Genomics of expanded avian sex chromosomes shows that certain chromosomes are predisposed towards sex-linkage in vertebrates

Running head: Genomics of lark sex chromosomes

Hanna Sigeman[a,#], Suvi Ponnikas[a], Pallavi Chauhan[a], Elisa Dierickx[b], M. de L. Brooke[b], Bengt Hansson[a,#]

[a] Department of Biology, Lund University, Ecology Building, 223 62 Lund, Sweden
[b] Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

[#] Correspondence: hanna.sigeman@biol.lu.se; bengt.hansson@biol.lu.se.

1

ABSTRACT

Sex chromosomes have evolved from the same autosomes multiple times across vertebrates, suggesting that certain genomic regions are predisposed towards sex-linkage. However, to test this hypothesis detailed studies of independently originated sex-linked regions and their gene content are needed. Here we address this problem through comparative genomics of birds where multiple chromosomes appear to have formed neo-sex chromosomes: larks (Alaudidae; Sylvioidea). We detected the largest known avian sex chromosome (195.3 Mbp) and show that it originates from fusions between (parts of) four avian chromosomes (Z, 3, 4A and 5). We found evidence of five evolutionary strata where recombination has been suppressed at different time points, and that these time points correlate with the level of Z–W gametolog differentiation. We show that there is extensive homology to sex chromosomes in other vertebrate lineages: three of the fused chromosomes (Z, 4A, 5) have independently evolved into sex chromosomes in fish (Z), turtles (Z, 5), lizards (Z, 4A) and mammals (Z, 4A). Moreover, we found that the fourth chromosome, chromosome 3, was significantly enriched for genes with predicted sex-specific functions. These results support a key role of chromosome content in the evolution of sex chromosomes in vertebrates.

Sex chromosome | fusion | resequencing | Sylvioidea | larks | birds

1. INTRODUCTION

Sex chromosomes have evolved from autosomes many times independently across the tree of life. The most generic hypothesis for the evolution of sex chromosomes invokes a selective advantage of linkage between sex-determining and sexually antagonistic genes (Haldane 1922; Fisher 1931; Lewis & John 1968; D. Charlesworth & B. Charlesworth 1980). Accordingly, this hypothesis suggests that the chromosomes harbouring these genes should frequently be involved in the formation, transition and turnover of sex

40 chromosomes (Bachtrog et al. 2011; Ross et al. 2009; O'Meally et al. 2012). Indeed, there are a limited

41 number of genes across vertebrates that have taken on the sex-determining role, and evidence is

42 accumulating that the chromosomes on which these reside have evolved into sex chromosomes several

43 times independently in different lineages (reviewed in O'Meally et al. 2012; Ezaz et al. 2006; Ezaz et al.

44 2016). Furthermore, it has been suggested that some autosomes fuse with existing sex chromosomes into

45 neo-sex chromosomes more easily than others due to their gene content (Kitano et al. 2009; Ross et al.

46 2009; Pala, Hasselquist, et al. 2012; Kitano & Peichel 2011; Yoshida et al. 2014; O'Meally et al. 2012). The

47 importance of chromosome content for the evolutionary dynamics of sex chromosomes is, however, still

48 poorly supported due to a limited number of well-studied origins of sex-linkage.

49

50 The sex chromosomes of birds are highly stable with a Z chromosome size of approximately 73 Mbp

51 (Ellegren 2010). One of the few known exceptions to this stability is the neo-sex chromosome of passerines

52 in the superfamily Sylvioidea, which has been formed by a fusion between half of chromosome 4A

53 (according to the zebra finch *Taeniopygia guttata* nomenclature) and the ancestral sex chromosome (Pala,

54 Naurin, et al. 2012; Pala, Hasselquist, et al. 2012; Sigeman et al. 2018). Moreover, two independent findings

55 suggest that at least some species of larks (Alaudidae), a family within Sylvioidea, have acquired autosome–

56 sex chromosome fusions in addition to the one including parts of 4A. Firstly, heavily enlarged sex

57 chromosome karyotypes were found in the bimaculated lark (*Melanocorypha bimaculata*) and the horned lark

58 (*Eremophila alpestris*; Bulatova 1981), and, secondly, genetic markers located on chromosome 3 and 5 were

59 found to have sex-specific inheritance in the Raso lark (*Alauda razae*; Brooke et al. 2010).

60

61 Here, we use comparative genomics to study this neo-sex chromosome system across three Alaudidae

62 species – two *Alauda* species (Raso lark and Eurasian skylark *A. arvensis*) and one *Eremophila* species (horned

63 lark) – and their sister species in the family Panuridae (the bearded reedling *Panurus biarmicus*). We use whole-

64 genome sequence data of these species to characterise in detail which genomic regions have become fused

65 to the sex chromosomes, and use phylogenetic information to determine the appearance and age of several

66 evolutionary strata where recombination has been suppressed at different points in time (*sensu* Lahn & Page

67   1999). Then, we test whether the age of the strata explains the degree of Z-W divergence of gametologous

68   genes, as predicted by sex chromosome theory (Rice 1994; Lahn & Page 1999). Next, we test the

69   hypothesised importance of chromosome content for the formation of sex chromosomes by analysing

70   whether the fused chromosomal regions are enriched for genes with sex-specific functions. Lastly, we

71   evaluate signs of repeated sex chromosome evolution by searching for homologies between the fused sex

72   chromosomes in larks and the sex chromosomes of other vertebrate lineages. Our study supports a key role

73   of chromosome content in the evolution of sex chromosomes and highlights the importance of studying

74   independently originated sex chromosomes for understanding how and why vertebrate sex chromosomes

75   evolve.

2. MATERIALS & METHODS

76

2.1 Sequence data

77

78   We extracted DNA from blood samples using a phenol–chloroform protocol (Sambrock & W Russel, 2001)

79   of one female and one male individual from each of our four study species: Raso lark (*Alauda razae*, from

80   Cape Verde), Eurasian skylark (*A. arvensis cantarella*, from Italy), horned lark (*Eremophila alpestris flava*, from

81   Sweden) and bearded reedling (*Panurus b. biarmicus*, from Sweden). According to *cytochrome b* sequence data,

82   the two *Alauda* species diverged c. 6 million years ago (Mya), *Alauda* and *Eremophila* c. 14 Mya (Alström et

83   al. 2013), and Alaudidae and Panuridae c. 17 Mya (Moyle et al. 2016). All samples were collected non-

84   destructively and with permission from the relevant authorities (Direcção Geral do Ambiente, Cape Verde,

85   and Malmö/Lund Ethical Committee for scientific work on animals, Sweden, no. 17277-18). DNA was

86   sequenced with Illumina HiSeqX (150 bp, paired-end) by SciLifeLab Sweden.

2.2 *De novo* assembly and mapping

87

88　To obtain reference genomes, we created *de novo* genome assemblies from the male sequence data for each

89　of the study species. The resequencing data was trimmed using nesoni clip (https://github.com/Victorian-

90　Bioinformatics-Consortium/nesoni) with a minimum read quality of 20 and minimum read length of 20 bp.

91　The trimmed reads were then assembled with Spades v3.5.0 (Bankevich et al. 2012) using 5 different kmer

92　lengths (21, 33, 55, 77 and 127) and the setting "careful". Scaffolds shorter than 1 kbp were discarded.

93　Quality statistics from the assemblies were calculated using Quast v4.5.4 (Gurevich et al. 2013). Assemblies

94　with N50 over 50 kbp were kept for further analysis. This threshold meant keeping the assemblies of the

95　Raso lark (N50=103 kbp, 28304 scaffolds) and the bearded reedling (N50=68 kbp, 36455 scaffolds), while

96　the assemblies for the Eurasian skylark (N50=8 kbp, 256822 scaffolds) and horned lark (N50=22 kbp,

97　109088 scaffolds) were discarded.

98

99　Reads from all of the samples (n = 8) were cleaned for adaptor sequences with Trimmomatic v.0.3.6 (Bolger

100　et al. 2014) using the adaptor file TruSeq3-PE and options seedMismatches: 2, palindromeClipThreshold:

101　30 and simpleClipThreshold: 10. Trimming of low-quality bases was done using a quality threshold of 15

102　from the leading end and 30 from the trailing end. The reads were further trimmed for a minimum quality

103　of 20 over sliding windows of 4 bp. Lastly, any reads shorter than 90 bp were excluded from further analyses.

104　The number of remaining reads in our samples ranged from 176 to 404 million. The samples were quality

105　checked using fastqc v0.11.7 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/).

106

107　The male and female samples of Raso lark and bearded reedling were aligned to their respective genome

108　assemblies, while the Eurasian skylark and horned lark individuals were aligned to the genome assembly of

109　their closest relative, the Raso lark. The alignment was done with bwa mem v0.7.17 (Li & Durbin 2009),

110　marking shorter split hits as secondary (option -M) for downstream compatibility. The aligned reads were

111　sorted with samtools v1.7 (Li et al. 2009), and duplicated reads were removed using picardtools v2.18.0

112　(http://broadinstitute.github.io/picard). Assembly and alignment statistics are provided in Suppl. Tables

113　S1,2.

2.3 Chromosome anchoring

The scaffolds in the genome assemblies were grouped into different chromosomes and ordered into chromosome-level using the genome assembly of the zebra finch (*Taeniopygia guttata*; (Warren et al. 2010). The zebra finch genome assembly taeGut.3.2.4 was downloaded from Ensembl (Cunningham et al. 2015) and transformed into a database using the last v876 (Kiełbasa et al. 2011) program lastdb. The two genome assemblies, of the Raso lark and bearded reedling, were aligned to the zebra finch genome using the program lastal and converted to psl format using the script maf-convert, both from the same software suite last v876. From there, we extracted chromosome anchoring coordinates based on the longest match to the zebra finch genome for each 5 kbp window in the Raso lark and bearded reedling assemblies, with a minimum requirement of 500 matching base pairs per 5 kbp window. The assembly positions from the output files of the coverage and single nucleotide variant (SNV) analyses were then translated to the starting positions of the match to the zebra finch genome assembly (see section below).

2.4 Identification of sex-linked genomic regions

We identified sex-linked regions using two different kinds of genomic signatures: (i) differential mapping success in males and females (i.e. sex-specific genome coverage), and (ii) an excess or deficit of female-specific genetic variation. Sex chromosomes almost invariably evolve recombination suppression in the heterogametic sex so that regions that have been sex-linked for a long time (such as the sex chromosomes that formed in the ancestor of all birds) will show pronounced sequence divergence and degeneration in the non-recombining chromosome (the W in birds; Zhou et al. 2014). The regions belonging to the ancestral sex chromosomes can thus be identified by lower female coverage, and fewer female-specific genetic variants compared to males. This is because reads from the female-specific W-chromosome will either not map to the male reference genome due to substantial differentiation between the Z and W or because of deletions on the W chromosome. More-recently formed sex-linked regions may be identified by lower mapping success in females, although we expect a subtler difference as the W-linked genomic region may

6

139   not have yet developed substantial differentiation from the Z-linked homologous region. Here, we also

140   expect a higher amount of female-specific mutations compared to males, due to differentiation between the

141   Z and W sex chromosome copies.

142

143   To uncover differential mapping success in males and females for each species, we parsed the alignment

144   files for reads with more than two mismatches and calculated genome-wide coverage for 5 kbp windows

145   with bedtools v2.71.1 (Quinlan & Hall 2010). All genome coverage values were normalised between the

146   female and the male sample, based on the number of reads in the trimmed and adaptor-free fastq files. To

147   identify sex-linked regions, we binned the female-to-male coverage values for every 1 Mbp genomic region

148   and extracted the mean value from each bin. Following discovery of sex-linked genomic regions, we

149   performed another normalisation step by dividing the median female-to-male coverage ratio for each 5 kbp

150   window by the genome-wide median female-to-male coverage ratio counting only chromosomes without

151   sex-linked regions. We then binned the data into 0.1 Mbp windows and calculated the mean female-to-male

152   coverage ratio within each bin.

153

154   To analyse female-specific variation, we called variants in the alignment files (with all mismatches allowed)

155   with freebayes v1.1.0 for each species separately (n = 2 in each analysis) using freebayes-parallel (Garrison

156   & Marth 2012) and parallel v20180322 (Tange 2018). The output was then parsed for any SNP that had

157   been marked with a flag other than PASS (--remove-filtered-all), a minimum quality of 20 and minimum

158   depth of 3x using vcftools v0.1.15 (Danecek et al. 2011). Private alleles (minor alleles occurring only in one

159   sample in a heterozygous state) were extracted with vcftools using option --singletons. We calculated the

160   difference between the number of female-specific private alleles and male-specific private alleles for each 5

161   kbp window and extracted the average difference across 1 Mbp and 0.1 Mbp windows.

162 ## 2.5 Analysis of divergence between gametologous genes

163   We used the whole-genome synteny aligner program SatsumaSynteny v. 2.0 (Grabherr et al. 2010) to align

164   the Raso lark assembly to the zebra finch assembly (taeGut.3.2.4), and then used kraken (Zamani et al. 2014,

7

165   downloaded 18 June 2018) to make a lift-over of the zebra finch annotations to the Raso lark assembly. Of

166   the 18204 transcripts and 17488 genes in the zebra finch annotation, 14466 transcripts (79 %) from 13764

167   genes (79 %) were annotated in the Raso lark. We used Freebayes v.1.1.0 (Garrison & Marth 2012) (--report-

168   monomorphic) to call variants for every base pair within all exons in the Raso lark based on the genome

169   coordinates from the lift-over.

170

171   We *in silico* extracted gametologous gene sequences from sex-linked regions using an in-house script (code

172   provided as Supplementary Code S1; general methodology described in (Sigeman et al. 2018) based on the

173   genotypes of the female and male samples. The script uses sex-specific genotype information to phase the

174   data into a Z and W sequence. We replaced a site by "N" if it had a quality score or sequence depth below

175   20 or if the genotype in either sample was not called. Any site without variants in either sample was extracted

176   as such, and remaining variants between the male and female were phased based on sex-specific allele

177   compositions provided in Suppl. Table S3.

178

179   To confirm that both Z and W gametologs were present in the data, we calculated genome coverage values

180   for every exon using bedtools v.2.27.1 multicov (Quinlan & Hall 2010) and normalised the values between

181   the two samples of each species. Exons with female coverage less than 75 % of the male sample were

182   masked with N:s, as this suggests absence of a W gametolog.

183

184   We used TransDecoder v3.0.1 (https://github.com/TransDecoder/TransDecoder/) to find the longest

185   open reading frames for each of the sequences (--retain_pfam_hits, --retain_blastp_hits, --single-best-orf).

186   Orthologous genes from the zebra finch, collared flycatcher (*Ficedula albicollis*) and chicken (*Gallus gallus*)

187   were downloaded through BioMart (database: Ensembl Genes 93), and the longest transcript from the

188   flycatcher and chicken corresponding to the zebra finch transcripts in the annotation was added as additional

189   sequences. These gene sequences were not used in further analyses but aided in evaluating the quality of the

190   alignments. The sequences were codon-aware aligned using *Prank* v.150803 (Loytynoja 2014). Each N in

191   the sequences was transformed into "-", and all sites including this character in any sequence were then

192 removed using gblocks v0.91b (Castresana 2000) . We calculated pairwise substitution rates between Z–W

193 gametologs using codeml (estimated kappa and codon frequency F3X4) from the PAML v4.9 package (Yang

194 2007). Gene sequences longer than 500 bp, and where the Z and W sequences within a species had a

195 synonymous substitution rate (dS) value above 0.01, were kept for further analysis. We used Kruskal-Wallis

196 tests to investigate potential differences in sequence divergence between the identified sex-linked genomic

197 regions (or strata, see below) as the data were not normally distributed. All values were $\log_2$-transformed

198 prior to these tests to fulfil the criteria of similar distributions between groups and a low number (0.00001)

199 was added to all synonymous substitution rate (dN) values as some contained zeros. Significance levels of

200 differences between groups were calculated using pairwise Wilcoxon tests with Benjamini & Hochberg

201 adjusted p values. Furthermore, we used ordered heterogeneity tests (OH; Rice & Gaines 1994) to evaluate

202 if the difference between Z–W nucleotide divergences between groups correlated with relative timing of

203 sex-linkage between the evolutionary strata. The OH value r_sp_c was calculated by multiplying the

204 complement of the p value from the Kruskal-Wallis tests (1 – p value) with the Spearman Rank rho value

205 from the median nucleotide divergence estimates for each stratum correlated with the order in which they

206 appear based on the phylogeny of the studied species. The relative timing of sex-linkage of stratum 5a and

207 5b (see below) could not be distinguished using our data and were thus given the same rank. P values were

208 extracted from the supplied figures in Rice & Gaines (1994), and multiplied by two for two-tailed hypothesis

209 testing.

## 2.6 Gene enrichment analysis

211 We downloaded Gene Ontogeny (GO) annotations for all the genes in the zebra finch genome from

212 Ensembl BioMart (taeGut3.2.4; accessed on 12 September 2018) and parsed the file for GO term names

213 with the following pattern matches: "sperm", "ovarian", "gonad", "estrogen", "testosterone", "sex

214 differentiation", "sex determination", "sexual characteristics", "sexual reproduction" and "oogenesis". The

215 zebra finch gene annotation had in total 323 genes matching to any of these patterns. We then counted the

216 number of genes matching any of these terms for i) the sex-linked region for each chromosome separately,

217 and ii) within each of the identified evolutionary strata (see below). We performed two-tailed binomial tests

9

218   to see if these genomic regions contained either more or fewer sex-related genes than would be expected if

219   the genes were randomly distributed across the genome. We calculated the proportion of the genome that

220   made up each studied sex-linked region (1.253 Gbp is the full genome size of the zebra finch genome

221   assembly used) and tested if the region contained a higher or lower proportion of sex-related genes than

222   would be expected based on this probability. The p values for the two categories of tests (i, ii) were adjusted

223   using the Benjamini & Hochberg method.

224   # 3. RESULTS

225   ## 3.1 Identification of sex-linked genomic regions

226   A genome-wide scan for sex-specific genetic variation (measured as the average difference in number of

227   female and male private SNVs per Mbp), and sex-specific genome coverage (measured as the average

228   female-to-male read coverage ratio per Mbp), revealed that four chromosomes stood out from the

229   autosomal pattern across the species (Figure 1; Suppl. Table S4). First, the ancestral Z chromosome showed

230   substantially lower female coverage, and no or a moderately lower amount of female-specific genetic

231   variation, in all four species (Figure 1). Secondly, the first half of chromosome 4A showed substantially

232   higher number of female-specific SNVs and moderately lower female coverage in all species. Thirdly, a

233   substantially higher amount of female-specific SNVs and no or moderate sex-specific coverage were found

234   over varying extents of chromosome 3 and 5 in the four species – small parts of chromosome 3 in the

235   bearded reedling, slightly larger parts of chromosome 3 in the horned lark, and the greater part of

236   chromosome 3 and 5 in the Raso lark and the Eurasian skylark (Figure 1). These results support the existence

237   of multiple autosome–sex chromosome fusions and several evolutionary strata where recombination has

238   been suppressed at different points in time across the avian and the lark phylogeny.

239
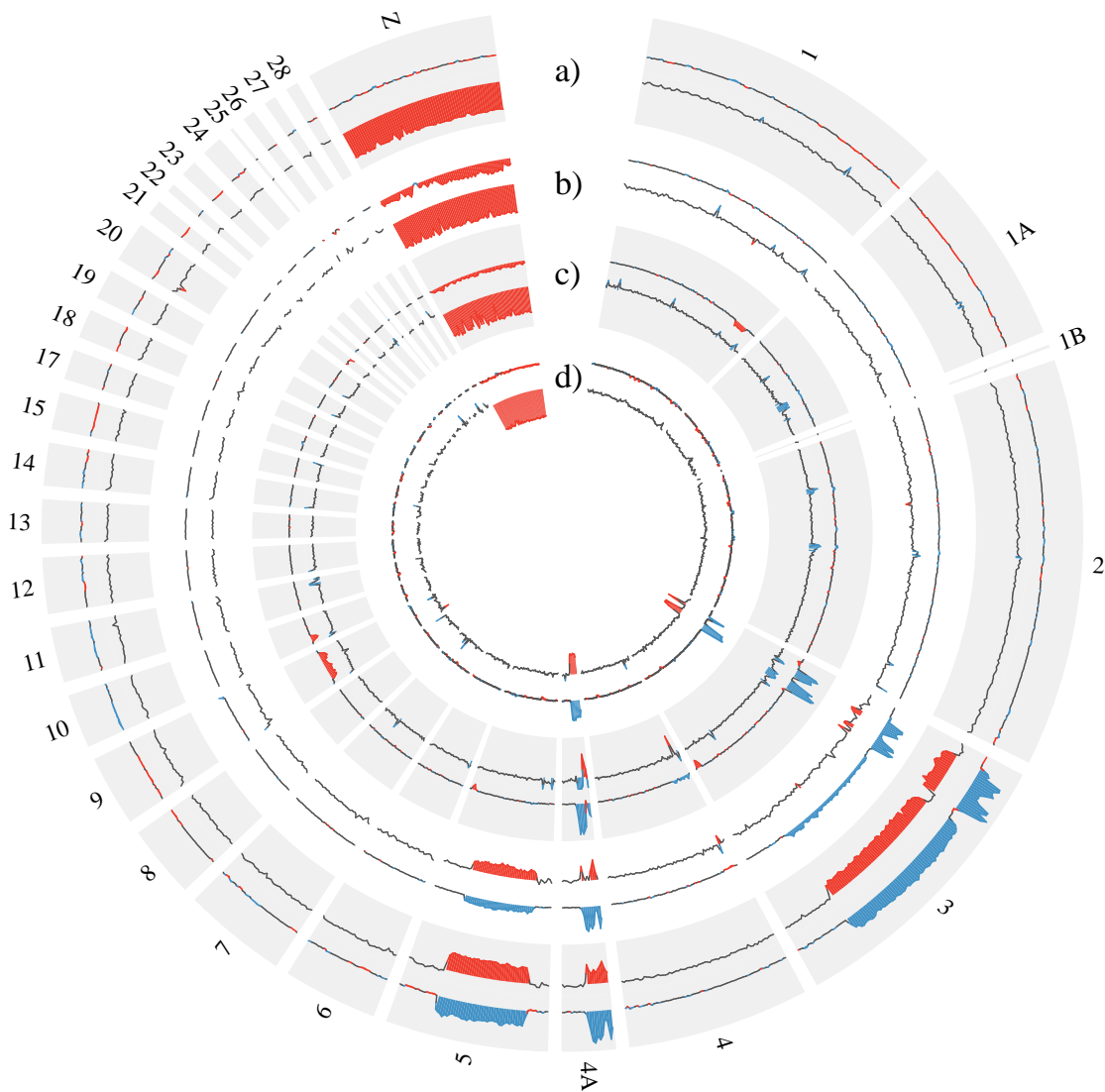
Figure 1. Genome-wide distribution of female-to-male difference in number of private single nucleotide variants (SNVs), and female-to-male coverage ratio, in the four study species. The background colours grey and white separate the data for each of the four species which are shown in the following order starting from the outer ring: a) Raso lark, b) Eurasian skylark, c) horned lark and d) 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 (with values > 500 in blue and values < -500 in red), and the inner line shows the average female-to-male coverage ratio across 1 Mbp windows (with values > 1.1 in blue and values < 0.9 in red). Chromosome-wide averages are provided in Suppl. Table S4.

To define the evolutionary strata more precisely, and to estimate their age, we assessed (i) the genomic signatures of sex-linkage (i.e., sex-specific genetic variation and genome coverage as above) at the scale of

11

250   0.1 Mbp regions in each species (Table 1; Suppl. Figure S1), and (ii) the most parsimonious order of

251   emergence based on dated phylogenies (Figure 2; Alström et al. 2013, Cortez et al. 2014, Moyle et al. 2016).

252   Stratum 1 corresponds to the entire ancestral Z chromosome (72.9 Mbp), which was formed c. 140 Myr

253   ago (Cortez et al. 2014), and stratum 2 corresponds to the previously known Sylvioidea neo-sex

254   chromosome, i.e. the first 9.6 Mbp of chromosome 4A (Pala et al. 2012), which was formed c. 21-19 Myr

255   ago when Sylvioidea split from other passerines (Moyle et al. 2016). For both these strata, all four study

256   species showed clear signatures of sex-linkage (Table 1; Suppl. Figure S1). Likewise, all four study species

257   showed sex-specific genetic variation and coverage in the regions between 8.4–10.4 and 18.1–24.1 Mbp on

258   chromosome 3 (with a total size of 8.0 Mbp): this constitutes stratum 3, which is the earliest stratum being

259   unique to the Alaudidae/Panuridae clade. The age of stratum 3 is defined by the split of

260   Alaudidae/Panuridae from other Sylvioidea families c. 19-17 Myr ago (Moyle et al. 2016). The next stratum,

261   stratum 4, has a size of 3.6 Mbp and spans the region between 10.4–14.0 Mbp on chromosome 3. It occurs

262   in all three lark species, but not in the bearded reedling, which gives an estimated age of c. 17-14 Myr as

263   defined by the split between Alaudidae and Panuridae (Alström et al. 2013). Finally, the Raso lark and the

264   Eurasian skylark had two additional strata on chromosome 3 (stratum 5a: 64.9 Mbp spanning the regions

265   between 5.8–8.4, 14.0-18.1 and 29.8-88.0 Mbp) and chromosome 5 (stratum 5b: 36.3 Mbp, the region

266   between 9.1-45.4 Mbp), respectively, that were not shared with the other two species. These strata form the

267   youngest sex chromosome layer, estimated to be c. 14-6 Myr based on the split between the horned lark

268   and the *Alauda* larks (Alström et al. 2013). Together, these sex-specific regions on chromosome Z, 4A, 3

269   and 5 amounts to 195.3 Mbp of the genome of the Raso lark and Eurasian skylark (16.3% of the genome

270   based on a genome size estimation of 1.2 Gbp) (Table 1; Suppl. Figure S1). The horned lark had 94.1 Mbp

271   (7.8%), and the bearded reedling 90.5 Mbp (7.5%), sex-linked genomic material (Table 1; Suppl. Figure S1).

272   Figure 2a shows a summary of the chromosomal position of the evolutionary strata and their most

273   parsimonious order of appearance based on presence and absence in the studied species (also shown are

274   silhouettes indicating vertebrate groups where homologs to these chromosomes act as sex chromosomes;

275   see Discussion). Figure 2b summarises the timings of sex-linkage based on dated phylogenies (Cortez et al.

276   2014, Alström et al. 2013, Moyle et al. 2016).

277

Table 1. Sex-linked regions in each of the four species. The strata are numbered according to the most parsimonious order of emergence based on phylogenetic analyses. Stratum 1 (homologous to chromosome Z in zebra finch) acts as the sex chromosome in all birds, and sex-linkage of stratum 2 (chromosome 4A in zebra finch) is common for all birds belonging to the superfamily Sylvioidea, which includes larks and the bearded reedling. Stratum 3 was seen to be sex-linked in all four studied species and is therefore older than stratum 4, which appears as sex-linked in all larks but not the bearded reedling. Stratum 5a and 5b appear as sex-linked in the Raso lark and Eurasian skylark, but not in horned lark or bearded reedling, and is therefore the youngest stratum but here divided into two substrata as they originated from different autosomes. The table provides mean values for the male-to-female coverage ratio (coverage) and female-to-male difference in number of private SNVs across 0.1 Mbp windows for each stratum. Bearded reedling and horned lark have no sex-linkage for stratum 5a and 5b, and bearded reedling not for stratum 4, and are therefore marked with NA in addition to the corresponding values.

| Stratum	Chromosome	Genomic region (Mbp)	Stratum Size (Mbp)		Raso lark	Eurasian skylark	Horned lark	Bearded reedling
1	Z	0-72.9	72.9	Coverage	0.52	0.55	0.54	0.52
				SNV	-17.77	-992.10	-375.85	-160.48
2	4A	0-9.6	9.6	Coverage	0.73	0.89	1.06	0.66
				SNV	2957.35	2227.36	2518.46	2413.16
3	3	8.4-10.4, 18.1-24.1	8.0	Coverage	0.74	0.95	1.20	0.68
				SNV	3386.68	2435.65	3404.79	2787.17
4	3	10.4-14.0	3.6	Coverage	0.69	0.84	0.97	NA (1.00)
				SNV	3019.06	2161.91	2512.71	NA (-10.54)
5a	3	5.8-8.4, 14.0-18.1, 29.8-88.0	64.9	Coverage	0.74	0.98	NA (1.00)	NA (0.99)
				SNV	1785.39	451.48	NA (1.73)	NA (1.77)
5b	5	9.1-45.4	36.3	Coverage	0.69	0.85	NA (1.00)	NA (0.99)
				SNV	1952.04	879.14	NA (20.19)	NA (-37.92)

289

290

13
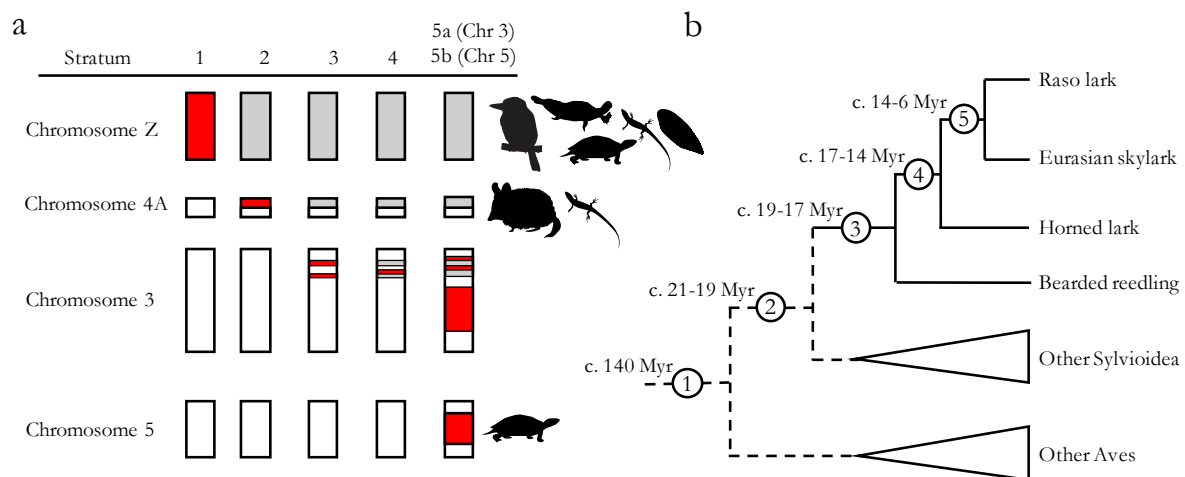
Figure 2. a) Cartoon representation of the evolutionary strata in the most parsimonious order of appearance (starting with stratum 1 and ending with stratum 5a and 5b). White colour represents autosomal regions and red colour marks the stratum being added. Grey colour represents sex-linked regions that already exist at the timing of the new stratum. Silhouettes indicate vertebrate groups where homologs to these chromosomes act as sex chromosomes (see Discussion): chromosome Z acts as the sex chromosome in all birds (Smith et al. 2009) and its homolog acts as sex chromosome in a flatfish (*Cynoglossus semilaevis*; Chen et al. 2014), a turtle (*Staurotypus triporcatus*; Montiel et al. 2016), a gekko (*Gekko hokouensis*; Kawai et al. 2008) and in the platypus (*Ornithorhynchus anatinus*; Grützner et al. 2004); the homolog of chromosome 4A is the sex chromosome in a lizard (*Takydromus sexlineatus*; Rovatsos et al. 2016) and in all eutherian mammals (Ross et al. 2005); and the homolog of chromosome 5 acts as sex chromosome in two turtle groups (*Glyptemys* wood turtles and *Siebenrockiella* marsh turtles; Montiel et al. 2016). (b) A cladogram showing when in time (Myr) the evolutionary strata were formed. The dating of stratum 1 is based on analyses of gametologous gene pairs in Cortez et al. (2014), stratum 2 and 3 on speciation estimates in Moyle et al. (2016), and stratum 4 and 5 (5a and 5b) on speciation estimates in Alström et al. (2013).

3.2 Analysis of divergence between gametologous genes

We analysed the Z–W divergence from the gametologous genes located on the different strata across the sex chromosome in the Raso lark and the Eurasian skylark, i.e. the species with all five evolutionary strata. We analysed the two different parts of stratum 5 on chromosome 3 and 5 separately in these analyses (i.e., stratum 5a and 5b, respectively). The degree of synonymous substitutions (median dS) between the Z and

14

310   W gametologs ranged from 0.22 at stratum 1 to 0.02 at stratum 5a and 5b in the Raso lark, and from 0.21

311   at stratum 1 to 0.02 at stratum 5a and 5b in the Eurasian skylark (Figure 3, Suppl. Table S5). In the Raso

312   lark, the $\log_2$-transformed dS values differed significantly between all strata except between stratum 2 and

313   stratum 3, and between stratum 5a and 5b (Kruskal-Wallis rank sum test: chi-squared = 154.0, df = 4, p <

314   0.001; for pairwise comparison values see Suppl. Table S6). In the Eurasian skylark, the $\log_2$-transformed

315   dS differed significantly between all strata, expect between stratum 5a and 5b (Kruskal-Wallis rank sum test:

316   chi-squared = 143.0, df = 4, p < 0.001; pairwise comparison values in Suppl. Table S6). The degree of non-

317   synonymous substitutions (dN) between the Z and W gametologs showed similar patterns, but less

318   pronounced (median dN ranged from 0.03 at stratum 1 to 0.002 at stratum 5a in the Raso lark, and from

319   0.03 at stratum 1 to 0.002 at stratum 5a in the Eurasian skylark; Figure 3, Suppl. Table S5). The dN values

320   were significantly different between stratum 5a and all other strata, and between 5b and all other strata in

321   both species, and between stratum 1 and stratum 4 in the Raso lark (Raso lark: chi-squared = 88.7, df = 4,

322   p < 0.001; Eurasian skylark: chi-squared = 143.0, df = 4, p < 0.001; for pairwise comparison values see

323   Suppl. Table S6). The dN/dS values were lowest at stratum 1 in both species (0.082 in the Eurasian skylark

324   and 0.097 in the Raso lark) and highest at stratum 4 (0.196 in the Eurasian skylark and 0.294 in the Raso

325   lark). The dN/dS values in stratum 4 differed significantly from strata 1, 2 and 5a, and stratum 5a differed

326   significantly from stratum 5b, in both species (Raso lark: chi-squared = 17.7, df = 4, p = 0.001; Eurasian

327   skylark: chi-squared = 14.0, df = 4, p = 0.007; pairwise comparisons in Suppl. Table S6).

328
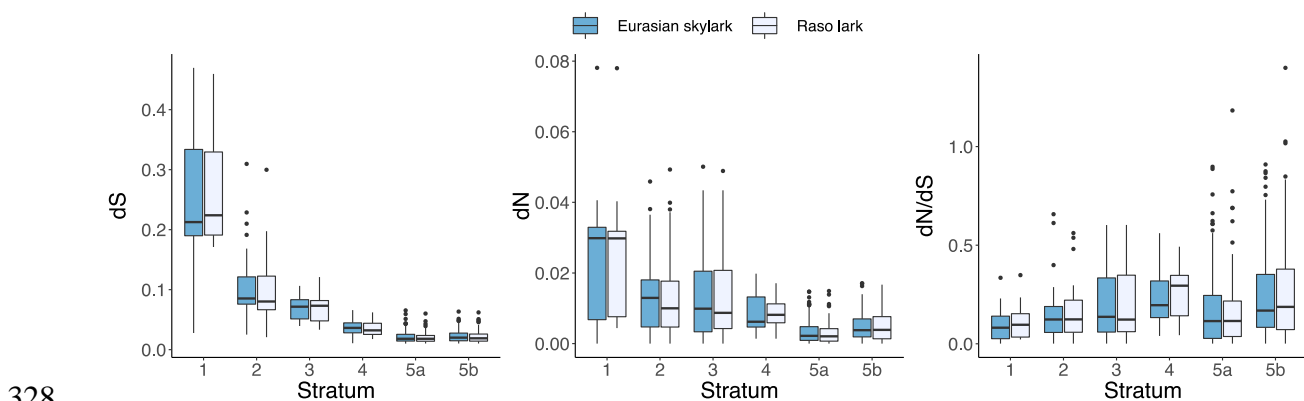
329   Figure 3. Z-to-W substitution rates for gametologous genes positioned within the different evolutionary strata in the

330   Eurasian skylark (dark blue) and the Raso lark (light blue). (a) Median dS values, (b) median dN values, and (c) median

331 dN/dS values are marked by the black line in each box, and the upper and lower hinges correspond to the first and

332 third quarantiles (the 25 and 75 percentiles). The whiskers extend to no more than 1.5× the interquartile range from

333 each hinge. See Suppl. Table S5 for details and main text for results from analyses of variances between strata based

334 on $\log_2$-transformed values.

335

336 In both Raso larks and Eurasian skylarks, the dS values correlated significantly with the relative age of the

337 strata (stratum 1 being oldest and strata 5a and 5b being youngest) (Ordered-heterogeneity test: $r_sP_c = 0.99$,

338 $k = 6$, $p < 0.001$ for both species). This was also true for the dN values ($r_sp_c = 0.99$, $k = 6$, $p < 0.001$ for

339 both species). However, the dN/dS values were not significantly associated with the age of the strata (Raso

340 lark: $r_sp_c = 0.41$, $k = 6$, $p \sim 0.11$; Eurasian skylark: $r_sP_c = 0.46$, $k = 6$, $p \sim 0.11$).

3.3 Gene enrichment analysis

342 Of the 323 genes with sex-related GO term names (see Methods section), 16 genes were located on the

343 ancestral avian sex chromosome Z (stratum 1), 5 within the sex-linked part of chromosome 4A (stratum 2),

344 33 within the sex-linked region of chromosome 3 (strata 3, 4 and 5a) and 5 within the sex-linked part of

345 chromosome 5 (stratum 5b) (Figure 4a; Suppl. Table S7). The different strata on chromosome 3 hold 5

346 (stratum 3), 0 (stratum 4) and 28 (stratum 5a) genes, respectively (Figure 4b; Suppl. Table S7). This means

347 that 59 genes with these hypothesised sex-related functions are sex-linked in the Raso lark and in the

348 Eurasian skylark, i.e. the species with all strata, compared to 16 in species only having the ancestral avian

349 sex chromosome.

350

351 Binomial tests showed that the sex-linked part of chromosome 3 (strata 3, 4 and 5a) had significantly more

352 sex-related genes than would be expected if the genes were randomly distributed over the genome (33

353 observed genes compared to 19.7 expected genes; Binomial test: adjusted $p = 0.019$; Figure 4a; Suppl. Table

354 S7). The sex-linked regions on chromosomes Z, 4A and 5 showed no statistical difference between observed

355 and expected number of genes (NS; Suppl. Table S7). When analysing each stratum separately, stratum 5a

356   was also significantly enriched for sex-related genes (28 observed genes compared to 16.7 expected genes;

357   adjusted p = 0.047; Figure 4b; Suppl. Table S7), while the other strata were not significantly enriched (NS;

358   Suppl. Table S7). Information about the analysed genes (chromosomal positions, GO terms and gene ID)

359   are provided in Supplementary Document S1.

360
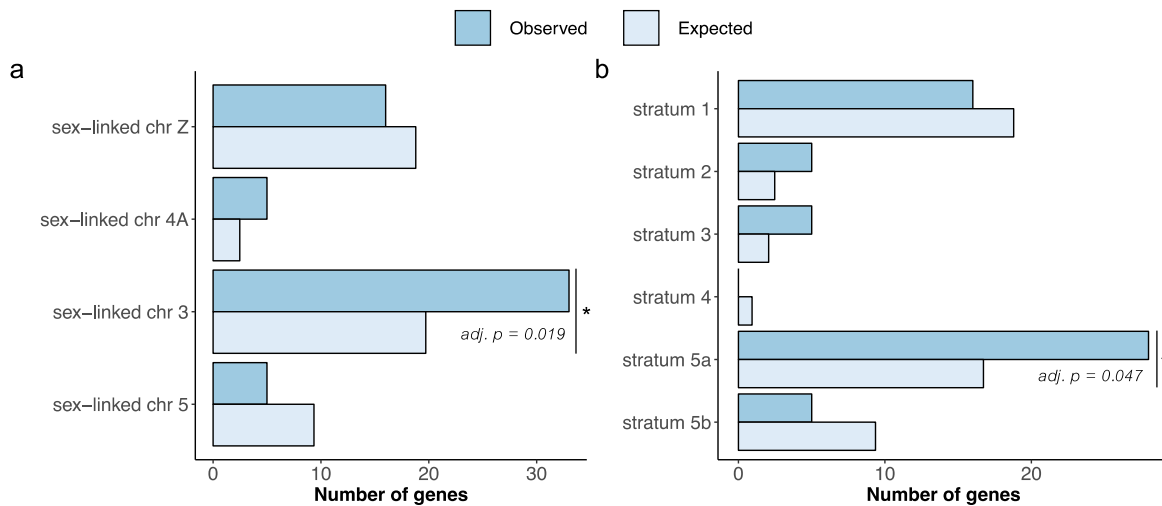
361

362   Figure 4. Observed and expected number of sex-related genes for (a) the sex-linked region of each chromosome and

363   (b) each stratum. Binomial tests (Suppl. Table S7) showed that the sex-linked region on chromosome 3 had significantly

364   more sex-related genes than would be expected if the genes were randomly distributed throughout the genome

365   (adjusted p = 0.019), and that this enrichment on chromosome 3 was mainly due to a high number of sex-related genes

366   on stratum 5a (adjusted p = 0.047).

4. DISCUSSION

368   By using comparative genomics in larks and their closest relative, the bearded reedling, we have identified

369   the precise genomic regions that have evolved sex-linkage across a group of birds with multiple autosome–

370   sex chromosome fusions. The extent of sex-linkage varied between species, and for two of them (Raso lark

371   and Eurasian skylark in the genus *Alauda*), we detected the largest known avian sex chromosome (195.3

372   Mbp) and show that it originates from fusions between (parts of) four avian chromosomes (Z, 3, 4A and

373   5). We found evidence of five evolutionary strata where recombination has been suppressed at different

17

374   time points; from c. 140 Myr ago (stratum 1, the sex chromosome common to all birds) to c. 14-6 Myr ago

375   (stratum 5a and 5b, the layer unique to *Alauda*; the strata and their age estimates are summarised in Figure

376   2). We further show that the substitution rates between Z and W gametologs increase with increasing age

377   of the strata. This gives support for a key prediction from sex chromosome theory: continuous

378   differentiation of non-recombining chromosomes (Rice 1994; Lahn & Page 1999).

379

380   With the available genomic data, we are not able to determine the chromosomal structure of this neo-sex

381   chromosome. However, the 195.3 Mbp sex-linked region in the two *Alauda* study species (Raso lark and

382   Eurasian skylark) constitutes c. 16.3% of their expected genome size (1.2 Gbp), which in turn corresponds

383   well to the genomic proportion of the sex chromosomes in the karyotypes of the bimaculated lark (i.e., c.

384   15-20%; Bulatova 1981). The bimaculated lark karyotypes further show that the Z and W chromosomes are

385   of similar size and drastically enlarged compared to the usual bird karyotype (Bulatova 1981). Therefore, we

386   believe that the additional chromosomes that are sex-linked in larks have fused to the Z as well as the W

387   chromosome. Moreover, the karyotype of a male horned lark (no female was karyotyped) indicates a Z

388   chromosome of similar size to that of the bimaculated lark (Bulatova 1981). Therefore, it is possible that

389   large parts of chromosome 3 and 5 are fused to the Z chromosome also in the horned lark, but as our results

390   show are still recombining in that species (with the exception of strata 3 and 4 on chromosome 3 that are

391   non-recombining). This reasoning is complemented by the results in E. Dierickx et al. (unpublished

392   manuscript) showing that different species and subspecies of *Alauda* (the oriental skylark *A. gulgula*, Eurasian

393   skylark sspp. [possibly *A. a. arvensis/dulcivox/intermedia*], and Raso lark), and a species of the genus *Gallerida*

394   (the crested lark *G. cristata*) – the sister genus of *Alauda* – have suppressed recombination on chromosome

395   5 (stratum 5b) but show recombination suppression to different degrees on chromosome 3 (stratum 5a).

396

397   A long-standing hypothesis in sex chromosome research posits that the evolutionary dynamics of sex

398   chromosomes are governed by selection acting on linked genes, in particular on sex-determining and

399   sexually antagonistic genes (Haldane 1922; Fisher 1931; Lewis & John 1968; D. Charlesworth & B.

400   Charlesworth 1980). This hypothesis states that chromosomes that harbour such genes should be over-

18

401 represented in sex chromosome formations, transitions and turnovers (Bachtrog et al. 2011; Ross et al. 2009;

402 O'Meally et al. 2012). In support of this hypothesis, we found that the sex-linked region of chromosome 3,

403 and one of the strata on that chromosome (stratum 5a), were significantly enriched for genes with sex-

404 related functions (adjusted p = 0.019 and 0.047, respectively). Among these genes is *ESR1* (*Estrogen receptor*

405 *1*), which has been suggested to be involved in sex reversal in the American alligator (*Alligator mississippiensis*;

406 Kohno et al. 2015). The other chromosomes (and strata) involved in the extensive neo-sex chromosome

407 formation in larks were not significant enriched for sex-related genes. However, all of them have

408 independently been recruited as sex chromosomes in other vertebrate lineages. The Z chromosome – the

409 sex chromosome in all birds, containing the putative sex determining gene *DMRT1* (*Doublesex and mab-3*

410 *related transcription factor 1*; Smith et al. 2009) – has independently been recruited as a sex chromosome in a

411 flatfish (the half-smooth tongue sole *Cynoglossus semilaevis*; Chen et al. 2014), a turtle (the Mexican musk turtle

412 *Staurotypus triporcatus*; Montiel et al. 2016), a gecko (the Kwangsi gecko *Gekko hokouensis;* Kawai et al. 2008),

413 and in the platypus (*Ornithorhynchus anatinus*; Grützner et al. 2004). Chromosome 4A is homologous to the

414 sex chromosome in a lizard (the Asian grass lizard *Takydromus sexlineatus*; Rovatsos et al. 2016) and in all

415 eutherian mammals (Ross et al. 2005), and contains several interesting genes, including *AR* (*Androgen receptor*)

416 and *SOX3* (*SRY-related HMG-box 3*) (note, however, that *SOX3* in not sex-linked in Sylvioidea; Pala, Naurin,

417 et al. 2012; this study). Finally, the homolog to chromosome 5 acts as sex chromosome in two turtle groups

418 (*Glyptemys insculpta* and *Siebenrockiella crassicollis*; Montiel et al. 2016). An alternative hypothesis is that fusions

419 may become fixed as a consequence of non-selective processes (see e.g. Pennell et al. 2015). However, we

420 have presented several strands of evidence – the enrichment of genes with sex-specific function and the

421 repeated homology to other vertebrate sex chromosomes – to support the hypothesis that the multiple

422 fusion events between chromosome Z, 4A, 3 and 5, that have formed these extraordinary neo-sex

423 chromosomes in larks and in the bearded reedling, have been driven by selective processes acting on their

424 gene content. This adds to the accumulating evidences that specific chromosomes are non-randomly

425 recruited as sex chromosomes in vertebrates (O'Meally et al. 2012; Ezaz et al. 2006; Ezaz et al. 2016).

426

427 Multiple autosome–sex chromosome fusions have also occurred in other clades, including gazelles and

428 snakes (Vassart et al. 1995; Pokorná et al. 2014). In snakes, all six neo-sex chromosome systems identified

429 so far are from a single family, Elapidae (Pokorná et al. 2014). These patterns, now also including those

430 found in the larks, hint at a phylogenetic component in neo-sex chromosome formation; that once an

431 autosome–sex chromosome fusion has occurred, additional ones may be expected (Pokorná et al. 2014). If

432 the fusions are being selected for because of any advantageous linkage that they may create, then previously

433 fused regions (with already suppressed recombination) provide more opportunities for beneficial linkage to

434 be formed between the sex-linked genes and new additional genomic regions.

435

Acknowledgements

446

Supplementary Material and Data Availability

448 Suppl. Figure S1 and Suppl. Tables S1-7 are provided together as a separate file. Suppl. Document S1 is

449 provided as a separate file. In-house scripts are given in Suppl. Code S1 (available upon request). Illumina

450 HiSeqX raw reads (150 bp, paired-end) will been deposited at NCBI sequence read archives upon

451 acceptance.

References

453 Alström, P. et al., 2013. Multilocus phylogeny of the avian family Alaudidae (larks) reveals complex
454   morphological evolution, non-monophyletic genera and hidden species diversity. *Molecular Phylogenetics*
455   *and Evolution*, 69(3), pp.1043–1056.

456 Bachtrog, D. et al., 2011. Are all sex chromosomes created equal? *Trends in genetics: TIG*, 27(9), pp.350–357.
457   Available at: http://dx.doi.org/10.1016/j.tig.2011.05.005.

458 Bankevich, A. et al., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell
459   sequencing. *Journal of computational biology: a journal of computational molecular cell biology*, 19(5), pp.455–477.

460 Bolger, A.M., Lohse, M. & Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.
461   *Bioinformatics*, 30(15), pp.2114–2120.

462 Brooke, M. de L. et al., 2010. Widespread translocation from autosomes to sex chromosomes preserves
463   genetic variability in an endangered lark. *Journal of Molecular Evolution*, 70(3), pp.242–246.

464 Bulatova, N.S., 1981. A comparative karyological study of passerine bird.

465 Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic
466   analysis. *Molecular Biology and Evolution*, 17(4), pp.540–552.

467 Charlesworth, D. & Charlesworth, B., 1980. Sex differences in fitness and selection for centric fusions
468   between sex-chromosomes and autosomes. *Genetical research*, 35(2), pp.205–214. Available at:
469   http://www.ncbi.nlm.nih.gov/pubmed/6930353.

470 Chen, S. et al., 2014. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome
471   evolution and adaptation to a benthic lifestyle. *Nature Genetics*, 46(3), pp.253–260. Available at:
472   http://www.ncbi.nlm.nih.gov/pubmed/24487278.

473 Cortez, D. et al., 2014. Origins and functional evolution of Y chromosomes across mammals. *Nature*,
474   508(7497), pp.488–493. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24759410.

475 Cunningham, F. et al., 2015. Ensembl 2015. *Nucleic Acids Research*, 43, pp.D662–9. Available at:
476   http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4383879.

477 Danecek, P. et al., 2011. The variant call format and VCFtools. *Bioinformatics*, 27(15), pp.2156–2158.

478 Ellegren, H., 2010. Evolutionary stasis: the stable chromosomes of birds. *Trends in Ecology & Evolution*,
479   25(5), pp.283–291.

480 Ezaz, T. et al., 2006. Relationships between Vertebrate ZW and XY Sex Chromosome Systems. *Current*
481   *biology: CB*, 16(17), pp.R736–R743.

482 Ezaz, T., Srikulnath, K. & Graves, J.A.M., 2016. Origin of Amniote Sex Chromosomes: An Ancestral
483   Super-Sex Chromosome, or Common Requirements? *Journal of Heredity*, 108(1), pp.94–105.

484　Fisher, R.A., 1931. THE EVOLUTION OF DOMINANCE. *Biological Reviews*, 6(4), pp.345–368.
485　　　Available at: http://doi.wiley.com/10.1111/j.1469-185X.1931.tb01030.x.

486　Garrison, E. & Marth, G., 2012. Haplotype-based variant detection from short-read sequencing. *arxiv.org*.

487　Grabherr, M.G. et al., 2010. Genome-wide synteny through highly sensitive sequence alignment: Satsuma.
488　　　*Bioinformatics*, 26(9), pp.1145–1151.

489　Grützner, F. et al., 2004. In the platypus a meiotic chain of ten sex chromosomes shares genes with the
490　　　bird Z and mammal X chromosomes. *Nature*, 432(7019), pp.913–917.

491　Gurevich, A. et al., 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8),
492　　　pp.1072–1075.

493　Haldane, J.B.S., 1922. Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, 12(2), pp.101–
494　　　109.

495　Kawai, A. et al., 2008. The ZW sex chromosomes of *Gekko hokouensis* (Gekkonidae, Squamata) represent
496　　　highly conserved homology with those of avian species. *Chromosoma*, 118(1), pp.43–51.

497　Kielbasa, S.M. et al., 2011. Adaptive seeds tame genomic sequence comparison. *Genome Research*, 21(3),
498　　　pp.487–493.

499　Kitano, J. & Peichel, C.L., 2011. Turnover of sex chromosomes and speciation in fishes. *Environmental*
500　　　*Biology of Fishes*, 94(3), pp.549–558.

501　Kitano, J. et al., 2009. A role for a neo-sex chromosome in stickleback speciation. *Nature*, 461(7267),
502　　　pp.1079–1083.

503　Kohno, S. et al., 2015. Estrogen Receptor 1 (ESR1; ERα), not ESR2 (ERβ), Modulates Estrogen-Induced
504　　　Sex Reversal in the American Alligator, a Species With Temperature-Dependent Sex Determination.
505　　　*Endocrinology*, 156(5), pp.1887–1899.

506　Lahn, B.T. & Page, D.C., 1999. Four Evolutionary Strata on the Human X Chromosome. *Science*,
507　　　286(5441), pp.964–967.

508　Lewis, K.R. & John, B., 1968. The Chromosomal Basis of Sex Determination. In International Review of
509　　　Cytology. Elsevier, pp. 277–379.

510　Li, H. & Durbin, R., 2009. Fast and accurate short read alignment with Burrows–Wheeler transform.
511　　　*Bioinformatics*, 25(14), pp.1754–1760.

512　Li, H. et al., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), pp.2078–
513　　　2079.

514　Loytynoja, A., 2014. Phylogeny-aware alignment with PRANK. *Methods in molecular biology (Clifton, N.J.)*,
515　　　1079(Chapter 10), pp.155–170.

516　Montiel, E.E. et al., 2016. Cytogenetic Insights into the Evolution of Chromosomes and Sex
517　　　Determination Reveal Striking Homology of Turtle Sex Chromosomes to Amphibian Autosomes.
518　　　*Cytogenetic and Genome Research*, 148(4), pp.292–304.

519　Moyle, R.G. et al., 2016. Tectonic collision and uplift of Wallacea triggered the global songbird radiation.
520　　　7, 12709 EP –.

521 O'Meally, D. et al., 2012. Are some chromosomes particularly good at sex? Insights from amniotes.
522   *Chromosome Research*, 20(1), pp.7–19.

523 Pala, I., Hasselquist, D., et al., 2012. Patterns of molecular evolution of an avian neo-sex chromosome.
524   *Molecular Biology and Evolution*, 29(12), pp.3741–3754.

525 Pala, I., Naurin, S., et al., 2012. Evidence of a neo-sex chromosome in birds. *Heredity*, 108(3), pp.264–272.
526   Available at: http://dx.doi.org/10.1038/hdy.2011.70.

527 Pennell, M. W. et al., 2015. Y Fuse? Sex Chromosome Fusions in Fishes and Reptiles. *PLOS Genetics*,
528   *11*(5), e1005237. Available at: http://doi.org/10.1371/journal.pgen.1005237

529 Pokorná, M., Altmanová, M. & Kratochvíl, L., 2014. Multiple sex chromosomes in the light of female
530   meiotic drive in amniote vertebrates. *Chromosome Research*, 22(1), pp.35–44.

531 Quinlan, A.R. & Hall, I.M., 2010. BEDTools: a flexible suite of utilities for comparing genomic features.
532   *Bioinformatics*, 26(6), pp.841–842.

533 Rice, W.R., 1994. Degeneration of a nonrecombining chromosome. *Science*, 263(5144), pp.230–232.

534 Rice, W.R. & Gaines, S.D., 1994. Extending nondirectional heterogeneity tests to evaluate simply ordered
535   alternative hypotheses. *Proceedings of the National Academy of Sciences*, 91(1), pp.225–226.

536 Rovatsos, M. et al., 2016. Mammalian X homolog acts as sex chromosome in lacertid lizards. *Heredity*,
537   *117*(1), 8–13. Available at: http://doi.org/10.1038/hdy.2016.18

538 Ross, J.A. et al., 2009. Turnover of Sex Chromosomes in the Stickleback Fishes (Gasterosteidae). *PLOS*
539   *Genetics*, 5(2), p.e1000391.

540 Ross, M.T et al. 2005. The DNA sequence of the human X chromosome. Nature, 434(7031), 325–337.
541 http://doi.org/10.1038/nature03440
542
543 Sambrock, J. & W Russel, D, 2001. *Molecular Cloning: A Laboratory Manual (3rd edition)*. Cold Spring Harbor
544 Laboratory Press.
545
546 Sigeman, H. et al., 2018. Insights into Avian Incomplete Dosage Compensation: Sex-Biased Gene
547   Expression Coevolves with Sex Chromosome Degeneration in the Common Whitethroat. *Genes*, 9(8),
548   p.373.

549 Smith, C.A. et al., 2009. The avian Z-linked gene DMRT1 is required for male sex determination in the
550   chicken. *Nature*, 461(7261), pp.267–271.

551 Tange, O., 2018. *GNU Parallel 2018*, Lulu.com.

552 Vassart, M., Séguéla, A. & Hayes, H., 1995. Chromosomal evolution in gazelles. *Journal of Heredity*, 86(3),
553   pp.216–227.

554 Warren, W.C. et al., 2010. The genome of a songbird. *Nature*, 464(7289), pp.757–762.

555 Yang, Z., 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution*,
556   24(8), pp.1586–1591.

557 Yoshida, K. et al., 2014. Sex Chromosome Turnover Contributes to Genomic Divergence between
558   Incipient Stickleback Species J. Zhang, ed. *PLOS Genetics*, 10(3), p.e1004223.

559 Zamani, N. et al., 2014. A universal genomic coordinate translator for comparative genomics. *BMC*
560   *Bioinformatics*, 15(1), p.227.

561 Zhou, Q. et al., 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science*,
562   346(6215), pp.1246338–1246338. Available at:
563   http://www.sciencemag.org/cgi/doi/10.1126/science.1246338.