Title

The gene expression profile of the song control nucleus HVC shows sex specificity, hormone responsiveness, and species specificity among songbirds

Author affiliation

Meng-Ching Ko[*] (ORCID 0000-0002-2234-9380)

Carolina Frankl-Vilches[*]

Antje Bakker[*]

Manfred Gahr[*]

[*]Department of Behavioural Neurobiology, Max Planck Institute for Ornithology, Eberhard-Gwinner Straße 6a, 82319 Seewiesen, Germany

Corresponding author

Meng-Ching Ko

E-mail: mcko@orn.mpg.de

Keywords

Sex differences; Songbirds; Gene expression; Brain; HVC; Testosterone

Abstract

Singing occurs in songbirds of both sexes, but some species show typical degrees of sex-specific performance. We studied the transcriptional sex differences in the HVC, a brain nucleus critical for song pattern generation, of the forest weaver (*Ploceus bicolor*), the blue-capped cordon-bleu (*Uraeginthus cyanocephalus*), and the canary (*Serinus canaria*), which are species that show low, medium, and high levels of sex-specific singing, respectively. We observed persistent sex differences in gene expression levels regardless of the species-specific sexual singing phenotypes. We further studied the HVC transcriptomes of defined phenotypes of canary, known for its testosterone-sensitive seasonal singing. By studying both sexes of canaries during both breeding and nonbreeding seasons, nonbreeding canaries treated with testosterone, and spontaneously singing females, we found that the circulating androgen levels and sex were the predominant variables associated with the variations in the HVC transcriptomes. The comparison of natural singing with testosterone-induced singing in canaries of the same sex revealed considerable differences in the HVC transcriptomes. Strong transcriptional changes in the HVC were detected during the transition from nonsinging to singing in canaries of both sexes. Although the sex-specific genes of singing females shared little resemblance with those of males, our analysis showed potential functional convergences. Thus, male and female songbirds achieve comparable singing behaviours with sex-specific transcriptomes.

Introduction

Most of the genomes of male and female individuals of the same species are the same but stark sex differences in physiological, phenotypical, or behavioural traits between the sexes are common and widespread in the animal kingdom. The songbird clade (a suborder of the perching birds) consists of more than 4000 extant avian species, and these exhibit a great diversity of sex differences in singing behaviour. Among domesticated canaries (*Serinus canaria*), males are known for their singing behaviour and have been selected for their sophisticated songs for centuries, whereas female canaries seldom sing (Hartley et al., 1997; Herrick and Harris, 1957; Ko et al., 2020; Pesch and Güttinger, 1985; Shoemaker, 1939; Vallet et al., 1996). Such substantial sex differences in singing behaviour are commonly found in the majority of northern temperate songbird species, even though females of many tropical and southern temperate species sing regularly, and their songs play an important role in inter-sexual communication (Hall et al., 2015; Price, 2019; Price et al., 2009; Slater and Mann, 2004). For example, female blue-capped cordon-bleus (*Uraeginthus cyanocephalus*) use their song as advertising signals and address their songs to their mates (Immelmann, 1968; Ota et al., 2018), although the female songs appear shorter and less complex than those of males (Geberzahn and Gahr, 2011). In another tropical songbird species, forest weavers (*Ploceus bicolor*), males and females develop their songs during pair binding and eventually learn to sing identical duets, which they use to defend their territories (Wickler and Seibt, 1980).

Although the extent of sex differences in singing behaviour varies greatly in the songbird clade (Ball, 2016), the song quality and occurrence in males and females of many songbird species are dependent on testosterone (Dittrich et al., 2014; Fusani et al., 2003; Ko et al., 2020; Leitner et al., 2001b; Nottebohm et al., 1987; Voigt and Leitner, 2008). In canaries that breed seasonally, the optimal breeding conditions are tightly associated with an increase in the day length, which initiates gonadal growth and testosterone production as well as various types of breeding activity, including singing (Leitner et al., 2001b; Nottebohm et al., 1987; Voigt and Leitner, 2008). The length and syllable repetition rate of songs during the breeding season are greater than those of songs in nonbreeding seasons (Leitner et al., 2001a; Leitner et al., 2001b; Voigt and Leitner, 2008). Castrated male canaries sing shorter songs than sham-operated males, and the subcutaneous implantation of testosterone results in the recovery of singing performance (Heid et al., 1985). The local administration of testosterone into the preoptic brain region increases the male canary singing rate by increasing motivation (Alward et al., 2013). Similarly, the systemic administration of testosterone reliably and repeatedly induces

singing behaviour in otherwise nonsinging female canaries (Fusani et al., 2003; Herrick and Harris, 1957; Leonard, 1939; Madison et al., 2015; Nottebohm, 1980; Shoemaker, 1939; Vellema et al., 2019b). Although female canaries rarely exhibit spontaneous singing (Hartley et al., 1997; Herrick and Harris, 1957; Ko et al., 2020; Pesch and Güttinger, 1985; Shoemaker, 1939; Vallet et al., 1996), its occurrence appears to be associated with the plasma androgen levels (Ko et al., 2020).

A set of interconnected neural circuits called the song control system controls the production and learning of singing behaviour (Nottebohm et al., 1976; Wild, 2004). The premotor nucleus HVC (used as the proper name), which is a sensorimotor integration centre in the song control system, is involved in the frequency and temporal modulation of songs in male and female songbirds (Hahnloser et al., 2002; Hoffmann et al., 2019) and in the sexual preferences regarding conspecific song displays (Brenowitz, 1991; Del Negro et al., 1998). The HVC is the only brain nucleus within the song control system that expresses receptors for both androgens and estrogens (Frankl-Vilches and Gahr, 2018; Gahr, 2001). Intriguingly, the anatomical properties of the HVC, such as volume (Figure 1), neuron number, and dendrite complexity, are male-biased (greater, higher and more complex, respectively, in males than in females) in all songbird species that have been examined, irrespective of the existence of sex differences in singing behaviour (Brenowitz et al., 1985; Gahr et al., 2008; Gahr et al., 1998; Gurney and Konishi, 1980; Hall et al., 2010; Lobato et al., 2015; MacDougall-Shackleton and Ball, 1999; Nixdorf et al., 1989; Nottebohm and Arnold, 1976; Schwabl et al., 2015). Testosterone treatment increases the delineable volume of the HVC in both male and female canaries (Fusani et al., 2003; Madison et al., 2015; Nottebohm, 1980) and many other species (Bernard and Ball, 1997; Dittrich et al., 2014; Gulledge and Deviche, 1999; Smith et al., 1997; Van Meir et al., 2004). Nevertheless, the HVC volume of female canaries implanted with testosterone remains markedly smaller than that of males (Nottebohm, 1980) (Figure 1). Testosterone clearly regulates singing behaviour and the anatomy of the HVC in both male and female songbirds. However, an intrinsic limit to the alterations induced by testosterone appears to exist, and this limit prevents female songbirds from reaching the same levels of "maleness" in terms of song characteristics and HVC anatomy.

How are such sex-specific differences in singing and HVC anatomy achieved in different species? We hypothesize the existence of a fundamental difference in gene regulation and expression between male and female conspecific songbirds. Based on this hypothesis, sex differences in the HVC transcriptomes of male and female conspecifics should always be observable, even though both sexes share an almost identical genome.

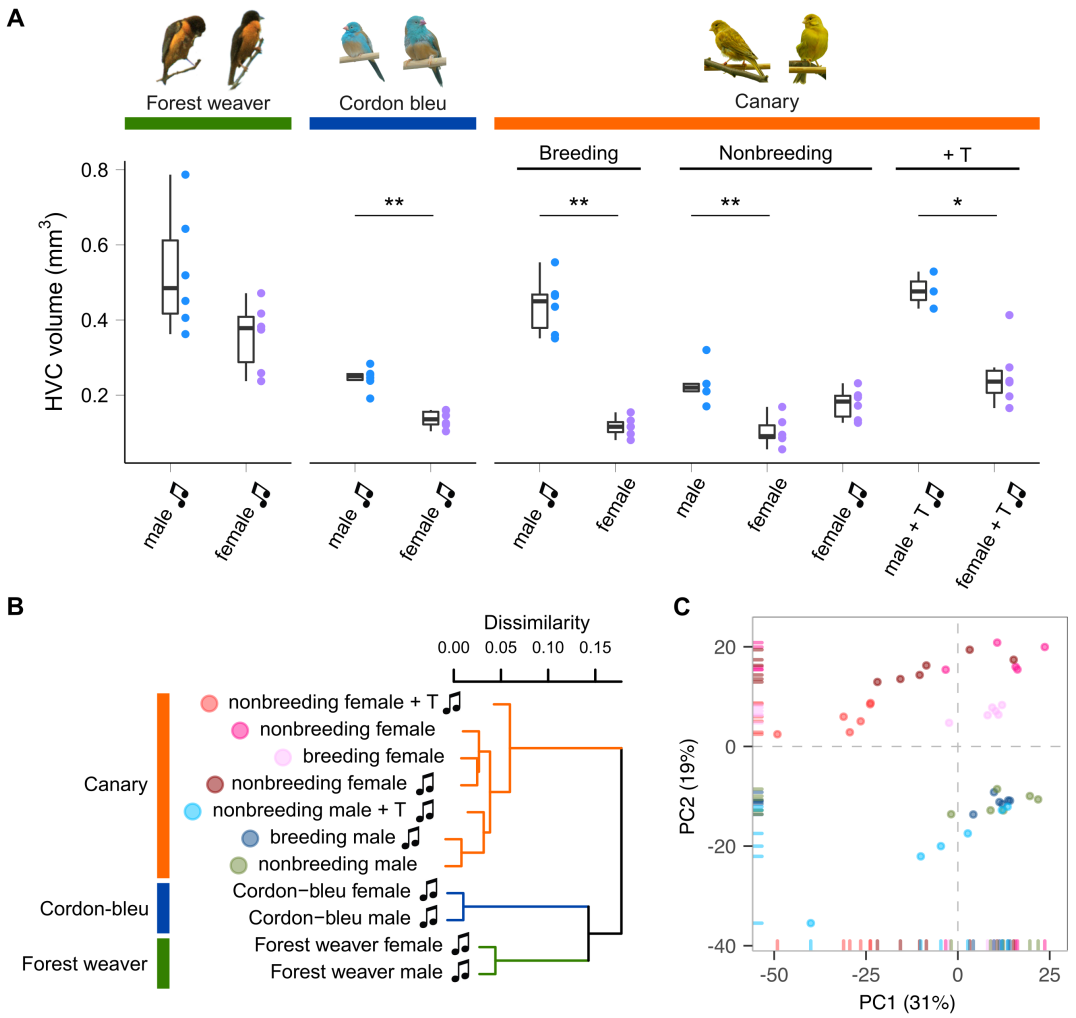
The female-specific W chromosome harbours 30-50 genes, whereas one and two copies of the Z chromosome (harbouring approximately 1,080 genes) are present in female and male birds, respectively (Frankl-Vilches et al., 2015; Sayers et al., 2020; Smeds et al., 2015). We further hypothesize that such intrinsic differences cannot be overwritten by manipulating the testosterone level in females. To test our hypotheses, we compared the transcriptomes of HVCs microdissected from three songbird species—the forest weaver, the blue-capped cordon-bleu, and the canary. These species represent three categories of sex differences in singing behaviour (low, medium, and high). Furthermore, we compared male and female canaries with and without testosterone implantation to evaluate the effects of testosterone (Supplementary Table 1). The results revealed sex differences in the HVC transcriptome between- and within-species comparisons. Our results indicate that although the extent of sex-biased gene expression is context-dependent, it is persistent regardless of hormonal manipulation, behavioural phenotypes, and distinct genetic backgrounds.

Results and discussion

The species identity distinguishes its HVC transcriptomes

We quantified the total mRNA from the HVC of male and female birds belonging to three songbird species (forest weavers, blue-capped cordon-bleus, and canaries) during the breeding season (Supplementary Table 1). For intraspecies sex comparisons, we included male and female canaries during the nonbreeding seasons (both nonsinging), and subgroups were treated with testosterone to induce singing behaviour in both sexes (Supplementary Figure 1). In addition, we obtained a rare group of nonbreeding female canaries that sang spontaneously without exogenous testosterone manipulation. However, their plasma androgen levels were intrinsically higher than those of nonbreeding and nonsinging females (Ko et al., 2020). To visualize the similarities in the gene expression patterns among and within the three species, we calculated Spearman's rank correlation coefficient σ for 12,360 gene expression levels and calculated the distance matrix based on the coefficient σ. The resulting cladogram indicated that the HVC transcriptomes of the eleven groups were clustered primarily by phylogenetic relatedness (Figure 1B and Supplementary Figure 2). We obtained a similar result with the transcriptomes of another tissue, the entopallium (Supplementary Figure 3), which is an area of the avian visual system that is functionally equivalent to the primary visual cortex of mammals. The within-species comparison of the canary HVC showed that the individual transcriptomes clustered well by sex, with the exception of the nonbreeding

141 canaries implanted with testosterone (Figure 1B and Supplementary Figure 2). The
142 subclusters of forest weaver and cordon-bleu HVC transcriptomes were less sexually
143 differential than those of canaries, i.e., male and female birds were intermittent within the
144 subcluster of the two species (Supplementary Figure 2 and Supplementary Figure 3).



145

146 **Figure 1. Species, sex, and plasma androgen level are the major determinants of the**
147 **HVC gene expression patterns in songbirds.**

A, HVC volume of birds used in this study. Forest weavers: males 0.528 mm^3 (mean), females 0.357 mm^3. Mann-Whitney Test, U = 7, P value = 0.0931. Cordon-bleus: males 0.245 mm^3, females 0.136 mm^3. Mann-Whitney Test, U = 0, P value = 0.00217. Breeding canaries: singing males 0.439 mm^3, nonsinging females 0.116 mm^3. Mann-Whitney Test, U = 0, P value = 0.00217. Nonbreeding canaries: nonsinging males 0.228 mm^3, nonsinging females 0.103 mm^3, singing females 0.176 mm^3. Mann-Whitney Test (nonsinging males vs. nonsinging females), U = 0, P value = 0.00492. Mann-Whitney Test (nonsinging males vs. singing females), U = 8, P value = 0.127. Testosterone-implanted canaries: males 0.478 mm^3, females 0.253 mm^3. Mann-Whitney Test, U = 0, P value = 0.0238. * P value < 0.05; ** P value < 0.01. Each colour-coded dot indicates the measurement from one bird. The boxes indicate the 25th/50th/75th percentiles (bottom/middle/top bar), and the extent of the whiskers indicates the most

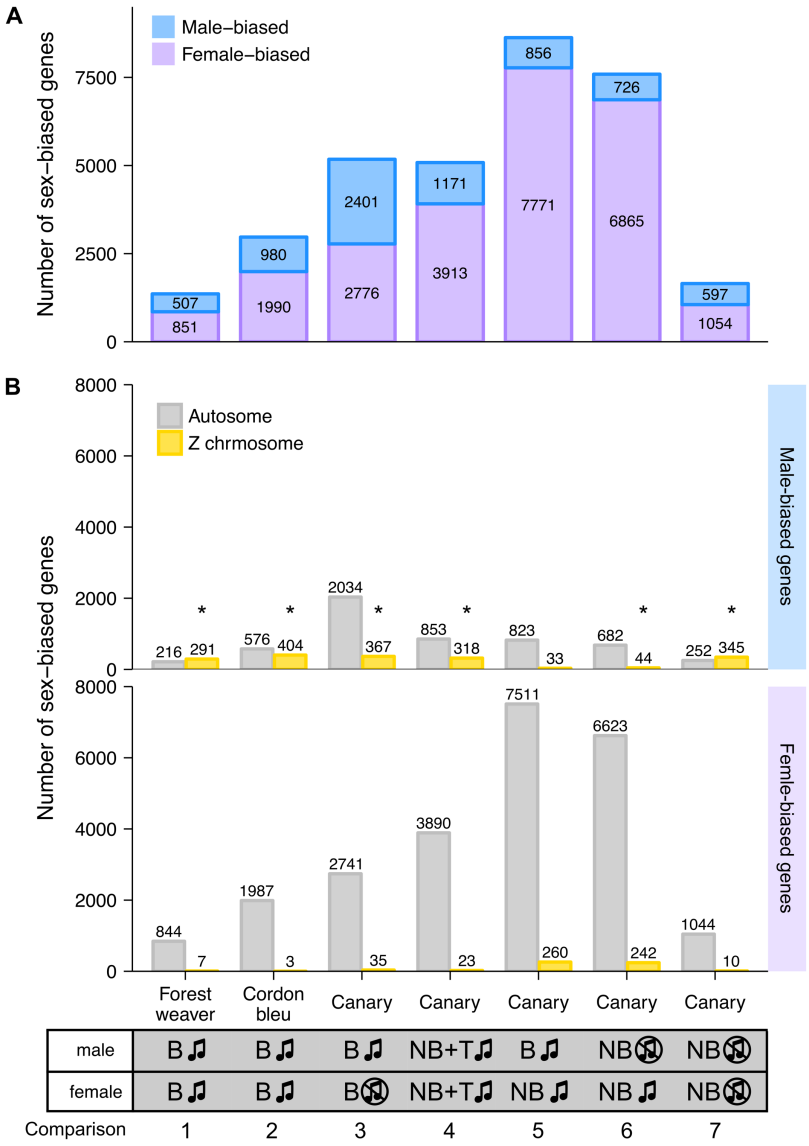
extreme values that are within 1.5 times the IQR (interquartile range) of the hinge. T: testosterone; music note: presence of singing behaviour. **B**, Hierarchical clustering showing that phylogenetic relatedness accounts for the most variation in the HVC transcriptomes of females and males of the forest weaver, cordon-bleu, and canary. Among the seven canary groups, the testosterone-treated animals were the least similar to the untreated canaries. **C**, A PCA of the HVC transcriptomes of the seven canary groups distinguished male birds (PC2 < 0) from female birds (PC2 > 0). Each point is colour-coded by group (see B) and represents an HVC sample.

148 **Sex-biased gene expression in the HVC of songbirds**

149 We defined the degree of sex differences in HVC gene expression levels as the number of
150 sex-biased genes whose expression levels showed significant differences between males
151 and females of the same species (male-biased genes showed higher expression levels in
152 males than in females, $\log_2$ (fold change) ≥ 0.5; female-biased genes presented higher
153 expression levels in females than in males, $\log_2$ (fold change) ≤ -0.5). For canary, we
154 quantified the sex differences from five distinct male-to-female comparisons (Figure 2A,
155 comparisons 3 to 7).

156 Interestingly, we found persistent sex differences in HVC gene expression levels
157 from all the comparisons, regardless of the degree of sex differences in singing behaviour
158 (Figure 2A). The majority (> 90%) of sex-biased genes were lowly biased (0.5 ≤ |$\log_2$ (fold
159 change)| < 1, Supplementary Figure 5). The comparison of forest weaver (the female birds
160 can sing identical songs to the males; least differences in the HVC volumes between the
161 sexes; Figure 1A) showed the least prominent sex differences in the HVC transcriptomes
162 across all the comparisons performed in this study. Nevertheless, the extent of sex
163 differences in gene expression levels was extensive (> 1,300 sex-biased genes, Figure 2A,
164 comparison 1). The comparison showing the second-lowest degree of sex-differences was
165 the comparison of nonbreeding male and nonbreeding female canaries (> 1,500 sex-
166 biased genes, Figure 2A, comparison 7), both lacking singing behaviour, but HVC volume
167 was male biased (Figure 1A). Testosterone implantation induced singing behaviour and
168 increased the HVC volume of both male and female canaries, although a markedly smaller
169 HVC volume continued to be observed in these females (Figure 1A and Supplementary
170 Figure 1C). Such processes in females are referred to as "masculinization" of the female
171 brain and behaviour (Arnold and Gorski, 1984; Wade, 2001). However, the gene
172 expression levels of the testosterone-stimulated singing female canaries were very
173 different from those of the male canaries administered the same treatment (> 5,000 sex-
174 biased genes, Figure 2A, comparison 4). The comparisons showing the most striking sex
175 differences were the two that included spontaneously singing female canaries. As

176 revealed in these comparisons, many genes showed markedly different regulation in
177 spontaneously singing female canaries, as reflected by expression levels, compared with
178 breeding singing male canaries (> 8,600 sex-biased genes, comparison 5) and
179 nonbreeding nonsinging male canaries (> 7,500 sex-biased genes, comparison 6, Figure
180 2A).



181

182 **Figure 2. Female- and male-biased sex differences in HVC gene expression levels in**
183 **songbirds.**

A, The bar graph summarizes the number of sex-biased genes in the HVC transcriptome identified from each male-to-female comparisons. The phenotypes of the groups being compared in each male-to-female comparison are listed at the bottom of the graph. **B,** The bar graph summarizes the number of autosomal and Z chromosomal genes classified as male- and female-biased genes in the HVC transcriptome based on each male-to-female comparison. Abbreviations: B, breeding; NB, nonbreeding; T, testosterone implantation. The music note indicates the presence or absence of singing behaviour; the asterisk (*)

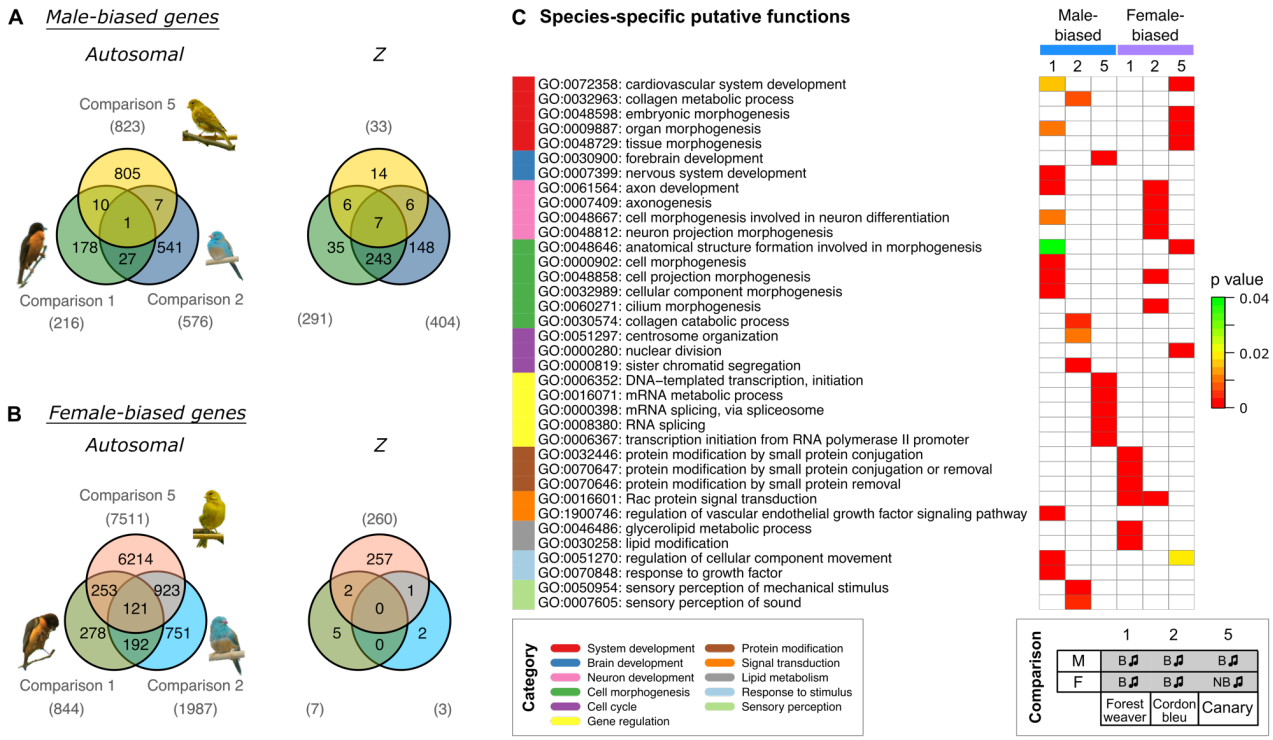
indicates enrichment for Z chromosomal genes (adjusted $p < 0.05$, Fisher's exact test followed by the Bonferroni correction for multiple testing).

184 For each comparison, we decomposed the sex-biased genes into two classes,
185 autosomal genes and Z sex chromosomal genes, to assess whether the sex-biased genes
186 were mainly concentrated on the sex chromosomes. Annotation for W chromosomal
187 genes was unfortunately not available in our study. We used Fisher's exact test to evaluate
188 whether a particular chromosome was enriched in sex-biased genes. We observed
189 enrichments of Z chromosomal genes among the male-biased genes identified in most of
190 the comparisons (Figure 2B and Supplementary Table 3). This observation was not
191 surprising because males are the homogametic sex (have two copies of Z), whereas
192 females are the heterogametic sex (have one copy of Z and one copy of W) in birds, and
193 the dosage compensation for Z-linked genes is less complete in birds than that in
194 mammals (Itoh et al., 2007; Mank, 2013; Nätt et al., 2014; Uebbing et al., 2015; Wolf and
195 Bryk, 2011). However, one canary comparison (breeding singing males and nonbreeding
196 singing females, Figure 2B, comparison 5) was an exception (Supplementary Table 3). In
197 this comparison, the spontaneously singing female canaries expressed a high number of
198 Z chromosomal genes (260 genes, approximately 24% of all Z chromosomal genes) at
199 higher levels than the breeding singing male canaries, which eliminated the male
200 enrichment of Z chromosomal genes.

201 Sex-biased genes were not only Z chromosomal genes; in fact, sizeable numbers
202 of the autosomal genes were found to be sex-biased genes, particularly female-biased
203 genes (male-biased genes: >40%; female-biased genes: >96%, Figure 2B). Many studies
204 have reported sex differences in the gene expression levels in multiple tissues and animal
205 species (Frésard et al., 2013; Itoh et al., 2007; Nätt et al., 2014; Uebbing et al., 2015; Wolf
206 and Bryk, 2011; Yang et al., 2006). Gene expression levels are generally under tight
207 regulation, and perturbed expression levels might result in functional consequences. For
208 example, several genes, including transcription factors, modulate distinct gene sets
209 depending on their expression levels (Birchler et al., 2001; Doghman et al., 2013; Schulz,
210 2017).

211 We compared the male-biased genes in a pairwise manner across the
212 comparisons and found that the male-biased autosomal genes showed low similarity to
213 each other (Supplementary Figure 6A), whereas the male-biased Z chromosomal genes
214 showed some similarity across the groups (Supplementary Figure 6B). Pairwise
215 comparison of the female-biased genes across groups showed low similarity, and this
216 finding was obtained for both autosomal genes (Supplementary Figure 6C) and Z

217 chromosomal genes (Supplementary Figure 6D). Fisher's exact test showed that
218 chromosome 2 (comparisons 1, 4, and, 5) and chromosome 3 (comparisons 1, 2, and 4)
219 are hotspots for female-biased genes; the enrichments of these chromosomes were
220 observed in four out of the seven pairwise comparisons (Supplementary Table 3).



221

222 **Figure 3. Cross-species comparisons of sex-biased genes expressed in the HVC of**
223 **birds with singing phenotypes shows high species specificity.**

Venn diagrams of male-biased (**A**) and female-biased (**B**) autosomal and Z chromosomal genes. The numbers of sex-biased genes are indicated in parentheses. **C,** GO term enrichment analysis predicting the biological functions of sex-biased (autosomal and Z chromosomal) genes. The GO terms were categorized and are colour coded. Bonferroni-adjusted P values are shown by colour scales in the heatmap. Only five GO terms with the lowest Bonferroni-adjusted P values are shown for each set (see Supplementary Table 4 for the complete results). The phenotypes of the groups used for comparisons 1, 2, and 5 are listed at the bottom. Abbreviations: B, breeding; NB, nonbreeding. The music note indicates the presence or absence of singing behaviour

224 Further examination of the sex-biased genes (comparisons 1, 2, and 5) identified
225 in the groups of naturally singing birds revealed that almost none of the male-biased
226 autosomal genes (Figure 3A) and none of the female-biased Z chromosomal genes (Figure
227 3B) were shared across the three songbird species. In addition, relatively low numbers of
228 male-biased Z chromosomal genes (Figure 3A) and female-biased autosomal genes
229 (Figure 3B) were shared across the three songbird species. This observation indicates
230 that only a small set of genes were regulated in the same manner (male- or female-biased)

231 between species even though all birds exhibit singing behaviours. Moreover, a Gene
232 Ontology (GO) enrichment analysis suggested that male- and female-biased genes did not
233 functionally converge to similar pathways in the three studied species (Figure 3C).

234 **The plasma androgen levels substantially alter the HVC transcriptomes of canaries**

235 With the aim of understanding the features that distinguish the HVC transcriptomes in
236 within-species contexts, we performed a principal component analysis (PCA) of the HVC
237 transcriptomic data from the seven canary groups to identify variables that would explain
238 the most variation in the data. By calculating the correlation coefficient between the PCs
239 and the variables (plasma testosterone levels (Supplementary Figure 1), HVC volume
240 (Figure 1A), sex, and singing), we identified the variables that were highly correlated with
241 the most important PCs. PC1 explained 32% of the data variance (Figure 1C and
242 Supplementary Figure 4) and was strongly correlated with the blood plasma androgen
243 concentrations and the presence of singing activity (plasma androgen level: Pearson's r =
244 -0.57, Bonferroni-adjusted $p = 1.0 \times 10^{-4}$; singing: Pearson's r = -0.51, Bonferroni-adjusted
245 p = 0.001, Supplementary Table 2). Sex and HVC volume were strongly correlated with
246 PC2 (sex: Pearson's r = -0.91, Bonferroni-adjusted $p = 4.4 \times 10^{-16}$; HVC volume: Pearson's
247 r = -0.73, Bonferroni-adjusted $p = 2.1 \times 10^{-7}$), which explained 19% of the variance in the
248 data (Supplementary Table 2). Taken together, the hierarchical clustering and PCA results
249 suggest that although phylogenetic relationships dominate the variation between
250 songbird species, the circulating testosterone levels, the presence of singing activity, and
251 sex identity dominate the HVC gene expression patterns within a single species.

252 **The application of testosterone does not yield an HVC transcriptome that mimics**
253 **that of natural singing canaries**

254 To investigate the molecular mechanisms underlying the singing phenotype, we
255 performed a differential gene expression analysis of the HVC transcriptomes of the two
256 groups of singing canary females (nonbreeding spontaneously singing females and
257 nonbreeding testosterone-stimulated singing females) against the nonsinging
258 nonbreeding female canaries. Similarly, a differential gene expression analysis of the HVC
259 transcriptomes of the two groups of singing male canaries (breeding singing males and
260 nonbreeding testosterone-stimulated singing males) with nonsinging nonbreeding male
261 canaries was performed. A substantial number of genes were differentially expressed
262 between singing and nonsinging birds (Figure 4A, nonbreeding singing females: 4,125
263 genes; nonbreeding testosterone-stimulated singing females: 7,702 genes; breeding

264 singing males: 3,359 genes; nonbreeding testosterone-stimulated singing males: 5,106
265 genes). Approximately 65% of the differentially expressed genes found in the naturally
266 singing birds overlapped with those found in the testosterone-treated birds of the same
267 sex (male: 65%; female: 62%). However, both testosterone-treated groups had markedly
268 higher numbers of differentially expressed genes than the naturally singing groups of the
269 same sex (Figure 4A). Thus, although testosterone implantation induced singing in
270 females and males, most testosterone-responsive genes (female: 67%; male: 57%) might
271 not be necessary for singing behaviour *per se* but rather a response to nonphysiological
272 levels of testosterone (Supplementary Figure 1). Alternatively, mechanisms that lead to
273 the first song in life might be very different from those that reinduce singing in animals
274 that sang before (Vellema et al., 2019a). Because male canaries sing regularly starting
275 from approximately 50 days of age (Nottebohm et al., 1986) while most female canaries
276 never sing (Ko et al., 2020), the induction of singing in females by testosterone likely
277 activated genes related to first-time singing, whereas in males, this treatment activated
278 genes related to reinduced singing.

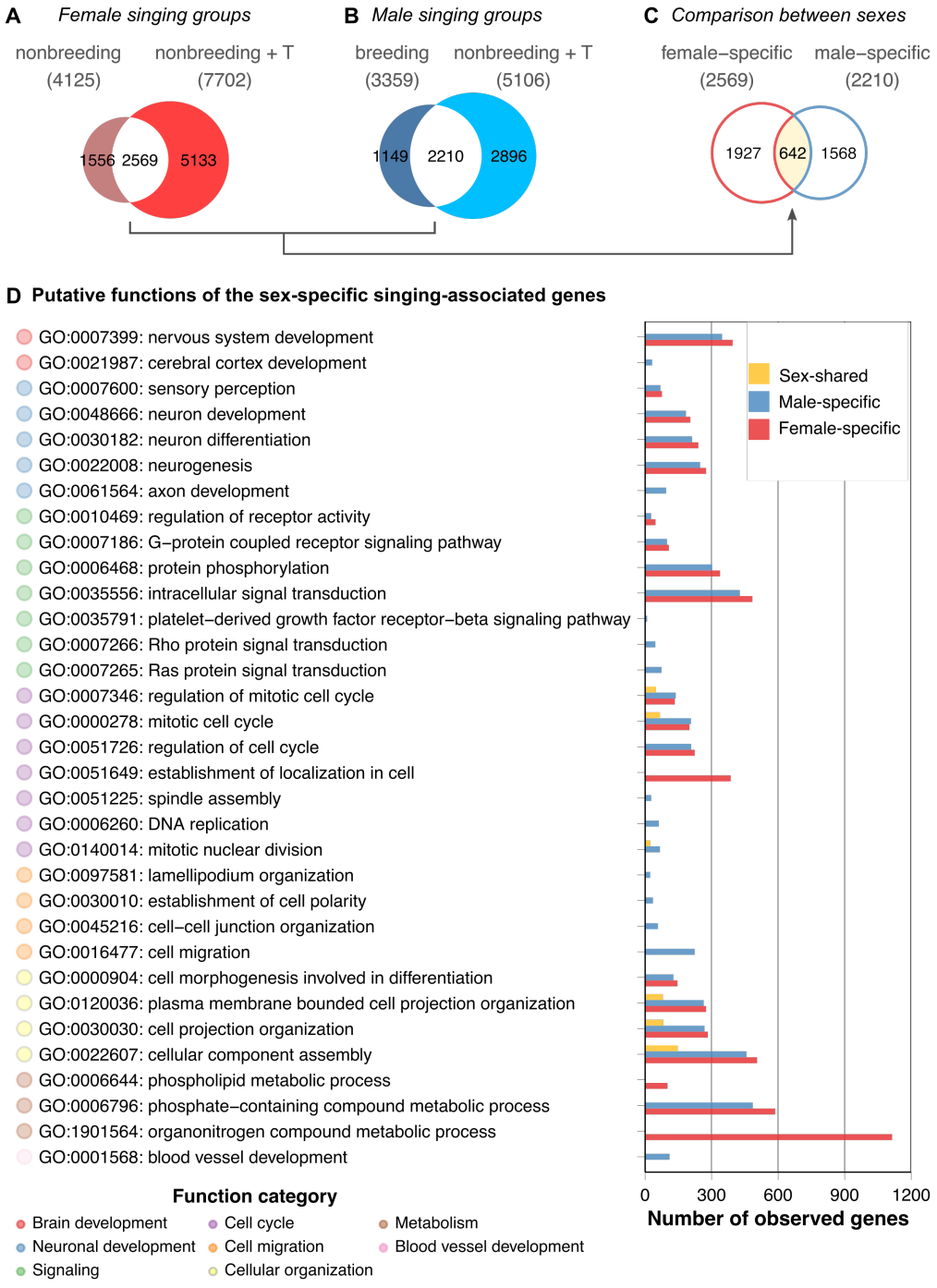
279 **The majority of differentially expressed genes between singing and nonsinging**
280 **canaries are sex-specific**

281 We examined whether the overlapping gene sets of the female canary signing groups
282 were similar to the overlapping gene sets of the singing males (Figure 4A). The Venn
283 diagram indicated that only approximately 25% (642 of 2569 genes) of the female-
284 specific expressed genes and 29% (642 of 2210 genes) of the male-specific expressed
285 genes were shared by both sexes of singing canaries.

286 **The sex-specific genes show functional overlap**

287 To understand the putative biological functions of the female-specific, male-specific, and
288 sex-shared expressed genes in the HVC, we performed a GO term enrichment analysis.
289 Interestingly, the results suggested that the female- and male-specific expressed genes
290 largely overlapped at the functional level (Figure 4B and Supplementary Table 5). GO
291 terms such as nervous system development (GO:0007399), neuron development
292 (GO:0048666), and intracellular signal transduction (GO:0035556) were shared among
293 the sexes. Female-specific GO terms were mainly related to cellular maintenance, such as
294 organonitrogen compound metabolic process (GO:1901564) and phospholipid metabolic
295 process (GO:0006644). In contrast, the male-specific GO terms were axon development
296 (GO:0061564), DNA replication (GO:0006260), cell migration (GO:0016477), and blood

297 vessel development (GO:0001568). In summary, although only approximately one-fourth
298 to one-third of the sex-specific expressed genes are shared among the sexes in terms of
299 their identities, most of the predicted functions of the sex-specific expressed genes are
300 nevertheless sex-shared in singing canaries.



301

302 **Figure 4. HVC transcriptomes of naturally singing canaries are dissimilar from**
303 **testosterone-stimulated singing canaries.**

Venn diagrams comparing (**A**) the two female singing groups, (**B**) the two male singing groups, and (**C**) the female- and male-biased genes resulting from comparisons A and B. Note that only approximately

20% of the genes were shared among the singing females and approximately 26% were shared among the singing males. The majority of these overlapping genes were not sex-shared; approximately 75% of female intersections were specific to female groups, whereas 71% of male intersections were specific to male groups. **D**, GO term enrichment analysis of female-specific, male-specific, and sex-shared expressed genes derived from comparisons A to C. Because many putative functions show similarities between male- and female-specific expressed genes, the results suggest functional intraspecies convergence based on sex-specific gene expression in the HVC. See Supplementary Table 5 for the full results of the GO term enrichment analysis.

Testosterone can be converted to 5α-dihydrotestosterone (5α-DHT) and 17β-estradiol, which activate androgen receptor (AR) and estrogen receptors (ERα and ERβ), respectively. Both AR (encoded by *AR*) and ERs (ERα encoded by *ESR1* and ERβ encoded by *ESR2*) are transcription factors that play important roles in the transcription of numerous genes (Bourdeau et al., 2004; Pihlajamaa et al., 2015; Takayama et al., 2007; Wilson et al., 2016). Interestingly, *ESR1* was specifically expressed in males, whereas *ESR2* was female-specifically expressed in the HVC of canaries, which suggests that estrogen receptor paralogues could provide finer-tuned mechanisms for sex-specific regulation. Empirical results have shown that ERα and ERβ bind to the same estrogen response element motifs (Zhao et al., 2010) and might functionally overlap in some tissues. Moreover, ER paralogues show sex differences in expression levels and tissue specificity (Zhang et al., 2017). In quail, the administration of an agonist specific to ERβ on embryonic day 7 demasculinizes male sexual behaviour and midbrain nuclei characteristics in Japanese quails, whereas an agonist specific to ERα does not exert this effect (Court et al., 2020). The specific roles of each ER paralogue in the adult HVC of male and female songbirds warrant further investigation. Until now, ERα but not ERβ was expected to regulate the function of HVC neurons of adult songbirds in addition to the AR (Frankl-Vilches and Gahr, 2017).

Conclusion

In this study, we investigated the sex differences in the gene expression patterns in the HVC of three songbird species with different levels of sex-specific singing and several different song-related phenotypes between male and female canaries. Our inter- and intraspecies comparisons yielded large-scale transcriptional sex differences regardless of singing behaviour. Instead, fundamental sex differences in gene expression levels were found in the HVC, and these differences were highly species-specific. By leveraging several experimental groups of canaries, we found that the plasma androgen levels and sex were the major contributors to the variations in the HVC transcriptome. Although

331 testosterone reliably induced singing in both female and male canaries, testosterone
332 treatment did not alter the transcriptome to imitate that of natural singing birds. Our
333 results suggest that female and male canaries rely on different gene networks for singing
334 behaviour, but the sex-specific gene networks might show functional convergence.

335 Materials and methods

336 Animals

337 Forest weaver (*Ploceus bicolor*) pairs were observed in their breeding territories in
338 eastern South Africa; the songs of the pairs were recorded to ensure that both mates were
339 singing. Subsequently, the animals were caught and sacrificed in accordance with permits
340 issued by the local authorities (Chief Professional Officer for Research at the Natal Parks,
341 Game and Fish Preservation Board, P. O. B. 662, Pietermaritzburg 3200). Blue-capped
342 cordon-bleus (*Uraeginthus cyanocephalus*) and canaries (*Serinus canaria*) were bred at
343 the animal facilities at the Max Planck Institute for Ornithology in Seewiesen, Germany.
344 The procedures used for animal housing and welfare complied with the European
345 directives for the protection of animals used for scientific purposes (2010/63/EU), and
346 the protocols were approved by the Government of Upper Bavaria (AZ No. 55.2-1-54-
347 2532-181-12). Adult canaries and blue-capped cordon-bleus (aged at least 1 year) were
348 housed in pairs under long-day conditions (light:dark = 14:10 hours), and their breeding
349 activities were monitored. The birds (male canaries and cordon-bleus) were sacrificed
350 after singing activity was observed during the breeding season. Nonbreeding canaries
351 were housed pairwise with a 9-hour light:15-hour dark schedule for at least 8 weeks
352 before their singing activity was monitored (generally starting in late September).
353 Nonbreeding male and female canaries were maintained alone in sound-attenuated boxes
354 (70 × 50 × 50 cm), recorded continuously during the song monitoring phase (two weeks)
355 and sacrificed after confirmation of no singing activity. Additional groups of male and
356 female nonbreeding and nonsinging canaries were implanted with testosterone for 2
357 weeks and sacrificed after observation of singing activity. Water and food were available
358 *ad libitum*. The sex was confirmed by PCR using the P2 and P8 primers for CHD genes
359 (Griffiths et al., 1998) and by visual inspection of the reproduction system after sacrifice.
360 We previously found that six nonbreeding female canaries exhibited singing behaviour
361 during long-term monitoring, and their songs have been described in detail (Ko et al.,
362 2020). In this study, we included the transcriptomes of these birds. All the birds were
363 sacrificed via an overdose of isoflurane followed by decapitation, the body weights were
364 recorded, and the brains were dissected, weighed, snap-frozen on dry ice and stored at -

365 80°C until further use. Supplementary Table 1 summarizes the information of the
366 experimental groups used in this study.

367 **Testosterone implantation**

368 A Silastic™ tube (Dow Corning; 1.47-mm inner diameter, 1.96-mm outer diameter, 0.23-
369 mm thickness) was cut to a length of 7 mm and loaded with testosterone (86500, Fluka)
370 as densely as possible. The two ends of the Silastic™ tube were sealed with silicone
371 elastomer (3140, Dow Corning). After closure, the implants were cleaned with 100%
372 ethanol to remove testosterone particles and then immersed in ethanol overnight in a
373 hood to ensure no leakage at either end. Implants with apparent dampness were
374 discarded. One day prior to the implantation, the implants were incubated in 0.1 M
375 phosphate buffered saline (PBS) overnight to enable the immediate release of
376 testosterone upon implantation (Rasika et al., 1994). We started implantation at
377 approximately 8:30 am (immediately after the light was turned on in the morning), which
378 resulted in a 20-minute interval between birds based on the scarification time. A small
379 incision was made on the back of the bird over the pectoral musculature, and one
380 testosterone implant was placed subcutaneously. The skin was closed by the application
381 of tissue glue. After 2 weeks, the animals were sacrificed. The testosterone implants were
382 checked, and this inspection revealed that the implants were all in place and were not
383 empty at the end of the experiments.

384 **Measurement of the HVC volume**

385 One series of 20-μm sections was mounted on RNase-free Superfrost slides, subjected to
386 Nissl staining with 0.1% thionin (Sigma-Aldrich), dehydrated, immersed in xylene and
387 cover-slipped with Roti-Histokitt II mounting medium (Roth). The HVC areas (typically 8-
388 10 slices) were measured with a Leica DM6000 B microscope connected to a computer-
389 based image-analysis system (IMAtec). All brains were coded such that the observers
390 were blind towards any additional information about the sections they measured during
391 the delineations. The volumes were derived from the summed area measurements
392 multiplied by the section thickness and the intersection distance.

393 **Radioimmunoassay of plasma testosterone**

394 Blood was sampled (< 150 μl) at the time sacrifice. All blood samples were taken between
395 8 and 11 am and were taken within 3 min after caught to avoid the effect of handling
396 (Wingfield et al., 1982). Blood samples were centrifuged (5,000 rpm, 10 min) to separate

397 the plasma from blood cells. Testosterone metabolites were measured with a
398 radioimmunoassay using a commercial antiserum against testosterone (T3-125,
399 Endocrine Sciences, Tarzana, USA) as previously described (Goymann et al., 2002).
400 Standard curves and sample concentrations were calculated with Immunofit 3.0
401 (Beckman Inc. Fullerton, CA) using a four-parameter logistic curve fit and corrected for
402 individual recoveries.

403 The testosterone concentrations were assayed in duplicate in five separate assays. The
404 mean extraction efficiency for plasma testosterone was 85.0 ± 3.9% (mean ± SD, N = 42).
405 The lower detection limits of the testosterone assays were 0.34, 0,35, 0.36, 0.38 and 0.35
406 pg per tube, and all the samples were above the detection limit. The intra-assay
407 coefficients of variation of a chicken plasma pool were 8.7%, 3.4%, 12.8%, 1.9%, and
408 4.4%. The interassay coefficient of variation as determined by the variation in the chicken
409 plasma pool between all the assays was 5.1%. Because the testosterone antibody used
410 shows significant cross-reactions with 5α-dihydrotestosterone (44%), our measurement
411 might include a fraction of 5α-DHT.

412 **Microarray procedures and annotation**

413 The birds were killed by an overdose of isoflurane, and their brains were snap-frozen on
414 dry ice. The brains were sectioned sagittally into four series of 40-µm sections and two
415 series of 20-µm sections with a cryostat (Jung CM3000 Leica). The 40-µm sections were
416 mounted on glass slides for subsequent tissue dissection, whereas the thin sections were
417 mounted on RNase-free Superfrost slides for Nissl staining. All sections were stored at -
418 80°C until further processing. For total RNA extraction, the song control nucleus HVC and
419 the visual area of the entopallium were dissected from the abovementioned 40-µm
420 sections under a stereomicroscope (typically 24-32 slices for the HVC and 16-20 slices for
421 the entopallium) and transferred into an Eppendorf tube containing 340 µl of RLT buffer
422 mixture (containing DTT, Qiagen). This dissection procedure using rather thin sections
423 reduces the contamination of HVC tissue with surrounding tissue. RNA was then
424 extracted using the RNeasy® Micro Kit (Qiagen) with the optional DNase digest step. The
425 RNA quality was assessed using the Agilent Model 2100 Bioanalyzer (Agilent
426 Technologies), and the RNA concentrations were assessed using a Nanodrop 1000
427 spectrometer (Thermo Fisher Scientific). The RNA quality of all samples was good (RIN >
428 7). The purified total RNA samples (at least 100 ng per sample) were subsequently
429 processed and hybridized using the Ambion WT Expression Kit and the Affymetrix WT
430 Terminal Labelling and Controls Kit. The resulting cDNA was hybridized to the Custom

431 Affymetrix Gene Chip® MPIO-ZF1s520811 Exon Array, which has been used successfully
432 and validated in cross-species hybridization studies (Dittrich et al., 2014; Frankl-Vilches
433 et al., 2015). The 5.7 million male zebra finch-specific probes spotted on this array
434 correspond to approximately 4,711,133 probe sets and hence to 25,816 transcripts
435 published in public databases (NCBI and Ensembl). We annotated more than 90% of the
436 transcripts to 12,729 human orthologous genes using several publicly available databases
437 (Ensembl, GenBank, UniProt, and DAVID (Benson et al., 2005; Consortium, 2015; Flicek
438 et al., 2014; Huang et al., 2008, 2009; Yates et al., 2016)) and commercial databases (El
439 Dorado, Genomatix, Precigen Bioinformatics Germany GmbH (PBG), Munich, Germany).
440 Hybridization was performed for 16 hours at 45°C and 60 rpm in a GeneChip
441 Hybridization Oven 640 (Affymetrix). The arrays were washed, stained, and scanned
442 using the Affymetrix GeneChip Fluidics Station 450 and the Affymetrix GeneChip scanner
443 3000 7G. The CEL files were generated using Affymetrix® GeneChip® Command
444 Console® Software (AGCC), and for quality control, the success of individual
445 hybridizations was assessed using Affymetrix® Expression Console™ software.

446 Differential expression was calculated using ChipInspector software version 21 (El
447 Dorado Database version E28R1306 (Genomatix)). ChipInspector is a single probe-based
448 analysis tool for microarray data that can show increased sensitivity compared with that
449 obtained with conventional probe set-based analyses, such as robust multiarray analysis.
450 ChipInspector consists of four steps: single probe-transcript annotation (ensuring up-to-
451 date annotation), total intensity normalization, SAM (significance analysis of microarrays,
452 adapted to single-probe handling) analysis (Tusher et al., 2001), and transcript
453 identification based on significantly changed probes (Cohen et al., 2008). We set the delta
454 threshold to 0 (to control the false positive rate) and used the groupwise exhaustive
455 comparison tool in ChipInspector. The minimum coverage for each transcript was set to
456 10 significant probes, and the minimum expression difference was $|\log_2(\text{fold change})| \geq$
457 0.5. The significantly differentially expressed transcripts obtained were annotated to
458 human orthologous genes as described above. For transcripts belonging to the same
459 genes, the average expression was calculated if all transcripts were regulated in the same
460 manner (e.g., all upregulated or all downregulated). If a gene contains both up- or
461 downregulated transcripts, the transcripts that showed changes in expression in the
462 minority direction were discarded (<40% of the total transcripts), and the average
463 expression was instead calculated from the remaining transcripts (60% or more).
464 Transcripts without human orthologous gene annotation were removed prior to
465 subsequent analyses. We were cautious about possible cross-species bias in
466 hybridization, and differential expression analyses were only performed between two

467 conspecific groups. The microarray data discussed in this publication have been
468 deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible
469 through the GEO Series accession number GSE83674.

470 **Hierarchical clustering**

471 The normalized gene expression levels (across all transcriptome samples) were
472 calculated using the "justRMA" function of the R package "affy" v1.66.0 (Gautier et al.,
473 2004), and the expression levels were further collapsed to gene levels using the
474 "collapseRows" functions of the R package "WGCNA" v1.69 (Langfelder and Horvath,
475 2008, 2012). The group-level expression levels were calculated by averaging the gene
476 expression levels. Spearman's ρ was calculated between samples (or groups) using the
477 "cor" function, and a Euclidean distance matrix was calculated (dist(1-cor)) and used for
478 hierarchical clustering analysis (method: complete) with the "hclust" function. This
479 computation was performed with R v4.0.2 (R_Core_Team, 2020) using the "dendextend"
480 v1.14.0 package for visualization (Galili, 2015).

481 **Principal component analysis (PCA)**

482 PCA (Ringner, 2008) was performed using the normalized and collapsed HVC gene
483 expression data (described in the hierarchical clustering section) from 41 canaries
484 (Supplementary Table 1). The computation was performed using the "pca" function of the
485 R package "pcaMethods" (v1.60.0 (Stacklies et al., 2007), method = svd). The data were
486 centred but not scaled because the expression data had already been normalized.

487 **Fisher's exact testing for chromosomal enrichment**

488 Fisher's exact test was used to test whether sex-biased genes were enriched on a
489 chromosome by comparing the gene lists of interest to the zebra finch annotation
490 (https://www.ncbi.nlm.nih.gov/genome/?term=zebra+finch). We used the "fisher.test"
491 function in R (R_Core_Team, 2020) with the alternative set to "greater" for enrichment.
492 The male- and female-biased genes identified from each comparison as well as the male-
493 specific, female-specific, and sex-shared genes were tested separately. The P values were
494 adjusted using the Bonferroni correction to account for multiple comparisons using
495 "p.adjust" in R (R_Core_Team, 2020).

496 **Gene Ontology (GO) term enrichment analysis**

497 We used ClueGO v2.2.4, an application built in the Cytoscape environment v3.3.0 (Bindea
498 et al., 2009; Shannon et al., 2003), to predict the putative biological functions of the genes
499 of interest. This software performs GO term enrichment hierarchical analyses and fuses
500 GO terms with similar functions. The enrichment was determined by the right-sided test
501 and corrected using the Bonferroni step-down method considering multiple
502 comparisons.

503 Data accessibility

504 The microarray CEL files (GEO Series accession number GSE83674) are available on
505 NCBI's Gene Expression Omnibus. The processed microarray data, plasma androgen
506 levels, HVC volume measures, and all scripts used for analysis and visualization are
507 available in GitHub (https://github.com/maggieMCKO/SongbirdSexDiff).

508 Authors' contributions

509 M-CK collected the nonbreeding testosterone-implanted canaries, analysed and
510 visualized the transcriptomic data and was a major contributor to the writing of the
511 manuscript. AB and M-CK performed the RNA extractions and microarray hybridizations.
512 MG and M-CK isolated the tissues for microarray. MG collected the forest weavers and
513 cordon-bleus. MG conceived the study and obtained financial support for the study. M-CK,
514 CF-V and MG drafted the manuscript. All the authors read and approved the final
515 manuscript.

516 Acknowledgements

517 We thank Dieter Schmidl for catching the forest weavers, Stefan Leitner for providing the
518 breeding canaries, Roswitha Brighton and David Witkowski for maintaining the breeding
519 colonies of the blue-capped cordon-bleus and the canaries, Doris Walcher for providing
520 data from the breeding male and female canaries as well as the nonbreeding nonsinging
521 male and female canaries, and Wolfgang Goymann and Monika Trappschuh for the RIA of
522 the plasma testosterone. Furthermore, we express our appreciation to Maude Baldwin,
523 Glenn Cockburn, Falk Dittrich, Lisa Gill, Vincent Van Meir, Keren Sadanandan, and Michiel
524 Vellema for their comments on previous versions of this manuscript and to the
525 International Max Planck Research School for Organismal Biology for the training and
526 support provided.

Competing interests

The authors declare no competing financial interests.

Funding

This research was supported by Max-Planck-Gesellschaft to MG. The funder had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Alward, B.A., Balthazart, J., and Ball, G.F. (2013). Differential effects of global versus local testosterone on singing behavior and its underlying neural substrate. Proc Natl Acad Sci U S A *110*, 19573-19578.

Arnold, A.P., and Gorski, R.A. (1984). Gonadal steroid induction of structural sex differences in the central nervous system. Annu Rev Neurosci *7*, 413-442.

Ball, G.F. (2016). Species variation in the degree of sex differences in brain and behaviour related to birdsong: adaptations and constraints. Philos Trans R Soc Lond B Biol Sci *371*.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., and Wheeler, D.L. (2005). GenBank. Nucleic Acids Research *33*, D34-D38.

Bernard, D.J., and Ball, G.F. (1997). Photoperiodic Condition Modulates the Effects of Testosterone on Song Control Nuclei Volumes in Male European Starlings. General and Comparative Endocrinology *105*, 276-283.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pagès, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics *25*, 1091-1093.

Birchler, J.A., Bhadra, U., Bhadra, M.P., and Auger, D.L. (2001). Dosage-Dependent Gene Regulation in Multicellular Eukaryotes: Implications for Dosage Compensation, Aneuploid Syndromes, and Quantitative Traits. Developmental Biology *234*, 275-288.

Bourdeau, V.r., Deschênes, J., Métivier, R.l., Nagai, Y., Nguyen, D., Bretschneider, N., Gannon, F., White, J.H., and Mader, S. (2004). Genome-Wide Identification of High-Affinity Estrogen Response Elements in Human and Mouse. Molecular Endocrinology *18*, 1411-1427.

Brenowitz, E. (1991). Altered perception of species-specific song by female birds after lesions of a forebrain nucleus. Science *251*, 303 - 305.

Brenowitz, E.A., Arnold, A.P., and Levin, R.N. (1985). Neural correlates of female song in tropical duetting birds. Brain Res *343*, 104-112.

Cohen, C.D., Lindenmeyer, M.T., Eichinger, F., Hahn, A., Seifert, M., Moll, A.G., Schmid, H., Kiss, E., Grone, E., Grone, H.J., *et al.* (2008). Improved elucidation of biological processes linked to diabetic nephropathy by single probe-based microarray data analysis. PLoS One *3*, e2937.

Consortium, T.U. (2015). UniProt: a hub for protein information. Nucleic Acids Research *43*, D204-D212.

568 Court, L., Vandries, L., Balthazart, J., and Cornil, C.A. (2020). Key role of estrogen
569 receptor beta in the organization of brain and behavior of the Japanese quail. Horm
570 Behav *125*, 104827.

571 Del Negro, C., Gahr, M., Leboucher, G., and Kreutzer, M. (1998). The selectivity of sexual
572 responses to song displays: effects of partial chemical lesion of the HVC in female
573 canaries. Behavioural Brain Research *96*, 151-159.

574 Dittrich, F., Ramenda, C., Grillitsch, D., Frankl-Vilches, C., Ko, M.C., Hertel, M., Goymann,
575 W., Ter Maat, A., and Gahr, M. (2014). Regulatory mechanisms of testosterone-
576 stimulated song in the sensorimotor nucleus HVC of female songbirds. BMC Neurosci *15*,
577 128.

578 Doghman, M., Figueiredo, B.C., Volante, M., Papotti, M., and Lalli, E. (2013). Integrative
579 analysis of SF-1 transcription factor dosage impact on genome-wide binding and gene
580 expression regulation. Nucleic Acids Research *41*, 8896-8907.

581 Edgar, R., Domrachev, M., and Lash, A. (2002). Gene Expression Omnibus: NCBI gene
582 expression and hybridization array data repository. Nucleic Acids Res *30*, 207 - 210.

583 Flicek, P., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D.,
584 Clapham, P., Coates, G., Fitzgerald, S.*, et al.* (2014). Ensembl 2014. Nucleic Acids
585 Research *42*, D749-D755.

586 Frankl-Vilches, C., and Gahr, M. (2017). Androgen and estrogen sensitivity of bird song: a
587 comparative view on gene regulatory levels. Journal of Comparative Physiology A.

588 Frankl-Vilches, C., and Gahr, M. (2018). Androgen and estrogen sensitivity of bird song: a
589 comparative view on gene regulatory levels. Journal of Comparative Physiology A *204*,
590 113-126.

591 Frankl-Vilches, C., Kuhl, H., Werber, M., Klages, S., Kerick, M., Bakker, A., de Oliveira, E.H.,
592 Reusch, C., Capuano, F., Vowinckel, J.*, et al.* (2015). Using the canary genome to decipher
593 the evolution of hormone-sensitive gene regulation in seasonal singing birds. Genome
594 Biol *16*, 19.

595 Frésard, L., Morisson, M., Brun, J.-M., Collin, A., Pain, B., Minvielle, F., and Pitel, F. (2013).
596 Epigenetics and phenotypic variability: some interesting insights from birds. Genetics
597 Selection Evolution *45*, 16.

598 Fusani, L., Metzdorf, R., Hutchison, J., and Gahr, M. (2003). Aromatase inhibition affects
599 testosterone-induced masculinization of song and the neural song system in female
600 canaries. J Neurobiol *54*, 370 - 379.

601 Gahr, M. (2001). Distribution of sex steroid hormone receptors in the avian brain:
602 Functional implications for neural sex differences and sexual behaviors. Microscopy
603 Research and Technique *55*, 1-11.

604 Gahr, M., Metzdorf, R., Schmidl, D., and Wickler, W. (2008). Bi-directional sexual
605 dimorphisms of the song control nucleus HVC in a songbird with unison song. PLoS One
606 *3*, e3073.

607 Gahr, M., Sonnenschein, E., and Wickler, W. (1998). Sex difference in the size of the
608 neural song control regions in a dueting songbird with similar song repertoire size of
609 males and females. Journal of Neuroscience *18*, 1124-1131.

610 Galili, T. (2015). dendextend: an R package for visualizing, adjusting and comparing
611 trees of hierarchical clustering. Bioinformatics *31*, 3718-3720.

612 Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). affy--analysis of Affymetrix
613 GeneChip data at the probe level. Bioinformatics *20*, 307-315.

614 Geberzahn, N., and Gahr, M. (2011). Undirected (solitary) birdsong in female and male
615 blue-capped cordon-bleus (Uraeginthus cyanocephalus) and its endocrine correlates.
616 PLoS One *6*, e26485.

617 Genomatix (http://www.genomatix.de: Precigen Bioinformatics Germany GmbH
618 (PBG)).

619 Goymann, W., Möstl, E., and Gwinner, E. (2002). Non-invasive methods to measure
620 androgen metabolites in excrements of European stonechats, Saxicola torquata rubicola.
621 General and Comparative Endocrinology *129*, 80-87.

622 Griffiths, R., Double, M.C., Orr, K., and Dawson, R.J.G. (1998). A DNA test to sex most
623 birds. Molecular Ecology *7*, 1071-1075.

624 Gulledge, C.C., and Deviche, P. (1999). Photoperiod and testosterone independently
625 affect vocal control region volumes in adolescent male songbirds. Journal of
626 Neurobiology *36*, 550-558.

627 Gurney, M.E., and Konishi, M. (1980). Hormone-induced sexual differentiation of brain
628 and behavior in zebra finches. Science *208*, 1380-1383.

629 Hahnloser, R.H.R., Kozhevnikov, A.A., and Fee, M.S. (2002). An ultra-sparse code
630 underliesthe generation of neural sequences in a songbird. Nature *419*, 65-70.

631 Hall, M.L., Rittenbach, M.R.D., and Vehrencamp, S.L. (2015). Female song and vocal
632 interactions with males in a neotropical wren. Frontiers in Ecology and Evolution *3*.

633 Hall, Z.J., MacDougall-Shackleton, S.A., Osorio-Beristain, M., and Murphy, T.G. (2010).
634 Male bias in the song control system despite female bias in song rate in streak-backed
635 orioles (Icterus pustulatus). Brain Behav Evol *76*, 168-175.

636 Hartley, R.S., Chinn, M.S., and Ullrich, N.F.E. (1997). Left Syringeal Dominance in
637 Testosterone-Treated Female Canaries. Neurobiology of Learning and Memory *67*, 248-
638 253.

639 Hartley, R.S., and Suthers, R.A. (1989). Airflow and pressure during canary song: direct
640 evidence for mini-breaths. Journal of Comparative Physiology A *165*, 15-26.

641 Heid, P., Güttinger, H.R., and Pröve, E. (1985). The Influence of Castration and
642 Testosterone Replacement on the Song Architecture of Canaries (Serinus canaria).
643 Zeitschrift für Tierpsychologie *69*, 224-236.

644 Herpin, A., and Schartl, M. (2015). Plasticity of gene-regulatory networks controlling sex
645 determination: of masters, slaves, usual suspects, newcomers, and usurpators. EMBO
646 Rep *16*, 1260-1274.

647 Herrick, E.H., and Harris, J.O. (1957). Singing Female Canaries. Science *125*, 1299-1300.

648 Hoffmann, S., Trost, L., Voigt, C., Leitner, S., Lemazina, A., Sagunsky, H., Abels, M.,
649 Kollmansperger, S., Maat, A.T., and Gahr, M. (2019). Duets recorded in the wild reveal
650 that interindividually coordinated motor control enables cooperative behavior. Nat
651 Commun *10*, 2577.

652 Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2008). Systematic and integrative
653 analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols *4*,
654 44-57.

655 Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Bioinformatics enrichment tools:
656 paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids
657 Research *37*, 1-13.

658 Immelmann, K. (1968). Zur biologischen Bedeutung des Estrildidengesanges, Vol 109.

Irwin, D.E. (2018). Sex chromosomes and speciation in birds and other ZW systems. Molecular Ecology *0*.

Itoh, Y., Melamed, E., Yang, X., Kampf, K., Wang, S., Yehya, N., Van Nas, A., Replogle, K., Band, M.R., Clayton, D.F., *et al.* (2007). Dosage compensation is less effective in birds than in mammals. Journal of Biology *6*, 1-15.

Ko, M.-C., Van Meir, V., Vellema, M., and Gahr, M. (2020). Characteristics of song, brain-anatomy and blood androgen levels in spontaneously singing female canaries. Hormones and Behavior *117*, 104614.

Kreutzer, M., Vallet, E., and Doucet, S. (1992). Sexual responses of female canaries at the onset of song stimuli. Experientia *48*, 679-682.

Lambeth, L.S., Raymond, C.S., Roeszler, K.N., Kuroiwa, A., Nakata, T., Zarkower, D., and Smith, C.A. (2014). Over-expression of DMRT1 induces the male pathway in embryonic chicken gonads. Developmental Biology *389*, 160-172.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics *9*, 1-13.

Langfelder, P., and Horvath, S. (2012). Fast R Functions for Robust Correlations and Hierarchical Clustering. Journal of statistical software *46*, i11.

Leitner, S., Voigt, C., and Gahr, M. (2001a). Seasonal Changes in the Song Pattern of the Non-Domesticated Island Canary (Serinus Canaria) a Field Study. Behaviour *138*, 885-904.

Leitner, S., Voigt, C., Garcia-Segura, L., Van't Hof, T., and Gahr, M. (2001b). Seasonal activation and inactivation of song motor memories in wild canaries is not reflected in neuroanatomical changes of forebrain song areas. Horm Behav *40*, 160 - 168.

Leonard, S.L. (1939). Induction of Singing in Female Canaries by Injections of Male Hormone. Experimental Biology and Medicine *41*, 229-230.

Lobato, M., Vellema, M., Gahr, C., Leitão, A., de Lima, S.M.A., Geberzahn, N., and Gahr, M. (2015). Mismatch in sexual dimorphism of developing song and song control system in blue-capped cordon-bleus, a songbird species with singing females and males. Frontiers in Ecology and Evolution *3*, 117.

MacDougall-Shackleton, S.A., and Ball, G.F. (1999). Comparative studies of sex differences in the song-control system of songbirds. Trends in Neurosciences *22*, 432-436.

Madison, F.N., Rouse, M.L., Jr., Balthazart, J., and Ball, G.F. (2015). Reversing song behavior phenotype: Testosterone driven induction of singing and measures of song quality in adult male and female canaries (Serinus canaria). Gen Comp Endocrinol *215*, 61-75.

Mank, J.E. (2013). Sex chromosome dosage compensation: definitely not for everyone. Trends in Genetics *29*, 677-683.

Nätt, D., Agnvall, B., and Jensen, P. (2014). Large Sex Differences in Chicken Behavior and Brain Gene Expression Coincide with Few Differences in Promoter DNA-Methylation. PLoS ONE *9*, e96376.

Nixdorf, B.E., Davis, S.S., and DeVoogd, T.J. (1989). Morphology of Golgi-impregnated neurons in hyperstriatum ventralis, pars caudalis in adult male and female canaries. J Comp Neurol *284*, 337-349.

Nottebohm, F. (1980). Testosterone triggers growth of brain vocal control nuclei in adult female canaries. Brain Res *189*, 429-436.

Nottebohm, F., and Arnold, A.P. (1976). Sexual dimorphism in vocal control areas of the songbird brain. Science *194*, 211-213.

Nottebohm, F., Nottebohm, M., Crane, L., and Wingfield, J. (1987). Seasonal changes in gonadal hormone levels of adult male canaries and their relation to song. Behav Neural Biol *47*, 197 - 211.

Nottebohm, F., Nottebohm, M.E., and Crane, L. (1986). Developmental and seasonal changes in canary song and their relation to changes in the anatomy of song-control nuclei. Behavioral and neural biology *46*, 445-471.

Nottebohm, F., Stokes, T.M., and Leonard, C.M. (1976). Central control of song in the canary, Serinus canarius. The Journal of Comparative Neurology *165*, 457-486.

Odom, K.J., Hall, M.L., Riebel, K., Omland, K.E., and Langmore, N.E. (2014). Female song is widespread and ancestral in songbirds. Nat Commun *5*, 3379.

Ota, N., Gahr, M., and Soma, M. (2018). Couples showing off: Audience promotes both male and female multimodal courtship display in a songbird. Sci Adv *4*, eaat4779.

Pesch, A., and Güttinger, H.-R. (1985). Der Gesang des weiblichen Kanarienvogels. Journal für Ornithologie *126*, 108-110.

Pihlajamaa, P., Sahu, B., and Jänne, O.A. (2015). Determinants of Receptor- and Tissue-Specific Actions in Androgen Signaling. Endocrine Reviews *36*, 357-384.

Price, J.J. (2019). Sex differences in song and plumage color do not evolve through sexual selection alone: new insights from recent research. Journal of Ornithology *160*, 1213-1219.

Price, J.J., Lanyon, S.M., and Omland, K.E. (2009). Losses of female song with changes from tropical to temperate breeding in the New World blackbirds. Proceedings of the Royal Society B: Biological Sciences *276*, 1971-1980.

Price, T.D. (2002). Domesticated Birds as a Model for the Genetics of Speciation by Sexual Selection. Genetica *116*, 311-327.

Pryke, S., R. (2010). Sex chromosome linkage of mate preference and color signal maintains assortative mating between interbreeding finch morphs. Evolution *64*, 1301-1310.

R_Core_Team (2020). R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; (R Foundation for Statistical Computing).

Rasika, S., Nottebohm, F., and Alvarez-Buylla, A. (1994). Testosterone increases the recruitment and/or survival of new high vocal center neurons in adult female canaries. Proceedings of the National Academy of Sciences *91*, 7854-7858.

Ringner, M. (2008). What is principal component analysis? Nat Biotechnol *26*, 303-304.

Robinson, C.M., Snyder, K.T., and Creanza, N. (2019). Correlated evolution between repertoire size and song plasticity predicts that sexual selection on song promotes open-ended learning. Elife *8*.

Sæther, S.A., Sætre, G.-P., Borge, T., Wiley, C., Svedin, N., Andersson, G., Veen, T., Haavie, J., Servedio, M.R., Bureš, S.*, et al.* (2007). Sex Chromosome-Linked Species Recognition and Evolution of Reproductive Isolation in Flycatchers. Science *318*, 95.

Sayers, E.W., Beck, J., Brister, J.R., Bolton, E.E., Canese, K., Comeau, D.C., Funk, K., Ketter, A., Kim, S., Kimchi, A.*, et al.* (2020). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res *48*, D9-D16.

750 Schulz, E.G. (2017). X-chromosome dosage as a modulator of pluripotency, signalling
751 and differentiation? Philosophical Transactions of the Royal Society B: Biological
752 Sciences *372*.

753 Schwabl, H., Dowling, J., Baldassarre, D.T., Gahr, M., Lindsay, W.R., and Webster, M.S.
754 (2015). Variation in song system anatomy and androgen levels does not correspond to
755 song characteristics in a tropical songbird. Animal Behaviour *104*, 39-50.

756 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N.,
757 Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for
758 integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.

759 Shoemaker, H.H. (1939). Effect of Testosterone Propionate on Behavior of the Female
760 Canary. Proceedings of the Society for Experimental Biology and Medicine *41*, 299-302.

761 Slater, P.J.B., and Mann, N.I. (2004). Why do the females of many bird species sing in the
762 tropics? Journal of Avian Biology *35*, 289-294.

763 Smeds, L., Warmuth, V., Bolivar, P., Uebbing, S., Burri, R., Suh, A., Nater, A., Bures, S.,
764 Garamszegi, L.Z., Hogner, S.*, et al.* (2015). Evolutionary analysis of the female-specific
765 avian W chromosome. Nat Commun *6*, 7330.

766 Smith, C.A., Roeszler, K.N., Ohnesorg, T., Cummins, D.M., Farlie, P.G., Doran, T.J., and
767 Sinclair, A.H. (2009). The avian Z-linked gene DMRT1 is required for male sex
768 determination in the chicken. Nature *461*, 267-271.

769 Smith, G.T., Brenowitz, E.A., and Wingfield, J.C. (1997). Roles of photoperiod and
770 testosterone in seasonal plasticity of the avian song control system. J Neurobiol *32*, 426-
771 442.

772 Stacklies, W., Redestig, H., Scholz, M., Walther, D., and Selbig, J. (2007). pcaMethods—a
773 bioconductor package providing PCA methods for incomplete data. Bioinformatics *23*,
774 1164-1167.

775 Takayama, K., Kaneshiro, K., Tsutsumi, S., Horie-Inoue, K., Ikeda, K., Urano, T., Ijichi, N.,
776 Ouchi, Y., Shirahige, K., Aburatani, H.*, et al.* (2007). Identification of novel androgen
777 response genes in prostate cancer cells by coupling chromatin immunoprecipitation and
778 genomic microarray analysis. Oncogene *26*, 4453.

779 Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays
780 applied to the ionizing radiation response. Proc Natl Acad Sci U S A *98*, 5116-5121.

781 Uebbing, S., Konzer, A., Xu, L., Backström, N., Brunström, B., Bergquist, J., and Ellegren, H.
782 (2015). Quantitative Mass Spectrometry Reveals Partial Translational Regulation for
783 Dosage Compensation in Chicken. Molecular Biology and Evolution.

784 Vallet, E., Kreutzer, M., and Gahr, M. (1996). Testosterone Induces Sexual Release Quality
785 in the Song of Female Canaries. Ethology *102*, 617-628.

786 Van Meir, V., Verhoye, M., Absil, P., Eens, M., Balthazart, J., and Van der Linden, A. (2004).
787 Differential effects of testosterone on neuronal populations and their connections in a
788 sensorimotor brain nucleus controlling song production in songbirds: a manganese
789 enhanced-magnetic resonance imaging study. Neuroimage *21*, 914-923.

790 Vellema, M., Diales Rocha, M., Bascones, S., Zsebők, S., Dreier, J., Leitner, S., Van der
791 Linden, A., Brewer, J., and Gahr, M. (2019a). Accelerated redevelopment of vocal skills is
792 preceded by lasting reorganization of the song motor circuitry. eLife *8*, e43194.

793 Vellema, M., Rocha, M., Bascones, S., Zsebők, S., Dreier, J., Leitner, S., Van der Linden, A.,
794 Brewer, J., and Gahr, M. (2019b). Accelerated redevelopment of vocal skills is preceded
795 by lasting reorganization of the song motor circuitry. eLife *8*, e43194.

796 Voigt, C., and Leitner, S. (2008). Seasonality in song behaviour revisited: Seasonal and
797 annual variants and invariants in the song of the domesticated canary (Serinus canaria).
798 Hormones and Behavior *54*, 373-378.

799 Wade, J. (2001). Zebra finch sexual differentiation: The aromatization hypothesis
800 revisited. Microscopy Research and Technique *54*, 354-363.

801 Wickler, W., and Seibt, U. (1980). Vocal Dueting and the Pair Bond II. Unisono dueting of
802 the African forest weaver, Symplectes bicolor. Zeitschrift für Tierpsychologie *52*, 217-
803 226.

804 Wild, J.M. (2004). Functional Neuroanatomy of the Sensorimotor Control of Singing.
805 Annals of the New York Academy of Sciences *1016*, 438-462.

806 Wilson, S., Qi, J., and Filipp, F.V. (2016). Refinement of the androgen response element
807 based on ChIP-Seq in androgen-insensitive and androgen-responsive prostate cancer
808 cell lines. Scientific Reports *6*, 32611.

809 Wingfield, J.C., Smith, J.P., and Farner, D.S. (1982). Endocrine Responses of White-
810 Crowned Sparrows to Environmental Stress. The Condor *84*, 399-409.

811 Wolf, J.B., and Bryk, J. (2011). General lack of global dosage compensation in ZZ/ZW
812 systems? Broadening the perspective with RNA-seq. BMC Genomics *12*, 91.

813 Yang, X., Schadt, E.E., Wang, S., Wang, H., Arnold, A.P., Ingram-Drake, L., Drake, T.A., and
814 Lusis, A.J. (2006). Tissue-specific expression and regulation of sexually dimorphic genes
815 in mice. Genome Res *16*, 995-1004.

816 Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C.,
817 Clapham, P., Fitzgerald, S., Gil, L., *et al.* (2016). Ensembl 2016. Nucleic Acids Research *44*,
818 D710-D716.

819 Zhang, R., Yin, Y., Sun, L., Yan, P., Zhou, Y., Wu, R., and Wu, X. (2017). Molecular cloning of
820 ESR2 and gene expression analysis of ESR1 and ESR2 in the pituitary gland of the
821 Chinese alligator (Alligator sinensis) during female reproductive cycle. Gene *623*, 15-23.

822 Zhao, C., Dahlman-Wright, K., and Gustafsson, J.-Å. (2010). Estrogen Signaling via
823 Estrogen Receptor β. Journal of Biological Chemistry *285*, 39575-39579.

824 # Supplementary Tables

825 **Supplementary Table 1. Experimental groups and sample sizes.**

Species	Group	Sex	Singing	Tissue	Sample size
Forest weaver	Breeding	Female	Yes	HVC	6
		Male	Yes	HVC	6
Blue-capped cordon-bleu	Breeding	Female	Yes	HVC	6
		Male	Yes	HVC	6
Canary	Breeding	Female	No	HVC	6
		Male	Yes	HVC	6
	Nonbreeding	Female	Yes	HVC	6
		Female	No	HVC	5
		Male	No	HVC	6
		Female	Yes	HVC	6

Nonbreeding + testosterone	Male	Yes	HVC	6

Supplementary Table 2. Pearson's correlation analysis of principal components and variables (plasma androgen levels, HVC volume, sex, and singing).

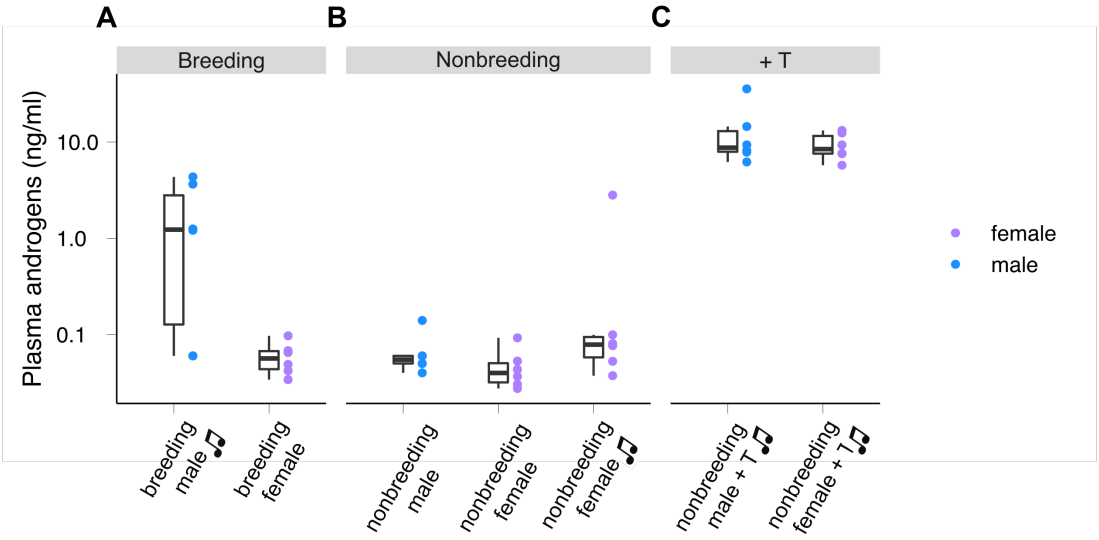
Supplementary Table 3. Fisher's exact tests of sex-biased genes for chromosome enrichment.

Supplementary Table 4. GO term enrichment analysis of sex-biased genes identified from the forest weaver (comparison 1), cordon-bleu (comparison 2), and canary (comparison 5) comparisons.

Abbreviations: FWm, forest weaver male-biased genes; CBm, cordon-bleu male-biased genes; Cm, canary male-biased genes; FWf, forest weaver female-biased genes; CBm, cordon-bleu female-biased genes; Cm, canary female-biased genes.

Supplementary Table 5. GO term enrichment analysis of female-specific, male-specific and sex-shared genes.

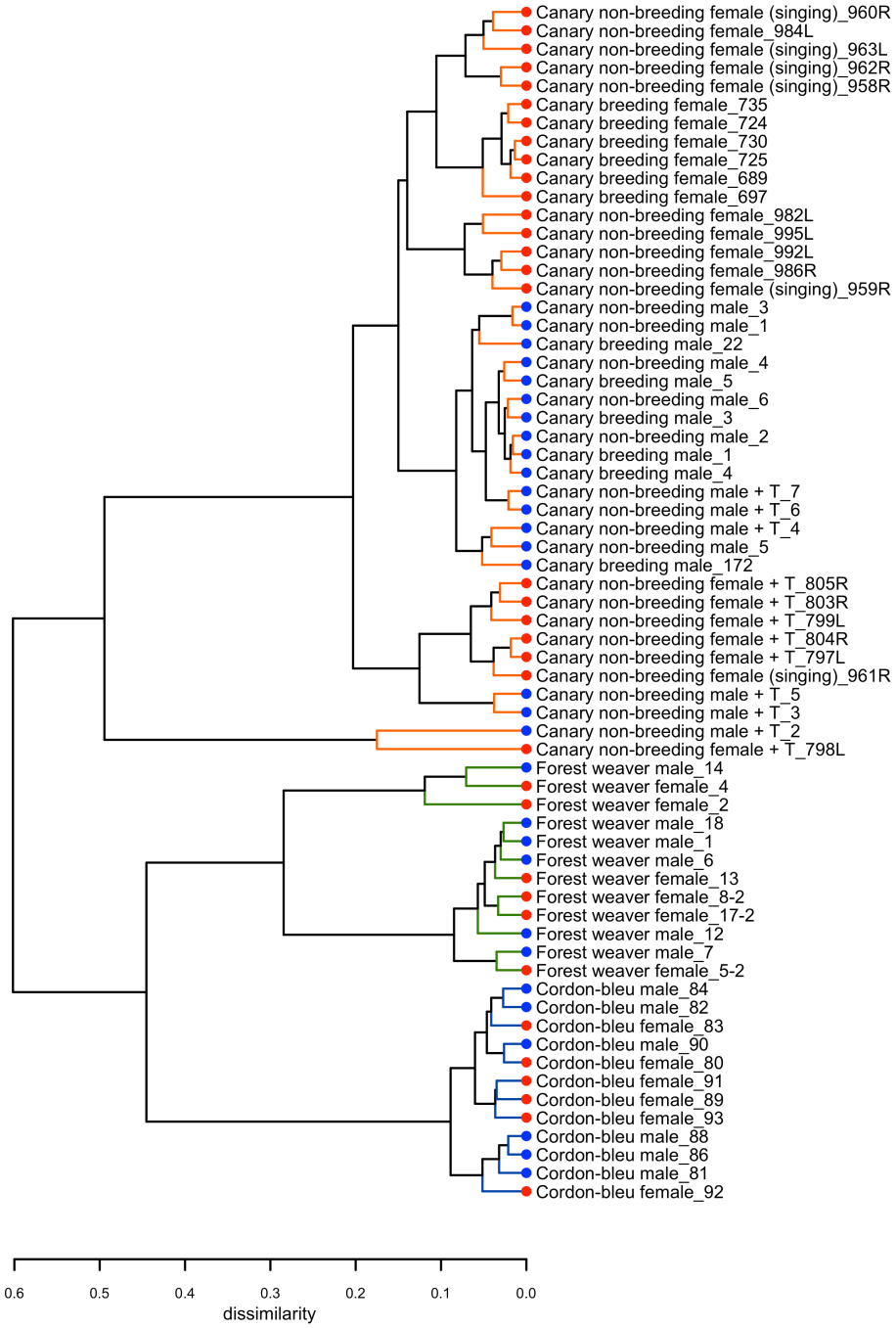
Supplementary Figures



Supplementary Figure 1. Plasma androgen levels of seven groups of canaries on the day of sacrifice.

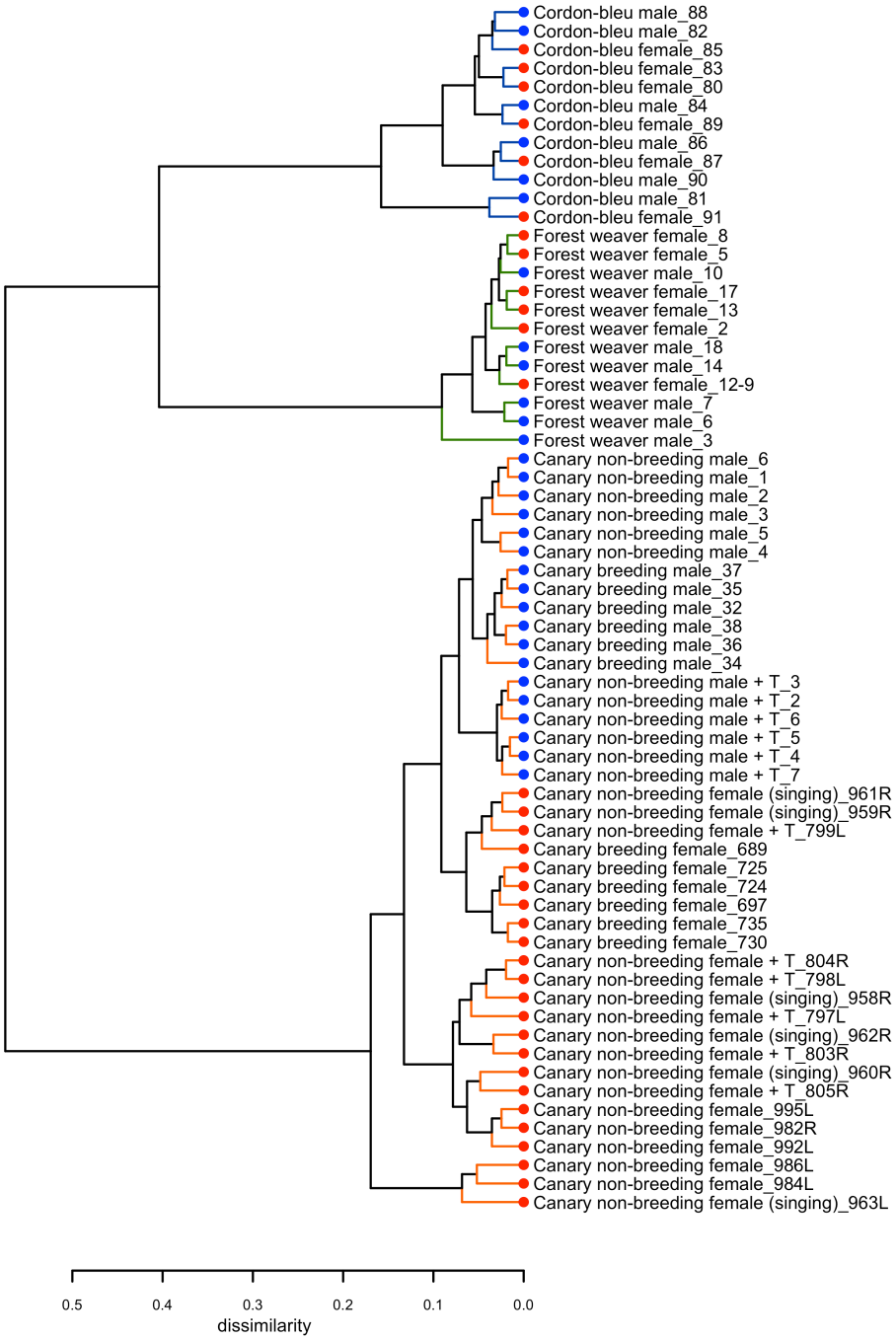
A, Canaries in breeding conditions. Singing males 1.76 ng/ml (mean), nonsinging females 59.2 pg/ml. Mann-Whitney Test, U = 6, P value = 0.07. **B,** Canaries in nonbreeding conditions. Nonsinging males 66.7 pg/ml (mean), nonsinging females 47.2 pg/ml, singing females 52.6 pg/ml. Mann-Whitney Test (nonsinging males vs. nonsinging females), U = 9, P value = 0.172. Mann-Whitney Test (nonsinging males vs. singing females), U = 24, P value = 0.377. **C,** Testosterone-implanted nonbreeding canaries. Males 13.5 ng/ml, females 9.31 ng/ml. Mann-Whitney Test, U = 13, P value = 0.485. Testosterone implantation significantly increased the plasma androgen levels of both nonbreeding males and females

Mann-Whitney Test (nonbreeding nonsinging males vs. testosterone-implanted singing males), U = 0, P value = 0.00492. Mann-Whitney Test (nonbreeding nonsinging females vs. testosterone-implanted singing females), U = 0, P value = 0.00217. The plasma androgen levels were higher in the breeding males than in the nonbreeding nonsinging males (Mann-Whitney Test, U = 32, P value = 0.0275). The boxes indicate the 25th/50th/75th percentiles (bottom/middle/top bar), and the extent of the whiskers indicates the most extreme values that are within 1.5 times the IQR (interquartile range) of the hinge. Each colour-coded dot indicates the measurement from one bird. The nonbreeding singing female canaries data were obtained from Ko et al. (2020).



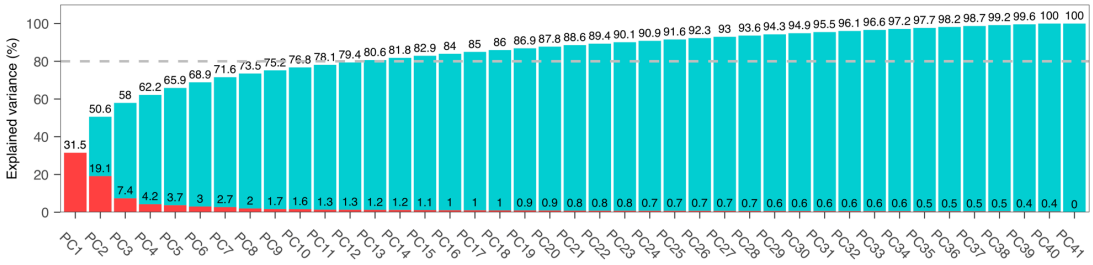
Supplementary Figure 2. Hierarchical clustering of the HVC transcriptomes of 65 birds used in this study.

Hierarchical clustering showed that the HVC transcriptomes were first clustered based on the phylogenetic relationship; among canaries, nonbreeding females implanted with testosterone showed the most distinctive patterns. +T: testosterone implantation.

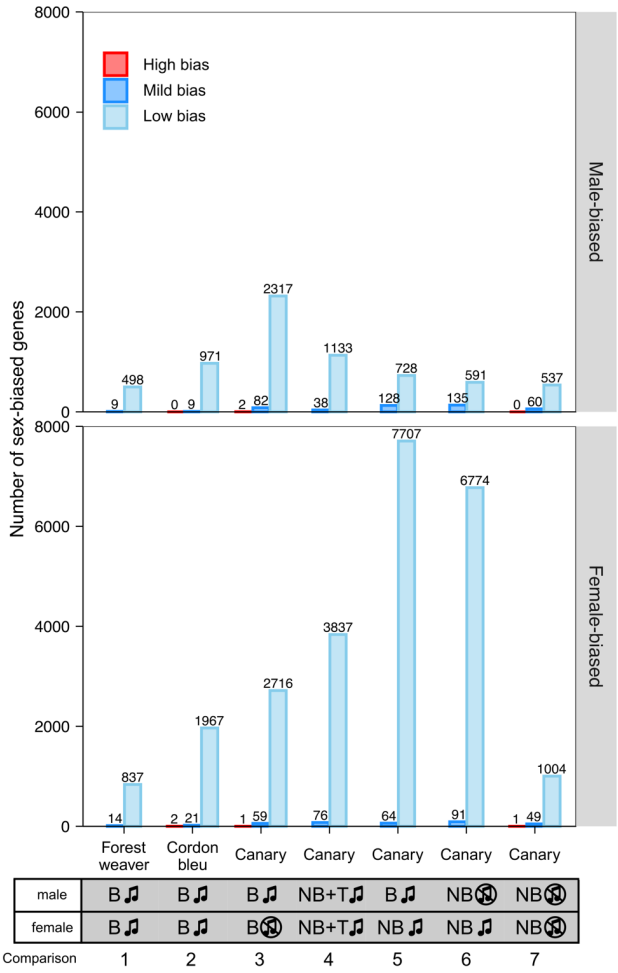


Supplementary Figure 3. Hierarchical clustering of the entopallium transcriptomes of 65 birds used in this study.

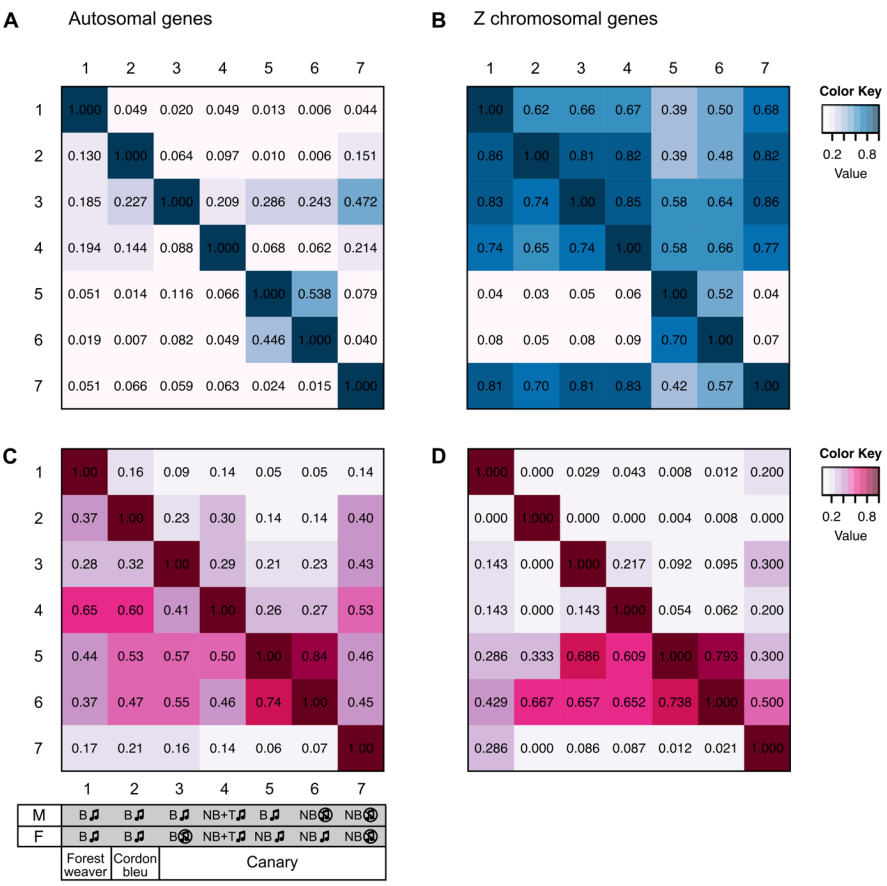
Hierarchical clustering showed that the entopallium transcriptomes were first clustered based on the phylogenetic relationship. The canaries were clustered by sex. +T: testosterone implantation.



Supplementary Figure 4. PCA scree plot showing the percentage of variance explained by each principal component.



Supplementary Figure 5. The majority of sex-biased genes showed low sex bias. The bar graph summarizes the number of sex-biased genes in the HVC transcriptome identified from each male-to-female comparison. High bias: $|\log_2(\text{fold change})| \geq 2$; moderate bias: $1 \leq |\log_2(\text{fold change})| < 2$; low bias: $0.5 \leq |\log_2(\text{fold change})| < 1$. The phenotypes of the groups being compared in each male-to-female comparison are listed at the bottom of the graph.



Supplementary Figure 6. Pairwise comparisons between autosomal and Z chromosomal sex-biased genes.

A, Autosomal male-biased genes. **B,** Z chromosomal male-biased genes. **C,** Autosomal female-biased genes. **D,** Z chromosomal female-biased genes. The numbers indicated in the matrices are the proportions of identical genes identified from each pair. The phenotypes of the groups used for each comparison are listed at the bottom. B: breeding; NB: nonbreeding; T: testosterone implantation; the music note indicates the presence or absence of singing behaviour.