



# Occasional paternal inheritance of the germline-restricted chromosome in songbirds

Yifan Pei<sup>a,1</sup>, Wolfgang Forstmeier<sup>a,1</sup>, Francisco J. Ruiz-Ruano<sup>b,c,1</sup>, Jakob C. Mueller<sup>a</sup>, Josefa Cabrero<sup>d</sup>, Juan Pedro M. Camacho<sup>d</sup>, Juan D. Alché<sup>e</sup>, Andre Franke<sup>f</sup>, Marc Hoepfner<sup>f</sup>, Stefan Börno<sup>g</sup>, Ivana Gessara<sup>h</sup>, Moritz Hertel<sup>h</sup>, Kim Teltscher<sup>a</sup>, Ulrich Knief<sup>i</sup>, Alexander Suh<sup>b,c,1,2</sup>, and Bart Kempnaers<sup>a,2</sup>

<sup>a</sup>Department of Behavioral Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology 82319 Seewiesen, Germany; <sup>b</sup>School of Biological Sciences, University of East Anglia, Norwich NR4 7TU, United Kingdom; <sup>c</sup>Department of Organismal Biology, Evolutionary Biology Centre, Science for Life Laboratory, Uppsala University SE-752 36 Uppsala, Sweden; <sup>d</sup>Department of Genetics, University of Granada E-18071 Granada, Spain; <sup>e</sup>Department of Biochemistry, Cell and Molecular Biology of Plants, Estación Experimental del Zaidín, Spanish National Research Council E-18008 Granada, Spain; <sup>f</sup>Institute of Clinical Molecular Biology, Christian-Albrechts-Universität zu Kiel 24118 Kiel, Germany; <sup>g</sup>Sequencing Core Facility, Max Planck Institute for Molecular Genetics 14195 Berlin, Germany; <sup>h</sup>Department of Behavioral Neurobiology, Max Planck Institute for Ornithology 82319 Seewiesen, Germany; and <sup>i</sup>Division of Evolutionary Biology, Faculty of Biology, Ludwig Maximilian University of Munich D-82152 Planegg-Martinsried, Germany

Edited by Scott Edwards, Department of Organismic and Evolutionary Biology, Museum of Comparative Zoology, Harvard University, Cambridge, MA; received March 3, 2021; accepted November 6, 2021

**Songbirds have one special accessory chromosome, the so-called germline-restricted chromosome (GRC), which is only present in germline cells and absent from all somatic tissues. Earlier work on the zebra finch (*Taeniopygia guttata castanotis*) showed that the GRC is inherited only through the female line—like the mitochondria—and is eliminated from the sperm during spermatogenesis. Here, we show that the GRC has the potential to be paternally inherited. Confocal microscopy using GRC-specific fluorescent *in situ* hybridization probes indicated that a considerable fraction of sperm heads (1 to 19%) in zebra finch ejaculates still contained the GRC. In line with these cytogenetic data, sequencing of ejaculates revealed that individual males from two families differed strongly and consistently in the number of GRCs in their ejaculates. Examining a captive-bred male hybrid of the two zebra finch subspecies (*T. g. guttata* and *T. g. castanotis*) revealed that the mitochondria originated from a *castanotis* mother, whereas the GRC came from a *guttata* father. Moreover, analyzing GRC haplotypes across nine *castanotis* matriline, estimated to have diverged for up to 250,000 y, showed surprisingly little variability among GRCs. This suggests that a single GRC haplotype has spread relatively recently across all examined matriline. A few diagnostic GRC mutations that arose since this inferred spreading suggest that the GRC has continued to jump across matriline boundaries. Our findings raise the possibility that certain GRC haplotypes could selfishly spread through the population via occasional paternal transmission, thereby out-competing other GRC haplotypes that were limited to strict maternal inheritance, even if this was partly detrimental to organismal fitness.**

germline-restricted chromosome | paternal spillover | elimination efficiency | selfish DNA | zebra finch

In sexually reproducing eukaryotes, the stable inheritance of the nuclear DNA typically requires the recombination and segregation of pairs of homologous chromosomes that come from both parents. The songbird germline-restricted chromosome (GRC) is an intriguing exception (1–3). As the name indicates, the GRC is only found in cells of the germline in all songbirds examined to date but is absent from any somatic tissue (1, 4–6), presumably due to its elimination from somatic cells during early embryogenesis. While the functional significance of the GRC is still largely unknown (4), this special chromosome appears to be far more than just an accumulation of highly repetitive DNA as it may initially have appeared (7). To the contrary, the zebra finch *Taeniopygia guttata castanotis* GRC is rich in genes that are expressed in testes or ovaries (4), and the gene content of the GRC appears to be evolving rapidly (4, 8), leading to a remarkable variation in GRC size between species (5, 9). This rapid evolution is puzzling

because the GRC's adaptive value for passerines is not at all obvious [compared to all other birds that lack a GRC (5)]. Rapid evolution often takes place when a genetic element successfully manipulates its mode of inheritance to its own advantage [so called “selfish genetic elements” (10–12)]. Hence, as a key step toward understanding the evolution and the function of the GRC, it is essential to fully understand how the GRC is inherited.

In most studies to date, the GRC was observed in two copies in female (primary) oocytes (5, 7, 9, 13) and as a single chromosome in male spermatogonia (1, 4–6, 9, 13). Cytogenetic investigations in males (predominantly *T. g. castanotis*) found that this single-copy GRC is eliminated from nuclei during meiosis I and expelled from spermatids in late spermatogenesis (1, 6, 7, 9, 13–15). Based on observations from multiple species (6, 9, 13), it has been concluded that the avian GRC is exclusively maternally inherited. In this study, we used the two zebra finch subspecies *T. g. castanotis* of Australia (hereafter *castanotis*) and *T. g. guttata* of the

## Significance

Most if not all songbirds possess a germline-restricted chromosome (GRC) which is believed to be exclusively maternally inherited. However, we show that, in the zebra finch, the GRC can also be paternally inherited and that the potential for paternal inheritance may differ between families. We further show that the genetic diversity of the GRC is extremely reduced compared to the high diversity of mitochondrial DNA lineages, suggesting that a single GRC haplotype has spread through the Australian zebra finch population relatively recently via occasional paternal inheritance. Our study therefore suggests that the GRC has the potential to evolve in a selfish manner, which could result in intragenomic conflict.

Author contributions: Y.P., W.F., F.J.R.-R., A.S., and B.K. designed research; Y.P. performed research; Y.P., F.J.R.-R., J.C.M., J.C., J.P.M.C., J.D.A., A.F., M. Hoepfner, S.B., I.G., M. Hertel, K.T., U.K., and A.S. contributed new reagents/analytic tools; Y.P. analyzed data; Y.P., W.F., F.J.R.-R., J.C.M., J.C., J.P.M.C., M. Hertel, U.K., A.S., and B.K. discussed the results; Y.P., W.F., A.S., and B.K. wrote the paper; W.F., A.S., and B.K. provided supervision; and A.S. and B.K. provided funding acquisition.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

<sup>1</sup>To whom correspondence may be addressed. Email: ypei@orn.mpg.de, forstmeier@orn.mpg.de, francisco.ruiz-ruano@ebc.uu.se, or alexander.suh@ebc.uu.se.

<sup>2</sup>A.S. and B.K. contributed equally to this work.

This article contains supporting information online at <http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2103960119/-DCSupplemental>.

Published January 20, 2022.

Indonesian islands such as Timor (hereafter *guttata*) to test whether inheritance of the GRC is strictly maternal. Specifically, we integrated cytogenetic data with various genomic datasets collected in recent studies (*SI Appendix, Behind-the-paper*) to study a) the elimination efficiency of the GRC during spermatogenesis, b) the strictness of the proposed matrilineal inheritance (i.e., expected coinheritance with the mitochondrial genome, “mtDNA”), and c) the genetic variation of the GRC and the coevolutionary history between GRC and the associated mtDNA haplotypes within the *castanotis* subspecies.

## Results

**GRC-Specific Sequences in Ejaculates: Repeatability and Differences between Families.** In principle, GRC-specific sequences might be found in ejaculates in three different forms: 1) expelled free-floating GRC micronuclei (1, 15); 2) small, digested DNA fragments (14, 15); and 3) nonexpelled GRCs or parts thereof in sperm heads. According to current knowledge (5, 6, 9, 13), primary spermatocytes contain a single copy of the GRC that is expelled as a micronucleus during early meiosis. As each primary spermatocyte results in four haploid spermatozoa and all chromosomes have two chromatids, we expect that ejaculates contain up to 25 free-floating GRC micronuclei per 100 spermatozoa in case of 100% expulsion. We examined seven natural ejaculate samples from five different *castanotis* males using a probe for the GRC-linked high-copy number gene *dph6* (4) for fluorescent in situ hybridization (FISH) to label the GRC in the ejaculate. Contrary to the expectation, we frequently found the FISH signal for the GRC inside sperm heads (mean = 9% of the sperm heads, range across five samples: 1 to 19%) and only a few free-floating micronuclei (mean = 1 micronucleus per hundred sperm heads, range: 0 to 2; Fig. 1 and *Dataset S1*). The GRC-containing spermatozoa (*dph6* positive) showed no visible morphological differences to the GRC-negative ones (Fig. 1). A subsequent confocal microscopy analysis of the *dph6*-positive sperm heads showed that the signal came from inside the sperm nucleus (Fig. 1C and *Video S1*). These results show that, in the zebra finch, the GRC is not completely eliminated during spermatogenesis and imply that a nonnegligible number of spermatozoa can potentially transmit the GRC.

Given the observed variability in the proportion of sperm heads that contained the GRC, we estimated individual repeatability and between-family variation in the GRC content of ejaculates. Based on comparisons of coverage of high-copy number GRC sequences between testes and soma and between ejaculates and soma (*Dataset S2*), we found that ejaculates from the same individual male zebra finch were remarkably repeatable in their GRC content (Fig. 2B and C;  $R_{\text{male}} = 0.98$ ,  $P < 0.001$ ; *Dataset S3*). The majority of the repeatable variation was due to a between-matriline difference ( $R_{\text{matriline}} = 0.96$ ,  $P = 0.012$ ; *Dataset S3*). Ejaculates from males of matriline B had significantly higher amounts of GRC (Fig. 2B) than those from males of matriline A (Fig. 2C;  $b_{\text{matriline B}} = 2.4$ ,  $SE = 0.28$ ,  $P < 0.001$ ; *SI Appendix, Figs. S1 C and E and S2 and Dataset S3*), qualitatively confirming the results from the cytogenetic analysis (1 versus 9% of sperm heads were GRC positive, respectively; Fig. 2D and E).

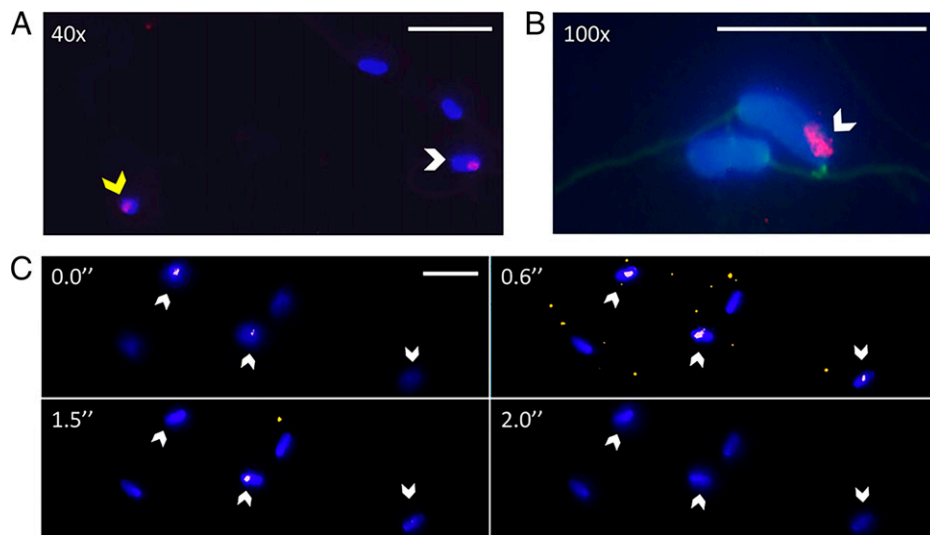
These results indicate that certain GRC haplotypes (e.g., those in matriline B) may be more likely to be transmitted via sperm than others (e.g., those in matriline A), and hence, more likely to potentially spread in a “selfish” manner. However, our current data cannot explain why the two families (matrilines) differ in their fractions of GRC-carrying sperm. The high individual repeatability and consistency within matrilines raises the question whether the observed difference in the elimination efficiency of the GRC during spermatogenesis might have a relatively simple genetic or epigenetic basis.

Future studies should test whether this is due to permanent environmental effects, to the GRC haplotype itself (including epigenetic marks such as histone or DNA modifications), to the A-chromosomal genotype (i.e., autosomal and sex-chromosomal), or to epistatic interactions between the GRC and the A-chromosomal genome.

**Paternally Inherited GRC in a Hybrid Individual.** Given the observed occurrence of the GRC in spermatozoa, we looked for cases of paternal inheritance of the typically maternally transmitted GRC (1, 7, 9, 13) among all sequenced testes samples. This included 12 *castanotis* males from nine matrilines (Fig. 2A–C) and one presumed *guttata* male as an outgroup for the *castanotis* GRCs. However, this putative *guttata* male turned out to be a descendant of a female *castanotis* x male *guttata* hybridization followed by backcrossing with *guttata* males. The sequencing data revealed that this male’s A-chromosomal genome consisted of 95% *guttata* and only 5% *castanotis* DNA (hence the *guttata* phenotype), whereas its mtDNA was from the *castanotis* matriline B (Fig. 3; see *SI Appendix, Results* for details).

If the GRC would have been inherited exclusively through the matriline (7, 9, 13), we expected that this hybrid male would show the same GRC haplotype as is typical for *castanotis* matriline B. However, we found a GRC haplotype that is uniquely and highly diverged from all the 12 GRC haplotypes examined so far (Fig. 3B and C and *SI Appendix, Fig. S3*). Given this high degree of divergence relative to the 12 *castanotis* GRCs, we hypothesized that this might be the hitherto unknown GRC of *guttata* (or at least a recombinant *guttata* x *castanotis* GRC). The high divergence was apparent in terms of a) a high number of private testis-specific single-nucleotide polymorphisms (SNPs), i.e., SNPs that were only detected in the testis of this individual in GRC-amplified regions that were shared among all 13 GRCs [hence GRC-linked variants;  $N_{\text{high-confidence SNP}} = 312$  in Fig. 3B; *SI Appendix, Fig. S3*] and b) four independent regions that appear to be GRC linked in high-copy number in this individual but are absent from all other 12 *castanotis* GRCs [ $N_{\text{high-confidence SNP}} = 169$  in Fig. 3C; *SI Appendix, Fig. S3*; including 13 genes with GRC paralogs that were not found previously (4), see *Datasets S4 and S5*]. These results suggest that the GRC was inherited from at least one of the *guttata* males during back crossing (Fig. 3A). We assumed strict maternal inheritance of mtDNA haplotypes, which is likely because a) paternal inheritance of the mtDNA is known to be rare (16) and has usually been detected in individuals with mitochondrial diseases and heteroplasmy (17, 18), and b) the avian W chromosome and mtDNA are in (near-)perfect linkage disequilibrium (16). Note that the *castanotis* x *guttata* testis sample had extremely low coverage on the GRC, presumably because of underdeveloped testes (*Materials and Methods*), and that no individuals from a previous generation could be retrieved (*SI Appendix, Behind-the-paper*). Lacking additional information from other individuals in this pedigree, we assume that the *guttata* GRC coexisted with, recombined with, or replaced the *castanotis* GRC and was stably coinherited with the *castanotis* mitochondria thereafter (Fig. 3).

**Discordant Evolutionary Histories between the mtDNA and the Associated GRC Haplotypes.** We focused on males with testis sequencing data ( $n = 12$  *castanotis* males; A-I, A0, A1, and B1 in Fig. 2) to study the *castanotis* GRC haplotypes. Males for which we only had data from ejaculates were used for confirmation (males B2 and B3) or were dropped from further analysis because of the extremely low GRC content in their ejaculates (males A2 and A3 in Fig. 2B). We found surprisingly little variation among the 12 *castanotis* GRCs from nine matrilines, some of which were old matrilines as judged from mtDNA divergence (Fig. 4 and *SI Appendix, Figs. S4 and S5*). In sharp contrast to



**Fig. 1.** Cyto-genetic evidence for the presence of the GRC in the nucleus of zebra finch *Taeniopygia guttata castanotis* sperm. The GRC-amplified probe *dph6* (see *Materials and Methods*) indicates the presence of the GRC (pink) inside some sperm heads (white arrow in A–C) and in free-floating micronuclei (yellow arrow in A). Blue DAPI stain without red indicates sperm heads without GRC. Green autofluorescence shows the sperm flagellum in B. (A) 40× magnification. (B) 100× magnification. (C) Individual z-sections under a confocal microscope show the sequential appearance and disappearance of the *dph6* signal along consecutive sections, indicating the location of the GRC within the nucleus of the spermatozoa. Time (in seconds) refers to [Video S1](#). The video consisted of 24 sections representing a total of 6.0 μm in depth. (Scale bars, 20 μm.)

the overall highly diverged *guttata* GRC, which harbored hundreds of private testis-specific SNPs (*Top* row in Fig. 3 B and C, *SI Appendix*, Fig. S34, and *Dataset S5*), such SNPs that were private to one of the 12 *castanotis* individuals were exceedingly rare (*Middle* and *Bottom* rows in Fig. 3 B and C and *SI Appendix*, Fig. S3 B–M). Note that in our full *castanotis* testis dataset of 12 GRCs, most of the males that shared the same mtDNA were brothers that can be expected to also share the same GRC haplotype (only one male, B, is not closely related; *Dataset S2*). Indeed, among the 14 males for which we obtained loci of nine single-copy GRC genes, the *castanotis* males that had the same mtDNA haplotype also had the same GRC haplotype (males A, A0, and A1 of matriline A and males B and B1 through B3 of matriline B; *SI Appendix*, Fig. S6).

As an additional test for discordance between the GRC and the mtDNA lineages, we compared the evolutionary histories of GRC haplotypes and their associated mitogenomes in nine *castanotis* matriline. For the GRC, we focused on nine single-copy GRC loci, *trim71<sub>GRC</sub>*, *bicc1<sub>GRC</sub>*, *pim3<sub>GRC</sub>*, *msh4<sub>GRC</sub>*, *surf4<sub>GRC</sub>*, *262<sub>intergenicGRC</sub>*, *aldh18a1<sub>GRC</sub>*, *rfc1<sub>GRC</sub>*, and *eppk1<sub>GRC</sub>* (4) (*SI Appendix*, Fig. S7), that were highly diverged from their A-chromosomal paralogs (for consensus sequence construction and additional details, see *Materials and Methods* and *SI Appendix*, *Materials and Methods* and *Results*).

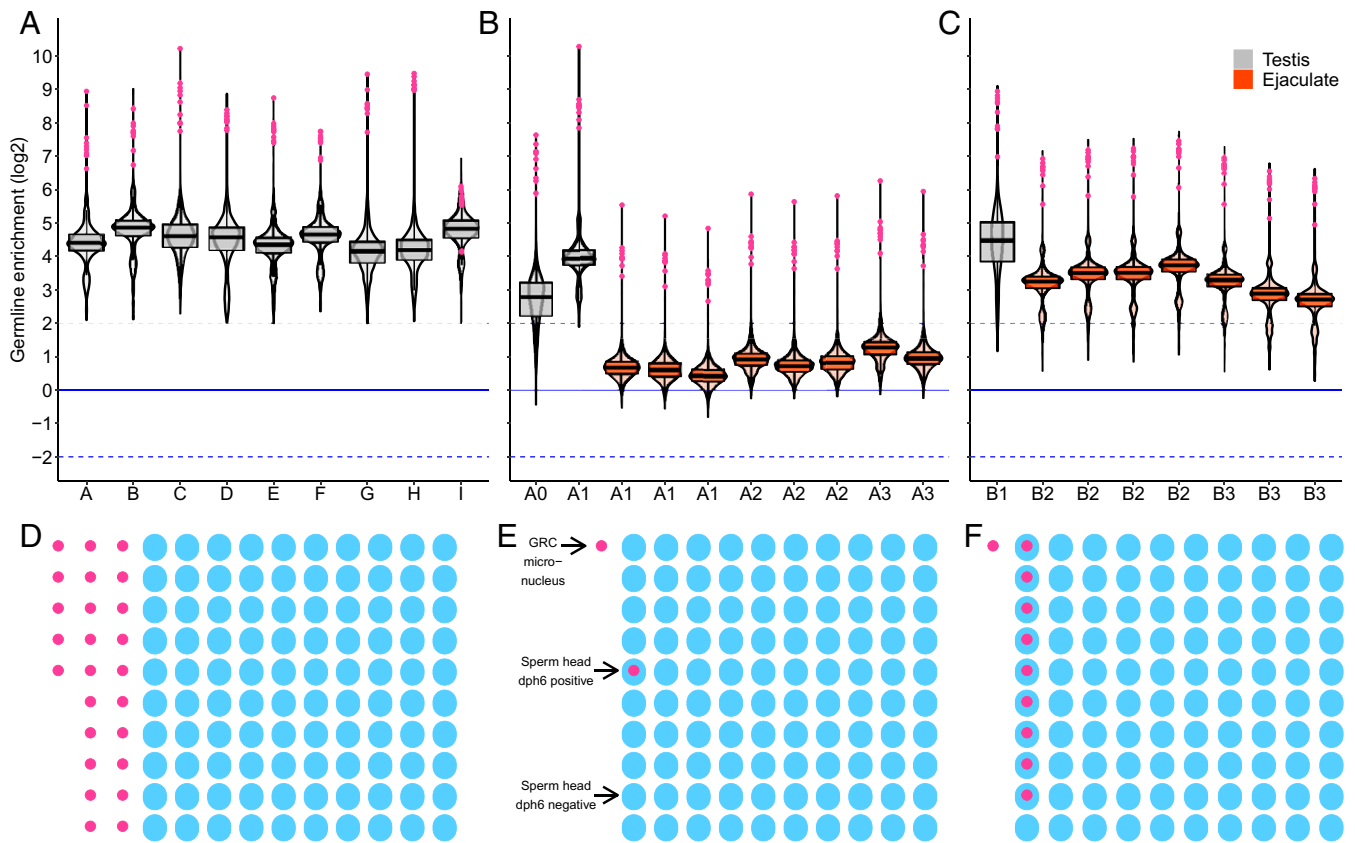
When comparing the phylogenetic tree of the mtDNA with that of the GRC haplotypes from the same *castanotis* males (based on the concatenated sequences of the nine GRC loci), we found no positive correlation between the two pairwise distance matrices (Fig. 4A;  $r = -0.04$ ; *Dataset S6*). Strikingly, multiple males from clearly diverged mtDNA lineages shared closely related GRC haplotypes, while males from closely related mtDNA lineages contained distantly related GRC haplotypes (Fig. 4 and *SI Appendix*, Figs. S6 and S8). For instance, we estimated that the mtDNA haplotypes E and H diverged about 250,000 y ago based on 0.5% substitutions per site and assuming a molecular clock with 2% substitutions per site per million years (19). However, the GRCs of males E and H are closely related to each other, as they both carry the same four derived GRC mutations in the contig containing *bicc1<sub>GRC</sub>* (Fig. 4B). This is difficult to explain unless GRCs are able to cross matriline

boundaries. In striking contrast to the considerable genetic diversity observed in mtDNA (13 SNPs/kilobase [kb]; Fig. 4C), we found little genetic diversity in the single-copy GRC-linked genes (0.4 SNPs/kb; Fig. 4C, *SI Appendix*, Fig. S8, and *Dataset S7*). This lack of diversity in the GRC even across deeply diverged matriline suggests that older GRC haplotypes were probably replaced by a more recent type that was able to spread across matriline boundaries. The GRC haplotype differences that have evolved since then still show no association with the structure of the mtDNA tree (Fig. 4A), suggesting that the crossing of matriline boundaries is still ongoing. Alternatively, the lack of GRC diversity could be due to an extremely low mutation rate or strong purifying selection on the GRC, although this cannot explain the shared GRC SNPs between distantly related matriline and the highly discordant tree topologies. Unfortunately, we were unable to obtain a *guttata* GRC haplotype from the hybrid *castanotis* × *guttata* male for these single-copy GRC loci because the sequenced testes were underdeveloped (*SI Appendix*, Fig. S7). Using a high-quality *guttata* testis sample would allow better relative estimates of mtDNA and GRC sequence evolution, which presumably would shed light on the observed low GRC diversity in *castanotis* compared to the high level of mtDNA diversity.

## Discussion

Our results demonstrate that GRCs, in addition to their regular maternal transmission, can be and have been transmitted from the father (Fig. 3) via sperm (Fig. 1). Our sequencing and microscopy results suggest that the likelihood of paternal inheritance of the GRC may be family-specific (Fig. 2) and may thus be heritable. This begs the question whether the between-family difference is due to variation in the GRC, the A-chromosomal genome, or to epigenetic effects. Although some of our findings are based on a single case (Fig. 3) or on 15 ejaculates from two families (Fig. 2), taken together, they clearly indicate that the GRC can be paternally inherited.

It remains unclear what happens when two GRCs (i.e., one maternally and one paternally inherited) enter the embryo. However, we did not observe any GRC-heterozygous birds (Fig. 4B and *SI Appendix*, Fig. S6), suggesting that one GRC is

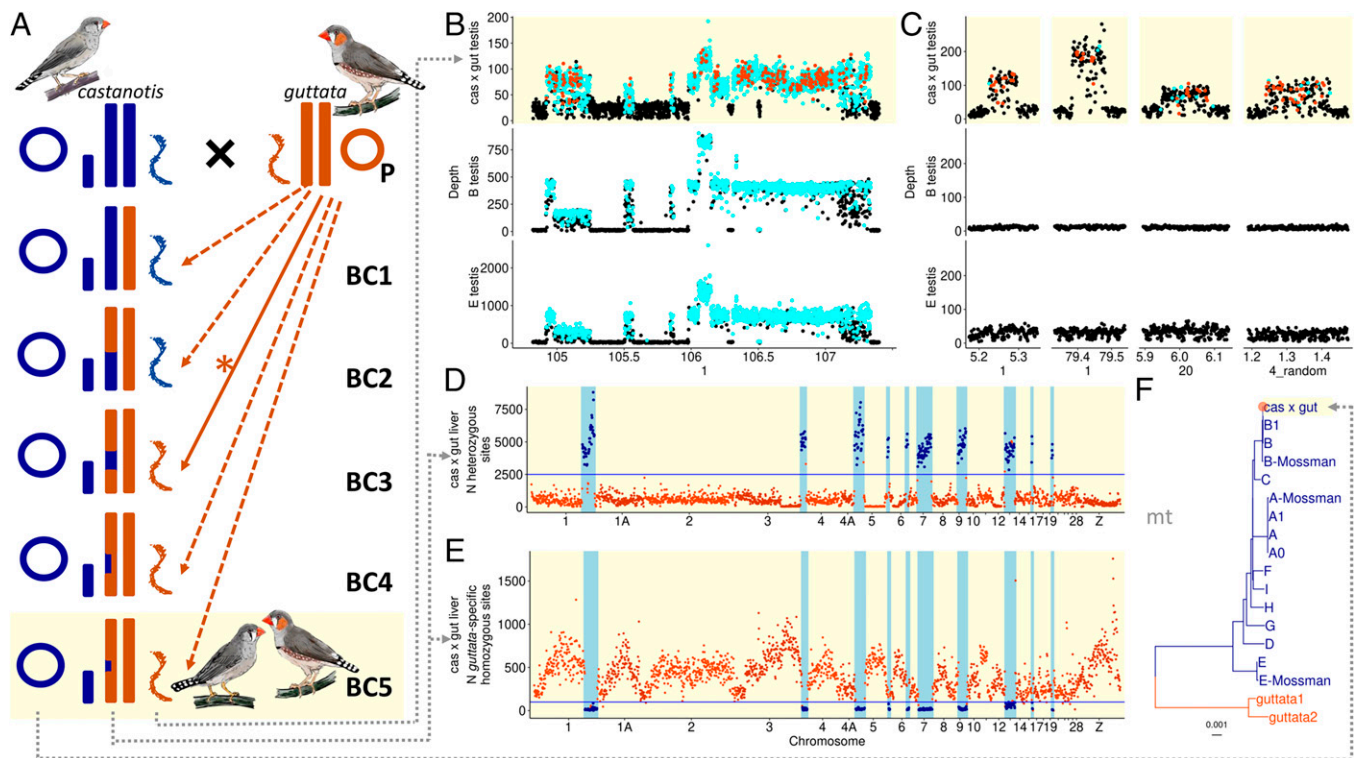


**Fig. 2.** The elimination efficiency of the GRC differs between *castanotis* matriline A and B. (A–C) Comparison of sequencing coverage of GRC-containing (testis, indicated in gray, and ejaculate, orange) and GRC-free tissue (liver) identifies sequences that are GRC linked in high copy number (4). Male identities are shown on the x-axis. The solid blue line refers to a log<sub>2</sub> germline-to-soma coverage ratio = 0 (i.e., no germline enrichment); the dashed blue line refers to a fourfold increase (Top) of coverage in germline compared to soma tissue. Pink dots highlight the 1-kb windows on *dph6*. (A) Violin box plots show coverage ratios of the selected windows with more than fourfold (log<sub>2</sub> > 2) enrichment in testes in comparison to soma in all nine *castanotis* males (A to I, of different matriline). The thick horizontal lines show the median, and boxes indicate the 25th and 75th percentiles. (B) Coverage ratios of the selected windows in matriline A (three brothers A1 to A3 and uncle A0) show that ejaculates contain lower amounts of GRC-derived reads compared to testes (comparison of the median of eight ejaculates with two testis samples:  $b = -2.56$ ,  $SE = 0.29$ ,  $P < 0.001$ ; Dataset S3) as expected from previous work (14, 15). (C) Coverage ratios of the selected windows in matriline B (three brothers B1 to B3) with a higher GRC content in ejaculates and hence a smaller difference with testis (comparison of the median of seven ejaculates with one testis sample:  $b = -1.21$ ,  $SE = 0.39$ ,  $P = 0.02$ ; Dataset S3). (D–F) Illustration of the expected (D) and observed (E and F) number of sperm heads that are free of GRC (blue ovals) and that contain GRC (blue ovals with a pink circle inside; i.e., with a positive *dph6* signal; Fig. 1) as well as the number of free-floating GRC-micronuclei (pink circles). (D) Error-free expulsion of GRC from spermatocytes (expected based on previous work) should result in up to 25 free-floating GRC micronuclei per 100 sperm heads in the ejaculate (1, 6, 13, 15). (E) Ejaculates from matriline A showed 1% of GRC-positive sperm heads ( $n = 677$  scored sperm). (F) Ejaculates from matriline B showed 9% of GRC-positive sperm heads ( $n = 1,533$  scored sperm; Dataset S1).

removed (either always the same one, i.e., an efficient driver, or with a 50:50 probability). In either case, a GRC that can escape elimination during spermatogenesis has an evolutionary advantage by having two possible routes of inheritance (as opposed to a single route for other GRC haplotypes that are only maternally transmitted) (10). Such unequal opportunities of inheritance favor the evolution of the more efficiently spreading DNA (a “selfish GRC”). The observation of low diversity on the *castanotis* GRCs even across highly diverged matriline (Fig. 4) is consistent with the idea that such a selfish GRC may have spread relatively recently. Strong purifying selection on the GRC (for other reasons than its ability to spread) could also explain the low diversity, but it cannot explain the observed discordance in tree topologies (Fig. 4A) and the distribution of derived GRC SNPs (Fig. 4B), which more plausibly suggests that the crossing of matriline boundaries may be still ongoing. The highly diverged *guttata* GRC (Fig. 3B) argues against the alternative explanation that the GRC has an unusually low mutation rate, which is also contradicted by the observation of the highly dynamic evolution of GRC gene content in general (4, 5). Hence, derived SNPs, tree topologies, and the shortage

of GRC variation in *castanotis* support the hypothesis of repeated lateral transfer via occasional paternal inheritance of a recently evolved GRC across all or nearly all *castanotis* matriline.

We thus found evidence for paternal GRC inheritance in both zebra finch subspecies (Figs. 1, 2, and 4 for *castanotis* GRC and Fig. 3 for *guttata* GRC). Interestingly, Malinovskaya et al. (9) reported a GRC copy number mosaicism in spermatogonia and pachytene spermatocytes in males of the pale martin (*Riparia diluta*). Because the GRC behaves similarly during meiosis in all songbirds examined to date (6, 9, 13) (i.e., elimination during male spermatogenesis and recombination in female oocytes), we hypothesize that the spermatogonia and spermatocytes of the pale martin that possess an extra copy of the GRC (9) might result in GRC-carrying sperm. The GRC-carrying sperm thereby give the GRC a “selfish advantage,” meaning that such GRCs might be able to spread even if they were mildly deleterious to the organism (e.g., to a certain sex or developmental stage). These examples suggest that the selfish evolution of the GRC might be more widespread, which deserves further study across songbirds.

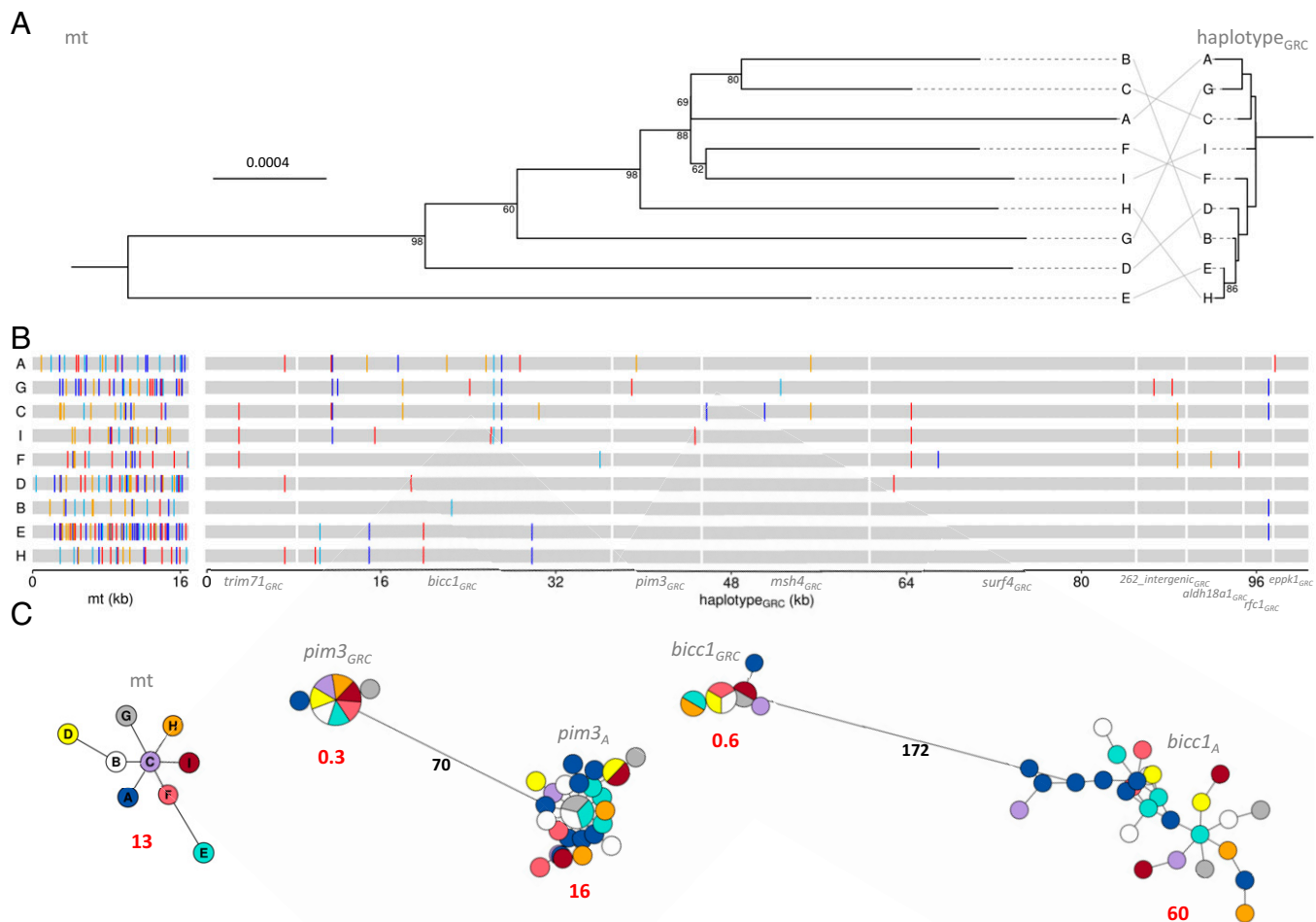


**Fig. 3.** A case of paternal inheritance of the GRC in a captive-bred hybrid *castanotis* × *guttata* population. (A) Reconstructed hypothetical breeding history during domestication of the recently wild-derived *guttata* subspecies. Presumably because females of wild-derived *guttata* birds do not easily reproduce in captivity, we hypothesized that *guttata* males were crossed with an already domesticated *castanotis* female (from Europe), and the resulting female hybrids were back crossed with pure *guttata* males for about five generations (P-BC5) until the population was phenotypically *guttata*-like. This reconstruction is based on the genotyping of one male (49) of the resulting hybrid population (yellow background, *Bottom*), which is characterized by a *castanotis* mother contributing the mtDNA (dark blue circle), the female-specific W chromosome (dark blue short rectangle; for details, see *SI Appendix, Results*), and 5% of the A-chromosomal DNA (blue fragments in the long rectangles) and *guttata* males contributing 95% of the A-chromosomal DNA (orange-brown long rectangles) and the GRC (orange-brown sigmoid symbol). Note that the paternal inheritance of the GRC must have happened sometime between generations P and BC5 (solid arrow with asterisk). (B and C) Raw read depth of selected genome regions of testis libraries [mapped to somatic reference *taeGut1* (36)] for the *castanotis* × *guttata* hybrid (*Top* row of each panel, yellow background) and two representative *castanotis* individuals (*Bottom* two rows). GRC-specific windows are distinguishable from somatic reference-like windows by high read depth values and by testis-specific SNPs (cyan). Each dot represents a 1-kb window. Each cyan dot indicates a 1-kb window that contains ≥3 germline-specific SNPs. Each orange-brown dot shows a 1-kb window that contains private testis-specific SNPs. To visually emphasize highly diverged sequences, windows with private GRC SNPs are only shown (in orange brown) if there are at least three such SNPs within a 10-kb window. Note that only the hybrid has that many private testis-specific SNPs (B and C) as well as private GRC-linked sequences that are absent in *castanotis* individuals (C), suggesting that it carries a *guttata* GRC. See *SI Appendix, Fig. S3* for genome-wide plots and more individuals. (D and E) Number of heterozygous sites (D) and number of fixed *guttata*-specific homozygous sites (E) in autosomal and Z-linked windows from somatic tissue of the *castanotis* × *guttata* hybrid compared to the pooled DNA from 100 wild-caught *castanotis* zebra finches (43). The horizontal blue line indicates the cutoff of 2,500 heterozygous sites (D) or 100 fixed *guttata*-specific homozygous sites per 500-kb window [i.e., each dot; (E)]. Dark blue dots show windows that contain an excess number of heterozygous sites (D), or a reduced number of *guttata*-specific homozygous sites (E), and indicate 10 introgressed *castanotis* segments. (F) Phylogenetic tree of the mitogenomes, showing that the hybrid's mtDNA (“cas x gut,” orange-brown dot at the top of the tree) clusters with typical captive European *castanotis* zebra finches (dark blue) rather than with mitogenome assemblies of two published *guttata* datasets [SRA accession numbers SRR2299402 (47) and SRR3208120 (48), respectively; orange brown]. The scale bar shows the number of substitutions per site.

We observed remarkable variation in the efficiency of GRC elimination from sperm, and we expect that this will be mirrored in the ability of the GRC to spread paternally (Fig. 2). Such variation could only be evolutionarily stable if the obtained advantages via selfish spreading would be compensated by other disadvantages, for example, if paternal inheritance would reduce fertility or embryo survival (antagonistic pleiotropy). A paternally spreading GRC haplotype may also have been fixed in the population as indicated by the low genetic diversity (Fig. 4) and the potential ability to spread via sperm (Figs. 1 and 2C). After a certain GRC haplotype has successfully spread to the entire population (going to fixation), its ability to spread through this second route (i.e., paternal transmission) may lose its adaptive value because there is no alternative haplotype with which to compete. However, a second variant that lacks this paternal spreading ability could then only have invaded if it conveyed another advantage (e.g., to organismal fitness).

Variation in the efficiency of elimination of the GRC during spermatogenesis (Fig. 2) might also explain why previous cytogenetic work on the testes of songbirds failed to detect GRC-positive spermatozoa (1, 4–6, 9). By chance, the examined individuals might resemble those in matriline A (Fig. 2B) in which the GRC elimination efficiency during spermatogenesis is high.

Recent studies suggest highly dynamic evolution of the gene content of the GRC. Across species, there is remarkable variation in the size and content of the GRC (5, 9), especially when compared to the highly syntenic A chromosomes in birds (20–22). In the Australian zebra finch (*T. g. castanotis*), much of the content of its GRC appears to have been derived from A-chromosomal paralogs only recently (4). The zebra finch (*T. g. castanotis*) GRC is enriched for genes showing gonad-specific expression (4), and some genes show signals of strong positive or purifying selection (4, 8), suggesting an essential role for the GRC in sexual reproduction. The genetic diversity of two of the



**Fig. 4.** Tanglegram (A), haplotype sequences (B), and networks (C) showing the different phylogenies of the mtDNA haplotypes and the GRC-linked genes in the same *castanotis* zebra finch individuals. (A) Phylogenetic trees were built from gap-free alignments of all nine haplotypes (A through I) of the *castanotis* mitogenome (Left) and the nine associated GRC haplotypes (concatenated sequence of the nine single-copy GRC loci; Right). The mtDNA tree was rooted by the mitogenome assemblies of two published *guttata* datasets [SRA accession numbers SRR2299402 (47) and SRR3208120 (48) (not shown here, see Fig. 3F)], whereas the tree of GRC haplotypes was rooted at the midpoint of the two most distantly related haplotypes. Node support bootstrap values (in %) are shown if >60 (based on 1,000 bootstraps). The scale bar indicates 0.0004 substitutions per site. (B) The alignments of the mitogenomes and the nine single-copy loci on the GRC (4) from the nine individuals used in A. Gray indicates consensus among the nine haplotypes. Orange, red, blue, and dark blue indicate a mutation toward A, T, C, and G comparing to the consensus, respectively. (C) Haplotype networks built from gap-free alignments (used in A) of the *castanotis* mitogenomes (Left), *pim3<sub>GRC</sub>* (Middle), and *bicc1<sub>GRC</sub>* (Right) with their A-chromosomal paralogs (i.e., *pim3<sub>A</sub>* and *bicc1<sub>A</sub>*). A-chromosomal paralog haplotypes were constructed from all *castanotis* somatic libraries (Dataset S2). Colors represent the different mitogenome haplotypes. The size of each circle indicates the number of samples of each haplotype, and the length of the black lines corresponds to the number of mutational steps between haplotypes. Red numbers refer to the number of SNPs per kb for each cluster of haplotypes; black numbers refer to the number of mutations per kb between the GRC-linked and A-chromosomal paralogs. Also see SI Appendix, Fig. S8 for the haplotype networks of all germline samples and for the double-copy GRC paralog *elavl4<sub>GRC</sub>*. Note the highly reduced genetic diversity (short branch length in A and little variation in B and C) in the GRC genes in comparison to the associated mitogenome ( $P < 0.0001$ ). Further note that different mitogenome haplotypes from different clades may share the same GRC haplotype, indicating different evolutionary histories between mtDNA and GRC (Dataset S6).

examined single-copy genes on the GRC is 56-fold lower than the diversity of their A-chromosomal paralogs (Fig. 4C, Dataset S7, SI Appendix, Results and Figs. S8 and S9, and ref. 23). Other non-recombining sex-specific chromosomes also show highly reduced genetic diversity compared to their autosomes, possibly because of strong sexual selection [e.g., the human Y chromosome shows a 23- to 41-fold reduction in the mean number of pairwise differences per site, i.e.,  $\pi$  (24)] or because of a Hill–Robertson interference with the mitogenome [e.g., the avian W chromosome shows a 46- to 104-fold reduction of  $\pi$  in four *Ficedula* flycatcher species (16) and a 90-fold reduction in chicken *Gallus gallus* (25)]. To increase our understanding of the genetic variation and evolutionary history of the zebra finch GRC, we suggest that future efforts should focus on completing a GRC reference assembly and on studying a true *guttata* GRC.

Another implication of the GRC's strong linkage disequilibrium with the two other nonrecombining elements in songbirds (i.e., the W chromosome and the mitogenome) is that it reduces the efficiency of positive or negative selection on them, especially in small populations (i.e., Hill–Robertson interference) (26, 27). However, this is only true if these three elements are always coinherited. With paternal spillover demonstrated here, the GRC is immediately decoupled from the other two and can create a new optimal combination of the three elements—the GRC, the mitogenome, and the W chromosome. Thus, the gene richness and highly dynamic nature of the GRC (4, 5), in combination with our observation that it can be paternally inherited, offers novel opportunities to study genetic compatibility and the evolution of the W chromosome and the mitogenome (16) of songbirds.

## Materials and Methods

**Samples.** All germline samples (i.e., ejaculate or testis) and their corresponding soma samples (liver of the same individual or blood of the parents of that individual) are described in [Dataset S2](#) (*SI Appendix, Behind-the-paper*). In brief, ejaculate samples were collected from individuals of a domesticated zebra finch *T. g. castanotis* population (“Seewiesen”) kept at the Max Planck Institute for Ornithology since 2004 (no. 18 in ref. 28). Using a dummy female, we collected natural ejaculates from eight brothers from two families (mtDNA haplotype A,  $N_{\text{ejaculates}} = 9$ , from males A1 to A3; mtDNA haplotype B,  $N_{\text{ejaculates}} = 11$ , from males B2 to B6). Males were 100 to 1,036 d old. The study was carried out under license (permit no. 311.4-si and 311.5-gr, Landratsamt Starnberg, Germany). For details on the method of collection, see *SI Appendix, Materials and Methods*.

We collected testis samples and their corresponding soma from one captive *guttata* zebra finch (“Timor zebra finch”), which turned out to be a *castanotis* × *guttata* hybrid (*SI Appendix, Results*), one wild-caught (male I, from Western Australia), and 11 captive *castanotis* zebra finches (“Australian zebra finch”). The *castanotis* samples were chosen to a) include nine major mtDNA haplotypes (i.e., males A to I from matriline A to I) and b) provide positive controls of GRC content for ejaculate samples (i.e., males A0, A1, and B1 from matriline A and B; male A0 was an uncle of males A1 through A3, and males B1 through B6 were brothers; *SI Appendix, Figs. S4 and S5*; see *SI Appendix, Materials and Methods* for mtDNA sequencing and assembly). Note that the ejaculates of males B4 through B6 were only sampled for cytogenetic analysis ([Dataset S1](#) and *SI Appendix, Behind-the-paper*). Matrilines A, B, and E were first described in ref. 29, whereas the others were described in this study and named C, D, and F to I for simplicity. The 11 captive *castanotis* zebra finches were sampled from two recently wild-derived populations (“Bielefeld” no. 19 in ref. 28 and “Melbourne” in ref. 30) and three domesticated populations (“Krakow” no. 11 in ref. 28, “Seewiesen,” and “Spain” in ref. 4). The sampled individuals were 238 to 1,882 d old ([Dataset S2](#)). All 11 captive *castanotis* zebra finch testes were large (longest diameter: 3 to 5 mm) compared to the testis of the *castanotis* × *guttata* zebra finch (longest diameter: ~1 mm), suggesting that the *castanotis* × *guttata* male might have been sexually inactive, which might have resulted in the low coverage of GRC sequences in its testis sample (*SI Appendix, Figs. S1F, S3A, and S7*) compared to the *castanotis* samples.

**Cytogenetics.** To determine the presence or absence of GRC in mature sperm, we conducted FISH on the ejaculates of one brother of matriline A and four brothers of matriline B (see sampling and [Datasets S1 and S2](#)) following a protocol modified from ref. 31. In brief, we collected fresh ejaculate in 10  $\mu\text{L}$  phosphate-buffered saline (PBS) with a 20- $\mu\text{L}$  pipette and osmotic shocked the sample by adding 250  $\mu\text{L}$  (~20-fold the volume of the sample) of 1% sodium-citrate solution for 20 min. We then spread the sample on a microscopy glass slide placed on a 60°C heating plate, following the Meredith’s technique (31), and let it dry on the heating plate. We used a FISH probe (4) for the gene *dph6*, which is present in about 300 copies on the GRC but only represented as a single-copy paralog on the A chromosomes (4). We amplified the probe by PCR from DNA extracted from a *castanotis* zebra finch testis using the primer sequences F-ACGCTCTTGCCTGACCCCTTTCAGA, R-TGCATAGAGTTCTCCATCAGACAGACA, taken from ref. 4, and then labeled it with tetramethylrhodamine-5-dUTP via nick translation (32). FISH was then performed on the ejaculate preparations (32). The hybridization mix consisted of 12  $\mu\text{L}$  formamide, 6  $\mu\text{L}$  dextran sulfate, 1.5  $\mu\text{L}$  20× saline-sodium citrate, 0.5  $\mu\text{L}$  salmon sperm, 0.5  $\mu\text{L}$  sodium dodecyl sulfate, 3  $\mu\text{L}$  *dph6* probe, and 6.5  $\mu\text{L}$  H<sub>2</sub>O. We applied 10 min of denaturation at 70°C.

FISH preparations were analyzed along the z-axis using a DSD2 confocal unit fitted to an Eclipse Ti inverted microscope (Nikon) with a Zyla sCMOS camera (Andor) under near-ultraviolet (405 nm) and green (550 nm) sequential excitation, obtained with a pE-4000 (CoolLED) device, to determine whether the GRC signal from the GRC-carrying sperm was inside the nucleus. To quantify the fraction of GRC-carrying sperm in the ejaculate, we sampled eight fields (2 rows × 4 columns) from each FISH preparation in which there were more than 20 nonoverlapping spermatozoa (hence, we did not use prior information on a GRC signal) using a Zeiss Axio Imager M2 microscope equipped with DAPI (465 nm), red (572 nm), and green (519 nm) filters at 40× magnification. We took photos with a Zeiss Axioacam 512 color camera using ZEN blue 3.1 software. We then counted the total number of spermatozoa, the number of GRC-carrying spermatozoa, and the number of expelled free-floating GRC micronuclei for all sampled fields. Additional images were generated under Leica DM 6000/HX PC APO 100×-Oil immersion for visualization. Images were processed using Fiji (33).

**Whole-Genome Sequencing.** We extracted genomic DNA from testis, ejaculate, and liver samples using a phenol chloroform extraction (for details, see

*SI Appendix, Material and Methods*) and from blood using the NucleoSpin Blood QuickPure Kit (from the company Macherey-Nagel) following the manufacturer’s instructions. Details of library preparation methods and library size for each sample can be found in [Dataset S2](#). Note that the raw sequencing reads of *castanotis* males A (SR00100), B (Spain\_1), and F (Spain\_2) were taken from ref. 4 (Sequence Read Archive accession numbers SRX6431677 to SRX6431681, SRX6431686, and SRX6431688 to SRX6431693).

All PCR-free libraries were constructed and sequenced with 20-fold coverage on the Illumina HiSeq 3000/4000 (ejaculate and blood samples) or the NovaSeq SP platforms (testis and liver; 2× 150 basepairs [bp], paired-end reads) at the Institute of Clinical Molecular Biology of Kiel University, Germany.

To study individual repeatability of GRC elimination patterns and to compare between-family differences, we sequenced DNA from 15 ejaculates (see *Samples*) and from the blood of the four parents (founders of families A and B as somatic baseline) using PCR-free Illumina libraries. We also sequenced DNA from the testis of one son per family (A1 and B1 in Fig. 2) and a pair of testis and liver from an uncle of males in family A (A0 in Fig. 2) to compare the GRC content between the testis and the ejaculate, which provides a measure of the remaining GRC content in the ejaculate after elimination during spermatogenesis.

To study the genetic diversity and the coevolutionary history of the GRC and its associated mtDNA, we sequenced (PCR-free) DNA from the testes and liver samples of a single male from each of the four major mtDNA haplotypes in captivity (i.e., males C, D, and G to H; see *Samples*).

Additionally, we generated 10X Chromium linked-read data for DNA samples that were extracted from testis and liver samples using magnetic beads on a Kingfisher robot (for details, see ref. 4) from one captive *guttata* × *castanotis* hybrid with *guttata* phenotype (*SI Appendix, Results* and Fig. 3), one domesticated *castanotis* zebra finch (male E) from population “Seewiesen” with a different mtDNA haplotype [E *sensu* (29)], and one wild-caught *castanotis* zebra finch from Western Australia (male I). All Chromium libraries were constructed and sequenced to high coverage (for details, see [Dataset S2](#)) as paired-end reads (2 × 150 bp) using the Illumina NovaSeq 6000 S4 platform at SciLifeLab Stockholm.

**Raw Reads Processing.** Raw reads from all whole-genome sequencing libraries (PCR-free and Chromium) were processed following a modified version of the “Genome Analysis Toolkit Best Practices Pipeline” (34). We filtered the raw reads using “BBduk” (35) and trimmed the last base of each sequence, putative adapter sequences, and bases with low quality and only kept high-quality reads that were more than 50 bp long. Because no GRC reference assembly is available, we only considered GRC-linked regions that have an A-chromosomal paralog (4). We then mapped each library (paired-end reads) against the reference somatic genome, *taeGut1* (36), using BWA-MEM v0.7.17 (37) with the default settings while marking shorter split hits as secondary. In a next step, we used the “Picard (38) MarkDuplicates” option to mark mapped reads that might result from PCR duplication to reduce PCR bias in the abundance of certain DNA fragments during sequencing. Finally, we analyzed coverage and called SNPs (see *Coverage Analysis and SNP Analysis* below) for downstream analysis. For a detailed pipeline and the corresponding scripts, see *Data Accessibility*.

**Coverage Analysis.** To quantify and compare the amount of GRC in the ejaculates, we applied an analysis of sequencing coverage that was adapted from ref. 4. As GRC-linked sequences are often difficult to distinguish from their ancestral A-chromosomal paralogs (because the latter vary substantially between individuals), GRC content is most easily quantified by focusing on sequences that reside on the GRC in numerous copies (compared to just two A-chromosomal copies) (4).

To do this, we calculated the ratios of sequencing coverage of pairs of germline samples (ejaculate or testis) over their corresponding somatic samples (liver or blood, averaged for the two parents when applicable) as GRC-free control tissue in nonoverlapping adjacent windows of 1 kb in width across the entire genome. For each library, we calculated read coverage using “SAMtools v1.6 (39) depth” per bp and used average values for each 1-kb window. For each germline sample, we then calculated the coverage ratio between the germline and its corresponding soma library and log<sub>2</sub> transformed the values after correcting for variation in library size (i.e., by dividing the coverage per window by the total number of bp sequenced for that library).

To quantify testis enrichment in coverage (compared to soma), we first removed windows with too low coverage (i.e., those in which both soma and germline samples had < threefold coverage). Second, we calculated mean and SD of coverage for all 1-kb windows of each somatic library and removed

windows with coverage  $> 2$  SD above the mean of a given library. Such high coverage values indicate duplications on the A-chromosomal paralog, which makes the quantification of copy number enrichment in testis difficult. Third, we centered the  $\log_2$ -transformed germline-to-soma coverage ratios of the high-quality windows on the median of the filtered windows.

To compare the amount of the GRC-linked DNA that remained in the ejaculates between the two focal families (matrilines A and B), we selected those 1-kb windows ( $n = 1,742$ ) that showed significant testis coverage enrichment (i.e.,  $\log_2$  testis-to-soma coverage ratio  $\geq 2$ ) using nine individuals for which we sequenced testis DNA (males A through I, see *Samples* and Fig. 2A). The nine mtDNA haplotypes represent most of the genetic variation of the mitochondria in the wild and in captivity (all males except "I" are from captivity; *SI Appendix*, Figs. S4 and S5). Then, we calculated the median  $\log_2$  ejaculate-to-soma coverage ratios of the selected windows for the 15 ejaculates from five additional *castanotis* males stemming from two families (i.e., matrilines A and B), including one male per family for which we had obtained cytogenetic data (ejaculates in Fig. 2B and C). As GRC-positive controls for the ejaculate samples, and to quantify family-specific elimination efficiency, we also examined three testis libraries from families A and B (testes in Fig. 2B and C).

We estimated the individual repeatability of the median  $\log_2$  ejaculate-to-soma coverage ratios of the selected windows (response variable) among the 15 ejaculate samples using a mixed-effect model with the "lmer" function in the "lme4" (40) package in R v4.0.3 (41) in which we fitted individual identity and ejaculate as random effects. Matriline repeatability was estimated by fitting matriline identity as an additional random effect in the previous mixed-effect model. To estimate the between-family difference in the GRC amount in ejaculates, we added "matriline" as a fixed effect in the mixed-effect model of individual repeatability. To estimate the amount of reduction in the GRC content in ejaculate samples compared to the testis in each family (A and B), we used two linear models (one for each family; Fig. 2) using the "lm" function in the R package "stats." Here, we used the median of the  $\log_2$  germline-to-soma coverage ratio (of the 1,742-selected windows) for each germline sample as a response variable because it is less sensitive to potential copy number variation in the GRC-linked high-copy number genes compared to the mean. In these two linear models, we only fitted the type of germline tissue (ejaculate or testis) as a fixed effect. For model structures and outputs, see also *Dataset S3*.

**SNP Analysis.** We used mpileup in SAMtools v1.6 (39) and call in bcftools v1.9 (42) to call SNPs and a customized R script to filter for high-confidence SNPs of interest (see *Data Accessibility*) as follows. To study the overall between-individual variation in GRC haplotypes, we called SNPs for all testis/soma pairs (one hybrid *castanotis*  $\times$  *guttata* and 12 *castanotis* males) simultaneously and selected high-confidence sites (see *SI Appendix, Materials and Methods* for details). We then identified high-confidence testis-specific alleles by selecting sites for which a) the soma library had more than 10 reads, b) the allele was found in  $\geq 3$  reads in the germline sample, and c) the allele was only present in the testis sample but not in the corresponding soma sample. We also identified those testis-specific SNPs that were private to only one of the 13 sequenced individuals (i.e., those testis-specific alleles that were absent from all soma and testis libraries except for the focal one). To reduce false positives in the private testis-specific SNPs, we focused on regions that contained multiple private testis-specific SNPs defined as 10-kb nonoverlapping adjacent windows with at least three such private SNPs (*SI Appendix*, Fig. S3 and Fig. 3B and C).

To study the GRC content in ejaculates, we called SNPs simultaneously for mapped reads from the 15 ejaculates, the two testes of the brothers from A1 and B1 in Fig. 2B and C, the four blood samples of their parents, and one pair of testis and liver from one uncle of family A (A0 in Fig. 2B). We filtered for high-quality, germline-specific alleles following the same procedure as described in the paragraph above. We then identified the 1-kb nonoverlapping adjacent windows that contained at least 15 germline-specific SNPs (*SI Appendix*, Fig. S2).

To study the extent of A-chromosomal introgression of *castanotis* DNA into the captive population of *castanotis*  $\times$  *guttata* hybrids, we called SNPs for the combined soma libraries of the hybrid (liver) and a pool of 100 wild-caught *castanotis* zebra finches (43) (blood). We then filtered for those high-quality SNPs that were homozygous in the (predominantly *guttata*) hybrid but absent from the 100 wild *castanotis* zebra finches. Additionally, we filtered for SNPs that were heterozygous in the soma (liver) library of the hybrid individual. We calculated the number of fixed (i.e., homozygous) *guttata* SNPs and the number of heterozygous sites for nonoverlapping adjacent windows of 500 kb. We considered windows with a low number of fixed *guttata*-specific SNPs as signals of *castanotis* introgression. We determined the copy number of those *castanotis*-introgressed sequences by their level of heterozygosity: a run of homozygosity

would indicate two copies of one *castanotis* haplotype, a similar level of heterozygosity compared to the nonintrogressed regions suggests two *castanotis* haplotypes, and an extremely elevated heterozygosity level implies that one copy of the *castanotis*-haplotype segregates with a *guttata* haplotype.

**Haplotype Analysis.** All existing analyses, including assembly subtraction, read mapping, and coverage-based analyses, have shown that the *castanotis* GRC is mostly composed of high-copy number genes that originated from the A chromosomes at various evolutionary timepoints (4, 8). Furthermore, no extended GRC-linked sequences are known that do not have an A-chromosomal paralog (4, 8). Our read mapping analyses confirmed that all identified GRC regions contain A-chromosomal paralogs. Therefore, to study the phylogenetic relationships between GRC haplotypes and mtDNA haplotypes, we focused on single-copy GRC genes that were highly diverged from their A-chromosomal paralogs. The latter is necessary because only high divergence ensures that all reads map without error to their correct origin, being either from the GRC or from the A-chromosomal paralog. We screened the published list of 267 GRC-linked genes (4) for a high number and high density of testis-specific SNPs and in single- or low-copy numbers following a customized pipeline (for details, see [https://github.com/fjrui/ruano/In\\_Silico\\_SeqCap\\_10xG](https://github.com/fjrui/ruano/In_Silico_SeqCap_10xG)). In brief, we first mapped the 10X reads from the published testis sample against the published transcripts of the focal genes of its GRC and the A-chromosomal paralogs (4). We extracted 10X barcodes from the reads that mapped only to the GRC paralogs. Then, we selected and assembled a GRC consensus from all reads with the same barcodes using Supernova v2 (44). To extend these GRC contigs while avoiding sequences from A-chromosomal paralogs, we masked the positions of those contigs that showed reads mapped from the liver library. We then repeated the mapping, barcode selection, and assembly of the testis sample against the masked GRC contigs. We thus generated nine single-copy GRC loci: *trim71<sub>GRC</sub>*, *bicc1<sub>GRC</sub>*, *pim3<sub>GRC</sub>*, *msh4<sub>GRC</sub>*, *surf4<sub>GRC</sub>*, *262<sub>intergenic<sub>GRC</sub></sub>*, *aldh18a1<sub>GRC</sub>*, *rfc1<sub>GRC</sub>*, and *eppk1<sub>GRC</sub>*. Gene names were based on BLAT hits with at least one of the annotated genomes (zebra finch, chicken [*Gallus gallus*], or human [*Homo sapiens*]).

Using BWA-MEM v0.7.17 (37), we separated the reads from the GRC genes and their A-chromosomal paralogs by mapping reads against both the A-chromosomal [i.e., sequence on *taeGut1* (36)] and the GRC paralogs. For single-copy GRC genes, this allowed us to generate naturally phased GRC haplotypes for each sample and to check for heterozygosity in terms of GRC haplotypes. Then, we used mpileup in SAMtools v1.6 (39) and call in bcftools v1.9 (42) to call SNPs for each library of those mapped reads. For the nine single-copy GRC loci (Fig. 4 and *SI Appendix*, Figs. S6 and S7), we then generated one GRC haplotype for each germline sample by substituting the called alternative allele from the reference consensus allele using customized R scripts (see *Data Accessibility*). Unfortunately, the coverage of GRC-linked reads was too low for the *castanotis*  $\times$  *guttata* hybrid male with the *guttata* GRC (*SI Appendix*, Fig. S7), so we were unable to construct the *guttata* version of these low-copy GRC-linked genes. Hence, we analyzed the GRC haplotypes of the *castanotis* zebra finches only.

The nine GRC loci were clearly in single copy because all testis mappings were homozygous (*SI Appendix*, Fig. S6A) and showed 41% read coverage compared to the genomic background of the corresponding sample (*SI Appendix*, Fig. S7 A–C). This confirms previous work showing that male germline cells carry a single copy of the GRC (1, 4, 5, 7, 9, 13) (i.e., in libraries of developed testes, the GRC is expected to have less than 50% of the read coverage of A chromosomes).

Then, for phylogenetic analysis, we concatenated the nine single-copy GRC loci to represent GRC haplotypes. Finally, we used DnaSP v6.12.01 (45) to calculate the mean number of pairwise differences per site ( $\pi$ ) for each GRC haplotype.

**Phylogenetic Analysis.** All phylogenetic trees were built using RAxML-NG v1.0.2 (46) assuming a general time-reversible model and a discrete gamma model of rate heterogeneity with 100 randomized parsimony starting trees and 1,000 bootstrap replicates (for details, see *Data Accessibility*). To demonstrate that the matriline of the hybrid *castanotis*  $\times$  *guttata* male was *castanotis*-B, we constructed one mtDNA tree using the gap-free alignment of mtDNA sequences from all testis samples used in this study (Fig. 3F).

To compare the evolutionary histories of the mitogenome and the associated GRC haplotypes from the same *castanotis* individuals, we constructed one best-supported tree for each gap-free alignment of the mtDNA and GRC haplotypes. The mtDNA tree was rooted by two *guttata* mtDNA assemblies [SRA accession numbers SRR2299402 (47) and SRR3208120 (48)], whereas the GRC haplotype tree was rooted at the midpoint of the two most distantly related haplotypes (Fig. 4A; for haplotype network analysis, see *SI Appendix, Results* and Fig. S8).



We extracted the pairwise distance matrices of the two phylogenetic trees in R and tested for similarity using a linear mixed-effect model. Here, we fitted the pairwise distance of the mtDNA haplotypes as response variable and the pairwise distance of the GRC haplotypes as fixed effect and included row and column identities as two random factors to control for overall position of each sample in the matrices (Dataset S6). We z-standardized the response variable and the covariate to account for the drastic difference in units between the two types of sequences.

**Data Accessibility.** All sequencing data have been deposited in the Sequence Read Archive (BioProject accession number PRJNA741250) (50). All alignments have been deposited in Figshare ([https://figshare.com/articles/dataset/Additional\\_datasets\\_for\\_Occasional\\_paternal\\_inheritance\\_of\\_the\\_germline-restricted\\_chromosome\\_in\\_songbirds\\_/14845026](https://figshare.com/articles/dataset/Additional_datasets_for_Occasional_paternal_inheritance_of_the_germline-restricted_chromosome_in_songbirds_/14845026)) (51). Additional supporting data have been deposited in the Open Science Framework (<https://doi.org/10.17605/OSF.IO/N9X2G>) (52). Previously published data were used for this work (4). All supporting pipelines and scripts have been deposited in the Open Science Framework (<https://doi.org/10.17605/OSF.IO/N9X2G>) and in GitHub [https://github.com/frjuizruano/ln\\_Silico\\_SeqCap\\_10xG](https://github.com/frjuizruano/ln_Silico_SeqCap_10xG).

- M. I. Pigozzi, A. J. Solari, Germ cell restriction and regular transmission of an accessory chromosome that mimics a sex body in the zebra finch, *Taeniopygia guttata*. *Chromosome Res.* **6**, 105–113 (1998).
- B. Hansson, On the origin and evolution of germline chromosomes in songbirds. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11570–11572 (2019).
- J. J. Smith, V. A. Timoshevskiy, C. Saraceno, Programmed DNA elimination in vertebrates. *Annu. Rev. Anim. Biosci.* **9**, 173–201 (2021).
- C. M. Kinsella et al., Programmed DNA elimination of germline development genes in songbirds. *Nat. Commun.* **10**, 5468 (2019).
- A. A. Torgasheva et al., Germline-restricted chromosome (GRC) is widespread among songbirds. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11845–11850 (2019).
- L. del Priore, M. I. Pigozzi, Histone modifications related to chromosome silencing and elimination during male meiosis in Bengalese finch. *Chromosoma* **123**, 293–302 (2014).
- Y. Itoh, K. Kampf, M. I. Pigozzi, A. P. Arnold, Molecular cloning and characterization of the germline-restricted chromosome sequence in the zebra finch. *Chromosoma* **118**, 527–536 (2009).
- M. K. Biederman et al., Discovery of the first germline-restricted gene by subtractive transcriptomic analysis in the zebra finch, *Taeniopygia guttata*. *Curr. Biol.* **28**, 1620–1627.e5 (2018).
- L. P. Malinovskaya et al., Germline-restricted chromosome (GRC) in the sand martin and the pale martin (Hirundinidae, Aves): Synapsis, recombination and copy number variation. *Sci. Rep.* **10**, 1058 (2020).
- J. H. Werren, Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10863–10870 (2011).
- A. Burt, R. Trivers, *Genes in Conflict: The Biology of Selfish Genetic Elements* (Harvard University Press, 2006).
- L. D. Hurst, A. Atlan, B. O. Bengtsson, Genetic conflicts. *Q. Rev. Biol.* **71**, 317–364 (1996).
- M. I. Pigozzi, A. J. Solari, The germ-line-restricted chromosome in the zebra finch: Recombination in females and elimination in males. *Chromosoma* **114**, 403–409 (2005).
- C. Goday, M. I. Pigozzi, Heterochromatin and histone modifications in the germline-restricted chromosome of the zebra finch undergoing elimination during spermatogenesis. *Chromosoma* **119**, 325–336 (2010).
- S. Schoenmakers, E. Wassenaar, J. S. E. Laven, J. A. Grootegoed, W. M. Baarends, Meiotic silencing and fragmentation of the male germline restricted chromosome in zebra finch. *Chromosoma* **119**, 311–324 (2010).
- L. Smeds et al., Evolutionary analysis of the female-specific avian W chromosome. *Nat. Commun.* **6**, 7330 (2015).
- S. Luo et al., Biparental inheritance of mitochondrial DNA in humans. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 13039–13044 (2018).
- M. Alexander et al., Mitogenomic analysis of a 50-generation chicken pedigree reveals a rapid rate of mitochondrial evolution and evidence for paternal mtDNA inheritance. *Biol. Lett.* **11**, 20150561 (2015).
- A. Arcones, R. Ponti, D. R. Vieites, Mitochondrial substitution rates estimation for divergence time analyses in modern birds based on full mitochondrial genomes. *Ibis* **163**, 1463–1471 (2021).
- H. Ellegren, Evolutionary stasis: The stable chromosomes of birds. *Trends Ecol. Evol.* **25**, 283–291 (2010).
- A. Kapusta, A. Suh, Evolution of bird genomes—a transposon’s-eye view. *Ann. N. Y. Acad. Sci.* **1389**, 164–185 (2017).
- T. M. Degrandi et al., Introducing the bird chromosome database: An overview of cytogenetic studies in birds. *Cytogenet. Genome Res.* **160**, 199–205 (2020).
- C. N. Balakrishnan, S. V. Edwards, Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* **181**, 645–660 (2009).
- M. A. Wilson Sayres, K. E. Lohmueller, R. Nielsen, Natural selection reduced diversity under human y chromosomes. *PLoS Genet.* **10**, e1004064 (2014).
- S. Berlin, H. Ellegren, Chicken W: A genetically uniform chromosome in a highly variable genome. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 15967–15969 (2004).
- W. G. Hill, A. Robertson, The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).
- J. M. Comeron, A. Williford, R. M. Kliman, The Hill-Robertson effect: Evolutionary consequences of weak selection and linkage in finite populations. *Heredity* **100**, 19–31 (2008).
- W. Forstmeier, G. Segelbacher, J. C. Mueller, B. Kempnaers, Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Mol. Ecol.* **16**, 4039–4050 (2007).
- J. A. Mossman, T. R. Birkhead, J. Slate, The whole mitochondrial genome sequence of the zebra finch (*Taeniopygia guttata*). *Mol. Ecol. Notes* **6**, 1222–1227 (2006).
- S. Jerónimo et al., Plumage color manipulation has no effect on social dominance or fitness in zebra finches. *Behav. Ecol. Evol.* **29**, 459–467 (2018).
- R. Meredith, A simple method for preparing meiotic chromosomes from mammalian testis. *Chromosoma* **26**, 254–258 (1969).
- J. Camacho, J. Cabrero, M. López-León, D. Cabral-de-Mello, F. Ruiz-Ruano, “Grasshoppers (Orthoptera)” in *Protocols for Cytogenetic Mapping of Arthropod Genomes*, I. V. Sharakhov, Ed. (CRC Press, 2014), pp. 381–438.
- J. Schindelin et al., Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
- G. A. Van der Auwera et al., From FastQ data to high-confidence variant calls: The Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* **43**, 11.10.1–11.10.33 (2013).
- B. Bushnell, BBMap. <https://sourceforge.net/projects/bbmap/> (2015). Accessed 19 September 2018.
- W. C. Warren et al., The genome of a songbird. *Nature* **464**, 757–762 (2010).
- H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [Preprint]* (2013). <https://sourceforge.net/projects/bio-bwa/>. Accessed 13 September 2018.
- Broad Institute, Picard Tools. *Broad Institute, GitHub Repos.* (2019).
- H. Li et al., 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
- R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing Vienna, Austria, 2020). <https://www.R-project.org/>. Accessed 14 January 2022.
- H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
- U. Knief et al., Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biol.* **17**, 199 (2016).
- N. I. Weisenfeld, V. Kumar, P. Shah, D. M. Church, D. B. Jaffe, Direct determination of diploid genome sequences. *Genome Res.* **27**, 757–767 (2017).
- J. Rozas et al., DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **34**, 3299–3302 (2017).
- A. M. Kozlov, D. Darriba, T. Flouri, B. Morel, A. Stamatakis, RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).
- S. Singhal, et al., Stable recombination hotspots in birds. *Science* **350**, 928–932 (2015).
- J. H. Davidson, C. N. Balakrishnan, Gene regulatory evolution during speciation in a songbird. *Genes Genomes Genetics* **6**, 1357–1364 (2016).

49. U. Knief, W. Forstmeier, Y. Pei, J. Wolf, B. Kempenaers, A test for meiotic drive in hybrids between Australian and Timor zebra finches. *Ecol. Evol.* **10**, 13464–13475 (2020).
50. Y. Pei, W. Forstmeier, F. J. Ruiz-Ruano, A. Suh, B. Kempenaers, Zebra finch ejaculates (*Taeniopygia guttata castanotis*), testis and somatic tissue (*T. g. castanotis* and one *castanotis* x *guttata* hybrid) sequencing. Sequence Read Archive (SRA). <https://www.ncbi.nlm.nih.gov/sra/PRJNA741250>. Deposited 21 June 2021.
51. Y. Pei, Additional datasets for “Occasional paternal inheritance of the germline-restricted chromosome in songbirds”. Figshare. [https://figshare.com/articles/dataset/Additional\\_datasets\\_for\\_Occasional\\_paternal\\_inheritance\\_of\\_the\\_germline-restricted\\_chromosome\\_in\\_songbirds\\_/14845026](https://figshare.com/articles/dataset/Additional_datasets_for_Occasional_paternal_inheritance_of_the_germline-restricted_chromosome_in_songbirds_/14845026). Deposited 2 December 2021.
52. Yifan Pei, Supporting information for “Occasional paternal inheritance of the germline-restricted chromosome in songbirds.” Open Science Framework. <https://doi.org/10.17605/OSF.IO/N9X2G>. Deposited 30 December 2021.