1
2
3 **RATES OF KARYOTYPIC EVOLUTION IN ESTRILDID FINCHES DIFFER BETWEEN**
4   **ISLAND AND CONTINENTAL CLADES**
5

6 Daniel M. Hooper[1,2] and Trevor D. Price[3]

7

8 [1]Commitee on Evolutionary Biology, University of Chicago, Chicago, Illinois 60637

9   [2] E-mail: dhooper1@uchicago.edu

10 [3]Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637

11

12

13   Sunday, January 18, 2015

14

15

16 Running head: Chromosome inversions in finches

17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34

35   Reasons why chromosomal rearrangements spread to fixation and frequently distinguish

36   related taxa remain poorly understood. We used cytological descriptions of karyotype to

37   identify large pericentric inversions between species of Estrildid finches (family

38   *Estrildidae*) and a time-dated phylogeny to assess the genomic, geographic, and

39   phylogenetic context of karyotype evolution in this group. Inversions between finch species

40   fixed at an average rate of one every 2.26 My. Inversions were twice as likely to fix on the

41   sex chromosomes compared to the autosomes, possibly a result of their repeat density, and

42   inversion fixation rate for all chromosomes scales with range size. Alternative mutagenic

43   input explanations are not supported, as the number of inversions on a chromosome does

44   not correlate with its length or map size. Inversions have fixed 3.3× faster in three

45   continental clades than in two island chain clades, and fixation rate correlates with both

46   range size and the number of sympatric species pairs. These results point to adaptation as

47   the dominant mechanism driving fixation and suggest a role for gene flow in karyotype

48   divergence. A review shows that the rapid karyotype evolution observed in the Estrildid

49   finches appears to be more general across birds, and by implication other understudied

50   taxa.

51

54

55

56

57

58

59   Chromosome inversions are often fixed between closely related species as well as found

60   segregating within species (Coyne and Orr 2004; Hoffmann and Rieseberg 2008; Faria and

61   Navarro 2010). Once established, inversions can foster divergence between populations

62   (Anderson et al. 2005; Lowry and Willis 2010; Jones et al. 2012) and aid in the speciation

63   process (Rieseberg et al. 2001; Noor et al. 2001; Brown et al. 2004; Joron et al. 2011; Ayala

64   et al. 2013; Fishman et al. 2013; Poelstra et al. 2014) by creating barriers to gene flow

65   through the suppression of recombination and/or the induction of structural

66   underdominance in heterokaryotypes. It remains unclear why rearrangements evolve so

67   frequently in some taxa and by what mechanisms, especially as new inversions are often

68   thought to have deleterious fitness effects due to structural underdominance (Stebbins

69   1958; White 1969; 1978; King 1987; Rieseberg 2001).

70   Mechanisms proposed to explain the fixation of novel underdominant

71   rearrangements involve both genetic drift and natural selection. As described by Lande

72   (1979), drift within a deme may counter selection and raise the rearrangement above a

73   frequency of 50%, whereupon selection drives it to fixation. Fixation rate in a species is

74   affected by deme size, but not the population size of the species itself. The reason

75   population size does not affect fixation rate is that, after becoming established in one deme,

76   the probability that a rearrangement will fix is equal to the inverse of the total number of

77   demes, but the more demes there are the greater the chance that an inversion becomes

78   established in at least one deme. As in classic models of genetic drift (Kimura 1962) the two

79   effects cancel out and the fixation rate for a species is approximately independent of its

80   population size, although this result depends on assumptions about population structure

81   and the deme where the rearrangement originated (Lande 1979; 1985; Hedrick 1981;

82   Walsh 1982; Spirito 1998). In contrast to the spread of an underdominant rearrangement,

83   if a rearrangement is inherently deleterious its fixation probability within a species should

84   decrease with population size and only realistically occur when population sizes are small

85   (Kimura 1962; Ohta 1972; Whitlock et al. 2004).

86   Chromosome rearrangements may also become established entirely by positive

87   selection. In this case, a rearrangement can spread to fixation if (1) its breakpoints

88   favorably alter gene expression (Wesley and Eanes 1994), (2) it increases linkage between

89   epistatically interacting sets of genes (Dobzhansky 1951), (3) it increases linkage between

90   locally adapted alleles within a population experiencing gene flow from a divergently

91   adapted population (Charlesworth and Charlesworth 1979; Kirkpatrick and Barton 2006;

92   Feder et al. 2011). In this scenario, gene flow plays a creative role in karyotype evolution

93   because of the selective advantage in keeping locally adapted alleles together. Adaptive

94   models include those external to the individual, as implied in the mechanisms described

95   above, as well as those resulting from genomic conflicts. For example, meiotic drive may

96   favor the spread of a rearrangement if it is associated with a set of interacting alleles that

97   together affect segregation distortion (White 1978; King 1993). Whenever adaptive

98   mechanisms operate, the fixation rate of a novel rearrangement should increase with

99   population size, as the process is mutation-limited. Here, we evaluate fixation rates of large

100   chromosomal inversions in a family of birds, and consider the possible roles of population

101   size and gene flow.

102   Birds have long been used as models of speciation and are perhaps the best-studied

103   group with respect to how behavior and ecology contribute to population divergence (Price

104   2008; Grant and Grant 2011) but little attention has been given to the possible role of

105   chromosome rearrangements in bird speciation (Price 2008; Ellegren 2013). This may be

106   because avian genomes are often considered to be highly conserved, and inter-

107   chromosomal rearrangements such as fusions, fissions, and translocations between species

108   appear to be rare (Takagi and Sasaki 1974; Griffin et al. 2007; Ellegren 2010; 2013; Zhang

109   et al. 2014). For example, diploid chromosome number (2N) for birds typically varies

110   between 76 and 80. However, in some clades 2N is much more variable, notably for finches

111   in the genus *Pytilia* (Christidis 1983), parrots (Nanda et al. 2007), birds of prey (Bed'Hom

112   et al. 2003; de Oliveira et al. 2005; Nanda et al. 2006; Nishida et al. 2008; 2014), and one

113   shorebird, the stone curlew (*Burhinus oedicunemus*; Nie et al. 2009; Hansmann et al. 2009).

114   In contrast to the general conservation of chromosome number and lack of inter-

115   chromosomal rearrangements, the cytological literature suggests that intra-chromosomal

116   rearrangements, notably inversions, occur more frequently (reviewed in Shields 1982;

117   Christidis 1990; Price 2008, ch. 16; Völker et al. 2010; Skinner and Griffin 2011; Lithgow et

118   al. 2014, and see Tables 3 and S6 of this paper). These findings are supported by recent

119   comparative genomics studies that report large numbers of inversions (10s to 100s)

120   between distantly related species (Stapley et al. 2008; Hansson et al. 2009; Aslam et al.

121   2010; Backström et al. 2010; Kawakami et al. 2014; Zhang et al. 2014). After whole genome

122   alignment Kawakami et al. (2014) identified 140 inversions at a median size of ~2 Mb that

123   distinguish zebra finch (*Taeniopygia guttata*) from collared flycatcher (*Ficedula albicollis*),

124   two species that last shared a common ancestor ~30 Ma (divergence date inferred from

125   Price et al. 2014). From this, we estimate a fixation rate as fast as one every 0.5 My along

126   the lineage leading to the zebra finch.

127   Typically a fifth of avian chromosomes are large in size (macro-chromosomes 160-

128   60 Mb) and easily distinguishable cytogenetically. The remaining chromosomes are smaller

129   (micro-chromosomes <40 Mb) and the smallest of which are often indistinguishable using

130   standard cytological approaches (Masabanda et al. 2004; Ellegren 2013). Here, we examine

131   the fixation of large cytologically detectable inversions across 32 species of Estrildid

132   finches (order *Passeriformes*). We use published cytological analyses of the Estrildids,

133   which have identified inversion differences between species and polymorphisms within

134   species (Prasad & Patnaik 1977; Christidis 1983; 1986a; 1986b; 1987; Itoh & Arnold 2005).

135   2N is either 76 or 78 for all Estrildid finches examined except one (the red-winged pytilia,

136   *Pytilia phoenicoptera*, 2N = 56). Seven pairs including the Z can be considered macro-

137   chromosomes and the remaining pairs micro-chromosomes. The Z chromosome is the 4th

138   largest while the W is typically 8th in size. The degree of karyotype similarity between

139   species varies considerably and some genera appear to have more labile karyotypes than

140   others (Christidis 1986a; 1986b). A preliminary analysis using a phylogeny for a subset of

141   the karyotyped Estrildid species estimated that the fixation of cytologically detectable

142   chromosomal inversions occurred every 2.7 My along a lineage (Price 2008, pp. 386-388).

143   We expand on this analysis using a time-dated tree to ask how variation in the rate of

144   inversion fixation differs among chromosomes and is associated with demographic and

145   geographic differences between species.

146

147   *Methods*

148   The family *Estrildidae* contains approximately 140 species distributed across the Old World

149   tropics and southern hemisphere temperate zone (Goodwin 1982; Sorenson et al. 2004).

150   The Estrildid finches originated in Africa (Sorenson et al. 2004; Sorenson *pers. comm.*) and

151   their subsequent dispersion and diversification across the Old World have produced 7

152   discrete radiations that vary extensively in the average population size and degree of

153   sympatry between member species (Goodwin 1982).

154

155   **IDENTIFYING INVERSIONS**

156   We use the data of Christidis (1983; 1986a; 1986b; 1987) supplemented by others (Hirschi

157   et al. 1972; Ray-Chaudhuri 1976; Ray-Chaudhuri 1976; Prasad and Patnaik 1977) who

158   describe gross karyotype structure (Tables S1 and S2). These authors used Giemsa staining

159   and C- and G-banding techniques to document large inversions, both pericentric (those

160   encompassing the centromere) and paracentric (not encompassing the centromere).

161   Sample size varied across species (maximum of 20 individuals for *Bathilda ruficauda* and

162   just one for three species) with an average of 5.7 karyotyped individuals per taxon (Table

163   S1).

164   We converted centromere position and banding pattern for the 6 autosomal macro-

165   chromosomes, the 5 largest autosomal micro-chromosomes, and both sex chromosomes

166   into character state data for each species (Table S2). Using published figures, we identified

167   homologous chromosomes between species by chromosome-diagnostic banding patterns

168   and scored them for approximate centromere position (i.e., whether they were

169   metacentric, sub-metacentric, sub-telocentric, or telocentric), following the naming

170   conventions established by Levan et al. (1964). We treated chromosome conformations as

171   distinct when species shared the same general classification (e.g., both were sub-

172   metacentric) but the author of the paper noted that the banding pattern flanking the

173   centromere differed. We identified pericentric inversions between taxa from changes in

174   both the location of the centromere and the orientation of the banding pattern immediately

175   flanking the centromere. We identified paracentric inversions by re-orientation of banding

176   pattern alone. However, as we detected only three paracentric inversions in our dataset

177   and this is likely to be a significant underestimate of their true diversity (Kawakami et al.

178   2014; Zhang et al. 2014), we therefore only included pericentric inversions in our analyses.

179   Centromere repositioning can result from processes other than pericentric

180   inversions, such as the redistribution of heterochromatin (Krasikova et al. 2009; Zlotina et

181   al. 2012) and the evolution of neo-centromeres (Marshall et al. 2008). We are able to

182   exclude these alternative explanations because centromere repositioning was matched

183   with inversion of proximal banding patterns.

184   Five species had polymorphic chromosome conformations due to pericentric

185   inversions found segregating within the individuals karyotyped (Table S2). We did not

186   include them in our analyses because we are only interested in the drivers of inversion

187   fixation. As such, an inversion on a polymorphic chromosome was only counted as a fixed

188   difference between species if the ancestral conformation, determined by Bayesian

189   approach in SIMMAP v1.5 (Bollback 2006) – see below, was not one of the forms found

190   segregating.

191

192   **PHYLOGENETIC ANALYSES**

193   Several phylogenetic surveys are available (Sorenson and Payne 2001; Sorenson et al.

194   2004; Arnaiz-Villena et al. 2009) but there is no complete published phylogeny of the

195   *Estrildidae*. We built a time-dated phylogeny for the 32 karyotyped finches using

196   mitochondrial data from GenBank (Table S3). We dated the tree using the divergence time

197   between the families *Estrildidae* and *Ploceidae* inferred from the multiple fossil-calibrated

198   passerine tree of Price et al. (2014), setting a uniform prior of 17.5-22.1 Ma, based on the

199   95% confidence limits from the posterior distribution of that tree. Phylogenetic analyses

200   were conducted using BEAST v1.8.0 (Drummond and Rambaut 2007). Data from the

201   mitochondrial loci *cytb* (1143bp), *ND2* (1041bp), and *ND6*+control region (1100bp) were

202   partitioned by locus, each with its own uncorrelated lognormal relaxed clock, and GTR + Γ

203   + I model of sequence evolution (estimated to be the optimal model for each locus using

204   Jmodeltest v0.1.1; Posada 2008). Algorithms were run for 50 million generations and

205   sampled every 5,000 for a total of 10,000 trees of which the first 1,000 were discarded as

206   burn in. We assessed run length and appropriate sampling for each parameter using Tracer

207   v1.5 (Rambaut and Drummond 2007). using TreeAnnotator v1.7.2 (Drummond and

208   Rambaut 2007), we extracted the maximum clade credibility tree, with associated

209   confidence intervals for median node heights.

210

211   **INVERSION FIXATION ANALYSES**

212   We estimated the ancestral centromere position (up to 4 possible states: metacentric, sub-

213   metacentric, sub-telocentric, or telocentric) for each chromosome at each node in the tree

214   using the Bayesian approach in SIMMAP v1.5 (Bollback 2006). To account for phylogenetic

215   uncertainty, we simulated over 1,000 randomly drawn trees from the BEAST output post

216   burn in. We obtained the posterior probability estimate for each ancestral centromere

217   position for each chromosome at every node. Inversions were inferred to have occurred

218   upon branches where the karyotype of an internal node differed from subsequent nodes or

219   the tips and was supported by a posterior probability, p > 0.75. The inferred number of

220   inversions per chromosome was concordant with results from reconstructions based on a

221   maximum likelihood model in Mesquite v2.7.5 (Maddison and Maddison 2001; results not

222   shown).

223   We used ancestral karyotype estimates to calculate the total number of inferred

224   inversions that have occurred on each chromosome separately. Here, we had to exclude

225   three chromosomes from a single species (*Pytilia phoenicoptera*) that have fused and thus

226   precluded identification of chromosome homology (Table S2; excluding this species

227   entirely would not affect conclusions). We next assessed the degree to which the genomic

228   distribution of inversions could be explained by mutagenic input. Under the assumption

229   that the inversion mutation rate is constant per DNA base, the probability of a new

230   inversion on a given chromosome or polytene chromosome arm should be proportional to

231   the chromosome's physical size. We also assessed the relevance of a chromosome's map

232   length and GC content based on an alternative process-driven assumption, conditioned on

233   the fact that chromosomal rearrangements are mediated by double-stranded meiotic

234   breaks, that predicts inversions should be proportional to cross-overs (Baudat and de

235   Massy 2007; de Massy 2013). Given that at least one cross-over per chromosome arm

236   appears to be required for proper pairing of homologous chromosomes during meiosis, we

237   assessed whether the fixation rate of inversions was constant per chromosome. Finally, as

238   chromosome breakpoints are often enriched in repeat dense regions (Skinner and Griffin

239   2011; Kawakami et al. 2014), we assessed whether variation in the number of inversions

240   per chromosome could be explained by variation in repeat content. We estimated

241   chromosome physical size and GC content from the zebra finch genome assembly (Warren

242   et al. 2010), repeat content per chromosome from a RepeatMasker annotation of the zebra

243   finch genome (Smit et al. 1996-2010; http://www.repeatmasker.org), and map distance

244   from two independent zebra finch linkage maps (Stapley et al. 2008; Backström et al.

245   2010). Hence, we are assuming that the general features of the Estrildid finch genome and

246   recombination landscape (chromosome size, GC content, map distances, and repeat

247   content) are conserved across species with respect to the zebra finch. Comparative studies

248   indicate that chromosome size, GC content, and repeat content varies little even between

249   distantly related avian species (Ellegren 2013; Kawakami et al. 2014; Poelstra et al. 2014;

250   Zhang et al. 2014) and that the recombination landscape has a phylogenetic signal,

251   suggesting our extension of genomic parameters from the zebra finch to this confamilial set

252   of Estrildid finches is warranted (Dumont and Payseur 2008; 2011; Smukowski and Noor

253   2011). We correlated the number of inversions with chromosome size and map distance,

254   using each chromosome as a replicate. We subsequently compared alternative hypotheses

255   using multiple regression.

256   In order to examine the influence of demography and geography on the rate of

257   fixation of inversions, we assigned 27 of the 32 Estrildid finches that have cytological data

258   into 5 clades comprising distinct geographic radiations with complete posterior probability

259   support (Figure 1, Table 1, and Figures S1-S5). We focus on clades as they represent

260   distinct monophyletic groupings at a deep timeline. Five species were not assigned to any

261   clade because they were either the only species with cytological data belonging to an

262   independent geographic radiation or were singleton species that did not group with other

263   Estrildid species in their region. For example, the black-and-white mannakin (*Spermestes*

264   *bicolor*) is the single species with cytological data belonging to a clade that radiated after

265   re-colonizing Africa from Asia, precluding an assessment of inversion evolution within this

266   interesting group. M. Sorenson (*pers. comm.)* provided information on the total number and

267   identity of Estrildid species within each clade (i.e. including species without cytological

268   data), based on unpublished phylogenetic results for the *Estrildidae* (Table 1). The total

269   number of inversions fixed in each clade was summed across chromosomes. We estimated

270   the fixation rate for the clade as the total number of inversions divided by the total branch

271   length connecting all karyotyped species within the clade. We also estimated the

272   pericentric inversion fixation rate, hereafter referred to as the inversion fixation rate, as

273   the total number of pericentric inversions divided by the number of nodes, as a measure of

274   the minimum number of speciation events. This may scale with the number of

275   opportunities for secondary contact between partially reproductively isolated forms

276   between which gene flow may promote the spread of an inversion (Kirkpatrick and Barton

277   2006).

278   To assess the relative influences of demography and geography, we extracted

279   species' range sizes and pairwise range overlaps from natureserve.org using the programs

280   Sp and PBSmapping in R (R Core Team 2014; Tables S4 and S5). We first scored the

281   influence of region (continental *vs.* island taxa) for each clade as the proportion of species

282   whose ranges were predominately continental. Species were scored as continental or

283   island depending on where the majority of their range lay (Tables 2 and S4). Of the 90

284   species considered within our 5 clades, only 3 had ranges that were between 25-75%

285   continental (Table S4). Classifying these species alternately as continental or as island did

286   not alter the results in any way.

287 Across the *Estrildidae*, range size varies by over 3 orders of magnitude from the red-

288 billed firefinch (*Lagonosticta senegala*), the most widely distributed species ($1x10^7$ km^2), to

289 the red-headed parrotfinch (*Erythrura cyaneovirens*), the most narrowly distributed

290 species (3000 km^2, Table S4). We use range size as a proxy for population size and assigned

291 a range size score to each clade based on the average range size of all species within it.

292 Range size generally correlates with nucleotide diversity within species, which in turn is

293 correlated with effective population size (Nevo et al. 1984; Cole 2003; Leffler et al. 2012).

294 In the Estrildid finches, estimates of nucleotide diversity are approximately an order of

295 magnitude greater for zebra finch populations upon the Australian continent compared to

296 zebra finch populations upon the Lesser Sunda Islands, consistent with the two order of

297 magnitude difference in their range size (Balakrishnan and Edwards 2009). It is worth

298 noting that the fixation probability of an adaptive mutation depends on variance in family

299 size (Peischl and Kirkpatrick 2012) and not N_e per se. Variance in family size is only one of

300 several factors that may cause N_e to differ from N and we have no reason to suspect that

301 variance in family size varies greatly across Estrildid species.

302 We define the range overlap of species A with species B as the proportion of species

303 A's range that is shared with B. Species pairs within clades were scored as sympatric if

304 their average range overlap was greater than 15% (Table S5). Each clade was assigned a

305 sympatry score as the proportion of all possible species pairs that were sympatric. Scoring

306 species' distributions as sympatric with an average of 10% or 3% range overlap did not

307 qualitatively change the results for the impact of sympatry on inversion fixation rate.

308 For each factor (region, range size, and range overlap), we calculated the correlation

309 with inversion fixation rate using clade as the replicate. By studying at the clade level (N =

310   5), we eliminate possibilities for pseudoreplication as far as possible, but the tests are

311   conservative. We fit a multiple linear regression model using fixation rate as the response

312   variable to determine the extent to which variation in karyotype evolution between

313   Estrildid finch clades is affected by demographic and geographic differences. We focus on

314   the rate of inversion fixation (number of inversions / total clade branch length connecting

315   karyotyped species) but also consider number of inversions as a function of the total

316   number of nodes in a clade. Neither of the two alternative methods of defining inversion

317   fixation rate correlates with total branch length (both p > 0.1).

318

319   *Results*

320   The time-calibrated phylogeny for those Estrildid finches with cytological data is presented

321   in Figure 1. Our topology matches that of Sorenson et al. (2004) and an unpublished

322   topology of all Estrildid species (Sorenson *pers. comm.*) for all shared nodes. Cytological

323   sampling of each of the 5 fully-supported clades, as determined from a phylogeny of all

324   Estrildids (Sorenson *pers. comm.*), averaged 38% (minimum of 13% in the firefinches,

325   pytilias, and waxbills – clade D – and maximum of 86% in the grassfinches – clade A, Table

326   1).

327   The average rate of pericentric inversion fixation across all 32 karyotyped species is

328   one inversion fixed every 2.26 million years (199 My total tree branch length / 88 inferred

329   inversions). The rate varies considerably among clades (Table 1). For example, gross

330   karyotype has changed as fast as one inversion fixed every 175,000 years of evolution (5.72

331   inversions per million years, Figure 1, clade A) within the Australian grassfinches but has

332   remained identical for over 11 million years of evolution between two species of Indo-

333   Malayan *Lonchura* munias (clade B).

334   Overall, the rate of inversion fixation differs 13× between the fastest and slowest

335   evolving chromosomes (Table 2). The number of inversions across the 11 autosomes

336   departs from a Poisson distribution (Goodness of fit test, χ^2_{10} = 93.4) implying that

337   inversion fixation does not occur at equal rates per chromosome. Across the autosomes, the

338   number of inversions per chromosome is not significantly associated with physical size

339   (Pearson's r = 0.29, p > 0.1), GC content (r = -0.61, p > 0.1), or map length regardless of

340   linkage map used (Stapley et al. 2008: r = 0.38, p > 0.1; Backström et al. 2010: r = 0.3, p >

341   0.1). Inclusion of the Z further weakened any association.

342   The two sex chromosomes have fixed inversions at a rate on average twice as fast as

343   the autosomes (two sample *t*-tests, using the autosomes as replicates: Z chromosome: t_{10} =

344   7.1, p < 0.01; W chromosome: t_{10} = 5.0, p <0.01). The only significant difference between

345   the autosomes and the Z with regard to the parameters we examined is that the Z

346   chromosome has an order of magnitude greater density of repeats than observed average

347   of the autosomes (two sample *t*-test: t_{10} = 37.4, p < 0.01; Table 2). There is currently no

348   comparable estimate for repeat content on the W. The W chromosome, however, stands out

349   when comparing chromosomes by the number of inversions per base pair (Table 2). We

350   observe one inversion fixed for every 2.5 Mb on the W, which is over twice as many as on

351   the Z chromosome or the closest autosome TGU5 (both one for every 5.7 Mb).

352   The fixation rate for inversions is on average 3.3× higher in the three continental

353   than the two island clades (*t*-test: t_3 = 16.9, p < 0.01) and is significantly correlated with the

354   proportion of species within clades that have continental distributions (N =5, r = 0.99, p <

355   0.001). Continental taxa have on average larger ranges and a higher percentage of

356   sympatric species than island taxa (Figures S1-S5). Both average range size (N =5, r = 0.91,

357   one-tailed p = 0.02; Figure 2) and proportion of sympatric pairs (N = 5, r = 0.92, one-tailed

358   p = 0.01) are significantly correlated with inversion fixation rate. In a multiple regression

359   on inversion fixation rate, the partial regression coefficients of range size and proportion of

360   sympatric pairs are both significant (range size: p = 0.042; proportion sympatric: p =

361   0.037), suggesting that both contribute independently.

362   Analyses conducted using inversion fixation rate alternatively defined as the

363   number of inversions within each clade divided by the number of nodes in that clade

364   yielded results similar to those reported above, but they were not as strong and not

365   significant at the P < 0.05 level.

366

367   *Discussion*

368   Previous cytological analyses of karyotype structure in Estrildid finches revealed multiple

369   fixed inversion differences between species (Hirschi et al. 1972; Ray-Chaudhuri 1976; Ray-

370   Chaudhuri 1976; Prasad and Patnaik 1977; Christidis 1983; 1986a; 1986b; 1987). We

371   found these inversions have accumulated disproportionately on the sex chromosomes and

372   are much more common within continentally distributed clades, which tend to have larger

373   ranges and a higher proportion of sympatric species. Assuming range size correlates with

374   population size, this is not in accord with the expectations for the fixation of

375   underdominant or deleterious rearrangements because these processes should be

376   independent of (Lande 1979; 1985; Walsh 1982; Spirito 1998) or negatively correlated

377   with (Kimura 1962) a species' population size. Instead, the correlation with range size

378   suggests that inversions have fixed by positive selection (Whitlock et al. 2004; Vicoso and

379   Charlesworth 2009; Mank et al. 2010b). The positive relationship between range size and

380   inversion fixation rate is present on the autosomes combined (across the 5 clades, r = 0.82)

381   and each of the sex chromosomes (Z: r = 0.44 and W: r = 0.84)

382   Overall, our results highlight the rapid rate of karyotype evolution in Estrildid

383   finches; with one pericentric inversion fixed every 2.26 My on average. This value is a

384   minimum estimate of the true rate of chromosomal rearrangement. First, many paracentric

385   inversions may have been missed because of the limited number of C- and G- bands per

386   chromosome arm required to infer them and, second, small inversions that cannot be

387   detected cytologically are likely to occur frequently (Kawakami et al. 2014). Paracentric

388   inversions may be less structurally underdominant than pericentric inversions because; in

389   some groups (e.g. *Drosophila*), crossing over within them produces acentric and dicentric

390   recombinants that are removed to the polar bodies during meiosis (i.e. they do not affect

391   female fertility). Whether similar mechanisms to affect the degree of structural

392   underdominance between paracentric and pericentric inversions exist in birds is not

393   known. Small inversions may be subject to similar selective pressures as those examined

394   here but this awaits genomic analysis (e.g. Kunte et al. 2014; Kawakami et al. 2014;

395   Poelstra el al. 2014).

396   We find that degree of sympatry between species is positively associated with the

397   rate of inversion accumulation, even after range size is accounted for, which may be

398   expected if gene flow between incipient species has contributed to inversion fixation, as in

399   models where rearrangements that capture locally adapted alleles are favored (Kirkpatrick

400   and Barton 2006; Feder et al. 2011). We first consider the genomic distribution of

401   inversions and then the demographic and geographic context of inversion fixation.

402   We find mixed support for the idea that mutational input has influenced the

403   genomic distribution of chromosome inversions. Chromosome breakpoints are often

404   located within regions that are repeat dense and this appears to be supported by the

405   enrichment of inversions we observe on the sex chromosomes. The repeat content of the Z

406   chromosome is an order of magnitude greater than the autosomal average and has the

407   greatest number of inversions fixed of any chromosome – perhaps suggesting the Z has a

408   greater structural mutation rate. While we do not know the repeat density of the W

409   chromosome, it is believed to be repeat rich due to the reduced power of selection for

410   fixing DNA replication errors (Dalloul et al. 2010; Völker et al. 2010; Kawakami et al. 2014).

411   Alternative mutagenic explanations for the genomic distribution of inversions, however,

412   bear little support. First, chromosome size is not correlated with inversion fixation rate,

413   which is the naïve expectation if structural mutations have a constant rate per base pair.

414   Second, neither a chromosomes map length nor its GC content are correlated with its rate

415   of inversion fixation suggesting an increased number of cross-overs per se does not

416   necessarily result in an increased rate of inversion fixation. While inversion breakpoints

417   may be more prevalent in areas of high cross-over (Skinner and Griffin 2011; Kawakami et

418   al. 2014) and inversions originate from errors during meiotic crossing-over, in the Estrildid

419   finches the fixation rate of inversions cannot be explained by differences between the

420   recombination landscapes of chromosomes alone.

421   These results contrast with a recent comparative genomic survey of inversion

422   differences between the zebra finch and the collared flycatcher, which found a positive

423   correlation between chromosome size and number of inversions (Kawakami et al. 2014).

424   The difference may reflect the level of resolution possible between studies. Our study only

425   considered pericentric inversions because of the limitations of the cytological methods

426   used to detect them, while the majority of the inversions found by Kawakami et al. (2014)

427   are small (median size of 2.62 Mb and 0.78 Mb in the lineages leading to zebra finch and

428   collared flycatcher, respectively). Smaller rearrangements potentially come with both

429   fewer deleterious effects and smaller selective advantages, resulting in a more even

430   distribution across the genome. Our results are perhaps striking because, in contrast to

431   small inversions, large pericentric inversions seem more likely to carry an intrinsic

432   selective disadvantage due to underdominance and thus suggest that selection was critical

433   in driving their fixation.

434   A strong result in our dataset is that inversions accumulated twice as rapidly on the

435   sex chromosomes as on the autosomes. While meiotic drive on the sex chromosomes (i.e.

436   sex chromosome drive) is often associated with the establishment of inversion

437   polymorphisms it is unlikely to lead to their fixation because of the deleterious effects of

438   high sex ratio skew (Jaenike 2001). Rather, the scaling of inversion fixation rate on both the

439   Z and W chromosomes with range size suggests positive selection. On the Z chromosome,

440   immediate exposure of recessive mutations to selection and the preferential accumulation

441   of sexually antagonistic genes may make for generally faster rates of adaptive evolution

442   (Charlesworth et al. 1987). This "faster-Z effect" is well documented in birds yet has been

443   attributed to genetic drift rather than selection (Mank and Ellegren 2007; Mank et al.

444   2010a; Ellegren 2009; Yan et al. 2010; Zhang et al. 2014). However, explanations crediting

445   the predominant role of drift towards the rapid rate of functional divergence on the Z are

446   flawed as they rely on large assumptions regarding the evolution of sex-biased gene

447   expression and avian effective population sizes (Mank and Ellegren 2007; Mank et al.

448   2010a). As such, the relative contributions of selection and drift towards the faster-Z effect

449   have yet to be properly examined in birds. Regarding the relative rate of inversion fixation,

450   because Z-linked genes diverge in function more rapidly than autosomal genes, at any point

451   in time before reproductive isolation is complete an inversion on the Z chromosome should

452   be more likely to capture two or more alleles locally adapted—either to that population's

453   habitat or genomic background—than an inversion on an autosome. Thus, inversions on

454   the Z may be more strongly selected for if gene flow is an important mechanism driving

455   their selective advantage.

456   The W chromosome has more inversions per Mb than any other chromosome we

457   consider and a fixation rate strongly correlated with range size (r = 0.84, p = 0.038). This is

458   surprising as much of the W is devoid of genic content (in the chicken, of the 1,000 active

459   genes on the Z only 10-100 are thought to remain active on the W chromosome, Chen et al.

460   2012; Ayers et al. 2013), suggesting few adaptive advantages for a recombination modifier.

461   Under positive selection, the fixation rate is $\sim2N\mu s$, where N is population size, μ is the

462   mutation rate, and s the selective advantage of the heterozygote (Haldane 1927; Peischl

463   and Kirkpatrick 2012). This formula assumes that variance in family size is Poisson. Among

464   the three parameters (N, μ, and s), μ and s may help explain the W chromosome's relatively

465   high fixation rate. We can dismiss an explanation based solely on N, as the population size

466   of the W is ¼ that of the autosomes, which should result in a lower fixation rate. While the

467   W chromosome's mutation rate at the nucleotide level is estimated to be significantly lower

468   than on the autosomes, due to its female-restricted mechanism of germ-line transmission

469 (Ellegren 2013), the structural mutation rate (μ) on the W may be greater because of its

470 elevated repeat content (Dalloul et al. 2010; Völker et al. 2010; Kawakami et al. 2014).

471 While the idea that the high number of inversions on the W is due to its repeat density is

472 plausible, chromosome breakpoints are also more often to be found located in areas that

473 have high recombination rates and a high GC content—both of which are lower on the W

474 than the autosomes (Völker et al. 2010; Kawakami et al. 2014; Ellegren 2013; Graves

475 2014). Finally, selection (s) is a composite variable summing across both positive and

476 deleterious effects on fitness. Possibly, s is higher on the W chromosome not because

477 inversions have a greater selective advantage than on the autosomes but rather because

478 the costs of structural rearrangements on the W are reduced.

479   Species ranges are on average 3.6× larger in continental versus island chain clades

480 and this difference is matched by a 3.3× faster rate of inversion fixation. Adaptive models

481 predict this association, but models where drift operates on structurally underdominant

482 deleterious rearrangements do not (Lande 1979; 1985). The usual explanation for the

483 correlation of population size with fixation rate is that the number of rearrangements (or

484 mutations) arising each generation positively correlates with population size (Kimura

485 1962; Whitlock et al. 2004). However, our analyses across chromosomes based on physical

486 and map size, as described above, argue for an important role for selection beyond raw

487 mutational input. Besides coming into contact more regularly, enabling gene flow that

488 promotes the spread of inversions in adaptive models, species on continents with larger

489 ranges are also more likely to encompass a wider variety of landscapes and climates across

490 which selection might favor alternate allelic combinations between populations connected

491 by gene flow. These are the theoretical starting conditions under which a novel

492   rearrangement that captures sets of locally adapted alleles may spread to fixation

493   (Kirkpatrick and Barton 2006; Feder et al. 2011).

494   Both the number of inversions and the rate of inversion fixation appear to be

495   correlated with the proportion of sympatric species within Estrildid clades even after

496   accounting for the effects of time and range size. This could be construed as support for

497   inversion fixation as a consequence of gene flow between partially differentiated

498   populations, if extant sympatry between closely related taxa does indeed reflect parapatric

499   divergence. In an extension of the local adaptation model of Kirkpatrick and Barton (2006),

500   chromosome rearrangements are selected to fix between populations upon secondary

501   contact when reproductive isolation is partial but incomplete (Kirkpatrick 2010). Genome-

502   scale sequence data from recently diverged taxa suggests that gene flow between incipient

503   species may occur regularly during the speciation process (e.g., Jones et al. 2012; Ayala et

504   al. 2013; Lowery et al. 2013; Eaton and Ree 2013; Cui et al. 2013) but excluding alternative

505   explanations based on incomplete lineage sorting or post-speciation selection is difficult

506   (reviewed in Sousa and Hey 2013; Cruickshank and Hahn 2014).

507   As in our study, a greater number of chromosome rearrangements have been found

508   in sympatric compared to allopatric taxa in *Drosophila* (Noor et al. 2001), *Helianthus*

509   sunflowers (Rieseberg 2001), *Anopheles* mosquitoes (Ayala and Coluzzi 2005), *Agrodiaetus*

510   butterflies (Kandul et al. 2007), and rodents in the families *Cricetidae* and *Muridae*

511   (Castiglia 2014). One hypothesis, as stated above, is that sympatry reflects historical gene

512   flow between partially reproductively isolated populations that promoted the fixation of

513   these inversions. An alternative is that the greater number of inversion differences

514   observed between sympatric, relative to allopatric, sister taxa occurs because incipient

515   species are more likely to fuse upon secondary contact when they share karyotype

516   structure (Noor et al. 2001; Rieseberg et al. 2001). Previous studies of the phylogeographic

517   context of rearrangement evolution have not evaluated these alternatives (Ayala and

518   Coluzzi 2005; Kandul et al. 2007; Castiglia 2014). If rearrangements accumulate at a

519   constant rate, and incipient species that differ by chromosome rearrangements are far

520   more likely to persist upon secondary contact than those that differ not, then the extent of

521   karyotypic differentiation between sympatric taxa should comprise the upper end of the

522   distribution of karyotypic differentiation between allopatric ones. We do not observe this

523   pattern in the Estrildid finches. Species pairs from island clades, which contain

524   predominately allopatric taxa, have significantly fewer inversions than species pairs from

525   continental clades, which consist of predominately sympatric taxa, regardless of the age of

526   species compared. However, this is confounded with population size. Future examination

527   into the extent of karyotype differentiation between allopatric sister taxa on continents

528   (e.g. the firetails of southwestern and southeastern Australia, gen. *Stagonopleura*),

529   sympatric sister taxa on islands (e.g. the munia species complex of Papua New Guinea, gen.

530   *Lonchura*), and active speciation in hybrid zones should elucidate the extent to which gene

531   flow facilitates inversion fixation more explicitly.

532   We conclude that karyotype structure across the Estrildid finches is highly variable

533   and appears to evolve rapidly under certain demographic and geographic conditions.

534   Considering that karyotype divergence has been considered unimportant to avian

535   diversification (Ellegren 2010), one question is whether the Estrildid finches are

536   representative of or an exception to the avian rule. In Table 3 we summarized the past 40+

537   years of cytological research in songbirds, the avian order comprising ~45% of all bird

538   species, which suggests that chromosome inversions, while variable in number between

539   families, are a pervasive feature of karyotype evolution in songbirds and often involve the

540   sex chromosomes (considering data from 351 taxa across 56 families, see Tables 3 and S6).

541   The upshot appears to be that the *Estrildidae* are not an exception so much in terms of the

542   raw variation in genomic structure between species as they are an exception with respect

543   to the breadth and intensity of taxa so far examined, likely because many Estrildid species

544   are available from aviculture stocks. Our results suggest that alterations of genomic

545   structure may be as important to bird speciation as has been proposed for other better-

546   studied taxa (reviewed in Hoffmann and Rieseberg 2008; Faria and Navarro 2010). That

547   karyotype evolution by chromosome inversion might be a common feature of avian

548   diversification is an exciting prospect for speciation studies and one that should serve to

549   stimulate future research into the extent of genomic rearrangement between species and

550   the evolutionary context in which these changes occur.

551

LITERATURE CITED:

Anderson, A. R., Hoffmann, A. A., Mckechnie, S. W., Umina, P. A., & Weeks, A. R. 2005. The latitudinal cline in the In (3R) Payne inversion polymorphism has shifted in the last 20 years in Australian *Drosophila melanogaster* populations. Molecular Ecology, 14: 851-858.

Arnaiz-Villena, A., Ruiz-del-Valle, V., Gomez-Prieto, P., Reguera, R., Parga-Lozano, C., & Serrano-Vela, I. 2009. Estrildinae Finches (Aves, Passeriformes) from Africa, South Asia and Australia: a Molecular Phylogeographic Study. The Open Ornithology Journal, 2, 29-36.

Aslam, M. L., Bastiaansen, J. W., Crooijmans, R. P., Vereijken, A., Megens, H. J., & Groenen, M. A. 2010. A SNP based linkage map of the turkey genome reveals multiple intrachromosomal rearrangements between the turkey and chicken genomes. Bmc Genomics, 11: 647.

Ayala FJ, & Coluzzi M. 2005. Chromosome speciation: Humans, Drosophila, and mosquitoes. Proc. Natl. Acad. Sci. USA 102:6535–42

Ayala, D., Guerrero, R. F., & Kirkpatrick, M. 2013. Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito. Evolution, 67: 946-958.

Backström, N., Forstmeier, W., Schielzeth, H., Mellenius, H., Nam, K., Bolund, E., ... & Ellegren, H. 2010. The recombination landscape of the zebra finch *Taeniopygia guttata* genome. Genome research, 20: 485-495.

Balakrishnan, C. N., & Edwards, S. V. 2009. Nucleotide variation, linkage disequilibrium and

578   founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia*

579   *guttata*). Genetics, 181: 645-660.

580   Baudat, F., & de Massy, B. 2007. Regulating double-stranded DNA break repair towards

581   crossover or non-crossover during mammalian meiosis. Chromosome research, 15:

582   565-577.

583   Bed'Hom, B., Coullin, P., Guillier-Gencik, Z., Moulin, S., Bernheim, A., & Volobouev, V. 2003.

584   Characterization of the atypical karyotype of the black-winged kite *Elanus caeruleus*

585   (Falconiformes: *Accipitridae*) by means of classical and molecular cytogenetic

586   techniques. Chromosome Research, 11: 335-343.

587   Bollback, J. P. 2006. SIMMAP: stochastic character mapping of discrete traits on

588   phylogenies. BMC bioinformatics, 71, 88.

589   Brown K.M., Burk L.M., Henagan L.M., Noor M.A.F. 2004. A test of the chromosomal

590   rearrangement model of speciation in *Drosophila pseudoobscura*. Evolution 58:

591   1856–60.

592   Castiglia, R. 2014. Sympatric sister species in rodents are more chromosomally

593   differentiated than allopatric ones: implications for the role of chromosomal

594   rearrangements in speciation. Mammal Review, 44: 1-4.

595   Charlesworth, D., & Charlesworth, B. 1979. Selection on recombination in clines. Genetics,

596   91: 581-589.

597   Charlesworth, B., Coyne, J. A., & Barton, N. H. 1987. The relative rates of evolution of sex

598   chromosomes and autosomes. American Naturalist, 113-146.

599   Christidis, L. 1983. Extensive chromosomal repatterning in two congeneric species: *Pytilia*

600   *melba*, and *Pytilia phoenicoptera* (Swainson *Estrildidae*); Aves. Cytogenetic and

601   Genome Research, 364: 641-648.

602   —. 1986a. Chromosomal evolution within the family *Estrildidae* Aves I. The Poephilae.

603   Genetica, 712: 81-97.

604   —. 1986b. Chromosomal evolution within the family *Estrildidae* Aves II. The Lonchurae.

605   Genetica, 712: 99-113.

606   —. 1987. Chromosomal evolution within the family *Estrildidae* Aves III. The Estrildae

607   waxbill finches. Genetica, 722: 93-100.

608   —.1990. Aves, Pages 116 in B. John, ed., Animal Cytogenetics. Berlin and Stuttgart,

609   Gebrüder Bornträger.

610   Cole, C. T. 2003. Genetic variation in rare and common plants. Annual Review of Ecology,

611   Evolution, and Systematics, 213-237.

612   Coyne, J. A., & Orr, H. A. 2004. Speciation (Vol. 37). Sunderland, MA: Sinauer Associates.

613   Cruickshank, T. E., & Hahn, M. W. 2014. Reanalysis suggests that genomic islands of

614   speciation are due to reduced diversity, not reduced gene flow. Molecular ecology.

615   Cui, R., Schumer, M., Kruesi, K., Walter, R., Andolfatto, P., & Rosenthal, G. G. 2013.

616   Phylogenomics reveals extensive reticulate evolution in Xiphophorus fishes.

617   Evolution, 67: 2166-2179.

618   de Massy, B. 2013. Initiation of meiotic recombination: how and where? Annu. Rev. Genet.

619   110711-155423

620   de Oliveira EHC, Habermann FA, Lacerda O, Sbalqueiro IJ, Wienberg J, Muller S. 2005.

621   Chromosome reshuffling in birds of prey: the karyotype of the worlds largest eagle

622   (Harpy eagle, *Harpia harpyja*) compared to that of the chicken (*Gallus gallus*).

623   Chromosoma 114: 338Y343.

624   Dobzhansky, T. 1951 Genetics and the origin of species (eds N. Eldredge & S. J. Gould), 3rd

625   edn. New York, NY: Columbia University Press.

626   Drummond, A. J., & Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling

627   trees. BMC evolutionary biology, 71: 214.

628   Dumont, B. L., & Payseur, B. A. 2008. Evolution of the genomic rate of recombination in

629   mammals. Evolution, 62: 276-294.

630   Dumont, B. L., & Payseur, B. A. 2011. Evolution of the genomic recombination rate in murid

631   rodents. Genetics, 187: 643-657.

632   Eaton, D. A., & Ree, R. H. 2013. Inferring phylogeny and introgression using RADseq data:

633   an example from flowering plants (*Pedicularis*: *Orobanchaceae*). Systematic biology,

634   62: 689-706.

635   Ellegren, H. 2009. Genomic evidence for a large-Z effect. Proceedings of the Royal Society B:

636   Biological Sciences, 276: 361-366.

637   —. 2010. Evolutionary stasis: the stable chromosomes of birds. Trends in ecology &

638   evolution, 255: 283-291.

639   —. 2013. The evolutionary genomics of birds. Annual Review of Ecology, Evolution, and

640   Systematics, 44: 239-259.

641   Faria, R., and A. Navarro. 2010. Chromosomal speciation revisited: rearranging theory with

642   pieces of evidence. Trends in Ecology & Evolution 25: 660-669.

643   Feder, J. L., Gejji, R., Powell, T. H., & Nosil, P. 2011. Adaptive chromosomal divergence

644   driven by mixed geographic mode of evolution. Evolution, 658: 2157-2170.

645   Fishman, L., Stathos, A., Beardsley, P. M., Williams, C. F., & Hill, J. P. 2013. Chromosomal

646   rearrangements and the genetics of reproductive barriers in *Mimulus* (monkey

647   flowers). Evolution, 67: 2547-2560.

648 Goodwin, D., & Woodcock, M. 1982. Estrildid finches of the world. London: British Museum

649   Natural History.

650 Grant, P. R., & Grant, B. R. 2011. How and why species multiply: the radiation of Darwin's

651   finches. Princeton University Press.

652 Graves, J. A. M. 2014. Avian sex, sex chromosomes, and dosage compensation in the age of

653   genomics. Chromosome Research, 22: 45-57.

654 Haldane, J., 1927 A mathematical theory of natural and artificial selection. Part V: selection

655   and mutation. Math. Proc. Camb. Philos. Soc. 3: 607–615.

656 Hansmann, T., Nanda, I., Volobouev, V., Yang, F., Schartl, M., Haaf, T., & Schmid, M. 2009.

657   Cross-species chromosome painting corroborates microchromosome fusion during

658   karyotype evolution of birds. Cytogenetic and genome research, 126: 281-304.

659 Hansson, B., Ljungqvist, M., Dawson, D. A., Mueller, J. C., Olano-Marin, J., Ellegren, H., &

660   Nilsson, J. Å. 2009. Avian genome evolution: insights from a linkage map of the blue

661   tit (*Cyanistes caeruleus*). Heredity, 104: 67-78.

662 Hedrick, P. W. 1981. The establishment of chromosomal variants. Evolution: 322-332.

663 Hirschi, M., Hauschteck-Jungen, E. & Zigwild, V., 1972. Karyotypen von zwei Vogelarten,

664   *Euodice cantans* (*Estrildidae*) und *Emberiza flaviventris* (*Pyrrhuloxiidae*,

665   Emberizinae). Cytologia 37: 525-529.

666 Hoffmann, A. A., & Rieseberg, L. H. 2008. Revisiting the impact of inversions in evolution:

667   from population genetic markers to drivers of adaptive shifts and speciation?.

668   Annual review of ecology, evolution, and systematics, 39: 21-42.

669   Itoh, Y., & Arnold, A. P. 2005. Chromosomal polymorphism and comparative painting

670   analysis in the zebra finch. Chromosome Research, 131: 47-56.

671   Jaenike, J. 2001. Sex chromosome meiotic drive. Annual Review of Ecology and Systematics,

672   25-49.

673   Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. 2012. The global diversity of

674   birds in space and time. Nature, 491: 444-448.

675   Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M.

676   Pirun, M. C. Zody, S. White, et al. 2012. The genomic basis of adaptive evolution in

677   threespine sticklebacks. Nature 484: 55–61.

678   Joron, M., L. Frezal, R. T. Jones, N. L. Chamberlain, S. F. Lee, C. R. Haag, A. Whibley, M.

679   Becuwe, S. W. Baxter, L. Ferguson, et al.. 2011.Chromosomal rearrangements

680   maintain a polymorphic supergene controlling butterfly mimicry. Nature 477: 203–

681   U102.

682   Kandul, N. P., Lukhtanov, V. A., & Pierce, N. E. 2007. Karyotypic diversity and speciation in

683   *Agrodiaetus* butterflies. Evolution, 613: 546-559.

684   Kawakami, T., Smeds, L., Backström, N., Husby, A., Qvarnström, A., Mugal, C. F., ... & Ellegren,

685   H. 2014. A high-density linkage map enables a second-generation collared flycatcher

686   genome assembly and reveals the patterns of avian recombination rate variation

687   and chromosomal evolution. Molecular ecology.

688   Kimura, M. 1962. On the probability of fixation of mutant genes in a population. Genetics,

689   47: 713.

690   King, M. 1987. Chromosomal rearrangements, speciation and the theoretical approach.

691   Heredity 59: 1–6.

692   Kirkpatrick, M., and N. H. Barton. 2006. Chromosome inversions, local adaptation and

693   speciation. Genetics 173: 419-434.

694   Kirkpatrick, M. 2010. How and why chromosome inversions evolve. PLoS biology, 8(9),

695   e1000501.

696   Krasikova, A., Daks, A., Zlotina, A., & Gaginskaya, E. 2009. Polymorphic heterochromatic

697   segments in Japanese quail microchromosomes. Cytogenetic and genome research,

698   126: 148-155.

699   Kunte, K., Zhang, W., Tenger-Trolander, A., Palmer, D. H., Martin, A., Reed, R. D., ... &

700   Kronforst, M. R. 2014. doublesex is a mimicry supergene. Nature, 507: 229-232.

701   Lande, R. 1979. Effective deme sizes during long-term evolution estimated from rates of

702   chromosomal rearrangement. Evolution, 33: 234–251.

703   —. 1985 The fixation of chromosomal rearrangements in a subdivided population with

704   local extinction and recolonisation. Heredity, 54: 323–332.

705   Leffler, E. M., Bullaughey, K., Matute, D. R., Meyer, W. K., Segurel, L., Venkat, A., ... &

706   Przeworski, M. 2012. Revisiting an old riddle: what determines genetic diversity

707   levels within species?. PLoS biology, 10, e1001388.

708   Levan, A., Fredga, K., & Sandberg, A. A. 1964. Nomenclature for centromeric position on

709   chromosomes. Hereditas, 52: 201-220.

710   Lithgow, P. E., O'Connor, R., Smith, D., Fonseka, G., Al Mutery, A., Rathje, C., ... & Griffin, D. K.

711   2014. Novel tools for characterising inter and intra chromosomal rearrangements in

712   avian microchromosomes. Chromosome Research, 22: 85-97.

713   Lowery, R. K., Uribe, G., Jimenez, E. B., Weiss, M. A., Herrera, K. J., Regueiro, M., & Herrera, R.

714   J. 2013. Neanderthal and Denisova genetic affinities with contemporary humans:

715   Introgression versus common ancestral polymorphisms. Gene, 530: 83-94.

716 Lowry, D. B., and J. H. Willis. 2010. A widespread chromosomal inversion polymorphism

717   contributes to a major life-history transition, local adaptation, and reproductive

718   isolation. PLoS Biol. 8:1–14.

719 Maddison, W. P., & Maddison, D. 2001. Mesquite: a modular system for evolutionary

720   analysis.

721 Mank, J. E., & Ellegren, H. 2007. Parallel divergence and degradation of the avian W sex

722   chromosome. Trends in ecology & evolution, 22: 389-391.

723 Mank, J. E., Nam, K., & Ellegren, H. 2010a. Faster-Z evolution is predominantly due to

724   genetic drift. Molecular biology and evolution, 27: 661-670.

725 Mank, J. E., Vicoso, B., Berlin, S., & Charlesworth, B. 2010b. Effective population size and the

726   Faster-X effect: empirical results and their interpretation. Evolution, 64: 663-674.

727 Marshall, O. J., Chueh, A. C., Wong, L. H., & Choo, K. H. 2008. Neocentromeres: new insights

728   into centromere structure, disease development, and karyotype evolution. The

729   American Journal of Human Genetics, 82: 261-282.

730 Masabanda, J. S., Burt, D. W., O'Brien, P. C., Vignal, A., Fillon, V., Walsh, P. S., ... & Griffin, D. K.

731   2004. Molecular cytogenetic definition of the chicken genome: the first complete

732   avian karyotype. Genetics, 166: 1367-1373.

733 Nanda I, Karl E, Volobouev V, Griffin, D.K., Schartl M., Schmid M. 2006. Extensive gross

734   genomic rearrangements between chicken and Old World vultures (Falconiformes:

735   *Accipitridae*). Cytogenet. Genome Res. 112: 286Y295.

736 Nanda, I., Karl, E., Griffin, D. K., Schartl, M., & Schmid, M. 2007. Chromosome repatterning in

737   three representative parrots (Psittaciformes) inferred from comparative

738   chromosome painting. Cytogenetic and genome research, 117: 43-53.

739   Navarro, A., and N. H. Barton. 2003. Accumulating postzygotic isolation genes in parapatry:

740   a new twist on chromosomal speciation. Evolution 57: 447-459.

741   Nevo, E., Beiles, A., & Ben-Shlomo, R. 1984. The evolutionary significance of genetic

742   diversity: ecological, demographic and life history correlates (pp. 13-213). Springer

743   Berlin Heidelberg.

744   Nie, W., O'Brien, P. C., Ng, B. L., Fu, B., Volobouev, V., Carter, N. P., ... & Yang, F. 2009. Avian

745   comparative genomics: reciprocal chromosome painting between domestic chicken

746   (*Gallus gallus*) and the stone curlew (*Burhinus oedicnemus*, Charadriiformes)—An

747   atypical species with low diploid number. Chromosome research, 17: 99-113.

748   Nishida, C., Ishijima, J., Kosaka, A., Tanabe, H., Habermann, F. A., Griffin, D. K., & Matsuda, Y.

749   2008. Characterization of chromosome structures of Falconinae (*Falconidae*,

750   Falconiformes, Aves) by chromosome painting and delineation of chromosome

751   rearrangements during their differentiation. Chromosome research, 16: 171-181.

752   Nishida, C., Ishishita, S., Yamada, K., Griffin, D. K., & Matsuda, Y. 2014. Dynamic chromosome

753   reorganization in the osprey (Pandion haliaetus, Pandionidae, Falconiformes):

754   relationship between chromosome size and the chromosomal distribution of

755   centromeric repetitive DNA sequences. Cytogenetic and genome research, 142: 179-

756   189.

757   Noor, M. A. F., L. A. Grams, L. A. Bertucci, Y. Almendarez, J. Reiland, and K. R. Smith. 2001.

758   The genetics of reproductive isolation and the potential for gene exchange between

759   *Drosophila pseudoobscura* and *D. persimilis* via backcross hybrid males. Evolution

760   55: 512-521.

761   Ohta, T. 1972. Population size and rate of evolution. Journal of molecular evolution, 1: 305-

762   314.

763   Ortiz-Barrientos D, Reiland J, Hey J, Noor MAF. 2002. Recombination and the divergence of

764   hybridizing species. Genetica 116: 167–78.

765   Peischl, S., & Kirkpatrick, M. 2012. Establishment of new mutations in changing

766   environments. Genetics, 191: 895-906.

767   Poelstra, J. W., Vijay, N., Bossu, C. M., Lantz, H., Ryll, B., Mueller, I., ... & Wolf, J. B. 2014. The

768   genomic landscape underlying phenotypic integrity in the face of gene flow in

769   crows. Science, 344: 1410-1414.

770   Prasad, R., & Patnaik, S. C. 1977. Karyotypes of five passerine birds belonging to family

771   *Ploceidae*. Caryologia, 30.

772   Price, T. 2008. Speciation in birds. Roberts. Boulder CO.

773   Price, T. D., Hooper, D. M., Buchanan, C. D., Johansson, U. S., Tietze, D. T., Alström, P., ... &

774   Mohan, D. 2014. Niche filling slows the diversification of Himalayan songbirds.

775   Nature, 509: 222-225.

776   R Core Team 2013. R: A language and environment for statistical computing. R

777   Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0,

778   URLhttp://www.R-project.org/.

779   Ray-Chaudhuri, R. 1976. Cytotaxonomy and chromosome evolution in Passeriformes

780   (Ayes). Journal of Zoological Systematics and Evolutionary Research, 14: 299-320.

781   Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. Trends in Ecology &

782   Evolution 16: 351-358.

783   Shields, G. F. 1982. Comparative avian cytogenetics: a review. Condor, 45-58.

784   Skinner, B. M., & Griffin, D. K. 2011. Intrachromosomal rearrangements in avian genome

785   evolution: evidence for regions prone to breakpoints. Heredity, 1081: 37-41.

786   Smit A., Hubley R., & Green, P. 1996–2010. RepeatMasker Open-3.0. www.repeatmasker.

787   org.

788   Smukowski, C. S., & Noor, M. A. F. 2011. Recombination rate variation in closely related

789   species. Heredity, 107: 496-508.

790   Sorenson, M. D., & Payne, R. B. 2001. A single ancient origin of brood parasitism in African

791   finches: Implications for host-parasite coevolution. Evolution, 55: 2550-2567.

792   Sorenson, M. D., Balakrishnan, C. N., & Payne, R. B. 2004. Clade-limited colonization in

793   brood parasitic finches (*Vidua* spp.). Systematic Biology, 53: 140-153.

794   Sousa, V., & Hey, J. 2013. Understanding the origin of species with genome-scale data:

795   modelling gene flow. Nature Reviews Genetics, 14: 404-414.

796   Stapley, J., Birkhead, T. R., Burke, T., & Slate, J. 2008. A linkage map of the zebra finch

797   *Taeniopygia guttata* provides new insights into avian genome evolution. Genetics,

798   179: 651-667.

799   Stebbins, G. L. 1958. The inviability, weakness and sterility of interspecific hybrids. Adv.

800   Gen. 9: 147–215.

801   Takagi, N., & Sasaki, M. 1974. A phylogenetic study of bird karyotypes. Chromosoma, 46:

802   91-120.

803   Vicoso, B., & Charlesworth, B. 2009. Effective population size and the faster-X effect: an

804   extended model. Evolution, 63: 2413-2426.

805   Völker, M., Backström, N., Skinner, B. M., Langley, E. J., Bunzey, S. K., Ellegren, H., & Griffin, D.

806   K. 2010. Copy number variation, chromosome rearrangement, and their association

807   with recombination during avian evolution. Genome research, 20: 503-511.

808   Walsh, J. B. 1982. Rate of accumulation of reproductive isolation by chromosome

809   rearrangements. Amer. Nat. 120: 510–532.

810   Warren, W. C., Clayton, D. F., Ellegren, H., Arnold, A. P., Hillier, L. W., Künstner, A., ... &

811   Lancet, D. 2010. The genome of a songbird. Nature, 464: 757-762.

812   Wesley, C. S., & Eanes, W. F. 1994. Isolation and analysis of the breakpoint sequences of

813   chromosome inversion In (3L) Payne in Drosophila melanogaster. Proceedings of

814   the National Academy of Sciences, 91: 3132-3136.

815   White, M. J. D. 1969. Chromosomal rearrangements and speciation in animals. Annu. Rev.

816   Genet. 3: 75–98.

817   —. 1978 Modes of speciation. San Francisco, CA: W.H. Freeman.

818   Whitlock, M. C., Bürger, R., & Dieckmann, U. 2004. Fixation of new mutations in small

819   populations. Evolutionary conservation biology, 155-170.

820   Yan, L., Liqing, Z., Dexiang, Z., Xiquan, Z., & Xuemei, L. 2010. Faster evolution of Z-linked

821   duplicate genes in chicken. Journal of Genetics and Genomics, 37: 695-702.

822   Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., ... & Ganapathy, G. 2014. Comparative

823   genomics reveals insights into avian genome evolution and adaptation. Science, 346:

824   1311-1320.

825   Zlotina, A., Galkina, S., Krasikova, A., Crooijmans, R. P., Groenen, M. A., Gaginskaya, E., &

826   Deryusheva, S. 2012. Centromere positions in chicken and Japanese quail

827   chromosomes: de novo centromere formation versus pericentric inversions.

828   Chromosome Research, 20: 1017-1032.

829

Table 1 – Phylogenetic distribution of chromosome inversions in the *Estrildidae*

Clade:	Inversion fixation rate (Inv/My)	Number of inversions	Combined branch length (My)	Species karyotyped (Total species)	Crown age (Ma)	Proportion continental species	Avg. Range Size (Range) (10^6 km²)	Proportion sympatric*
A) Grassfinches	0.502	31	61.72	11 (14)	7.5	0.93	1.5 (5×10^{-2} - 6.3)	0.57
B) Munias	0.176	4	22.66	6 (26)	5.6	0.23	0.99 ($10x^{-3}$ - 7.8)	0.18
C) Parrotfinches	0.137	2	14.59	2 (12)	7.3	0.09	0.26 (3×10^{-3} - 1.6)	0.11
D) Pytilias, Cordon-bleus, & Firefinches	0.526	16	30.4	4 (31)	8.5	1.0	2.3 (10^{-2} - 10.3)	0.49
E) Adavats, Amadinas, & Quailfinches	0.545	14	25.68	4 (8)	9.4	1.0	2.7 (0.5 - 4.0)	0.39

830   *Proportion sympatric is the proportion of all pairs of species within each clade that overlapped in range more than 15%.

831

Table 2 – Genomic distribution of pericentric chromosome inversions

Chromosome ID	Size (Mb)	Map Length (cM)*		GC Content (%)	Repeat Content (%)	Invers
TGU2	156.4	34.7	(76)	39.0	0.38	3
TGU1	118.6	63.7	(70)	39.2	0.23	8
TGU3	112.6	22	(69)	39.4	0.38	8
TGU1A	73.7	63.3	(91)	39.7	0.14	9
TGU4	69.8	31.9	(18)	39.2	0.28	10
TGU5	62.4	89.8	(64)	40.8	0.12	11
TGU7	39.8	27.8	(41)	41.1	0.14	5
TGU6	36.3	44.6	(60)	41.6	0.19	4
TGU8	28.0	44.9	(47)	41.3	0.05	4
TGU9	27.2	60.5	(52)	43.1	0.15	4
TGU12	21.6	47.5	(34)	43.7	0.07	1
Z	74.6	32.8	(29)	39.2	1.5	13
W	27.0	-		-	-	11

832 *Map length values are from the framework map for zebra finch given in Table 2 in St

833 et al. (2008) and those shown in parentheses from the framework map zebra finch m

834 Backström et al. (2010)

835

836 **Table 3** – Summary of karyotypic variation in *Passeriformes* from cytological analyse

837 Songbird families, based on the taxonomy of Jetz (2012), are listed from the top, roug

838 from basal to most derived. Cell entries refer to the number of autosomal chromoson

839 carrying at least one inversion; A refers to an absence of inversions detected. For the

840 chromosomes: P indicates inversions present and A indicates their absence. The *Estr*

841 are in bold. Only families with karyotype descriptions for two or more taxa are show

842 (see Table S6 for a complete dataset listing taxa examined and references).

Family	Number of taxa karyotyped	Autosomal chromosomes	Chr. Z	Chr. W
Suborder Tyranni:				
Tyrannidae	13	8	P	P
Furnariidae	2	2	A	A
Thamnopphilidae	4	5	P	A
Suborder Passeri:				
Parvorder Corvida:				
Corvidae	16	4	P	P
Laniidae	5	4	P	P
Dicruridae	2	A	A	A
Campephagidae	2	2	P	P
Oreolidae	3	1	P	A
Vireonidae	4	3	A	A
Parvorder Passerida:				
Sittidae	2	2	P	A
Sturnidae	6	4	P	P
Mimidae	3	2	P	A
Muscicapidae	29	7	P	P
Turdidae	16	6	P	P
Bombycillidae	2	1	A	A
Chloropsidae	2	A	P	P
Ploceidae	3	3	A	A
Estrildidae	**32**	**11**	**P**	**P**
Motacillidae	5	4	P	P
Passeridae	6	4	P	P
Emberizidae	16	8	P	P
Icteridae	3	3	P	P
Parulidae	6	2	A	A
Passerelidae	17	8	P	P
Cardinalidae	7	2	P	P
Thraupidae	42	6	P	P
Fringillidae	21	11	P	P
Paridae	5	A	A	A
Hirundinidae	3	1	A	A
Pycnonotidae	8	6	P	P
Locustellidae	3	3	P	A
Cisticolidae	2	2	A	A
Acrocephalidae	3	5	P	A
Aegithalidae	3	2	P	P
Phylloscopidae	9	6	P	P
Sylviidae	5	5	P	A
Zosteropidae	4	1	A	A
Timaliidae	3	1	P	A
Pellorneidae	2	2	P	P
Leiothrichidae	15	6	A	P

843

844   **Figure Legends**

845

846   **Figure 1** – Phylogenetic distribution of chromosome inversions

847   Phylogeny of 32 Estrildid finch species with karyotype described and time-dated using the

848   split between families *Estrildidae* and *Ploceidae*. Continental Estrildid clades are boxed in

849   red and island chain clades in blue. The number of inversions inferred to have occurred,

850   based on ancestral karyotype reconstruction, are shown on each branch. The total number

851   of inversions inferred in each of the five clades and a species representative are shown to

852   the right of each clade. All 5 clades have 100% posterior probability support. From top to

853   bottom: (A) grassfinches – *Stitzoptera bichenovii*, (B) munias – *Lonchura malacca*, (C)

854   parrotfinches – *Cholebia gouldiae*, (D) pytilias, cordon-bleus, and firefinches – *Lagonosticta*

855   *senegala*; (E) adavats, amadina, and quailfinches – *Amadina erythrocephala*.

856

857   **Figure 2** – Rate of inversion fixation (number of inversions / total clade branch length in

858   millions of years) against the average range size of all taxa within each clade. One-tailed

859   Pearson's correlation: r = 0.91, p = 0.02. Continental Estrildid clades are colored in red and

860   island chain clades in blue. Clades are arranged from left to right (following Table 1): (C)

861   parrotfinches, (B) munia, (A) grassfinches, (D) pytilias, cordon-bleus, and firefinches, (E)

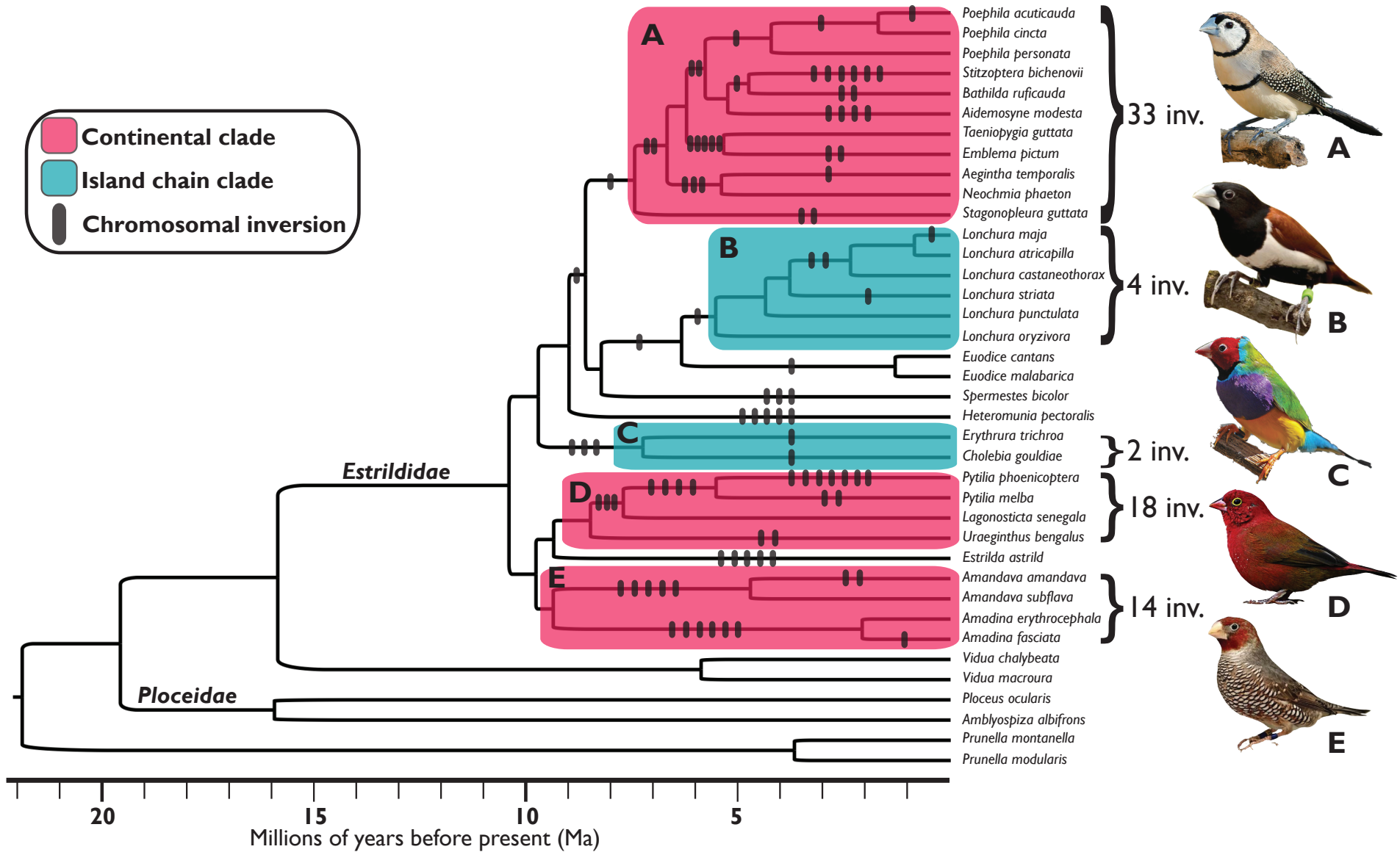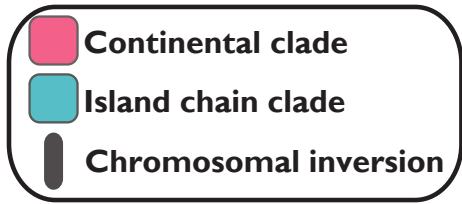862   adavats, amadina, and quailfinches.

Figure 1

Figure 2