than females (Naurin et al. 2010). These sex differences, which are predicted to be especially pronounced in polygynous species (Charlesworth & Charlesworth 2000; Graves 2006; Bachtrog 2013), should result in higher rates of degeneration on Y chromosomes compared to W chromosomes. To evaluate how these predictions from theory fit with empirical data, additional studies from young heterogametic systems are needed (Bergero & Charlesworth 2009; **Thesis aim 6**).

Dosage compensation - the largest discrepancy?

The process of dosage compensation is perhaps the area of sex chromosome evolution where results from recent empirical studies are most at odds with predictions from early theory. Ohno (1967) predicted that the heterogametic sex should upregulate expression levels of the X (or Z) chromosome to compensate for the loss of functional genes on the Y (or W) chromosome. This was supported by early empirical work from the model-systems of the time; fruit flies, eutherian mammals and the nematode *Caenorhabditis elegans*, which all displayed equal expression levels between sexes, despite having highly heteromorphic XY systems (or XO in *C. elegans*) (Mank 2013). Since then, results from gene expression studies across a broad range of taxa have overthrown the notion of complete dosage compensation as a necessary, or even common, step in the evolution of heteromorphic sex chromosomes (Mank 2013).

Instead, dosage compensation has been shown to be highly variable across taxa, and even within vertebrates, with dosage compensation being achieved to different degrees and through different mechanisms (reviewed in Mank 2013; Disteche 2012). Across monotreme mammals, reptiles, fish and birds, dosage compensation seems to be only partial and tissue-specific (Graves 2016). Recent studies have even revaluated the status of eutherian mammals as having complete dosage compensation. Instead of a compensatory upregulation of the X chromosome in males to its ancestral state, which was the original definition of dosage compensation, one of the X chromosomes is instead globally inactivated in females (Julien et al. 2012). This produces equal gene expression levels between sexes (dosage balance), but should, according to a strict definition, not be considered complete dosage compensation (Gu & Walters 2017). The picture is further complicated by the possibility of regulatory changes of the autosomal genes that are interacting with sex-linked genes as a response to evolving sex chromosome expression patterns (Julien et al. 2012).

Dosage compensation is thus an area of sex chromosome evolution where the need for theoretical advances, as well as data from additional independent origins of sex chromosomes is large. Highly valuable would be empirical studies that focus on sex chromosomes that were recently evolved *de novo*, either by replacing a previous sex chromosome system or from systems with environmental sex determination. While studies of dosage compensation in neo-sex chromosomes are informative, any changes to the gene expression levels in the recently added part are likely to be influenced by the pre-existing mode of dosage compensation (Gu & Walters 2017).

The presence or absence of dosage compensation mechanisms inherently affects other aspects of sex chromosome evolution. For example, a recent model of Y chromosome evolution shows that simultaneous Y degeneration and dosage compensation could lead to a higher rate of degeneration than predicted by standard theory, and a better fit to observations from gene-poor Y chromosomes (Lenormand et al. 2020). How well this XY model fits with data from ZW systems will be interesting, as a striking pattern so far is that ZW systems are almost never associated with complete dosage compensation (Parsch & Ellegren 2013). Studies of rates of Y/W chromosome degeneration from systems with different modes of dosage compensation may tell us more about these processes.

Study system

My thesis work is based on species belonging to the Sylvioidea superfamily, a large clade of songbirds consisting of around 1200 species (Alström et al. 2014). This large clade, which includes more than 10% of all extant birds, originated 20-25 million years ago and is represented on all the continents of the world except Antarctica (Oliveros et al. 2019). Prior to the start of my PhD project, evidence of a neo-sex chromosome in the form of sex-linked genetic markers located on a chromosome known to be autosomal in other birds (chromosome 4A) was found in several species of Sylvioidea (Pala et al. 2012a; Pala et al. 2012b).

Several factors make the Sylvioidea an ideal system for studying sex chromosome evolution. Firstly, the finding of a neo-sex chromosome was unexpected since sex chromosomes in birds, which formed over 100 million years ago, were generally believed to be extremely stable (Nanda et al. 2008; Zhou et al. 2014). Secondly, neo-sex chromosomes enable us to study the early stages of sex chromosome evolution in birds, where most of the ancestral W chromosome is degenerated. Thirdly, as birds are both female heterogametic (ZW) and devoid of dosage balance, this constitutes a promising study system for generating results that can be compared with the better studied XY systems. Fourthly, chromosome 4A (the autosome now fused to the sex chromosomes in Sylvioidea) is partly homologous to the mammalian X chromosome and contains a gene called the "androgen receptor" (*ar*), which is involved in the development of male primary and secondary sexual characteristics in some species.

As my PhD progressed, I discovered that there have been multiple autosome-sex chromosome fusions within Sylvioidea. While the fusion involving chromosome 4A is common to the entire group (and will be referred to as the "Sylvioidea-shared fusion"), other fusions have formed later in specific Sylvioidea lineages. The lineage-specific fusions mean that we can study several independent origins of sex-linkage in this group. Furthermore, as these fusions occurred at distinct time points (between ~25-8 million years ago), we can use this study system to explore temporal patterns of early sex chromosome evolution. This variation in sex-linked genetic material between studied species also means that we can use species in which each region is not sex-linked as a proxy for its original (autosomal) state. The Sylvioidea study system is, therefore, ideal for answering a wide range of questions related to the different steps in the evolution of sex chromosomes.

Thesis aims

The overall aim of my thesis is to provide a deeper understanding of sex chromosome evolution, using genomic analysis of neo-sex chromosomes in the Sylvioidea superfamily of songbirds. The specific aims of this thesis can be broadly categorized under three themes.

The first theme, "Discovery", relates to the discovery of novel sex chromosome diversity (**papers I-IV**). In the first paper (**paper I**) I present a computational pipeline for identifying sex chromosomes using genome sequencing data, and in the other papers (**papers II-IV**) I apply these methods to characterize sex chromosome diversity within Sylvioidea birds. My second theme is "Characterization", in which I conduct detailed analyses of the structure of the autosome-sex chromosome fusions, and their effect on the ancestral PAR (**papers V** and **VI**). My third theme, "Testing theory", uses the Sylvioidea study system to test different aspects of sex chromosome evolution theory (**papers V-VIII**). Specifically, I studied; (*a*) rates and spatial patterns of recombination suppression (**papers V-VI**), (*b*) rates and gene-specific patterns of W chromosome degeneration (**papers V** and **VIII)** and (*c*) how sequence divergence between gametologous gene pairs (Z and W) and gene expression coevolve (**paper VII**). In Table 1, I have categorized the specific questions that I address in this thesis (**Thesis aims 1-6**) under these three research themes.

Table 1. Questions that I adress in this thesis, categorized under different research themes and aims. Papers that explore each of these questions are specified.

Theme	Aim	Question	Paper
Discovery	1	How can sex chromosomes be identified in a user-friendly way using minimal genomic resources?	**I**
	2	What is the extent of sex chromosome diversity within Sylvioidea songbirds?	**II-IV**
Characterization	3	What is the chromosomal structure of the Sylvioidea sex chromosome fusions?	**V-VI**
	4	How was the ancestral pseudoautosomal region (PAR) affected by the fusions?	**VI**
Testing theory	5	How does recombination suppression evolve along the added part of a neo-sex chromosome?	**V-VI**
	6	How does sex-linkage affect W chromosomes in terms of repeat accumulation, gene loss and gene expression?	**V, VII-VIII**

General methodology

I studied sex chromosome evolution in Sylvioidea songbirds by analysing DNA and RNA sequencing data using bioinformatic software and methods. Thanks to the long tradition of bird research at Lund university, blood samples from several species within Sylvioidea were available in the department freezers. Additional samples were obtained from colleagues and collaborators at other universities. I am grateful to my collaborators and co-authors who extracted DNA and RNA from different study species, which allowed me to focus on the computational analysis of this data. Details of DNA/RNA extraction and sequencing technologies can be found in each paper. I also use published genomic data or genome assemblies in some form in all papers. Details regarding this is specified in each paper.

The type of data and the bioinformatic methods used varied between projects. While methodological details are described in each paper, below I have grouped them according to similarities in the types of data and methods used. The order in which I present these groups does not strictly follow the order in which the papers occur later (**paper I-VIII**). This is done intentionally with the aim of reflecting the journey of this thesis work, which has changed in response to new findings about this system. I will not give details of the specific bioinformatic software used in each paper here, since this information can be found in the methods section of each paper.

Throughout this thesis, I refer to Sylvioidea chromosome regions by their homology to regions in the zebra finch (*Taeniopygia guttata*) genome, which is the most well-studied songbird genome and is commonly used for this purpose. Birds have, in general, highly stable karyotypes with few interchromosomal rearrangements. A linkage map analysis I was involved in (Ponnikas et al. 2020 and Ponnikas et al. unpublished) revealed, in support of this, that the chromosome structure of one Sylvioidea species (the great reed warbler; *Acrocephalus arundinaceus*), is essentially identical to the zebra finch with the exception of the fusion between chromosomes Z and 4A.

Papers I-IV

A main aim of this thesis was to sequence and study the sex chromosomes of species-representatives of as many of the ~22 families (Fregin et al. 2012) within the Sylvioidea group as possible. Our sampling and sequencing strategy was to target one male and one female individual of at least one species per family, using Illumina paired-end sequencing technology. Following the discovery of additional

fusions in some families, we also sequenced additional species within these groups, where possible. Subsets of this data were reused in all papers in this thesis.

In **papers II-IV**, I used this genome data to identify and characterize sex chromosomes across different families and genera. The general methodology was to construct a reference genome based on the male (homogametic; ZZ) sample of each species. I then aligned genome data from males and females to these reference genomes, and conducted scans for genome regions with sex-specific differences in either genome coverage or heterozygosity.

While working on **papers II-IV**, I developed a computational pipeline for discovering sex-linked genomic regions using this kind of data. This pipeline was further developed and resulted in a separate paper (**paper I**), with the hope that this method will be useful to other researchers. In this paper, I validated the pipeline using publicly available genomic data from different taxonomic groups, including insects, fish, reptiles, birds and mammals.

Paper V

Some of my thesis aims were to study the chromosomal structure of the previously discovered (Sylvioidea-shared) neo-sex chromosome involving chromosomes Z and 4A and to study how the evolution of recombination suppression and subsequent W chromosome degeneration had progressed along the added sex chromosome region (the region that is homologous to chromosome 4A). The core of this project (**paper V**) involved constructing a high-quality reference genome of a female (ZW) great reed warbler (*Acrocephalus arundinaceus*). Long-read sequencing data was used as a backbone for the assembly, which was then polished using short-read sequence data and scaffolded using linked reads and optical mapping data. Sex-limited chromosomes (Y and W) are challenging to assemble using short-read data due to their repeat rich and haploid nature (Tomaszkiewicz et al. 2017). The use of long-read sequencing technology as a base for the genome assembly was therefore crucial, as I wanted to study the structure of the W chromosomes (which requires highly contiguous scaffolds). RNA expression data from the same female was used to generate a gene annotation for this genome assembly, followed by extensive manual curation of all gametologous (Z and W) gene pairs.

This annotated reference genome was used to study (1) the chromosome structure, (2) repeat-landscape and (3) nucleotide diversity of the great reed warbler sex chromosomes. Furthermore, we evaluated (4) the timing and spatial progression of recombination suppression, and (5) differences in the characteristics of genes where the W-linked gametolog was retained versus those where it was not. The analyses of nucleotide diversity were based on an additional five male and five female great reed warblers sequenced using short-read technology. The recombination suppression analysis was based on gametologous (Z and W) gene pairs from the great reed warbler, in addition to orthologous Z-linked gene sequences from five

additional Sylvioidea species (generated from one short-read sequenced male of each species) and orthologous genes from non-Sylvioidea species. These were downloaded from the public repository Ensembl (Cunningham et al. 2019). The results from a linkage map analysis of a large pedigree of great reed warblers (Ponnikas et al. 2020) were used to order Z-linked scaffolds into a complete neo-Z chromosome.

Paper VII

In **paper VII**, I studied the process of W chromosome degeneration by exploring the relationship between sex chromosome-specific (Z and W) gene sequence evolution and sex-specific gene expression levels in one Sylvioidea species, the common whitethroat (*Sylvia communis*). Z and W gametolog sequences were extracted based on whole-genome sequence data from two male and two female common whitethroats. I then calculated substitution rates and differences in GC-content between these Z and W gametolog gene pairs, and between each gametolog sequence and orthologous genes in the zebra finch genome. Lastly, I correlated the substitution rates to results from a differential gene expression analysis based on mRNA data from six male and six female common whitethroats.

Paper VI and VIII

Towards the end of my PhD research, I had discovered that at least five independent autosome-sex chromosome fusions had occurred within the Sylvioidea superfamily. Each of these independent origins of neo-sex chromosomes, and the availability of sequence data from many different Sylvioidea species, led me to look more broadly at certain aspects of sex chromosome evolution.

In **paper VI**, I was interested in whether some of the discovered fusions were associated with full sex-linkage of the ancestral pseudoautosomal region (PAR), which is small (~0.5 Mb) and conserved among songbirds. I also wanted to study whether full sex-linkage of this ancestral PAR was associated with fusions to the PAR-end of the sex chromosome. To do this, I utilized published, high-quality, reference genomes as well as the same short-read whole genome data as in **papers II-IV**. First, the short-read sequence data were used to phase out sex chromosome-specific gene sequences of genes localized on the ancestral PAR (as identified in other songbirds). Next, I constructed phylogenetic trees and calculated between-sex pairwise distances between tips. I evaluated if the ancestral PAR genes had evolved full sex-linkage in the Sylvioidea species by having either large between-sex distances or grouped according to sex chromosome copies (Z and W), rather than by species. I then went on to study the chromosomal structure of the Sylvioidea fusion events by searching for scaffolds in the high-quality reference genomes indicative of either: fusion or fission sites, or novel PAR boundaries. Lastly, I studied relationships between female-to-male sequence divergence (phylogenetic

tree distances) and position on the ancestral PAR, according to the gene order of the zebra finch.

In **paper VIII**, I studied the rate of W gene degeneration across species where more than one autosome-sex chromosome fusion had occurred. This was done using short-read genome data from a male and female of eight different Sylvioidea species, which were aligned to the zebra finch genome. Based on (1) genome coverage, (2) heterozygosity and (3) presence or absence of loss-of-function mutations, I categorized all W-linked gametologs in sex-linked genome regions as either being retained, lost or uncertain. I compared relative frequencies of these gene categories to estimate the level of W chromosome degeneration on each sex-linked region. I then evaluated when W-linked gametologs disappeared across a phylogeny, to study the rate of W chromosome degeneration through time. Lastly, I tested the expectation from theory that retained W-linked genes should be more dosage sensitive and evolutionary constrained than those which have disappeared.

Results and discussion

Paper I: XYZWfinder: a snakemake pipeline for detecting and visualising sex chromosomes using short read data

Sex chromosomes can be identified using different types of DNA and RNA sequencing data, sampling strategies and computational methods (recently reviewed in Palmer et al. 2019). The core idea is that sequencing data generated from females and males are expected to differ for sex chromosomes but not for autosomes. The development of user-friendly and cost-effective methods for identifying sex chromosomes is important, as a deeper understanding of sex chromosome variation can help us better understand crucial processes about their formation and continued evolution. Furthermore, knowledge of the sex chromosome system of one's study species is important for the interpretation of the results of analyses undertaken for other purposes, such as population genetic studies.

Identifying sex-linked genomic regions is a major theme of this thesis. Searching for sex-specific differences in genome coverage and heterozygosity are commonly used methods for identifying sex chromosomes using genomic data (Vicoso et al. 2013; Palmer et al. 2019). These genomic signatures were used, although with small variations between studies, in **papers II-IV**. In **paper I**, this method was developed further into an automated computational workflow, which can be used to identify and visualize sex chromosomes using whole-genome sequencing data. This is the first pipeline that utilizes (unassembled) whole-genome sequencing data for this purpose, and the first dedicated pipeline for finding differences in both genome coverage and heterozygosity (to my knowledge). This method complements other published software that utilizes other kinds of data, such as RAD-seq and RNA expression data (Muyle et al. 2016; Feron et al. 2021). One important component of my pipeline is the option to translate the genome coordinates of the reference genome of the studied species to a second reference genome. This allows direct identification of homologies between sex-linked genomic regions across species, which is a major priority for the research field of sex chromosome evolution. This workflow can also be used to verify the sex of sampled individuals, which may be highly challenging to do in the field with sexually monomorphic species.

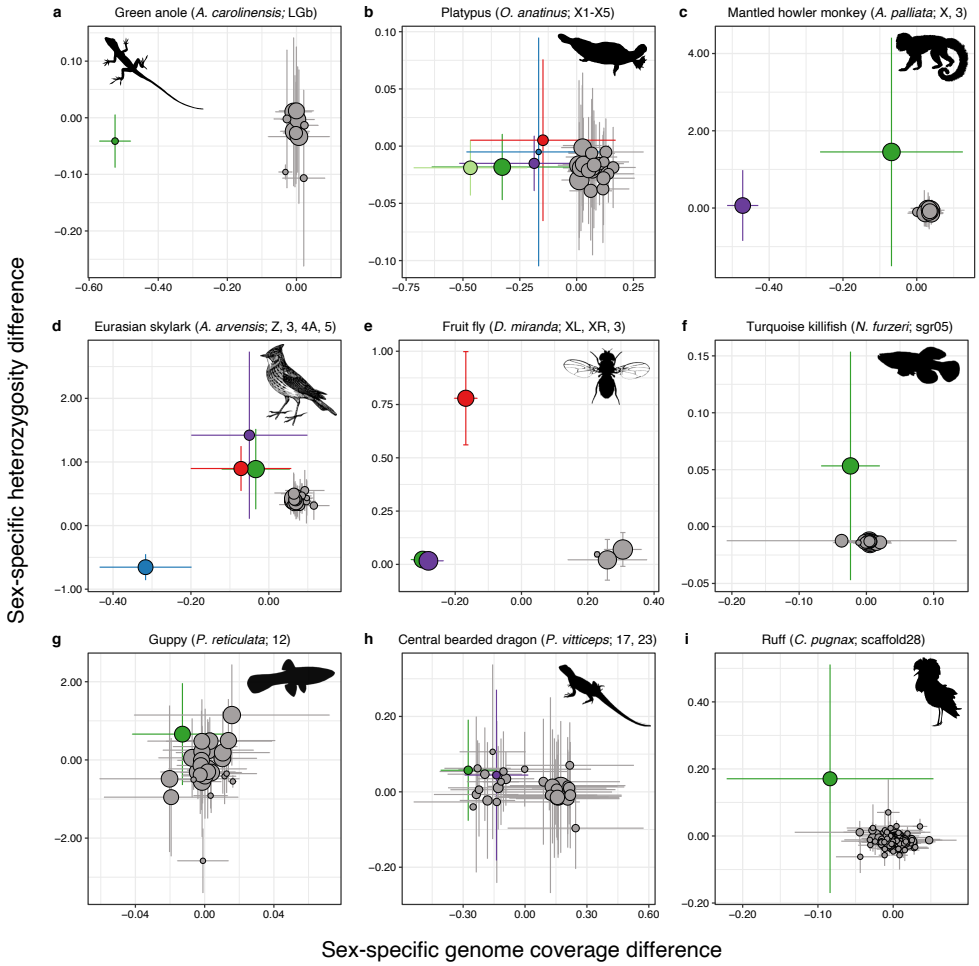


Figure 3. Genome coverage and heterozygosity values (mean ± standard deviation across 1 Mb windows) per chromosome or scaffold for all studied species in paper I. The chromosome/scaffold names in parentheses next to species names are the expected sex chromosomes/inversion polymorphism scaffolds (see paper I for details). The data points associated with these chromosomes/scaffolds are coloured differently from other chromosomes/scaffolds (which are grey). Figure from paper I.

I validated my method by identifying sex chromosomes in species across variable taxonomic groups (mammals, birds, fish, reptiles and insects), and demonstrated its effectiveness in identifying homologies of these sex-linked regions in outgroup reference genomes (Figure 3). I also demonstrated the usefulness of this pipeline in identifying other divergent haplotype blocks by including data from the ruff (*Calidris pugnax*), a bird species where a large inversion polymorphism determines the phenotype of males (Figure 3i). The pipeline was successful in identifying the expected chromosome(s) or scaffold(s) in seven out of nine species (Figure 3a-f,i), but had only limited or no success in the central bearded dragon and guppy (Figure

3g,h). These species have either extremely homomorphic sex chromosomes (guppies), or small sex chromosomes for which homologies to chromosomes in other species has proven challenging (central bearded dragon). In cases such as these, more careful analyses may be needed. This could be done using a larger number of samples, samples from the same family (in which SNP segregation patterns could reveal sex-linkage), or analyses over smaller genome windows (the default for the pipeline is across 100 kb and 1 Mb windows).

In some species, the sex chromosomes are genetically differentiated only across a very small proportion of the sex chromosomes. An extreme example are the sex chromosomes of some pufferfish, where sex is determined by a single nucleotide polymorphism (Kamiya et al 2012). As this pipeline is built to scan across relatively large genome windows (100 kb and 1 Mb), the identification of extremely small sex-linked regions is naturally limited. Another drawback is that, as opposed to some cytogenetic methods, we cannot immediately distinguish the chromosomal structure of fused chromosomes, nor detect formerly autosomal regions that are fused to sex chromosomes but have not evolved recombination suppression. A clear advantage of this method compared to cytogenetics is that it reveals the extent of recombination suppression.

Paper II: Repeated sex chromosome evolution in vertebrates supported by expanded avian sex chromosomes

Two strands of evidence led me to explore the sex chromosome diversity in the Sylvioidea families Alaudidae (larks) and Panuridae (the sister lineage of larks, where the bearded reedling (*Panurus biarmicus*) is the only member). Firstly, a previous study found that some genetic markers that were expected to be autosomal in the Raso lark (*Alauda razae*) appeared to be sex-linked (Brooke et al. 2010). Secondly, a karyotype study of the bimaculated lark (*Melanocorypha bimaculate*) and the horned lark (*Eremophila alpestris*) revealed highly enlarged sex chromosomes compared to other birds (Bulatova 1981). I identified sex-linked genomic regions across four species: Eurasian skylark (*Alauda arvensis*), Raso lark (*A. razae*), horned lark (*E. alpestris*) and bearded reedling (*P. biarmicus*), using whole-genome sequencing data from one individual per sex and species.

As expected, I found that the ancestral bird sex chromosome (chromosome Z) and chromosome 4A (the Sylvioidea-shared fusion) were sex-linked across all studied species. I also found sex-linkage across parts of chromosome 3 in all the studied species, although to different extents, and that chromosome 5 was sex-linked in the two *Alauda* larks but not in the others (Figure 4). The two Alauda larks displayed sex-linkage across almost 200 Mb; corresponding to a fifth of the total genome size and almost three times the size of the sex chromosomes in most birds (chromosome Z; 73 Mb).

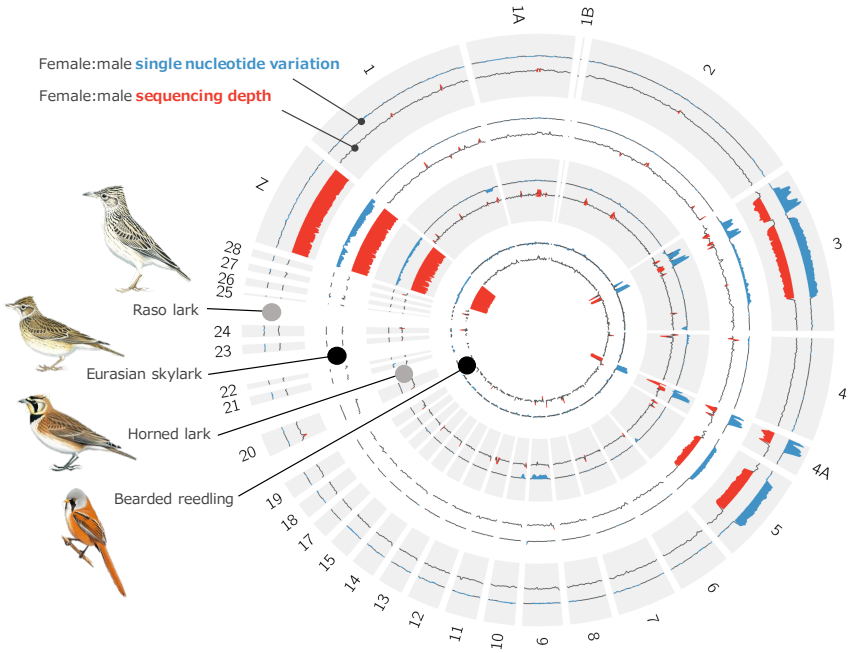


Figure 4. Genome-wide distribution of female-to-male difference in number of private single nucleotide variants (SNVs), and female-to-male coverage ratio in the four study species. The background colours grey and white separate the data for each of the species which are shown in the following order starting from the outer ring: (i) Raso lark, (ii) Eurasian skylark, (iii) horned lark, and (iv) bearded reedling. Within each ring, the outer line shows the average difference in number of private SNVs between females and males across 1 Mbp windows (blue: values greater than 500 or less than −500), and the inner line shows the average female-to-male coverage ratio across 1 Mbp windows (red: values greater than 1.1 or less than 0.9). From paper II. Reprinted with permission from The Royal Society.

Note that for all sex-linked chromosomes, except for the ancestral Z chromosome, the sex-specific genetic differentiation only extends across parts of the chromosomes. This pattern could be caused by either, or both, of the following scenarios: if only parts of these chromosomes (3, 4A and 5) fused to the ancestral sex chromosome, or if the entire chromosomes were fused but recombination suppression has evolved only across parts of these fused regions. In **paper V** and **paper VI,** I explore these hypotheses in more detail.

I divided the sex-linked regions into five evolutionary strata, according to their likely relative timing of recombination suppression. Specifically, I considered sex-linked genome regions shared between species (i.e., regions with shared boundaries between sex-linked and (pseudo)autosomal regions) to have originated before speciation events. I hypothesized that sex-linkage evolved on the different chromosomes in the following order: (i) chromosome Z, (ii) chromosome 4A, (iii-v) different regions of chromosome 3 and (v) chromosome 5, and found that the level of synonymous substitutions between Z and W gametolog gene pairs was

correlated with this relative order (ordered-heterogeneity test: rs=0.99, rsPc = 0.99, k = 6, p < 0.001).

It has been suggested that certain chromosomes are predisposed to evolve into sex chromosomes because of specialized gene content (Graves & Peichel 2010; Bachtrog et al. 2011; Ross et al. 2009; O'Meally et al. 2012). It has also been suggested that some autosomes more readily become part of neo-sex chromosomes because of their gene content, for example sexually antagonistic genes or genes involved in the sex determination pathway (Ross et al. 2009; Kitano et al. 2009; Pala, et al. 2012a; Kitano & Peichel 2011; Yoshida et al. 2014; O'Meally et al. 2012). I show that the chromosome regions involved in the autosome-sex chromosome fusions within Alaudidae and Panuridae have in fact been recruited as sex chromosomes repeatedly across different vertebrate taxa. I also found that chromosome 3 had a significant excess of genes involved in sex-specific functions (binomial test: 33 observed genes, 19.7 expected genes, adjusted p = 0.019). However, to rigorously test hypotheses of predisposition towards sex-linkage, analyses across a wider range of sex chromosome systems are required.

Paper III: Whole-genome analysis across 10 songbird families within Sylvioidea reveals a novel autosome-sex chromosome fusion

In **paper III**, I sought to characterize sex chromosome diversity across a wider taxonomic range, now including species from ten different Sylvioidea families. This revealed yet another autosome-sex chromosome fusion within the group; chromosome 4 was sex-linked in the zitting cisticola (*Cisticola juncidis*), belonging to the Cisticolidae family (Figure 5). The level of genetic differentiation appeared similar between chromosome 4A and chromosome 4, indicating that the Cisticola-specific fusion did not occur extremely recently.

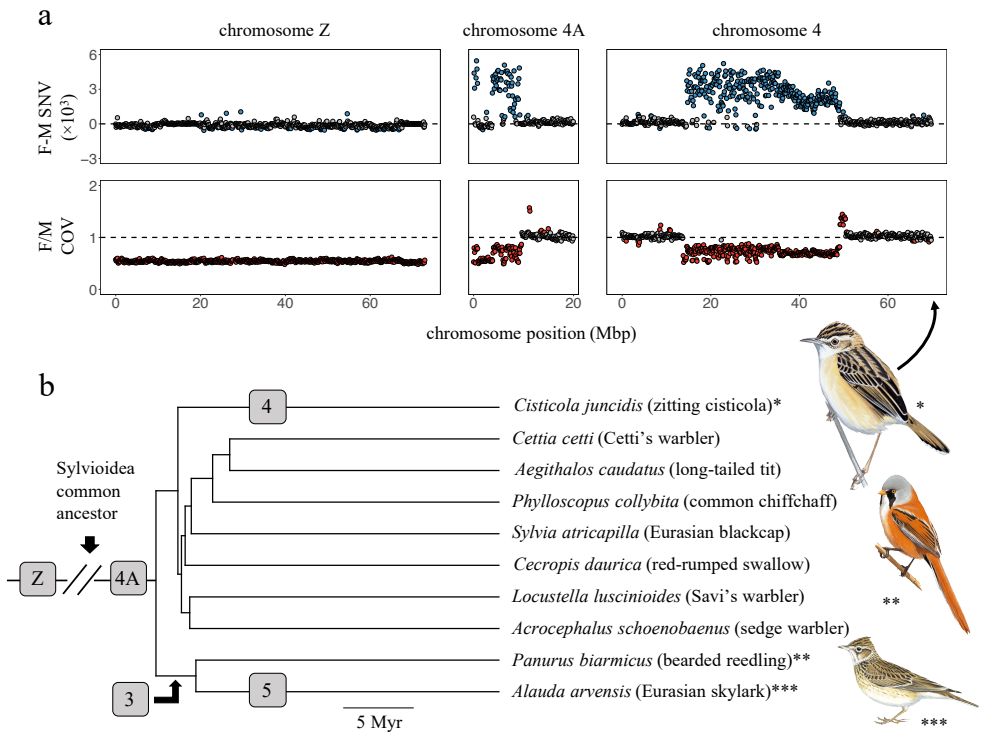


Figure 5. (a) Female-to-male difference in number of SNVs, and read coverage ratio, along the chromosomes showing signatures of sex linkage in *Cisticola juncidis* (i.e., Z, 4A and 4) across 100 kbp windows (SNV outlier values in blue and genome coverage outlier values in red; see electronic supplementary material, Methods for details). (b) A phylogeny of the studied species and inferences of when the autosome–sex chromosome fusions occurred, based on the species in which they have been found. From paper III. Reprinted with permission from The Royal Society.

Paper IV: A novel neo-sex chromosome in *Sylvietta brachyura* (Macrosphenidae) adds to the unprecedented avian sex chromosome diversity among Sylvioidea songbirds

In **paper IV**, I included genome data from species-representatives of three additional Sylvioidea families: common bulbul (*Pycnonotus barbatus*; Pycnonotidae), northern crombec (*Sylvietta brachyura*; Macrosphenidae) and Iraq babbler (*Argya altirostris*; Leiothrichidae). I found yet another autosome-sex chromosome fusion involving chromosome 8 in the northern crombec. I also found that the sex-linked part of chromosome 4A was smaller in this species than in all others studied so far (Figure 6). This implies that recombination suppression has evolved (at least) twice, independently, across the first ~5Mb of chromosome 4A.

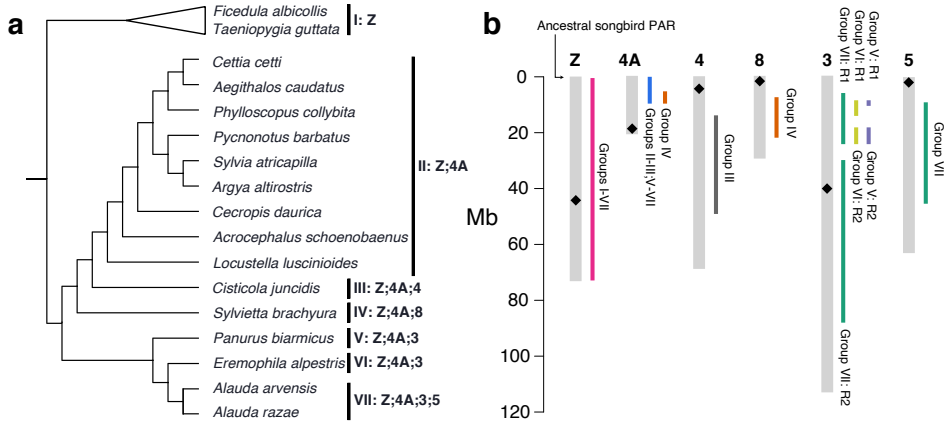


Figure 6. Sex chromosome diversity in Sylvioidea songbirds. Figure from **paper VI**. (a) Phylogenetic tree of some of the studied species in **papers II-IV**, grouped according to shared sex-linked genome regions. Group I are non-Sylvioidea passerines, Groups II-VII belong to Sylvioidea and share the fusion between chromosome Z and 4A. Within Groups II-VII, additional chromosomes have evolved sex-linkage. (b) Sex-linked chromosome regions identified across Groups I-VII. If more than one consecutive genome region was found on one chromosome, they are specified as either R1 or R2 (chromosome 3). Note that the sex-linked region on chromosome 4A is smaller in *Sylvietta brachyura* (Group IV) than in other Sylvioidea species (Groups II-III,V-VII).

We also found that parts of chromosome 25 might be sex-linked in the common bulbul (*Pycnonotus barbatus*). However, the region deviating from the autosomal pattern was small (< 1 Mb) and further study is needed to confirm this finding.

Paper V: Avian neo-sex chromosomes reveal dynamics of recombination suppression and W degeneration

In **paper V**, I wanted to reconstruct the chromosomal structure of the sex chromosomes of one Sylvioidea species: the great reed warbler (*Acrocephalus arundinaceus)*. In this species, chromosome Z ("the ancestral sex chromosome region") and chromosome 4A ("the added sex chromosome region") are sex-linked. We constructed a high-quality annotated reference genome from a female (ZW) great reed warbler (N50: 21.4 Mb), and used whole-genome data from five male and five female great reed warblers to identify all Z-linked (n = 22, mean length 22 Mb) and W-linked scaffolds (n = 50, mean length 0.6 Mb) in the genome.

The Z-linked scaffolds were ordered along the Z chromosome using data from a linkage map analysis of a pedigree of great reed warblers (Ponnikas et al. 2020). I found the added (4A) sex chromosome region to be physically attached to both the Z and W chromosome copy in this species (Figure 7). The same result was found in another Sylvioidea species (Leroy et al. 2019), but with a different W-4A fusion site. I found that the fusion between the ancestral Z chromosome and chromosome 4A occurred at the distal end of the ancestral Z chromosome, concordant with results from Leroy et al. (2019), and that the gene order had remained collinear with its ancestral state (Figure 7a,c). I found many links between the ancestral and added W

chromosome, indicative of complex rearrangements since the fusion event (Figure 7b,c). The linkage map analysis further confirmed that the scaffold containing the distal part of chromosome 4A (9.6-20 Mb; Figure 7a) segregates as an independent entity.

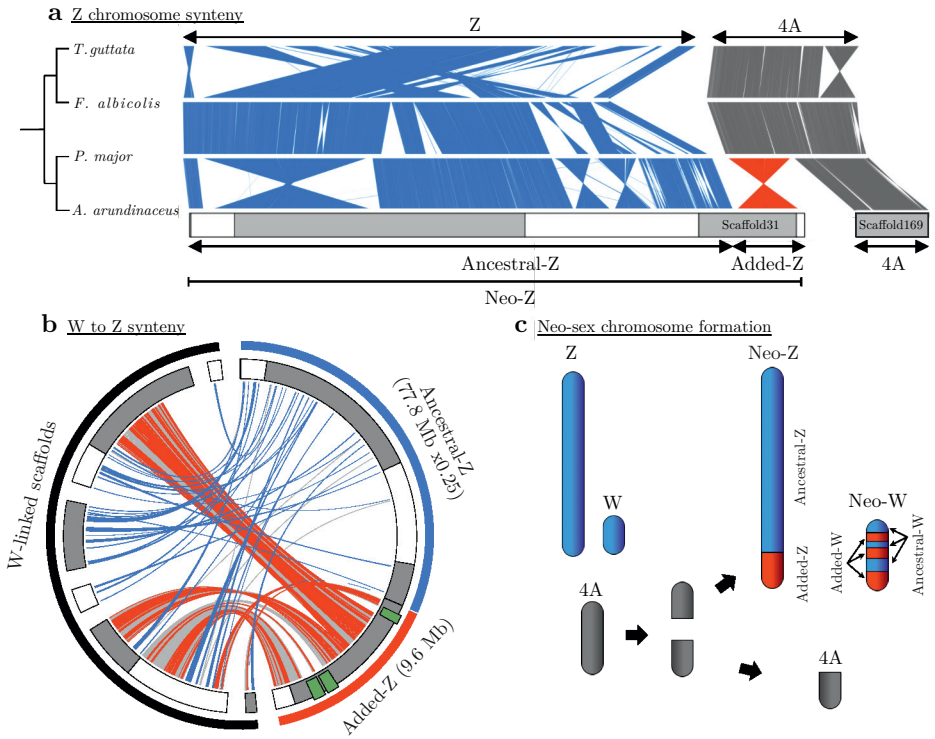


Figure 7. From Paper V. Structure of the great reed warbler Z and W neo-sex chromosomes. (a) Chromosome alignments of chromosome Z (blue) and 4A (grey) in four species of passerines, with the part of 4A representing the added-Z region in the great reed warbler indicated (red). (b) Syntenic regions between the great reed warbler W-linked and Z-linked (grey and white alternating) scaffolds. The Z-linked scaffolds belonging to the ancestral sex chromosome region (blue) are scaled to 25% of their true size for illustrative purposes. The grey links show syntenic information on a genomic level, while chromosomal positions of gametologous (ZW) gene pairs are shown as blue links for ancestral sex chromosome genes and red links for added sex chromosome genes. Note that four W-linked scaffolds have genes with orthologs on both the ancestral and added sex chromosome regions, strongly suggesting that the ancestral-W and added-W are physically connected. (c) A graphic representation of the fusion event forming the Sylvioidea neo-sex chromosome system.

How and why recombination suppression evolves are key questions in sex chromosome research. Here, we inferred the timing and spatial progression of recombination suppression across this neo-sex chromosome using a multi-species approach. Specifically, we identified and manually curated gametologous (Z and W) gene pairs in the great reed warbler annotation, and analysed the resulting gene sequences with Z-linked or autosomal gene sequences from other Sylvioidea species (all Z-linked) and outgroups (either Z-linked or autosomal). We found no clear pattern of evolutionary strata across the added sex chromosome region. Instead, we

found a mosaic pattern of recombination suppression timings, where some genes evidently continued to recombine long after the fusion event. Continued recombination long after the fusion event is further supported by evidence of retained recombination across parts of chromosome 4A in *Sylvietta brachyura* (**paper IV**). Note that the androgen receptor (*ar*) gene, which is involved in development of sexual characteristics, was not among the first genes to stop recombining, according to our analysis. This means that selection for linkage between this gene and the sex determining gene was probably not a driver of this fusion event.

Following recombination suppression, the W chromosome is expected to accumulate repeat elements due to inefficient selection against deleterious mutations. In line with this hypothesis, both the ancestral (64% of the total DNA sequence) and added (35%) W chromosome region were heavily enriched for repeat elements compared to the genome-wide average (19%). Due to this repeat accumulation, the combined length of the scaffolds from the added W chromosome region was longer than scaffolds from the added Z chromosome region. Furthermore, there was a distinct peak in repeat-richness at the fusion site between the ancestral and added sex chromosome region on the Z chromosome copy. The nucleotide divergence levels on the W chromosomes were also extremely low, both in the ancestral (0.0002) and added (0.0001) region, compared to the genome average (0.0030).

I compared gene characteristics of genes where the W-linked gametolog had been retained versus those where the W-linked gametolog had been lost. I found that genes in which the W-linked gametolog was retained were, on average, more dosage sensitive than those where the W-linked gametolog was lost. This difference was statistically significant on the ancestral sex chromosome region (Mann-Whitney U test: $U = 5308$, $p < 0.001$) but not on the added sex chromosome region (Mann-Whitney U test: $U = 335$, $p = 0.11$). I also found that genes with retained W-linked gametologs evolved slower than others (as estimated by dN/dS values between Z-linked gametologs and gene sequences from the zebra finch), both on the ancestral (Mann-Whitney U test: $U = 6460$, $p < 0.001$) and added (Mann-Whitney U test: $U = 1161$, $p < 0.001$) sex chromosome region. These results are in line with those from mammals (Bellott et al. 2014) and other (non-Sylvioidea) birds (Bellott et al. 2017).

Lastly, I correlated dN/dS ratios between the great reed warbler gametologs to those between different pairs of species (for example, great reed warbler Z-to-W dN/dS correlated against dN/dS values between human and mouse) for the same genes. I found that the rate of evolution of genes on the sex chromosomes in the great reed warbler was highly correlated with gene evolution rates in species where these genes are autosomal. This suggests that the relative selection pressures on autosomal genes are retained after becoming sex-linked.

Paper VI: Evolutionary dynamics of enlarged sex chromosomes and novel pseudoautosomal regions in Sylvioidea songbirds

In **paper VI**, I asked how the ancestral pseudoautosomal region ("the songbird PAR"), which has been shown to be conserved among songbirds (Smeds et al 2014; Zhou et al. 2014; Singhal et al. 2015), had been affected by the fusions in the Sylvioidea group. Specifically, I wanted to know if any of the fusion events were associated with recombination suppression of the songbird PAR.

I found that the songbird PAR, a ~0.5 Mb region at the beginning of the Z chromosome that contains ~20 genes, was retained as pseudoautosomal in all species with fusions only involving chromosome 4A, and in species with fusions involving chromosomes 4A and 3. In three genera in three families, however, the songbird PAR had evolved full sex-linkage (as shown by pronounced between-sex differentiation; Figure 8). These genera were all associated with lineage-specific fusions: *Alauda* (chromosome 5), *Cisticola* (chromosome 4) and *Sylvietta* (chromosome 8).

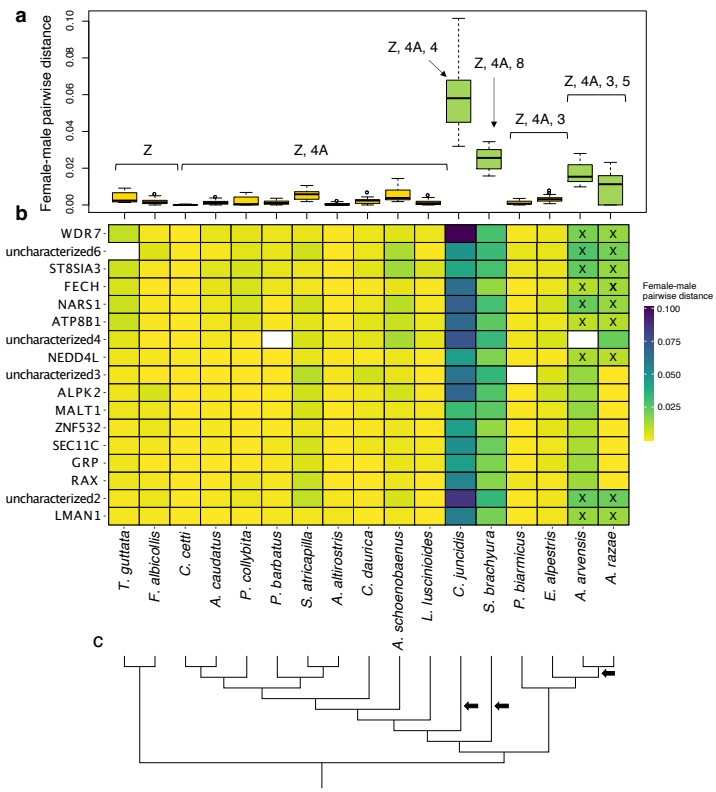


Figure 8. Assessment of recombination suppression of the ancestral PAR in Sylvioidea species (and outgroups). (a) Boxplots showing pairwise distances between female and male sequences of ancestral PAR genes, calculated from phylogenetic tree branch lengths. High female-to-male sequence divergence-values indicate highly diverged Z and W gene copies as a consequence of recombination suppression. The combination of chromosomes with known sex-linkage for each species is indicated above the boxplots. (b) Pairwise distance between female and male sequences

plotted for each gene separately. The crosses inside tiles represent genes where the sequences from Alauda arvensis and A. razae cluster by sex rather than species, suggesting that these genes ceased to recombine before the split of the species. (c) Cladogram showing the phylogenetic relationships between the studied species. Black arrows point to branches leading up to species where ancestral PAR genes have evolved recombination suppression. From paper VI.

Based on these patterns, I hypothesized that the two fusions not associated with sex-linkage of the songbird PAR (chromosome 3 and 4A) fused to the non-PAR end of the Z chromosome while the other three fusions (chromosomes 4, 5 and 8) fused to the PAR-end. To test this hypothesis, I studied the Z chromosome structure of Sylvioidea species with more fusions than the one common to the entire group, using published reference genomes of higher contiguity than those generated based on the short-read genome data (see General methodology). As all of these published genomes were based on data from male individuals (ZZ), I cannot conclude anything about the chromosome structure of the W chromosome in these species. I found support for a fusion site of chromosome 5 in the PAR-end of the Z chromosome in *Alauda arvensis* where chromosome 5 is sex-linked, but unexpectedly also in *Eremophila alpestris* where chromosome 5 shows no signs of genetic differentiation between sexes (Figure 6; Figure 9). I also found weak evidence to support that chromosome 3 fused to the distal end of the Sylvioidea neo-sex chromosome (holding chromosome 4A; Figure 9).

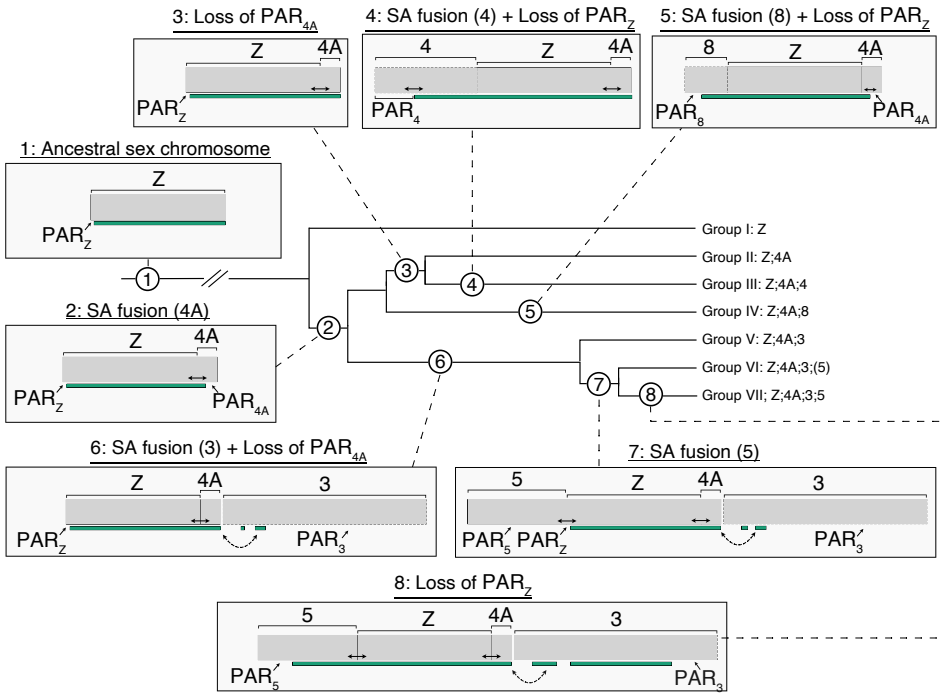


Figure 9. Schematic of the hypothesized structure of neo-Z chromosomes in Sylvioidea songbirds. Figure from paper VI. The studied species were grouped (I-V) according to those sharing identical sex-linked genome regions. The species associated with each group can be found in Figure 6.

Theory predicts that Z-to-W sequence divergence in the PAR should increase with proximity to the PAR boundary. We found significant correlations between female-to-male sequence divergence and proximity to the PAR boundary in six of the eleven Sylvioidea-species with retained ancestral PAR (Spearman's test: rho ranging between 0.532 and 0.706, n = 16-17 genes, p ranging between 0.002 and 0.036), and in one of the two outgroup species (zebra finch; Spearman's test: rho = 0.832, n = 16 genes, $p < 0.001$).

The PAR has remained stable in songbirds for many millions of years. One explanation for why is that the PAR is needed for proper meiosis. Alternatively, recombination suppression could be selected against, for example if the genes residing in the songbird PAR are highly dosage sensitive. Yet another explanation could simply be that neither of the genes are sexually antagonistic, meaning that there may be no selection for recombination suppression.

Paper VII: Insights into Avian incomplete dosage compensation: sex-biased gene expression coevolves with sex chromosome degeneration in the common whitethroat

In this paper, I studied relationships between Z and W chromosome gene sequence evolution (substitution rates and GC-content) and gene expression differences in the common whitethroat (*Sylvia communis*). Like the great reed warbler (**paper V**), this species only has the "Sylvioidea-shared" fusion between chromosome Z and parts of chromosome 4A.

First, I show that the Z and W gene sequences have diverged significantly from each other, both on the ancestral and added sex chromosome region (dS: 0.01-0.83). On the added sex chromosome region, I could infer which chromosome copy (Z or W) that had diverged most since becoming sex-linked by comparing dN and dS differences between each gene sequence to the orthologous gene sequence in the zebra finch. This analysis revealed that while the Z and W sequence had diverged similarly in terms of synonymous substitutions (paired t-test: t = 1.18, p = 0.24), the W sequences had diverged significantly more in terms of non-synonymous substitutions (paired t-test: t = 6.87, p < 0.001). Weakened selection against amino acid-changing (dN) mutations is an expected consequence of the lack of recombination on W chromosomes, suggesting that the added sex chromosome region is undergoing W chromosome degeneration. I also found overall low dN/dS values on the added sex chromosome region (all but two genes had dN/dS values < 0.2). This suggests that these genes are constrained by purifying selection. However, I did not perform tests of selection types to confirm this.

The two chromosomes with highest gene expression differences between sexes were, unsurprisingly, chromosome Z and 4A. On both of these chromosomes the gene expression level was significantly lower in females than in males, but

especially so on the ancestral sex chromosome. I found that the female-to-male gene expression ratio was significantly negatively correlated with the level of dN divergence between gametologous gene pairs on the added sex chromosome region, but not on the ancestral sex chromosome region. The expression ratio also correlated negatively with the level of dN/dS on the ancestral sex chromosome region but not the added one. The expression ratio did not correlate with dS values on either sex chromosome region.

Based on these results, it is likely that gene expression of W-linked genes on the added region (chromosome 4A) is being adaptively downregulated in females in response to accumulations of deleterious mutations on the W chromosome (measured as the level of functional divergence, dN).

Paper VIII: The rate of W chromosome degeneration across multiple autosome-sex chromosome fusions in birds

In **paper VIII**, I studied the extent of W chromosome degeneration across eight different Sylvioidea species in which multiple autosome-sex chromosome fusions have occurred. I determined, based on genome coverage, heterozygosity and loss-of-function mutations, whether the W-linked gametolog was retained or not for all genes occurring in a sex-linked region in any of the studied species.

I found that the ancestral sex chromosome was highly degenerated; only 5-8% of the original genes from chromosome Z (chrZ) remained on the W chromosome as W-linked gametologs. This is similar to values from other studies of W chromosome degeneration in birds (Smeds et al. 2015; Bellott et al. 2017). The more recently sex-linked regions, formed through autosome-sex chromosome fusions, displayed a range of degeneration levels; from intermediate levels (55%) to almost no W degeneration (98%). I inferred the likely timing of W gene loss across a phylogeny of the studied species and found that most W-linked gametologs on the ancestral sex chromosome (chrZ) disappeared before the Sylvioidea evolved (Figure 10a), but that only a small proportion (6%) of the W-linked gametologs on the part of chromosome 4A inferred to have become sex-linked first (chr4A-a) disappeared in a common ancestor (Figure 10a). Furthermore, I found that dosage sensitive and evolutionary constrained genes had been retained to a higher degree on the W chromosome of these species than others.

The different evolutionary processes underlying Y/W chromosome degeneration (such as Mullers ratchet, background selection and genetic hitchhiking) are expected to vary over time and with the number of functional Y/W genes (Bachtrog 2008). Datasets like ours could therefore help to untangle the relative importance of these processes. However, we need a better understanding of how and when recombination was suppressed across each stratum to be able to study this process

in detail. Furthermore, denser taxonomic sampling would also increase the preciseness of these analyses.

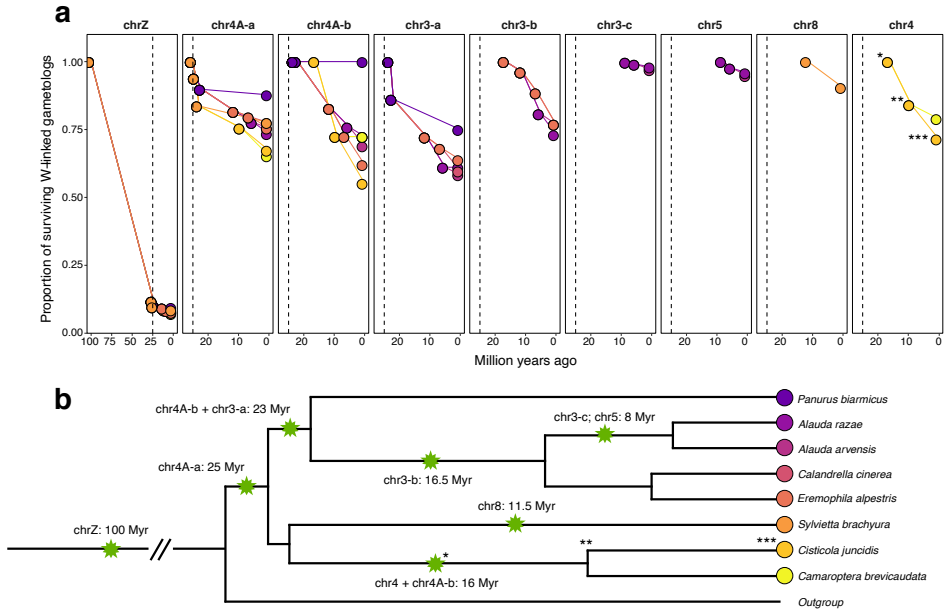


Figure 10. (a) Inferences of the rate of W gene loss in eight Sylvioidea species across nine evolutionary strata. The colours of the data points in (a) follow those in the phylogenetic tree in (b). (b) Phylogenetic tree showing the estimated timing of recombination suppression of each strata, which correspond to the first data point on the x-axis in (a). From paper VIII.

Conclusions and future studies

Here, I address the specific questions that were asked in my Thesis aims (Table 1).

How can sex chromosomes be identified in a user-friendly way using minimal genomic resources?

The combined use of genome coverage and heterozygosity is a powerful approach to identifying sex-linked genomic regions (Vicoso et al. 2013; Palmer et al. 2019). I show that our computational workflow (**paper I**) can successfully utilize these genomic signatures in whole-genome sequencing data using only one sample from each sex, to identify sex-linked genome regions in a broad range of taxa with sex chromosome systems of varying sizes and degrees of differentiation.

What is the extent of sex chromosome diversity within Sylvioidea songbirds?

Within Sylvioidea songbirds, we have now discovered five separate autosome-sex chromosome fusions (Pala et al. 2012a; Pala et al. 2012b; Sigeman et al. 2019 (**paper II**); Leroy et al. 2019; Dierickx et al. 2020; Sigeman et al. 2020 (**paper III**); **paper IV**). So far, species from 14 of the ~22 families in Sylvioidea have been studied (Fregin et al. 2012; Oliveros et al. 2019), meaning that there may be additional sex chromosome diversity yet to be discovered (and this could of course apply to unstudied species within these 14 families). Regardless of any undetected variation, the existence of five separate fusion events within Sylvioidea is exceptional for birds, where the sex chromosomes have previously been found to be extremely stable, even between distantly related species (Nanda et al. 2008).

The large discrepancy between Sylvioidea birds compared to others is likely to be influenced by study-bias. Although there has been a massive burst of bird genomes being sequenced during the last decade (e.g., Zhou et al. 2014; Feng et al. 2020), genomic data from both sexes is still not available for more than a limited number of species. In my PhD, we generated a large amount of such data that was only focused on the Sylvioidea clade. To truly estimate the difference in sex chromosome variability between Sylvioidea birds and others, more systematic testing is needed. Nevertheless, it is clear that the Sylvioidea system is highly useful for testing theories of sex chromosome evolution, owing to its already established variability.

In line with the argument that additional avian sex chromosome diversity is to be expected are the recent reports of two additional autosome-sex chromosome fusions in non-Sylvioidea birds; one in the eastern yellow robin (*Eopsaltria australis*;

involving chromosomes Z and 1A; Gan et al. 2019), and another in the smooth-billed ani (*Crotophaga ani*; between chromosome Z and 17; Kretschmer et al. 2021). Future studies will hopefully reveal if there is additional diversity in close relatives of these two species. A multiple sex chromosome system was also recently found in the Adélie penguin (*Pygoscelis adeliae*; Gunski et al. 2018). Whether this was the result of a fusion with an autosome or due to a fission event is currently not clear. Preliminary analyses (by me) of published whole-genome data using the pipeline from **paper I** suggest that there are no sex-linked genomic regions in this species that are not homologous to the Z chromosome. This suggests that the multiple sex chromosome system in this species was caused by a fission.

What is the chromosomal structure of the Sylvioidea sex chromosome fusions?

We show that the fusion between chromosome Z and chromosome 4A occurred in the distal end of chromosome Z (**paper V, VI**) and that chromosome 5 is fused to the opposite end of the Z chromosome (**paper VI**). The ancestral W appears to be physically attached to chromosome 4A in the great reed warbler (**paper V**), yet the exact structure of this fusion is unclear due to the level of fragmentation among the W-scaffolds in our reference genome. Future studies will hopefully reveal more details about the chromosomal structure of the Sylvioidea sex chromosomes. One possibility is to create high-quality *de novo* reference genomes from female individuals of additional species, in combination with cytogenetic work.

How was the ancestral pseudoautosomal region (PAR) affected by the fusions?

The results from **paper VI** show a clear association between certain autosome-sex chromosome fusions and full sex-linkage of the ancestral PAR. I find it likely that this is a result of relaxed selection against recombination suppression of the ancestral PAR (which is needed for proper segregation of chromosomes during meiosis), as the new fusions created extended PARs in the fused regions.

How does recombination suppression evolve along the added part of a neo-sex chromosome?

The results from **paper V** show that recombination suppression spread across the added (4A) region in a gradual and mosaic manner over the course of several million years. This is further confirmed by results in **paper IV**, showing that parts of this added region are still recombining in the northern crombec (*Sylvietta brachyura*). The results from **paper VI** show additional variation in the extent of recombination suppression following a fusion event. Specifically, the fusion with chromosome 5 was followed by recombination suppression in some lineages (Eurasian skylark, *Alauda arvensis*, and Raso lark, *A. razae*) but not in others (horned lark, *Eremophila alpestris*). This variation has created excellent opportunities to study when and why recombination suppression spread on sex chromosomes.

How does sex-linkage affect W chromosome sequences in terms of repeat accumulation, gene loss and gene expression?

We found that the recombination suppression is followed by W chromosome degeneration in the form of gene loss (**paper V, VIII**) and the accumulation of repeats (**paper V**). In line with theory, this gene loss is non-random: evolutionary constrained and dosage sensitive genes lose functionality to a lesser degree than other genes (**paper V, VIII**). Our results also suggest that female gene expression is lowered in response to W chromosome mutations (**paper VII**). This is also in line with theoretical expectations of adaptive silencing of malfunctioning genes during W chromosome degeneration (Lenormand et al. 2020).

What's next?

In this thesis, I investigated the evolution of neo-sex chromosomes across a large part of the Sylvioidea phylogeny, from a variety of angles. A major goal was to analyse data across as many Sylvioidea families as possible. The sampling strategy of targeting one male and one female per species was ideal for this goal, but this is not enough to study some aspects of sex chromosome evolution. For example, data from one female and one male is not sufficient for examining genomic signatures of sexual antagonism, for determining the precise PAR boundary, or for evaluating if there is intraspecific PAR variation, as is observed in some mammals (Cotter et al. 2016) and plants (Qiu et al. 2016). Furthermore, we only produced long-range sequencing data from one species, the great reed warbler (**paper V**). An obvious next step is to generate data from multiple individuals per species, and generate high-quality reference genomes from females of additional species.

By also adding ecological data to the genomic data, we can begin to answer a whole other range of questions related to the evolution of sex chromosomes. For example, according to Haldane's rule, the heterogametic sex is more likely to suffer from hybrid sterility than the homogametic sex (Haldane 1922). Furthermore, chromosomal rearrangements are known to promote reproductive isolation (e.g., Hou et al. 2014). In line with this, neo-sex chromosomes have been connected to reproductive isolation in several species (Kitano et al. 2009; Smith et al. 2016; Bracewell et al. 2017). Does this apply to the Sylvioidea fusions as well? And what are the fitness consequences of the continuous W chromosome degeneration observed across Sylvioidea species (**papers V, VII, VIII**)? Finally, are there associations between sex chromosome diversity and sexual dimorphism? These are questions for the future, for which the work in this thesis has laid a foundation.

References

Abbott JK, Nordén AK, Hansson B. 2017. Sex chromosome evolution: historical insights and future perspectives. *Proceedings of the Royal Society B*, **284**, 20162806.

Adolfi MC, Nakajima RT, Nóbrega RH, Schartl M. 2019. Intersex, hermaphroditism, and gonadal plasticity in vertebrates: evolution of the Müllerian duct and Amh/Amhr2 signaling. *Annual Review of Animal Biosciences*, **7**, 149–172.

Alström P, Hooper DM, Liu Y, Olsson U, Mohan D, Gelang M, et al. 2014. Discovery of a relict lineage and monotypic family of passerine birds. *Biology Letters*, **10**, 20131067.

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

Bachtrog D. 2006. A dynamic view of sex chromosome evolution. *Current Opinion in Genetics & Development*, **16**, 578–585.

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

Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nature Reviews Genetics*, **14**, 113–124.

Bachtrog D, Kirkpatrick M, Mank JE, McDaniel SF, Pires JC, Rice W, et al. 2011. Are all sex chromosomes created equal? *Trends in Genetics*, **27**, 350–357.

Bachtrog D, Mank JE, Peichel CL, Kirkpatrick M, Otto SP, Ashman TL, et al. 2014. Sex determination: why so many ways of doing it? *PLoS Biology*, **12**, e1001899.

Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho TJ, et al. 2014. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature*, **508**, 494–499.

Bellott DW, Skaletsky H, Cho TJ, Brown L, Locke D, Chen N, 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, et al. 2010. Convergent evolution of chicken Z and human X chromosomes by expansion and gene acquisition. *Nature* **466**, 612–616.

Bergero R & Charlesworth D. 2009. The evolution of restricted recombination in sex chromosomes. *Trends in Ecology & Evolution*, **24**, 94–102.

Beukeboom L & Perrin N. 2014. *The Evolution of Sex Determination*. Oxford University Press, Oxford, United Kingdom.

Blackmon H, Ross L, Bachtrog D. 2016. Sex determination, sex chromosomes, and karyotype evolution in insects. *Journal of Heredity*, **108**, 78–93.

Bowen ST. 1965. The genetics of *Artemia salina*. V. Crossing over between the X and Y chromosomes. *Genetics*, **52**, 695–710.

Bracewell RR, Bentz BJ, Sullivan BT, Good JM. 2017. Rapid neo-sex chromosome evolution and incipient speciation in a major forest pest. *Nature Communications*, **8**, 1593.

Branco S, Badouin H, de la Vega RCR, Gouzy J, Carpentier F, Aguileta G, et al. 2017. Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *Proceedings of the National Academy of Sciences*, **114**, 7067–7072.

Brooke MdeL, Welbergen JA, Mainwaring MC, Van Der Velde M, Harts AMF, Komdeur J, et al. 2010. Widespread translocation from autosomes to sex chromosomes preserves genetic variability in an endangered lark. *Journal of Molecular Evolution*, **70**, 242–246.

Bulatova NS. 1981. A comparative karyological study of passerine bird. *Academiae Scientiarum Bohemoslovacae Brno*, **15**, 1–44.

Bull JJ. 1983. *Evolution of Sex Determining Mechanisms*. The Benjamin/Cummings Publishing Company.

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 & Charlesworth D. 2000. The degeneration of Y chromosomes. *Philosophical Transactions of the Royal Society B*, **355**, 1563–1572.

Charlesworth B, Charlesworth D. 2020. Evolution: a new idea about the degeneration of Y and W chromosomes. *Current Biology*, **30**, R871–R873.

Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, **130**, 113–146.

Charlesworth B & Wall JD. 1999. Inbreeding, heterozygote advantage and the evolution of neo-X and neo-Y sex chromosomes. *Proceedings of the Royal Society B*, **266**, 51–56.

Charlesworth D. 2018. The guppy sex chromosome system and the sexually antagonistic polymorphism hypothesis for Y chromosome recombination suppression. *Genes*, **9**, 264.

Charlesworth D. 2021. When and how do sex-linked regions become sex chromosomes? *Evolution*, 10.1111/evo.14196.

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.

Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A, Waters PD, et al. 2014. Origins and functional evolution of Y chromosomes across mammals. *Nature*, **508**, 488–493.

Cotter DJ, Brotman SM, Wilson Sayres MA. 2016. Genetic diversity on the human X chromosome does not support a strict pseudoautosomal boundary. *Genetics*, **203**, 485–492.

Cunningham F, Achuthan P, Akanni W, Allen J, Amode MR, Armean IM, et al. 2019. Ensembl 2019. *Nucleic Acids Research*, **47**, D745–D751.

Dierickx EG, Sin SYW, van Veelen HPJ, Brooke MdeL, Liu Y, Edwards SV, et al. 2020. Genetic diversity, demographic history and neo-sex chromosomes in the critically endangered Raso lark. *Proceedings of the Royal Society B*, **287**, 20192613.

Disteche CM. 2012. Dosage compensation of the sex chromosomes. *Annual Review of Genetics*, **46**, 537–560.

Ezaz T, Srikulnath K, Graves JAM. 2016. Origin of amniote sex chromosomes: an ancestral super-sex chromosome, or common requirements? *Journal of Heredity*, **108**, 94–105.

Feng S, Stiller J, Deng Y, Armstrong J, Fang Q, Reever AH et al. 2020. Dense sampling of bird diversity increases power of comparative genomics. *Nature*, **587**, 252–257

Feron R, Pan Q, Wen M, Imarazene B, Jouanno E, Anderson J, et al. 2021. RADSex: a computational workflow to study sex determination using Restriction Site-Associated DNA Sequencing data. *Molecular Ecology Resources*, 10.1111/1755-0998.13360.

Fisher RA. 1931. The evolution of dominance. *Biological Rev*iews, **6**, 345–368.

Fregin S, Haase M, Olsson U, Alström P. 2012. New insights into family relationships within the avian superfamily Sylvioidea (Passeriformes) based on seven molecular markers. *BMC Evolutionary Biology*, **12**, 157.

Fry JD. 2010, The genomic location of sexually antagonistic variation: some cautionary comments. *Evolution*, **64**, 1510–1516.

Furman BLS, Metzger DCH, Darolti I, Wright AE, Sandkam BA, Almeida P, et al. 2020. Sex chromosome evolution: so many exceptions to the rules. *Genome Biology and Evolution*, **12**, 750–763.

Gamble T, Coryell J, Ezaz T, Lynch J, Scantlebury DP, Zarkower D. 2015. Restriction Site-Associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems. *Molecular Biology and Evolution*, **32**, 1296–1309.

Gammerdinger WJ & Kocher TD. 2018. Unusual diversity of sex chromosomes in african cichlid fishes. *Genes*, **9**, 480.

Gan HM, Falk S, Morales HE, Austin CM, Sunnucks P, Pavlova A. 2019. Genomic evidence of neo-sex chromosomes in the eastern yellow robin. *GigaScience*, **8**, giz111.

Gorelick R. 2003. Evolution of dioecy and sex chromosomes via methylation driving Muller's ratchet. *Biological Journal of the Linnean Society*, **80**, 353–368.

Graves JAM. 2006. Sex chromosome specialization and degeneration in mammals. *Cell*, **124**, 901–914.

Graves JAM. 2016. Evolution of vertebrate sex chromosomes and dosage compensation. *Nature Reviews Genetics*, **17**, 33–46.

Graves JAM & Peichel CL. 2010. Are homologies in vertebrate sex determination due to shared ancestry or to limited options? *Genome Biology*, **11**, 205.

Griffin DK, Robertson LBW, Tempest HG, Skinner BM. 2007. The evolution of the avian genome as revealed by comparative molecular cytogenetics. *Cytogenetic and Genome Research*, **117**, 64–77.

Grützner F, Rens W, Tsend-Ayush E, et al. 2004. In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature*, **432**, 913–917.

Gu L & Walters JR. 2017. Evolution of sex chromosome dosage compensation in animals: a beautiful theory, undermined by facts and bedeviled by details. *Genome Biology and Evolution*, **9**, 2461–2476.

Gunski RJ, Cañedo AD, del Valle Garnero A, Ledesma MA, Coria N, Montalti D, Degrandi TM. 2017. Multiple sex chromosome system in penguins (*Pygoscelis*, Spheniscidae). *Comparative Cytogenetics*, **11**, 541–552.

Haldane JBS. 1922. Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, **12**, 101–109.

Handley LJ, Ceplitis H, Ellegren H. 2004. Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. *Genetics*, **167**, 367–76.

Henking H. 1891. Uber spermatogenese und deren beziehung zur entwicklung bei *Pyrrhocoris apterus*. *L. Z. Wiss. Zool*. **51**, 685–736.

Hou J, Friedrich A, de Montigny J, Schacherer J. 2014. Chromosomal rearrangements as a major mechanism in the onset of reproductive isolation in *Saccharomyces cerevisiae*. *Current Biology*, **24**, 1153–1159.

Iannucci A, Altmanová M, Ciofi C, et al. 2019. Conserved sex chromosomes and karyotype evolution in monitor lizards (Varanidae). *Heredity*, **123**, 215–227.

Ironside JE. 2010. No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *BioEssays*, **32**, 718–726.

Iwase M, Satta Y, Hirai Y, Hirai H, Imai H, Takahata N. 2003. The amelogenin loci span an ancient pseudoautosomal boundary in diverse mammalian species. *Proceedings of the National Academy of Sciences*, **100**, 5258–5263.

Jeffries DL, Lavanchy G, Sermier R, Sredl MJ, Miura I, Borzee A, et al. 2018. A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. *Nature Communications*, **9**, 4088.

Jordan CY & Charlesworth D. 2012. The potential for sexually antagonistic polymorphism in different genome regions. *Evolution*, **66**, 505–516.

Julien P, Brawand D, Soumillon M, Necsulea A, Liechti A, Schütz F, et al. 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, 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.

Kawai A, Nishida-Umehara C, Ishijima J, Tsuda Y, Ota H, Matsuda Y. 2007. Different origins of bird and reptile sex chromosomes inferred from comparative mapping of chicken Z-linked genes. *Cytogenetic and Genome Research*, **117**, 92–102.

Kirkpatrick M & Hall DW. 2004. Male-biased mutation, sex linkage, and the rate of adaptive evolution. *Evolution*, **58**, 437–440.

Kitano J & Peichel CL. 2011. Turnover of sex chromosomes and speciation in fishes. *Environmental Biology of Fishes*, **94**, 549–558.

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

Kondo M, Nanda I, Schmid M, Schartl M. 2009. Sex determination and sex chromosome evolution: insights from medaka. *Sexual Development*, **3**, 88–98.

Kuroiwa A, Ishiguchi Y, Yamada F, Shintaro A, Matsuda Y. 2010. The process of a Y-loss event in an XO/XO mammal, the Ryukyu spiny rat. *Chromosoma*, **119**, 519–526.

Kretschmer R, Gunski RJ, del Valle Garnero A, de Freitas TRO, Toma GA, de Bello Cioffi M, et al. 2021. Chromosomal analysis in *Crotophaga ani* (Aves, Cuculiformes) reveals extensive genomic reorganization and an unusual z-autosome Robertsonian translocation. *Cells*, **10**, 4.

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

Lande R. 1979. Effective deme sizes during long-term evolution estimated from rates of chromosomal rearrangement. *Evolution*, **33**, 234–251.

Lenormand T, Fyon F, Sun E, Roze D. 2020. Sex chromosome degeneration by regulatory evolution. *Current Biology*, **30**, P3001–3006.e5.

Leroy T, Anselmetti Y, Tilak M-K, Berard S, Csukonyi L, Gabrielli M, et al. 2019. A bird's white-eye view on neosex chromosome evolution. *bioRxiv*:505610.

Lewis KR & John B. 1968. The chromosomal basis of sex determination. *International Review of Cytology*, **23**, 277–379.

Liu Z, Moore PH, Ma H, Ackerman CM, Ragiba M, Yu Q, et al. 2004. A primitive Y chromosome in papaya marks incipient sex chromosome evolution. *Nature*, **427**, 348–352.

Livernois AM, Graves JAM, Waters PD. 2012. The origin and evolution of vertebrate sex chromosomes and dosage compensation. *Heredity*, **108**, 50–58.

Mank JE. 2013. Sex chromosome dosage compensation: definitely not for everyone. *Trends in Genetics*, **29**, 677–683.

Matsubara K, Tarui H, Toriba M, Yamada K, Nishida-Umehara C, Agata K, et al. 2006. Evidence for different origin of sex chromosomes in snakes, birds, and mammals and step-wise differentiation of snake sex chromosomes. *Proceedings of the National Academy of Sciences*, **103**, 18190–18195.

Mueller JL, Skaletsky H, Brown LG, Zaghlul S, Rock S, Graves T, et al. 2013. Independent specialization of the human and mouse X chromosomes for the male germ line. *Nature Genetics*, **45**, 1083–1087.

Muller HJ. 1914. A gene for the fourth chromosome of *Drosophila*. *Journal of Experimental Zoo*logy, **17**, 325–336.

Muller HJ. 1918. Genetic variability, twin hybrids and constant hybrids, in a case of balanced lethal factors. *Genetics*, **3**, 422–499.

Muyle A, Käfer J, Zemp N, Mousset S, Picard F, Marais GA. 2016. SEX-DETector: a probabilistic approach to study sex chromosomes in non-model organisms. *Genome Biology and Evolution*, **8**, 2530–2543.

Myosho T, Takehana Y, Hamaguchi S, Sakaizumi M. 2015. Turnover of sex chromosomes in celebensis group medaka fishes. *G3*, **5**, 2685–2691.

Nanda I, Schlegelmilch K, Haaf T, Schartl M, Schmid M. 2008. Synteny conservation of the Z chromosome in 14 avian species (11 families) supports a role for Z dosage in avian sex determination. *Cytogenetic and Genome Research*, **122**, 150–156.

Natri HM, Merilä J, Shikano T. 2019. The evolution of sex determination associated with a chromosomal inversion. *Nature Communications*, **10**, 145.

Natri HM, Shikano T, Merilä J. 2013. Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. *Molecular Biology and Evolution*, **30**, 1131–1144.

Naurin S, Hansson B, Bensch S, Hasselquist D. 2010. Why does dosage compensation differ between XY and ZW taxa? *Trends in Genetics*, **26**, 15–20.

Nicolas M, Marais G, Hykelova V, Janousek B, Laporte V, Vyskot B, et al. 2004. A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants. *PLoS Biology*, **3**, e4.

Ogata M, Hasegawa Y, Ohtani H, Mineyama M, Miura I. 2008. The ZZ/ZW sex-determining mechanism originated twice and independently during evolution of the frog, *Rana rugosa*. *Heredity*, **100**, 92–99.

Ohno S. 1967. *Sex Chromosomes and Sex-linked Genes*. Springer, Berlin, Germany.

Oliveros CH, Field DJ, Ksepka DT, Barker FK, Aleixo A, Andersen MJ, et al. 2019. Earth history and the passerine superradiation. *Proceedings of the National Academy of Sciences*, **116**, 7916–7925.

O'Meally D, Ezaz T, Georges A, Sarre SD, Graves JAM. 2012. Are some chromosomes particularly good at sex? Insights from amniotes. *Chromosome Research*, **20**, 7–19.

Otto SP, Pannell JR, Peichel CL, Ashman TL, Charlesworth D, Chippindale AK, et al. 2011. About PAR: the distinct evolutionary dynamics of the pseudoautosomal region. *Trends in Genetics*, **27**, 358–367.

Pala I, Hasselquist D, Bensch S, Hansson B. 2012a. Patterns of molecular evolution of an avian neo-sex chromosome. *Molecular Biology and Evolution*, **29**, 3741–3754.

Pala I, Naurin S, Stervander M, Hasselquist D, Bensch S, Hansson B. 2012b. Evidence of a neo-sex chromosome in birds. *Heredity*, **108**, 264–272.

Palmer DH, Rogers TF, Dean R, Wright AE. 2019. How to identify sex chromosomes and their turnover. *Molecular Ecology*, **28**, 4709–4724.

Papageorgiou L, Eleni P, Raftopoulou S, Mantaiou M, Megalooikonomou V, Vlachakis D. 2018. Genomic big data hitting the storage bottleneck. *EMBnet Journal*, **24**, e910.

Parsch J & Ellegren H. 2013. The evolutionary causes and consequences of sex-biased gene expression. *Nature Reviews Genetics*, **14**, 83–87.

Pennell MW, Kirkpatrick M, Otto SP, Vamosi JC, Peichel CL, Valenzuela N, et al. 2015. Y Fuse? Sex chromosome fusions in fishes and reptiles. *PLOS Genetics*, **11**, e1005237.

Pennell MW, Mank JE, Peichel CL. 2018. Transitions in sex determination and sex chromosomes across vertebrate species. *Molecular Ecology*, **27**, 3950–3963.

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

Ponnikas S, Sigeman H, Abbott JK, Hansson B. 2018. Why do sex chromosomes stop recombining? *Trends in Genetics*, **34**, 492–503.

Ponnikas S, Sigeman H, Lundberg M, Hansson B. 2020. Extreme variation in recombination rate and genetic variation along the Sylvioidea neo-sex chromosome. *bioRxiv*, 2020.09.25.314054.

Qiu S, Bergero R, Guirao-Rico S, Campos JL, Cezard T, Gharbi K, Charlesworth D. 2016. RAD mapping reveals an evolving, polymorphic and fuzzy boundary of a plant pseudoautosomal region. *Molecular Ecology*, **25**, 414–430.

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

Rice WR. 1987. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution*, **41**, 911–914.

Rodrigues N, Studer T, Dufresnes C, Perrin N. 2018. Sex-chromosome recombination in common frogs brings water to the fountain-of-youth. *Molecular Biology and Evolution*, **35**, 942–948.

Ross JA, Urton JR, Boland J, Shapiro MD, Peichel CL. 2009. Turnover of sex chromosomes in the stickleback fishes (Gasterosteidae). *PLOS Genetics*, **5**, e1000391.

Rovatsos M, Pokorná M, Altmanová M, Kratochvíl L. 2014. Cretaceous park of sex determination: sex chromosomes are conserved across iguanas. *Biology Letters*, **10**, 20131093.

Sandell L & Otto SP. 2016. Probing the depths of biological diversity during the second century of genetics. *Genetics*, **204**, 395–400.

Sigeman H, Ponnikas S, Chauhan P, Dierickx E, Brooke MdeL, Hansson B. 2019. Repeated sex chromosome evolution in vertebrates supported by expanded avian sex chromosomes. *Proceedings of the Royal Society B*, **286**, 20192051.

Sigeman H, Ponnikas S, Hansson B. 2020. Whole-genome analysis across 10 songbird families within Sylvioidea reveals a novel autosome-sex chromosome fusion. *Biology Letters*, **16**, 20200082.

Singhal S, Leffler EM, Sannareddy K, Turner I, Venn O, Hooper DM, et al. 2015. Stable recombination hotspots in birds. *Science*, **350**, 928–932.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*, **423**, 825–837.

Smeds L, Skaletsky H, Huang X, Mank JE, Kimura M, Yanai I, et al. 2015. Evolutionary analysis of the female-specific avian W chromosome. *Nature Communications*, **6**, 7330.

Smith DA, Gordon IJ, Traut W, Herren J, Collins S, Martins DJ, et al. 2016. A neo-W chromosome in a tropical butterfly links colour pattern, male-killing, and speciation. *Proceedings of the Royal Society B*, **283**, 20160821.

Soh YQ, Alföldi J, Pyntikova T, Brown LG, Graves T, Minx PJ, et al. 2014. Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes. *Cell*, **159**, 800–813.

Stevens NE. 1906. *Studies in Spermatogenesis: A Comparative Study Of The Heterochromosomes In Certain Species Of Coleoptera, Hemiptera And Lepidoptera, With Especial Reference To Sex Determination*. Carnegie Institute of Washington, Publication no. 36, 33–74.

Tomaszkiewicz M, Makova KD, Medvedev P. 2017. Y and W chromosome assemblies: approaches and discoveries. *Trends in Genetics*, **33**, 266–282.

The Tree of Sex Consortium. 2014. Tree of Sex: A database of sexual systems. *Scientific Data*, **1**, 140015.

Úbeda F, Patten MM, Wild G. 2014. On the origin of sex chromosomes from meiotic drive. *Proceedings of the Royal Society B*, **282**, 20141932.

Vicoso B. 2019. Molecular and evolutionary dynamics of animal sex-chromosome turnover. *Nature Ecology & Evolution*, **3**, 1632–1641.

Vicoso B & Bachtrog D. 2013. Reversal of an ancient sex chromosome to an autosome in *Drosophila*. *Nature*, **499**, 332–335.

Vicoso B & Bachtrog D. 2015. Numerous transitions of sex chromosomes in Diptera. *PLoS Biology*, **13**, e1002078.

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.

White M. 1973. *Animal Cytology and Evolution*. Cambridge University Press, Cambridge, United Kingdom.

Wilson EB. 1905. Studies on chromosomes II. The paired microchromosomes, idiochromosomes and heterotropic chromosomes in Hemiptera. *Journal of Experimental Zoology*, **2**, 507–545.

Wright AE, Dean R, Zimmer F, Mank JE. 2016. How to make a sex chromosome. *Nature Communications*, **7**, 12087.

Xu L, Auer G, Peona V, Suh A, Deng Y, Feng S, et al. 2019. Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds. *Nature Ecology & Evolution*, **3**, 834–844.

Yoshida K & Kitano J. 2012. The contribution of female meiotic drive to the evolution of neo-sex chromosomes. *Evolution*, **66**, 3198–3208.

Zhou Q, Zhang J, Bachtrog D, An N, Huang Q, Jarvis ED, et al. 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science*, **346**, 1246338.

Acknowledgements

Many people have been involved in the making of this thesis. My supervisors and co-authors most of all, but I also want to acknowledge my other colleagues at the Department of Biology in Lund. Thank you for making it a friendly and inspiring place to work, and a place to miss during this last year which we have spent mostly at home.

Bengt, my main supervisor, thank you for your support and guidance over these years. You have been a great supervisor, and I am happy I had the opportunity to work with you on this project. Thank you for taking so much of your time to help me whenever I needed it, for your one million ideas and for always being on the lookout for new species for me to sequence and study.

Jessica, my assistant supervisor, I'm really glad you and Bengt formed the SexGen research group at the beginning of my PhD. Thank you for maintaining its fun and inspiring atmosphere, for offering advice and guidance during our weekly meetings and for keeping us reading and discussing science. I also want to thank you for making us do weird sports!

Dennis, my other assistant supervisor, thanks for your great reed warbler enthusiasm and for your great reed-ing and commenting on manuscripts.

Suvi, you have been my office mate and friend for four years, and I could not have asked for a better one. Thank you for always taking the time to look at my plots and talk about sex chromosomes, and Beyoncé. You have contributed immensely to my progress as a PhD student, and I have learned a lot from you.

Julian, I feel a bit like we're siblings too! I'm so glad we were on this PhD journey together.

Ann-Kathrin and **Linus**, thank you for all the fun times and for being such good friends.

Hongkai, thank you for the amazing EPA analyses, for babysitting me in the lab and for all the PlayStation recommendations.

Bella, your python expertise really took the pipeline to the next level and it was really fun working with you. Thanks!

Thank you to all other members of **SexGen**, past and present. Special thanks to **Kat**, **Colin**, **Qinyang, Vignesh, Aivars** and **Yesbol**, for all the great discussions, support, knowledge sharing and fun times.

I was lucky enough to be part of two research groups: Thank you also to all **MEEL**ers for the great presentations and weekly meetups.

The great reed warbler genome was done in collaboration with the MHC-people. **Maria**, thank you for your enthusiasm and for being such a great colleague. I don't know anyone who has more "favourite scaffolds" than you. **Helena**, during my first year as a PhD student we spent a lot of time in meetings discussing optimal genome assembly strategies. Thank you for being so engaged with this work and for always trying to find the best way forward.

Staffan, thank you for being a great head of **MEMEG**.

Max, you put a lot of time and effort into supervising my Master thesis project, for which I am grateful.

Jacob, thank you for all the Copenhagen dinners and for showing me around the Kvismaren field site.

Anders, thank you for keeping track of my PhD progress during these years.

All my fellow PhD students, thank you for making the PhD journey such a fun experience. **Kristaps**, **Carlos**, **Elsie, Victor, David**, **Ainara, Fredrik, Juan Pablo**, **Micaela**, **Elsa**, **Hamid**, **Alex**, **Samantha**, **Leidys**, **William**, **Martin**, **Pierre, and many others.**

Thank you to my other **colleagues** in the Ecology building for creating such a nice workplace: **Nathalie, Tobias, Arne, Mads, Georgina, Charlie, Emily, Stephen, Jane, Maarit, Amanda, Lars, Johan, Hannah W, Hanna L, Andreas, Anna R, Anna D, Maria, Maja, Utku, Violetta, Pallavi, and many others.**

Thank you **Björn** for coordinating the bioinformatic mentor program, and to **Dag** for being a great mentor. I really appreciate all the work that you put into this program and feel lucky to have had the opportunity to take part.

I would also like to acknowledge **SciLifeLab** for the amazing courses, keeping servers running, but mostly for sending **Verena** and **Estelle** our way. Thank you both so much for all the help over the years. **Ignas** and **Lucile**, thank you as well!

Thank you also to the coordinators of the **GENECO** research school, for providing opportunities for us PhD students to meet up and discuss our science, and for organizing great courses.

Emma, **Per**, and the rest of **FUN**, thank you for the work that you put into making the PhD education at our department great**.**

Torbjörn, thank you for being my half-time opponent, and for the discussions on the mysteries of chromosome fusions. And thank you **Bengt-Olle** for taking an interest in my research.

A big thank you to all my other **co-authors**, for your work and input on these manuscripts.

Philip, thank you for your love and support. Namaste for keeping me sane during these last months of frantic thesis work, for all the writing advice and late-night proof reading. You have also contributed to my scientific development as much as anyone, and this thesis would not have been the same without you.

Tack till min familj, speciellt till **mamma** och **pappa** som alltid finns där för mig. Tack för det fina omslaget till min avhandling **Martin**, och för den alternativa titeln: "Analyser av fågelmos". **Ingrid** och **Henrik**, tack till er med! **Farmor**, tack för att du finns och tror på mig.

List of papers

I. Sigeman H, Sinclair B, Hansson B. 2021. XYZWfinder: a snakemake pipeline for detecting and visualising sex chromosomes using short read data. Manuscript.

II. Sigeman H, Ponnikas S, Chauhan P, Dierickx E, Brooke MdeL, Hansson B. 2019. Repeated sex chromosome evolution in vertebrates supported by expanded avian sex chromosomes. Proceedings of the Royal Society B, 286, 20192051 (doi: 10.1098/rspb.2019.2051).

III. Sigeman H, Ponnikas S, Hansson B. 2020. Whole- genome analysis across 10 songbird families within Sylvioidea reveals a novel autosome-sex chromosome fusion. Biology Letters, 16, 20200082 (doi: 10.1098/rsbl.2020.0082).

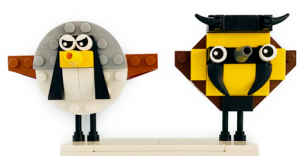
IV. Sigeman H, Zhang H, Adeb SA, Hansson B. 2021. A novel neo-sex chromosome in *Sylvietta brachyura* (Macrosphenidae) adds to the unprecedented avian sex chromosome diversity among Sylvioidea songbirds. Manuscript.

V. Sigeman H, Strandh M, Proux-Wéra E, Kutschera VE, Ponnikas S, Zhang H, Lundberg M, Soler L, Bunikis I, Tarka M, Hasselquist D, Nystedt B, Westerdahl H, Hansson B. 2020. Avian neo-sex chromosomes reveal dynamics of recombination suppression and W degeneration. bioRxiv, 2020.09.25.314088 (doi: 10.1101/2020.09.25. 314088). Submitted manuscript.

VI. Sigeman H, Hansson B. 2021. Evolutionary dynamics of enlarged sex chromosomes and novel pseudoautosomal regions in Sylvioidea songbirds. Manuscript.

VII. Sigeman H, Ponnikas S, Videvall E, Zhang H, Chauhan P, Naurin S, Hansson B. 2018. Insights into avian incomplete dosage compensation: sex-biased gene expression coevolves with sex chromosome degeneration in the common whitethroat. Genes, 9, 373 (doi: 10.3390/genes9080373).

VIII. Sigeman H, Hansson B. 2021. The rate of W chromosome degeneration across multiple autosome-sex chromosome fusions in birds. Manuscript.



Printed by Media-Tryck, Lund 2021 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Faculty of Science
Department of Biology
ISBN 978-91-7895-782-8