

**Evolutionary genetics of reproductive performance  
in the zebra finch**

**Dissertation**

**Fakultät für Biologie**

**Ludwig-Maximilians-Universität München**

**durchgeführt am**

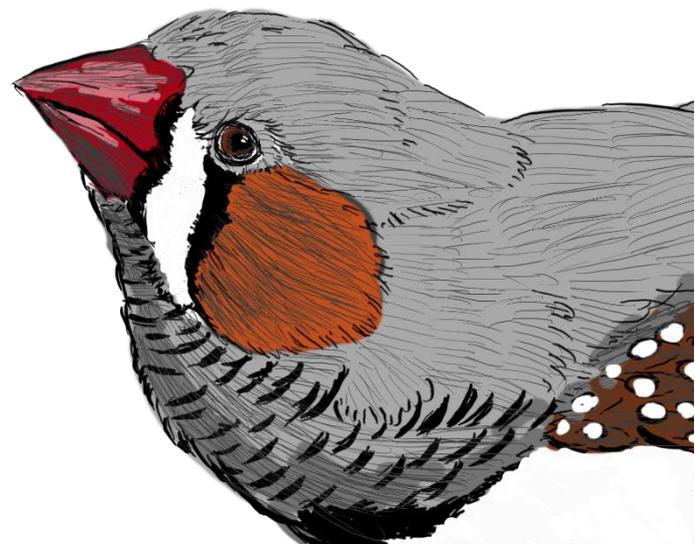
**Max-Planck-Institut für Ornithologie**

**Seewiesen**

**vorgelegt von**

**Yifan Pei**

**August 2021**



Erstgutachter: Prof. Dr. Bart Kempenaers

Zweitgutachter: Prof. Dr. Jochen Wolf

Tag der Abgabe: 30.08.2021

Tag der mündlichen Prüfung: 27.01.2022

Diese Dissertation wurde unter der Leitung von Dr. **Wolfgang Forstmeier** angefertigt.

Cover illustration: Yifan Pei

To my grandparents, Wenhua Fan & Shiquan Guo,  
who history has forgotten.



# Contents

<b>Summary</b> .....	<b>1</b>
<b>General introduction</b> .....	<b>5</b>
<b>Chapter 1</b> .....	<b>21</b>
<i>Proximate causes of infertility and embryo mortality in captive zebra finches</i> .....	22
<b>Chapter 2</b> .....	<b>62</b>
<i>Offspring performance is well buffered against stress experienced by ancestors</i> .....	63
<b>Chapter 3</b> .....	<b>90</b>
<i>Overdominance effects of a microchromosomal inversion on multiple fitness components in the zebra finch</i> .....	91
<b>Chapter 4</b> .....	<b>134</b>
<i>Maintaining an inversion polymorphism by antagonistic effects on fitness</i> .....	135
<b>Chapter 5</b> .....	<b>162</b>
<i>Occasional paternal inheritance of the germline restricted chromosome</i> .....	163
<b>General discussion</b> .....	<b>226</b>
<b>Addresses of co-authors</b> .....	<b>237</b>
<b>Author contributions</b> .....	<b>239</b>
<b>Acknowledgement</b> .....	<b>241</b>
<i>Curriculum vitae</i> .....	<b>243</b>
<b>Statutory declaration and statement</b> .....	<b>245</b>

Nothing in biology makes sense except in the light of evolution.

- Theodosius Dobzhansky, 1973

## Summary

Having difficulties conceiving a child or experiencing reproductive failure is not only an emotional and personal concern for a family, but also an evolutionary puzzle in a wide range of species, including birds. Reproduction is one of the most complex traits directly targeted by natural selection. It involves a wide range of fitness-related traits that may differ between males and females, including fertility, fecundity, offspring survival and individual survival. On the one hand, repeatable variation in reproductive performance between individuals and between combinations of partners suggests that reproductive failures may have a genetic or epigenetic basis (**Chapter 1**). On the other hand, natural selection favors mutations that make an individual better adapted to the environment and produce more offspring. Such mutations are expected to rise in frequency over time, which in turn should reduce the phenotypic and (additive) genetic variation in reproductive performance within the population. In my dissertation, I use the zebra finch *Taeniopygia guttata* as a model species, because the species shows remarkably high rates of infertility and embryo mortality that vary greatly between individuals and populations. As the persistence of such variability poses an interesting challenge for evolutionary geneticists, I here attempt to shed light on the proximate causes and the genetic architecture of reproductive failure (**Chapters 1 - 4**). I also focus on large chromosomal inversions, i.e. segments of chromosomes that are reversed in sequence from end to end, which can physically link hundreds of genes into non-recombining haplotypes. This facilitates the study of genetic effects on variation in fitness traits. I therefore investigate the evolutionary mechanisms that appear to maintain the observed genetic polymorphisms (**Chapters 3 & 4**).

My dissertation starts with a detailed description of the phenotypic variation in all reproductive performance traits and with a comparison of the importance of some proximate causes of reproductive failure (**Chapters 1 & 2**). My **first chapter** compiles a database of fitness-related traits based on the fate of 23,000 zebra finch eggs that were laid in captivity. Using a quantitative genetics approach, we found that poor early growth conditions, old age, and inbreeding (pairing of relatives) are significant factors in predicting rates of infertility, embryo and nestling mortality, but they only explain a small portion of the variation in fitness-related traits (**Chapter 1**). It has been widely assumed that the reproductive performance of an individual is sensitive to early developmental stress (e.g. poor early growth conditions) and even the early stress experienced by ancestors (e.g. parents and grandparents). Yet, in contrast to this view, my **second chapter** demonstrates that the reproductive performances of

individual zebra finches are remarkably robust against early developmental stress, particularly against stress experienced by their parents and grandparents. After decomposing the variance components for each fitness-related trait (**Chapter 1**), we found that infertility is mainly a male-specific problem whereas low fecundity and high embryo mortality are largely due to the mothers. The specific combination of genetic (i.e. biological) parents is important in predicting the survival of embryos whereas the combination of social parents is important in predicting the survival of nestlings. These analyses suggest a potential genetic basis for fitness-related traits. Hence, **Chapter 1** explores the heritability and the additive-genetic basis of the fitness-related traits. Despite the overall low heritability, we found evidence that male-specific and female-specific reproductive performance traits tend to be negatively correlated at the additive-genetic level. This suggests that some of the heritable variation in fitness can be maintained via alleles having opposing effects on male fertility and female fecundity plus offspring survival.

**Chapters 3 & 4** test how the genotypes of large chromosomal inversion polymorphisms influence male and female reproductive performance and individual survival. Zebra finches have at least four large inversion polymorphisms on chromosomes *Tgu5*, *Tgu11*, *Tgu13* and *TguZ* that all segregate at allele frequencies close to 50%. However, the mechanisms that maintain these polymorphisms are still unclear, except for the inversion on *TguZ*. It was recently discovered that the inversion on *TguZ* shows heterosis, whereby males that are heterozygous for the inversion had increased siring success. In **Chapter 3**, I scanned the whole genome for additional large chromosomal inversions, particularly on microchromosomes, using whole genome sequencing data from multiple wild and captive zebra finches. I found and characterized the two additional large inversion polymorphisms on the microchromosomes *Tgu26* and *Tgu27*. When studying the genotypic effects of these two inversions, I found that individuals putatively heterozygous for the inversion on *Tgu27* showed significant heterosis in nearly all fitness-related traits. After re-examining the genotypic effects of the previously identified inversion polymorphism on *Tgu11* (**Chapter 4**), I found that males that carry the derived inversion on *Tgu11* have higher siring success and females have higher fecundity, while individuals that are homozygous for the derived inversion type exhibit a reduced probability of survival. Using the estimated effect sizes on fitness-related traits that cover a full reproductive cycle, I simulated the change of allele frequency of the derived inversion over time. The results are consistent with the idea that the derived inversion on *Tgu11* may initially have spread due to its additive effects that increase male and female reproduction performance, and later may have been stopped from going to fixation by its opposing detrimental effects at the homozygous state that reduce offspring and individual

survival. However, all these genotypic effects (**Chapters 3 & 4**) are small and leave the majority of the repeatable variation in reproductive performance unexplained.

Taking the search for possible reasons for reproductive failure one step further, my **Chapter 5** investigates an accessory germline-restricted chromosome (GRC) in songbirds that is only present in germline cells but absent from all somatic tissues. The GRC was first identified in 1998 and was thought to be strictly maternally inherited, like the mitochondria. However, I found that the GRC can sometimes also be inherited paternally via sperm. This suggests that GRC haplotypes which, besides maternal inheritance, can also occasionally be transmitted by males may exhibit an advantage over strictly maternally inherited GRC haplotypes. Such biparentally spreading haplotypes may hence have spread through the population, even if they were somewhat detrimental to the organism. These findings suggest that the GRC may be a promising genetic candidate for the causation of infertility and embryo mortality in the zebra finch (**Chapter 5**).

Taken together, my dissertation illustrates that some of the genetic causes of reproductive failure may be due to heterozygote advantage or antagonistic pleiotropy on different fitness-related traits (**Chapters 1, 3 & 4**). In addition, the identified genetic factors include additive-genetic effects and segregating large inversion polymorphisms. However, the effects of the identified genetic and environmental causes are typically small (**Chapters 1 - 4**), suggesting a potential role for additional genetic factors that sit outside of the regular autosomes and sex chromosomes. Lastly, **Chapter 5** reveals a peculiar inheritance pattern of an accessory germline-restricted chromosome, suggesting it to be a new candidate to study the genetic causes of infertility and embryo mortality at least in the zebra finch.



## General introduction

Who survives better? Who reproduces more? Why would variation in reproductive performance exist? Is poor reproductive performance transmitted from parents to offspring? These are key concerns of the general public as much as of many (evolutionary) biologists. These questions boil down to the fundamental ingredients for evolution to take place: variation in reproductive success (often referred to as fitness), variation in genes and some form of causal connection between the two. One expects that evolution would constantly select against genotypes that lead to poor reproductive performance, but infertility and embryo mortality are commonly observed in many species, including humans (Miyamoto et al. 2012; Jarvis 2017).

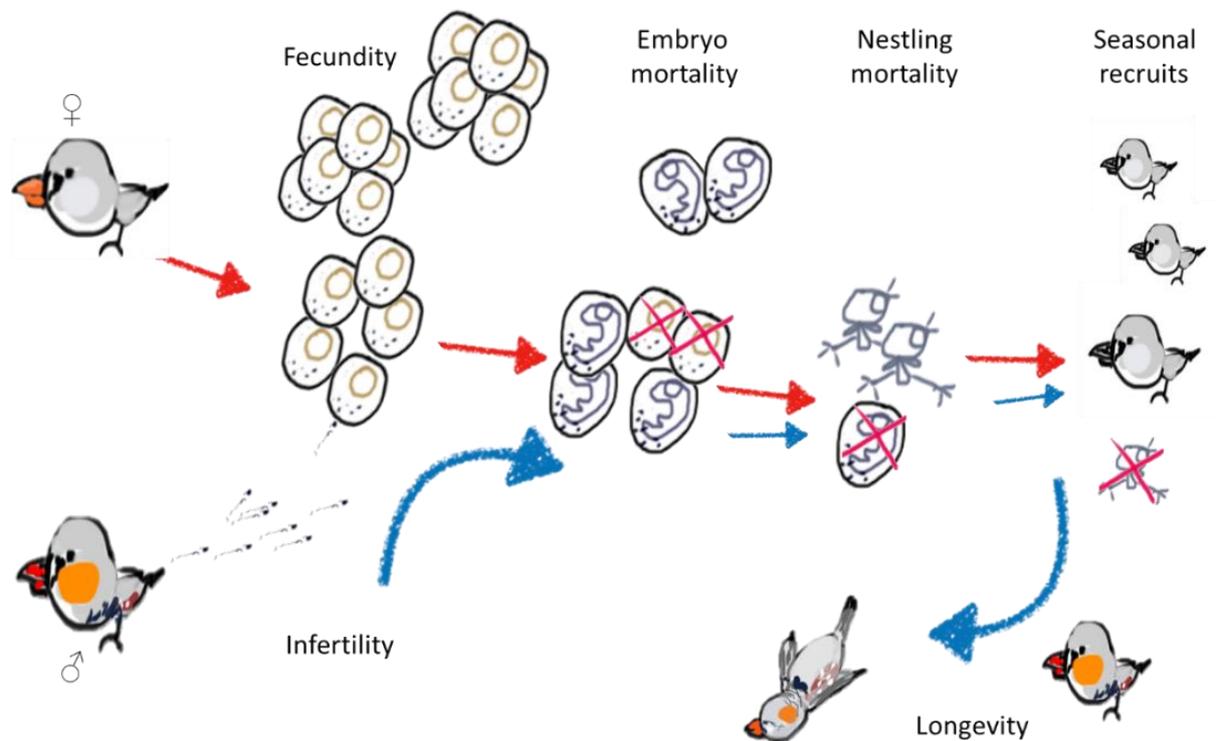
To explore potential answers of these questions, researchers have focused on identifying the proximate (intrinsic) causes of poor performance, such as aging (Harely 1990; Speroff 1994; Lecomte et al. 2010; Charmantier et al. 2015), poor early growth conditions (Monaghan 2008; Tschirren et al. 2009; Drummond and Ancona 2015; Kraft et al. 2019) and pairing between relatives (inbreeding) (Amos et al. 2001; Charlesworth and Willis 2009; Forstmeier et al. 2012; Hemmings et al. 2012), as well as on linking the variation in performance to certain physical or behavioral characteristics (Dingemanse and Réale 2005; Cote et al. 2008; Dingemanse et al. 2010; Schielzeth et al. 2011). A few factors were found to correlate with reproductive failure, including infertility and embryo mortality. Yet, the majority of the variation in reproductive performance remains largely unexplained (Schielzeth et al. 2011; Hemmings et al. 2012; Tan et al. 2013). Individuals seem to be repeatable in their reproductive failure and such failure segregates in the population over generations, suggesting a role of genetic or epigenetic elements (Merilä and Sheldon 2000; Kosova et al. 2010). However, the genetic architecture behind the reproductive performance variation is unclear, even in humans and well-studied species like the zebra finch *Taeniopygia guttata*.

To study the genetic causes of reproductive failure it is essential to monitor a large number of individuals for their reproductive performance, while acquiring sufficient genetic and genomic information. Given the fast advancement in sequencing technology and reduction in sequencing costs in the past decade, it has been possible to sequence longer DNA molecules (e.g. linked-reads) and (re-)sequence more (different individuals and different tissue types). In both model and non-model organisms, there are newly identified genetic and genomic variants, including large chromosomal structural variants (Küpper et al. 2015; Lamichhaney et al. 2015; Knief et al. 2016; Mérot et al. 2021)

and tissue-restricted genetic elements, e.g. the songbird germline-restricted chromosome (Biederman et al. 2018; Makunin et al. 2018; Kinsella et al. 2019; Torgasheva et al. 2019). However, there have only been a few cases that directly linked the genetic variation, in forms of large chromosomal inversions, to the fitness variation (Lee et al. 2016; Kim et al. 2017; Knief et al. 2017; Mérot et al. 2020). This is perhaps due to the complexity of fitness measure itself, which includes a wide-range of quantitative traits. Some of the traits are difficult to measure and easily confounded by each other, especially in mammals. For instance, in humans, early embryo mortality before the first sign of pregnancy is almost impossible to notice, and it might often be difficult to differentiate the infertility of men from recurrent early pregnancy loss of his partner (Jarvis 2017).

### **Quantifying fitness variation in captivity**

To assess fitness in sexually reproducing animals, one has to quantify a wide range of complex traits, including survival and different traits of reproductive performance from both males and females. Wild populations are typically preferred to use for studies of fitness consequences because these animals live in their natural habitat. However, in the wild, fitness measurements are often confounded by the spatial and temporal heterogeneity of the environment and stochastic events, such as predation, undetected emigration and immigration. Captive animals are interesting complementary systems to repeatedly measure fitness-related traits under standardized conditions. One may be able to observe ongoing adaptation to the new captive environment, and one can trace the phenotypic variation through complete pedigrees. Therefore, I used captive zebra finches as my model system, so that I can finely partition and quantify fitness-related traits covering the full zebra finch life cycle (**Fig. 1; Chapter 1**), as well as repeatedly sample natural ejaculates from specific males (**Chapter 5**).



**Figure 1.** Male and female zebra finch reproductive performance and survival traits quantified in this dissertation (**Chapter 1**).

### Candidate genetic causes of reproductive failure

There exists a large variety of forms of genetic mutations, including point mutations (e.g. single nucleotide polymorphism or insertion/deletion), small (a few kilobases) insertions, deletions, copy number variation, and large (up to megabases) structural rearrangements. Typically, very few mutations are beneficial, and these are likely to increase in frequency and become fixed in the population quickly. Most of the mutations are neutral (e.g. synonymous changes, mutations in the non-coding regions), or (slightly) deleterious to the organism (e.g. indels or rearrangement that disrupt the coding sequences). Among the deleterious mutations, the ones with large effects are typically selected against by natural selection; although the forces of drift and recombination may complicate matters, strongly deleterious variants are typically effectively removed from the population. However, selection is inefficient when the deleterious mutations are recessive, especially in populations when inbreeding is rare. This is likely to be very common especially when the deleterious effect is small. Hence, we see that offspring that were produced by related individuals would suffer from reduced survival or low reproductive performance (i.e. inbreeding depression) (Charlesworth and Willis 2009). Moreover, slightly deleterious mutations (i.e. alleles) could segregate in the population via some balancing effects, maintaining genetic variation, e.g. heterozygote advantage where only individuals that are

homozygous suffer from reduced fitness, or antagonistic pleiotropy where the focal mutation shows opposing effects on the two sexes or on different components of fitness (Schluter et al. 1991). The segregating genetic variants might involve (slightly) deleterious alleles, in other words, the genetic causes of reproductive failure may explain the maintenance of some of the genetic variation. In this dissertation, I combine the classic quantitative genetic approach (**Chapters 1 & 2**) and the genetic variation-based approach (**Chapters 3 & 4**) to study the link between genes and fitness.

#### *Additive genetic components*

As quantitative traits, each fitness component is likely to be influenced by numerous loci. Among those loci, only a few are likely to show large effects. The vast majority of loci that make up the additive genetic component behind a quantitative trait like fitness is practically impossible to identify and study individually. A quantitative genetics approach can statistically investigate the relative importance of the overall additive genetic component of quantitative traits based on the pedigree information, without necessarily knowing the genomic and genetic information of the individual (Hill 2010). The maintenance of additive genetic variation is typically thought to be due to antagonistic pleiotropic effects on different fitness components or life-history trade-offs, such as trade-offs between the two sexes or between performance and survival (Stearns 1989).

#### *Supergenes - chromosomal inversions*

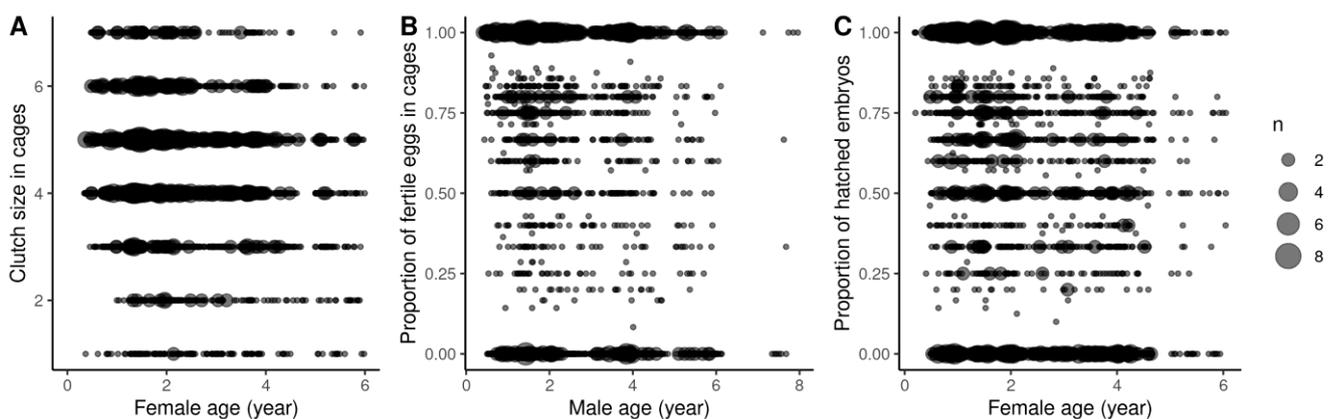
A chromosomal inversion is a type of structural variant where a segment of DNA is reversed in sequence by 180°. Inversions range from fractions of kilobases to several megabases and can physically link hundreds of genes (Hoffmann et al. 2004; Conrad and Hurler 2007; Wellenreuther and Bernatchez 2018). The most parsimonious scenario for a segregating polymorphism assumes that the inversion arose in a single individual and afterwards this inverted haplotype rose in frequency. The ancestral and the inverted types then may diverge over time via reduced recombination within the inverted region and via the accumulation of private mutations. Eventually the inverted haplotype will segregate in the population as a ‘supergene’. Recently, large inversion polymorphisms have been increasingly recognized to be important genetic elements that explain cases of extraordinary morphological and behavioral variants in plants (Lowry and Willis 2010; Lee et al. 2016), insects (Le Poul et al. 2014; Pracana et al. 2017; Jay et al. 2018; Mérot et al. 2020) and birds (Küpper et al. 2015; Lamichhaney et al. 2015; Tuttle et al. 2016; Kim et al. 2017; Knief et al. 2017). However, many segregating inversion polymorphisms may still be overlooked particularly on the microchromosomes, and the mechanisms that maintain the many identified inversion polymorphisms are still unknown (Knief et al. 2016; Wellenreuther and Bernatchez 2018; Mérot et al. 2021). My **Chapter 3** makes use

of the linked-read sequencing data from multiple zebra finches and the most recently published reference genome of the zebra finch to identify and characterize additional inversion polymorphisms. **Chapters 3 & 4** study the possible mechanisms that maintain the observed inversion polymorphisms.

### *Selfish genetic elements*

According to Mendel's law, the inheritance of the two alleles at a specific locus should only be a matter of fair chance (Mendel 1865). However, selfish genetic elements can manipulate the gametes to achieve a higher rate of inheritance than 50% (Werren et al. 1988). Selectively neutral selfish elements that do not affect the fitness of the organism are normally hard to detect because they can easily drive to fixation via the advantage they have during inheritance alone. In most cases, a maintained polymorphism requires the driver allele to be linked to some deleterious effect (Lindholm et al. 2016). Examples include the SD-element in *Drosophila melanogaster* where males that are homozygous for the derived SD allele are largely infertile (Sandler et al. 1959); the t-complex in mice where t-allele homozygous individuals die during development or are sterile (Lyon 1986; Safronova and Chubykin 2013); the selfish R2d2 in mice where female carriers have reduced litter size (Didion et al. 2016).

To address the genetic basis of variation in infertility and embryo mortality, in this dissertation, I explore the genetic architecture of individual survival and reproductive performance (**Chapters 1, 3 & 4**), along with the mechanisms that maintain some of the genetic polymorphisms (**Chapters 3 & 4**) in the zebra finch.



**Figure. 2** Variation in clutch size, rates of infertility and embryo mortality in relation to the age of zebra finches used in this dissertation. Sizes of the dots represent the number of clutches.

### The zebra finch as a model species – state of the art

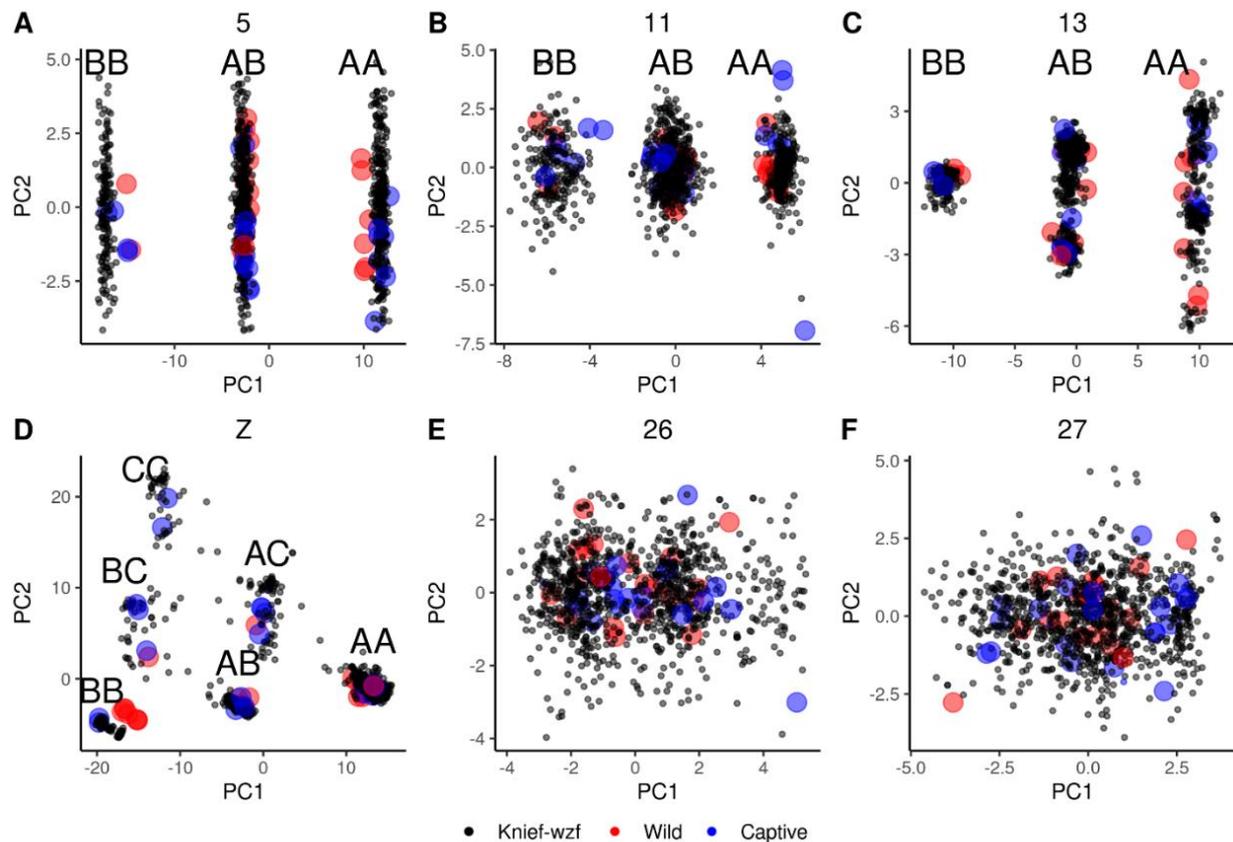
The Australian zebra finch *Taeniopygia guttata castanotis* is a small estrildid finch that commonly inhabits the arid areas in Australia. It has been domesticated and commonly used for neuroscience and behavioral science research since the early 19<sup>th</sup> century, due to the fact that they are opportunistic breeders and easy to breed in captivity (Zann 1996). Many laboratories, including my lab during my PhD, maintain them in large numbers and monitor their breeding performance over generations, providing a great system for studying detailed fitness-related traits.

The zebra finch is also a favored model species in genomics research. The zebra finch genome was the second avian genome sequenced after chicken (Warren et al. 2010), consisting of 14 macrochromosomes *Tgu1-Tgu12*, *Tgu1A* and *TguZ* (20-152 Mb) and 26 microchromosomes (1-20 Mb). Subsequently, the zebra finch has been studied intensively for its genetic variation and molecular genetics. Many wild and domesticated individuals have been genotyped or sequenced using various technologies (Frankl-Vilches et al. 2015; Singhal et al. 2015; Knief et al. 2016, 2017; Kim et al. 2017; Kinsella et al. 2019; Rhie et al. 2021).

The rich phenotypic and genomic information makes the zebra finch the perfect system to study the potential genetic basis of fitness-related traits. Otherwise, generating the same amount of knowledge for other non-model species would require a tremendous amount of effort from multiple fields.

### *Inversion polymorphisms*

The zebra finch has at least four large inversion polymorphisms on chromosomes *Tgu5*, *Tgu11*, *Tgu13* and *TguZ* that all segregate at about 0.5 in allele frequency both in the wild and in captivity (Itoh and Arnold 2005; Itoh et al. 2011; Knief et al. 2016) (**Fig. 3A-D**). Additionally, two microchromosomes, *Tgu26* and *Tgu27*, have shown weak signals of linkage possibly suggesting segregating inversion polymorphisms (**Fig. 3E-F**). Yet the latter two were dropped from further analysis due to the insufficient number of SNPs (Knief et al. 2016). In the initial study from our laboratory, it was unclear how the four large inversions could be kept polymorphic (Knief et al. 2016). It was later found that the inversion on the sex-chromosome *TguZ* had large phenotypic effects on the characteristics of sperm (Kim et al. 2017; Knief et al. 2017, 2019). A subsequent analysis of the effects of the *TguZ* inversion types on male infertility showed that heterozygous males had higher siring success, indicating that the *TguZ* inversion polymorphism may be maintained via heterosis effects on male fitness (Kim et al. 2017; Knief et al. 2017). However, despite large effects on sperm morphology, the effect sizes on male infertility and siring success were comparatively small (Kim et al. 2017; Knief et al. 2017).



**Figure 3.** Principal component analysis based on the 4553 randomly selected SNPs (genome-wide) genotypes from 948 wild zebra finches (small black dots) used (Knief et al. 2016) for chromosomes that contain inversion polymorphisms, i.e. *Tgu5*, *Tgu11*, *Tgu13*, *TguZ* (A-D) and for the two microchromosomes *Tgu26* and *Tgu27* that show weak signals of segregating inversions (E-F). Red and blue dots indicate wild and captive birds, respectively, used in **Chapter 3**. The PC scores of the 43 zebra finches with whole-genome sequencing data (red and blue) were predicted by the loadings of the 4553 SNPs from the 948 wild birds. Black letters A, B and C indicate the inversion types of the major, minor and the least abundant alleles defined in (Knief et al. 2016).

### Meiotic driver

In our captive population of zebra finches, a weak transmission distorter on *Tgu2* was found to be acting on both male and female birds (Knief et al. 2015), where the major allele was the driving allele with a higher inheritance ratio of 0.567 comparing to the losing allele (0.433). The authors ruled out postzygotic viability selection and biased gene conversion to be the cause of the drive. For example, within individuals that are heterozygous for the meiotic driver, one would expect to see the under-transmitted allele in the form of (apparently) infertile eggs or dead embryos. However, there was no apparent higher number of the losing allele in the dead embryos, and the carriers of the meiotic driver didn't lay more apparently infertile eggs (Knief et al. 2015). Consequently, the authors suggested that

the transmission distortion happens prezygotically (Knief et al. 2015). Yet, it is still unclear how could this drive be stably maintained in the population.

### *The songbird germline-restricted chromosome*

Ever since its first discovery (Pigozzi and Solari 1998), the germline-restricted chromosome (GRC) was thought to be a zebra-finch-specific oddity that was only present in the germline cells but absent from all somatic tissues. It was mostly observed as two copies in the female oocytes but as only a single copy in male spermatogonia, which was later expelled during spermatogenesis. The GRC was hence thought to be strictly maternally inherited, like the mitochondria DNA (Pigozzi and Solari 1998; Itoh et al. 2009; Goday and Pigozzi 2010; Schoenmakers et al. 2010). Recently when studying the Bengalese finch (Del Priore and Pigozzi 2014) and a wider sample of other bird species (Torgasheva et al. 2019; Malinovskaya et al. 2020), it became clear that the GRC is widespread in all songbirds (at least in those examined to date). Recent genetic studies showed that the zebra finch *T. g. castanotis* GRC is evolving rapidly (Biederman et al. 2018; Kinsella et al. 2019), and contains genes that are expressed in the ovary and testis (Kinsella et al. 2019). Cytogenetic analysis using a whole-GRC probe found that the GRC shows little homology across species (Torgasheva et al. 2019). All these findings suggest that the GRC could have a fundamental biological function that makes it indispensable in songbirds, presumably in germline and/or early embryo development. Yet, the lack of knowledge renders the discussion on the GRC evolution speculative. It calls for systematic research on, for instance, the strictness of the maternal inheritance, the intraspecific genetic variation and ultimately the fitness consequences behind any existing variation.

### **Aims**

In this dissertation, I aimed to explore the genetic basis of reproductive performance, particularly infertility and embryo mortality, using the captive zebra finch as my model. To do so, I compiled a comprehensive database of all fitness-related traits that can readily be measured in captivity (**Chapter 1**), including individual longevity, female fecundity, male infertility, embryo mortality, nestling mortality and the number of seasonal recruits produced by male and female zebra finches. This database was based on the fate of >23,000 eggs that were collected over 14 years of breeding experiments. Then I mainly work with this database to study the proximate and genetic causes of reproductive failure (**Chapters 1 - 4**). Specially, **Chapters 1 & 2** try to describe and study the observed variation at the phenotypic level and at the additive genetic level, **Chapters 3 & 4** study the effects of

specific genes (i.e., inversions), and **Chapter 5** explores the inheritance pattern and the intraspecific genetic variation of a rather unusual genetic element – the songbird germline-restricted chromosome.

Before diving into the search of genetic causes, in **Chapter 1** I first test and compare the effects of possible proximate causes of reproductive failure including aging, inbreeding and early growth conditions, as well as other confounding factors such as the laying and hatching order of the offspring. Then, as the first step to understand the genetic architecture, I use a quantitative genetic approach to study the individual repeatability and heritability of reproductive performance. If the variation in fitness-related traits has an additive genetic basis, one can study how traits are correlated at the additive genetic level. Given that selection constantly removes additive genetic variation in fitness, the additive genetic variation could be maintained via antagonistic pleiotropic effects on different fitness components (i.e. reflected by negative covariance between traits at the additive genetic level). In sum, **Chapter 1** describes the variation in zebra finch reproductive failure and studies the overall additive genetic basis for fitness-related traits.

A common concern of the general public is that negative experiences made by ancestors may influence the performance of subsequent generations, e.g. inheritance via epigenetic marks. **Chapter 2** uses the compiled comprehensive database of zebra finch reproductive performance to test how individual performance is influenced by possible transgenerational effects of early developmental stress. **Chapter 2** summarizes and compares the effects of about 1,000 combinations of 23 performance traits and six potential early developmental stressors from seven different sources: the individual itself, the two parents and the four grandparents.

After ruling out additive genetic effects and developmental conditions as the main causes of reproductive failure (**Chapters 1 & 2**), I focus on the genotypic effects of candidate genetic elements on reproductive performance. Here I chose to study inversion polymorphisms (known as supergenes) that are maintained at intermediate allele frequencies. Because inversions may capture hundreds of genes that diverged overtime, it is plausible to assume that the maintenance of large inversion polymorphisms is due to balanced effects on different fitness-related traits. Zebra finches have at least four large inversion polymorphisms that are mostly on macrochromosomes, but only the one on *TguZ* was found to show significant heterotic effects. The zebra finch microchromosomes (especially those that are smaller than 10 Mb) were often ignored in studies due to a range of methodological difficulties. Therefore, in my **Chapter 3**, I scan the zebra finch genome for additional inversion polymorphisms, especially focusing on microchromosomes, using whole-genome sequencing data of multiple wild and

captive zebra finches. I use the linked-read sequencing data to characterize the inversion breakpoints in the most recent zebra finch genome assembly. To study the evolutionary history of an inversion, it is important to differentiate the ancestral and the derived inversion types. Here I combine genetic variation and phylogenetic analyses using information from SNPs for all sequenced individuals. My expectation is that individuals that are homozygous for the ancestral type should contain higher genetic variation and cluster closer to outgroup when compared to the individuals that are homozygous for the derived inversion type.

**Chapters 3 & 4** study the evolutionary mechanisms (i.e. the fitness consequences) of the detected inversion polymorphisms. I used SNPs to tag the inversion types, and genotyped all individuals whose fitness-related traits are known (**Chapter 1**). Then I tested the effects of inversion types on the fitness-related traits of the individual. To draw general conclusions about the genotypic effects of the inversion on fitness components, I meta-summarized the estimated effect sizes across different traits. Besides the simplest heterosis scenario that maintains the inversion polymorphisms, the initial spread and the maintenance of the inversion types are less obvious, e.g. in the case of antagonistic pleiotropy. For this, I simulated the change in allele frequency over time using the estimated effect sizes of the genotypes on different fitness-related traits that cover the full zebra finch life cycle (**Chapter 4**).

After exploring proximate (**Chapters 1 & 2**) and potential autosomal causes (**Chapters 1, 3 & 4**) of reproductive failure, none of the identified factors explained a large amount of variation in infertility and embryo mortality, suggesting that there still must be other reasons for the observed high rates of reproductive failure in this species. Hence, the germline-restricted chromosome (GRC) became my next candidate to search for the genetic causes of reproductive failure in the zebra finch. Currently there is only limited knowledge on this particular chromosome. In the beginning of my **Chapter 5**, I found unexpectedly that the GRC presented in some sperm heads. This unexpected discovery violates the commonly believed strict matrilineal inheritance, which could favor the evolution of selfish DNA. Therefore, my **Chapter 5** aims at exploring the inheritance pattern and intra-specific genetic variation of the GRC. To test for the inheritance pattern, ideally one would like to trace the GRC haplotypes through the pedigree to look for mismatches of GRC haplotype and markers of matriline (e.g. mitochondrial haplotype). However, this is currently impossible, because there is no knowledge regarding the within-population variation of GRC haplotypes, and genotyping haplotypes would require killing of the individual. Therefore, I first tested how efficient the GRC elimination is in males, combining both cytogenetic and bioinformatics approaches. Specifically, I used a GRC-specific probe to label GRC-specific DNA in the ejaculates using fluorescent in situ hybridization (FISH), and

sequenced multiple natural ejaculates from multiple males to quantify the amount of GRCs in the ejaculate. To study the intra-specific variation of the GRC, I first clarified the matriline of all founder females by sequencing four amplicons that cover the whole mitochondrial genome. Then I sampled nine most-common matriline from five different populations. Lastly, I assembled and analyzed the nine GRC haplotypes using two GRC-linked genes that are in single copy.

Taken together, my dissertation explored a wide-range of intrinsic and genetic causes for reproductive failure and the possible mechanism that maintains the genetic variation. I hope my study may shed light on the link between genetic variation and variation in reproductive performance in animals beyond zebra finches.

## References

- Amos, W., J. Worthington Wilmer, K. Fullard, T. M. Burg, J. P. Croxall, D. Bloch, and T. Coulson. 2001. The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268:2021–2027.
- Biederman, M. K., M. M. Nelson, K. C. Asalone, A. L. Pedersen, C. J. Saldanha, and J. R. Bracht. 2018. Discovery of the first germline-restricted gene by subtractive transcriptomic analysis in the zebra finch, *Taeniopygia guttata*. *Current Biology* 28:1620–1627.e5.
- Charlesworth, D., and J. H. Willis. 2009. The genetics of inbreeding depression. *Nature Reviews Genetics* 10:783–796.
- Charmantier, A., G. Sorci, C. Teplitsky, A. Robert, G. Leveque, Y. Hingrat, S. Chantepie, et al. 2015. Quantitative Genetics of the Aging of Reproductive Traits in the Houbara Bustard. *Plos One* 10:e0133140.
- Conrad, D. F., and M. E. Hurles. 2007. The population genetics of structural variation. *Nature gene* 39:S30–36.
- Cote, J., A. Dreiss, and J. Clobert. 2008. Social personality trait and fitness. *Proceedings of the Royal Society B: Biological Sciences* 275:2851–2858.
- Del Priore, L., and M. I. Pigozzi. 2014. Histone modifications related to chromosome silencing and elimination during male meiosis in Bengalese finch. *Chromosoma* 123:293–302.
- Didion, J. P., A. P. Morgan, L. Yadgary, T. A. Bell, R. C. McMullan, L. O. De Solorzano, J. Britton-Davidian, et al. 2016. R2d2 Drives Selfish Sweeps in the House Mouse. *Molecular Biology and Evolution* 33:1381–1395.
- Dingemanse, N. J., A. J. N. Kazem, D. Réale, and J. Wright. 2010. Behavioural reaction norms: animal personality meets individual plasticity. *Trends in Ecology and Evolution* 25:81–89.
- Dingemanse, N. J., and D. Réale. 2005. Natural selection and animal personality. *Behaviour* 142:1159–1184.
- Drummond, H., and S. Ancona. 2015. Observational field studies reveal wild birds responding to early-life stresses with resilience, plasticity, and intergenerational effects. *The Auk* 132:563–576.
- Forstmeier, W., H. Schielzeth, J. C. Mueller, H. Ellegren, and B. Kempenaers. 2012. Heterozygosity–fitness correlations in zebra finches: Microsatellite markers can be better than their reputation. *Molecular Ecology* 21:3237–3249.
- Frankl-Vilches, C., H. Kuhl, M. Werber, S. Klages, M. Kerick, A. Bakker, E. H. de Oliveira, et al. 2015. Using the canary genome to decipher the evolution of hormone-sensitive gene regulation in seasonal singing birds. *Genome biology* 16:19.
- Goday, C., and M. I. Pigozzi. 2010. Heterochromatin and histone modifications in the germline-restricted chromosome of the zebra finch undergoing elimination during spermatogenesis. *Chromosoma* 119:325–336.
- Harely, D. 1990. Aging and reproductive performance in langur monkeys (*Presbytis entellus*). *American Journal of Physical Anthropology* 83:253–261.
- Hemmings, N. L., J. Slate, and T. R. Birkhead. 2012. Inbreeding causes early death in a passerine bird. *Nature Communications* 3:1–4.
- Hill, W. G. 2010. Understanding and using quantitative genetic variation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365:73–85.
- Hoffmann, A. A., C. M. Sgrò, and A. R. Weeks. 2004. Chromosomal inversion polymorphisms and adaptation. *Trends in Ecology and Evolution* 19:482–488.
- Itoh, Y., and A. P. Arnold. 2005. Chromosomal polymorphism and comparative painting analysis in the zebra finch. Pages 47–56 in *Chromosome Research* (Vol. 13).
- Itoh, Y., K. Kampf, C. N. Balakrishnan, and A. P. Arnold. 2011. Karyotypic polymorphism of the zebra finch Z chromosome. *Chromosoma* 120:255–264.
- Itoh, Y., K. Kampf, M. I. Pigozzi, and A. P. Arnold. 2009. Molecular cloning and characterization of the germline-restricted chromosome sequence in the zebra finch. *Chromosoma* 118:527–536.

- Jarvis, G. E. 2017. Early embryo mortality in natural human reproduction: What the data say. *F1000Research* 5:2765.
- Jay, P., A. Whibley, L. Frézal, M. Á. Rodríguez de Cara, R. W. Nowell, J. Mallet, K. K. Dasmahapatra, et al. 2018. Supergene Evolution Triggered by the Introgression of a Chromosomal Inversion. *Current Biology* 28:1839-1845.e3.
- Kim, K.-W., C. Bennison, N. Hemmings, L. Brookes, L. L. Hurley, S. C. Griffith, T. Burke, et al. 2017. A sex-linked supergene controls sperm morphology and swimming speed in a songbird. *Nature Ecology & Evolution* 1:1168–1176.
- Kinsella, C. M., F. J. Ruiz-Ruano, A.-M. Dion-Côté, A. J. Charles, T. I. Gossmann, J. Cabrero, D. Kappei, et al. 2019. Programmed DNA elimination of germline development genes in songbirds. *Nature Communications* 10:5468.
- Knief, U., W. Forstmeier, B. Kempnaers, and J. B. W. Wolf. 2019. A sex chromosome inversion is associated with copy number variation of mitochondrial DNA in zebra finch sperm. *bioRxiv*.
- Knief, U., W. Forstmeier, Y. Pei, M. Ihle, D. Wang, K. Martin, P. Opatová, et al. 2017. A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nature Ecology & Evolution* 1:1177–1184.
- Knief, U., G. Hemmrich-Stanisak, M. Wittig, A. Franke, S. C. Griffith, B. Kempnaers, and W. Forstmeier. 2016. Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biology* 17:199.
- Knief, U., H. Schielzeth, H. Ellegren, B. Kempnaers, and W. Forstmeier. 2015. A prezygotic transmission distorter acting equally in female and male zebra finches *Taeniopygia guttata*. *Molecular Ecology* 24:3846–3859.
- Kosova, G., M. Abney, and C. Ober. 2010. Heritability of reproductive fitness traits in a human population. *Proceedings of the National Academy of Sciences of the United States of America* 107:1772–1778.
- Kraft, F.-L. O. H., S. C. Driscoll, K. L. Buchanan, and O. L. Crino. 2019. Developmental stress reduces body condition across avian life-history stages: A comparison of quantitative magnetic resonance data and condition indices. *General and Comparative Endocrinology* 272:33–41.
- Küpper, C., M. Stocks, J. E. Risse, N. Dos Remedios, L. L. Farrell, S. B. McRae, T. C. Morgan, et al. 2015. A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics* 48:79–83.
- Lamichhaney, S., G. Fan, F. Widemo, U. Gunnarsson, D. S. Thalmann, M. P. Hoepfner, S. Kerje, et al. 2015. Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nature Genetics* 48:84–88.
- Le Poul, Y., A. Whibley, M. Chouteau, F. Prunier, V. Llaurens, and M. Joron. 2014. Evolution of dominance mechanisms at a butterfly mimicry supergene. *Nature Communications* 5:1–8.
- Lecomte, V. J., G. Sorci, S. Cornet, A. Jaeger, B. Faivre, E. Arnoux, M. Gaillard, et al. 2010. Patterns of aging in the long-lived wandering albatross. *Proceedings of the National Academy of Sciences* 107:6370–6375.
- Lee, Y. W., L. Fishman, J. K. Kelly, and J. H. Willis. 2016. A segregating inversion generates fitness variation in yellow monkeyflower (*Mimulus guttatus*). *Genetics* 202:1473–1484.
- Lindholm, A. K., K. A. Dyer, R. C. Firman, L. Fishman, W. Forstmeier, L. Holman, H. Johannesson, et al. 2016. The Ecology and Evolutionary Dynamics of Meiotic Drive. *Trends in Ecology and Evolution* 31:315–326.
- Lowry, D. B., and J. H. Willis. 2010. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology* 8.
- Lyon, M. F. 1986. Male sterility of the mouse t-complex is due to homozygosity of the distorter genes. *Cell* 44:357–363.
- Makunin, A. I., S. A. Romanenko, V. R. Beklemisheva, P. L. Perelman, A. S. Druzhkova, K. O. Petrova, D. Y. Prokopov, et al. 2018. Sequencing of supernumerary chromosomes of red fox and raccoon dog confirms a non-random gene acquisition by B chromosomes. *Genes* 9:1–14.

- Malinovskaya, L. P., K. S. Zadesenets, T. V. Karamysheva, E. A. Akberdina, E. A. Kizilova, M. V. Romanenko, E. P. Shnaider, et al. 2020. Germline-restricted chromosome (GRC) in the sand martin and the pale martin (Hirundinidae, Aves): synapsis, recombination and copy number variation. *Scientific Reports* 10:1–10.
- Mendel, G. 1865. Experiments in Plant Hybridisation, trans. Castle, WE, *Genetics and Eugenics* 281–321.
- Merilä, J., and B. C. Sheldon. 2000. Lifetime Reproductive Success and Heritability in Nature. *The American Naturalist* 155:301–310.
- Mérot, C., E. L. Berdan, H. Cayuela, H. Djambazian, A.-L. Ferchaud, M. Laporte, E. Normandeau, et al. 2021. Locally Adaptive Inversions Modulate Genetic Variation at Different Geographic Scales in a Seaweed Fly. (D. Charlesworth, ed.) *Molecular Biology and Evolution*.
- Mérot, C., V. Llaurens, E. Normandeau, L. Bernatchez, and M. Wellenreuther. 2020. Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. *Nature Communications* 11:670.
- Miyamoto, T., A. Tsujimura, Y. Miyagawa, E. Koh, M. Namiki, and K. Sengoku. 2012. Male infertility and its causes in human. *Advances in Urology* 2012:1–7.
- Monaghan, P. 2008. Early growth conditions, phenotypic development and environmental change. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363:1635–1645.
- Pigozzi, M. I., and A. J. Solari. 1998. Germ cell restriction and regular transmission of an accessory chromosome that mimics a sex body in the zebra finch, *Taeniopygia guttata*. *Chromosome Research* 6:105–113.
- Pracana, R., I. Levantis, C. Martínez-Ruiz, E. Stolle, A. Priyam, and Y. Wurm. 2017. Fire ant social chromosomes: Differences in number, sequence and expression of odorant binding proteins. *Evolution Letters*.
- Rhie, A., S. A. McCarthy, O. Fedrigo, J. Damas, G. Formenti, S. Koren, M. Uliano-Silva, et al. 2021. Towards complete and error-free genome assemblies of all vertebrate species. *Nature* 592:737–746.
- Safronova, L. D., and V. L. Chubykin. 2013. Meiotic drive in mice carrying t-complex in their genome. *Russian Journal of Genetics* 49:885–897.
- Sandler, L., Y. Hiraizumi, and I. Sandler. 1959. Meiotic drive in natural populations of *Drosophila melanogaster*. I. the cytogenetic basis of segregation-distortion. *Genetics* 44:233–50.
- Schielzeth, H., E. Bolund, B. Kempnaers, and W. Forstmeier. 2011. Quantitative genetics and fitness consequences of neophilia in zebra finches. *Behavioral Ecology* 22:126–134.
- Schluter, D., T. D. Price, and L. Rowe. 1991. Conflicting selection pressures and life history trade-offs. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 246:11–17.
- Schoenmakers, S., E. Wassenaar, J. S. E. Laven, J. A. Grootegoed, and W. M. Baarends. 2010. Meiotic silencing and fragmentation of the male germline restricted chromosome in zebra finch. *Chromosoma* 119:311–324.
- Singhal, S., E. M. Leffler, K. Sannareddy, I. Turner, O. Venn, D. M. Hooper, A. I. Strand, et al. 2015. Stable recombination hotspots in birds. *Science* 350:928–932.
- Speroff, L. 1994. The effect of aging on fertility. *Current opinion in obstetrics & gynecology* 6:115–20.
- Stearns, S. C. 1989. Trade-Offs in Life-History Evolution. *Function Ecology* 3:259–268.
- Tan, C. K. W., T. Pizzari, and S. Wigby. 2013. Parental age, gametic age, and inbreeding interact to modulate offspring viability in *Drosophila melanogaster*. *Evolution* 67:3043–3051.
- Torgasheva, A. A., L. P. Malinovskaya, K. S. Zadesenets, T. V. Karamysheva, E. A. Kizilova, E. A. Akberdina, I. E. Pristiyazhnyuk, et al. 2019. Germline-restricted chromosome (GRC) is widespread among songbirds. *Proceedings of the National Academy of Sciences* 116:201817373.
- Tschirren, B., A. N. Rutstein, E. Postma, M. Mariette, and S. C. Griffith. 2009. Short- and long-term consequences of early developmental conditions: A case study on wild and domesticated zebra finches. *Journal of Evolutionary Biology* 128:103–115.
- Tuttle, E. M., A. O. Bergland, M. L. Korody, M. S. Brewer, D. J. Newhouse, P. Minx, M. Stager, et al. 2016. Divergence and functional degradation of a sex chromosome-like supergene. *Current Biology* 26:344–350.

Warren, W. C., D. F. Clayton, H. Ellegren, A. P. Arnold, L. W. Hillier, A. Künstner, S. Searle, et al. 2010. The genome of a songbird. *Nature* 464:757–762.

Wellenreuther, M., and L. Bernatchez. 2018. Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends in Ecology and Evolution* 33:427–440.

Werren, J. H., U. Nur, and C. I. Wu. 1988. Selfish genetic elements. *Trends in Ecology and Evolution* 3:297–302.

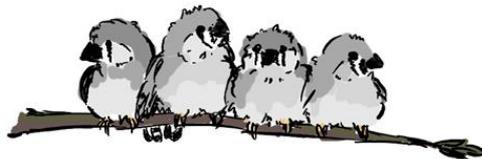
Zann, R. 1996. *The Zebra Finch: A synthesis of field and laboratory Studies*. Oxford University Press, New York, USA.



# Chapter 1

## Infertility & embryo mortality

"We are siblings of four!"



"I am the only child."



## Chapter 1

### **Proximate causes of infertility and embryo mortality in captive zebra finches**

Yifan Pei, Wolfgang Forstmeier, Daiping Wang, Katrin Martin, Joanna Rutkowska, Bart Kempenaers

Some species show high rates of reproductive failure, which is puzzling because natural selection works against such failure in every generation. Hatching failure is common in both captive and wild zebra finches (*Taeniopygia guttata*), yet little is known about its proximate causes. Here we analyze data on reproductive performance (fate of >23,000 eggs) based on up to 14 years of breeding of four captive zebra finch populations. We find that virtually all aspects of reproductive performance are negatively affected by inbreeding (mean  $r = -0.117$ ), by an early-starting, age-related decline (mean  $r = -0.132$ ), and by poor early-life nutrition (mean  $r = -0.058$ ). However, these effects together explain only about 3% of the variance in infertility, offspring mortality, fecundity and fitness. In contrast, individual repeatability of different fitness components varied between 15% and 50%. As expected, we found relatively low heritability in fitness components (median: 7% of phenotypic, and 29% of individually repeatable variation). Yet, some of the heritable variation in fitness appears to be maintained by antagonistic pleiotropy (negative genetic correlations) between male fitness traits and female and offspring fitness traits. The large amount of unexplained variation suggests a potentially important role of local dominance and epistasis, including the possibility of segregating genetic incompatibilities.

#### Published as:

Pei, Y., Forstmeier, W., Wang, D., Martin, K., Rutkowska, J., & Kempenaers, B. (2020). Proximate causes of infertility and embryo mortality in captive zebra finches. *The American Naturalist*, 196(5), 577-596.

# Proximate Causes of Infertility and Embryo Mortality in Captive Zebra Finches

Yifan Pei (裴一凡),<sup>1</sup> Wolfgang Forstmeier,<sup>1,\*</sup> Daiping Wang (王代平),<sup>1</sup> Katrin Martin,<sup>1</sup> Joanna Rutkowska,<sup>2</sup> and Bart Kempenaers<sup>1</sup>

1. Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany;

2. Institute of Environmental Sciences, Faculty of Biology, Jagiellonian University, 30387 Krakow, Poland

Submitted November 18, 2019; Accepted May 19, 2020; Electronically published September 28, 2020

Online enhancements: supplemental PDF, Excel file.

**ABSTRACT:** Some species show high rates of reproductive failure, which is puzzling because natural selection works against such failure in every generation. Hatching failure is common in both captive and wild zebra finches (*Taeniopygia guttata*), yet little is known about its proximate causes. Here we analyze data on reproductive performance (the fate of >23,000 eggs) based on up to 14 years of breeding of four captive zebra finch populations. We find that virtually all aspects of reproductive performance are negatively affected by inbreeding (mean  $r = -0.117$ ); by an early-starting, age-related decline (mean  $r = -0.132$ ); and by poor early-life nutrition (mean  $r = -0.058$ ). However, these effects together explain only about 3% of the variance in infertility, offspring mortality, fecundity, and fitness. In contrast, individual repeatability of different fitness components varied between 15% and 50%. As expected, we found relatively low heritability in fitness components (median: 7% of phenotypic variation and 29% of individually repeatable variation). Yet some of the heritable variation in fitness appears to be maintained by antagonistic pleiotropy (negative genetic correlations) between male fitness traits and female and offspring fitness traits. The large amount of unexplained variation suggests a potentially important role of local dominance and epistasis, including the possibility of segregating genetic incompatibilities.

**Keywords:** inbreeding, senescence, early nutrition, reproductive failure, quantitative genetics, sexual antagonism.

## Introduction

Reproductive performance, including offspring survival, is subject to strong directional selection in every genera-

tion. Such strong selection works not only on individuals that live in their natural habitat but also on those that live in captivity, unless artificial selection counters it. Thus, it is puzzling that some populations (or species) have substantial difficulties with successful reproduction, shown as high rates of infertility or embryo mortality. Prominent examples of frequent reproductive failure include humans (De Braekeleer and Dao 1991; Sierra and Stephenson 2006; Miyamoto et al. 2012) and other animals both in natural environments (Lyon 1986; Grossen et al. 2012) and in captive conditions (Ayalon 1978; Bunin et al. 2008; Gwaza et al. 2016; Griffith et al. 2017). Given that selection constantly removes genetic variants that lead to poor performance, one might suspect that reproductive failure typically results from inbreeding (Briskie and Mackintosh 2004), because selection against recessive deleterious mutations is inefficient, or from environmental factors (Jurewicz et al. 2009), such as pollutants (Jackson et al. 2011). However, as explained below, the range of possible explanations is much wider.

Reproductive failure and individual survival are complex traits and hence may be influenced by multiple genetic components that can be evolutionarily stable. For instance, reproductive failure and mortality may be caused by selfish genetic elements that are self-promoting at the cost of organismal fitness (Sandler et al. 1959; Lyon 1986; Safronova and Chubykin 2013; Lindholm et al. 2016). Additive genetic variants can also be preserved under intralocus sexual antagonism, where genes that are beneficial to one sex impose detrimental effects on the other (Foerster et al. 2007; Van Doorn 2009; Innocenti and Morrow 2010). Furthermore, there might be evolutionary trade-offs between traits, such that individuals that invest more in reproduction might show lower survival rates (Stearns 1989; Schluter et al. 1991). A few recent genetic and genomic studies detected genetic variants (e.g., specific genes) involved in dominance effects or rare variants that show main effects

\* Corresponding author; email: forstmeier@orn.mpg.de.

**ORCID:** Pei, <https://orcid.org/0000-0002-2411-4454>; Forstmeier, <https://orcid.org/0000-0002-5984-8925>; Wang, <https://orcid.org/0000-0002-9045-051X>; Martin, <https://orcid.org/0000-0003-2477-8397>; Rutkowska, <https://orcid.org/0000-0003-0396-1790>; Kempenaers, <https://orcid.org/0000-0002-7505-5458>.

Am. Nat. 2020. Vol. 196, pp. 577–596. © 2020 by The University of Chicago. 0003-0147/2020/19605-59613\$15.00. All rights reserved. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits non-commercial reuse of the work with attribution. For commercial use, contact [journalpermissions@press.uchicago.edu](mailto:journalpermissions@press.uchicago.edu). DOI: 10.1086/710956

on reproductive traits (e.g., Christians et al. 2000; Safronova and Chubykin 2013; Kim et al. 2017; Knief et al. 2017). As an extreme example, a balanced system of two nonrecombining lethal alleles was identified in crested newts *Triturus cristatus*, where all embryos that are homozygous for chromosome 1 (about 50% of all embryos) die during development (Sims et al. 1984; Grossen et al. 2012).

Despite the development of new genomic tools, it remains difficult to identify and examine the genetic components that show antagonistic effects or to involve more than one locus, that is, intra- and interlocus genetic incompatibilities (Dobzhansky 1936; Fishman and Willis 2006; Johnson 2008; Eroukhmanoff et al. 2016). This difficulty is likely due to the complexity of interactions between multiple loci and between the genotype and the environment (Carrell and Aston 2011; Krausz and Riera-Escamilla 2018). If animals in captivity show high rates of reproductive failure because they are not adapted to a given artificial environment, selection can act on the standing genetic variance. This would result in a transient phase where fitness is heritable until the population is better able to cope with the new environment (e.g., as a result of behavioral and physiological adaptations to captivity). In general, the genetic basis of reproductive failure and variation in survival remains largely unclear in most species.

The zebra finch is a good model species to study how survival and reproductive performance of the two sexes are correlated at the additive genetic level. The zebra finch is a short-lived songbird that easily breeds in captivity (Zann 1996), and its reproductive performance varies extensively among individuals under controlled breeding conditions in both domesticated and recently wild-derived populations (Griffith et al. 2017; Wang et al. 2017). In the wild, the rate of hatching failure (infertile eggs and dead embryos) was estimated to be >15% (table 1). This excludes clutches that failed completely, because nest desertion cannot be ruled out as the reason of failure. In lab stocks, the average proportion of eggs remaining apparently unfertilized ranged from 17% in aviary breeding to 30%–35% in cage breeding (table 1), while average embryo mortality rates varied between 24% and 75% (table 1). Average nestling mortality rates were also high (table 1). Although some of the variation has been explained by specific treatment effects (e.g., inbreeding, force pairing, maternal stress; Hemmings et al. 2012; Ihle et al. 2015; Khan et al. 2016), the high baseline levels of infertility and embryo and nestling mortality remain largely unexplained.

To better understand this variation in reproductive performance and individual survival, we here report on a comprehensive quantitative genetic analysis of life span, fecundity, infertility, offspring mortality, and other fitness-related traits that cover most phases of reproduction for the two sexes (table 2). We quantified the effects of in-

breeding, age, and an individual's early nutritional condition on all measured aspects of reproductive performance and survival.

Wild zebra finches have a remarkably large effective population size (Balakrishnan and Edwards 2009), where inbreeding is almost completely absent (Knief et al. 2015a). In contrast, in captivity, mating between related individuals is practically inevitable in the long run (Knief et al. 2015a). The level of inbreeding typically correlates negatively with offspring survival, individual fitness, and various morphological and life-history traits (Charlesworth and Charlesworth 1987; Keller and Waller 2002), for instance, in *Drosophila* (Garcia et al. 1994; Bechsgaard et al. 2013; Tan et al. 2013), in wild populations of lizards (Michaelides et al. 2016), and in mammals (Hoffman et al. 2014; Huisman et al. 2016). This is also true for captive zebra finches, whereby the estimated effect sizes of inbreeding depression vary widely among studies (Bolund et al. 2010a; Forstmeier et al. 2012; Hemmings et al. 2012).

Aging, or senescence, typically leads to a decline in reproductive function at old age, for example, in birds (Bouwhuis et al. 2009; Lecomte et al. 2010) and humans (Speroff 1994; Shirasuna and Iwata 2017). In zebra finches breeding in cages, male and female fertility declined when individuals became older (Knief et al. 2017). More generally, the relationship between age and reproductive performance is often quadratic, with an initial increase in performance due to gained experience that may mask any early-starting decline caused by deterioration of the body (Harely 1990; Bouwhuis et al. 2009; Lecomte et al. 2010).

The conditions that an individual experienced during early development may also affect fitness later in life. Such permanent environmental effects have been demonstrated using brood size manipulations, and they may affect individual behavior and reproductive investment (Gorman and Nager 2004; Tschirren et al. 2009; Rickard et al. 2010; Boersma et al. 2014). In zebra finches, being raised in enlarged broods apparently did not affect later performance (Tschirren et al. 2009). However, a nonexperimental measure of individual early-growth condition, namely, body mass measured at 8 days of age (which ranges from 2 to 12 g), had a significant but small effect on fitness later in life (Bolund et al. 2010b).

For this study, we used systematically recorded data on individual body mass at 8 days of age and on reproductive parameters and survival for four captive populations of zebra finches with an error-free pedigree. The aims of this study were (1) to estimate and compare the effect sizes of inbreeding, early nutritional condition, and age on reproductive performance traits; (2) to estimate the relative importance of individual and pair identity (i.e., repeatability) on reproductive performance; (3) to quantify the heritability of individual reproductive performance;

**Table 1:** Summary of rates of hatching failure, infertility, and embryo and offspring mortality reported in the literature on zebra finches

Population	Sample description	Hatching failure (%)	Infertility (%)	Embryo mortality (%)	Nestling mortality (%)	Reference
Wild	1,156 eggs; clutches that produced no nestlings were removed	>17	...	...	...	Zann 1996
Wild	872 eggs; clutches that produced no nestlings were removed	16	...	...	9	Griffith et al. 2008
La Trobe University, Australia, domesticated	31 untreated and 25 CORT-treated pairs; clutches that produced no nestlings and all first eggs were removed	Untreated: 24; treated: 45	Untreated: 7; treated: 15	Untreated: 10; treated: 29	...	Khan et al. 2016
Max Planck Institute for Ornithology, Germany, domesticated (from Sheffield, UK)	11,617 eggs	...	30	...	...	Knief et al. 2015b
Max Planck Institute for Ornithology, Germany, recently wild derived (from Bielefeld, Germany)	852 eggs; aviary	...	17	24	45	Ihle et al. 2015
Sheffield University, UK	161 eggs for infertility; 2,884 eggs for hatching failure and nestling mortality	52	35	...	31	Kim et al. 2017
Sheffield University, UK	1,524 eggs; 77 unrelated and 20 sib-sib pairs	...	Unrelated: 9; sib-sib: 11	Unrelated: 59; sib-sib: 75	Unrelated: 55; sib-sib: 67	Hemmings et al. 2012

Note: For the population from La Trobe University, Australia, in treated pairs females were given a corticosterone (CORT) mix after laying the first egg. The CORT mix was made of 0.5 mg of crystalline corticosterone dissolved by 10  $\mu$ L of ethanol, then mixed with 990  $\mu$ L of peanut oil (Khan et al. 2016). Hatching failure indicates the proportion of eggs that do not hatch. Infertility indicates the proportion of eggs that show no sign of development. Embryo mortality indicates the proportion of fertilized eggs where the embryo died before hatching. Nestling mortality indicates the proportion of nestlings that died before fledging or independence.

and (4) to test whether some of the heritable components can be maintained by antagonistic pleiotropy, by analyzing the additive genetic correlations between reproductive performance traits and life span across the two sexes.

### Methods

Zebra finches are opportunistic breeders that are abundant throughout most of Australia. Individuals become sexually mature around the age of 90 days and then form pairs for life through mutual mate choice. Breeding pairs cooperatively incubate and raise nestlings until they reach independence around the age of 35 days (Zann 1996). Captive zebra finches live for about 4.5 years on average and maximally for 10 years (Zann 1996). The studied zebra finches originated from four populations held at the

Max Planck Institute for Ornithology, Seewiesen, Germany. Details about the population background, rearing conditions, and breeding seasons are provided in the supplemental material (tables S1, S2; tables S1–S11 are available online). Housing in captivity implies that birds are supplied with food ad lib., which is known to maximize their reproductive performance (Lemon and Barth 1992). In brief, we compiled and analyzed up to 14 years of zebra finch reproductive performance data from (1) population Seewiesen, a domesticated population derived from the University of Sheffield with a nine-generation-long error-free pedigree (population 18 in Forstmeier et al. 2007b); (2) population Krakow, a domesticated population that was generated by hybridizing between Krakow (population 11 in Forstmeier et al. 2007b) and Seewiesen populations; (3) population Bielefeld, which was derived from the wild in the late 1980s (population 19 in Forstmeier

**Table 2:** Description of reproductive performance traits in our zebra finch study

Trait	Fixed effects for	Random effects	BLUPs calculated for	Description
Female				
Clutch size cage	Female	Female	Female	Number of eggs consecutively laid by a single female in a cage (containing one male and one female), allowing for laying gaps of maximally 4 days between subsequent eggs; for 2% (65 of 3,694) clutches that had >7 eggs, they were counted as 7
	...	Male	...	
	...	Pair	...	
Clutch size aviary	Female	Female	Female	Number of eggs consecutively laid by a female in a communal breeding aviary, allowing for laying gaps of maximally 4 days between subsequent eggs; for 5% (173 of 3,663) clutches that had >7 eggs, they were counted as 7
Fecundity aviary	Female	Female	Female	Total number of eggs laid by a female in a communal breeding aviary over the course of a breeding season (35–83 days), where no offspring rearing was allowed
Seasonal recruits	Female	Female	Female	Total number of genetic offspring that survived to independence in a communal breeding aviary, i.e., age 35 days, within a breeding season (83–113 days for egg laying plus about 50 days for rearing)
Male				
Fertility cage	Female	Female	...	Whether an egg was fertilized by the male in the cage (containing one male and one female)
	Male	Male	Male	
	...	Pair	...	
	Egg	...	...	
Within-pair paternity	Female	Female	...	Whether an egg laid by the social partner of the male in a communal breeding aviary was fertilized by the male (infertile eggs and extrapair fertilizations count as failed within-pair paternity)
	Male	Male	Male	
	...	Pair	...	
Siring success	Male	Male	Male	Total number of eggs fertilized by a male in a communal breeding aviary over the course of a breeding season (35–113 days)
Seasonal recruits	Male	Male	Male	Total number of genetic offspring that survived to independence in a communal breeding aviary, i.e., age 35 days, within a breeding season (83–113 days for egg laying plus about 50 days for rearing)
Offspring				
Embryo survival	Female	Female	Female	Whether a fertilized egg that was incubated by an individual in a cage (containing one male and one female) or a communal breeding aviary hatched
	Male	Male	...	
	...	Pair	...	
	Embryo	...	...	
Nestling survival	Female	Female	Female	Whether a nestling that hatched in a cage (containing one male and one female) or a communal breeding aviary survived to independence, i.e., age 35 days
	Male	Male	Male	
	...	Pair	...	
	Nestling	...	...	
Individual				
Life span	Individual	...	Individual	Number of days from the date of hatching to the date of natural death; some missing values were replaced by life expectancy

Note: Traits were measured in the context of either single pairs breeding in a small cage or multiple pairs breeding communally in a large aviary. Fixed effects (focal) are inbreeding coefficient, age, and early condition (mass at day 8). Random effects (focal) are the variance components explained by female, male, or pair identity. Best linear unbiased predictions (BLUPs) are estimated from univariate models where we controlled for significant fixed and random effects. For the offspring trait of embryo survival, female, male, and pair identities refer to the genetic parents of the embryo, whereas for nestling survival, female, male, and pair identities refer to the social parents that raised the nestling. Cage dimensions, before 2012: 60 cm × 40 cm × 45 cm (length × width × height); after 2012: 120 cm × 40 cm × 45 cm. For details of housing conditions, see Bolund et al. (2007). A semioutdoor aviary measured 500 cm × 200 cm × 200 cm (length × width × height).

et al. 2007b); and (4) population Melbourne, which was derived from the wild in the early 2000s (see Jerónimo et al. 2018). All data underlying this study have been deposited in the Open Science Framework (<https://doi.org/10.17605/OSF.IO/TGSZ8>; Pei 2020).

Birds from the two recently wild-derived populations were smaller (~11 g) compared to domesticated birds (~15–16 g, because of selective breeding by aviculturists) and shier, so we bred them only in large semioutdoor aviaries (rather than in small cages; see table 2 for sizes of cage and aviary). Between 2004 and 2017, we bred zebra finches in four settings with various treatments (see tables S1 and S2 for details of breeding seasons): (1) cage breeding, (2) cage laying, (3) aviary breeding, and (4) aviary laying. In cages, single pairs were kept, and hence, partners were assigned. In aviaries, groups of birds were kept together, and individuals could freely form pairs. Group size was typically 12 but ranged from 10 to 42, with sex ratio (proportion of males) ranging from 0.4 to 0.6. In a breeding setup, pairs were allowed to rear their offspring, whereas in a laying setup, all eggs were collected for paternity assignment and replaced by plastic eggs that were removed after 7 or 10 days of incubation. The proportion of individuals that participated in more than one breeding season ranged from 0.23 to 0.84 (mean: 0.47).

In this study, we focus on general effects on reproductive performance in zebra finches, not on population-specific effects. Therefore, in all analyses, we controlled statistically only for between-population differences in reproductive performance (main effects only, no interactions; see below for model details).

#### *Measures of the Focal Fixed Effects: Inbreeding, Age, and Early Nutrition*

We used the pedigree-based inbreeding coefficient  $F_{\text{ped}}$ , calculated using the R package pedigree version 1.4 (Coster 2015), as a measure of the degree of inbreeding of an individual (Wright 1922; Knief et al. 2016b);  $F_{\text{ped}}$  reflects the proportion of an individual's genome that is expected to be identical by descent. Hence,  $F_{\text{ped}}$  can be used to estimate without bias the slope of the regression of fitness over inbreeding (Howrigan et al. 2011; Knief et al. 2016b). For instance, full-sibling mating produces inbred offspring that are expected to have 25% of the genome identical by descent ( $F_{\text{ped}} = 0.25$ ). For practical reasons, all founders were assumed to be unrelated ( $F_{\text{ped}} = 0$ ; Forstmeier et al. 2004), even though their true level of identity by descent is likely about 5% (judging from runs of homozygosity; Knief et al. 2015a).

For all birds, we recorded their exact hatch date. Thus, for models of reproductive performance at the level of

eggs, clutches, and breeding rounds (as the unit of analysis), we used the exact age (in days) of the female or the male when an egg was laid, a clutch started, or a breeding round started, respectively. At the start of reproduction, individuals were 69–2,909 days old (fig. S1; figs. S1–S9 are available online).

On the day of hatching, we individually marked all nestlings on the back using waterproof marker pens (randomly using red, blue, and green and pairwise combinations of these colors if there were more than three nestlings). We checked survival almost daily (daily on weekdays, occasionally during weekends) until offspring became independent (age 35 days). As a measure of early-growth condition, we determined body mass of each nestling to the nearest 0.1 g at 8 days of age (hereafter, condition). Despite the fact that high-quality food was available to all parents ad lib., nestling body mass at this age ranged from about 1.5 to 12.6 g (mean =  $7.1 \pm 1.7$  SD). For 297 of 6,190 nestlings, body mass was measured on day 6, 7, or 9. For those individuals, we estimated their mass on day 8 as follows. We constructed a linear mixed effects model, with nestling body mass as the dependent variable, actual age of the mass measurement and  $F_{\text{ped}}$  as two continuous covariates, and population (1–4; see above) as a fixed factor. We also included the identity of the genetic mother as a random effect. Using the slope of daily mass gain, we estimated mass at day 8 for those 297 individuals by adding or subtracting 0.97 g per day of measuring too early or too late. Because the four populations differ in body mass, we normalized (Z scaled) all measured or estimated values of mass at day 8 within each population before further analysis.

We report effects of inbreeding, age, and early condition always with a negative sign, such that negative values of greater magnitude reflect stronger detrimental effects of being inbred, old, or poorly fed. This allows us to meta-summarize the results and to directly compare the strength of the focal fixed effects on reproductive performance.

#### *Measures of Life Span and Reproductive Performance Traits*

Table 2 provides an overview of all traits included in this study. To allow direct comparison and easy interpretation of the fixed effects and additive genetic correlations, we scored all traits such that higher positive values reflect better reproductive performance.

Life span was analyzed in the following subset of birds: five generations of birds from the Seewiesen population (referred to as generations P,  $F_1$ – $F_3$ , and  $S_3$ ;  $N = 1,855$  individuals) and four generations of birds from the Bielefeld population ( $F_1$ – $F_4$ ;  $N = 1,067$  individuals). Among those

birds, we used the four most complete generations, P and F<sub>1</sub>–F<sub>3</sub> Seewiesen, for which we recorded the exact life span for all ( $N = 1,175$  individuals) as a pool to impute missing life spans. For 219 S<sub>3</sub> Seewiesen birds and for 663 Bielefeld birds, no date of natural death was available (e.g., because individuals were still alive or because their fate was unknown). For these individuals, we used imputed life expectancy in all analyses, defined as the average life span of individuals from the same pool that lived longer than the focal bird when last observed alive.

In aviaries, we identified social pairs by behavior (clumping, allopreening, and visiting a nest together). All parentage assignments were based on conventional microsatellite genotyping using 10–15 microsatellite markers on up to 13 chromosomes (Wang et al. 2017), following Forstmeier et al. (2007a). We assigned every fertilized egg to its genetic mother ( $N = 11,704$  eggs). When the egg appeared infertile (no visible embryo; Birkhead et al. 2008), we assigned it to the social female that was attending the clutch ( $N = 3,630$  cases). In 36 cases where two females used the same nest to lay eggs, we assigned the unfertilized eggs to the female that laid the most similar eggs (in size and shape), based on eggs that were certainly laid by a given female (e.g., fertilized eggs and eggs in other clutches laid by that female). In cases where birds were not allowed to rear offspring, we quantified female fecundity as the total number of eggs laid by the focal female during the breeding period (see tables S1, S2).

In breeding experiments, we opened all unhatched eggs to check for visible signs of embryo development and classified them as either infertile or embryo mortality. In experiments in which all eggs were incubated artificially for a few days to collect DNA from embryos, we classified eggs as infertile or not but discarded information on embryo viability. Visual inspection of opened eggs has the disadvantage that early embryo mortality may get misclassified as infertility if it occurred before any visible signs of development. Misclassification cannot be avoided entirely, even with more time-consuming examination of eggs, which would be challenging to do for thousands of eggs (Birkhead et al. 2008; Murray et al. 2013). However, genotyping the germinal disk and counting sperm on the perivitelline membranes of 76 freshly laid eggs revealed 22 apparently infertile eggs. Only one of those (5%) had more than 20 sperm on the perivitelline membrane, suggesting early embryo mortality (fig. S2; see also Birkhead and Fletcher 1998). In contrast, among 37 eggs with more than 20 sperm on the perivitelline membrane, 36 (97%) developed diploid tissue. Thus, we expect only a small fraction of misclassification.

In cages, we measured male fertility as a binary trait, that is, whether an egg was fertilized. Because extrapair copulations can be excluded in cages, we only genotyped

all surviving offspring with the same set of microsatellites used in aviaries as confirmation (Wang et al. 2017). In 12 cases, one to five eggs (median: one egg) were fertilized by the previous partner of the female, and those were counted as infertile eggs of the focal male. In aviaries, we assessed for each egg whether it was sired by the social male of the female who laid the egg. We refer to this as male within-pair paternity, a trait that reflects a male's ability to defend his paternity against extrapair males. We also quantified male siring success as the total number of fertilized eggs sired by a focal male. This includes males that remained unpaired (without a social female).

For each fertilized egg that was incubated by the social parents, we recorded whether it hatched (binomial trait for the genetic parents). For each hatched egg that was reared, we recorded whether the nestling survived to independence (day 35; binomial trait for the social parents). We quantified the number of seasonal recruits as the number of genetic offspring that survived to independence within a given breeding season. The number of seasonal recruits was square root transformed to approach normality.

#### *Statistical Models*

All mixed effects models were run in R version 4.0.0 (R Core Team 2020), using the R package lme4 version 1.1-23 (Bates et al. 2015). All animal models were run using VCE6 (Neumaier and Groeneveld 1998) because (a) it allows running a 12-trait multivariate animal model that consists of 2,346 individuals with at least one trait value per individual and (b) it has a reasonable running time. To check the consistency of model outputs, we repeated all animal models in the R packages pedigreemm version 0.3-3 (Vazquez et al. 2010; univariate animal models only) and MCMCglmm (Hadfield 2010; univariate and bivariate animal models). All model details, with the supporting data and R scripts, have been deposited in the Open Science Framework (<https://doi.org/10.17605/OSF.IO/TGSZ8>; Pei 2020). Model outputs of all methods are given in the supplemental Excel file (available online). The heritability and additive genetic correlation estimates were highly correlated between methods ( $r > 0.65$ ,  $P < .002$ ). We report the VCE6 estimates, unless otherwise stated. Figure S3 shows the exact range of each focal fixed effect and each performance trait value. Here, we Z transformed all covariates and response variables across populations to allow direct comparison of the effect sizes for inbreeding, age, and condition across all models. The 95% confidence intervals of fixed effects from mixed effects models were calculated using the function glht from the R package multcomp version 1.4-13 while controlling

for multiple testing (Hothorn et al. 2008). Data analysis involved four consecutive steps (fig. 1).

*Step 1: estimation of fixed effects and variance decomposition.* The goal of step 1 was to estimate (a) all fixed effects on reproductive performance and (b) individual repeatability of performance traits (fig. 1). All fixed and random effects of models used in step 1 are listed in tables S3 and S4. In brief, we first fitted all models with a Gaussian error distribution to compare and metasummarize the estimated effect sizes of the fixed effects and to estimate the variance components for the random effects. We used all observations with information on the three fixed effects (age,  $F_{\text{ped}}$ , and early condition of the male, female, and the individual egg if applicable) and included population (fixed effect) and female, male, and pair identity (random effects). We analyzed traits that were measured at either egg, clutch, or season level. As applicable, we fitted as fixed effects the laying sequence of eggs within a clutch, the order of hatching of offspring within a brood, the order of the clutches that were laid by a female over the course of a season, the sex ratio in the aviary, and the duration of the season (table S1). For models of embryo survival, we also controlled for whether the eggs were incubated in a nest that still contained offspring from a previous brood (7% of embryos). For models of nestling survival, we added as fixed effect pair type (pair formed through mate choice or through force pairing; Ihle et al. 2015). For models of egg-based fertility, within-pair paternity, and embryo and nestling survival, we also tested the effect of egg volume on egg fate (we calculated volume as  $V = (1/6) \times \pi \times \text{width}^2 \times \text{length}$ , where egg length and width had been measured to the nearest 0.1 mm). For this analysis, we fitted the mean egg volume of each female and the centered egg volumes (centered within individual females) to distinguish between the effects of between- and within-female variation in egg size (van de Pol and Wright 2009). We estimated the variance components for male, female, and pair identity and further controlled for clutch identity and identity of the setup (see tables S1, S2), as applicable, by adding them as random effects. Life span had no repeated measurement; therefore, we included only individual identity as a dummy random effect for practical reasons when running the model and extracting estimates in R. For this lm model, the correlation between the residuals and the dummy random effect equals 1, and the fixed effect estimates were unaffected by the dummy variable. Table 2 shows for which group of individuals, that is, female, male, or the offspring itself, we tested which focal fixed and random effects.

To allow direct comparison of the magnitude of fixed effects at the same level of measurement, we also aggregated data within clutches (e.g., proportion of infertile eggs within a clutch) and within individuals over the course of a season. Models on aggregated data were weighted by

the number of eggs within a clutch or by the number of eggs or clutches for an individual within a season (fig. 1). As expected, the proportion of variance explained by male, female, and pair identity increased from the egg level to the season level (see “Results”). However, the relative proportions explained by female, male, and pair identity did not change notably. Therefore, we focus on the analyses of fixed effect estimates at the breeding season level.

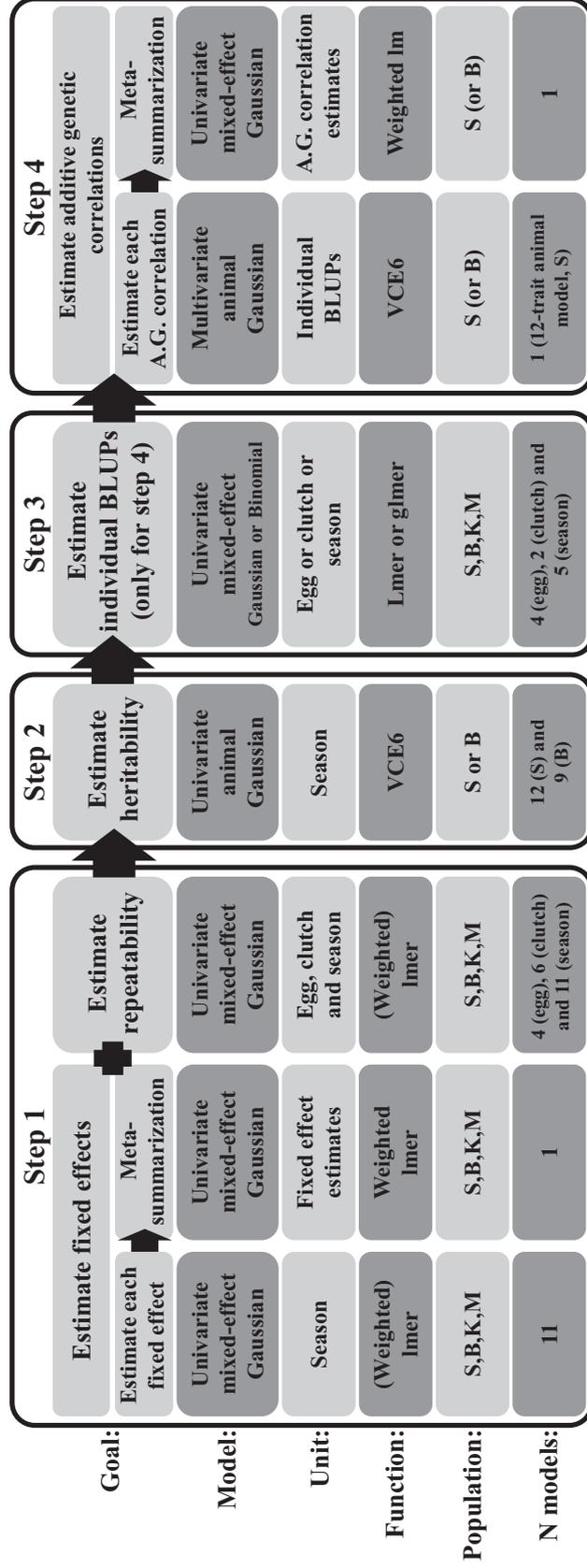
To compare the overall effect sizes between the focal fixed effects, we metasummarized the estimated effect sizes for inbreeding, age, and condition using the weighted lmer function from the R package lme4 (fig. 1, step 1, metasummarization of estimated effect sizes). The uncertainty of each estimate was accounted for by using the multiplicative inverse of the standard error (1/SE) of the response variable as weight. In this metamodel, we used effect size estimates from models that had been aggregated at the season level as the dependent variable. Note that effects of inbreeding of the egg on fertility in cage breeding and nestling survival were taken from egg-based models because they cannot be aggregated by clutch or season. Additionally, we tested whether effect sizes differed among males, females, and offspring (fixed effect with three levels) or among traits (random effect with 11 levels; as listed in table 2).

Additionally, we tested for early-starting aging effects by selecting reproductive performance data for males and females that were <2 years old when reproducing. We then metasummarized the mean age effect estimates using the R function lm, weighted by the multiplicative inverse of the standard error.

We calculated the amount of variance explained by each fixed effect (Nakagawa and Schielzeth 2010) as the sum of squares of the fixed effect divided by the number of observations ( $N - 1$ ; Henderson 1953). In weighted models, we divided the variance components of the fixed effects and the residual by the mean weight value (Bates et al. 2015).

*Step 2: estimation of heritability of fitness-related traits.* The goal of step 2 was to estimate the heritability of reproductive performance traits using univariate Gaussian animal models (fig. 1). Because quantitative genetic models require large amounts of data, we restrict our analyses to the populations Seewiesen and Bielefeld. Note that the pedigrees of our four captive populations are not connected, so it was not useful to analyze them jointly.

We kept the general model structure from step 1 but excluded the fixed effects of egg volume on male fertility, embryo, and offspring survival (to avoid removing biological variation that is potentially heritable and hence of interest; note that the effect sizes of egg volume are small; see “Results”). For the embryo survival model, we excluded the nonsignificant fixed effects of male age, inbreeding,



**Figure 1:** Steps of data analysis from univariate mixed models to multivariate animal models. Shown are the goals of the analysis, the model properties, the unit of analysis (i.e., whether rows in the data represent single eggs, clutches, individuals in a breeding season, single fixed effect estimates, or individuals overall), the software functions used for analysis (for models on aggregated levels, weight stands for the number of eggs or clutches used for each aggregation, whereas in metasummarization models, weight stands for the multiplicative inverse of the standard error of each estimate), and the population abbreviations for data used for the analysis (S = Seewiesen; B = Bielefeld; K = Krakow; M = Melbourne). Shown are the number of models conducted within each step with their specific details (e.g., unit, population, or model type) used for analysis. A.G. = additive genetic.

and condition. For the model on male fertility from cage breeding, we excluded the nonsignificant effect of the level of inbreeding of the egg itself. To most effectively use the available information on reproductive performance, we included individuals with missing values for condition ( $N = 231$  founder individuals and  $N = 23$  individuals of the  $F_2$  generation; i.e., 7% of Seewiesen birds). These missing values were replaced by the population mean. Individual identity was fitted twice, once linked to the individual correlation matrix (pedigree) to estimate the amount of variance from additive genetic effects ( $V_A$ ) and once to estimate the remaining amount of variance from permanent environmental effects ( $V_{PE}$ ; Kruuk and Hadfield 2007). Animal models on nestling mortality were run twice, once for the mother and once for the father. We calculated heritability based on the total phenotypic variance,  $V_{Ph}$ , as  $h^2 = (V_A/V_{Ph})$ , and we also quantified  $V_A$  relative to individual repeatability as  $(V_A/(V_A + V_{PE}))$ .

We compared the estimates of heritability (and  $V_A$  relative to the individual repeatability) between the domesticated population Seewiesen and the recently wild-derived population Bielefeld using the R function `lmer`. We used the multiplicative inverse of the standard error as weight to control for variation in uncertainty of each estimate. We used the estimates of heritability as the response variable and fitted population as a fixed effect (two levels) and trait as a random effect (nine levels, including only traits that were measured in both populations).

*Step 3: calculation of mean individual fitness-related trait values using best linear unbiased predictions (BLUPs).* The only goal of step 3 was to extract individual estimates of reproductive performance needed for step 4. We kept the model structure from step 1, except that we used a binomial error structure for binary traits, that is, male fertility in cages and aviaries and embryo and nestling survival. Missing values for condition (mostly founders of each population; 6% of all birds of the four populations) were replaced with population means as in step 2. For the embryo survival model, we again excluded the nonsignificant effects of male inbreeding, age, and condition. We also excluded (a) effects of egg volume from all egg-based models and (b) the effect of the level of inbreeding of the egg itself from the model of male fertility measured in cages (see step 2). We extracted the BLUPs for female or male identity (as applicable) as the estimated life-history trait value of that individual (table 2) for step 4.

*Step 4: estimation of additive genetic correlations.* The goal of step 4 was to estimate additive genetic correlations between different performance traits using multivariate animal models. Before fitting a 12-trait animal model that estimates for each matrix (genetic and residual) all 12 variances and 66 covariances simultaneously, we aggregated

the raw data to one phenotypic value per individual for each trait (fig. 1, step 3). This was necessary because we are not aware of software that can handle the full complexity of the underlying raw data (involving more than 26 different fixed effects). Because simple averages of multiple measures can result in outliers when sample size is small, we used the phenotypic BLUPs described above. BLUPs do not produce outliers and account for all considered fixed and random effects (Robinson 1991; Houslay and Wilson 2017). Breeding values (genetic BLUPs) suffer from nonindependence because the phenotype of one individual influences the breeding values of all its relatives (Hadfield et al. 2010). Note that this is not the case for the phenotypic BLUPs we use here. However, the uncertainty that is inherent to each BLUP is not taken into account, which may lead to underestimation of standard errors (Houslay and Wilson 2017). To check the robustness of our results, we compared our estimates with those obtained (a) using a smaller data set from another population (Bielefeld) with the same method and (b) using bivariate animal models in MCMCglmm version 2.29 (Hadfield 2010; population Seewiesen). The latter approach is presumably less powerful than a full 12-trait animal model.

For each of the 12 traits, we fitted an intercept and the pedigree as the only random effect to separate additive genetic from residual variance. We ran these models for the largest and most comprehensive data set (population Seewiesen;  $N = 2,346$  individuals with at least one trait value, BLUPs for 12 traits, and 66 covariances) and for the more limited data set (population Bielefeld;  $N = 1,134$  individuals, BLUPs for 9 traits, and 36 covariances; see “Results”; fig. 1, step 4, estimate additive genetic correlations).

We used the weighted `lm` function in the R package `stats` to summarize the estimated additive genetic correlations within and between the major categories of traits, that is, female, male, offspring traits, and life span, for each population separately (table 2; fig. 1, step 4, metasummarization of estimated additive genetic correlations). We fitted the estimates of additive genetic correlations (for each pair of traits, weighted by the multiplicative inverse of the standard error of each estimate) as the dependent variable, with trait class combination as a predictor with seven levels. We removed the intercept to estimate the mean additive genetic correlation for each pairwise combination of classes. We then computed the eigenvectors of the additive genetic variance-covariance matrix of traits, using the R function `eigen`, and visualized the orientation of the traits in the additive genetic variation space defined by the principle components PC1 and PC2 (Berner 2012). The proportion of variance explained by the first two principle components was calculated using the functions `summary` and `prcomp` in the R packages `base` and `stats`, respectively.

## Results

### *Effects of Laying and Hatching Order, Clutch Order, and Egg Volume on Egg and Embryo Fate*

The fate of an egg and its embryo depended on the order of laying within a clutch, the order of hatching within a brood, and the order of consecutive clutches within a breeding season (fig. S4; table S3, models at the egg level; see also fig. 1, step 1). First-laid eggs in a clutch were significantly more likely to be infertile or to contain a dead embryo. Fertility and embryo viability were the highest for the third egg (fig. S4). Male fertility significantly increased over the first three clutches and stayed high afterward. In contrast, clutch order did not affect the probability of embryo and nestling survival.

The average effect of egg volume on measures of egg fate was small (mean:  $r = 0.040 \pm 0.016$  SE; fig. S5). Effects of egg volume were largest for nestling survival after hatching and smallest for embryo survival (table S3; fig. S5). Despite large sample size ( $N = 9,785$  eggs), embryo survival was not significantly influenced by egg volume (between-female variation:  $r = 0.015 \pm 0.017$  SE,  $P = .37$ ; within-female variation:  $r = 0.018 \pm 0.010$  SE,  $P = .08$ ; table S3). Additionally, embryos in clutches that were incubated in the presence of nestlings from previous breeding attempts were more likely to die before hatching ( $b = 0.192 \pm 0.048$  SE,  $P < .0001$ ; table S3). Overall, the total amount of variance explained by laying and hatching order, clutch order, and egg volume on egg fate was less than 5% (table S4).

### *Effects of Inbreeding, Age, and Early Condition*

Individuals performed worse in virtually all studied reproductive traits when they were more inbred, as they became older, and when they weighed less at 8 days of age (figs. 2, S3, S6; table S3; see also fig. 1, step 1). Interestingly, reproductive performance did not show an initial increase at a young age (metasummarized effect size of age among birds younger than 2 years:  $r = -0.013 \pm 0.011$  SE; figs. 2C, 2F, 3, A3). Inbred eggs were equally as likely to be infertile as outbred eggs, while inbred embryos and offspring were more likely to die (fig. 3C). Together, this suggests that most infertile eggs were not cases of undetected early embryo mortality. Individuals lived shorter lives when they were inbred and when they had low weight at day 8 (fig. 3; table S3). However, the fixed effects of inbreeding, age, and condition together explained, on average, only 2% of the variance across all traits (fig. 4; table S5).

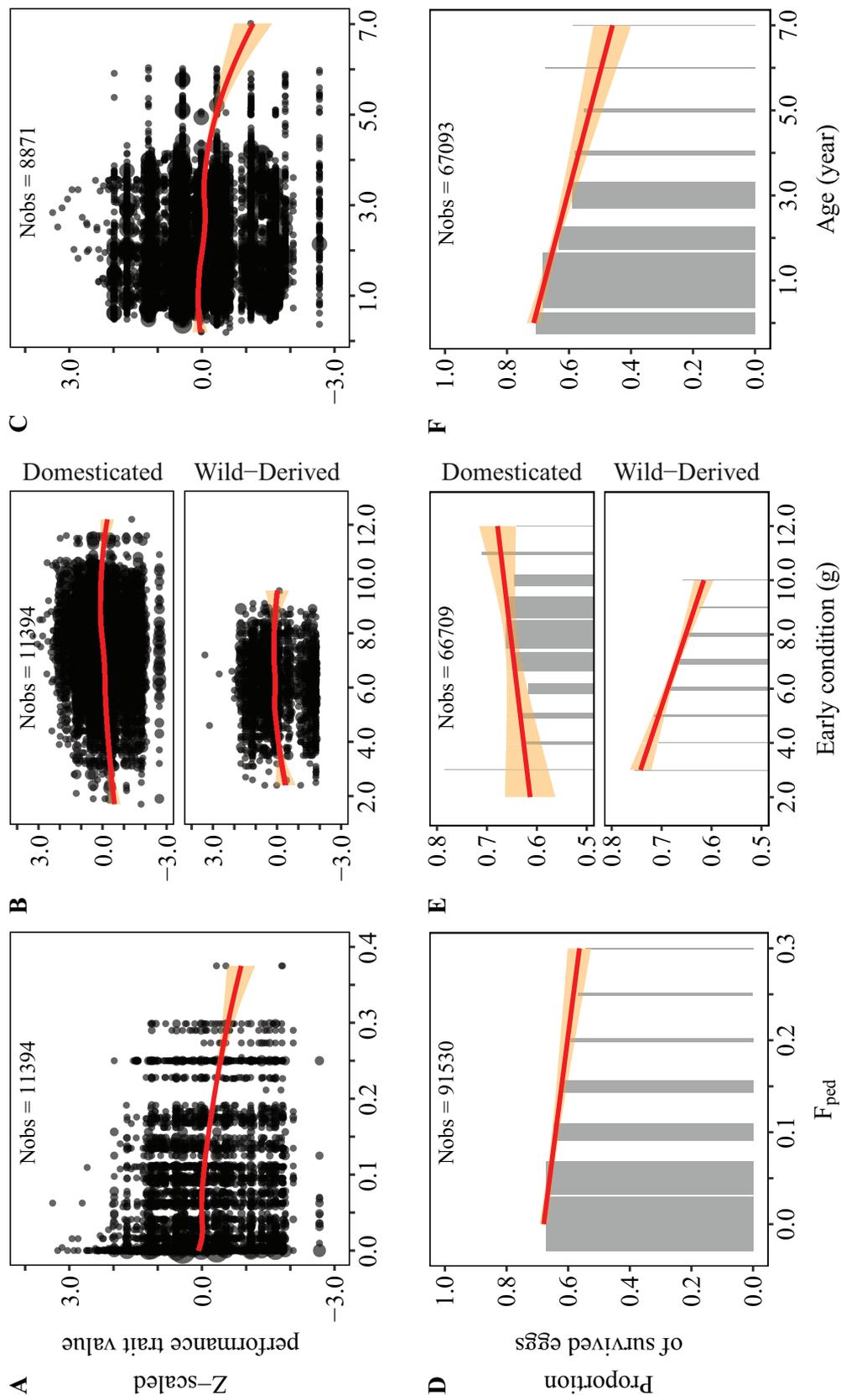
Metasummarized effect sizes of inbreeding ( $r = -0.117 \pm 0.024$  SE) and age ( $r = -0.132 \pm 0.032$  SE) were similar in magnitude and were about twice as large as the remarkably small effect of early condition ( $r =$

$-0.058 \pm 0.029$  SE; fig. 3; table S4; see also fig. 1, step 1, metasummarization of estimated fixed effects). There was no significant difference among males, females, and offspring in how strongly they were affected by these three factors ( $b \leq 0.012 \pm 0.028$  SE,  $P = .63$ ; table S4). Fitting trait (fitness component, 11 levels) as a random effect explained 1.5% of the variance in effect sizes ( $P = .02$ ; table S4), suggesting that some components might be less sensitive than others (fig. 3; table S3). Female traits significantly predicted offspring survival and male fertility (independent of whether they were measured in a cage or in an aviary), whereas male traits showed no effect on offspring survival (fig. 3).

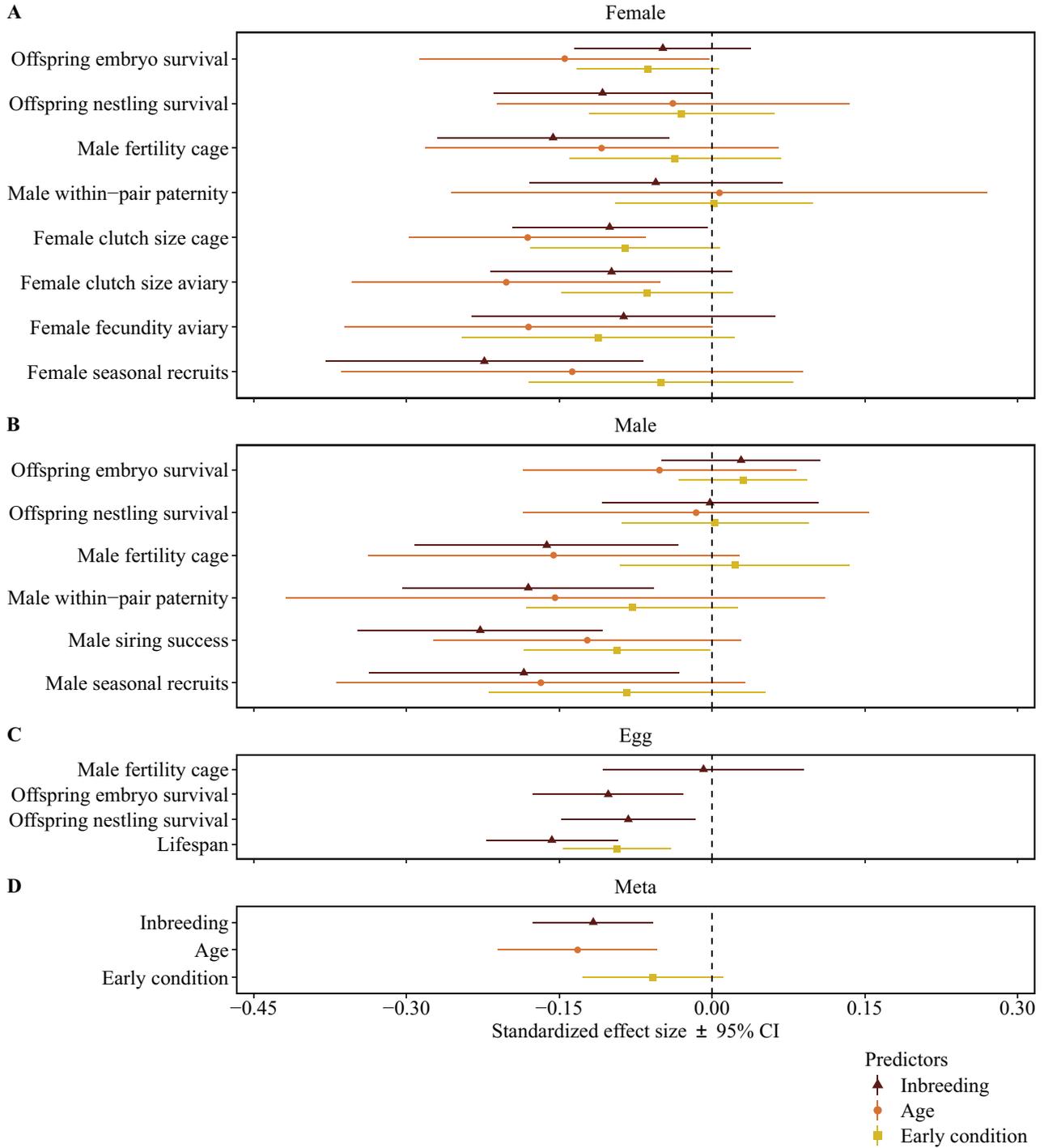
### *Variance Components and Heritability*

Variance components for all reproductive performance traits are shown in figure 4 (see also table S4; fig. 1, step 1, estimate repeatability). Overall, individual reproductive performance traits were significantly repeatable (median  $R = 0.28$ , range: 0.15–0.50). Female reproductive performance traits (clutch size, fecundity, and female seasonal recruits) showed reasonably high repeatability for individual females ( $R \sim 0.26$ –0.40). Likewise, male fertility, male siring success, and male seasonal recruits were highly repeatable for individual males ( $R \sim 0.24$ –0.50). Female reproductive traits from aviary breeding were analyzed independently of whether the focal female had a partner (table 2), but female clutch size measured in a cage showed no contribution from the male partner or from pair identity. In contrast, male fertility depended on all three random effects and was repeatable for males ( $R > 0.23$ ,  $P < .0001$ ) but less so for females ( $R < 0.18$ ,  $P < .1$ ) or for the particular pair combinations ( $R < 0.23$ ,  $P < .05$ ). The model on embryo survival showed significant female and pair identity (genetic parents) effects that were similar in size (both  $R = 0.20$ ,  $P < .0002$ ), while genetic male identity explained no variance (fig. 4). In contrast, social female ( $R = 0.17$ ,  $P = .017$ ) and social male ( $R = 0.15$ ,  $P = .039$ ) identity explained significant amounts of the variance in nestling survival, while the effect of pair identity (parents that raised the brood) was less clear ( $R = 0.14$ ,  $P = .11$ ).

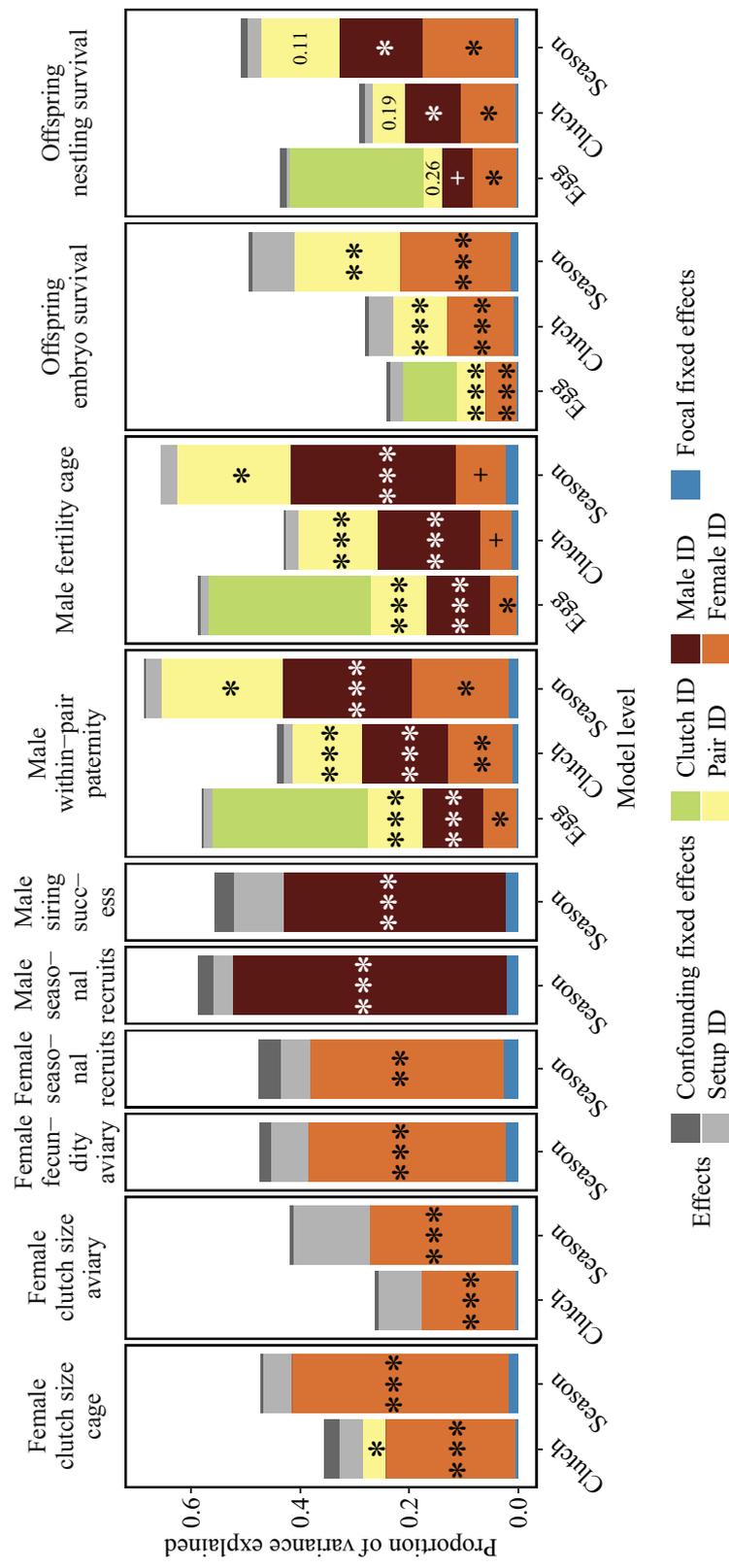
Reproductive performance traits and life span in general had low narrow-sense heritability ( $V_A/V_{Ph}$ ; Seewiesen: median  $h^2 = 0.07$ ; Bielefeld: median  $h^2 = 0.11$ ) and explained only a limited amount of the individual repeatability ( $V_A/(V_A + V_{PE})$ ; Seewiesen: median = 0.29; Bielefeld: median = 0.32; see all heritability estimates in tables S6 and S7; fig. 1, step 2). Heritability estimates from the recently wild-derived population Bielefeld were similar to those from the domesticated Seewiesen population (for nine traits measured in both populations; mean difference in  $h^2 = 0.02$ , range:  $-0.10$  to 0.13, metasummarized



**Figure 2:** Reproductive performance traits (A–C, continuous or count traits; D–F, binomial traits) as a function of inbreeding coefficient ( $F_{ped}$ ; A, D); early condition (mass at day 8), separately for populations that differ in body size (B, E); and age (C, F). Clutch size, fecundity, siring success, seasonal recruits, and life span are continuous or count traits (Z scaled), whereas the proportions of eggs fertilized, embryos survived, and nestlings survived are binomial traits. Note that these are composite figures of all effects that were examined (see fig. S3, available online, for plots of single traits with absolute trait values), such that the fate of one embryo may be shown twice, once as a function of the embryo's own  $F_{ped}$  and once as a function of its mother's  $F_{ped}$  (hence the high sample sizes,  $N_{obs}$ ). The age category zero contains measurements until day 365. Shown are smoothed regressions (red lines) with 95% confidence intervals (orange areas). Circle size (A–C) and bar width (gray; D–F) reflect sample sizes.



**Figure 3:** Standardized effect sizes with their 95% confidence intervals for inbreeding ( $F_{ped}$ ), age, and early condition (mass at day 8) on zebra finch fitness components estimated in univariate Gaussian mixed effects models where all response variables were measured at the level of individuals within seasons and all measurements were Z scaled (table S3, available online). Note that the effect of inbreeding of the offspring on its own mortality was taken from egg-based models. Negative effects of condition indicate low fitness of relatively light-weight individuals at 8 days of age. Panels separate effects of condition, age, and inbreeding of the female (A), the male (B), and the individual egg itself (C). Panel D shows the metasummarized effect sizes for reproductive performance and life span (table S4, available online). The X-axes indicate effect sizes in the form of Pearson correlation coefficients.



**Figure 4:** Variance components estimated in univariate Gaussian mixed effects models (table S5, available online). Each dependent trait is shown in a separate panel. Within panels, the X-axis separates models according to the unit of analysis, based on egg fate (egg), values per clutch (clutch), or values per individual within a breeding season (season). The Y-axis indicates the proportion of variance explained by random effects after accounting for fixed effects. Focal fixed effects refer to the total variance explained by inbreeding, age, and early condition combined. For the key variance components, numbers show nonsignificant *P* values; otherwise,  $^+ P < .1$ ,  $^* P < .05$ ,  $^{**} P < .001$ ,  $^{***} P < .0001$ . Note that models of female clutch size aviary, female fecundity aviary, and female seasonal recruits were analyzed without male ID and pair ID and, likewise, male seasonal recruits and male siring success were analyzed without female ID and pair ID because not all birds form a pair bond; male ID explained no variance in models of clutch size cage and embryo survival, while pair ID explained no variance in clutch size cage model.

difference after controlling for the uncertainty of each estimate:  $\Delta b < 0.0001$ ; mean difference in  $V_A/(V_A + V_{PE}) = 0.20$ , range:  $-0.13$  to  $0.68$ , metasummarized difference:  $\Delta b = 0.0002$ ; table S8).

#### Additive Genetic Correlations

Reproductive performance traits were grouped into three classes: (1) aspects of male reproductive performance, (2) aspects of female reproductive performance, and (3) aspects of offspring survival (table 2). Traits within each of these classes were, on average, positively correlated with each other at the additive genetic level (for the Seewiesen population, female traits: mean  $r_A = 0.66$ ,  $P < .0001$ ; male traits: mean  $r_A = 0.67$ ,  $P < .0001$ ; offspring survival traits: mean  $r_A = 0.36$ ,  $P = .09$ ; fig. 5A; see also fig. 1, steps 3 and 4). Results for the Bielefeld population are shown in figure S7. The metasummarized results are given in table S9, and all additive genetic correlation estimates are listed in tables S10 and S11 (fig. 1, step 4). Estimates of the additive genetic correlations from bivariate animal models using MCMCglmm are shown in figures S8 (Seewiesen) and S9 (Bielefeld).

Male and female reproductive performance traits were weakly negatively correlated at the additive genetic level (mean  $r_A = -0.14$ ,  $P = .04$ ; see MF in figs. 5A, S8A). Accordingly, the eigenvectors for male and female fitness traits were pointing in different directions (figs. 5B, S8B). This pattern was somewhat consistent between the Seewiesen and Bielefeld populations (see figs. S7 and S9 for the Bielefeld population). However, the negative correlation between male and female fitness traits was no longer significant when estimated by the bivariate animal models in MCMCglmm and disappeared in the Bielefeld data set (table S9). The orientation of offspring survival traits relative to male and female fitness traits was less consistent. In the Seewiesen population, female fitness traits were positively correlated with offspring survival traits at the additive genetic level (mean  $r_A = 0.61$ ,  $P < .0001$ ), while male fitness traits were not aligned with offspring survival traits (mean  $r_A = -0.11$ ,  $P = .24$ ; fig. 5). In contrast, in the Bielefeld population, both female and male fitness traits were positively correlated with offspring survival traits (fig. S7). Life span tended to be positively correlated with all reproductive performance traits (Seewiesen: mean  $r_A = 0.19$ ,  $P = .02$ ; Bielefeld: mean  $r_A = 0.60$ ,  $P = .0006$ ; figs. 5, S7; table S9).

## Discussion

### *Effects of Inbreeding, Age, and Early Condition*

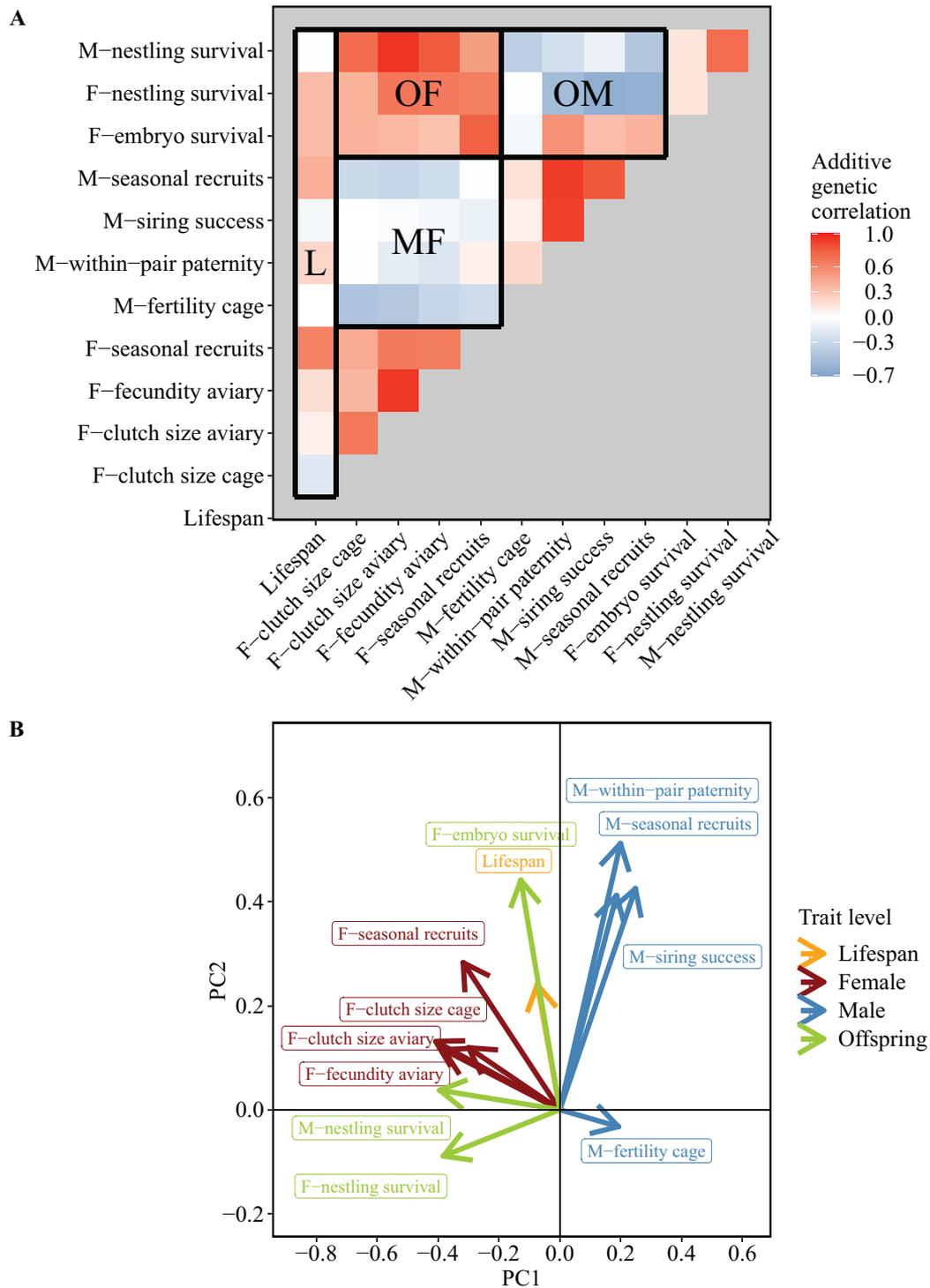
Many studies have shown that inbreeding depression significantly influences morphological, behavioral, and fitness-

related traits in zebra finches (Bolund et al. 2010a; Forstmeier et al. 2012; Hemmings et al. 2012; Opatová et al. 2016) and in other species (Amos et al. 2001; Reed and Frankham 2003; Williams et al. 2003; Michaelides et al. 2016). This study confirms that inbreeding negatively influenced all phases of offspring survival, reproductive performance, and life span. We found that the level of inbreeding of both genetic parents negatively influenced egg fertility, suggesting that this is a matter of not only sperm functionality (Opatová et al. 2016) but also female reproductive performance (e.g., egg quality or copulation behavior). Male and female fitness estimates (seasonal recruits) were most strongly affected by inbreeding (fig. 3), presumably because the successful rearing of offspring to independence requires proper functionality at every step of reproduction.

Age effects on reproductive performance typically show an initial increase in performance in both short- and long-lived species (e.g., in great tits *Parus major* [Bouwhuis et al. 2009], wandering albatrosses *Diomedea exulans* [Lecomte et al. 2010], Houbara bustards *Chlamydotis undulata* [Preston et al. 2011], Langur monkeys *Presbytis entellus* [Harely 1990], and red deer *Cervus elaphus* [Pemberton et al. 2009]). Interestingly, in our captive zebra finches we found that reproductive performance (especially male fertility, female clutch size, fecundity, and female effects on embryo survival) did not show an initial increase after birds reached sexual maturity at about 100 days of age (figs. 2C, 2F, A3). This could be because zebra finches are short-lived opportunistic breeders that reach sexual maturity earlier compared to most other birds (Zann 1996). Thus, zebra finches might have been selected to perform best early on. Alternatively, this effect may not be present in the wild, where experience might play a more important role in determining reproductive success.

Over the past decades, numerous studies focused on how early developmental conditions affect behavior, life history, and reproductive performance later in life (Tschirren et al. 2009; Rickard et al. 2010; Boersma et al. 2014). Here we show that even dramatic differences in early growth conditions of surviving offspring (see range of X-axis in fig. 2B) have remarkably small (though statistically significant) effects on adult reproductive performance.

Overall, the proportion of variance explained by inbreeding, age, and early condition (characteristics of conditions) was less than 3% (fig. 4; table S4). This indicates that individuals' robustness against poor conditions appears more noteworthy than their sensitivity. As will be discussed below, the majority of the individual repeatability in reproductive performance cannot be explained by such individual characteristics.



**Figure 5: G matrix of reproductive performance traits and life span estimated from a multivariate animal model for the Seewiesen population (shown are estimates from VCE6; for estimates of MCMCglmm bivariate models, see fig. S8; see also figs. S7 and S9 for estimates from the Bielefeld population; estimates are given in tables S10 and S11; all are available online). A, Heat map of additive genetic correlations between components of male (M), female (F), and offspring (O) fitness and life span (L). Red indicates a positive genetic correlation between traits, while blue indicates a negative correlation. Blocks marked in bold emphasize correlations between categories (e.g., MF stands for correlations between male and female fitness components). B, First two principal components of the G matrix, showing eigenvectors of the 12 fitness components. The amount of variance explained by PC1 and PC2 is 67% and 8%, respectively. Note that aspects of male fitness do not align with aspects of female and offspring fitness.**

*Repeatability and Heritability  
of Reproductive Performance*

Individual zebra finches were remarkably repeatable in their reproductive performance. Our variance-partitioning analysis showed that infertility is largely a male-specific trait, whereas embryo and offspring survival are mostly related to female identity (fig. 4; table S4). In contrast, in polyandrous crickets, egg hatching (primarily a matter of embryo survival) was mostly influenced by male identity (García-González and Simmons 2005; Ivy 2007). The effects of pair identity on infertility and offspring mortality in zebra finches may reflect behavioral incompatibility, while the pair effect on embryo mortality more likely reflects genetic incompatibility (Ihle et al. 2015).

Although male and female zebra finches are highly repeatable in their reproductive performance, the heritability of fitness traits was low. Heritability estimates were similar between the recently wild-derived Bielefeld population and the domesticated Seewiesen population. This contradicts the idea that ongoing adaptation to captivity would result in a higher heritability of fitness traits. Overall, our findings indicate that there are some additive genetic components underlying zebra finch reproductive performance.

*Evidence for Sexually Antagonistic Pleiotropy and  
Other Potential Causes of Reproductive Failure*

Some of the standing additive genetic variance in reproductive performance could be maintained by intralocus sexual antagonism between male fitness traits and female (and offspring) fitness traits (Cox and Calsbeek 2009). This has, for example, been suggested in quantitative genetic studies on *Drosophila* (Innocenti and Morrow 2010), red deer (Foerster et al. 2007), and the bank vole *Myodes glareolus* (Mills et al. 2012). We found that male fertility, siring success, and seasonal recruitment were overall negatively correlated with female fitness and offspring survival traits, suggesting that alleles that increase male fitness tend to reduce female and offspring fitness (fig. 5). In contrast, life span and reproductive performance tended to be positively correlated at the additive genetic level, which is suggestive of some overall good gene variation in our population (fig. 5). Some words of caution should be added to these observations. VCE6 (figs. 5, S7) yielded higher absolute values of estimates than those calculated with the R functions PedigreeMM (heritability estimates only) and MCMCglmm (see figs. S8, S9; also see tables S6, S7, S10, S11). Nevertheless, the additive genetic correlation estimates are highly correlated between the two methods ( $r > 0.7$ ,  $P < .0001$ ; see tables S10, S11). Estimating genetic correlations between traits with low heritability requires large data sets, especially on additive genetic correlations of between-sex reproductive performance

where the traits of male fertility and female clutch size in cages are missing ( $N$  performance traits: Seewiesen = 12, Bielefeld = 9;  $N$  birds have at least one entry of reproductive performance data: Seewiesen = 2,346, Bielefeld = 1,134; hence, these results are presented in fig. S7). Despite this lack of power in our second-largest data set of population Bielefeld, its overall orientation of traits in the additive genetic variation space of the principle components PC1 and PC2 is very similar to population Seewiesen (note that life span is in the center of all fitness traits and that aspects of female fitness do not align with male fitness in figs. 5B, S7B, S8B, and S9B).

Individual repeatability of fitness-related traits could arise from permanent environmental effects (e.g., early developmental conditions and long-lasting diseases) or from genetic effects. However, although food shortage experienced during early development (reflected in body mass at 8 days old) strongly predicted nestling mortality (Pei et al. 2020), it explained only <1% of variation in reproductive performance later in life (mean  $r = -0.058$ ; figs. 2B, 2E, 3, 4). Additionally, our captive zebra finches were raised and kept in a controlled environment with no obvious diseases detected. Additive genetic effects explained only about 30% of the large remaining unexplained individual repeatability in fitness-related traits, suggesting that reproductive performance might be (predominantly) dependent on genetic effects of local over- or underdominance and epistasis, that is, incompatibility between loci. For instance, high levels of reproductive failure could be maintained when alleles show nonadditive effects, with selection favoring the heterozygous genotype (see, e.g., Sims et al. 1984; Grossen et al. 2012). In zebra finches, males that are heterozygous for the inversion on the Z chromosome produced fast-swimming sperm and sired more offspring (Kim et al. 2017; Knief et al. 2017), while heterozygous males for inversions on chromosomes Z and 13 produced slightly more dead embryos, likely caused by unbalanced crossover during spermatogenesis (Knief et al. 2016a). However, these phenomena explain only a small fraction of infertility and embryo mortality. Overall, there is little over- or underdominance for fitness related to the major inversion polymorphisms that segregate in wild and captive zebra finch populations (Knief et al. 2016a).

Epistatic effects that involve several genes (e.g., incompatibility between nuclear loci or between mitochondrial and nuclear genomes; Zeh and Zeh 2005) could be evolutionarily stable when certain combinations of genotypes perform better than others, especially when combined with overdominance. Examples of incompatibilities are mostly known from hybrid systems (Arntzen et al. 2009; Hermansen et al. 2014; Eroukhanoff et al. 2016), but they could also be segregating within a species after the

mixing of two lineages that have evolved weak incompatibilities. Some studies on inbred lines in invertebrates found evidence of mitonuclear incompatibilities. For example, in the spider mite *Tetranychus evansi*, the eggs of F<sub>1</sub> hybrid females of two genetic lineages showed higher hatching failure compared to the pure parental lines (Knecht et al. 2017), and in *Drosophila melanogaster*, the interaction of mitonuclear background explained a small but significant amount of variation in female fitness (Downing et al. 2007).

Infertility, as one of the main and puzzling sources of reproductive failure, behaved as a male-specific trait that may also depend in part on behavioral compatibility between pair members (reflected in copulation behavior) and in part on the male's genotype at sexually antagonistic loci. The intrinsic male fertility, measured in a cage, that is, in the absence of sperm competition, correlated negatively with all female and offspring survival traits at the additive genetic level (sexual antagonism; median  $r_A = -0.30$ , range:  $-0.45$  to  $-0.01$ ; fig. 5; table S10). In contrast, in the presence of sperm competition (aviary breeding), high male within-pair paternity, siring success, and seasonal recruitment should also be influenced by the competitive ability of the individual, and this could explain why these traits correlated positively with life span and trade off less with female traits and offspring rearing ability at the additive genetic level (figs. 4, S7; tables S10, S11).

Embryo mortality, another main source of reproductive wastage, mostly depended on the identity of the genetic mother and the identity of the genetic pair members. A previous study using cross fostering of freshly laid eggs also showed that embryo mortality is a matter of the genetic parents rather than the foster environment (Ihle et al. 2015). The female component suggested an overall female genetic quality effect yet with limited heritability (pointing toward dominance variance or epistasis). The effect of the combination of parents on embryo mortality might reflect an effect of the genotype of the embryo itself, possibly involving multilocus incompatibilities (Corbett-Detig et al. 2013).

### Conclusions

Our results suggest that sexually antagonistic pleiotropy between male and female fitness plus offspring rearing traits may maintain some of the existing additive genetic variation in reproductive performance traits in captive zebra finches. Additionally, there appears to be some “good gene” (heritable) variation among reproductive performance traits and individual life span, which suggests an ongoing adaptation to the captive environment. We found that the level of inbreeding, age, and—to a lesser extent—early rearing conditions predicted a small but

statistically significant amount of variation in individual reproductive performance and life span. However, those three effects were so small that they cannot be the main causes of reproductive failure. Our results show that fertility is mostly influenced by the male, whereas embryo and nestling survival are mainly influenced by the female. Although individual zebra finches were moderately repeatable in their reproductive performance, the heritability of those traits was low. Overall, our results suggest that alleles that have additive effects on fitness might be maintained through sexually antagonistic pleiotropy and that the major genetic causes of reproductive failure might be determined by genetic incompatibilities or local dominance effects.

### Acknowledgments

We thank M. Schneider for molecular work; K. Teltscher for counting sperm on the perivitelline membranes of the eggs; H. Milewski for advice on linear algebra; H. Schielzeth, E. Bolund, M. Ihle, U. Knief, C. Burger, S. Janker, J. Schreiber, T. Aronson, and A. Edme for help with data collection; T. Birkhead, T. Krause, J. Lesku, and K. Roberts for providing birds; and S. Bauer, E. Bodendorfer, J. Didsbury, A. Grötsch, J. Hacker, A. Kortner, J. Minshull, P. Neubauer, C. Scheicher, F. Weigel, and B. Wörle for animal care and help with breeding. We are grateful to editor in chief Daniel Bolnick, associate editor Daniel Matute, and two anonymous reviewers for constructive comments on the manuscript. This research was supported by the Max Planck Society (to B.K.). Y.P. is part of the International Max Planck Research School for Organismal Biology.

### Statement of Authorship

W.F. and B.K. designed and planned the study. W.F., D.W., and K.M. collected reproductive performance data. Y.P. and W.F. analyzed the data, with input from J.R. Y.P., W.F., and B.K. interpreted the results and wrote the manuscript, with input from J.R. All authors contributed to the final manuscript.

### Data and Code Availability

Supporting data, model structures, and R scripts have been deposited in the Open Science Framework (<https://doi.org/10.17605/OSF.IO/TGSZ8>; Pei et al. 2020).

### Literature Cited

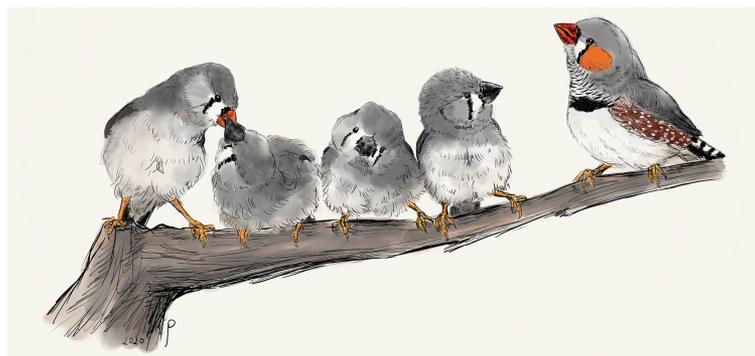
Amos, W., J. Worthington Wilmer, K. Fullard, T. M. Burg, J. P. Croxall, D. Bloch, and T. Coulson. 2001. The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society B* 268:2021–2027.

- Arntzen, J. W., R. Jehle, F. Bardakci, T. Burke, and G. P. Wallis. 2009. Asymmetric viability of reciprocal-cross hybrids between crested and marbled newts (*Triturus cristatus* and *T. marmoratus*). *Evolution* 63:1191–1202.
- Ayalon, N. 1978. A review of embryonic mortality in cattle. *Reproduction* 54:483–493.
- Balakrishnan, C. N., and S. V. Edwards. 2009. Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* 181:645–660.
- Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1–48.
- Bechsgaard, J. S., A. A. Hoffmann, C. Sgró, V. Loeschcke, T. Bilde, and T. N. Kristensen. 2013. A comparison of inbreeding depression in tropical and widespread *Drosophila* species. *PLoS ONE* 8: e51176.
- Berner, D. 2012. How much can the orientation of G's eigenvectors tell us about genetic constraints? *Ecology and Evolution* 2:1834–1842.
- Birkhead, T. R., and F. Fletcher. 1998. Sperm transport in the reproductive tract of female zebra finches (*Taeniopygia guttata*). *Reproduction* 114:141–145.
- Birkhead, T. R., J. Hall, E. Schut, and N. Hemmings. 2008. Unhatched eggs: methods for discriminating between infertility and early embryo mortality. *Ibis* 150:508–517.
- Boersma, G. J., T. L. Bale, P. Casanello, H. E. Lara, A. B. Lucion, D. Suchecki, and K. L. Tamashiro. 2014. Long-term impact of early life events on physiology and behaviour. *Journal of Neuroendocrinology* 26:587–602.
- Bolund, E., K. Martin, B. Kempenaers, and W. Forstmeier. 2010a. Inbreeding depression of sexually selected traits and attractiveness in the zebra finch. *Animal Behaviour* 79:947–955.
- Bolund, E., H. Schielzeth, and W. Forstmeier. 2007. Intrasexual competition in zebra finches, the role of beak colour and body size. *Animal Behaviour* 74:715–724.
- . 2010b. No heightened condition dependence of zebra finch ornaments: a quantitative genetic approach. *Journal of Evolutionary Biology* 23:586–597.
- Bouwhuis, S., B. C. Sheldon, S. Verhulst, and A. Charmantier. 2009. Great tits growing old: selective disappearance and the partitioning of senescence to stages within the breeding cycle. *Proceedings of the Royal Society B* 276:2769–2777.
- Briskie, J. V., and M. Mackintosh. 2004. Hatching failure increases with severity of population bottlenecks in birds. *Proceedings of the National Academy of Sciences of the USA* 101:558–561.
- Bunin, J. S., I. G. Jamieson, and E. D. 2008. Low reproductive success of the endangered Takahe *Porphyrio mantelli* on offshore island refuges in New Zealand. *Ibis* 139:144–151.
- Carrell, D. T., and K. I. Aston. 2011. The search for SNPs, CNVs, and epigenetic variants associated with the complex disease of male infertility. *Systems Biology in Reproductive Medicine* 57:17–26.
- Charlesworth, D., and B. Charlesworth. 1987. Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics* 18:237–268.
- Christians, E., A. A. Davis, S. D. Thomas, and I. J. Benjamin. 2000. Maternal effect of *Hsf1* on reproductive success. *Nature* 407:693–694.
- Corbett-Detig, R. B., J. Zhou, A. G. Clark, D. L. Hartl, and J. F. Ayroles. 2013. Genetic incompatibilities are widespread within species. *Nature* 504:135–137.
- Coster, A. 2015. R package pedigree version 1.4. <https://cran.r-project.org/web/packages/pedigree/pedigree.pdf>.
- Cox, R. M., and R. Calsbeek. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. *American Naturalist* 173:176–187.
- De Braekeleer, M., and T. N. Dao. 1991. Cytogenetic studies in male infertility: a review. *Human Reproduction* 6:245–250.
- Dobzhansky, T. 1936. Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. *Genetics* 21:113–135.
- Dowling, D. K., U. Friberg, F. Hailer, and G. Arnqvist. 2007. Intergenomic epistasis for fitness: within-population interactions between cytoplasmic and nuclear genes in *Drosophila melanogaster*. *Genetics* 175:235–244.
- Eroukhmanoff, F., M. Rowe, E. R. A. Cramer, F. Haas, J. S. Hermansen, A. Runemark, A. Johnsen, et al. 2016. Experimental evidence for ovarian hypofunction in sparrow hybrids. *Avian Research* 7:3.
- Fishman, L., and J. H. Willis. 2006. A cytonuclear incompatibility causes another sterility in *Mimulus* hybrids. *Evolution* 60:1372–1381.
- Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. B. B. Kruuk. 2007. Sexually antagonistic genetic variation for fitness in red deer. *Nature* 447:1107–1110.
- Forstmeier, W., D. W. Coltman, and T. R. Birkhead. 2004. Maternal effects influence the sexual behavior of sons and daughters in the zebra finch. *Evolution* 58:2574–2583.
- Forstmeier, W., H. Schielzeth, J. C. Mueller, H. Ellegren, and B. Kempenaers. 2012. Heterozygosity–fitness correlations in zebra finches: microsatellite markers can be better than their reputation. *Molecular Ecology* 21:3237–3249.
- Forstmeier, W., H. Schielzeth, M. Schneider, and B. Kempenaers. 2007a. Development of polymorphic microsatellite markers for the zebra finch (*Taeniopygia guttata*). *Molecular Ecology Notes* 7:1026–1028.
- Forstmeier, W., G. Segelbacher, J. C. Mueller, and B. Kempenaers. 2007b. Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular Ecology* 16:4039–4050.
- Garcia, N., C. Lopez-Fanjul, and A. Garcia-Dorado. 1994. The genetics of viability in *Drosophila melanogaster*: effects of inbreeding and artificial selection. *Evolution* 48:1277–1285.
- García-González, F., and L. W. Simmons. 2005. The evolution of polyandry: intrinsic sire effects contribute to embryo viability. *Journal of Evolutionary Biology* 18:1097–1103.
- Gorman, H. E., and R. G. Nager. 2004. Prenatal developmental conditions have long-term effects on offspring fecundity. *Proceedings of the Royal Society B* 271:1923–1928.
- Griffith, S. C., O. L. Crino, S. C. Andrew, F. Y. Nomano, E. Adkins-Regan, C. Alonso-Alvarez, I. E. Bailey, et al. 2017. Variation in reproductive success across captive populations: methodological differences, potential biases and opportunities. *Ethology* 123:1–29.
- Griffith, S. C., S. R. Pryke, and M. Mariette. 2008. Use of nest-boxes by the zebra finch (*Taeniopygia guttata*): implications for reproductive success and research. *Emu-Austral Ornithology* 108:311–319.
- Grossen, C., S. Neuenschwander, and N. Perrin. 2012. The balanced lethal system of crested newts: a ghost of sex chromosomes past? *American Naturalist* 180:E174–E183.
- Gwaza, D. S., N. I. Dim, and O. M. Momoh. 2016. Evaluation of sire and dam for genetic improvement of fertility, embryonic mortality and hatchability in Nigerian local chicken populations. *Advancements in Genetic Engineering* 5:147–153.

- Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* 33:1–22.
- Hadfield, J. D., A. J. Wilson, D. Garant, B. C. Sheldon, and L. E. B. Kruuk. 2010. The misuse of BLUP in ecology and evolution. *American Naturalist* 175:116–125.
- Harely, D. 1990. Aging and reproductive performance in langur monkeys (*Presbytis entellus*). *American Journal of Physical Anthropology* 83:253–261.
- Hemmings, N. L., J. Slate, and T. R. Birkhead. 2012. Inbreeding causes early death in a passerine bird. *Nature Communications* 3:1–4.
- Henderson, C. R. 1953. Estimation of variance and covariance components. *Biometric* 9:226–252.
- Hermansen, J. S., F. Haas, C. N. Trier, R. I. Bailey, A. J. Nederbragt, A. Marzal, and G.-P. Saetre. 2014. Hybrid speciation through sorting of parental incompatibilities in Italian sparrows. *Molecular Ecology* 23:5831–5842.
- Hoffman, J. I., F. Simpson, P. David, J. M. Rijks, T. Kuiken, M. A. S. Thorne, R. C. Lacy, et al. 2014. High-throughput sequencing reveals inbreeding depression in a natural population. *Proceedings of the National Academy of Sciences of the USA* 111:3775–3780.
- Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous inference in general parametric models. *Biometrical Journal* 50:346–363.
- Houslay, T. M., and A. J. Wilson. 2017. Avoiding the misuse of BLUP in behavioural ecology. *Behavioral Ecology* 28:948–952.
- Howrigan, D. P., M. A. Simonson, and M. C. Keller. 2011. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC Genomics* 12:460.
- Huisman, J., L. E. B. Kruuk, P. A. Ellis, T. Clutton-Brock, and J. M. Pemberton. 2016. Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of the National Academy of Sciences of the USA* 113:3585–3590.
- Ihle, M., B. Kempnaers, and W. Forstmeier. 2015. Fitness benefits of mate choice for compatibility in a socially monogamous species. *PLoS Biology* 13:1–21.
- Innocenti, P., and E. H. Morrow. 2010. The sexually antagonistic genes of *Drosophila melanogaster*. *PLoS Biology* 8:e1000335.
- Ivy, T. M. 2007. Good genes, genetic compatibility and the evolution of polyandry: use of the diallel cross to address competing hypotheses. *Journal of Evolutionary Biology* 20:479–487.
- Jackson, A. K., D. C. Evers, M. A. Etterson, A. M. Condon, S. B. Folsom, J. Detweiler, J. Schmerfeld, et al. 2011. Mercury exposure affects the reproductive success of a free-living terrestrial songbird, the Carolina wren (*Thryothorus ludovicianus*). *Auk* 128:759–769.
- Jerónimo, S., M. Khadraoui, D. Wang, K. Martin, J. A. Lesku, K. A. Robert, E. Schlicht, et al. 2018. Plumage color manipulation has no effect on social dominance or fitness in zebra finches. *Behavioral Ecology* 29:459–467.
- Johnson, N. 2008. Hybrid incompatibility and speciation. *Nature Education* 1:20.
- Jurewicz, J., W. Hanke, M. Radwan, and J. Bonde. 2009. Environmental factors and semen quality. *International Journal of Occupational Medicine and Environmental Health* 22:305–329.
- Keller, L. F., and D. M. Waller. 2002. Inbreeding effects in wild populations. *Trends in Ecology and Evolution* 17:230–241.
- Khan, N., R. A. Peters, and K. Robert. 2016. Compensating for a stressful start: maternal corticosterone, offspring survival, and size at fledging in the zebra finch, *Taeniopygia guttata*. *Emu* 116:262–272.
- Kim, K.-W., C. Bennison, N. Hemmings, L. Brookes, L. L. Hurley, S. C. Griffith, T. Burke, et al. 2017. A sex-linked supergene controls sperm morphology and swimming speed in a songbird. *Nature Ecology and Evolution* 1:1168–1176.
- Knegt, B., T. Potter, N. A. Pearson, Y. Sato, H. Staudacher, B. C. J. Schimmel, E. T. Kiers, et al. 2017. Detection of genetic incompatibilities in non-model systems using simple genetic markers: hybrid breakdown in the haplodiploid spider mite *Tetranychus evansi*. *Heredity* 118:311–321.
- Knief, U., W. Forstmeier, Y. Pei, M. Ihle, D. Wang, K. Martin, P. Opatová, et al. 2017. A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nature Ecology and Evolution* 1:1177–1184.
- Knief, U., G. Hemmrich-Stanisak, M. Wittig, A. Franke, S. C. Griffith, B. Kempnaers, and W. Forstmeier. 2015a. Quantifying realized inbreeding in wild and captive animal populations. *Heredity* 114:397–403.
- . 2016a. Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biology* 17:199.
- Knief, U., B. Kempnaers, and W. Forstmeier. 2016b. Meiotic recombination shapes precision of pedigree- and marker-based estimates of inbreeding. *Heredity* 118:239–248.
- Knief, U., H. Schielzeth, H. Ellegren, B. Kempnaers, and W. Forstmeier. 2015b. A prezygotic transmission distorter acting equally in female and male zebra finches *Taeniopygia guttata*. *Molecular Ecology* 24:3846–3859.
- Krausz, C., and A. Riera-Escamilla. 2018. Genetics of male infertility. *Nature Reviews Urology* 15:369–384.
- Kruuk, L. E. B., and J. D. Hadfield. 2007. How to separate genetic and environmental causes of similarity between relatives. *Journal of Evolutionary Biology* 20:1890–1903.
- Lecomte, V. J., G. Sorci, S. Cornet, A. Jaeger, B. Faivre, E. Arnoux, M. Gaillard, et al. 2010. Patterns of aging in the long-lived wandering albatross. *Proceedings of the National Academy of Sciences of the USA* 107:6370–6375.
- Lemon, W. C., and R. H. Barth. 1992. The effects of feeding rate on reproductive success in the zebra finch, *Taeniopygia guttata*. *Animal Behaviour* 44:851–857.
- Lindholm, A. K., K. A. Dyer, R. C. Firman, L. Fishman, W. Forstmeier, L. Holman, H. Johannesson, et al. 2016. The ecology and evolutionary dynamics of meiotic drive. *Trends in Ecology and Evolution* 31:315–326.
- Lyon, M. F. 1986. Male sterility of the mouse *t*-complex is due to homozygosity of the distorter genes. *Cell* 44:357–363.
- Michaelides, S. N., G. M. While, N. Zajac, F. Aubret, B. Calsbeek, R. Sacchi, M. A. L. Zuffi, et al. 2016. Loss of genetic diversity and increased embryonic mortality in non-native lizard populations. *Molecular Ecology* 25:4113–4125.
- Mills, S. C., E. Koskela, and T. Mappes. 2012. Intralocus sexual conflict for fitness: sexually antagonistic alleles for testosterone. *Proceedings of the Royal Society B* 279:1889–1895.
- Miyamoto, T., A. Tsujimura, Y. Miyagawa, E. Koh, M. Namiki, and K. Sengoku. 2012. Male infertility and its causes in human. *Advances in Urology* 2012:1–7.
- Murray, J. R., C. W. Varian-Ramos, Z. S. Welch, and M. S. Saha. 2013. Embryological staging of the zebra finch, *Taeniopygia guttata*. *Journal of Morphology* 274:1090–1110.
- Nakagawa, S., and H. Schielzeth. 2010. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biological Reviews* 85:935–956.

- Neumaier, A., and E. Groeneveld. 1998. Restricted maximum likelihood estimation of covariances in sparse linear models. *Genetics Selection Evolution* 30:3.
- Opatová, P., M. Ihle, J. Albrechtová, O. Tomášek, B. Kempnaers, W. Forstmeier, and T. Albrecht. 2016. Inbreeding depression of sperm traits in the zebra finch *Taeniopygia guttata*. *Ecology and Evolution* 6:295–304.
- Pei, Y. 2020. Data from: Proximate causes of infertility and embryo mortality in captive zebra finches. *American Naturalist*, Open Science Framework, <https://doi.org/10.17605/OSF.IO/TGSZ8>.
- Pei, Y., W. Forstmeier, and B. Kempnaers. 2020. Offspring performance is well buffered against stress experienced by ancestors. *Evolution* 74:1525–1539.
- Pemberton, J. M., L. E. B. Kruuk, A. Morris, T. H. Clutton-Brock, M. N. Clements, and D. H. Nussey. 2009. Inter- and intrasexual variation in aging patterns across reproductive traits in a wild red deer population. *American Naturalist* 174:342–357.
- Preston, B. T., M. Saint Jalme, Y. Hingrat, F. Lacroix, and G. Sorci. 2011. Sexually extravagant males age more rapidly. *Ecology Letters* 14:1017–1024.
- R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Reed, D. H., and R. Frankham. 2003. Correlation between fitness and genetic diversity. *Conservation Biology* 17:230–237.
- Rickard, I. J., J. Holopainen, S. Helama, S. Helle, A. F. Russell, and V. Lummaa. 2010. Food availability at birth limited reproductive success in historical humans. *Ecology* 91:3515–3525.
- Robinson, G. K. 1991. That BLUP is a good thing: the estimation of random effects. *Statistical Science* 6:15–51.
- Safronova, L. D., and V. L. Chubykin. 2013. Meiotic drive in mice carrying *t*-complex in their genome. *Russian Journal of Genetics* 49:885–897.
- Sandler, L., Y. Hiraizumi, and I. Sandler. 1959. Meiotic drive in natural populations of *Drosophila melanogaster*. I. the cytogenetic basis of segregation-distortion. *Genetics* 44:233–250.
- Schluter, D., T. D. Price, and L. Rowe. 1991. Conflicting selection pressures and life history trade-offs. *Proceedings of the Royal Society B* 246:11–17.
- Shirasuna, K., and H. Iwata. 2017. Effect of aging on the female reproductive function. *Contraception and Reproductive Medicine* 2:23.
- Sierra, S., and M. Stephenson. 2006. Genetics of recurrent pregnancy loss. *Seminars in Reproductive Medicine* 24:17–24.
- Sims, S. H., H. C. Macgregor, P. S. Pellatt, and H. A. Horner. 1984. Chromosome 1 in crested and marbled newts (*Triturus*): an extraordinary case of heteromorphism and independent chromosome evolution. *Chromosoma* 89:169–185.
- Speroff, L. 1994. The effect of aging on fertility. *Current Opinion in Obstetrics and Gynecology* 6:115–120.
- Stearns, S. C. 1989. Trade-offs in life-history evolution. *Functional Ecology* 3:259–268.
- Tan, C. K. W., T. Pizzari, and S. Wigby. 2013. Parental age, gametic age, and inbreeding interact to modulate offspring viability in *Drosophila melanogaster*. *Evolution* 67:3043–3051.
- Tschirren, B., A. N. Rutstein, E. Postma, M. Mariette, and S. C. Griffith. 2009. Short- and long-term consequences of early developmental conditions: a case study on wild and domesticated zebra finches. *Journal of Evolutionary Biology* 128:103–115.
- van de Pol, M., and J. Wright. 2009. A simple method for distinguishing within- versus between-subject effects using mixed models. *Animal Behaviour* 77:753–758.
- Van Doorn, G. S. 2009. Intra-locus sexual conflict. *Annals of the New York Academy of Sciences* 1168:52–71.
- Vazquez, A. I., D. M. Bates, G. J. M. Rosa, D. Gianola, and K. A. Weigel. 2010. Technical note: an R package for fitting generalized linear mixed models in animal breeding. *Journal of Animal Science* 88:497–504.
- Wang, D., N. Kempnaers, B. Kempnaers, and W. Forstmeier. 2017. Male zebra finches have limited ability to identify high-fecundity females. *Behavioral Ecology* 28:784–792.
- Williams, C. G., L. D. Auckland, M. M. Reynolds, and K. A. Leach. 2003. Overdominant lethals as part of the conifer embryo lethal system. *Heredity* 91:584–592.
- Wright, S. 1922. Coefficients of inbreeding and relationship source. *American Naturalist* 56:330–338.
- Zann, R. 1996. *The zebra finch: a synthesis of field and laboratory studies*. Oxford University Press, New York.
- Zeh, J. A., and D. W. Zeh. 2005. Maternal inheritance, sexual conflict and the maladapted male. *Trends in Genetics* 21:281–286.

Associate Editor: Daniel R. Matute  
Editor: Daniel I. Bolnick



A family of zebra finches (*Taeniopygia guttata*). The mother (left) is feeding one of the three fledglings while the father (right) is watching the surrounding. Drawing by Yifan Pei.

## Supplementary Materials

### Supplementary Methods

To confirm the estimates from animal models in VCE6 (Neumaier and Groeneveld 1998), we repeated Steps 2 and 4 (fig. 1 in the main text) using software packages ‘pedigreemm’ V 0.3-3 (Vazquez et al. 2010) and ‘MCMCglmm’ V 2.29 (Hadfield 2010). For all MCMCglmm models, we used 1,300,000 iterations, with a thinning interval of 1,000 and burn-in interval of 300,000. Priors are detailed as below.

#### *Step 2: Estimation of heritability of fitness-related traits*

In Step 2 (fig. 1 in the main text), heritability calculation was repeated in ‘pedigreemm’ and ‘MCMCglmm’ using the same model structure as used in VCE6 (in the main text) at the season level (fig. 1). For ‘MCMCglmm’ univariate animal models, we used a default non-informative prior, with the expected variance of random effects and residuals (‘V’) set to one, and the degree-of-believe parameter (‘nu’) set to 0.002. For ‘pedigreemm’ models, aggregated data were weighted by the number of eggs (male fertility cage, male within-pair paternity, female embryo and nestling survival and male nestling survival) or clutches for an individual within a season (female clutch size in cage and aviary; fig. 1).

Heritability estimates were highly correlated among the three software packages (VCE6, pedigreemm, MCMCglmm;  $h^2$ :  $r > 0.78$ ,  $P < 0.0001$ ;  $VA/(VA+VPE)$ :  $r > 0.65$ ,  $P < 0.002$ ; Supplementary tables S6-S7).

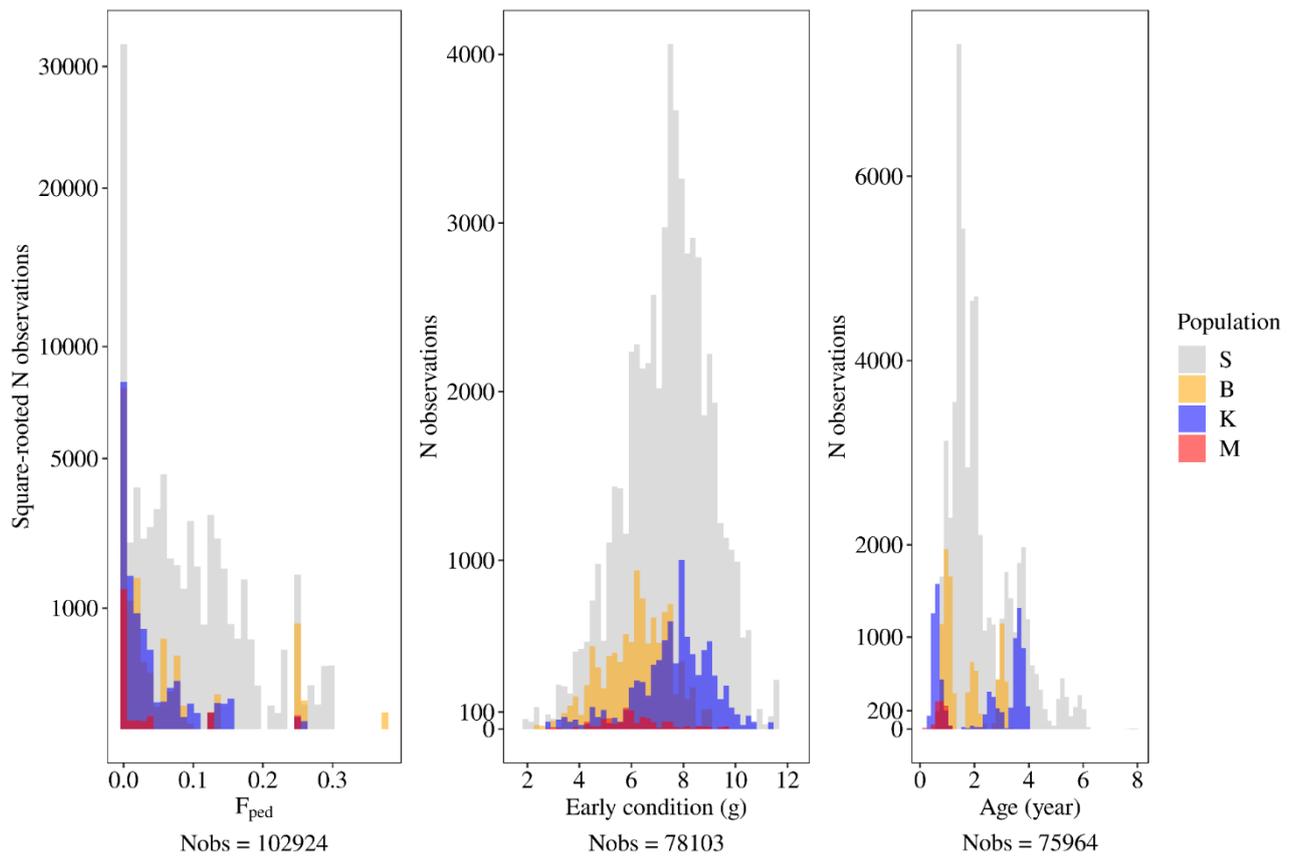
#### *Step 4: Estimation of additive genetic correlation*

In Step 4 (fig. 1 in the main text), for multivariate animal models, because a 12-trait (population ‘Seewiesen’) or 9-trait (population ‘Bielefeld’) model with  $> 1000$  individuals would be difficult to run, we instead fitted 66 bivariate models for ‘Seewiesen’ population (36 for ‘Bielefeld’ population) to obtain a second estimate for each genetic correlation. As priors, we set the expected additive genetic effect ‘VA’ to 0.2, the expected residual ‘VR’ to 0.8, and used ‘nu’ = 1.002. The estimates of the additive genetic correlations were highly correlated between the two software packages ( $r = 0.77$ ,  $N = 66$  estimates from the Seewiesen and  $r = 0.70$ ,  $N = 36$  from Bielefeld population,  $P < 0.0001$ ; Supplementary tables S10-S11).

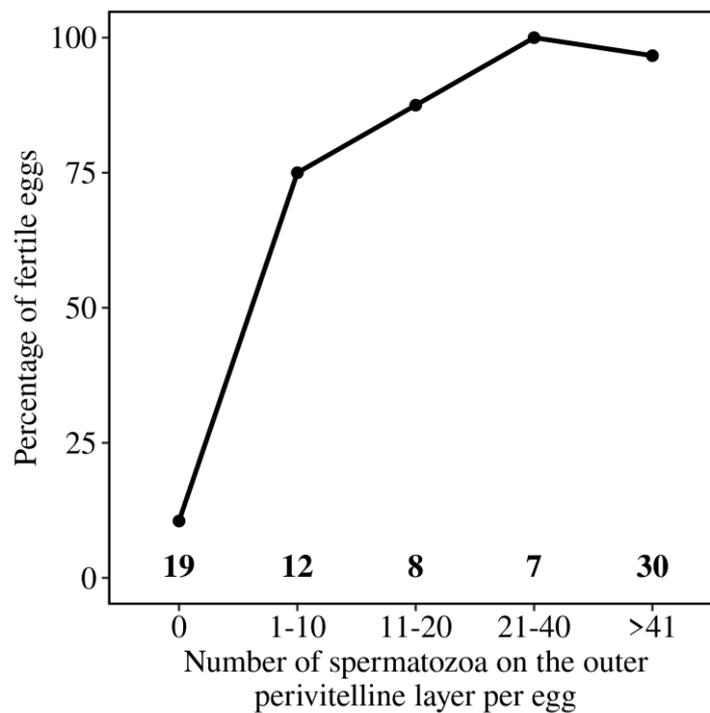
## References

- Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software* 33:1–22.
- Neumaier, A., and E. Groeneveld. 1998. Restricted maximum likelihood estimation of covariances in sparse linear models. *Genetics Selection Evolution* 30:3.
- Vazquez, A. I., D. M. Bates, G. J. M. Rosa, D. Gianola, and K. A. Weigel. 2010. Technical note: An R package for fitting generalized linear mixed models in animal breeding. *Journal of Animal Science* 88:497–504.

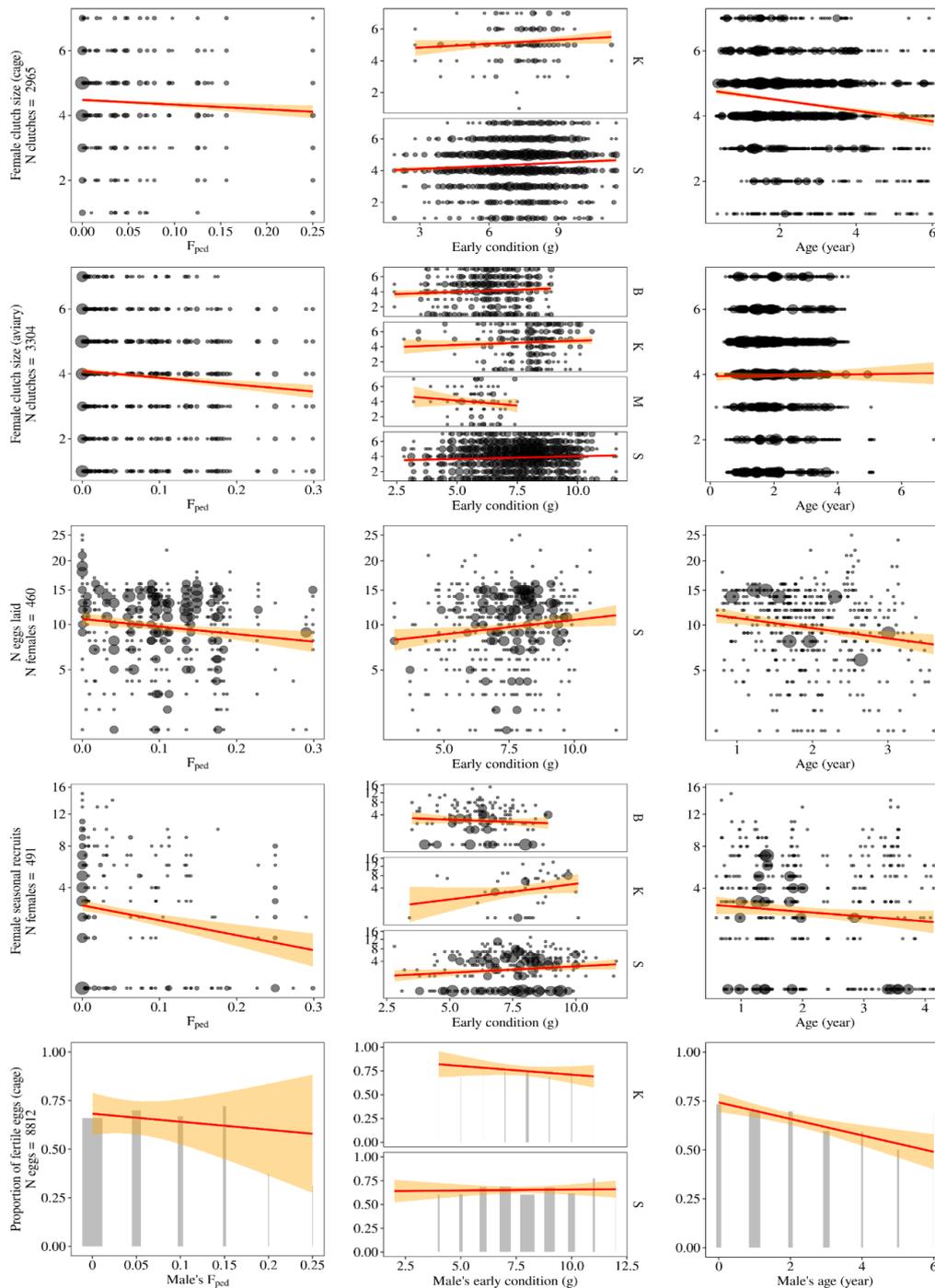
**Figure S1.** Overlaying histograms of inbreeding coefficient ( $F_{ped}$ , left), early condition (mass at day 8, centre), and age (right) of all observations of reproductive performance traits, for populations Seewiesen ('S', grey), Bielefeld ('B', orange), Krakow ('K', blue) and Melbourne ('M', red). Reproductive performance traits include clutch size (in cages and aviaries), fecundity, siring success, seasonal recruits, lifespan, fertility in cages, within-pair paternity in aviaries, embryo survival and nestling survival (Step 1, all observations used in models of 'estimate repeatability', in fig. 1; also see fig. 2 for the reproductive performance traits as a function of inbreeding coefficient, early condition and age). For better illustration, the y-axis of the inbreeding coefficient histogram (left) was square-root transformed. Note that a value for an individual is used repeatedly (for each dependent trait).



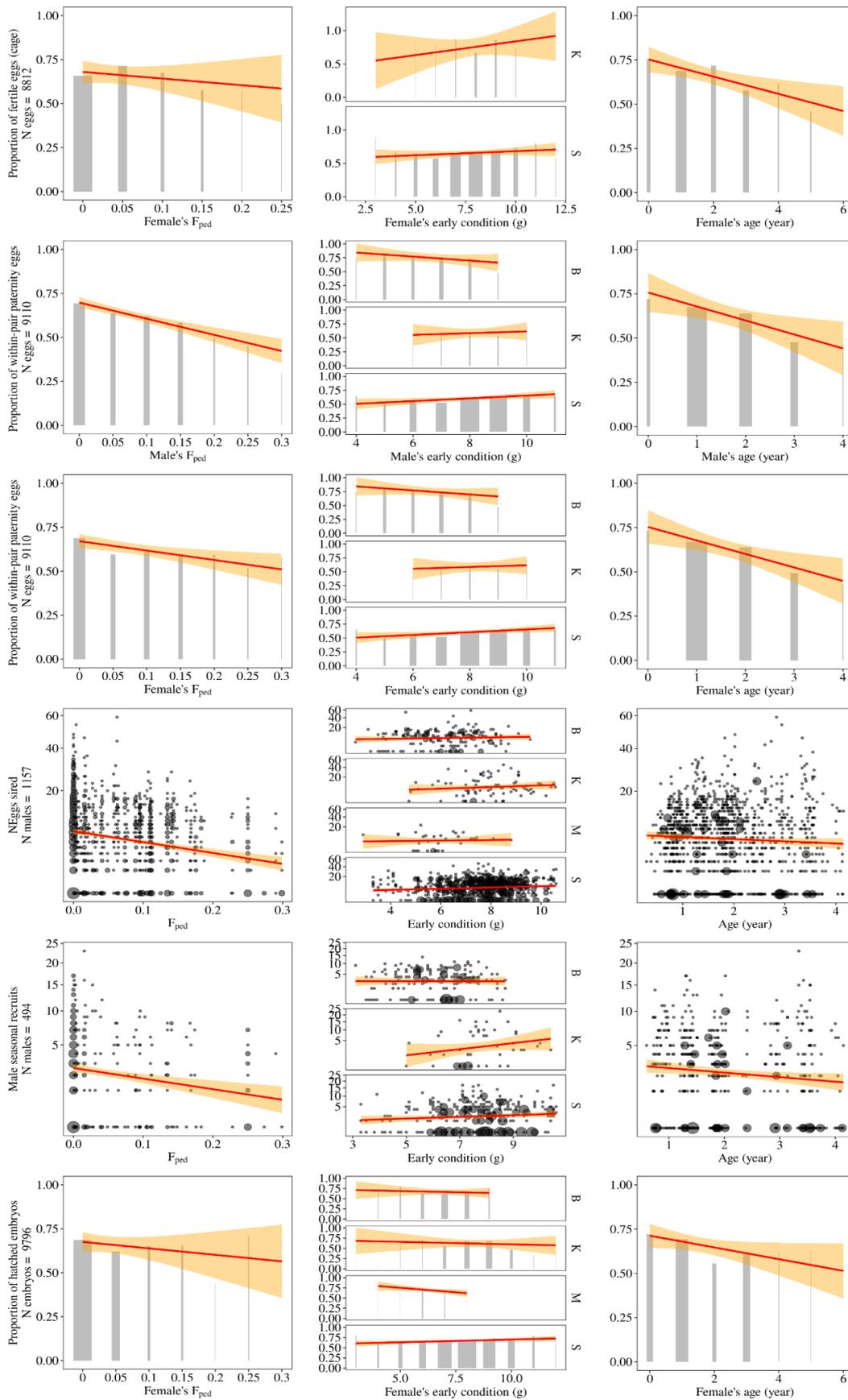
**Figure S2.** Relationship between the number of spermatozoa on the perivitelline layer (PVL) of the egg and the percentage of eggs being fertile. Numbers indicate the number of eggs in each category of sperm count. Here, 76 freshly laid eggs were placed in the incubator for 24 hours to acquire more embryonic cells. Then each egg was opened, and the germinal disc was collected using a flat gel loading pipette tip. The tip was then cut off and dropped into an Eppendorf tube with 70% ethanol for subsequent DNA extraction and genotyping. Fertile eggs are identified by successful genotyping of both maternal and paternal alleles from the germinal disc. To count the sperm that were trapped on the PVL, egg membranes were washed in water and then carefully placed flat on a microscope slide. To aid identification of sperm on PVL, a drop of Hoechst 33342 fluorescent dye was added to the slide and left to dry under room temperature. The number of sperms on the entire membrane was counted using a Zeiss Axio Imager.M2 microscope at 200x magnification.



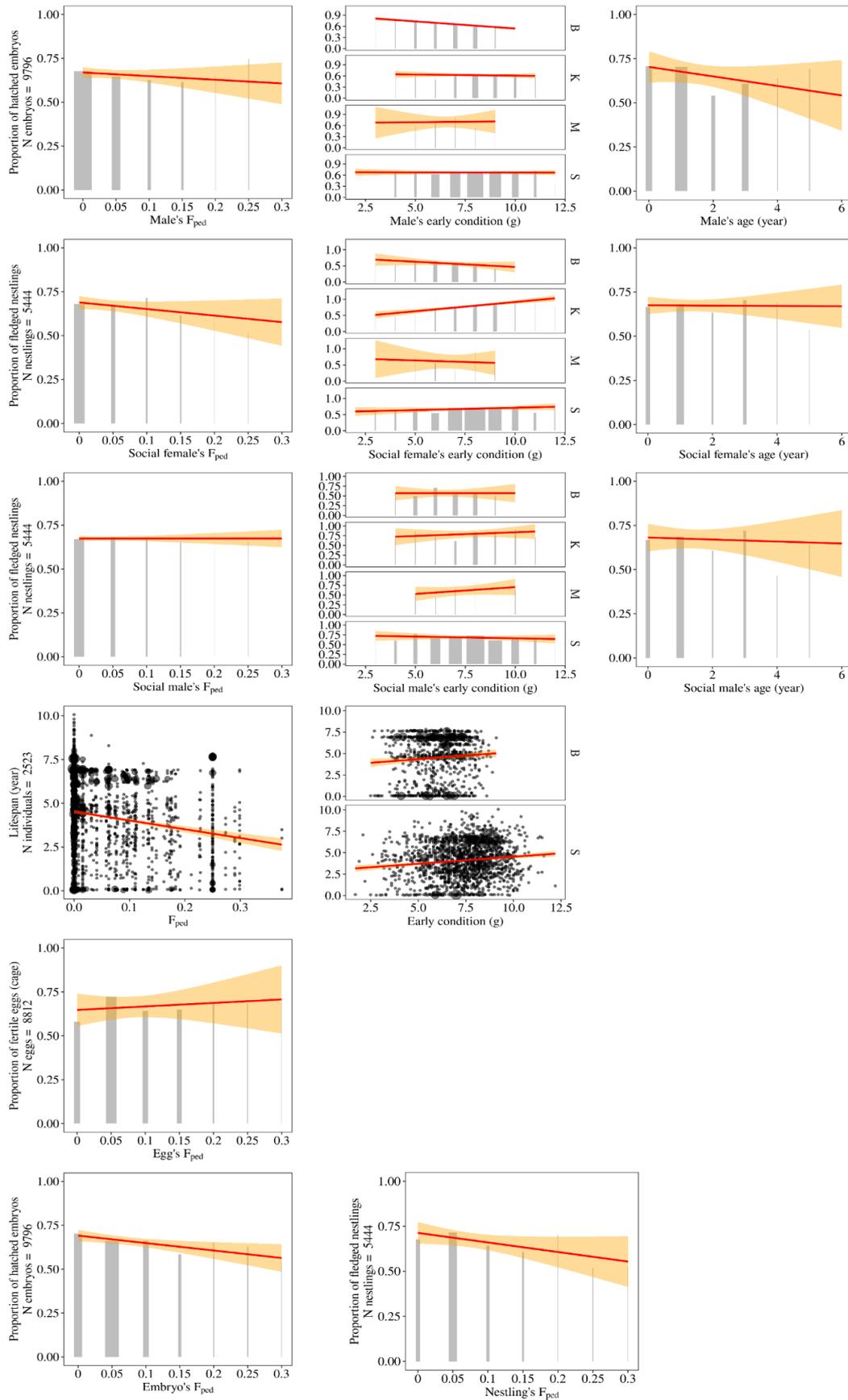
**Figure S3.** Reproductive performance traits (continuous or count traits as scatter plots, binomial traits as bar plots) as a function of the inbreeding coefficient ( $F_{ped}$ , left), early condition (mass at day 8, separately for populations that differ in size, centre), and age (right; Step 1 in fig. 1). The age category zero contains measurements between day 100 and day 365. Red lines and yellow areas show linear regressions with 95% CIs. Dot size and bar width reflect sample size. The total sample sizes are indicated on the Y-axis label. To better illustrate count data, the Y-axes were square-root transformed for ‘female seasonal recruits’, ‘N eggs laid’, ‘male seasonal recruits’ and ‘N eggs sired’.



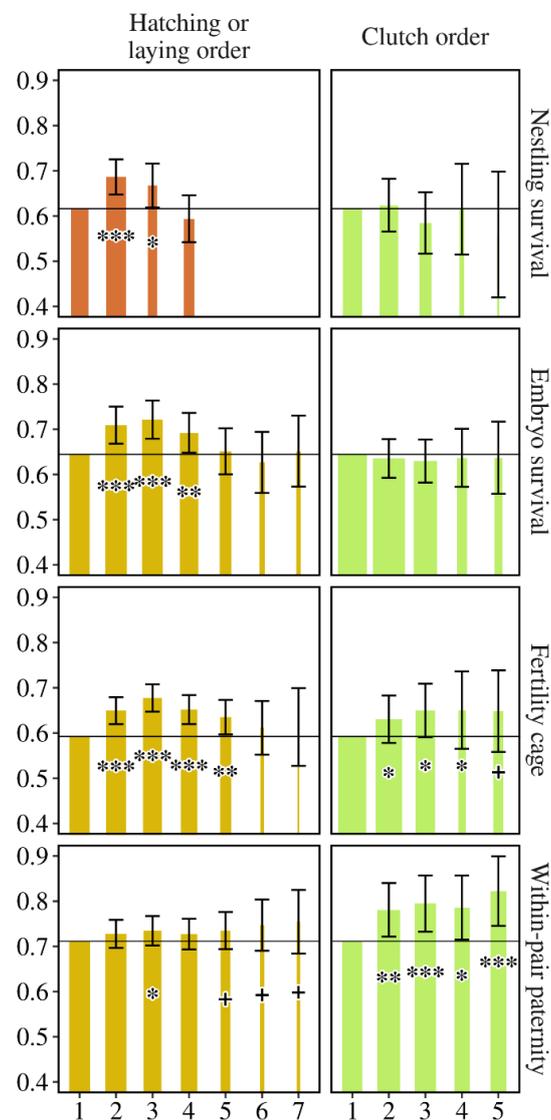
(Fig. S3 continued)



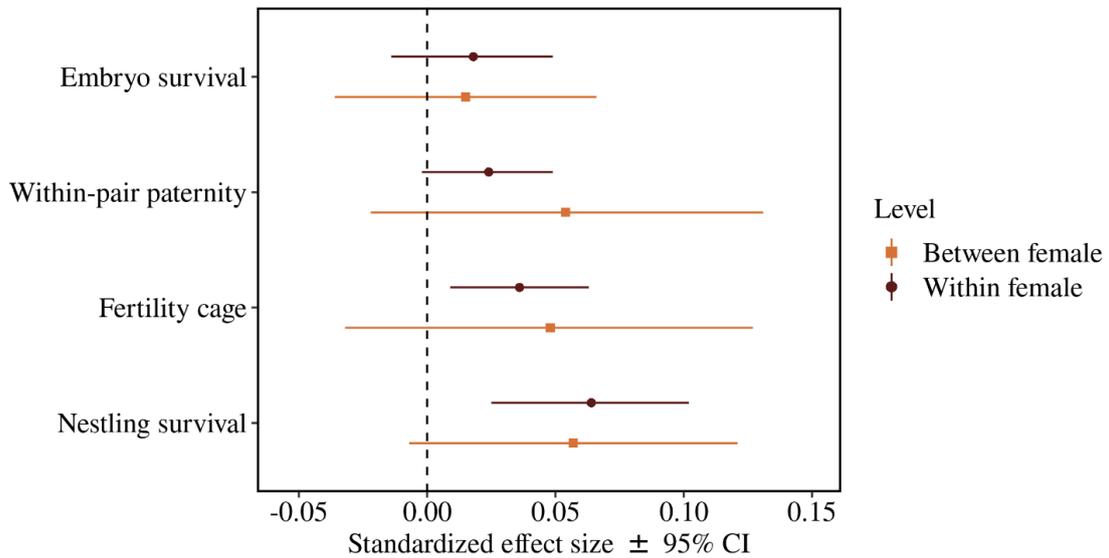
(Fig. S3 continued)



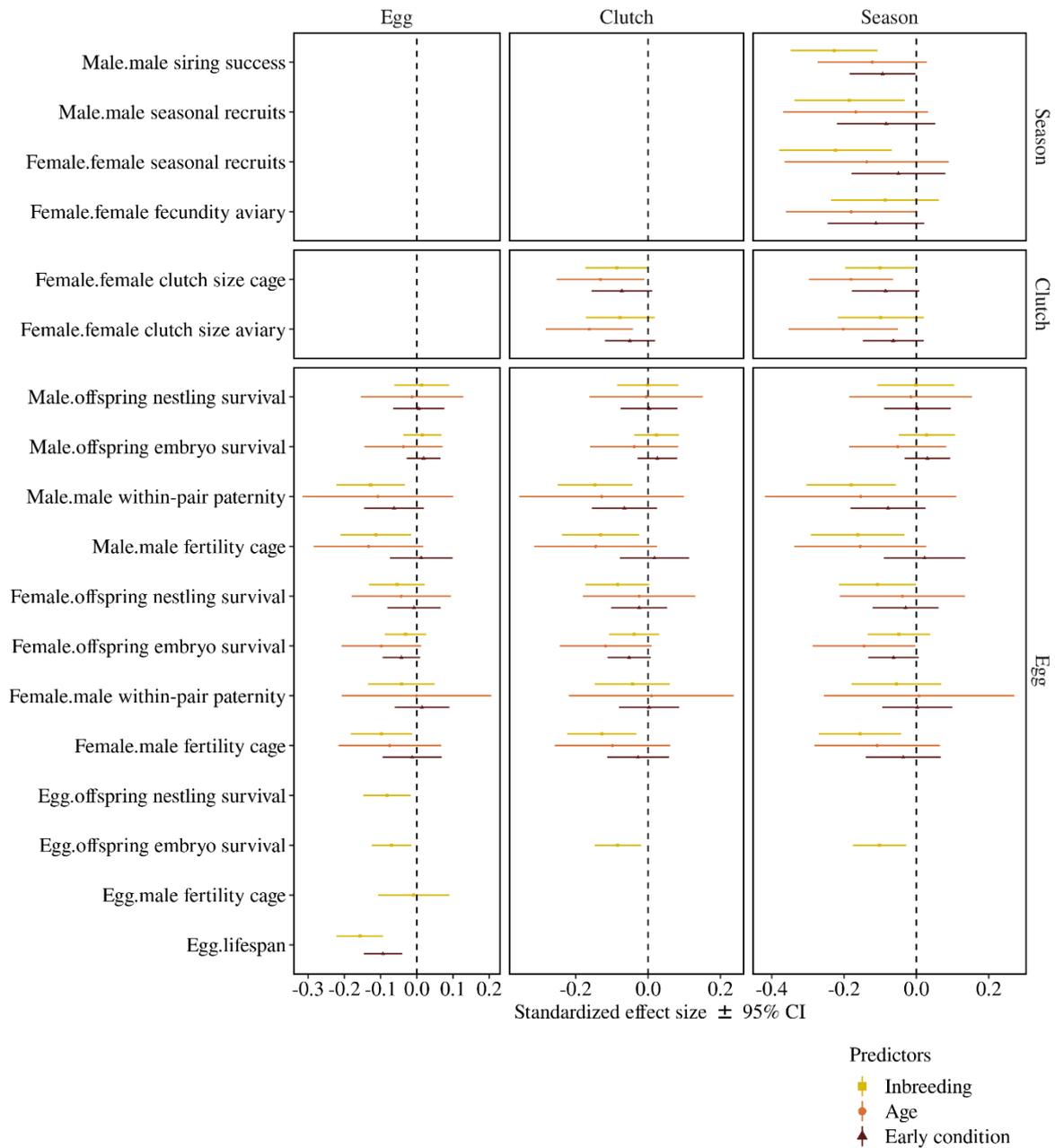
**Figure S4.** Proportion of developing eggs (fertility), surviving embryos and nestlings in relation to the laying position within a clutch (yellow), the hatching order in a brood (orange) and the clutch order within a breeding season (green) estimated in univariate Gaussian mixed-effect models where all response variables were measured at the egg level, and all covariates were Z-scaled for illustration. Error bars around the estimates are 95% CIs. Effects are shown relative to the baseline, i.e., position ‘1’, shown as a horizontal line. The 95% CIs of laying position ‘1’ were wide, and hence not shown in the figure. The 95% CIs were estimated while controlling for multiple testing using the function ‘glht’ from the R package ‘multcomp’ V1.4-13, whereas P values for a difference from the baseline were calculated as  $e(-0.717 \times |Z| - 0.416 \times Z \times Z)$ . Therefore, the 95% CIs can sometimes overlap with the baseline value, despite  $P < 0.05$ . ‘+’  $P < 0.1$ , ‘\*’  $P < 0.05$ , ‘\*\*’  $P < 0.001$ , ‘\*\*\*’  $P < 0.0001$ . Bar width reflects sample size. See table S3 for details of the models where all covariate and dependent variables were Z-scaled.



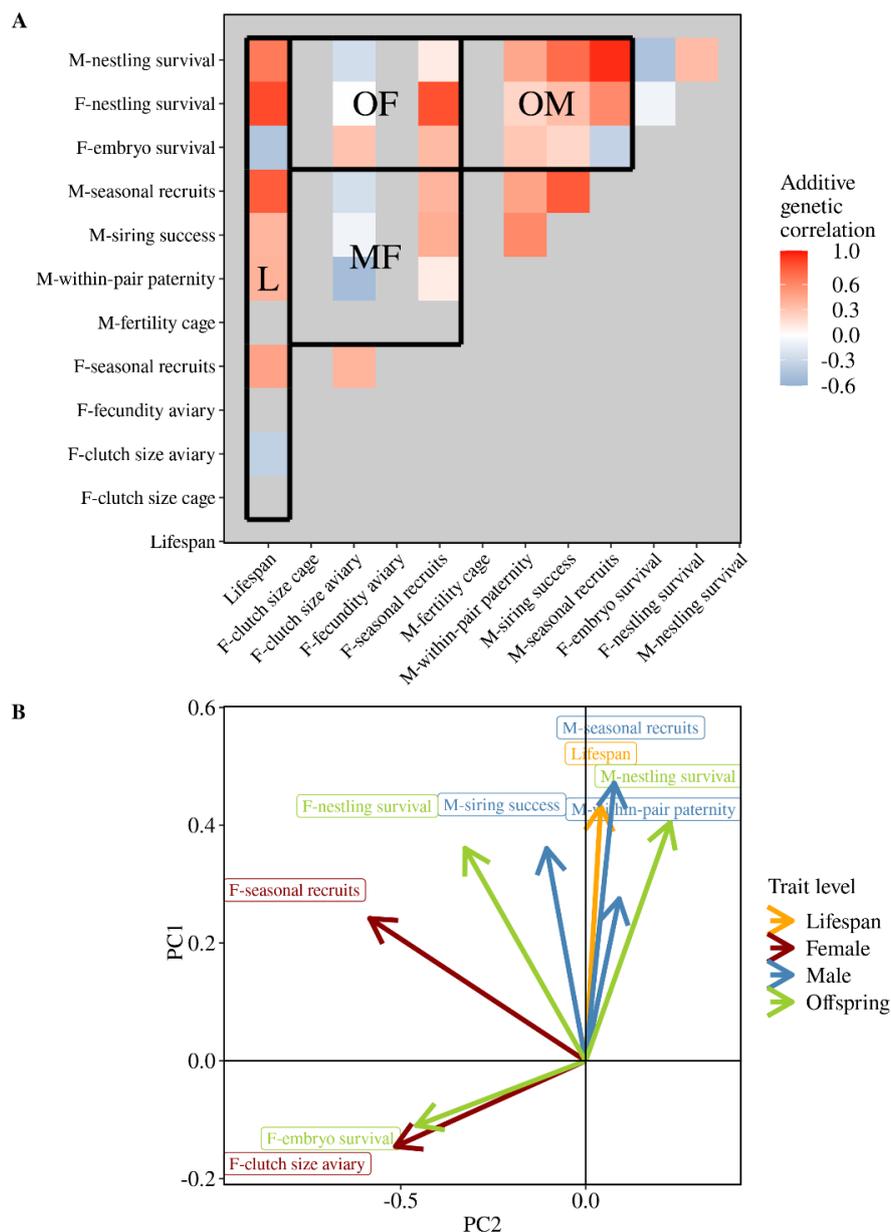
**Figure S5.** Effects ( $\pm 95\%$  CI) of egg volume on the probability that an egg will be fertilized by the social partner (fertility in cages, within-pair paternity in aviaries) and on the probability of offspring survival before (embryo survival) and after hatching (nestling survival). Variation in egg volume is split into variation between different females (Between female) and variation within females (Within female). Estimates are from univariate Gaussian mixed-effect models where all response variables were measured at the egg level (no data aggregation), and all measurements were Z-scaled. The X-axis indicates effect sizes in the form of Pearson correlation coefficients (see table S3 for details).



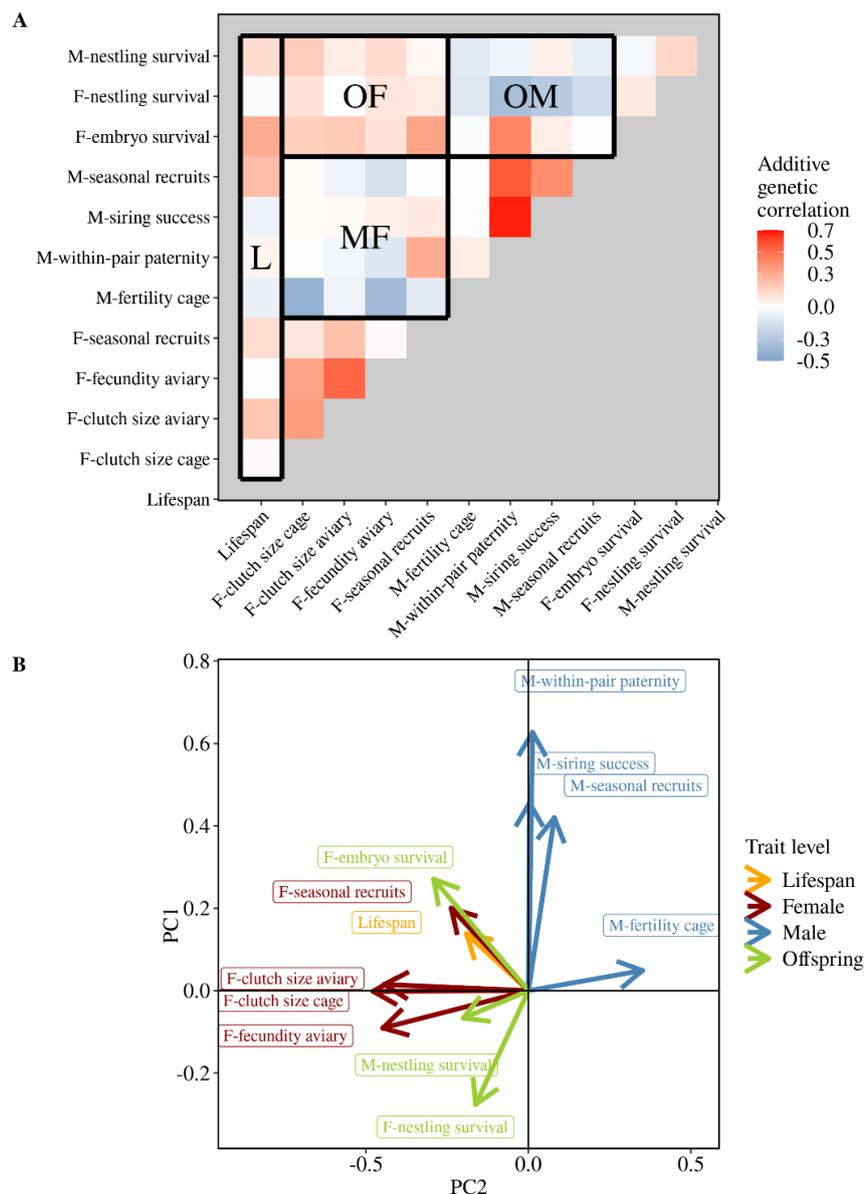
**Figure S6.** Standardized effect sizes with their 95% CIs for inbreeding ( $F_{ped}$ ), age and early condition (mass at day 8) on fitness components estimated in univariate Gaussian mixed-effect models (focal fixed effect estimates of Step 1 in fig. 1). Response variables were measured either at the level of single eggs (left panel), or at the clutch level (centre), or at the level of individuals within seasons (right panel), such that there is increasing aggregation of data from the left to the right. All measurements were Z-scaled, such that the X-axis indicates effect sizes in the form of Pearson correlation coefficients (see table S3 for details).



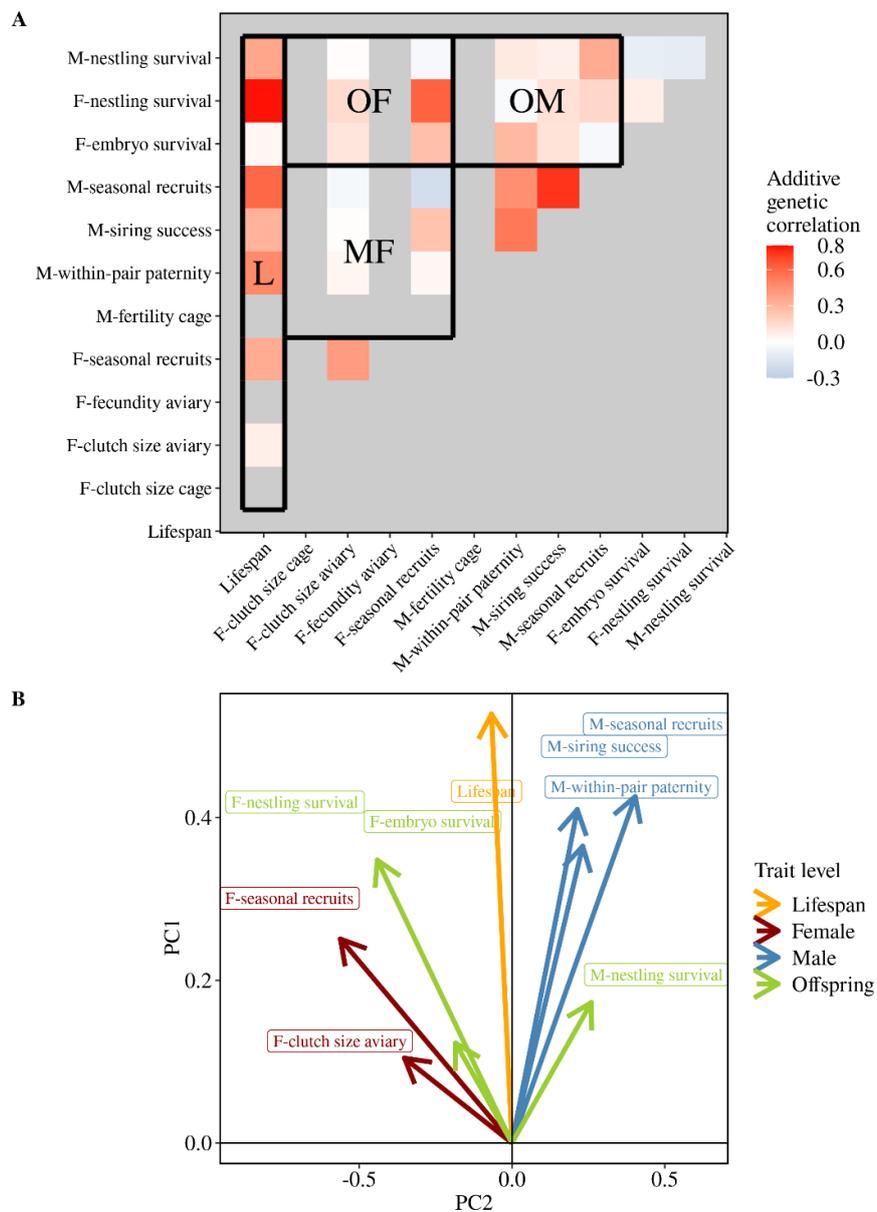
**Figure S7.** G-matrix of reproductive performance traits and lifespan estimated from multivariate animal models for the Bielefeld population (shown are estimates from VCE; table S11; Step 4 in fig. 1). (A) Heatmap of additive genetic correlations between components of male (M), female (F), and offspring (O) fitness, and life span (L). Note that some columns and rows of the matrix are empty because they were not measured for this population (no ‘cage’-breeding). Red indicates a positive genetic correlation between traits while blue indicates a negative correlation. Blocks marked in bold emphasize correlations between categories (e.g. MF stands for correlations between male and female fitness components). (B) The first two principal components of the G-matrix, showing eigenvectors of the 9 fitness components. The proportions of variance explained by PC1 and PC2 are 62% and 9%, respectively. Note that aspects of male fitness do not align with aspects of female fitness.



**Figure S8.** G-matrix of reproductive performance traits and lifespan estimated from bivariate animal models for the Seewiesen population (shown are estimates from MCMCglmm; table S10; repeated estimates of Step 4 in fig. 1). (A) Heatmap of additive genetic correlations between components of male (M), female (F), and offspring (O) fitness, and life span (L). Red indicates a positive genetic correlation between traits while blue indicates a negative correlation. Blocks marked in bold emphasize correlations between categories (e.g. MF stands for correlations between male and female fitness components). (B) The first two principal components of the G-matrix, showing eigenvectors of the 12 fitness components. The proportions of variance explained by PC1 and PC2 are 45% and 19%, respectively. Note that aspects of male fitness do not align with aspects of female fitness and offspring fitness.



**Figure S9.** G-matrix of reproductive performance traits and lifespan estimated from bivariate animal models for the Bielefeld population (shown are estimates from MCMCglmm; table S11; repeated estimates of Step 4 in fig. 1). (A) Heatmap of additive genetic correlations between components of male (M), female (F), and offspring (O) fitness, and life span (L). Note that some columns and rows of the matrix are empty because they were not measured for this population (no ‘cage’-breeding). Red indicates a positive genetic correlation between traits while blue indicates a negative correlation. Blocks marked in bold emphasize correlations between categories (e.g. MF stands for correlations between male and female fitness components). (B) The first two principal components of the G-matrix, showing eigenvectors of the 9 fitness components. The proportions of variance explained by PC1 and PC2 are 29% and 25%, respectively. Note that aspects of male fitness do not align with aspects of female fitness.



**Table S1:** Description of breeding seasons in aviaries

Population	Experiment	Treatment	Setup ID	Start	End	Duration (days)	Neggs	References
Seewiesen	Aviary breeding 2005	sex ratio	sex ratio	30 August 2005	21 November 2005	83	996	Schielzeth et al. 2011
Seewiesen	Aviary breeding 2006	sex ratio	sex ratio	31 March 2006	26 June 2006	87	1060	Schielzeth et al. 2011
Seewiesen	Aviary breeding 2007	inbreeding	inbreeding	11 September 2007	11 December 2007	91	606	Bolund et al. 2010
Seewiesen	Aviary laying 2008	dumping females	dumping females	02 December 2008	23 February 2009	83	490	Bolund et al. 2012
Seewiesen	Aviary breeding 2009	husk-ad libitum	husk-ad libitum	07 April 2009	28 July 2009	112	439	Bolund et al. 2012
Seewiesen	Aviary laying 2011	Assortative pairing	Assortative pairing	10 May 2011	08 July 2011	59	141	Unpublished
Seewiesen	Aviary laying 2014	SelectionLines S3 I-1	SelectionLines S3	27 January 2014	13 March 2014	45	539	Wang et al. 2017
Seewiesen	Aviary laying 2014	SelectionLines S3 I-2	SelectionLines S3	24 March 2014	08 May 2014	45	561	Wang et al. 2017
Seewiesen	Aviary laying 2014	SelectionLines S3 II-1	SelectionLines S3	26 May 2014	10 July 2014	45	534	Wang et al. 2017
Seewiesen	Aviary laying 2014	SelectionLines S3 II-2	SelectionLines S3	21 July 2014	04 September 2014	45	465	Wang et al. 2017
Seewiesen	Aviary laying 2014	SelectionLines S3 III-1	SelectionLines S3	06 October 2014	20 November 2014	45	495	Wang et al. 2017
Seewiesen	Aviary laying 2014	SelectionLines S3 III-2	SelectionLines S3	01 December 2014	15 January 2015	45	548	Wang et al. 2017
Seewiesen	Aviary laying 2015	SelectionLines S3 IV-1	SelectionLines S3	09 February 2015	26 March 2015	45	454	Wang et al. 2017
Seewiesen	Aviary laying 2015	SelectionLines S3 IV-2	SelectionLines S3	06 April 2015	21 May 2015	45	445	Wang et al. 2017
Seewiesen	Aviary laying 2016	S3 Song Mate Choice 1	S3 Song Mate Choice	11 January 2016	15 February 2016	35	162	<a href="https://osf.io/yzpm6">https://osf.io/yzpm6</a>
Seewiesen	Aviary laying 2016	S3 Song Mate Choice 2	S3 Song Mate Choice	22 February 2016	28 March 2016	35	177	<a href="https://osf.io/yzpm6">https://osf.io/yzpm6</a>
Seewiesen	Aviary breeding 2016	Breeding S3 I	Breeding S3	11 April 2016	09 July 2016	89	636	Unpublished
Seewiesen	Aviary breeding 2016	Breeding S3 II	Breeding S3	05 September 2016	03 December 2016	89	460	Unpublished
Seewiesen	Aviary breeding 2017	4Pop-CrossFoster_2017	4Pop-CrossFoster_2017	08 May 2017	25 June 2017	65	471	Wang et al. In prep.
Bielefeld	Aviary breeding 2012	force-pairing for choice	force-pairing for choice	21 May 2012	21 August 2012	92	767	Ihle et al. 2015
Bielefeld	Aviary breeding 2012	force-pairing for quality	force-pairing for quality	28 May 2012	21 August 2012	85	385	Schreiber 2012
Bielefeld	Aviary breeding 2013	force-pairing for choice s2	force-pairing for choice	21 May 2013	21 August 2013	92	713	Ihle et al. 2015

Bielefeld	Aviary breeding 2016	ColourManipulation	ColourManipulation	24 May 2016	14 September 2016	113	735	Jerónimo et al. 2018
Bielefeld	Aviary breeding 2017	4Pop-CrossFoster_2017	4Pop-CrossFoster_2017	08 May 2017	25 June 2017	48	437	Wang et al. In prep.
Krakow	Aviary breeding 2016	Breeding_H2_ColourManipulation	ColourManipulation	24 May 2016	14 September 2016	113	781	Jerónimo et al. 2018
Krakow	Aviary breeding 2016	Breeding_H1_ColourManipulation	ColourManipulation	24 May 2016	14 September 2016	113	552	Jerónimo et al. 2018
Krakow	Aviary breeding 2017	4Pop-CrossFoster_2017	4Pop-CrossFoster_2017	08 May 2017	25 June 2017	48	280	Wang et al. In prep.
Melbourne	Aviary breeding 2016	ColourManipulation	ColourManipulation	24 May 2016	14 September 2016	113	1341	Jerónimo et al. 2018
Melbourne	Aviary breeding 2017	4Pop-CrossFoster_2017	4Pop-CrossFoster_2017	08 May 2017	25 June 2017	48	377	Wang et al. In prep.

Note: Summarized from the table of female clutch size aviary in the supporting data. 'Experiment' and 'Treatment' are names used in the raw database. In 'Experiment', 'laying' and 'breeding' refers to aviary settings as described in the main text. Neggs: total number of eggs laid, including eggs that were broken and with no parentage assignment.

#### References:

- Bolund, E., K. Martin, B. Kempnaers, and W. Forstmeier. 2010. Inbreeding depression of sexually selected traits and attractiveness in the zebra finch. *Animal Behaviour* 79:947–955.
- Bolund, E., H. Schielzeth, and W. Forstmeier. 2012. Singing activity stimulates partner reproductive investment rather than increasing paternity success in zebra finches. *Behavioral Ecology and Sociobiology* 66:975–984.
- Ihle, M., B. Kempnaers, and W. Forstmeier. 2015. Fitness Benefits of Mate Choice for Compatibility in a Socially Monogamous Species. *PLoS Biology* 13:1–21.
- Jerónimo, S., M. Khadraoui, D. Wang, K. Martin, J. A. Lesku, K. A. Robert, E. Schlicht, et al. 2018. Plumage color manipulation has no effect on social dominance or fitness in zebra finches. *Behavioral Ecology*.
- Schielzeth, H., E. Bolund, B. Kempnaers, and W. Forstmeier. 2011. Quantitative genetics and fitness consequences of neophilia in zebra finches. *Behavioral Ecology* 22:126–134.
- Wang, D., W. Forstmeier, K. Martin, A. Wilson, and B. Kempnaers. 2020. The role of genetic constraints and social environment in explaining female extra-pair mating. *Evolution* evo.13905.
- Wang, D., N. Kempnaers, B. Kempnaers, and W. Forstmeier. 2017. Male zebra finches have limited ability to identify high-fecundity females. *Behavioral Ecology* 28:784–792.
- Schreiber Johannes. 2012. Effects of variation in genetic quality on sexual behaviour and fitness in zebra finches (*Taeniopygia guttata*). Master Thesis University Oldenburg.

**Table S2:** Description of breeding seasons in cages

Population	Experiment	Treatment or setup ID	Start	End	Duration (days)	Neggs	References
Seewiesen	Cage laying Sheffield	box	23 February 2003	17 November 2003	267	1000	Unpublished
Seewiesen	Cage breeding 2004	breeding F1	06 February 2004	29 July 2004	174	1621	Forstmeier 2005
Seewiesen	Cage breeding 2005	Foster pairs	28 June 2005	23 December 2005	178	1466	Schielzeth et al. 2011
Seewiesen	Cage breeding 2005	Genetic diversity	28 June 2005	14 December 2005	169	112	Unpublished
Seewiesen	Cage laying 2005	SU16 phase 1	08 October 2005	22 December 2005	75	233	Bolund et al. 2009, 2012
Seewiesen	Cage laying 2005	SU16 phase 2	09 January 2006	17 March 2006	67	210	Bolund et al. 2009, 2012
Seewiesen	Cage breeding 2006	Foster pairs	28 March 2006	27 July 2006	121	771	Schielzeth et al. 2011
Seewiesen	Cage breeding 2006	former undirected song pa	28 March 2006	07 July 2006	101	192	Unpublished
Seewiesen	Cage laying 2007	Differential Allocation 1	30 January 2007	13 March 2007	42	270	Bolund et al. 2009
Seewiesen	Cage laying 2007	Differential Allocation 2	04 May 2007	13 June 2007	40	251	Bolund et al. 2009
Seewiesen	Cage laying 2007	perevitalline layer	31 January 2007	19 December 2007	322	151	Unpublished
Seewiesen	Cage laying 2007	song recording	14 March 2007	03 May 2007	50	225	Unpublished
Seewiesen	Cage laying 2008	matched song	21 October 2008	20 February 2009	122	535	Unpublished
Seewiesen	Cage laying 2008	hormone pairs	15 October 2008	02 December 2008	48	219	Unpublished
Seewiesen	Cage laying 2008	Song similarity	13 March 2008	06 May 2008	54	346	Unpublished
Seewiesen	Cage laying 2009	undirected song F3	05 February 2009	23 December 2009	321	619	Bolund et al. 2010
Seewiesen	Cage breeding 2009	SelectionLines P-1	02 June 2009	25 November 2009	176	1471	Wang et al. 2020
Seewiesen	Cage breeding 2010	SelectionLines P-2	04 January 2010	21 July 2010	198	1312	Wang et al. 2020
Seewiesen	Cage breeding 2011	SelectionLines S1-1	25 January 2011	27 May 2011	122	1165	Wang et al. 2020
Seewiesen	Cage breeding 2011	SelectionLines S1-2	19 July 2011	28 November 2011	132	945	Wang et al. 2020
Seewiesen	Cage laying 2012	LineDifferences	21 February 2012	05 April 2012	44	132	Unpublished
Seewiesen	Cage laying 2012	ClutchSize_S1	22 March 2012	11 September 2012	173	741	Unpublished
Seewiesen	Cage breeding 2012	SelectionLines S2-1	21 June 2012	11 December 2012	173	1351	Wang et al. 2020
Seewiesen	Cage breeding 2013	SelectionLines S2-2	01 February 2013	01 July 2013	150	1312	Wang et al. 2020
Seewiesen	Cage laying 2013	ClutchSize_S2	10 June 2013	07 August 2013	58	316	Unpublished
Seewiesen	Cage laying 2013	MeioticDrive_Tgu2	24 October 2013	20 May 2014	208	701	Knief et al. 2015
Seewiesen	Cage laying 2014	Vitamins	21 March 2014	12 May 2014	52	207	Unpublished

Bielefeld	Aviary breeding 2011	Breeding pairs F1-F2	01 April 2011	14 November 2011	227	906	Ihle et al. 2013
Krakow	Cage breeding 2012	Breeding_Krakau_P	09 January 2012	17 February 2012	39	77	Unpublished
Krakow	Cage breeding 2012	Breeding_KS-Cross_P-1	23 March 2012	25 June 2012	94	158	Unpublished
Krakow	Cage breeding 2012	Breeding_KS-Hybrids_H1	10 October 2012	17 February 2013	130	618	Unpublished
Krakow	Cage breeding 2013	Breeding_KS-Cross_P-2	24 October 2013	11 February 2014	110	265	Unpublished
Krakow	Cage breeding 2014	Breeding_Krakau_P-2	11 March 2014	14 May 2014	64	95	Unpublished
Krakow	Cage breeding 2014	Breeding_Seewiesen_P-2	11 March 2014	25 April 2014	45	66	Unpublished
Krakow	Cage breeding 2016	Breeding_H2_for Lund	13 June 2016	09 December 2016	179	437	Unpublished

Note: 'Experiment' and 'Treatment or setup ID' are names used in the raw database. In 'Experiment', 'laying' and 'breeding' refers to cage settings as described in the main text. The experiment of 'Aviary breeding 2011' for population Bielefeld where single pairs were breeding in separate aviaries such that the social environment was identical to the 'cage' setup, except that there was more space available. Neggs: total number of eggs laid, including eggs that were broken and with no parentage assignment.

#### References:

- Bolund, E., K. Martin, B. Kempenaers, and W. Forstmeier. 2010. Inbreeding depression of sexually selected traits and attractiveness in the zebra finch. *Animal Behaviour* 79:947–955.
- Bolund, E., H. Schielzeth, and W. Forstmeier. 2009. Compensatory investment in zebra finches: Females lay larger eggs when paired to sexually unattractive males. *Proceedings of the Royal Society B: Biological Sciences* 276:707–715.
- Bolund, E., H. Schielzeth, and W. Forstmeier. 2012. Singing activity stimulates partner reproductive investment rather than increasing paternity success in zebra finches. *Behavioral Ecology and Sociobiology* 66:975–984.
- Forstmeier, W. 2005. Quantitative genetics and behavioural correlates of digit ratio in the zebra finch. *Proceedings of the Royal Society B: Biological Sciences* 272:2641–2649.
- Ihle, M., B. Kempenaers, and W. Forstmeier. 2013. Does hatching failure breed infidelity? *Behavioral Ecology* 24:119–127.
- Knief, U., H. Schielzeth, H. Ellegren, B. Kempenaers, and W. Forstmeier. 2015. A prezygotic transmission distorter acting equally in female and male zebra finches *Taeniopygia guttata*. *Molecular Ecology* 24:3846–3859.
- Schielzeth, H., E. Bolund, B. Kempenaers, and W. Forstmeier. 2011. Quantitative genetics and fitness consequences of neophilia in zebra finches. *Behavioral Ecology* 22:126–134.
- Wang, D., W. Forstmeier, K. Martin, A. Wilson, and B. Kempenaers. 2020. The role of genetic constraints and social environment in explaining female extra-pair mating. *Evolution* evo.13905.

**Table S3:** Fixed effect estimates from univariate Gaussian models (Step 1 in fig. 1)

This table contains 301 rows.

**Table S4:** Variance components from univariate Gaussian mixed effect models (Step 1 in fig. 1)

This table contains 113 rows.

**Table S5:** Meta-summarization model output of the focal fixed effect estimates (Step 1 in fig. 1)

This table contains 8 rows.

**Table S6:** Heritability estimates of reproductive performance traits based on phenotypic variation and permanent environmental variation, using raw trait estimates at season level in the Seewiesen population (Step 2 in fig. 1)

This table contains 16 rows.

**Table S7:** Heritability estimates of reproductive performance traits based on phenotypic variation and permanent environmental variation, using raw trait estimates at season level in the Bielefeld population (Step 2 in fig. 1)

This table contains 16 rows.

**Table S8:** Weighted linear (mixed-effect) model outputs of meta-summarized comparison of heritabilities between populations Seewiesen and Bielefeld (Step 2 in fig. 1)

This table contains 7 rows.

**Table S9:** Meta-summarized additive genetic correlations in 12-trait (Seewiesen) and 9-trait (Bielefeld) multivariate animal models (VCE6) and bivariate animal models (MCMCglmm; Step 4 in fig. 1)

This table contains 29 rows.

**Table S10:** Additive genetic correlation estimates of reproductive performance traits and lifespan based on BLUPs (best linear unbiased predictions) in the Seewiesen population (Step 4 in fig. 1)

This table contains 67 rows.

**Table S11:** Additive genetic correlation estimates of reproductive performance traits and lifespan based on BLUPs (best linear unbiased predictions) in the Bielefeld population (Step 4 in fig. 1)

This table contains 38 rows.

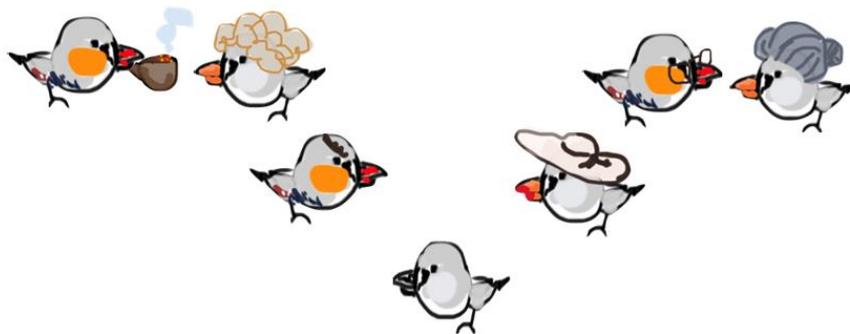
For details of Tables S3-S11 please see online Supplementary Material:

<https://www.journals.uchicago.edu/doi/suppl/10.1086/710956>



## Chapter 2

### Transgenerational effects



"I'm not doing so well, is it because of my parents and grandparents? "

## Chapter 2

### Offspring performance is well buffered against stress experienced by ancestors

Yifan Pei, Wolfgang Forstmeier, Bart Kempenaers

Evolution should render individuals resistant to stress and particularly to stress experienced by ancestors. However, many studies report negative effects of stress experienced by one generation on the performance of subsequent generations. To assess the strength of such transgenerational effects we propose a strategy aimed at overcoming the problem of type I errors when testing multiple proxies of stress in multiple ancestors against multiple offspring performance traits, and we apply it to a large observational data set on captive zebra finches (*Taeniopygia guttata*). We combine clear one-tailed hypotheses with steps of validation, meta-analytic summary of mean effect sizes, and independent confirmatory testing. We find that drastic differences in early growth conditions (nestling body mass 8 days after hatching varied 7-fold between 1.7 and 12.4 gram) had only moderate direct effects on adult morphology (95%CI:  $r=0.19-0.27$ ) and small direct effects on adult fitness traits ( $r=0.02-0.12$ ). In contrast, we found no indirect effects of parental or grandparental condition ( $r=-0.017-0.002$ ; meta-analytic summary of 138 effect sizes), and mixed evidence for small benefits of matching environments between parents and offspring, as the latter was not robust to confirmatory testing in independent data sets. This study shows that evolution has led to a remarkable robustness of zebra finches against undernourishment. Our study suggests that transgenerational effects are absent in this species, because confidence intervals exclude all biologically relevant effect sizes.

#### Published as:

Pei, Y., Forstmeier, W., & Kempenaers, B. (2020). Offspring performance is well buffered against stress experienced by ancestors. *Evolution*, 74(7), 1525-1539.

# Offspring performance is well buffered against stress experienced by ancestors

Yifan Pei,<sup>1</sup>  Wolfgang Forstmeier,<sup>1,2</sup>  and Bart Kempenaers<sup>1</sup> 

<sup>1</sup>Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen 82319, Germany

<sup>2</sup>E-mail: forstmeier@orn.mpg.de

Received March 30, 2020

Accepted May 17, 2020

Evolution should render individuals resistant to stress and particularly to stress experienced by ancestors. However, many studies report negative effects of stress experienced by one generation on the performance of subsequent generations. To assess the strength of such transgenerational effects we propose a strategy aimed at overcoming the problem of type I errors when testing multiple proxies of stress in multiple ancestors against multiple offspring performance traits, and we apply it to a large observational dataset on captive zebra finches (*Taeniopygia guttata*). We combine clear one-tailed hypotheses with steps of validation, meta-analytic summary of mean effect sizes, and independent confirmatory testing. We find that drastic differences in early growth conditions (nestling body mass 8 days after hatching varied sevenfold between 1.7 and 12.4 g) had only moderate direct effects on adult morphology (95% confidence interval [CI]:  $r = 0.19$ – $0.27$ ) and small direct effects on adult fitness traits ( $r = 0.02$ – $0.12$ ). In contrast, we found no indirect effects of parental or grandparental condition ( $r = -0.017$  to  $0.002$ ; meta-analytic summary of 138 effect sizes), and mixed evidence for small benefits of matching environments between parents and offspring, as the latter was not robust to confirmatory testing in independent datasets. This study shows that evolution has led to a remarkable robustness of zebra finches against undernourishment. Our study suggests that transgenerational effects are absent in this species, because CIs exclude all biologically relevant effect sizes.

**KEY WORDS:** Anticipatory effect, condition transfer, early developmental stress, life span, morphology, multiple testing, quantitative genetics, reproductive performance, resilience, transgenerational effect.

Early developmental stress experienced by an individual may have long-term negative effects on its morphology, physiology, behavior and reproductive performance later in life (i.e., “direct effect”; Lindström 1999; Tschirren et al. 2009; Bolund et al. 2010; Boersma et al. 2014; Eyck et al. 2019; Kraft et al. 2019; Pei et al. 2019). Effects of conditions in early life could also be transmitted to subsequent generations (Marshall and Uller 2007; Monaghan 2008; Engqvist and Reinhold 2016, 2018; Bonduriansky and Crean 2018), potentially via the inheritance of epigenetic markers (e.g., altered DNA methylation, transmission of small interference RNAs, or hormones; Holliday 1987; Colborn et al. 1993; Jablonka and Raz 2009). Such inheritance of acquired traits could exist either because of an inevitable transfer of condition from one generation to the next (i.e., “condition transfer,” “carry-over,” or “silver-spoon” effect,

e.g., low-condition mothers produce low-condition offspring; Bonduriansky and Head 2007; Hettinger et al. 2012; Franzke and Reinhold 2013; Burton and Metcalfe 2014; Bonduriansky and Crean 2018), or because organisms have evolved mechanisms of adaptive transgenerational plasticity, where offspring were “primed” by their parents and perform the best if they grow up in an environment similar to that of their parents (i.e., “anticipatory effect,” hypothesis of matching/mismatching environments; Krause and Naguib 2014; Engqvist and Reinhold 2016; Raveh et al. 2016). It is difficult to distinguish between the two types of transgenerational effects, that is, (1) “condition transfer” and (2) “anticipatory effects,” in experiments (e.g., match/mismatch), especially when the unavoidable intragenerational, (3) “direct effects” of early conditions experienced by the individual itself are substantial (Hettinger et al. 2012; Dey et al. 2016;

Engqvist and Reinhold 2016, 2018; Bonduriansky and Crean 2018).

From an evolutionary perspective, we would expect that natural selection acts to minimize the susceptibility of organisms to harmful direct and indirect, condition-transfer effects. Fitness-related traits in particular are selected to be well buffered against detrimental influences from the environment (evolution of stress tolerance, robustness, and developmental canalization; e.g., Waddington 1942; Siegal and Bergman 2002). Moreover, selection will disfavor mothers that handicap their own offspring. In general, detrimental carry-over effects may be inevitable to some extent, but selection will work against them. In contrast, “transgenerational anticipatory effects” are thought to have evolved for an adaptive function. Such “transgenerational anticipatory programming” of offspring may have evolved when the environments in which parents and offspring grow up are generally similar (e.g., Krause and Naguib 2014; Raveh et al. 2016), and when proximate mechanisms of epigenetic inheritance enable it (e.g., Holliday 1987; Colborn et al. 1993; Jablonka and Raz 2009; but see also Proulx and Teotónio 2017 for alternative scenarios for the evolution of adaptive anticipatory effects). Studies of epigenetic inheritance boomed since the early 1990s (Jablonka and Raz 2009; Jensen 2013), focusing mostly on organisms that are immobile or lack differentiation between soma and germ cells, such as fungi (Benkemon and Saupé 2006), plants (Cubas et al. 1999; Molinier et al. 2006; Henderson and Jacobsen 2007; Feng 2010), and nematodes (Bagijn et al. 2012; Rechavi et al. 2014; Dey et al. 2016; Lev et al. 2019). Meanwhile, sexually reproducing animals, such as fruit flies (Magiafoglou and Hoffmann 2003), birds (Naguib and Gil 2005; Monaghan 2008; Khan et al. 2016), mice (Morgan et al. 1999; Carone et al. 2010), rats (Anway et al. 2005), and humans (Colborn et al. 1993) also became popular study subjects for epigenetic inheritance. However, in the latter group, we still lack studies that show mechanistically how experiences made by the soma can be transferred to the germline. In sum, the widespread existence of both types of transgenerational effects seems somewhat unlikely, because condition transfer is selected against and anticipatory effects may lack a mechanism that could achieve such adaptation.

Although the mechanisms behind most of the observed epigenetic inheritance remain largely unclear (Jablonka and Raz 2009; Miska and Ferguson-Smith 2016), evolutionary biologists have studied transgenerational effects and have estimated the fitness consequence of stress experienced by one generation on individuals of subsequent generations in various animal systems (Ledón-Rettig et al. 2013; Guerrero-Bosagna et al. 2018), sometimes with individuals from the wild (Drummond and Ancona 2015), but mostly with captive-bred animals, for example (Naguib and Gil 2005; Uller et al. 2005; Alonso-Alvarez et al. 2007; Krause and Naguib 2014; Wilson et al. 2019). Such effects

have typically been investigated experimentally across two generations, that is, effects of increasing stress experienced by the parents on the offspring, using brood or litter-size manipulation (Naguib and Gil 2005; Alonso-Alvarez et al. 2007), restricted food supply during female pregnancy, or nestling or puppy stages (Bertram et al. 2008), restraint stress exposure during early life where individuals were intermittently deprived from social interactions (Goerlich et al. 2012), corticosterone intake during female pregnancy or early individual development (Khan et al. 2016), and cold or heat shock (mostly for insects, e.g., in *Drosophila* and *Tribolium*; Magiafoglou and Hoffmann 2003; Eggert et al. 2015). In general, the reported significant effects are often accompanied by numerous nonsignificant test results, and sometimes a significant effect with a sign opposite to expectations may still get interpreted as evidence for the existence of transgenerational effects. Moreover, transgenerational effects are sometimes being reported in a sex-specific way (interaction effect between the sex of the parent and that of the offspring). For example, in humans, effects from (grand)mother to (grand)daughters and from (grand)father to (grand)sons have been reported (Pembrey et al. 2006; Kaati et al. 2007). When studies examine multiple predictors and response variables in multiple ancestors, there is a risk of selective reporting of the strongest effects. Unbiased estimates can only be obtained when including all predictors and responses that appeared worth investigating at the start of a study (or when subsetting is not conditional on the results). Hence, for assessing the importance of transgenerational effects, we suggest that rigorous testing of one-tailed a priori hypotheses and meta-analytic summary of effect sizes is essential.

Here, we use observational data of more than 2000 captive zebra finches from a long-term, error-free pedigree to study the sex-specific effects of multiple stressors experienced during early development on later-life morphology and fitness-related traits. We consider both direct, intragenerational effects and effects of developmental stress experienced by parental and grandparental generations. For all individuals, we systematically recorded variables that have previously been used as indicators of early developmental conditions: brood size (Koskela 1998; Naguib and Gil 2005; Tschirren et al. 2009), hatching order (Saino et al. 2001; Ferrari et al. 2006; Wilson et al. 2019), laying order of eggs (Soma et al. 2007; Gilby et al. 2012), and clutches (Tomita et al. 2011), egg volume (Love and Williams 2011), and nestling body mass at 8 days old (Bolund et al. 2010). We also used five morphological traits as dependent variables: tarsus (Naguib and Gil 2005; Tschirren et al. 2009) and wing length (Naguib and Gil 2005; Krause and Naguib 2014; Wilson et al. 2019), body mass (Tschirren et al. 2009; Krause and Naguib 2014; Wilson et al. 2019), abdominal fat deposition (Bolund et al. 2010), and beak color (Tschirren et al. 2009; Bolund et al. 2007, 2010; Wilson et al. 2019). For a subset of birds, we also measured life span

and aspects of reproductive performance (female clutch size in cages and in aviaries, female fecundity, male infertility in cages, male within-pair paternity, male siring success, female embryo mortality, nestling mortality for a given social mother, and for a given social father, female, and male seasonal recruits (for details, see the “Methods” section), following, for example (Naguib et al. 2006; Tschirren et al. 2009; Krause and Naguib 2014; Khan et al. 2016; Wilson et al. 2019).

Regarding condition transfer, we focus our analyses on the a priori hypothesis that the stress that an individual’s parents and grandparents experienced in early life has detrimental effects on the morphology and reproductive performance of that individual as an adult. We assume that the direction of effects is independent of the sex of the focal individual. We further hypothesize a priori that if such transgenerational effects were sex-specific (e.g., epigenetics of sex chromosome by environment interaction), the environment experienced by mothers and grandmothers would affect daughters and granddaughters whereas the environment experienced by fathers and grandfathers would affect sons and grandsons (Pembrey et al. 2006; Kaati et al. 2007). Such one-tailed expectations have the advantage that trends which are opposite to the expectation can be quantified as negative effect sizes. If the null hypothesis is true, that is, if there is no effect, we expect a meta-analytic mean effect size that does not differ from zero.

Regarding anticipatory effects, we focus on the a priori, one-tailed hypothesis that offspring perform better as adults when they experienced similar early-life growth conditions as their parents did (measured as nestling body mass at 8 days old). Note that this only concerns variation in individual growth conditions, while the captive environment (aviary or laboratory) provides relatively stable conditions (Kuijper et al. 2014; Kuijper and Johnstone 2016).

First, we validate the six proxies of early developmental stress by examining their direct effects on the individual itself. Second, we use meta-analysis to average transgenerational effect sizes across multiple traits reflecting either morphology or reproductive performance of adult male and female zebra finches. Lastly, we use an independent dataset (i.e., additional birds from populations with shorter pedigrees, but otherwise equal data quality), to assess whether the significant findings from the initial tests can be replicated.

## Methods

### STUDY SYSTEM AND GENERAL PROCEDURES

The zebra finch is an abundant, opportunistic breeder in Australia in the wild (Zann 1996) that also breeds easily in captivity. We used birds from a domesticated zebra finch population with a 13-

generation error-free pedigree, maintained at the Max Planck Institute for Ornithology, Seewiesen, Germany (#18 in Forstmeier et al. 2007). Housing conditions have been described elsewhere (see Bolund et al. 2007; Ihle et al. 2015; Pei et al. 2019). The study was conducted under license (permit number 311.4-si and 311.5-gr, Landratsamt Starnberg, Germany).

We used all individuals ( $N = 2099$ ) from this study population for which complete information was available on laying and hatching dates, egg volume, and nestling mass at 8 days of age, both from the focal individual itself, but also from its nearest six ancestors (parents and grandparents). Breeding experiments were conducted either in cages with single pairs whereby the partners were assigned to each other, or in semioutdoor aviaries with groups of females and males whereby birds freely formed breeding pairs. During episodes of breeding, nests were checked daily on weekdays and occasionally during weekends to collect the required data (see below).

### PROXIES OF EARLY DEVELOPMENTAL STRESS

We examined six parameters that have been used in previous studies as potential proxies of nutritional status or stress experienced during early development: (i) the laying order of eggs within a clutch (range: 1–18, mean = 3.1, SD = 1.7, note that only five birds hatched from eggs with laying order > 10; a clutch was defined as eggs that were laid consecutively by a focal female allowing for laying gaps of maximally 4 days between subsequent eggs; Soma et al. 2007; Gilby et al. 2012), (ii) the order of clutches laid within a breeding season (range: 1–8, mean = 2.1, SD = 1.2; Tomita et al. 2011), (iii) the order of hatching within a brood (range: 1–6, mean = 2.1, SD = 1.1; Saino et al. 2001; Ferrari et al. 2006), (iv) brood size (number of nestlings reaching 8 days of age (range: 1–6, mean = 3.3, SD = 1.2; Koskela 1998; Tschirren et al. 2009), (v) relative egg volume (i.e., centered to the mean egg volume laid by a given female; range: –0.26 to 0.30, mean = 0.01, SD = 0.07; Love and Williams 2011), and (vi) nestling body mass at 8 days of age (range: 1.7–12.4 g, mean = 7.2 g, SD = 1.6; Bolund et al. 2010). Egg volume was calculated as  $V = (1/6) \times \pi \times \text{Width}^2 \times \text{Length}$ , whereby egg length and width were measured to the nearest 0.1 mm. Note that the first five parameters describe developmental conditions that are beyond the control of the developing organism, while the last one (nestling mass) is a trait that also depends on genetic variation in growth rate. In our population, a cross-fostering experiment revealed that nestling mass at day 8 has a heritability of 13% (Bolund et al. 2010). Thus, zebra finch nestling mass primarily reflects environmental conditions experienced by the individual during early growth. We decided to measure nestling mass only at 8 days of age, because at that age nestlings on average reach about half of their final mass, and hence we expected variation due to extrinsic growth conditions to be maximal.

We hypothesized that individuals (or their parents and grandparents) developed under more stressful conditions if they came from eggs later in the laying, clutch or hatching sequence, were raised in a larger brood, hatched from an egg that was relatively small, and had a lower body mass at 8 days of age. For the measure of similarity of parent-offspring early developmental condition (predictor of “anticipatory effect”), we calculated the absolute difference in nestling mass at 8 days between parent (mother or father) and offspring (mother-offspring range: 0–8.4 g, mean = 1.8 g, SD = 1.4; father-offspring range: 0–7.9 g, mean = 1.7 g, SD = 1.3).

To aid interpretation, we scored all stressors in such a way that all estimated effects are expected to be positive (multiplication by  $-1$  where necessary). Thus, positive effect sizes indicate detrimental effects of a stressor on a trait.

### ADULT PERFORMANCE TRAITS

We studied the following morphological traits, measured when the individual reached adulthood (median = 115 days of age, range 93–229 days, >95% of birds were 100–137 days old): (i) body mass (measured to the nearest 0.1 g using a digital scale,  $N = 947$  females and 1012 males), (ii) length of the right tarsus (measured from the bent foot to the rear edge of the tarsometatarsus, including the joint, using a wing ruler to the nearest 0.1 mm,  $N = 944$  females and 1008 males; see method 3 in Forstmeier et al. 2007), (iii) length of the flattened right wing (measured with a wing ruler to the nearest 0.5 mm,  $N = 939$  females and 1004 males), (iv) visible clavicular and abdominal fat deposition, scored from 0 to 5 in 0.5 increments ( $N = 932$  females and 989 males), and (v) redness of the beak (Bolund et al. 2007), scored by comparison to a color standard following the Munsell color scale from 0 to 5.5 in 0.1 increments ( $N = 947$  females and 1012 males). Male and female traits were analyzed separately, leading to a total of 10 morphological traits.

We also studied the following 13 fitness-related traits (data taken from Pei 2020b): (i) female clutch size measured in cages ( $N = 166$  females) or (ii) in aviaries ( $N = 274$  females); (iii) female fecundity, that is, total number of eggs laid in aviaries without nestling rearing ( $N = 230$  females); (iv) male infertility, measured in cages as the proportion of nondeveloping eggs ( $N = 132$  males); (v) male within-pair paternity, measured in aviaries as the proportion of eggs fertilized by the social male ( $N = 237$  males); (vi) male siring success, measured in aviaries as the total number of eggs sired (within and extra pair;  $N = 281$  males); (vii) female embryo mortality, measured as the proportion of a genetic mother’s embryos dying ( $N = 228$  genetic mothers); (viii) nestling mortality, measured as the proportion of hatchlings in a brood that died before day 35, for a given social mother ( $N = 233$ ); and (ix) for a given social fathers ( $N = 228$ ); (x) female and (xi) male seasonal recruits as the total number of independent

offspring produced (defined as offspring that survived until day 35;  $N = 126$  males and  $N = 125$  females); (xii) female ( $N = 409$ ); and (xiii) male life span ( $N = 412$ ). All measures of reproductive success in aviaries were based on genetically assigned parentage, including all dead embryos and all nestlings (see Ihle et al. 2015; Wang et al. 2017, Wang et al. 2020).

For infertility, within-pair paternity, embryo and nestling mortality, we used raw data based on the fate of single eggs, while controlling for pseudo-replication by adding male and female identities as random effects in all models (see the “Statistics” section). Female clutch size was analyzed at the clutch level, controlling for female identity, because 94% of females produced multiple clutches. For fecundity, siring success and seasonal recruits, we used the data from individuals within a given breeding season (96% of females and 78% of males had multiple measures for fecundity and siring success, while for seasonal recruits, females and males were only measured once). For easy interpretation of the results, we scored all fitness-related traits in such a way that high trait values refer to better reproductive performance (multiplication by  $-1$  where necessary).

The morphological and fitness-related traits are in general positively correlated within female and male zebra finches (Fig. S1 and Table S1).

### STATISTICS

We estimated the effect of each potential stressor experienced either by the individual itself (direct effects), or by one of its parents or grandparents (condition transfer) on each trait in a separate model (6 stressors  $\times$  7 sources  $\times$  23 traits = 966 models). To estimate anticipatory effects, we analyzed each of the 23 performance traits as a function of parent-offspring similarity in nestling mass (once for the mother, once for the father) while statistically controlling for the direct effect of the individual’s nestling mass (see the “Results” section; 2 sources  $\times$  23 traits = 46 models). We used mixed-effect models and animal models to control for the nonindependence of data points due to shared random effects including genetic relatedness. For animal models, we used the package “pedigreeMM” V0.3-3 (Vazquez et al. 2010) and for mixed-effect models we used “lme4” V1.1-23 (Bates et al. 2015) in R V4.0.0 (R Core Team 2020). The 95% CIs of estimated effect sizes were calculated using the “glht” function in the “multcomp” V1.4-13 R package while controlling for multiple testing (Hothorn et al. 2008), unless stated otherwise.

Morphological traits typically show high heritability, so we included the between-individual relatedness matrix (using pedigree information) as a random effect to control for the genetic relatedness of individuals. In contrast, fitness-related traits typically have low heritability (Pei et al. 2019), so we analyzed fitness-related traits in mixed-effect models while only controlling for repeated measurements from the same focal individual,

parent, or grandparent. To compare and summarize the effects of the variables indicating early-life conditions on different traits, we Z-scaled all dependent and all predictor variables (stressors), assuming a Gaussian distribution.

Details on model structures (see Tables S2–S4 for all fixed effects), all scripts and underlying data are provided in the Open Science Framework at <https://osf.io/wjg3q/> (Pei 2020a). In brief, for all morphological traits, we fitted sex (male and female), fostering experience (three levels: no cross-fostering, cross-fostered within or between populations), and inbreeding level (pedigree-based inbreeding coefficient,  $F_{\text{ped}}$ , where outbred birds have  $F_{\text{ped}} = 0$  and full-sib matings produce birds with  $F_{\text{ped}} = 0.25$ ) as fixed effects. For models with beak color, wing, and tarsus length as the dependent variable, we also fitted the identity of the observer that measured the trait as a fixed effect to control for between-observer variation. We included the identity of the peer group in which the individual grew up as a random effect. We fitted individual identity twice in the random structure, once linked to the pedigree to control for relatedness between individuals and once to estimate the permanent environmental effect. Additionally, for models with body mass, beak color, wing, and tarsus length and fat score as the dependent variable, we included the identity of the batch of birds that were measured together as a random effect (group ID) to control for batch effects between measurement sessions.

For models of fitness-related traits, we controlled for individual age, inbreeding level ( $F_{\text{ped}}$ ), the number of days the individual was allowed to breed (in aviaries), the sex ratio (i.e., the proportion of males), and pairing status (force-paired in cages or free-paired in aviaries) by including them as fixed effects, whenever applicable. Additionally, for egg-based models (male fertility, within-pair paternity, embryo and nestling survival), we controlled for clutch order and laying or hatching order of the egg that was laid/potentially sired by the focal female or male. For models on embryo and nestling survival, we also controlled for the inbreeding level of the offspring. In all models, we included individual identity, breeding season identity, clutch identity, identity of the partner of the focal individual and the pair identity, as appropriate.

We metasummarized effect sizes using the “lm” function in the R package “stats,” whereby we weighted each effect size by the inverse of the standard error of the estimate to account for the uncertainty of each estimate. Intercepts were removed to estimate the mean of each category unless stated otherwise. First, we metasummarized the direct effect of each of the six stressors on the individual’s own morphological versus fitness-related traits (“trait type,” two levels). In this model, we fitted the pairwise combination of the trait type and the potential stressor as a fixed effect with 12 levels (Table S5). Second, we summarized the di-

rect or transgenerational effects (from the individual, its parents and grandparents, seven levels, “stress experienced by a certain individual”) of the most powerful proxy of developmental stress (nestling body mass at 8 days old; see the “Results” section) on the morphological versus fitness-related traits (two levels) of males and females (two levels, “sex”). In this model, we fitted the pairwise combination of stress experienced by a certain individual, trait type, and sex as a fixed effect with 28 levels (Table S6). Third, we metasummarized the transgenerational anticipatory effect of the similarity between parent-offspring in their nestling mass (mother or father in combination with daughters or sons, four levels) on the offspring’s morphological versus fitness-related traits (two levels). Here we included the pairwise combination of parent, offspring sex and trait type as a fixed effect with eight levels (Table S7).

Then, we metasummarized the overall transgenerational effects of condition transfer and anticipatory effects in two mixed-effect models using the “lmer” function in the R package “lme4” (Bates et al. 2015), where we weighted each estimate by the multiplicative inverse of its standard error to account for their level of uncertainty. To account for the nonindependence between response variables (see Fig. S1), we fitted a random effect that reflects their dependencies. For this purpose, we grouped all 23 performance traits based on their pairwise correlation coefficients (Table S1) into 11 categories (see Table S8). The fitted random effect groups the performance traits into 11 categories separately for each ancestor (22 levels for the parents and 44 levels for grandparents). We metasummarized the overall transgenerational effects of condition transfer of mass at day 8 experienced by the ancestors (parents and grandparents) on the traits of individuals, by only including an intercept (Table S8). Last, we metasummarized the overall transgenerational anticipatory effect of similarity between parent-offspring in their mass at day 8 on the traits of offspring, by only including an intercept (Table S9).

For visualization, we calculated the expected Z-values with 95% CIs from a normal distribution given the number of Z-values for each group of effects due to each stressor experienced by the focal individual, its mother, its father and its grandparents formulas as follows: expected Z-values as “qnorm(ppoints( $N$  Z-values))” (i.e., the integrated quantiles assuming a uniformly distributed probability of a given number of observations) and 95% CIs of the expected Z-values as “qnorm(qbeta( $p = (1 \pm \text{CI})/2$ , shape1 = 1:  $N$  Z-values, shape2 =  $N$  Z-values:1))” (i.e., the integrated quantiles of quantiles of a uniformly distributed probability of a given number of observations from a beta distribution) in the R package “stats.” We visually inspected the ZZ-plots for the expected versus observed Z-values dependent on the direction of the effects. Z-values larger than 1.96 were considered to be significant.

## CONFIRMATORY ANALYSIS

For the confirmatory analysis, we used additional birds, including the remaining individuals from the main study population (referred to as “Seewiesen”) whose maternal nestling mass was known (but information from grandparents was missing), as well as birds from two other captive populations with short pedigrees: “Krakow” (interbreeding between populations “Krakow” #11 in (Forstmeier et al. 2007) and “Seewiesen”) and “Bielefeld” (wild-derived in the late 1980s, #19 in Forstmeier et al. 2007). These datasets are of equal quality as the main dataset, but have shorter pedigrees. To replicate the tests that showed significant effects of maternal early condition and the similarity between mother-daughter early condition on daughter fecundity-related traits (see the “Results” section), we used the following samples: (i) female clutch size measured in cages ( $N = 156$  “Seewiesen” and 30 “Krakow” females) or (ii) in aviaries ( $N = 84$  “Seewiesen,” 66 “Krakow,” and 53 “Bielefeld” females); (iii) female fecundity, measured in aviaries ( $N = 31$  “Seewiesen” females). We Z-scaled nestling body mass at 8 days of age within each population before further analysis because birds in the recently wild-derived population “Bielefeld” were smaller compared to those of the domesticated “Seewiesen” and “Krakow” populations. We used the “lmer” function from the R package “lme4” to estimate the maternal nestling mass effect on daughters’ fecundity-related traits. The same model structure was used as in the initial tests, but we additionally controlled for between-population differences by including the population where the female came from as a fixed effect. In the model of female fecundity, we removed the variable “number of days the female stayed in the experiment” (because there was no variation) and the random effect “female identity” (because each individual contributed only one data point). To replicate the tests that showed a significant effect of similarity of father-daughter early condition on the daughters’ size-related traits, we used (1) tarsus length of  $N = 447$  “Seewiesen,” 290 “Krakow” and 333 “Bielefeld” females and (2) wing length of  $N = 222$  “Seewiesen,” 287 “Krakow” and 332 “Bielefeld” females. We analyzed the animal model for each population separately, using the same model structure as in the initial test, using the R function “pedigreeMM” from package “PedigreeMM.” In the confirmatory models for the Seewiesen population, we removed “author identity” because all birds were measured by the same person.

We metasummarized the effects of (1) maternal mass at 8 days old, (2) similarity of mother-daughter early condition on her daughters’ fecundity-related traits, and (3) similarity of father-daughter early condition on his daughters’ size (see the “Results” section) in a “lm” model, by fitting the pairwise combination of test (initial or confirmatory) and the three effects as a fixed effect with six levels and the multiplicative reverse of the standard error of each estimate as “weight” (Table S10).

## Results

We examined the effects of six variables describing early developmental conditions (potential stressors) on 10 measures of morphology and on 13 aspects of reproductive performance, resulting in 138 predictor-outcome combinations ( $6 \times 23$  tests). We thus obtained 138 effect sizes for the direct effects (intragenerational; Table S2), 828 effect sizes for the intergenerational condition-transfer effects (i.e., effects of the early-life experiences of the six ancestors: two parents and four grandparents,  $6 \times 138$ ; Table S2) and 46 effect sizes for the anticipatory effects (i.e., effects of similarity in nestling mass between mother and offspring and between father and offspring for 23 performance traits; Table S3).

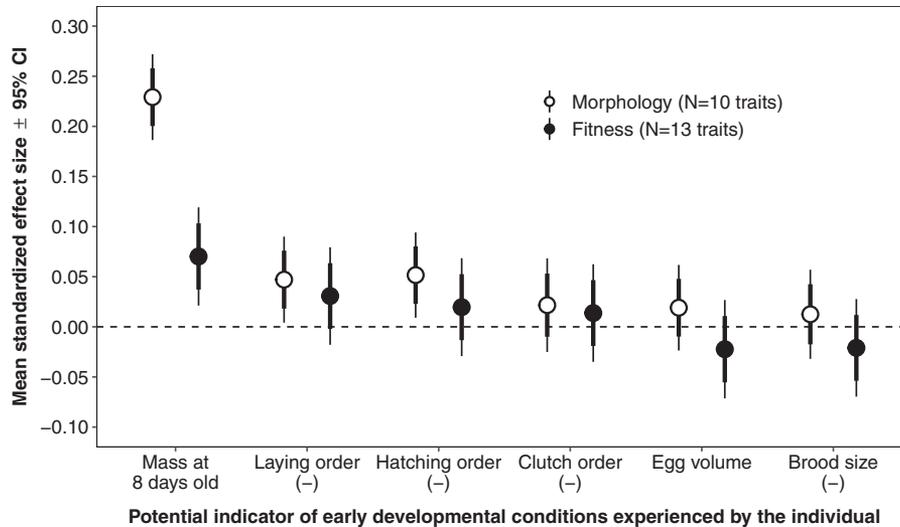
### VALIDATION OF STRESSORS USING DIRECT EFFECTS

Of the six putative indicators of early developmental conditions, only one measure had significant consequences for the adult individual (Fig. 1). Nestling body mass measured at 8 days of age affected both adult morphology (mean  $r = 0.229$ , 95% CI: 0.186–0.272,  $P < 0.0001$ ) and reproductive performance (mean  $r = 0.070$ , 95% CI: 0.021–0.119,  $P < 0.0001$ ; Fig. 1 and Table S5). The mass of nestlings when 8 days old varied by a factor of 10 (range: 1.2–12.4 g,  $N = 3525$  nestlings), and light-weight nestlings had a clearly reduced chance of survival to adulthood (see Fig. S2). Among the survivors ( $N = 3326$ ) and among those individuals included in the analyses of direct and transgenerational effects ( $N = 2099$ ), mass at day 8 still varied by a factor of 7 (range: 1.7–12.4 g).

Other indicators of developmental conditions, despite being widely used as proxies in the published literature, had little direct effect on the individual later in life. Therefore, in the following analyses we only use nestling body mass at day 8 as the proxy of early-life condition of parents and grandparents to assess the strength of the two types of transgenerational effects. Note that nestling body mass is a measure of an outcome of stress rather than the cause of stress. In contrast, the other five variables represent causes rather than outcomes of stress. However, none of them shows direct effects (Fig. 1), so there seems little scope for detecting transgenerational effects. Nevertheless, Table S2 lists all transgenerational effects (mean effect size  $r = -0.003$ ,  $N = 828$  effects).

### TRANSGENERATIONAL EFFECTS OF NESTLING MASS: CONDITION TRANSFER

We did not find any evidence for a transgenerational effect of nestling mass either of the parents or of the grandparents on the adult offspring (mean estimate of 138 transgenerational effects after accounting for some level of non-independence between the response variables  $r = -0.007$ , 95% CI:  $-0.017$  to  $0.002$ ; Table S8; see also Fig. 2B and D, and Table S6). Among



**Figure 1.** Average magnitude of direct effects of six potential indicators of early developmental conditions experienced by the individual itself on their adult morphology (averaged across 10 traits; open symbols) and fitness-related traits (13 traits; filled; Table S5). Error bars show two types of 95% CIs: thick lines refer to the single estimate and thin lines are Bonferroni adjusted for conducting 12 tests (figure-wide significance). Morphological traits (sex-specific body mass, tarsus length, wing length, fat score and beak color, median  $N = 944$  females and 1008 males) were measured when individuals were 93–229 days old. Fitness-related traits include male and female life span, male and female number of seasonal recruits, female clutch size (in cages and aviaries) and fecundity, male fertility, within-pair paternity, male siring success in aviaries, female embryo survival, and female and male nestling survival (median  $N = 228$  individuals). Four out of the six indicators of early-life conditions were multiplied by  $-1$  (indicated by  $(-)$ ) such that positive effect sizes reflect better performance under supposedly better conditions. Morphological and fitness-related traits as well as indicators of early-life conditions were Z-transformed to yield effect sizes in the form of Pearson correlation coefficients.

the many correlations examined, only one was significant: the nestling mass of the mother correlated positively with the reproductive performance (clutch size in cages and aviaries, fecundity in aviaries, embryo survival, nestling survival, seasonal recruits, and life span; Table S2) of her daughters (mean  $r = 0.071$ , 95% CI: 0.024–0.118,  $P = 0.003$ , without correction for multiple testing; Table S6; also see Fig. 2D). This finding was mostly driven by a large positive effect of maternal early growth on daughter fecundity (Fig. 3F), which was even larger than the direct effect of the daughters’ own nestling mass at day 8 (Fig. 3E). For all other dependent traits that were influenced by nestling mass, the direct effects (Fig. 3A and C) exceeded the indirect maternal effects (Fig. 3B and D).

The direct effects of nestling mass on the individual’s adult traits are clearly stronger than expected under a random distribution of effect sizes (Fig. 2E; see also Fig. 2A and C). In contrast, the positive effects of the early-life condition of the mother (Fig. 2F) are not much stronger than the presumably coincidental negative effects (opposite to expectations) of the early-life condition of the father and the grandparents (Fig. 2G and H; see also Fig. S3). The two significant maternal effects (upper right corner in Fig. 2F) are those on daughter fecundity (see Fig. 3F) and on daughter clutch size ( $r = 0.103$ , 95% CI: 0.025–0.181,  $P = 0.01$  without correction for multiple testing; Table S2). These

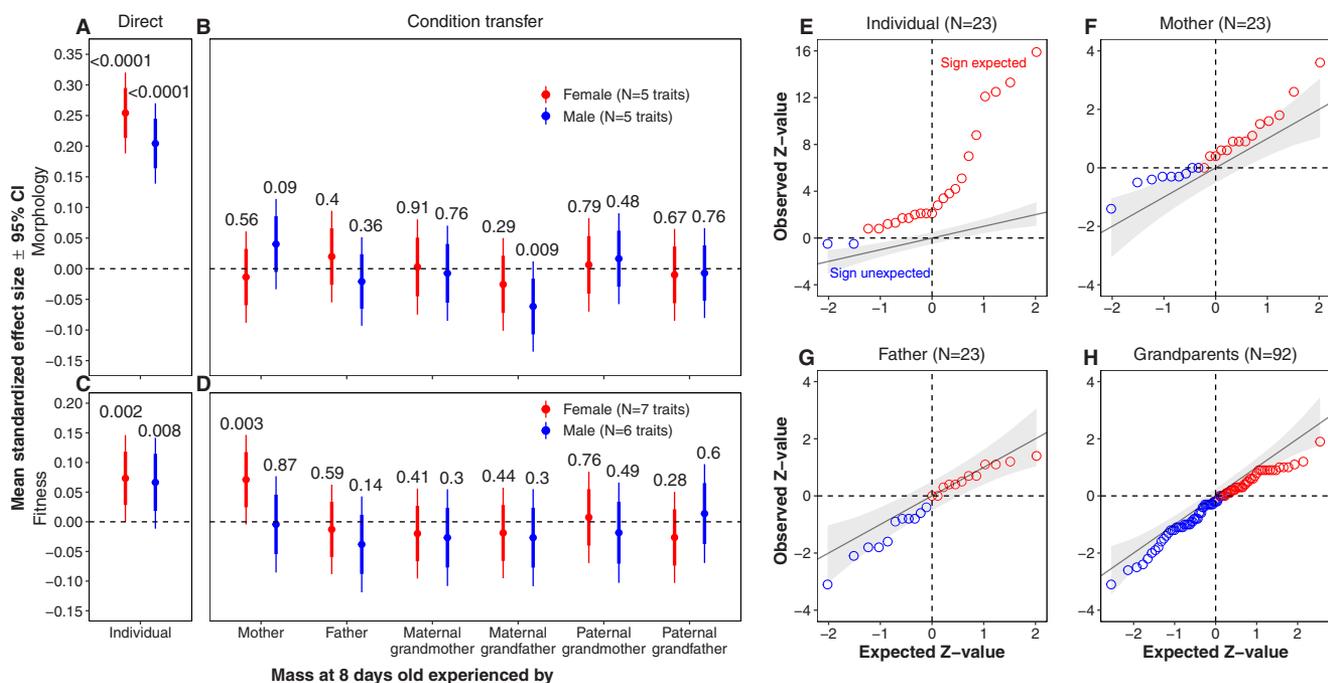
findings are not independent because clutch size and fecundity are strongly correlated ( $r = 0.71$ ,  $N = 230$  females; Fig. S1 and Table S1), partially due to the fact that they are measured in the same breeding season ( $N = 183$  females).

**TRANSGENERATIONAL EFFECTS OF NESTLING MASS: ANTICIPATORY EFFECTS**

Offspring performed significantly better when growing up under similar conditions as their parents (similarity in mass at day 8), but the effect size was small (mean estimate of 46 transgenerational effects after accounting for some level of non-independence between the response variables  $r = 0.028$ , 95% CI: 0.016–0.040; Table S9; see also Fig. 4 and Table S7). This was mainly driven by the positive effects of (1) father-daughter similarity on the daughters’ size (i.e., tarsus and wing length) and (2) mother-daughter similarity on the daughters’ fitness-related traits (Fig. 4 and Table S7; see also Table S3).

**CONFIRMATORY TESTS ON INDEPENDENT DATA**

To independently verify the strongest and most plausible findings of (1) condition transfer from the mother affecting daughter fecundity, (2) anticipatory effects of similarity between mother and daughter in their nestling body mass on daughter fecundity, and (3) anticipatory effects from the father-daughter similarity on



**Figure 2.** Transgenerational condition-transfer effects of early developmental conditions (measured as nestling body mass at 8 days of age). (A–D) Average magnitude of condition-transfer effects from six types of ancestors (B, D) in comparison to the direct effects of the experience of the individual itself (A, C) on morphological (mean of five traits; A, B) and fitness-related traits (mean of six or seven traits; C, D) for individual females (red) and males (blue; Table S6). Error bars show two types of 95% CIs: thick lines refer to the single estimate and thin lines are Bonferroni adjusted for conducting 28 tests (figure-wide significance among A–D). Indicated *P*-values refer to each average effect estimate without correction for multiple testing. For further explanations see legend of Figure 1. (E–H) ZZ-plots of expected versus observed Z-values of the effects of early developmental conditions (mass at 8 days) experienced by the focal individual itself (E), its mother (F), its father (G) and its four grandparents (H) on 10 morphological and 13 fitness-related traits. *N* indicates the number of tests. Red indicates that the sign of the estimate is in the expected direction, blue indicates that the sign is in the opposite direction. Lines of identity (where observation equals prediction) and their 95% CIs are shown.

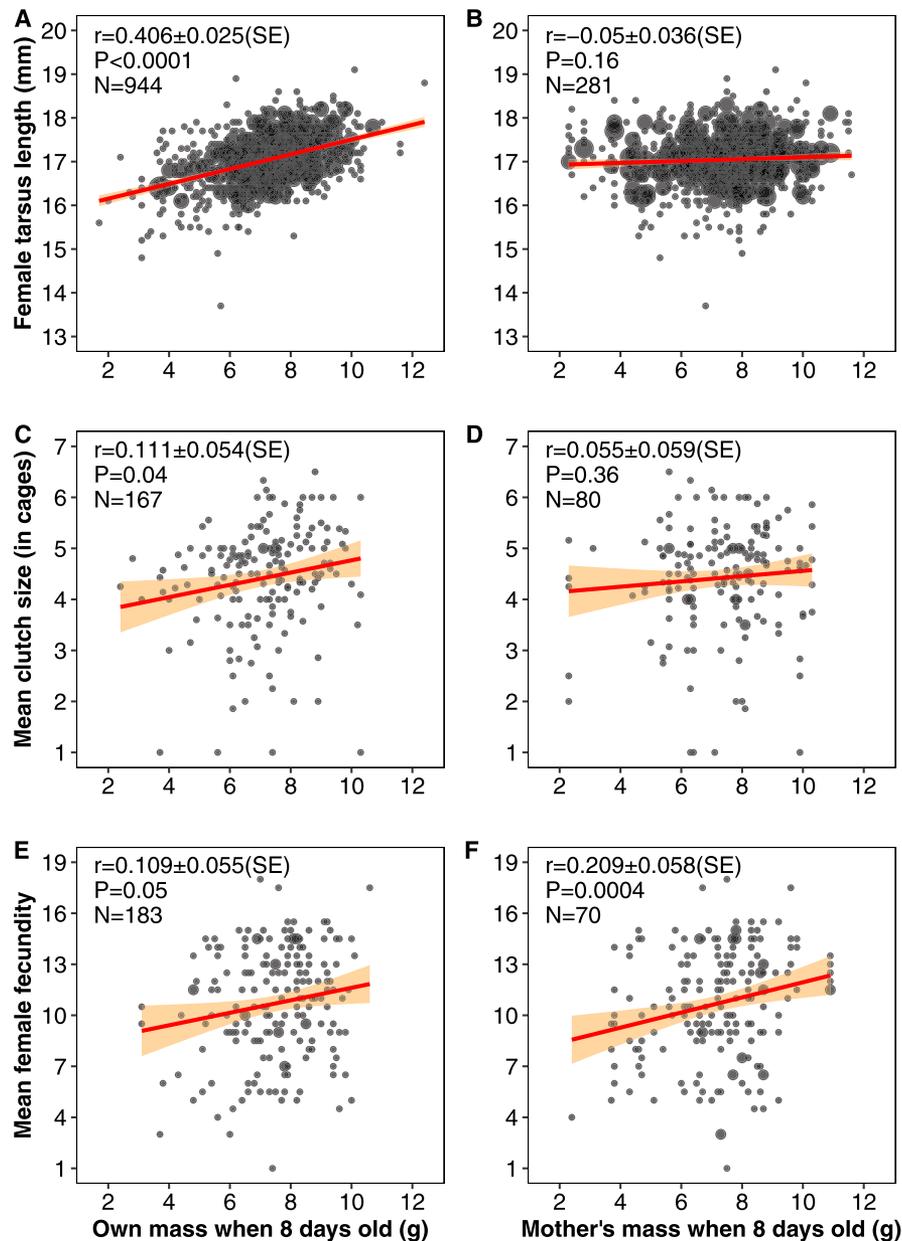
daughter body size, we examined an independent dataset (comprising data from several populations, see the “Methods” section for details). All effect sizes of the confirmatory analysis are listed in Table S10. For all three tests, the initial effect size was clearly larger than the independent verification effect size (exploratory vs. confirmatory, detailed in Fig. 5 and Table S10) and, apart from the effect of father-daughter similarity on daughter tarsus length in one of the three populations, none of the confirmatory tests was significant.

## Discussion

Our study supports the general idea that individuals are resilient to stress and particularly to stress experienced by ancestors. Even though individuals differed sevenfold in body mass when 8 days old, nestling mass only had small effects on morphology and reproductive success later in life. Our results clearly reject the hypothesis of condition transfer between generations, in line with the idea that selection acts against transmitting a handicap to the next generation. We found some evidence for transgenerational

anticipatory effects, but the mean effect was small ( $r = 0.028$ ), and did not hold up in an independent confirmatory test (Fig. 5B and C). These mixed results indicate that, in our study, the effect size for transgenerational anticipatory effects must be exceedingly small (Uller et al. 2013; Horsthemke 2018).

In conclusion, transgenerational effects were absent or miniscule, and direct effects on fitness traits were relatively small given that some of the offspring were seriously undernourished. Thus, at least in this study system, the notion of organismal robustness seems more noteworthy than the claim of sensitivity to early-life conditions within and across generations. Nevertheless, the latter dominates both the literature with studies focused on zebra finches (e.g., Naguib and Gil 2005; Monaghan 2008; Tschirren et al. 2009; Krause and Naguib 2014; Khan et al. 2016; Wilson et al. 2019) and the broader literature (e.g., Pembrey et al. 2006; Marshall and Uller 2007; Uller et al. 2013; Engqvist and Reinhold 2016; Zizzari et al. 2016). This begs the question whether the underrepresentation of studies emphasizing “robustness” in the literature is the result of the predominant framework of hypothesis testing, where the rejection of the null

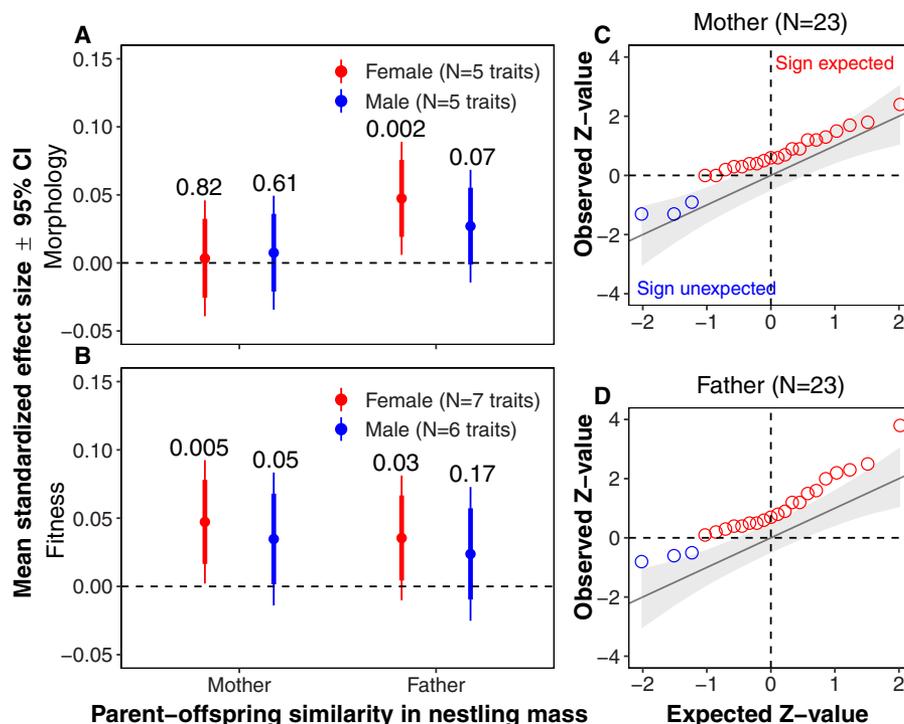


**Figure 3.** Comparison of direct effects (left column) and effects of condition transfer from mother to daughter (right column). Relationship between nestling mass at 8 days old experienced by the focal female (A, C, E) or by her mother (B, D, F) and tarsus length (A, B), mean clutch size in cage breeding (C, D) and mean female fecundity when breeding in an aviary (E, F). Dot size reflects sample size. Red lines are the linear regression lines ( $\pm 95\%$  CI, orange shading) on the data shown, while indicated effect sizes ( $r$  with SE,  $P$ ,  $N$ : number of females for A, C, E, and number of mothers for B, C, D) reflect estimates from mixed models. Note that each individual's own mass when 8 days old corresponds to one value of the dependent variable whereas each mother's mass at 8 days old can correspond to multiple values of the dependent variable (one for each of her daughters).

hypothesis is almost a pre-condition of getting published (Greenwald 1975).

We found that direct intragenerational effects of early environment on morphology were of moderate magnitude while effects on fitness-related traits were small, which is largely in line with previous findings (Tschirren et al. 2009; Eyck et al. 2019). Regarding transgenerational effects of early stress, we examined

the existing zebra finch literature that is mostly based on captive birds (Naguib and Gil 2005; Naguib et al. 2006; Alonso-Alvarez et al. 2007; Krause and Naguib 2014; Khan et al. 2016; Wilson et al. 2019) and found that studies typically report a large number of tests (median number of discussed combinations of stressors, traits and sex: 18, range: 7–150). Only 15% of all tests were statistically significant, which is not far from the random



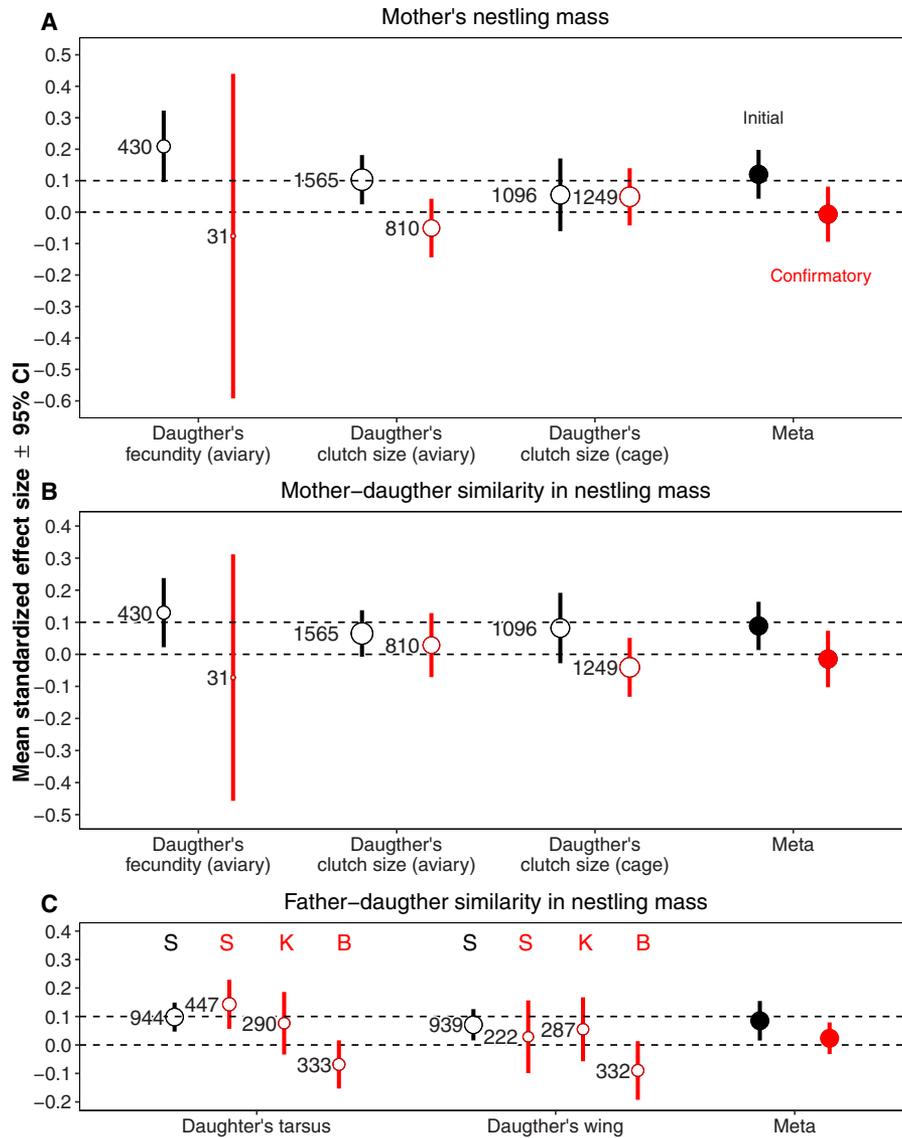
**Figure 4.** Transgenerational anticipatory effects (of similarity in early growth conditions between parents and offspring). (A, B) Average magnitude of anticipatory effect on morphological (mean of five traits; A) and fitness-related traits (mean of six or seven traits; B) for individual females (red) and males (blue; Table S7). Error bars show two types of 95% CIs: thick lines refer to the single estimate and thin lines are Bonferroni adjusted for conducting eight tests (figure-wide significance in A, B). Indicated *P*-values refer to each average effect estimate without correction for multiple testing. (C, D) ZZ-plots of expected versus observed Z-values of the effects of similarity in conditions (mass at 8 days old) between the focal individual itself and its mother (C), and between the focal individual itself and its father (D) on 10 morphological and 13 fitness-related traits (Table S3). For further explanations of A and B, see legends of Figure 1, and for C and D see legends of Figure 2.

expectation, especially if some nonsignificant findings were not reported. Additionally, an experimental study on zebra finches found no transgenerational anticipatory effect (Krause and Naguib 2014) and a meta-analysis of studies on plants and animals found no effect of transgenerational condition transfer (Uller et al. 2013).

Given the small (expected) effect sizes, we argue that transgenerational effects can sensibly only be studied within a framework that ensures a comprehensive reporting of all effect sizes and a meta-analytic summary of these effects. Focus on a subset of tests (e.g., those that are significant) leads to bias, but selective attention may be advisable in two situations. Firstly, when there is an independent selection criterion. For example, we limited our analysis of transgenerational effects to those involving only the most powerful indicator of early developmental conditions. In this case, the selection criterion (magnitude of direct effects; Fig. 1) was established independently of the outcome variable (magnitude of transgenerational effects). Second, when there is an independent dataset. For example, we selected the largest transgenerational effects from a first dataset, and assessed

them independently using the second dataset (Fig. 5). Consistent with the phenomenon of the winner's curse (Forstmeier and Schielzeth 2011), we found that selective attention to large effects yields inflated effect size estimates compared to the independent replication.

Selective attention to large effects makes the published effect size estimates unreliable. Thus, we propose to base conclusions on meta-analytic averages of all effect sizes that have been judged worth of investigation before any results were obtained. With this approach we shift our attention from identifying the supposedly best predictor and best response towards the quantification of the magnitude of an average predictor on an average response. Clearly, the latter is more reliable than the former, just as the average of many numbers is more robust than the maximum. Accordingly, the meta-analytic summary yields narrow confidence intervals (CIs) around the estimated mean effect size. Note, however, that the estimated 95% CI might be somewhat anticonservative (i.e., too narrow), because the summarized effect sizes are not fully independent of each other (multiple response variables are correlated; see Fig. S1 and Table S1). In the cases where we



**Figure 5.** Confirmatory analysis of transgenerational condition-transfer (A) and anticipatory (B, C) effects. Effect sizes (mean  $\pm$  95% CI without controlling for multiple testing) of mother's mass at 8 days old (A) and the similarity between mother and daughter in their mass at 8 days old (B) on daughters' fecundity, clutch size in cage and aviary (open) and the similarity of father and daughter in their mass at day 8 on the daughters' tarsus and wing length (open symbols; C). The filled symbols show the metasummarized effect sizes from the initial dataset (black) and from the confirmatory dataset (red). Numbers in the plots refer to the number of individuals (tarsus and wing length), clutches (clutch size in aviary and cage), or breeding seasons (fecundity). Daughter fecundity-related and size-related traits, mother's mass at 8 days old, and similarity between mother-daughter and father-daughter in nestling mass were Z-transformed to yield effect sizes in the form of Pearson correlation coefficients. Tarsus and wing length were analyzed by population due to the between-population difference in body size (C), where "S," "K," and "B" refer to populations "Seewiesen," "Krakow," and "Bielefeld." For additional details, see Table S10.

summarize a large number of effect sizes (138 estimates in Table S8 and 44 estimates in Table S9) we fitted a random effect that controls for some of this nonindependence, and this led to CIs that are about 20% wider (compared to dropping the random effect). This approach of modeling and quantifying the degree of nonindependence cannot be applied when summarizing only few effect size estimates (between 2 and 13 estimates in Figs. 1, 2,

4, and 5), meaning that the indicated CIs will be somewhat too narrow.

In our study, five of six putative indicators of early developmental stress had little or no direct effect on an individual's morphology and fitness later in life (Fig. 1). This suggests that it is not worth to examine these traits for transgenerational effects (Jablonka and Raz 2009; see also Fig. S3), unless one can

plausibly assume that some indicators of early stress cause direct effects, while others cause transgenerational effects. This differs from previous studies that showed various direct effects, but did not metasummarize all examined effects, for example, of brood size (Naguib et al. 2004, Naguib et al. 2008; Tschirren et al. 2009), laying order (Gorman and Nager 2004; Soma et al. 2007) and hatching order (Wilson et al. 2019). In contrast to the other five variables, nestling mass (8 days old) was clearly associated with both nestling survival (Fig. S2) and adult performance. However, its strongest effect was on morphology (highest  $r = 0.41$ ; Fig. 3A and Table S2), which is somewhat trivial. Food shortage during the developmental period reduces growth and this in turn affects body size later in life (Bolund et al. 2010). Because body size *per se* has little direct causal effect on fitness in zebra finches (Bolund et al. 2011), more complete developmental canalization for size-related traits may not have evolved. Indeed, despite large variation in mass at day 8 (1.7–12 g), the effect of nestling mass on reproductive performance and life span was weak, suggesting that fitness is remarkably resilient to variation in early-life conditions (Waddigton 1942; Drummond and Ancona 2015).

Note that our study is nonexperimental and on captive individuals. The latter implies that individuals were kept in a safe environment with *ad libitum* access to food (but with intense social interactions including competition for mates and nest sites). Direct and transgenerational effects on reproductive performance traits may be different in free-living populations, where individuals live and reproduce under potentially more stressful environmental conditions. Additionally, our dataset was not ideal to test “anticipatory parental effects.” This hypothesis predicts that offspring have higher fitness when the offspring environment matches the parental environment (e.g., Uller et al. 2013; Engqvist and Reinhold 2016). In an ideal experiment, one would manipulate the parents’ and the offspring’s breeding environments in a fully factorial design and examine the effects of matching versus mismatching on offspring performance (Monaghan 2008; Uller et al. 2013; Engqvist and Reinhold 2016, 2018). Our study only uses observational data and only regarding the similarity of the early growth environments (but not breeding environments). However, a meta-analysis of experimental studies on plants and animals only found a weak trend for small beneficial anticipatory parental effects (effect size  $d = 0.186$ , highest posterior density:  $-0.030, 0.393$ ) (Uller et al. 2013). Experimental studies are better suited to test causality, but when analyses of observational data suggest no effect, experiments may not provide much insight (note that the 95% CI for the mean effect excluded all biologically relevant effect sizes, for example, the estimated condition-transfer effects ranged from  $-0.017$  to  $0.002$ , and anticipatory effects ranged from  $0.016$  to  $0.040$ ). Our approach had the advantage that we could make use of the entire range of ob-

served growth conditions (sevenfold difference in mass at day 8), while experimental studies often only induce a 10–15% difference in nestling mass between treatment groups (because ethical concerns prohibit strong treatments; Naguib et al. 2004; Bolund et al. 2010). This then requires much larger sample sizes to detect similar phenotypic effects. Our additional confirmatory datasets had smaller sample sizes than the initial dataset (Fig. 5), and the data were more heterogeneous because they included individuals from different populations that differ in genetic background, body size, and domestication history.

Zebra finches in the wild might breed multiple times across a broad range of conditions (Zann 1996). One might thus question the suitability of zebra finches as a good model for testing “anticipatory parental effects.” Hence, the biological conclusions of our study should be taken with caution because our findings on zebra finches might not be representative for a broad range of organisms. In contrast, the meta-analytical method we propose here can be broadly applied—as an alternative to or in combination with preregistration—to ensure that effect sizes are not inflated. Biased reporting presumably occurs in most disciplines, and such biases could explain the discrepancy between our findings and the conclusions of the existing zebra finch literature on early-life and transgenerational effects.

In summary, for future studies on transgenerational effects, we suggest an approach that renders multiple testing a strength rather than a burden and that consists of four simple steps: (i) start with clear, one-tailed hypotheses (Ruxton and Neuhäuser 2010); (ii) validation by assessing the direct effects (Fig. 1); (iii) meta-analysis of all effects (Fig. 2) and—if feasible—(iv) verify the effects with an independent confirmatory dataset (Fig. 5). Using this approach, our study shows convincing evidence for small direct effects, and—at best—weak evidence for small transgenerational effects on morphology and fitness. Hence, our study supports the null hypothesis that selection buffers individual fitness against detrimental epigenetic effects, such that the detrimental effects due to stress experienced early in life by the ancestors are not carried on across generations (Waddigton 1942; Hallgrímsson et al. 2002).

#### AUTHOR CONTRIBUTIONS

WF and BK designed the study. WF collected the morphological data. YP and WF analyzed the data and interpreted the results with input from BK. YP and WF wrote the manuscript with help from BK.

#### ACKNOWLEDGMENTS

We thank M. Schneider for molecular work; K. Martin, S. Bauer, E. Boddendorfer, J. Didsbury, A. Grötsch, A. Kortner, P. Neubauer, F. Weigel, and B. Wörle for animal care and help with breeding; and Associate Editor Thomas Flatt and two anonymous reviewers for constructive comments on the manuscript. This research was supported by the Max Planck

Society (to BK). YP is part of the International Max Planck Research School for Organismal Biology.

## DATA ARCHIVING

Supporting data and R scripts can be found in the Open Science Framework at <https://osf.io/wjg3q/> (doi:10.17605/OSF.IO/WJG3Q) and the fitness-related data can be found in the Open Science Framework at <https://osf.io/tgsz8/> (doi:10.17605/OSF.IO/TGSZ8).

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## LITERATURE CITED

- Alonso-Alvarez, C., S. Bertrand, and G. Sorci. 2007. Sex-specific transgenerational effects of early developmental conditions in a passerine. *Biol. J. Linn. Soc.* 91:469–474.
- Anway, M. D., A. S. Cupp, M. Uzumcu, and M. K. Skinner. 2005. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308:1466–1469.
- Bagijn, M. P., L. D. Goldstein, A. Sapetschnig, E. M. Weick, S. Bouasker, N. J. Lehrbach, M. J. Simard, and E. A. Miska. 2012. Function, targets, and evolution of *Caenorhabditis elegans* piRNAs. *Science* 337:574–578.
- Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67:1–48.
- Benkemoun, L., and S. J. Saupé. 2006. Prion proteins as genetic material in fungi. *Fungal Genet. Biol.* 43:789–803.
- Bertram, C., O. Khan, S. Ohri, D. I. Phillips, S. G. Matthews, and M. A. Hanson. 2008. Transgenerational effects of prenatal nutrient restriction on cardiovascular and hypothalamic-pituitary-adrenal function. *J. Physiol.* 586:2217–2229.
- Boersma, G. J., T. L. Bale, P. Casanello, H. E. Lara, A. B. Lucion, D. Suchecki, and K. L. Tamashiro. 2014. Long-term impact of early life events on physiology and behaviour. *J. Neuroendocrinol.* 26:587–602.
- Bolund, E., H. Schielzeth, and W. Forstmeier. 2007. Intrasexual competition in zebra finches, the role of beak colour and body size. *Anim. Behav.* 74:715–724.
- Bolund, E., H. Schielzeth, and W. Forstmeier. 2010. No heightened condition dependence of zebra finch ornaments—a quantitative genetic approach. *J. Evol. Biol.* 23:586–597.
- Bolund, E., H. Schielzeth, and W. Forstmeier. 2011. Correlates of male fitness in captive zebra finches—a comparison of methods to disentangle genetic and environmental effects. *BMC Evol. Biol.* 11:327.
- Bonduriansky, R., and A. J. Crean. 2018. What are parental condition-transfer effects and how can they be detected? *Methods Ecol. Evol.* 9:450–456.
- Bonduriansky, R., and M. Head. 2007. Maternal and paternal condition effects on offspring phenotype in *Telostylinus angusticollis* (Diptera: Neriidae). *J. Evol. Biol.* 20:2379–2388.
- Burton, T., and N. B. Metcalfe. 2014. Can environmental conditions experienced in early life influence future generations? *Proc. R. Soc. B* 281:20140311.
- Carone, B. R., L. Fauquier, N. Habib, J. M. Shea, C. E. Hart, R. Li, C. Bock, C. Li, H. Gu, P. D. Zamore, et al. 2010. Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell* 143:1084–1096.
- Colborn, T., F. S. Vom Saal, and A. M. Soto. 1993. Developmental effects of endocrine-disrupting chemicals in wildlife and humans. *Environ. Health Perspect.* 101:378–384.
- Cubas, P., C. Vincent, and E. Coen. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401:157–161.
- Dey, S., S. R. Proulx, and H. Teotónio. 2016. Adaptation to temporally fluctuating environments by the evolution of maternal effects. *PLoS Biol.* 14:1–29.
- Drummond, H., and S. Ancona. 2015. Observational field studies reveal wild birds responding to early-life stresses with resilience, plasticity, and intergenerational effects. *Auk* 132:563–576.
- Eggert, H., M. F. Diddens-de Bühr, and J. Kurtz. 2015. A temperature shock can lead to trans-generational immune priming in the Red Flour Beetle, *Tribolium castaneum*. *Ecol. Evol.* 5:1318–1326.
- Engqvist, L., and K. Reinhold. 2016. Adaptive trans-generational phenotypic plasticity and the lack of an experimental control in reciprocal match/mismatch experiments. *Methods Ecol. Evol.* 7:1482–1488.
- Engqvist, L., and K. Reinhold. 2018. Adaptive parental effects and how to estimate them: a comment to Bonduriansky and Crean. *Methods Ecol. Evol.* 9:457–459.
- Eyck, H. J. F., K. L. Buchanan, O. L. Crino, and T. S. Jessop. 2019. Effects of developmental stress on animal phenotype and performance: a quantitative review. *Biol. Rev.* 94:1143–1160.
- Feng, S. 2010. Epigenetic reprogramming in plant. *Science* 622:622–627.
- Ferrari, R. P., R. Martinelli, and N. Saino. 2006. Differential effects of egg albumen content on barn swallow nestlings in relation to hatch order. *J. Evol. Biol.* 19:981–993.
- Forstmeier, W., G. Segelbacher, J. C. Mueller, and B. Kempenaers. 2007. Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular Ecology* 16:4039–4050.
- Forstmeier, W., and H. Schielzeth. 2011. Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behav. Ecol. Sociobiol.* 65:47–55.
- Franzke, A., and K. Reinhold. 2013. Transgenerational effects of diet environment on life-history and acoustic signals of a grasshopper. *Behav. Ecol.* 24:734–739.
- Gilby, A. J., E. Sorato, and S. C. Griffith. 2012. Maternal effects on begging behaviour: an experimental demonstration of the effects of laying sequence, hatch order, nestling sex and brood size. *Behav. Ecol. Sociobiol.* 66:1519–1529.
- Goerlich, V. C., D. Nätt, M. Elfving, B. Macdonald, and P. Jensen. 2012. Transgenerational effects of early experience on behavioral, hormonal and gene expression responses to acute stress in the precocial chicken. *Horm. Behav.* 61:711–718.
- Gorman, H. E., and R. G. Nager. 2004. Prenatal developmental conditions have long-term effects on offspring fecundity. *Proc. R. Soc. London. Ser. B* 271:1923–1928.
- Greenwald, A. G. 1975. Consequences of prejudice against the null hypothesis. *Psychol. Bull.* 82:1–20.
- Guerrero-Bosagna, C., M. Morisson, L. Liaubet, T. B. Rodenburg, E. N. de Haas, E. Košťál, and F. Pitel. 2018. Transgenerational epigenetic inheritance in birds. *Environ. Epigenet.* 4:1–8.
- Hallgrímsson, B., K. Willmore, and B. K. Hall. 2002. Canalization, developmental stability, and morphological integration in primate limbs. *Am. J. Phys. Anthropol.* 119:131–158.
- Henderson, I. R., and S. E. Jacobsen. 2007. Epigenetic inheritance in plants. *Nature* 447:418–424.
- Hettinger, A., E. Sanford, T. M. Hill, A. D. Russell, K. N. S. Sato, J. Hoey, M. Forsch, H. N. Page, and B. Gaylord. 2012. Persistent carry-over effects of planktonic exposure to ocean acidification in the Olympia oyster. *Ecology* 93:2758–2768.
- Holliday, R. 1987. Epigenetic defects. *Science* 238:163–170.
- Horsthemke, B. 2018. A critical view on transgenerational epigenetic inheritance in humans. *Nat. Commun.* 9:2973.

- Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous inference in general parametric models. *Biometrical J.* 50:346–363.
- Ihle, M., B. Kempnaers, and W. Forstmeier. 2015. Fitness benefits of mate choice for compatibility in a socially monogamous species. *PLoS Biol.* 13:1–21.
- Jablonka, E., and G. Raz. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q. Rev. Biol.* 84:131–176.
- Jensen, P. 2013. Transgenerational epigenetic effects on animal behaviour. *Prog. Biophys. Mol. Biol.* 113:447–454.
- Kaati, G., L. O. Bygren, M. Pembrey, and M. Sjöström. 2007. Transgenerational response to nutrition, early life circumstances and longevity. *Eur. J. Hum. Genet.* 15:784–790.
- Khan, N., R. A. Peters, E. Richardson, and K. A. Robert. 2016. Maternal corticosterone exposure has transgenerational effects on grand-offspring. *Biol. Lett.* 12:20160627.
- Koskela, E. 1998. Offspring growth, survival and reproductive success in the bank vole: a litter size manipulation experiment. *Oecologia* 115:379–384.
- Kraft, F.-L. O. H., S. C. Driscoll, K. L. Buchanan, and O. L. Crino. 2019. Developmental stress reduces body condition across avian life-history stages: a comparison of quantitative magnetic resonance data and condition indices. *Gen. Comp. Endocrinol.* 272:33–41.
- Krause, E. T., and M. Naguib. 2014. Effects of parental and own early developmental conditions on the phenotype in zebra finches (*Taeniopygia guttata*). *Evol. Ecol.* 28:263–275.
- Kuijper, B., and R. A. Johnstone. 2016. Parental effects and the evolution of phenotypic memory. *J. Evol. Biol.* 29:265–276.
- Kuijper, B., R. A. Johnstone, and S. Townley. 2014. The evolution of multi-variate maternal effects. *PLoS Comput. Biol.* 10:e1003550.
- Ledón-Rettig, C. C., C. L. Richards, and L. B. Martin. 2013. Epigenetics for behavioral ecologists. *Behav. Ecol.* 24:311–324.
- Lev, I., I. A. Toker, Y. Mor, A. Nitzan, G. Weintraub, O. Antonova, O. Bhonkar, I. Ben Shushan, U. Seroussi, J. M. Claycomb, et al. 2019. Germ Granules Govern Small RNA Inheritance. *Curr. Biol.* 29:2880–2891.e4.
- Lindström, J. 1999. Early development in birds and mammals. *Trends Ecol. Evol.* 14:343–348.
- Love, O. P., and T. D. Williams. 2011. Manipulating developmental stress reveals sex-specific effects of egg size on offspring phenotype. *J. Evol. Biol.* 24:1497–1504.
- Magiafoglou, A., and A. A. Hoffmann. 2003. Cross-generation effects due to cold exposure in *Drosophila serrata*. *Funct. Ecol.* 17:664–672.
- Marshall, D. J., and T. Uller. 2007. When is a maternal effect adaptive? *Oikos* 116:1957–1963.
- Miska, E. A., and A. C. Ferguson-Smith. 2016. Transgenerational inheritance: models and mechanisms of non-DNA sequence-based inheritance. *Science* 354:59–63.
- Molinier, J., G. Ries, C. Zipfel, and B. Hohn. 2006. Transgeneration memory of stress in plants. *Nature* 442:1046–1049.
- Monaghan, P. 2008. Early growth conditions, phenotypic development and environmental change. *Philos. Trans. R. Soc. B* 363:1635–1645.
- Morgan, H. D., H. G. E. Sutherland, D. I. K. Martin, and E. Whitelaw. 1999. Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* 23:314–318.
- Naguib, M., and D. Gil. 2005. Transgenerational body size effects caused by early developmental stress in zebra finches. *Biol. Lett.* 1:95–97.
- Naguib, M., C. Heim, and D. Gil. 2008. Early developmental conditions and male attractiveness in zebra finches. *Ethology* 114:255–261.
- Naguib, M., A. Nemitz, and D. Gil. 2006. Maternal developmental stress reduces reproductive success of female offspring in zebra finches. *Proc. R. Soc. B* 273:1901–1905.
- Naguib, M., K. Riebel, A. Marzal, and D. Gil. 2004. Nestling immunocompetence and testosterone covary with brood size in a songbird. *Proc. R. Soc. Lond. Ser. B* 271:833–838.
- Pei, Y. 2020a. Supporting information for “Offspring performance is well buffered against stress experienced by ancestors.” <http://doi.org/10.17605/OSF.IO/WJG3Q>.
- Pei, Y. 2020b. Supporting information for “Proximate causes of infertility and embryo mortality in captive zebra finches.” <http://doi.org/10.17605/OSF.IO/TGSZ8>.
- Pei, Y., W. Forstmeier, D. Wang, K. Martin, J. Rutkowska, and B. Kempnaers. 2019. Proximate causes of infertility and embryo mortality in captive zebra finches. <http://doi.org/10.1101/847103>.
- Pembrey, M. E., L. O. Bygren, G. Kaati, S. Edvinsson, K. Northstone, M. Sjöström, and J. Golding. 2006. Sex-specific, male-line transgenerational responses in humans. *Eur. J. Hum. Genet.* 14:159–166.
- Proulx, S. R., and H. Teotónio. 2017. What kind of maternal effects can be selected for in fluctuating environments? *Am. Nat.* 189:E118–E137.
- R Core Team. 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Raveh, S., D. Vogt, and M. Kölliker. 2016. Maternal programming of offspring in relation to food availability in an insect (*Forficula auricularia*). *Proc. R. Soc. B* 283:20152936.
- Rechavi, O., L. Houriz-Ze’evi, S. Anava, W. S. S. Goh, S. Y. Kerk, G. J. Hannon, and O. Hobert. 2014. Starvation-induced transgenerational inheritance of small RNAs in *C. elegans*. *Cell* 158:277–287.
- Ruxton, G. D., and M. Neuhäuser. 2010. When should we use one-tailed hypothesis testing? *Methods Ecol. Evol.* 1:114–117.
- Saino, N., M. Incagli, R. Martinelli, R. Ambrosini, and A. P. Moller. 2001. Immunity, growth and begging behaviour of nestling Barn Swallows *Hirundo rustica* in relation to hatching order. *J. Avian Biol.* 32:263–270.
- Siegal, M. L., and A. Bergman. 2002. Waddington’s canalization revisited: developmental stability and evolution. *Proc. Natl. Acad. Sci.* 99:10528–10532.
- Soma, M., D. S. Saito, T. Hasegawa, and K. Okanoya. 2007. Sex-specific maternal effect on egg mass, laying order, and sibling competition in the Bengalese finch (*Lonchura striata* var. domestica). *Behav. Ecol. Sociobiol.* 61:1695–1705.
- Tomita, N., K. Kazama, H. Sakai, M. Sato, A. Saito, M. Takagi, and Y. Niizuma. 2011. Within- and among-clutch variation in maternal yolk testosterone level in the black-tailed gulls *Larus crassirostris*. *Ornithol. Sci.* 10:21–25.
- Tschirren, B., A. N. Rutstein, E. Postma, M. Mariette, and S. C. Griffith. 2009. Short- and long-term consequences of early developmental conditions: a case study on wild and domesticated zebra finches. *J. Evol. Biol.* 128:103–115.
- Uller, T., J. Eklöf, and S. Andersson. 2005. Female egg investment in relation to male sexual traits and the potential for transgenerational effects in sexual selection. *Behav. Ecol. Sociobiol.* 57:584–590.
- Uller, T., S. Nakagawa, and S. English. 2013. Weak evidence for anticipatory parental effects in plants and animals. *J. Evol. Biol.* 26:2161–2170.
- Vazquez, A. I., D. M. Bates, G. J. M. Rosa, D. Gianola, and K. A. Weigel. 2010. Technical note: an R package for fitting generalized linear mixed models in animal breeding. *J. Anim. Sci.* 88:497–504.
- Waddington, C. H. 1942. Canalization of development and the inheritance of acquired characters. *Nature* 381:563–565.

- Wang, D., W. Forstmeier, K. Martin, A. Wilson, and B. Kempenaers. 2020. The role of genetic constraints and social environment in explaining female extra-pair mating. *Evolution* 74:544–558.
- Wang, D., N. Kempenaers, B. Kempenaers, and W. Forstmeier. 2017. Male zebra finches have limited ability to identify high-fecundity females. *Behav. Ecol.* 28:784–792.
- Wilson, K. M., A. Tatarenkov, and N. T. Burley. 2019. Early life and trans-generational stressors impact secondary sexual traits and fitness. *Behav. Ecol.* 30:830–842.
- Zann, R. 1996. *The zebra finch: a synthesis of field and laboratory studies.* Oxford Univ. Press, New York.
- Zizzari, Z. V., N. M. van Straalen, and J. Ellers. 2016. Transgenerational effects of nutrition are different for sons and daughters. *J. Evol. Biol.* 29:1317–1327.

Associate Editor: T. Flatt  
Handling Editor: T. Chapman

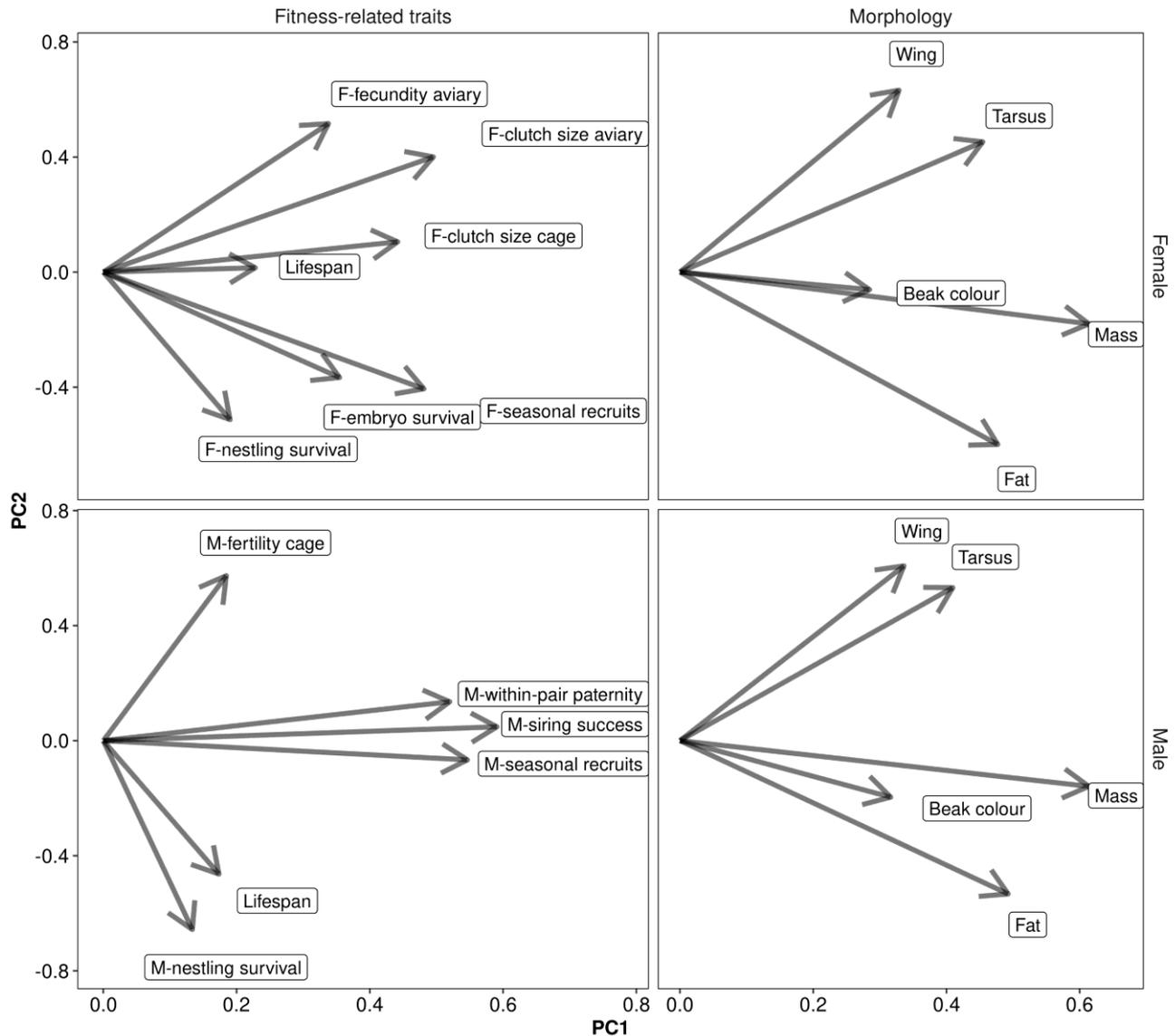
### *Supporting Information*

Additional supporting information may be found online in the Supporting Information section at the end of the article.

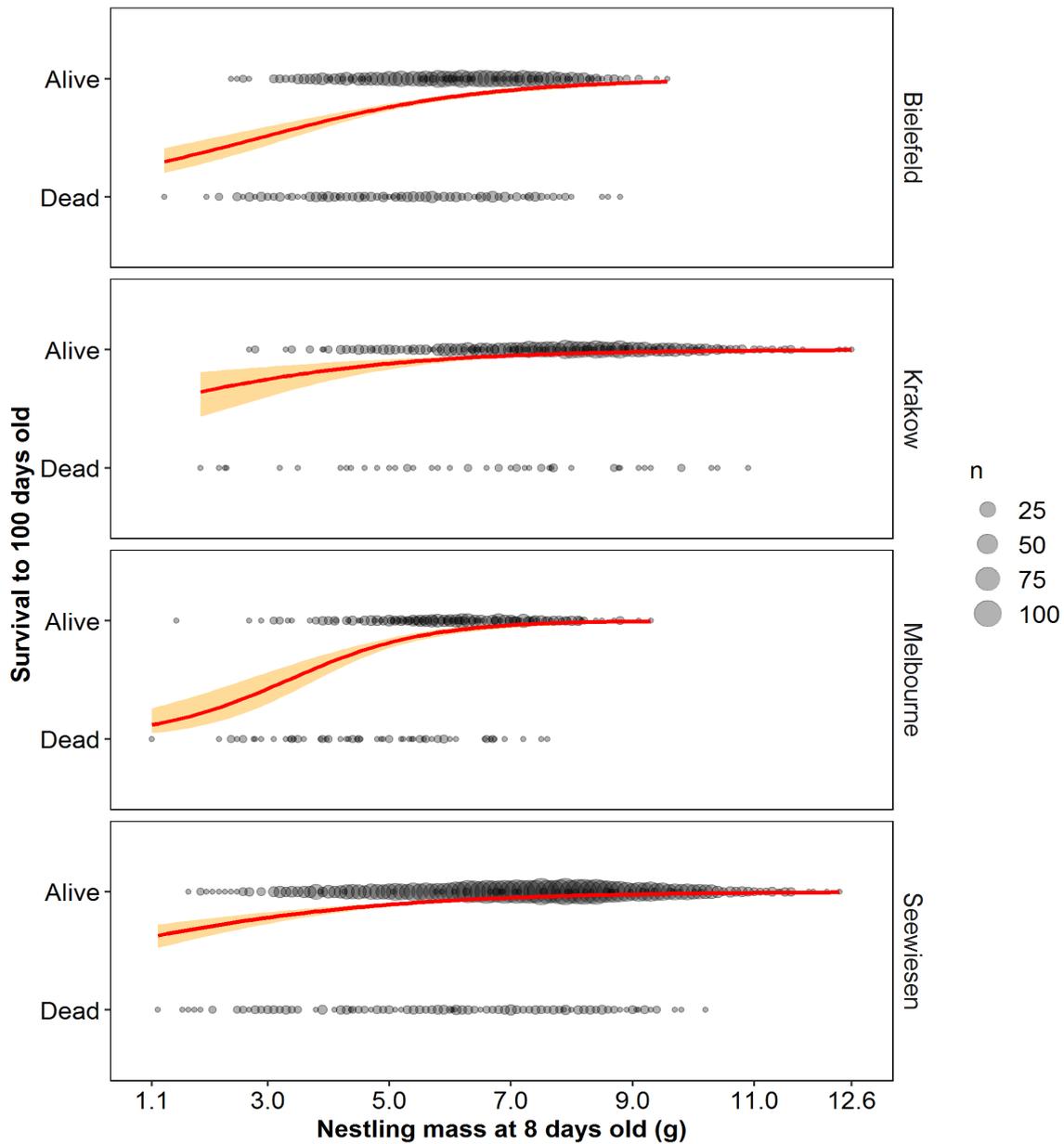
Supplementary material

## Supplementary Materials

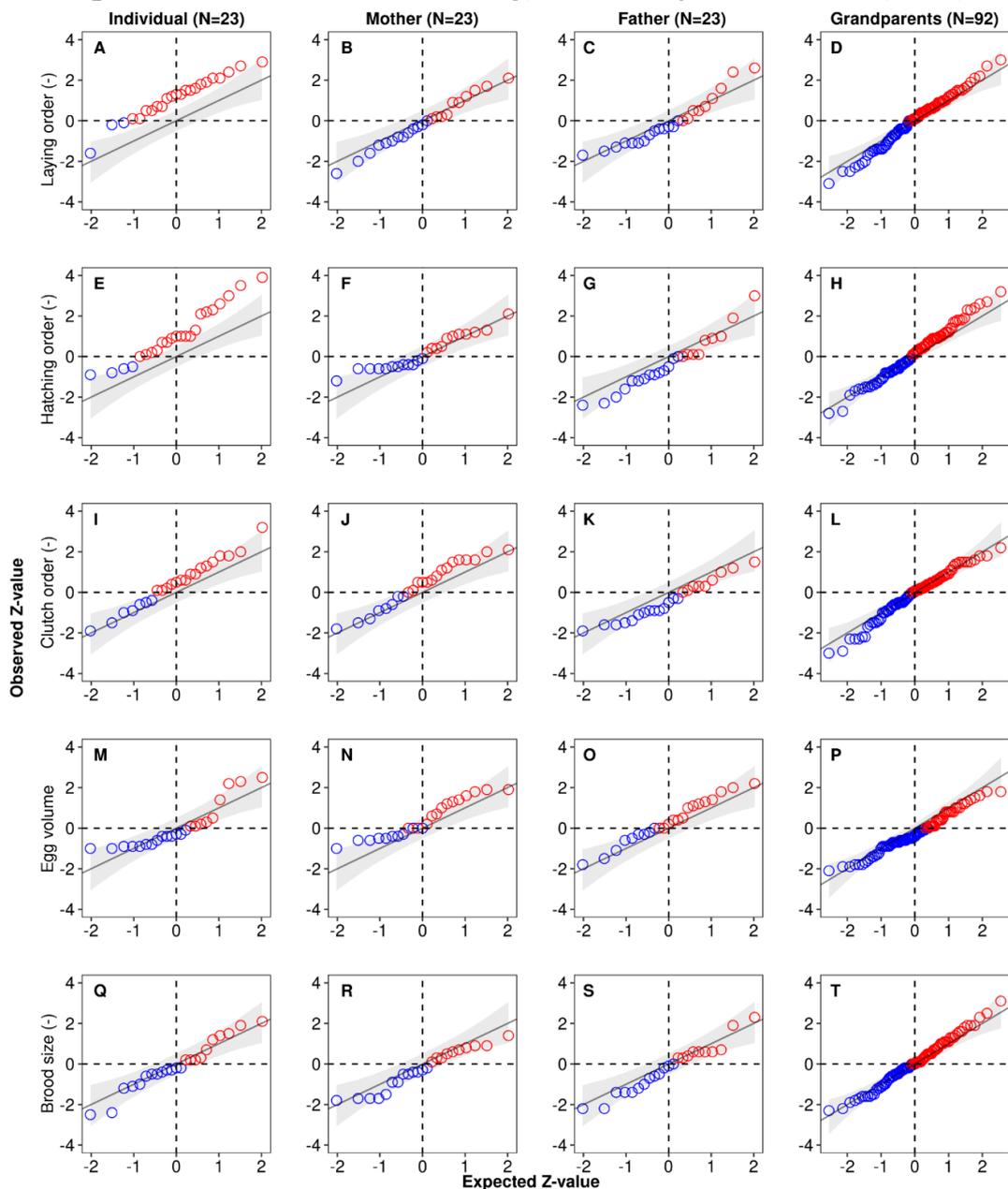
**Figure S1.** The first two principal components of the phenotypic correlation-matrices of fitness-related and morphological traits of female and male zebra finches. Shown are eigenvectors of the trait components, calculated from the correlation-matrices in Table S6. For visualization purpose only, the missing values were replaced by the median of all entries of each matrix. Note how identical the male and female morphological traits and how similar the male and female fitness-related traits were distributed in the PC1 and PC2 space. Also note that female fecundity in aviary, female clutch size in cage and aviary were clustered together, indicating strong correlations. Additional information sees Table S1.



**Figure S2.** Individual survival to adult (from 8 to 100 days old) as a function of nestling body mass at 8 days old for Bielefeld, Krakow, Melbourne and Seewiesen populations (from top to bottom). Birds in Seewiesen and Krakow populations are domesticated ones that are larger in body size whereas Bielefeld and Melbourne populations are recently wild-derived and smaller in body size. Seewiesen population is the focal population used in this study. Red lines are logistic regressions with 95% CIs, circle size shows sample size. Note the high mortality rate for light weighted nestlings (<5g) at 8 days old.



**Figure S3.** ZZ-plots of expected versus observed Z-values of effects of early developmental conditions, i.e. laying order (A-D), hatching order (E-H), clutch order (I-L), egg volume within female variation (M-P) and brood size (Q-T), experienced by the focal individual itself (A, E, I, M, Q), its mother (B, F, J, N, R), its father (C, G, K, O, S) and its four grandparents (D, H, L, P, T) on 10 morphological and 13 fitness-related traits of male and female zebra finches. Four out of the six indicators of conditions were multiplied by -1 (indicated by (-)) such that positive effect sizes reflect better performance under supposedly better conditions. N indicates the number of tests. Red indicates that the sign of the estimate is in the expected direction, blue indicates that the sign is in the opposite direction. Lines of identity (where observation equals prediction) and their 95% CIs are shown. See Fig 2E-2H in the main text for the ZZ-plots of mass at 8 days old. Note the number of tests (dots). Also note that only the effects of individual its own laying (A) and hatching order (E) have more effects that are significant (observed Z-values are larger than 2, above the line of identity) with the expected direction (in red).



**Table S1.** Pairwise correlation matrices of female and male fitness related and morphological traits for all birds used in the initial tests. Correlation coefficients were calculated using R function 'cor'. '-' indicates no data available. Additional descriptions see Fig. S1.

This table contains 27 rows.

**Table S2.** Focal fixed effect estimates for six early developmental stressors experienced by the individual itself, its parents and its grandparents, on morphological and fitness-related traits in the focal individual. Morphological traits are beak color, fat deposition, mass, tarsus and wing length of female and male zebra finches, estimated in animal models. Fitness related traits of male and female zebra finches were estimated in mixed effect models. All focal fixed effects and dependent variables were Z-scored, 95% CIs were calculated while accounting for multiple testing within each model (see Methods for details).

This table contains 966 rows.

**Table S3.** Focal fixed effect estimates of parent (mother or father) -offspring similarity in nestling mass when 8 days old, on morphological and fitness-related traits of the offspring. Morphological traits are beak color, fat deposition, mass, tarsus and wing length of female and male zebra finches, estimated in animal models. Fitness related traits of male and female zebra finches were estimated in mixed effect models. All focal fixed effects and dependent variables were Z-scored, 95% CIs were calculated while accounting for multiple testing within each model (see Methods for details).

This table contains 47 rows.

**Table S4.** Confounding fixed effect estimates for all models of conditional transfer and anticipatory effect on 10 morphological and 13 fitness related traits. Focal fixed effects of stressor see Tables S2 and S3. Morphological traits are beak colour, fat deposition, mass, tarsus and wing length of female and male zebra finches, estimated in animal models. Fitness related traits of male and female zebra finches were estimated in mixed effect models. All dependent variables and covariates were Z-scored, 95% CIs were calculated while accounting for multiple testing within each model (see Methods for details).

This table contains 7042 rows.

For details of Tables S1-S4 please see online Supporting information:

<https://onlinelibrary.wiley.com/doi/full/10.1111/evo.14026>

**Table S5.** Averaged effect size estimates (weighted by the uncertainty of each estimates, in Table S2) of early developmental stress experienced by the individual (direct effects) on morphology and fitness-related traits (Fig. 1). The 95% CIs were calculated while accounting for conducted 12 tests, see Methods for details.

Trait	Stressor	Estimate	SE	Z	95% CIs	P*
Morphological traits (N=10)	Mass at 8 days old	0.229	0.015	15.6	[0.186,0.272]	<0.0001
Fitness-related traits (N=13)	Mass at 8 days old	0.070	0.017	4.2	[0.021,0.119]	<0.0001
Morphological traits (N=10)	Clutch order (-)	0.022	0.016	1.3	[-0.025,0.068]	0.18
Fitness-related traits (N=13)	Clutch order (-)	0.014	0.017	0.8	[-0.035,0.062]	0.41
Morphological traits (N=10)	Laying order (-)	0.047	0.015	3.2	[0.004,0.09]	0.002
Fitness-related traits (N=13)	Laying order (-)	0.031	0.017	1.8	[-0.018,0.079]	0.07
Morphological traits (N=10)	Brood size (-)	0.013	0.015	0.8	[-0.032,0.057]	0.42
Fitness-related traits (N=13)	Brood size (-)	-0.021	0.017	-1.3	[-0.07,0.028]	0.21
Morphological traits (N=10)	Hatching order (-)	0.052	0.015	3.5	[0.009,0.094]	0.001
Fitness-related traits (N=13)	Hatching order (-)	0.020	0.017	1.2	[-0.029,0.068]	0.25
Morphological traits (N=10)	Egg volume	0.019	0.015	1.3	[-0.024,0.062]	0.20
Fitness-related traits (N=13)	Egg volume	-0.022	0.017	-1.3	[-0.071,0.027]	0.19

\* Estimated in the meta-analytic linear regression model.

**Table S6.** Averaged effect size estimates (weighted by the uncertainty of each estimates, in Table S2) of nestling mass at 8 days old experienced by the individual itself, its parents and grandparents on its morphology and fitness-related traits by sex (Fig. 2). 95% CIs were calculated while accounting for multiple testing within each model, see Methods.

Trait	Trait of	Stress experienced by	Estimate	SE	Z	95% CIs	P*
Morphological traits (N=5)	Female	Individual	0.254	0.021	12.2	[0.188,0.32]	< 0.0001
Morphological traits (N=5)	Female	Mother	-0.014	0.023	-0.6	[-0.088,0.061]	0.56
Morphological traits (N=5)	Female	Father	0.020	0.024	0.8	[-0.055,0.095]	0.40
Morphological traits (N=5)	Female	Maternal grandmother	0.003	0.024	0.1	[-0.075,0.081]	0.91
Morphological traits (N=5)	Female	Maternal grandfather	-0.026	0.024	-1.1	[-0.101,0.05]	0.29
Morphological traits (N=5)	Female	Paternal grandmother	0.006	0.024	0.3	[-0.07,0.083]	0.79
Morphological traits (N=5)	Female	Paternal grandfather	-0.010	0.024	-0.4	[-0.085,0.065]	0.67
Fitness-related traits (N=7)	Female	Individual	0.073	0.023	3.2	[0,0.146]	0.002
Fitness-related traits (N=7)	Female	Mother	0.071	0.024	3.0	[-0.004,0.147]	0.003
Fitness-related traits (N=7)	Female	Father	-0.013	0.024	-0.5	[-0.088,0.063]	0.59
Fitness-related traits (N=7)	Female	Maternal grandmother	-0.020	0.024	-0.8	[-0.096,0.056]	0.41
Fitness-related traits (N=7)	Female	Maternal grandfather	-0.019	0.024	-0.8	[-0.095,0.058]	0.44
Fitness-related traits (N=7)	Female	Paternal grandmother	0.007	0.024	0.3	[-0.07,0.085]	0.76
Fitness-related traits (N=7)	Female	Paternal grandfather	-0.026	0.024	-1.1	[-0.103,0.051]	0.28
Morphological traits (N=5)	Male	Individual	0.204	0.021	9.9	[0.139,0.27]	< 0.0001
Morphological traits (N=5)	Male	Mother	0.040	0.023	1.7	[-0.034,0.114]	0.09
Morphological traits (N=5)	Male	Father	-0.021	0.023	-0.9	[-0.093,0.051]	0.36
Morphological traits (N=5)	Male	Maternal grandmother	-0.007	0.024	-0.3	[-0.085,0.07]	0.76
Morphological traits (N=5)	Male	Maternal grandfather	-0.062	0.023	-2.7	[-0.135,0.012]	0.01
Morphological traits (N=5)	Male	Paternal grandmother	0.016	0.023	0.7	[-0.058,0.091]	0.48
Morphological traits (N=5)	Male	Paternal grandfather	-0.007	0.023	-0.3	[-0.08,0.066]	0.76
Fitness-related traits (N=6)	Male	Individual	0.067	0.025	2.7	[-0.012,0.145]	0.01
Fitness-related traits (N=6)	Male	Mother	-0.004	0.026	-0.2	[-0.086,0.077]	0.87
Fitness-related traits (N=6)	Male	Father	-0.038	0.025	-1.5	[-0.119,0.043]	0.14
Fitness-related traits (N=6)	Male	Maternal grandmother	-0.027	0.026	-1.0	[-0.108,0.055]	0.30
Fitness-related traits (N=6)	Male	Maternal grandfather	-0.027	0.026	-1.0	[-0.109,0.055]	0.30
Fitness-related traits (N=6)	Male	Paternal grandmother	-0.018	0.027	-0.7	[-0.103,0.066]	0.49
Fitness-related traits (N=6)	Male	Paternal grandfather	0.014	0.026	0.5	[-0.069,0.097]	0.60

\* Estimated in the meta-analytic linear regression model.

**Table S7.** Averaged effect size estimates (weighted by the uncertainty of each estimates, in Table S3) of parent-offspring similarity in nestling mass at 8 days old on its morphology and fitness-related traits (Fig. 4). 95% CIs were calculated while accounting for conducting 8 tests, see Methods for details.

Trait	Trait of	Nestling mass matching with	Estimate	SE	Z	95% CIs	P*
Morphological traits (N=5)	Female	Mother	0.003	0.015	0.2	[-0.039,0.046]	0.82
Morphological traits (N=5)	Female	Father	0.047	0.014	3.3	[0.006,0.089]	0.002
Fitness-related traits (N=7)	Female	Mother	0.047	0.016	3.0	[0.002,0.092]	0.005
Fitness-related traits (N=7)	Female	Father	0.035	0.016	2.2	[-0.01,0.081]	0.03
Morphological traits (N=5)	Male	Mother	0.008	0.015	0.5	[-0.034,0.049]	0.61
Morphological traits (N=5)	Male	Father	0.027	0.014	1.9	[-0.014,0.068]	0.07
Fitness-related traits (N=6)	Male	Mother	0.035	0.017	2.1	[-0.014,0.083]	0.047
Fitness-related traits (N=6)	Male	Father	0.024	0.017	1.4	[-0.025,0.073]	0.17

\* Estimated in the meta-analytic linear regression model.

**Table S8.** Meta-summarized effect size in a weighted mixed-effect model of transgenerational effect, i.e. parents and four grandparents, of mass at 8 days old on offspring morphological and fitness related traits. All focal fixed effects were taken from Table S2 (see Methods for details).

<b>N estimates</b>	<b>Predictor</b>	<b>Estimate</b>	<b>SE</b>	<b>Z</b>	<b>95% CIs*</b>
	Intercept	-0.007	0.005	-1.5	[-0.017,0.002]
138	Var <sub>Trait category x Ancestor</sub> <sup>†</sup> (66 levels)	0.001			
	Var <sub>Residual</sub> <sup>†</sup>	0.039			

\*The 95% CIs were calculated as estimate  $\pm$  1.96xSE.

<sup>†</sup>VarResidual: residual variation. Var<sub>Trait category x Ancestor</sub>: variance component of pairwise combinations of ancestors (6 levels, i.e. two parents and four grandparents) and categories of traits that are correlated (11 levels, i.e. female fecundity-related traits: female fecundity in aviary, female clutch size in cage and aviary; female offspring survival: female seasonal recruits, female's embryo and nestling survival; female lifespan; female size: female's tarsus and wing length, female mass and abdominal fat score; female beak colour; male siring ability: male seasonal recruits, male within-pair paternity and siring success in aviary; male fertility in cage; male lifespan; male nestling survival; male size: male's tarsus and wing length, male mass and abdominal fat score; male beak colour).

**Table S9.** Meta-summarized effect size in a weighted mixed-effect model of transgenerational anticipatory effect, i.e. similarity between parent-offspring in nestling mass at 8 days old, on offspring morphological and fitness related traits. All focal fixed effects were taken from Table S3 (see Methods for details).

<b>N estimates</b>	<b>Predictor</b>	<b>Estimate</b>	<b>SE</b>	<b>Z</b>	<b>95% CIs*</b>
	Intercept	0.028	0.006	4	[0.016,0.040]
46	Var <sub>Trait category x Ancestor</sub> † (22 levels)	0.0002			
	Var <sub>Residual</sub> †	0.032			

\*The 95% CIs were calculated as estimate  $\pm$  1.96xSE.

†Var<sub>Residual</sub>: residual variation. Var<sub>Trait category x Ancestor</sub>: variance component of pairwise combinations of offspring nestling mass matching with father of mother (2 levels for the Ancestor) and categories of traits that are correlated (11 levels for the Trait category, i.e. female fecundity-related traits: female fecundity in aviary, female clutch size in cage and aviary; female offspring survival: female seasonal recruits, female's embryo and nestling survival; female lifespan; female size: female's tarsus and wing length, female mass and abdominal fat score; female beak colour; male siring ability: male seasonal recruits, male within-pair paternity and siring success in aviary; male fertility in cage; male lifespan; male nestling survival; male size: male's tarsus and wing length, male mass and abdominal fat score; male beak colour).

**Table S10.** Effect estimates of mother's mass at 8 days old on her daughter's fecundity, clutch size cage and aviary from an independent confirmatory dataset and the meta-summarized effect sizes for the initial and confirmatory tests (Fig. 5). In trait-based models, all focal fixed effects and dependent variables were Z-scored (see Methods for details). The 95% CIs were calculated as estimate  $\pm$  1.96xSE, see Material and methods.

Trait	Stressor	Population	Nobs*	Estimate	SE	Z	95% CIs	Data set
Female clutch size aviary	Mother-daughter similarity in mass at 8 days old		810	0.029	0.051	0.6	[-0.117,0.174]	Confirmatory
Female clutch size cage	Mother-daughter similarity in mass at 8 days old		1249	-0.041	0.047	-0.9	[-0.173,0.092]	Confirmatory
Female fecundity	Mother-daughter similarity in mass at 8 days old		31	-0.073	0.196	-0.4	[-0.584,0.439]	Confirmatory
Female tarsus	Father-daughter similarity in mass at 8 days old	Bielefeld	333	-0.068	0.043	-1.6	[-0.184,0.047]	Confirmatory
Female tarsus	Father-daughter similarity in mass at 8 days old	Krakow	290	0.076	0.056	1.4	[-0.072,0.225]	Confirmatory
Female tarsus	Father-daughter similarity in mass at 8 days old	Seewiesen	447	0.143	0.044	3.3	[0.032,0.254]	Confirmatory
Female wing	Father-daughter similarity in mass at 8 days old	Bielefeld	332	-0.090	0.053	-1.7	[-0.23,0.051]	Confirmatory
Female wing	Father-daughter similarity in mass at 8 days old	Krakow	287	0.055	0.057	1.0	[-0.097,0.207]	Confirmatory
Female wing	Father-daughter similarity in mass at 8 days old	Seewiesen	222	0.029	0.065	0.4	[-0.139,0.197]	Confirmatory
Female clutch size cage	Mother's nestling mass		1249	0.049	0.046	1.0	[-0.081,0.178]	Confirmatory
Female clutch size aviary	Mother's nestling mass		810	-0.051	0.047	-1.1	[-0.185,0.084]	Confirmatory
Female fecundity	Mother's nestling mass		31	-0.076	0.263	-0.3	[-0.751,0.598]	Confirmatory
Meta-daughter fecundity†	Mother's nestling mass		3	-0.007	0.045	-0.1	[-0.142,0.129]	Confirmatory
Meta-daughter size†	Father-daughter similarity in nestling mass		6	0.024	0.028	0.8	[-0.062,0.109]	Confirmatory
Meta-daughter fecundity†	Mother-daughter similarity in nestling mass		3	-0.015	0.045	-0.3	[-0.15,0.121]	Confirmatory
Meta-daughter fecundity†	Mother's nestling mass		3	0.120	0.040	3.0	[0.0004,0.24]	Initial‡
Meta-daughter size†	Father-daughter similarity in nestling mass		2	0.085	0.035	2.4	[-0.022,0.192]	Initial‡
Meta-daughter fecundity†	Mother-daughter similarity in nestling mass		3	0.089	0.038	2.3	[-0.027,0.205]	Initial‡

\*For female clutch size: number of clutches. Fecundity: number of birds in breeding seasons. Female tarsus and wing length: number of females. Meta: number of estimates.

†Meta summarized effect sizes among female fecundity, clutch size cage and aviary; or among models of female tarsus and wing lengths.

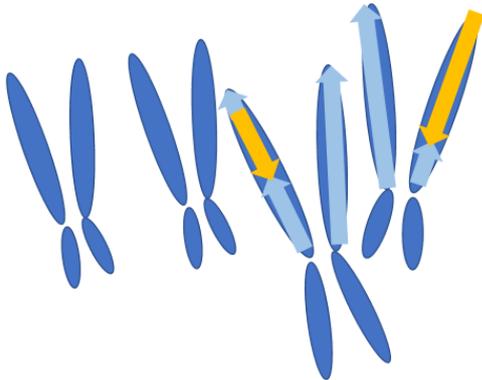
‡Initial scan of effects of mother's nestling mass on female fecundity, clutch size cage and aviary are in Table S2. Initial tests of effects of father-daughter similarity in nestling mass on female tarsus and wing and mother-daughter similarity in nestling mass on female fecundity, clutch size cage and aviary are in Table S3.



## Chapter 3

### Microchromosomal inversions

"We are macrochromosomes,  
we have *inversions!*"



"We are microchromosomes,  
we also have *inversions!*"



## Chapter 3

### Overdominance effects of a microchromosomal inversion on multiple fitness components in the zebra finch

Yifan Pei, Wolfgang Forstmeier, Alexander Suh, Anne-Marie Dion-Côté, Ulrich Knief, Jochen Wolf, Bart Kempenaers

A typical avian haploid genome contains about 10 macrochromosomes and 30 microchromosomes. The few macrochromosomes make up the majority of the genome, and these chromosomes have received much attention, particularly in studies of the evolution and fitness consequences of structural variants like inversion polymorphisms. In contrast, the numerous avian microchromosomes have received less attention, often due to methodological difficulties. Using genome-wide SNP-markers (single-nucleotide polymorphisms) and shared barcodes of linked-read sequences from multiple wild and captive zebra finches *Taeniopygia guttata*, we here describe two large inversions on the microchromosomes *Tgu26* and *Tgu27*. Both inversions are about 3 Mb long and, in our sample, show minor allele frequencies of 0.29 and 0.42, respectively. For both microchromosomes, phylogenetic and genetic variability analyses of individual-based haplotypes indicate that individuals that are homozygous for the minor allele cluster together and form a derived clade, showing fewer heterozygous sites compared to individuals homozygous for the major allele, which form an ancestral clade. Hence, the minor alleles appear to be the derived inversion types in both *Tgu26* and *Tgu27*. Further, we used two SNPs to tag each of the two inversions, and genotyped about 5,000 captive zebra finches for which a wide range of fitness-related traits had been measured. Based on these markers we found that individuals that were likely heterozygous for the inversion on *Tgu27* showed significant heterosis, including increased longevity, siring success, fecundity, and higher rates of embryo and nestling survival, while no such effects were found for *Tgu26*. Our study suggests that overdominance for fitness can explain why the inversion on the microchromosome *Tgu27* has spread and is being maintained at an allele frequency of about 0.42 (wild) to 0.52 (captivity).

Prepared as:

Pei, Y., Forstmeier, W., Suh, A., Dion-Côté, A.M., Knief, U., Wolf, J., & Kempenaers, B. Overdominance effects of a micro-chromosomal inversion on multiple fitness components in the zebra finch

## Introduction

Chromosomal inversions that physically link hundreds of genes have received much research attention (1), because they may explain extraordinary within-species variations in morphology, behavior and fitness, e.g. in birds (2–6), insects (7–10) and plants (11, 12). In avian species, such inversions were frequently found on macrochromosomes (1), while microchromosomes (which are typical for birds) were usually omitted from studies of intrachromosomal rearrangements, particularly inversions, e.g. (2, 6, 7). Microchromosomes are difficult to distinguish cytogenetically (13, 14), as they are small (usually <2.5  $\mu\text{m}$  or 20 Mb in length) and typically acrocentric, that is with the centromere located on one end. Moreover, in genome-wide scans, microchromosomes are often insufficiently covered with markers (15). This may explain why inversions on microchromosomes have rarely been found. Nevertheless, avian microchromosomes are known for being enriched in gene density, for showing conserved gene synteny, and for exhibiting high recombination rates compared to the macrochromosomes (13, 16–19). To better understand the evolutionary biology of a species and the evolution of avian genomes, systematic studies on inversions on microchromosomes are needed.

The zebra finch *Taeniopygia guttata* is one of the most studied bird species with a high quality reference genome available (20), consisting of seven macrochromosomes *Tgu1-Tgu12*, *Tgu1A* and *TguZ* (62 - 152 Mb) and 33 microchromosomes (seven of them range from 20 to 40 Mb and the rest range from 1 to 20 Mb). Zebra finches have at least four large (12 to 63Mb) intraspecific inversion polymorphisms that segregate at about 50% allele frequency in the wild, that are located on chromosomes *Tgu5*, *Tgu11*, *Tgu13* and *TguZ* (15, 17). In view of the high minor allele frequencies, it is a plausible hypothesis that these polymorphisms could be maintained by heterosis, but an initial study from our laboratory found no indications of heterozygote advantage (15). Later it was found that the polymorphism on *TguZ* is likely maintained by overdominance, as heterozygous males show increased sperm motility and siring success (5, 6). Moreover, the polymorphism on *Tgu11* might be maintained by antagonistic-pleiotropic effects on multiple life history traits, where the derived allele appears to additively increase female fecundity and male siring success, but to reduce individual survival in the homozygous state (Pei et al., In preparation). However, in our previous scan for inversion polymorphisms (15), there was insufficient power to detect and study inversion polymorphisms on the majority of the microchromosomes, especially those that are less than 10 Mb in length (N=19 (20, 21)). For instance, the two microchromosomes *Tgu26* and *Tgu27* were suspected to contain inversions but were dropped due to very few informative SNP-markers (15).

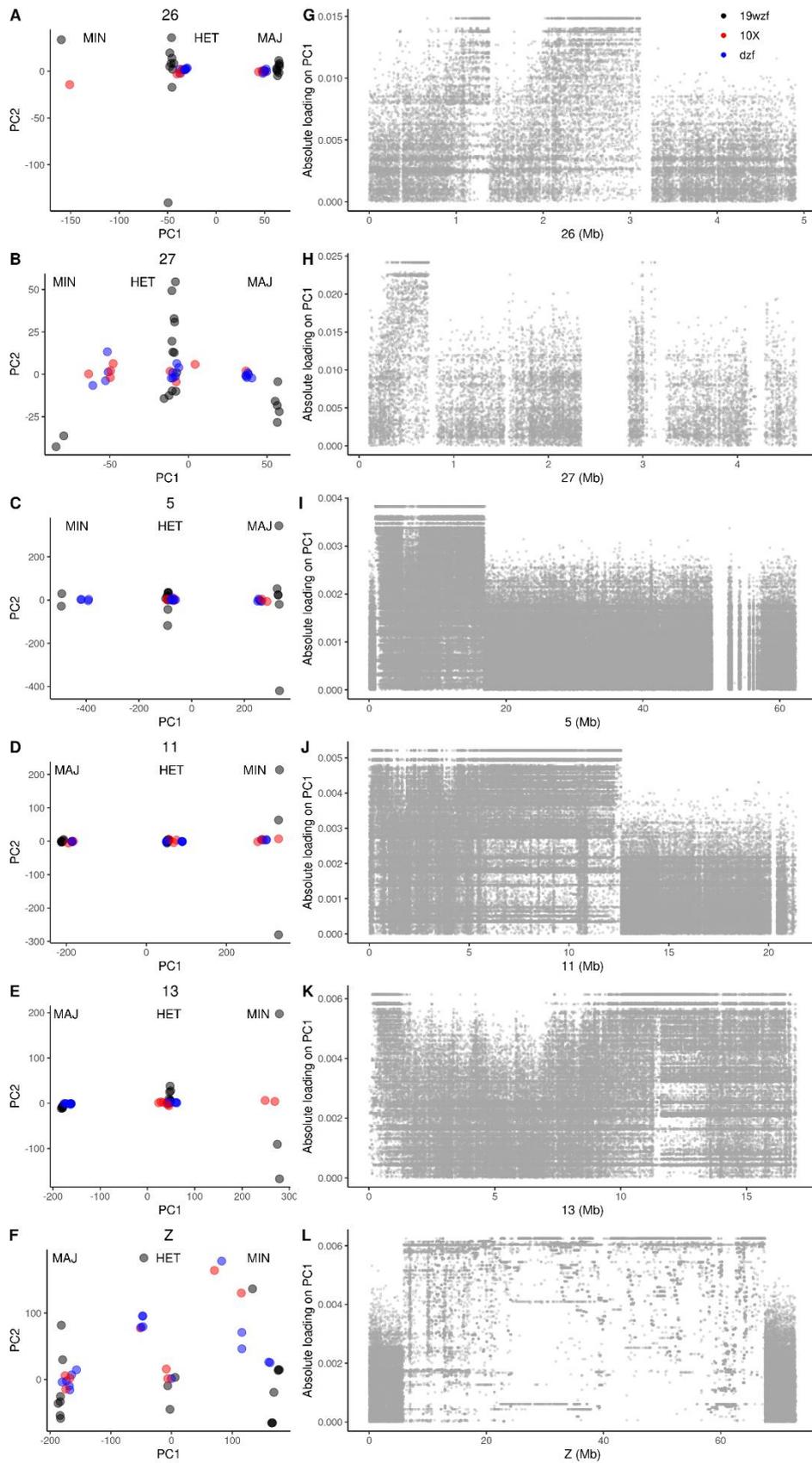
In the present study, we aim to investigate large inversion polymorphisms (spanning at least 1 Mb) on the zebra finch microchromosomes using conventional Illumina and 10X linked-read sequencing data of multiple wild and captive individuals. We first screened for large structural polymorphisms on all zebra finch chromosomes using principal component analysis (PCA) with SNPs called on Illumina data from 19 wild zebra finches taken from (22). Second, we verified the detected inversions using 10X linked-read sequencing data from eight additional individuals that mostly came from captivity. Then we defined the ancestral state for all identified chromosomal inversion polymorphisms using the 19 wild, the eight linked-read sequenced zebra finches, and 16 additional captive zebra finches with different genetic backgrounds. Lastly, we estimated the fitness consequences of the newly detected inversions on microchromosomes to study the evolutionary mechanism that might maintain these polymorphisms.

## Results

### *Detection of inversions on microchromosomes Tgu26 and Tgu27 using PCA*

Principal component analysis of SNPs revealed that the four chromosomes *Tgu5*, *Tgu11*, *Tgu13*, and *TguZ* with known inversion polymorphisms (15, 17) as well as the two microchromosomes *Tgu26* and *Tgu27* showed the typical patterns of inversion polymorphisms with two major haplotypes (**Fig. 1**, and **Fig. S1**). The 19 wild and the 24 mostly captive zebra finches (**Supplementary Table S1**) were grouped into three clusters along PC1, with presumed heterozygous individuals falling into the central cluster, and the two types of homozygotes on either side (**Fig. 1A-F**). For convenience and easy comparison with previous studies, we here named the haplotypes for each of the six chromosomes based on their allele frequencies, as minor and major alleles, among the 19 wild zebra finches. Minor allele frequencies were roughly comparable ( $r = 0.73$ ,  $n = 6$ ) between the 19 wild (range 0.29 – 0.50) and the 24 captive (0.27 – 0.50) zebra finches used in our study (**Fig. S1**; **Supplementary Table S2**). For the two microchromosomes *Tgu26* and *Tgu27*, we found their MAFs to be 0.29 and 0.42, respectively, among the 19 wild birds, and 0.27 and 0.52 among the 24 mostly captive birds (**Fig. 1A-B** and **Supplementary Table S2**; for details of population background of all birds see **Table S1**).

**Fig. 1.** Principal component analysis of genome-wide SNPs among 19 wild zebra finches (black dots; left) and the loadings of SNPs (grey) on PC1 (right) for chromosomes that show signs of segregating inversion polymorphisms, i.e. microchromosomes *Tgu26* (**A,G**), *Tgu27* (**B,H**) and for previously identified inversions on chromosomes *Tgu5* (**C,I**), *Tgu11* (**D,J**), *Tgu13* (**E,K**) and *TguZ* (**F,I**) (15). (**A-F**) Each dot represents an individual, saturated colors (black, red and blue) indicate multiple overlapping individuals. PC1 and PC2 scores of 24 additional birds (red and blue dots) were calculated from the SNP-loadings derived on the 19 wild zebra finches (black dots; for details see **Methods**). Red identifies the eight individuals that were sequenced using 10X linked-read technology, and blue depicts 16 captive zebra finches that were additionally included for haplotype and evolutionary analysis. Individuals homozygous for the major or the minor alleles and heterozygous individuals were defined based on clusters of individuals along PC1. Major (MAJ) and minor (MIN) alleles were defined based on the allele frequency in the 19 wild birds (black dots). (**G-I**) Each grey dot is a SNP, and the saturation of the grey color indicates overlapping SNPs. SNP positions are based on the old assembly TaeGut1 (21).

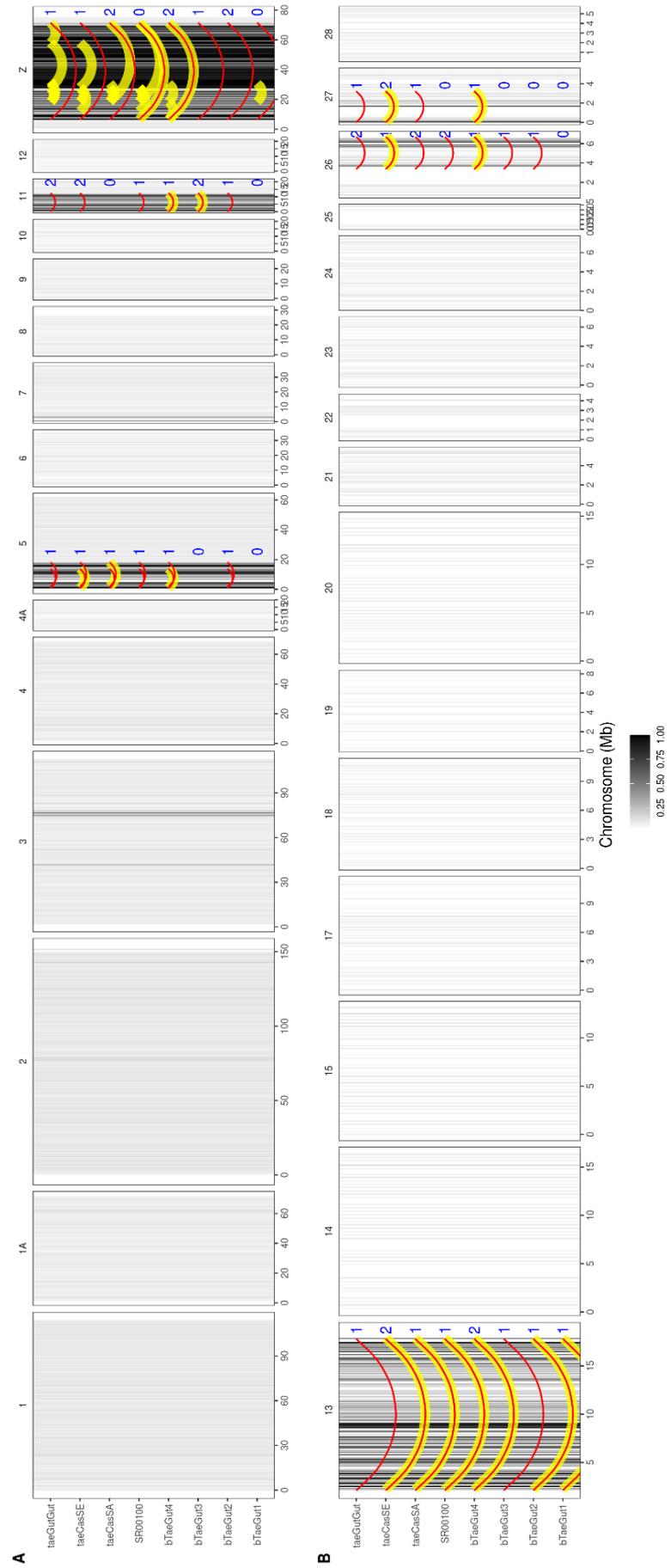


*Verification and characterization of inversion polymorphism using linked-read sequencing data*

To verify and characterize the putative inversions on the two microchromosomes, we first mapped the linked-read sequencing libraries of seven *castanotis* and one *castanotis* x *guttata* hybrid zebra finch against the new zebra finch assembly bTaeGut1.202104, where the assemblies of microchromosomes were most complete (20) (**Fig S2**). For these eight individuals and the reference, the (putative) inversion types of chromosomes *Tgu5*, *Tgu11*, *Tgu13*, *TguZ*, *Tgu26* and *Tgu27* were predicted via the SNP-loadings of PCA among the 19 wild zebra finches (**Fig. 1**; **Supplementary Fig. S1**; for details see **Methods**).

We found that large (>2 Mb) intra-chromosomal structural variant calls almost exclusively occurred on the aforementioned chromosomes with inversions (**Fig. 2** and **Fig. S3**). This enabled us to identify the breakpoints for the inversions on *Tgu5*, *Tgu11*, *Tgu13*, *Tgu26* and *Tgu27*, with an average resolution of 8 Kb (ranging from 43 to 58,094 bp) on the newest zebra finch assembly (20) (**Fig. 2**; **Supplementary Table S2**; also see **Supplementary Fig. S4**). For chromosomes *Tgu11* and *Tgu26*, each chromosome was called for only a single pair of distant barcode-interactions (**Figs. 2, S3, Supplementary Table S3**), suggesting a single structural change. Using a more relaxed cut-off (>1Mb and singletons are included), chromosomes *Tgu13* and *Tgu27* both had a second variant call with the smaller one sitting inside the larger one (**Supplementary Fig. S3**), suggesting that a second inversion variant was segregating within one of the major haplotypes (included inversions; **Supplementary Fig. S3**; e.g. **Fig. 1B,E** and **Supplementary Fig. S1C,F**; also see (15)). Chromosome *Tgu5* contained two overlapping structural variants (**Fig. 2A**). This could be either a mis-assembly of the regions that contain the breakpoints or *Tgu5* containing a double inversion. Additionally, we found that the sex chromosome *TguZ* had the greatest number of calls of unexpected distant barcode interactions (**Fig. 2A**), suggesting that *TguZ* contained many small structural variations. Because we are focusing on the major haplotypes of the inversion (**Fig. 1** and **Fig. S1**), in the following analyses, we only focused on the longest pairs of breakpoints with shared barcodes on *TguZ*.

**Fig. 2.** Genome-wide distribution of unexpected intra-chromosomal barcode sharing over long distances (yellow and red curves) in 8 sequenced individuals (rows) relative to the reference individual (i.e. the new zebra finch assembly bTaeGut1.202104 (20)). Yellow depicts high quality calls of intra-chromosomal barcode interactions detected by Long Ranger that are more than 2 Mb apart and that were found in more than one individual (for outlier calls from single individuals see **Supplementary Fig. S3**). Red depicts manually validated interactions based on raw barcode interaction plots (blindly called from **Supplementary Fig. S4** in chromosomes that showed a signal of an inversion in the PCA (**Fig.1**; also see **Supplementary Fig. S1**). Blue numbers indicate the expected number of alternative inversion types (relative to the reference; 0 = same as reference) (**Supplementary Fig. S1**). Each vertical bar is a SNP that was used in Knief *et al.* (2016) (15). SNP positions were lifted from the old assembly TaeGut1 (21) to the new assembly bTaeGut1.202104 (20). The blackness of each bar indicates the highest level of linkage disequilibrium (LD) of this SNP with another SNP that is at least 1 Mb apart within the same chromosome, among 948 wild zebra finches (15). The large blocks of SNPs on chromosomes *Tgu5*, *I1*, *I3*, *Z*, that are in high LD (black bars) indicate inversions, whereas the microchromosomes *Tgu26* and *Tgu27* showed only weak signals of LD suggesting the presence of inversions by Knief *et al.* (2016) (15). Note that all high-quality distant barcode interactions that were detected by Long Ranger were on chromosomes that contained inversions (i.e. N=39 yellow curves on chromosomes that show a signal of segregating inversions, i.e. *Tgu5*, *Tgu11*, *Tgu13*, *TguZ*, *Tgu26* and *Tgu27*).



To check for false-positive signals (e.g. due to mis-assembly or repetitive elements in the reference) and false-negative calls (e.g. coverage of barcodes below detection threshold), we manually called from **Supplementary Fig. S4** the absence (i.e. being identical to the reference haplotype) and presence (i.e. heterozygous or homozygous for the alternative haplotype) of barcode interactions in all eight individuals, for each of the large intra-chromosomal interactions called by Long Ranger. For all chromosomes but *TguZ*, the resulting calls (**Fig. 2**, shown in red) indeed matched the inversion types of each individual that were predicted from PCA (**Fig. 2** shown in blue; **Figs. 1,S1** and **Fig. S4A-F**). For the distant barcode-interactions on chromosome *TguZ*, the barcode interactions were found in all individuals regardless of their inversion type (**Figs. 2, S3G-H**). This suggests that the breakpoints on *TguZ* might contain repetitive sequences and that the reference is not correct. In general, the so-detected locations of inversion breakpoints for each chromosome were located flanking to and just outside of the linkage-blocks that had previously been identified among 948 wild caught zebra finches (15) (**Fig. 2; Supplementary Fig. S3**).

#### *Ancestral type reconstruction for inversions*

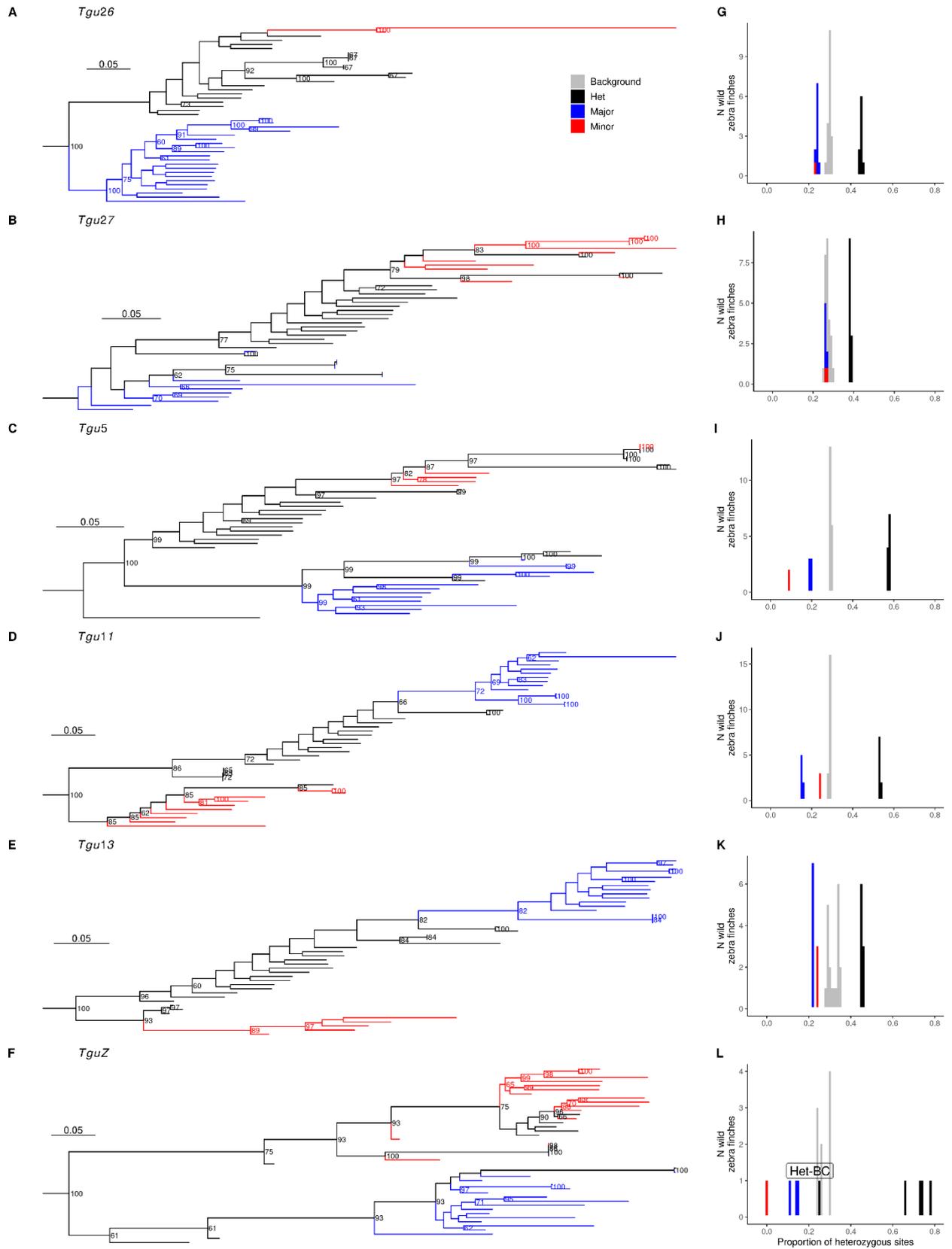
Assuming the simplest scenario where a segregating inversion polymorphism goes back to a single mutation event, we expect the derived inversion haplotype to be largely depleted of genetic variation, because all copies go back to the same founder and because recombination with the genetically diverse ancestral population is largely suppressed (23). Therefore, we expected to find that, within the inverted region the ancestral type would contain more genetic variation (i.e. higher proportion of heterozygous sites in homozygous individuals) and be older (representing ancestral diversity) in phylogenetic analyses of haplotypes than the derived inversion type.

As expected, for each of the six inversion polymorphisms, we found a clear distinction between individuals homozygous for the minor allele and individuals homozygous for the major allele (**Fig. 3**). Specifically, in the phylogenetic trees comprising all 43 sequenced zebra finches (19 wild + 24 mostly captive), we found that individuals homozygous for the minor allele on *Tgu5*, *Tgu26*, *Tgu27* and *TguZ*, and those homozygous for the major allele on *Tgu11* and *Tgu13* clustered as the most derived monophyletic clades (**Fig. 3A-E**). Moreover, these individuals had the least number of heterozygous sites in the putative inverted regions among the 19 wild

zebra finches (autosomal inversions; **Fig. 3G-K**) or among the nine wild male zebra finches (for *TguZ*; **Fig. 3L**). These patterns are characteristic for the derived inversion types. In contrast, individuals homozygous for the alternative allele, i.e. major alleles on *Tgu5*, *Tgu26*, *Tgu27* and *TguZ*, and the minor alleles of *Tgu11* and *Tgu13* were clustering at the basal positions in the trees (**Fig. 3A-E**), and had more heterozygous sites, suggesting that they constitute the ancestral types. Note that the minor allele inversion type on chromosome *TguZ* can be further differentiated into two subtypes, i.e. type B and C in **Fig. S1D**, see also (15). In a phylogenetic analysis with detailed Z inversion types, we found that the two sub-types were further clustering as two separate clades where the homozygous BB were more derived compared to homozygous CC (i.e. the least common haplotype; **Fig. S5**). The two major clades of *TguZ* haplotypes ‘MAJ’ (i.e. A) and ‘MIN’ (i.e. B and C) were both derived (**Fig. 3F**), and all males homozygous for any Z type had (on average) 1-fold lower proportion of heterozygous sites within the inversion as compared to outside the inversion and all other chromosomes (**Fig. 3L**). This suggests that all existing *TguZ* types were relatively derived and that the ancestral genetic diversity has been driven to extinction.

Overall, we found that individuals heterozygous for inversions had more heterozygous sites in the presumable inverted regions, and homozygous individuals had fewer heterozygous sites in this region compared to the (non-inverted) rest of the chromosome (i.e. background) (**Fig. 3G-L**). This is expected due to the reduced recombination and accumulated fixed-differences between the two inversion types (24). This further confirmed that the putative inverted regions that were selected by high SNP-loadings on PC1 (**Fig. S6**) were inside of the inversion. Interestingly, in the two inversions on the microchromosomes *Tgu26* and *Tgu27*, individuals homozygous for either of the two alleles were very similar in their levels of heterozygosity and both were close to the background level, particularly for *Tgu27* (**Fig. 3G-H**).

**Fig. 3.** Phylogenetic trees (**A-F**) and histograms of the proportion of heterozygous sites within the inverted regions (**G-L**) for chromosomes *Tgu26* (**A,G**), *Tgu27* (**B,H**), *Tgu5* (**C,I**), *Tgu11* (**D,J**), *Tgu13* (**E,K**) and *TguZ* (**F,L**) in relation to an individual's inversion type (predicted via PCA in **Fig. 1**, see also **Fig. S1**). Blue depicts individuals that are homozygous for the major allele, red depicts individuals that are homozygous for the minor allele (as defined from 19 wild zebra finches, see **Fig.1**) and black indicates heterozygous individuals. (**A-F**) For each chromosome, the phylogenetic tree was built using all SNPs that are both polymorphic among the wild and the domesticated zebra finches used in this study (N=43). Trees were rooted using the long-tailed finch (not shown). Bootstrap support values are shown if larger than 60. Scale bars show 0.05 substitutions per site. (**G-L**) For each of the 19 wild zebra finches, the proportion of heterozygous sites was calculated for SNPs within the inverted region (red, blue, and black), and outside of the inverted region as background (grey). Here, the inverted regions were inferred from blocks of SNPs with high loadings on PC1 (**Fig. S6**). Note that for all polymorphisms, one haplotype is older and adjacent to the outgroup (blue in **A-C** and red in **D-F**) and has higher genetic variability (blue in **G-I** and red in **J-L**) than the alternative haplotype (red in **A-C,G-I** and blue in **D-F, J-L**). The alleles that form the derived monophyletic groups (red in **A-C** and blue in **D-F** and have the lowest genetic variability (red in **G-I** and blue in **J-L**) are the derived (i.e. inverted) types. Note that both major haplotypes of the sex chromosome *TguZ* are relatively derived. Individuals that are heterozygous for the inversions clustered in between the ancestral and the derived types (black; **A-F**) and had the highest proportion of heterozygous sites (black; **G-L**). Note that one male in **L** was heterozygous for a third haplotype that was closely related to the homozygous major allele. Hence, it showed a low proportion of heterozygous sites (Het-BC in **L**; for detailed phylogenetic relationships between the three inversion types on *TguZ* see **Supplementary Fig S5**).



### *Fitness consequence of inversions on Tgu26 and Tgu27*

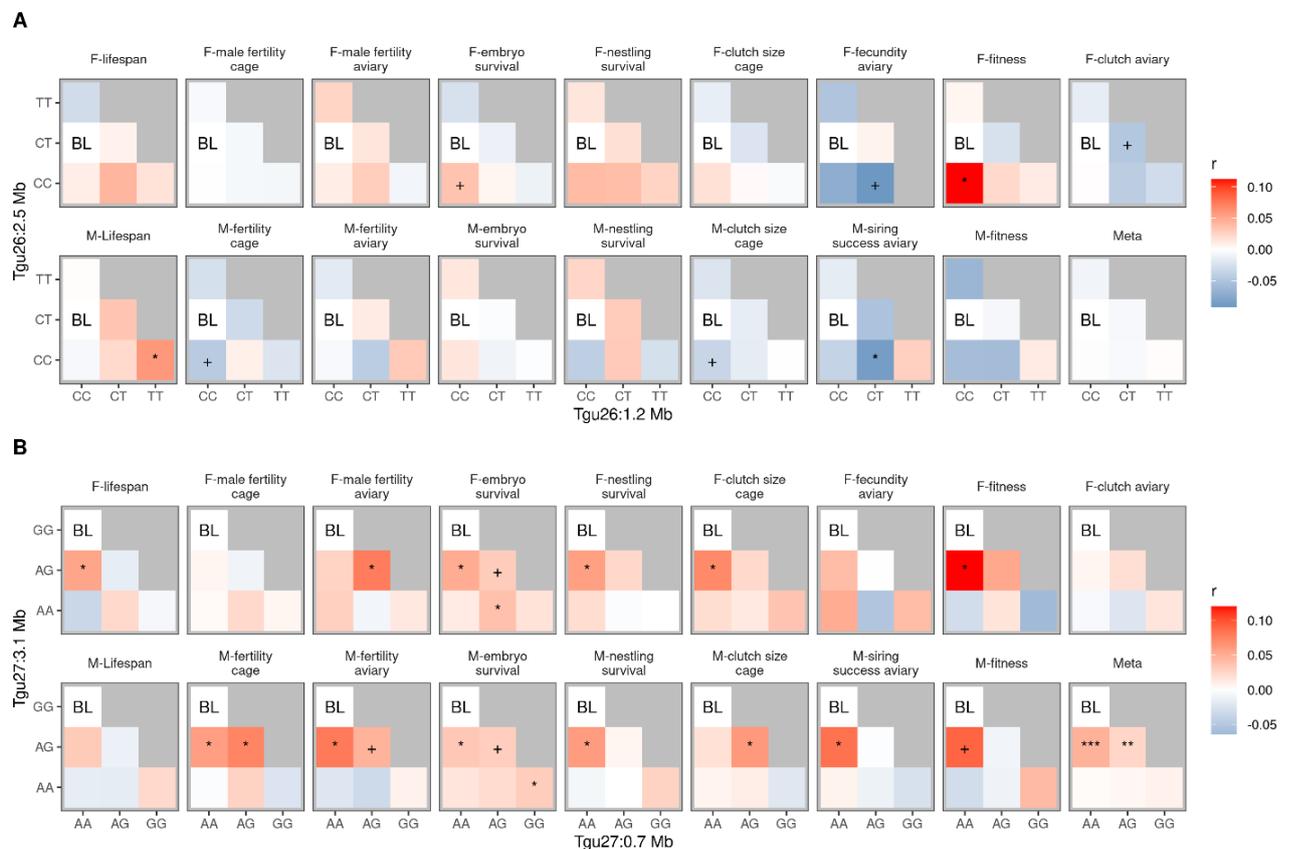
We used two SNP markers to tag the inversion types on each of the two microchromosomes and genotyped about 5,000 captive zebra finches (**Supplementary Figs. S7-S9**) to study the fitness consequences of the two inversions (for details on selection of tag SNPs see **Supplementary Fig. S7, Table S4** and **Methods**). Both markers on *Tgu26* showed no consistent effects on fitness related traits (**Fig. 4A** and **Supplementary Table S5**; mean  $r$  of different combinations of genotypes ranged from -0.009 to 0.002,  $P > 0.2$ ). In contrast, one of the markers on *Tgu27* (SNP at 3,097,302 bp on the old assembly) showed overdominance effects on most of the studied fitness-related traits. Heterozygous females laid more eggs and heterozygous males fertilized more eggs, heterozygous individuals had more offspring that survived till hatching, more hatchlings that survived to independence, heterozygous individuals had higher fitness and lived longer in captive female and male zebra finches (**Fig. 4B** and **Supplementary Table S6**; the heterozygous genotypes ‘AA-AG’ mean  $r = 0.048 \pm 0.0006SE$ ,  $P < 0.0001$ ; ‘AG-AG’ mean  $r = 0.026 \pm 0.006SE$ ,  $P < 0.0001$ ).

## **Discussion**

### *Characteristics of the six chromosomal inversions in zebra finch*

Traditionally, polymorphic large inversions were identified either via cytogenetic observations, e.g. *Tgu5* and *TguZ* in the zebra finch (17, 25), or from their large phenotypic effects (1), e.g. (2–4, 8). Our study shows that PCA on SNPs from whole genome sequencing (WGS) data effectively identified the segregating inversion polymorphisms on microchromosomes smaller than 7 Mb and with a minor allele frequency as low as 29% among 19 wild birds (**Fig. 1**). Using information on shared barcodes of linked-reads, we successfully identified the breakpoint regions for the two microchromosomal inversions on *Tgu26* and *Tgu27*, and the previously identified inversions on *Tgu5*, *Tgu11* and *Tgu13* (15, 17, 25), with an average resolution of 8 Kb on the new zebra finch assembly bTaeGut1.202104 (20) (**Fig. 2; Supplementary Table S3**). However, chromosome *TguZ* (15, 17, 25) was dropped from our analysis of inversion breakpoints, because the barcode coverage at the putative breakpoints was extremely high and present in all samples regardless of their inversion type (**Fig. 2; Supplementary Fig S4G-H**), indicating that the true breakpoints on *TguZ* contain highly repetitive elements.

**Fig. 4.** Estimated standardized effect sizes of genotypes of the tag SNPs on *Tgu26* (A) and *Tgu27* (B) relative to the most common genotype (i.e. baseline, ‘BL’) on female and male fitness traits as well as the meta-summarized effect sizes across all traits. Note that on each chromosome there are two tag SNPs which allow for 9 combinations of genotypes, only 6 of which were observed to occur (due to linkage; missing genotypes in grey). Red indicates positive effects of the genotype on fitness-related traits whereas blue indicates negative effects. For estimated effects of each genotype, ‘+’ indicates  $P < 0.1$ , ‘\*’ indicates  $P < 0.05$ , ‘\*\*\*’ indicates  $P < 0.001$ , ‘\*\*\*\*’ indicates  $P < 0.0001$ , otherwise not significant. Note there is no significant overall effect of genotypes on fitness on *Tgu26* (A). In contrast, there are significant heterotic effects of the two heterozygote genotypes (AA-AG and AG-AG) on *Tgu27* (B). The SNP-marker that shows heterosis (*Tgu27*: at 3.1 Mb based on the old assembly) is in strong LD with the inversion types on *Tgu27* ( $r^2 > 0.8$  both in the 19 wild and the 24 domesticated zebra finches; **Supplementary Fig. S7I,L** and **Table S4**; also see **Fig. S10** for sample size).



Our results show that the inversions on microchromosomes *Tgu26* and *Tgu27* both started close to one end of the chromosome (i.e. 0.5 Kb and 0.02 Mb away from one end in the new assembly, respectively) and spanned approximately half of the microchromosome (i.e. 3.4 Mb out of 6.8 Mb on *Tgu26* and 3.3 Mb out of the 5.8 Mb on *Tgu27* (20); **Supplementary Table S3**). In addition to the known pericentric inversions on chromosomes *Tgu5* and *TguZ* (17, 25, 26), we found that the inversion on the acrocentric microchromosome *Tgu26* also shifts the position of the centromere (26), also see **Supplementary Fig. S2**. As the position of the centromere on *Tgu27* is still unclear (26), the inversion on *Tgu27* may potentially also shift the position of the centromere. Future cytogenetic studies on individuals with different variants of *Tgu26* and *Tgu27* may shed more light on the centromere organization in these two microchromosomes.

Using phylogenetic analysis on individual-based haplotypes, we successfully clarified the ancestral states of the segregating inversion polymorphisms on the six chromosomes in the zebra finch (**Fig. 3, Supplementary Fig. S5 and Table S2**). Interestingly, within the inversions on the three smallest (<20 Mb) chromosomes *Tgu13*, *Tgu26* and *Tgu27*, particularly *Tgu27*, both the ancestral and the derived types contained high genetic variability, similar to the level of variability in the background outside the inversion (**Fig. 3G,H,K, also see Fig. S5**). Additionally, close to the centre of these three inversions, we found a decay in SNP loadings that separate individuals along PC1 (**Fig. S6A,B,E**). Closer to the breakpoints, all three inversions were highly diverged from their ancestral types as indicated by the high proportion of heterozygous sites in individuals that were heterozygous for the inversion (**Fig. 3G,H,K**). These reduced loadings in the centre could be explained by double cross-overs leading to recombination between the inversion types, or higher levels of gene conversion, presumably due to the high recombination rates in microchromosomes (19, 27). We also confirmed that all existing *TguZ* haplotypes were derived (**Supplementary Fig S5**) and potentially contained multiple secondary small inversions (**Fig. 2 and Supplementary Fig. S3**; also see (15)), implying a fast evolution of the zebra finch sex-chromosome.

Intriguingly, a recent study on avian ancestral chromosome structure found that avian microchromosomes 26-28 had a three-fold higher density of evolutionary breakpoint regions (e.g. break points of structural variations) than the genome-wide average during the ancestral avian chromosomal evolution (Damas *et al.* 2018 (18)). Moreover, the peculiar avian

microchromosome 27 was found to have the highest gene density in the reconstructed ancestral type (18).

Note that in this study we focused on large inversion polymorphisms (e.g. at least 2 Mb) that are high in minor allele frequency. Among the six identified large inversion polymorphisms, five (i.e. *Tgu5*, *Tgu11*, *Tgu13*, *Tgu27* and *TguZ*) appear to be maintained close to 50% allele frequency, while on *Tgu26* the derived inversion segregates at an allele frequency of about 27%-29% (**Fig. 1**, **Supplementary Fig S1** and **Tables S2, S4**).

#### *Heterosis – the simplest mechanism that maintains inversion polymorphisms*

Chromosomal inversions have long been appreciated as an important source of genetic diversity, local adaptation and speciation (1, 28). However, the mechanisms that maintain inversion polymorphisms remain largely unknown (1). The few cases of large inversions that have been well studied are often linked to large phenotypic effects. For instance, in the white throated sparrow *Zonotrichia albicollis*, the large (>100 Mb) inversion on chromosome 2 (2, 29), and in the ruff *Philomachus pugnax*, the inversion (>4.5 Mb) on chromosome 11 (3, 4) were identified by linking the genomic regions that underpin phenotypically distinct morphs, a suite of behavioural traits, and ultimately alternative breeding strategies. In these cases, the behavioural strategies are thought to be maintained via negative frequency-dependent selection and some reduction in fitness in individual that are homozygous for the derived type.

The maintenance of other polymorphisms, that do not involve interacting behavioural strategies, may most easily be explained by heterozygote advantage, where heterozygous individuals have increased reproductive performance (30), e.g. zebra finch *TguZ* (5, 6) and *Tgu27* (**Fig. 4B**). Similarly, antagonistic pleiotropic effects where alleles that show opposing effects on different life-history traits (9), e.g. zebra finch *Tgu11* (Pei et al. In preparation), may also often sum up to a net heterosis (Pei et al. In preparation). Finally, inversion polymorphisms may mediate adaptation to certain local conditions (11, 12, 31, 32). In the zebra finch, at least three out of the six inversion polymorphisms appear to be stably maintained close to 50% allele frequency (**Figs. 1B,D,F** and **Supplementary Figs. S1B,D,F**) by heterotic effects involving various fitness-related traits (**Fig. 4B**, Pei et al. In preparation and (5, 6)). However, the observed effect sizes on fitness-related traits were remarkably small (mean  $r = 0.05$  in **Fig. 4B**, Pei et al. In preparation and (5, 6)), suggesting that weak heterosis on multiple fitness-related traits could

be the most common mechanism that maintains many inversion polymorphisms and may be hard to measure.

## Materials and Methods

### *Whole genome sequencing data and reference genomes*

We used published whole-genome sequencing data on 43 zebra finches (both the Australian subspecies *Taeniopygia guttata castanotis* and the Timor subspecies *T. g. guttata*) to scan the genome for large inversion polymorphism. This included 19 wild and five domesticated *castanotis* (MP1-5) zebra finches from North America (Singhal *et al.* (2015) (22)), four captive *castanotis* zebra finches from North America (bTaeGut1-4) (Rhie *et al.* (2021) (20)), 13 captive *castanotis* zebra finches from five different captive populations in Europe, one wild caught individual (taeCasSA) and one captive *castanotis* x *guttata* hybrid (taeGutGut; also see (33, 34)) (Kinsella *et al.* (2019) (35) and Pei *et al.* (2021) (34)). Among these zebra finches, eight individuals were sequenced using linked-read technology (bTaeGut1-4 (20) and SR00100, taeCasSA, taeCasSE and taeCutGut (34, 35)) and the rest was sequenced using conventional Illumina WGS sequencing either from PCR-free or regular gDNA libraries (**Supplementary Table S1; Fig. 1**). Note that the hybrid individual (taeGutGut) had 95% of its genome coming from *T. g. guttata*. We included this hybrid individual to maximize our sample size in the linked-read dataset. For phylogenetic analyses of the inversions, we included published WGS data on 20 long-tailed finches as the outgroup (Singhal *et al.* (2015)). Details on each sample and sequencing technologies used are in **Supplementary Table S1**.

For SNP-based analyses, we used the taeGut1 (21) reference genome. For linked-read analyses, we used the most recent zebra finch assembly bTaeGut1.pri.cur.20210409 (aka bTaeGut1.2021, taken from [https://vgp.github.io/genomeark/Taeniopygia\\_guttata/](https://vgp.github.io/genomeark/Taeniopygia_guttata/)). Both assemblies were constructed using DNA from the same male. For simplicity, we refer to the taeGut1 (21) assembly as ‘old’ and the bTaeGut1.pri.cur.20210409 (20) as ‘new’.

### *Raw reads processing, SNP calling and SNP selection*

We obtained the filtered variant call files for the 19 wild and five domesticated zebra finches and the 20 long-tailed finches (22) from the European Nucleotide Archive (accession ID PRJEB10586). We used the raw variant call file for the 14 *castanotis* zebra finches and one

*castanotis* x *guttata* hybrid from (34). The above variants were all called against the old reference genome taeGut1 (21).

For the four *castanotis* individuals from (20), there was no variant call data available. Therefore, we downloaded the raw linked-read (10X) sequencing data from [https://vgp.github.io/genomeark/Taeniopygia\\_guttata/](https://vgp.github.io/genomeark/Taeniopygia_guttata/). We applied a fast raw-read processing and SNP calling pipeline adapted from (34), where we first mapped the raw reads against taeGut1 (21) using BWA-MEM v0.7.17 (36) with default settings and marked shorter split hits as secondary. Then we down-sampled each individual to 10 fold coverage and used the ‘Picard (37) MarkDuplicates’ option to mark mapped reads that might have resulted from PCR duplication. Lastly, we called SNPs for the four individuals simultaneously using SAMtools v1.6 mpileup (38) and bcftools v1.9 call (39).

For principal component and phylogenetic analyses, we filtered for high quality (quality score >100) bi-allelic SNPs that were polymorphic both among the 19 wild caught zebra finches (22) and among the additional 24 zebra finches with different genetic backgrounds (20, 22, 34, 35), using a customized R script (see **Code And Data Accessibility**).

#### *Principal component analysis (PCA)*

We performed PCA on the number of alternative alleles with the filtered SNPs in R V4.1.0, using the ‘prcomp’ function from the ‘stats’ package. Specifically, we screened each chromosome for large chromosomal structural variations (inversions) using PCA on all filtered SNPs on that chromosome (15, 40). In a first step, we used the 19 wild zebra finches (Singhal et al (2015) (22)) to detect segregating inversion polymorphisms in the wild. Particularly, we screened for chromosomes where the 19 wild zebra finches formed three clusters along PC1. We expected that individuals that are homozygous for the two inversion types would cluster at the two distal ends along PC1 while heterozygous individuals were intermediate. In a second step, we used the SNP loadings from the above PCA to predict the PC scores of the remaining 24 individuals to confirm the inversion polymorphisms and to infer their inversion types for downstream analyses.

#### *Verification of inversion polymorphisms using 10X linked-read sequencing data*

In order to verify and characterize the inversions, we mapped raw reads together with their barcodes from eight zebra finches (20, 34, 35) against the new zebra finch assembly (20) using the “wgs” function in Long Ranger. For each individual, we down-sampled its library to 35 Gb to speed up SV detection and to standardize library sizes.

The reference individual was homozygous for most inversions except for *Tgu13* and *TguZ* (**Fig. S1**). This is because in the new reference assembly bTaeGut1.2021, SNPs are phased within contigs, but not between contigs. Hence, the haploid genome of the reference individual clustered in the heterozygous group on *Tgu13* and *TguZ* (**Fig. S1C,D**). But this will not influence the identification of inversion breakpoints as long as the breakpoints are assembled within contigs (**Fig. 2** and **Supplementary Fig. S3, S4**). The reference was called as heterozygous of types B and C (i.e. the minor allele in **Fig. 1F**) for the inversion on *TguZ* (**Fig. S1D**), therefore we focused on identifying the breakpoints that differentiate haplotype A (the major allele) and B plus C. For the inversion type of the *Tgu13* reference we guessed the inversion types post-hoc (i.e. based on the calls of the eight 10X samples).

We expected to see long-range barcode interactions from the two inversion breakpoints in individuals heterozygous or homozygous for the alternative types as the reference individual. We expected to see no barcode interactions in individuals that have the same inversion genotype as the reference individual. We filtered the Long Ranger output to include only intra-chromosomal barcode interactions that (a) are more than two Mb apart, (b) had a minimum quality score of four, (c) were at least 10 Kb away from an ambiguous region (i.e. multi-Ns in the reference) and (d) were called from more than one individual to minimize false positives (but see **Supplementary Fig. S3** for the distribution of candidates based on relaxed criteria). Long Ranger called 24 candidates from the 8 samples on chromosome *TguZ*. Because *TguZ* is known to have a large number of inversions (15), we only focused on the call from the most distant barcode interactions on *TguZ*.

We checked the filtered calls of intra-chromosomal barcode interactions using the 10x Genomics Loupe genome browser v2.1.1. To minimize false positive and false negative rates from the above filtered large SVs that were called by ‘longranger’, a person manually scored each individual for the presence or absence of barcode interactions while being blind for individual inversion types that were inferred from PCA.

#### *Phylogenetic and genetic variability analyses of the inversions*

We studied the evolutionary history of the six chromosomal inversions in the zebra finch using the filtered high-quality SNPs from the 19 wild and the 24 additional zebra finches with different genetic backgrounds (**Supplementary Table S1**). For each individual, its inversion types were inferred using PCA (**Fig. 1** and **Supplementary Fig. S1**). We identified regions that were inside the inversions in the old assembly (**Fig. S6**) based on SNP loadings (21). For chromosomes *Tgu26* and *Tgu27*, we additionally used the alignment between the old and the new assembly (with the signals of barcode interactions based on the new assembly bTaeGut1.2021 (20)) to identify regions in the old assembly that are located inside the inversion (**Fig. S2**).

For the phylogenetic analyses we first created one haplotype sequence for each individual using all the above filtered variable sites within the inversions on each chromosome. Heterozygous sites were recoded following the IUPAC code. We built a phylogenetic tree for each chromosome using RAxML-NG v1.0.2 (41) assuming a general time-reversible model and a discrete GAMMA model of rate heterogeneity with 100 randomized parsimony starting trees and 1000 bootstrap replicates. To root the tree, we used the same genomic regions from 20 long-tailed finches. We assumed that individuals homozygous for the ancestral types would cluster closer to the outgroup, whereas individuals homozygous for the alternative types would form one derived monophyletic group. Heterozygous individuals would be positioned between the two homozygous types.

We analyzed the genetic variability within inversion genotypes (i.e. homozygous for the major allele, homozygous minor allele or heterozygous) to infer the ancestral and derived inversion types. For this we used only the 19 wild zebra finches because some of the domesticated birds were closely related (e.g. individuals bTaeGut2-4 (20), individuals MP1-5 (22) were family trios and quintets) or inbred. We calculated the proportion of heterozygous sites within the putative inversions for each individual (**Fig S6**).

#### *Fecundity, fertility and viability measurements*

We studied the effects of inversion genotypes on fecundity, fertility and viability data gathered on three captive populations of zebra finches maintained at the Max Planck Institute for Ornithology, Seewiesen, Germany. Among them, populations (i) ‘Seewiesen’ (population #18

in (42)) and (ii) ‘Bielefeld’ (population #19 in (42)) were genetically independent, whereas population (iii) ‘Krakow’ was established by cross-breeding individuals from the Krakow (#11 in (42)) and from ‘Seewiesen’ populations. For details see (43).

Data on reproductive performance and lifespan were taken from <https://osf.io/tgsz8/>. Overall, reproductive performance of birds was measured in (1) cages with a single pair of one male and one female, and (2) aviaries where a group of males and females could form breeding pairs freely. For detailed information on experiment setups, rearing conditions and measurements of traits see (43). In brief, for female fecundity, we measured the number of eggs that was laid per clutch in cages, in aviaries, and in aviaries where offspring rearing was not allowed and eggs were replaced by plastic eggs. For male fertility, we measured if an egg laid in a cage was fertilized by the male or not, if an egg laid in an aviary was fertilized by the social partner of the female or not and the total number of eggs that were fertilized by a male, including both within- and extra-pair eggs. For offspring survival, we measured for the female and male whether the embryo hatched or not, as well as if the social chick reared by them survived until independence (i.e. 35 days of age) or not. For fitness, we measured the total number of independent young produced by the genetic mother and the total number of independent young sired by the genetic father. Lifespan of an individual was measured as the duration from the date of hatching until death.

#### *Tag SNPs selection and genotyping*

Tag SNPs were selected before the generation of sequencing data (see also *Whole genome sequencing data*). Knief et al (2016) (15) genotyped 948 wild Australian zebra finches, 88 founders of the Seewiesen population, 74 founders of the Bielefeld population and 63 founders of the Krakow population (i.e. 25 from Krakow and 38 from Seewiesen population) using an Illumina Infinium iSelect HD Custom BeadChip. Among the 4553 genotyped SNPs, 34 and 36 were evenly spaced along *Tgu26* and *Tgu27*, respectively (**Supplementary Fig. S7A,G**) (15). Knief et al (2016) (15) detected weak signals of linkage disequilibrium on *Tgu26* and *Tgu27* among the 948 wild zebra finches through principal component analysis with maximal  $r^2$  values of 0.244 and 0.553 (**Fig. 2**), respectively. They concluded that there might have been polymorphic inversions on these two chromosomes that were either not tagged perfectly by the SNPs or that had exchanged material between inversion types through gene conversion (44) or

double crossovers (45, 46). For each of these two microchromosomes, two tag SNPs were selected based on their loadings on PC1 in the 948 wild zebra finches. These were located at chr26:1,249,190 bp, chr26:2,522,498 bp and at chr27:710,554 bp, chr27:3,097,302 bp (SNP positions in reference to the old assembly; **Supplementary Table S4**). The correlations between the genotypes of these SNPs with the inversion types among the 19 wild and the 24 additional zebra finches were high on *Tgu27* (Pearson's correlation coefficients  $r^2 = 0.81-0.84$ ;  $P < 0.0001$ ; **Supplementary Table S4** and **Fig. S7G-L**). On *Tgu26*, the tag SNP at 1,249,190 bp was in high LD (Pearson's correlation coefficients  $r^2 = 0.62-1$ ;  $P \leq 0.001$ ) whereas the tag SNP at 2,522,498 bp was in low LD with the inversion type (Pearson's correlation coefficients  $r^2 = 0.18-0.3$ ;  $P \geq 0.2$ ; **Supplementary Table S4** and **Fig. S7A-F**). For 4958 captive zebra finches from three populations (**Supplementary Figs. S8,S9**), the four tag SNPs were genotyped using the Sequenom MassARRAY iPLEX platform (47) at the Institute of Clinical Molecular Biology at Kiel University. Genotypes were called by the default parameters on the MassARRAY iPLEX platform (see (48) for details). The founders of the Seewiesen, Bielefeld and Krakow populations were genotyped on both platforms. We kept only those genotype calls that were consistent between the two platforms or that were called on the Sequenom platform only.

### *Statistical analyses*

We used mixed-effect models to estimate the effects *Tgu26* and *Tgu27* genotypes on all fitness-related traits. All models were fitted using the function 'lmer' in the R package 'lme4' (49). Links to detailed model structures, all scripts and data can be found in the **Code and data accessibility** statement (also see **Supplementary Tables S5,S6**). In order to compare and summarize the effect sizes of the genotypes on different reproductive performance and viability traits, we treated all response variables as normally distributed and Z-scaled them. We estimated the genotypic effects of *Tgu26* and *Tgu27* as continuous variables in separate models. Specifically, for each chromosome, instead of fitting individual genotype as a factor variable with multiple levels, we treated the most common genotype combination in the Seewiesen population as baseline (i.e. not estimating that effect; **Fig. 4; Figs. S8,S9**). Then each of the five other genotype combinations (with more than 10 observations) was treated as one continuous variable to reflect if the individual had that genotype or not (**Fig. 4; Figs. S8,S9**). Then the focal genotypes of each chromosome were Z-scored and fitted simultaneously in the models to estimate the effect of the genotype. We controlled for inbreeding (i.e. pedigree-based

inbreeding coefficient) and age effects by fitting them as additional covariates whenever possible. Additionally, we controlled for confounding fixed effects (factor) including (i) the laying order, (ii) hatching order and (iii) laying order of the clutch within the breeding season, (iv) if the clutch was laid while nestlings from a previous clutch were still in the nest (two levels) and (v) the population where the egg belonged to (three levels: Seewiesen, Bielefeld, Krakow), whenever applicable. We included the identities of the female, the male, the pair, the clutch and the breeding season as random intercepts, whenever applicable to control for pseudo-replication.

To meta-summarize the genotype effect on all fitness-related traits, we used two weighted linear models for *Tgu26* and *Tgu27* separately using the `lm` function in the R package `stats`. We treated the above estimated standardized effect sizes of the genotypes as response variable, and we fitted the genotype as a fixed factor variable with five levels (because the most common genotype was the baseline and was not estimated in the model). To account for the uncertainty of the estimated effect sizes, we included the multiplicative inverse of the standard error of each estimate as the weight. We removed the intercept to estimate the mean effect sizes for each genotype (i.e. pairwise combination of genotypes of the two tag SNPs).

### **Code and data accessibility**

Fitness-related data can be found at the Open Science Framework at <https://osf.io/tgsz8/>. Supporting data and R scripts will be uploaded to the Open Science Framework or available upon request.

### **Competing interests**

We have no competing interests.

### **Funding**

This research was supported by the Max Planck Society (to BK). YP was part of the International Max Planck Research School for Organismal Biology.

### **Acknowledgements**

We thank M. Schneider for molecular work, and K. Martin, S. Bauer, E. Bodendorfer, J. Didsbury, A. Grötsch, A. Kortner, P. Neubauer, F. Weigel and B. Wörle for animal care and help with breeding.

## References

1. M. Wellenreuther, L. Bernatchez, Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends Ecol. Evol.* **33**, 427–440 (2018).
2. E. M. Tuttle, *et al.*, Divergence and functional degradation of a sex chromosome-like supergene. *Curr. Biol.* **26**, 344–350 (2016).
3. S. Lamichhaney, *et al.*, Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nat. Genet.* **48**, 84–88 (2015).
4. C. Küpper, *et al.*, A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.* **48**, 79–83 (2015).
5. K.-W. Kim, *et al.*, A sex-linked supergene controls sperm morphology and swimming speed in a songbird. *Nat. Ecol. Evol.* **1**, 1168–1176 (2017).
6. U. Knief, *et al.*, A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nat. Ecol. Evol.* **1**, 1177–1184 (2017).
7. Y. Le Poul, *et al.*, Evolution of dominance mechanisms at a butterfly mimicry supergene. *Nat. Commun.* **5**, 1–8 (2014).
8. P. Jay, *et al.*, Supergene Evolution Triggered by the Introgression of a Chromosomal Inversion. *Curr. Biol.* **28**, 1839–1845.e3 (2018).
9. C. Mérot, V. Llaurens, E. Normandeau, L. Bernatchez, M. Wellenreuther, Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. *Nat. Commun.* **11**, 670 (2020).
10. R. Pracana, *et al.*, Fire ant social chromosomes: Differences in number, sequence and expression of odorant binding proteins. *Evol. Lett.* (2017) <https://doi.org/10.1002/evl3.22>.
11. D. B. Lowry, J. H. Willis, A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **8** (2010).
12. Y. W. Lee, L. Fishman, J. K. Kelly, J. H. Willis, A segregating inversion generates fitness variation in yellow monkeyflower (*Mimulus guttatus*). *Genetics* **202**, 1473–1484 (2016).
13. V. Fillon, The chicken as a model to study microchromosomes in birds: A review. *Genet. Sel. Evol.* **30**, 209–219 (1998).
14. M. I. Pigozzi, Relationship between physical and genetic distances along the zebra finch Z chromosome. *Chromosom. Res.* **16**, 839–849 (2008).
15. U. Knief, *et al.*, Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biol.* **17**, 199 (2016).
16. D. W. Burt, Origin and evolution of avian microchromosomes. *Cytogenet. Genome Res.* **96**, 97–112 (2002).
17. Y. Itoh, A. P. Arnold, Chromosomal polymorphism and comparative painting analysis in the zebra finch in *Chromosome Research*, (2005), pp. 47–56.
18. J. Damas, J. Kim, M. Farré, D. K. Griffin, D. M. Larkin, Reconstruction of avian ancestral karyotypes reveals differences in the evolutionary history of macro- and microchromosomes. *Genome Biol.* **19**, 1–16 (2018).
19. H. Ellegren, The evolutionary genomics of birds. *Annu. Rev. Ecol. Evol. Syst.* **44**, 239–259 (2013).
20. A. Rhie, *et al.*, Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **592**, 737–746 (2021).
21. W. C. Warren, *et al.*, The genome of a songbird. *Nature* **464**, 757–762 (2010).

22. S. Singhal, *et al.*, Stable recombination hotspots in birds. *Science* (80-. ). **350**, 928–932 (2015).
23. A. A. Hoffmann, C. M. Sgrò, A. R. Weeks, Chromosomal inversion polymorphisms and adaptation. *Trends Ecol. Evol.* **19**, 482–488 (2004).
24. M. Kirkpatrick, How and why chromosome inversions evolve. *PLoS Biol.* **8** (2010).
25. Y. Itoh, K. Kampf, C. N. Balakrishnan, A. P. Arnold, Karyotypic polymorphism of the zebra finch Z chromosome. *Chromosoma* **120**, 255–264 (2011).
26. U. Knief, W. Forstmeier, Mapping centromeres of microchromosomes in the zebra finch (*Taeniopygia guttata*) using half-tetrad analysis. *Chromosoma* **125**, 757–768 (2016).
27. N. Backström, *et al.*, The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Res.* **20**, 485–95 (2010).
28. A. A. Hoffmann, L. H. Rieseberg, Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? *Annu. Rev. Ecol. Evol. Syst.* (2008) <https://doi.org/10.1146/annurev.ecolsys.39.110707.173532>.
29. E. M. Tuttle, Alternative reproductive strategies in the white-throated sparrow: Behavioral and genetic evidence. *Behav. Ecol.* **14**, 425–432 (2003).
30. T. Dobzhansky, O. Pavlovsky, An extreme case of heterosis in a central American population of *Drosophila tropicalis*. *Proc. Natl. Acad. Sci.* **41**, 289–295 (1955).
31. E. Durmaz, C. Benson, M. Kapun, P. Schmidt, T. Flatt, An inversion supergene in *Drosophila underpins* latitudinal clines in survival traits. *J. Evol. Biol.* **31**, 1354–1364 (2018).
32. R. Faria, *et al.*, Multiple chromosomal rearrangements in a hybrid zone between *Littorina saxatilis* ecotypes. *Mol. Ecol.* **28**, 1375–1393 (2019).
33. U. Knief, W. Forstmeier, Y. Pei, J. Wolf, B. Kempnaers, A test for meiotic drive in hybrids between Australian and Timor zebra finches. *Ecol. Evol.*, ece3.6951 (2020).
34. Y. Pei, *et al.*, Occasional paternal inheritance of the germline restricted chromosome in songbirds. *bioRxiv* (2021) <https://doi.org/https://doi.org/10.1101/2021.01.28.428604>.
35. C. M. Kinsella, *et al.*, Programmed DNA elimination of germline development genes in songbirds. *Nat. Commun.* **10**, 5468 (2019).
36. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv Prepr. arXiv* (2013).
37. Broad Institute, Picard Tools. *Broad Institute, GitHub Repos.* (2019).
38. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (2009) <https://doi.org/10.1093/bioinformatics/btp352>.
39. 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).
40. J. Ma, C. I. Amos, Investigation of Inversion Polymorphisms in the Human Genome Using Principal Components Analysis. *PLoS One* **7**, e40224 (2012).
41. 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).
42. 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).
43. Y. Pei (裴一凡), *et al.*, Proximate Causes of Infertility and Embryo Mortality in Captive Zebra Finches. *Am. Nat.* **196**, 577–596 (2020).

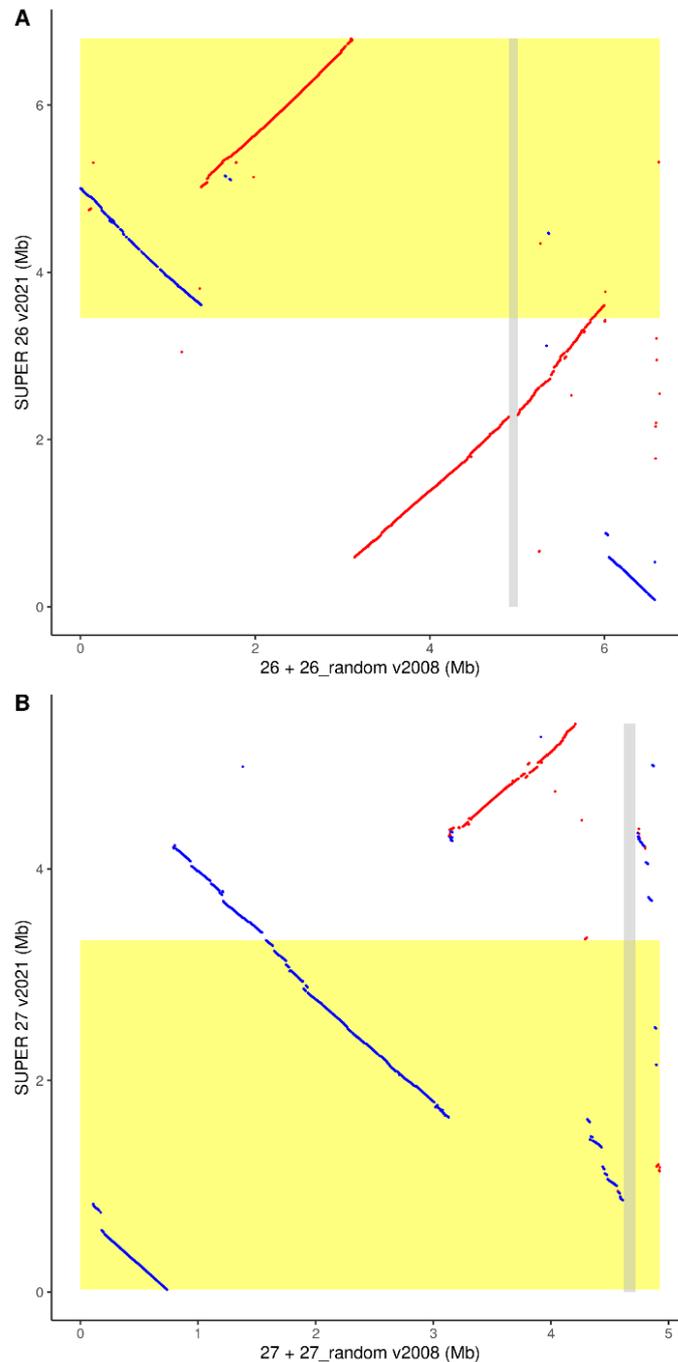
44. K. L. Korunes, M. A. F. Noor, Pervasive Gene Conversion in Chromosomal Inversion Heterozygotes. *Mol. Ecol.* (2018) <https://doi.org/10.1111/mec.14921> (December 1, 2018).
45. L. S. Stevison, K. B. Hoehn, M. A. F. Noor, Effects of inversions on within- and between-species recombination and divergence. *Genome Biol. Evol.* **3**, 830–841 (2011).
46. K. Ishii, B. Charlesworth, Associations between allozyme loci and gene arrangements due to hitch-hiking effects of new inversions. *Genet. Res., Camb*, 30 (2018).
47. U. Knief, *et al.*, Quantifying realized inbreeding in wild and captive animal populations. *Heredity (Edinb)*. **114**, 397–403 (2015).
48. U. Knief, *et al.*, Association mapping of morphological traits in wild and captive zebra finches: reliable within, but not between populations. *Mol. Ecol.* **26**, 1285–1305 (2017).
49. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).

### Supplementary Materials

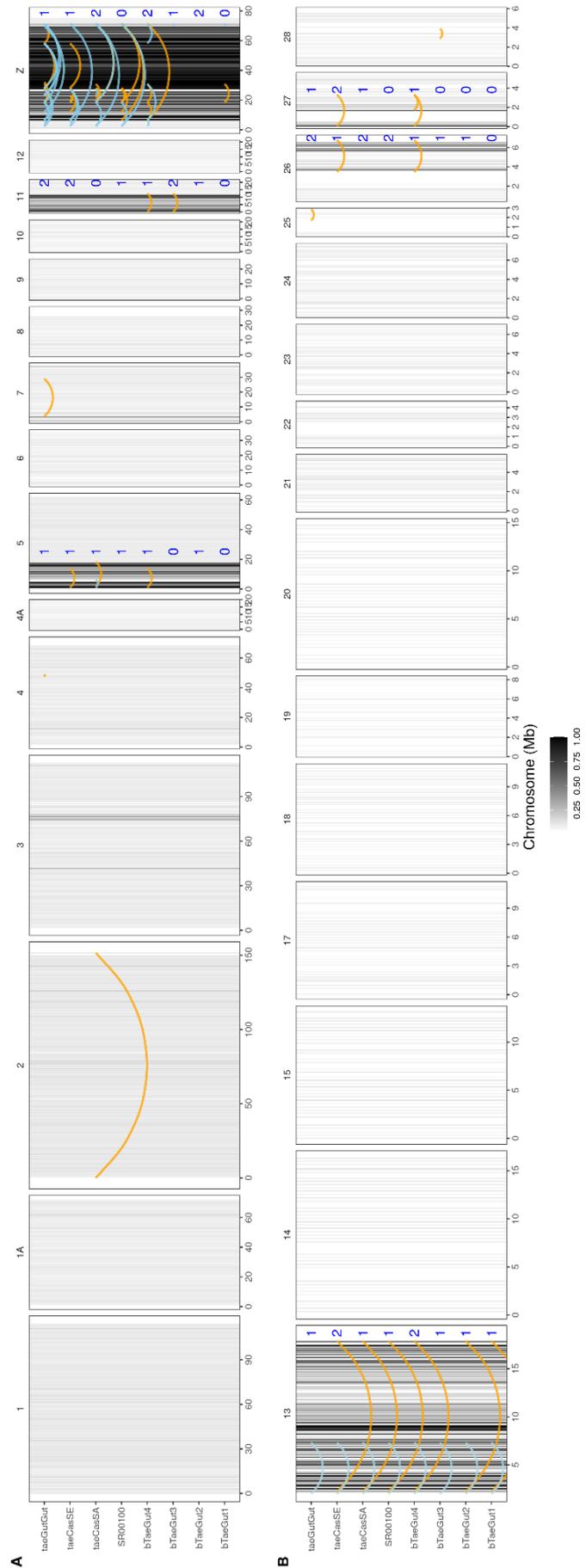
**Fig. S1.** Principal component analysis based on the 4553 randomly selected SNPs (genome-wide) genotypes from 948 wild zebra finches (small grey dots) used in Knief et al. 2016 (1) for chromosomes that contain inversion polymorphisms, i.e. *Tgu5*, *Tgu11*, *Tgu13*, *TguZ*, *Tgu26* and *Tgu27*. Dots represent wild-caught individuals whereas triangles are captive individuals (**Methods**). The PC values of the 43 zebra finches used in this study (red, blue and black) and the new reference genome bTaeGut1.202104 (2) (orange) were predicted by the loadings of the 4553 SNPs from the 948 wild birds. Red and blue depict individuals that are homozygous for the minor or major alleles, respectively (**Fig. 1**), whereas black indicates heterozygous individuals (also see **Fig. 3**). Grey letters A, B and C indicate the inversion types of the major, minor and the least abundant alleles defined in Knief et al. 2016 (1). Individuals used for linked-read analysis are highlighted by boxes around their IDs (**Figs. 2, S3**). For a detailed description of haplotypes defined in this study and in Knief et al. 2016 (1) see **Supplementary Table S2**.



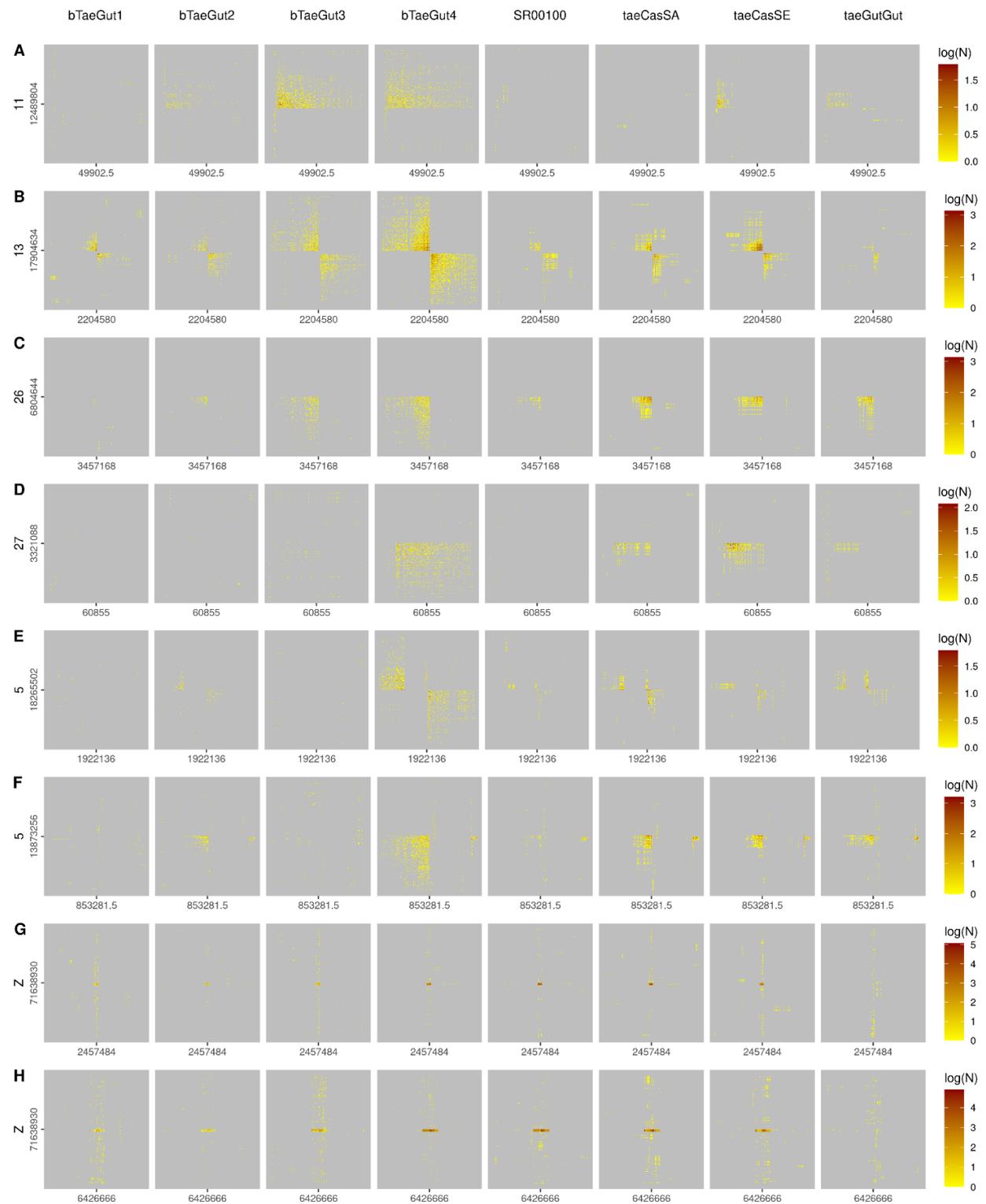
**Fig S2.** Dot-plots show the alignments of the old *taeGut1* (5) (x-axis) and the new *bTaeGut1.202104* (2) (y-axis) assemblies for chromosomes *Tgu26* (A) and *Tgu27* (B). Red indicates alignment at the forward strands whereas blue indicates the alignment of the reverse strand. Yellow background indicates the inverted regions identified by shared barcodes of linked-reads on the new assembly (2) (y-axis; see **Fig. 2, Supplementary Figs. S3,S4** and **Table S3**). Grey bars separate the major micro-chromosomal assembly and the chromosome random for *Tgu26* and *Tgu27* on the old assembly (5).



**Fig. S3.** Genome-wide distribution of unexpected intra-chromosomal barcode interactions that were detected by Long Ranger (orange and light blue curves). All individuals (rows) are *Taeniopygia guttata castanotis*, except for individual taeGutGut which is a hybrid between the two subspecies *T. g. castanotis* and *T. g. guttata* (see **Methods**; also see (3, 4)). Chromosomal positions are based on the new genome assembly bTaeGut1.202104 (2). Orange depicts all high-quality candidate calls of intra-chromosomal barcode interactions that are more than 1 Mb apart (also see **Fig. 2** and **Methods** in the main text). Light blue depicts candidate calls that are close to gaps in the reference. Blue numbers indicate the expected number of alternative inversion types based on PCA of SNPs (**Supplementary Fig. S1**). Each vertical bar is a SNP that was used in Knief *et al.* (2016) (1). SNP positions were lifted from the old TaeGut1 (5) to the new assembly bTaeGut1.202104 (2). The blackness of each bar indicates the highest level of linkage disequilibrium (LD) of this SNP with another SNP that is at least 1 Mb apart within the same chromosome, among 948 wild zebra finches (1). The large blocks of SNPs on chromosomes *Tgu5*, *I1*, *I3*, *Z*, that are in high LD (black bars) indicate the inversions, whereas the microchromosomes *Tgu26* and *Tgu27* contained weak signals of LD, which was suggestive of inversion (Knief *et al.* (2016) (1). Note that the distant barcode interactions detected by Long Ranger were enriched for chromosomes that contained inversions (i.e. N=40 orange curves on chromosomes that show a signal of segregating inversions, i.e. *Tgu5*, *I1*, *I3*, *Z*, 26 and 27 whereas N = 5 orange curves on *Tgu2*, *Tgu4*, *Tgu7*, *Tgu25* and *Tgu28*). Also note that three out of the five singleton structural variation calls that were found outside of the six focal chromosomes were present in individual taeGutGut (i.e. on chromosomes *Tgu4*, *Tgu7* and *Tgu25*), which is a hybrid with 95% of its genome coming from *T. g. guttata* (see **Methods**; also see (3, 4)).

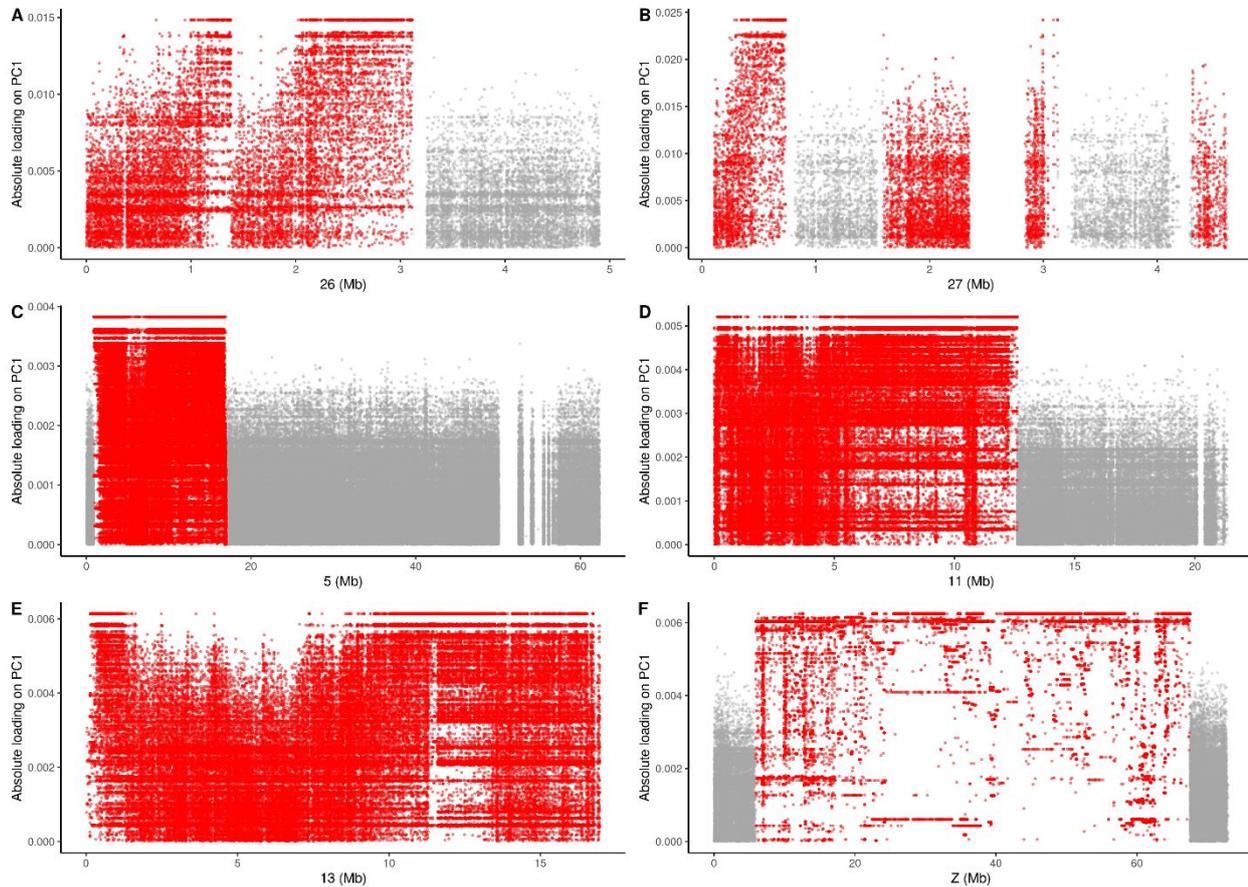


**Fig. S4.** The number of  $\log_2$ -transformed barcodes shared between distinct intra-chromosomal regions (i.e. putative inversion breakpoints) detected by Long Ranger (**Supplementary Table S3**; for details see **Methods**). Genomic positions were based on the new assembly bTaeGut1.202104 (2).

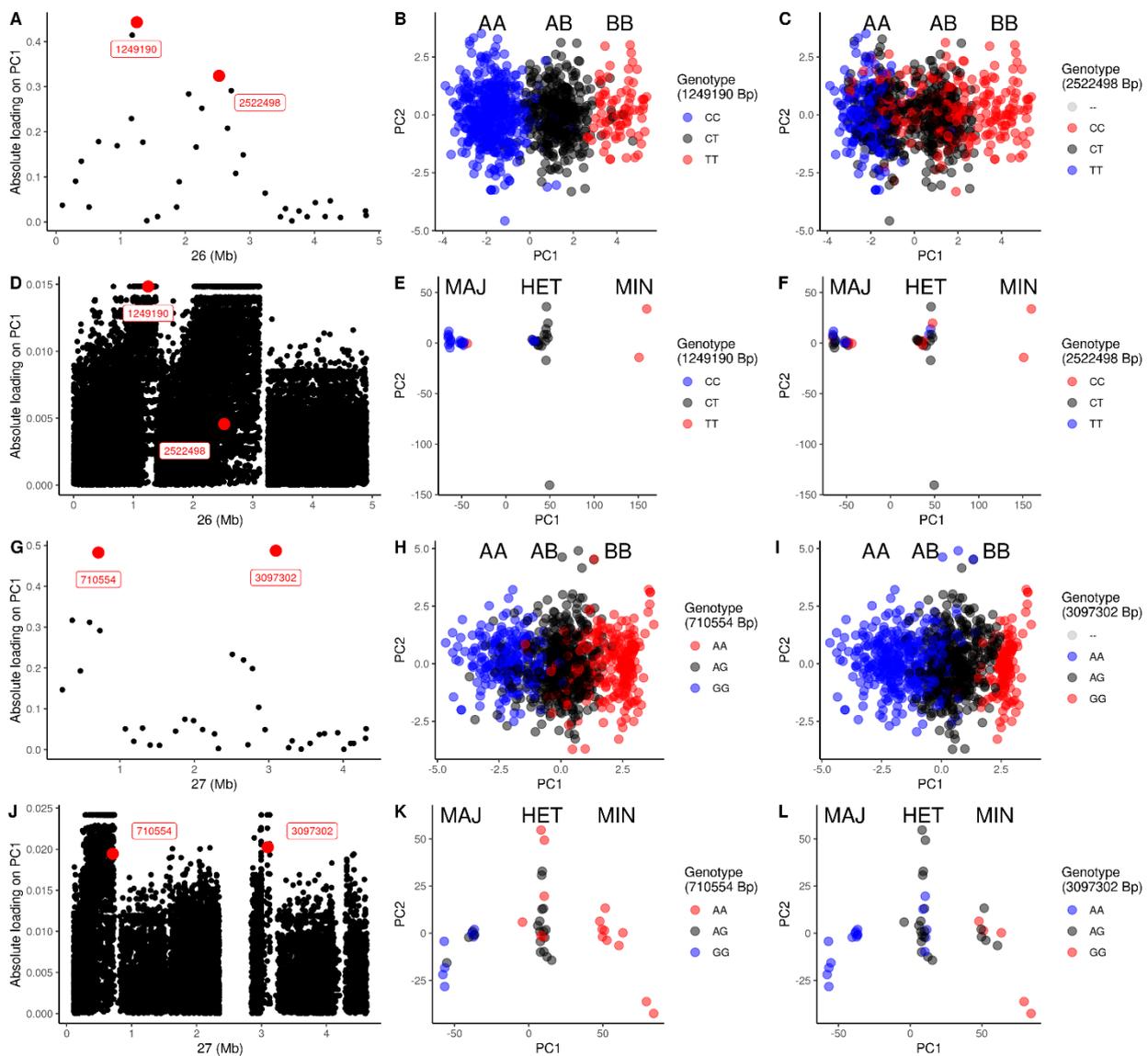




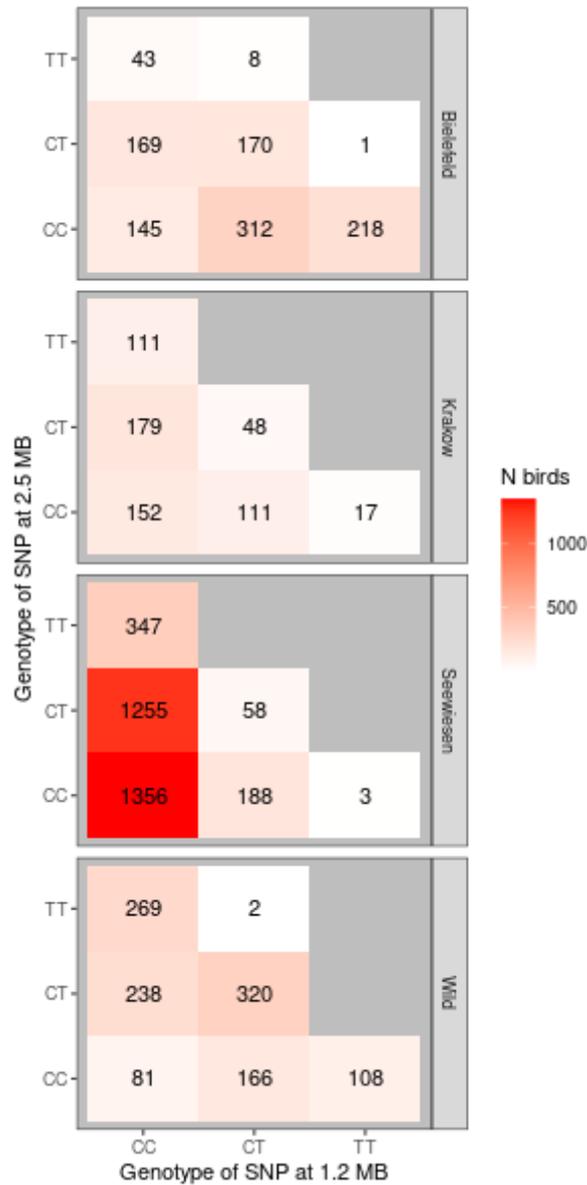
**Fig. S6.** SNP loadings on PC1 from PCA using the 19 wild zebra finches identified the putatively inverted regions (red). Dots are SNPs. Positions are based on the old zebra finch assembly TaeGut1 (5). Note that for the micro-chromosomes *Tgu26* (A) and *Tgu27* (B), the old assembly was fragmented and incomplete as compared to the new assembly bTaeGut1.202104 (2) (Fig. S2). Hence, we used the alignment of the two assemblies to guide the identification of the inverted regions for these two microchromosomes (Fig. S2). Also see Fig. 1 for the PCA plots.



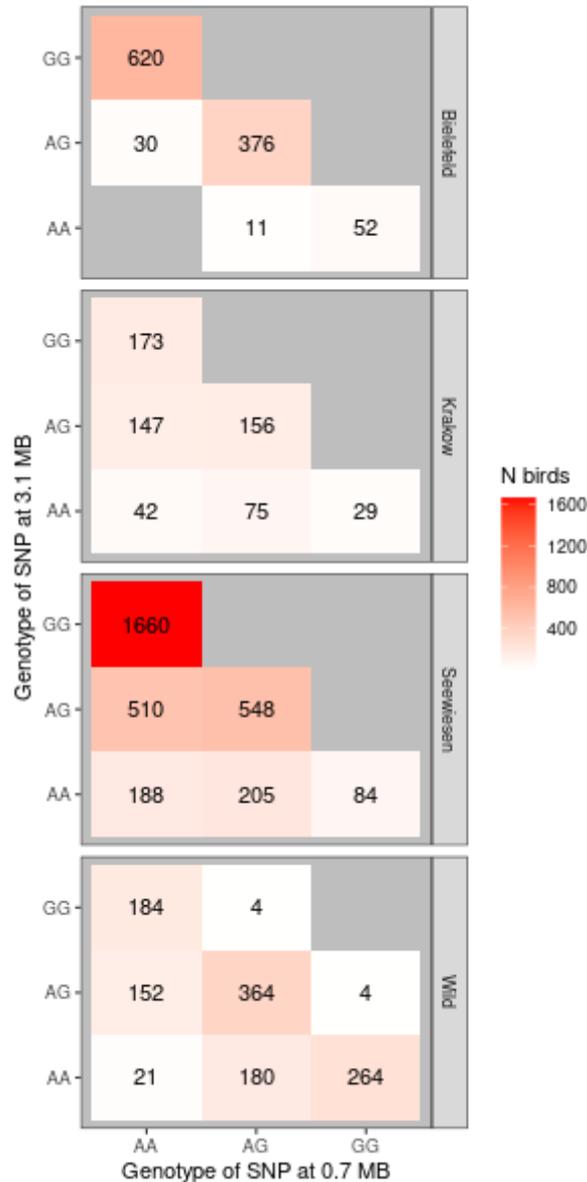
**Fig. S7.** Tag SNPs (red dots in **A,D,G,J**) for the inversions on microchromosomes *Tgu26* (**A-F**) and *Tgu27* (**G-L**) were selected based on the highest loadings on PC1 from PCA analysis of 34 (**A-C**) and 36 SNPs (**G-I**) among the 948 wild zebra finches used in Knief et al 2016 (1) (**A-C** and **G-I**). PCA analyses using SNPs from WGS data from the 19 wild zebra finches (6) (**D,J**) show that the selected tag SNPs (red dots in **A,D,G,J**) were in high LD with the defined inversion types (**E,K,L**) except for the tag SNP on *Tgu26* at 2.5 Mb that was in relatively low LD (**F**; for additional details see also **Supplementary Table S4**). Red boxes highlight the positions of the tag SNPs (**A,D,G,J**). (**B,C,E,F,H,I,K,L**) In PCA plots, blue depicts the tag SNP genotype of individuals that are homozygous for the major allele homozygous (MAJ), red depicts individuals homozygous for the minor allele (MIN) and black indicates heterozygous individuals (HET). A and B are based on the major and minor alleles among the 948 wild zebra finches in Knief et al. 2016 (1). Note that the SNP at 3,097,302 bp on *Tgu27* showed heterosis on all fitness-related traits (**Fig. 4B**).



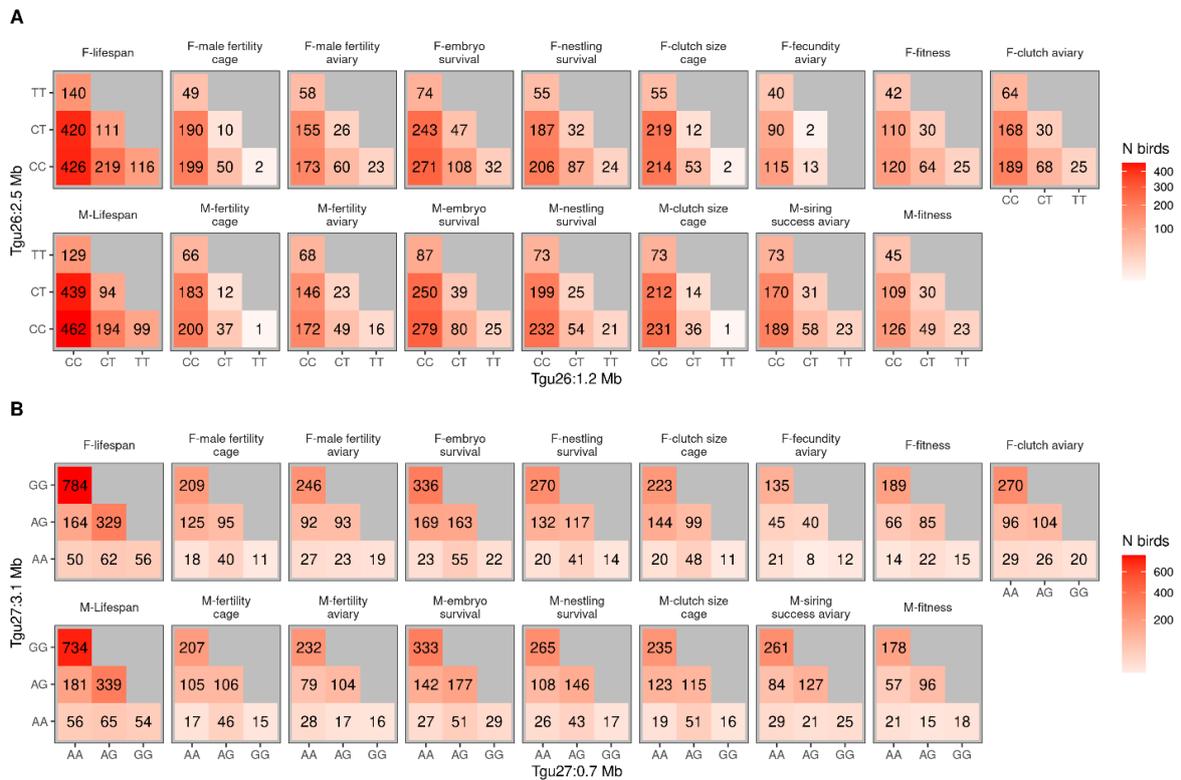
**Fig. S8.** Number of birds (older than 8 days) for all combinations of the two tag SNP genotypes on *Tgu26* (**Fig. S7A-F** and **Table S4**) in the 948 wild zebra finches (1), and in the three captive populations Bielefeld, Krakow and Seewiesen (7). X-axis shows the three genotypes for the SNP at 1,249,190 bp and y-axis shows the genotypes of the SNP at 2,522,498 bp, based on the old assembly (5). Grey indicates no birds carried such genotype. Note genotype CC-CC is the most common genotype in our captive populations.



**Fig. S9.** Number of birds (older than 8 days) for all combinations of the two tag SNP genotypes on *Tgu27* (**Fig. S7G-L** and **Table S7**) in the 948 wild zebra finches (1), and in the three captive populations Bielefeld, Krakow and Seewiesen (7). X-axis shows the three genotypes for the SNP at 710,554 bp and y-axis shows the genotypes of the SNP at 3,097,302 bp, based on the old assembly (5). Grey indicates no birds carried such genotype. Note genotype AA-GG is the most common genotype in our captive populations. Note that the SNP at 3097302 Bp showed heterosis on all fitness-related traits (**Fig. 4B**).



**Fig. S10.** Number of birds behind the estimated standardized effect sizes of all combinations of the two tag SNP genotypes on *Tgu26* (A) and *Tgu27* (B) on female and male fitness traits. For additional details see Fig. 4.



**Table S1.** Description of all WGS samples used in this study.

Individual ID	Sex	Tissue	Alternative individual ID in pedigree or database	Population	Library preparation‡	Reference
26462	Female	Blood	101	Wild	Illumina	Singhal et al. 2015
26516	Male	Blood	129	Wild	Illumina	Singhal et al. 2015
26721	Female	Blood	113	Wild	Illumina	Singhal et al. 2015
26733	Male	Blood	141	Wild	Illumina	Singhal et al. 2015
26781	Male	Blood	153	Wild	Illumina	Singhal et al. 2015
26792	Male	Blood	165	Wild	Illumina	Singhal et al. 2015
26795	Female	Blood	177	Wild	Illumina	Singhal et al. 2015
26820	Female	Blood	137	Wild	Illumina	Singhal et al. 2015
26881	Female	Blood	149	Wild	Illumina	Singhal et al. 2015
26896	Female	Blood	161	Wild	Illumina	Singhal et al. 2015
28016	Female	Blood	173	Wild	Illumina	Singhal et al. 2015
28078	Female	Blood	185	Wild	Illumina	Singhal et al. 2015
28313	Male	Blood	189	Wild	Illumina	Singhal et al. 2015
28339	Female	Blood	105	Wild	Illumina	Singhal et al. 2015
28353	Male	Blood	109	Wild	Illumina	Singhal et al. 2015
28402	Male	Blood	121	Wild	Illumina	Singhal et al. 2015
28404	Male	Blood	133	Wild	Illumina	Singhal et al. 2015
28456	Female	Blood	117	Wild	Illumina	Singhal et al. 2015
28481	Male	Blood	145	Wild	Illumina	Singhal et al. 2015
MP1	Female	-	MP1	American domesticated	Illumina	Singhal et al. 2015
MP2	Male	-	MP2	American domesticated	Illumina	Singhal et al. 2015
MP3	Male	-	MP3	American domesticated	Illumina	Singhal et al. 2015
MP4	Male	-	MP4	American domesticated	Illumina	Singhal et al. 2015
MP5	Male	-	MP5	American domesticated	Illumina	Singhal et al. 2015
bTaeGut1	Male	-	black17	American domesticated	10X	Rhie et al. 2021
bTaeGut2	Female	-	blue55	American domesticated	10X	Rhie et al. 2021
bTaeGut3	Male	-	father of blue55	American domesticated	10X	Rhie et al. 2021
bTaeGut4	Female	-	Mother of blue55	American domesticated	10X	Rhie et al. 2021
BR19117	Male	Liver	BR19117	Bielefeld recently wild-derived	PCR-free	Pei et al. 2021
CR19001	Male	Liver	CR19001	Krakow domesticated	PCR-free	Pei et al. 2021
CR19103	Male	Liver	CR19103	Krakow domesticated	PCR-free	Pei et al. 2021
MR19042	Male	Liver	MR19042	Melbourne recently wild-derived	PCR-free	Pei et al. 2021
SR12047	Male	Blood	SR12047	Seewiesen domesticated	PCR-free	Pei et al. 2021
SR12333	Female	Blood	SR12333§	Seewiesen domesticated	PCR-free	Pei et al. 2021
SR12483	Male	Liver	SR12483§	Seewiesen domesticated	PCR-free	Pei et al. 2021
SR15062	Female	Blood	SR15062	Seewiesen domesticated	PCR-free	Pei et al. 2021
SR15064	Male	Blood	SR15064	Seewiesen domesticated	PCR-free	Pei et al. 2021
tacCasSA	Male	Liver	B56513	Chimooli Dam, 8.3 km ENE Wyndham, Western Australia wild	10X	Pei et al. 2021
tacCasSE	Male	Liver	SR12451	Seewiesen domesticated	10X	Pei et al. 2021

taeGutGut	Male	Liver	TR13008	European captive <i>castanotis x guttata</i>	10X	Pei et al. 2021
SpainM1	Male	Muscle	Spain_1	Spain domesticated	PCR-free	Kinsella et al. 2019
SpainM2	Male	Muscle	Spain_2	Spain domesticated	PCR-free	Kinsella et al. 2019
SR00100	Male	Liver	SR00100	Seewiesen domesticated	10X	Kinsella et al. 2019

## References:

- S. Singhal, et al., Stable recombination hotspots in birds. *Science* (80-.). 350, 928–932 (2015)
- A. Rhie, et al., Towards complete and error-free genome assemblies of all vertebrate species. *Nature* 592, 737–746 (2021).
- Y. Pei, et al., Occasional paternal inheritance of the germline restricted chromosome in songbirds. *bioRxiv* (2021) <https://doi.org/https://doi.org/10.1101/2021.01.28.428604>.
- C. M. Kinsella, et al., Programmed DNA elimination of germline development genes in songbirds. *Nat. Commun.* 10, 5468 (2019).

**Table S2.** Genotype and allele frequencies and the ancestral state of the six large zebra finch chromosomal inversion polymorphisms.

Chr	Genotype (948 wzf in Knief et al 2016)	Genotype (19wzf)	Genotype frequency (19wzf)	Genotype frequency (24 zf)	Allele frequency (19wzf)	Allele frequency (24 zf)	N birds (19wzf)	N birds (24 zf)	Ancestral state (Fig. 3)
26	AA	maj	0.47	0.50	0.71	0.73	9	12	Ancestral
26	AB	het	0.47	0.46	-	-	9	11	-
26	BB	min	0.05	0.04	0.29	0.27	1	1	Derived
27	AA	maj	0.26	0.29	0.58	0.48	5	7	Ancestral
27	AB	het	0.63	0.38	-	-	12	9	-
27	BB	min	0.11	0.33	0.42	0.52	2	8	Derived
5	AA	maj	0.32	0.29	0.61	0.56	6	7	Ancestral
5	AB	het	0.58	0.54	-	-	11	13	-
5	BB	min	0.11	0.17	0.39	0.44	2	4	Derived
11	AA	maj	0.37	0.29	0.61	0.50	7	7	Derived
11	AB	het	0.47	0.42	-	-	9	10	-
11	BB	min	0.16	0.29	0.39	0.50	3	7	Ancestral
13	AA	maj	0.37	0.38	0.61	0.65	7	9	Derived
13	AB	het	0.47	0.54	-	-	9	13	-
13	BB	min	0.16	0.08	0.39	0.35	3	2	Ancestral
Z	AA	maj	0.33	0.25	0.50	0.53	3	6	Derived
Z	BB	min	0.11	0.04	0.43	0.21	1	1	Derived (slightly younger than A)
Z	CC	min	0.00	0.05	0.07	0.26	0	1	Derived from C
Z	AB	het	0.33	0.16	-	-	3	3	-
Z	AC	het	0.11	0.26	-	-	1	5	-
Z	BC	min	0.11	0.16	-	-	1	3	-
Z	AW*	-	0.4	0.6	-	-	4	3	-
Z	BW*	-	0.6	0.2	-	-	6	1	-
Z	CW*	-	0	0	-	-	0	1	-

\* Z inversion types for females.

**Table S3.** Putative inversion breakpoints detected via barcode interactions using Long Ranger on the new assembly bTaeGut1.202104 across eight zebra finch libraries. For each putative inversion breakpoint, the mean and the range of the predicted breakpoint were given in bp. Centromere information was inferred from Knief & Forstmeier (2016) and **Fig. S2**.

Chr	Break point 1 (bp)	Break point 2 (bp)	length (Mb)	Remarks
5	1893285 [1864238-1922332]	18265696 [18267704-18269712]	16.37	Pericentric
5	846418 [832197-853579]	13873812 [13872932-13875054]	13.03	Pericentric
11	8044 [7541-8548]	12484390 [12482869-12485911]	12.48	Paracentric
13	2205887 [2204762-2208730]	17902407 [17898043-17904840]	15.70	Paracentric
13	2003255 [2003209-2003383]	7314643 [7314624-7314667]	5.31	Paracentric
26	3450204 [3445346-3457364]	6804355 [6803332-6804894]	3.35	Pericentric
27	22176 [21947-22404]	3321144 [3321002-3321285]	3.30	-
Z*	[6412262-6412451]	[71635454-71635549]	65.22	Contain centromere; breakpoints unclear

\* Note that the two breakpoints on *TguZ* are uncertain. The barcode interaction signals on *TguZ* show repetitive elements (i.e. high coverage on Barcodes in all samples; Supplementary Fig. S3), hence we dropped it in the main text and we only report here for completeness of information.

Reference: U. Knief, W. Forstmeier, Mapping centromeres of microchromosomes in the zebra finch (*Taeniopygia guttata*) using half-tetrad analysis. *Chromosoma* 125, 757–768 (2016).

**Table S4.** Tag SNPs for the inversions on *Tgu26* and *Tgu27* (**Supplementary Fig. S7**).

SNP Name	Chr	Pos (old)	REF	ALT	REF freq*	Pearson's correlation with inversion type (19 wzf)	t (19 wzf)	P (19 wzf)	Pearson's correlation with inversion type (24 zf)	t (24 zf)	P (24 zf)
WZF00114346	26	1249190	T	C	0.285	1	>10	<0.0001	0.62	3.67	0.001
WZF00114455	26	2522498	C	T	0.527	0.3	1.30	0.2	0.18	0.85	0.4
WZF00114985	27	3097302	G	A	0.375	0.83	6.19	<0.0001	0.82	6.67	<0.0001
WZF00115358	27	710554	A	G	0.525	0.81	5.61	<0.0001	0.84	7.29	<0.0001

\*Reference allele frequency in 948 wild zebra finches (Knief et al. 2016).

**Table S5.** Estimated standardized effect sizes of genotypes of the tag SNPs on *Tgu26* on female and male fitness traits (**Fig. 4A**). For details of the SNPs see **Supplementary Table S4**.

This table contains 112 rows.

**Table S6.** Estimated standardized effect sizes of genotypes of the tag SNPs on *Tgu27* on female and male fitness traits (**Fig. 4B**). For details of the SNPs see **Supplementary Table S4**.

This table contains 113 rows.

Tables S5 & S6 will be uploaded with the publication.

## References

1. U. Knief, *et al.*, Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biol.* **17**, 199 (2016).
2. A. Rhie, *et al.*, Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **592**, 737–746 (2021).
3. Y. Pei, *et al.*, Occasional paternal inheritance of the germline restricted chromosome in songbirds. *bioRxiv* (2021) <https://doi.org/https://doi.org/10.1101/2021.01.28.428604>.
4. U. Knief, H. Schielzeth, H. Ellegren, B. Kempnaers, W. Forstmeier, A prezygotic transmission distorter acting equally in female and male zebra finches *Taeniopygia guttata*. *Mol. Ecol.* **24**, 3846–3859 (2015).
5. W. C. Warren, *et al.*, The genome of a songbird. *Nature* **464**, 757–762 (2010).
6. S. Singhal, *et al.*, Stable recombination hotspots in birds. *Science (80-. )*. **350**, 928–932 (2015).
7. Y. Pei (裴一凡), *et al.*, Proximate Causes of Infertility and Embryo Mortality in Captive Zebra Finches. *Am. Nat.* **196**, 577–596 (2020).



## Chapter 4

### Antagonistic effects on fitness

"Having two copies of the **inversion** is good for having more kids, "



"... but bad for them to survive."



## Chapter 4

### Maintaining an inversion polymorphism by antagonistic effects on fitness

Yifan Pei, Wolfgang Forstmeier, Ulrich Knief, Bart Kempenaers

Chromosomal inversions are commonly recognized as super genes that are maintained by various forms of balancing selection in the population. Chromosomal inversions typically create tight linkage between alleles at numerous loci. If the inversion links an additive beneficial allele with a recessive deleterious allele, the polymorphism can be maintained through antagonistic pleiotropy of the entire inversion haplotype. Here, we study the mechanisms involved in the maintenance of an inversion polymorphism that links 265 genes on chromosome *Tgu11* in the zebra finch (*Taeniopygia guttata*). Based on data from over 6000 captive birds, we estimated the effects of this inversion on a wide range of fitness-related traits, including male siring success, female fecundity, offspring and individual survival. We then compared the fitness between individuals that are homozygous for the ancestral A-allele, heterozygous individuals and individuals that are homozygous for the derived B-allele. We found that males with the derived inversion haplotype B had higher siring success and females with B were more fecund, while individuals that were homozygous for B had lower rates of embryo and nestling survival. An individual-based simulation model using the estimated fitness effects from a complete generation cycle shows that the polymorphism could have spread and be maintained over time. Our results highlight that inversion polymorphisms can become stabilized at an intermediate allele frequency if the alleles have antagonistic fitness effects.

Prepared as:

Pei, Y., Forstmeier, W., Knief, U., & Kempenaers, B. Maintaining an inversion polymorphism by antagonistic effects on fitness.

## Introduction

Inversions are segments of chromosomes that are arranged in reverse order. The inversion and its non-inverted ancestral type have the same gene content, but recombination between the two types is suppressed (Kirkpatrick 2010). Once an inversion has arisen, drift and directional selection (if it conveys positive effect on fitness) might cause it to spread and ultimately to get fixed in the population. However, the latter seems rather unlikely, given that beneficial mutations are relatively rare compared to mutations that are neutral or detrimental. Yet, many studies have shown evidence for stable inversion polymorphisms in a population (Wellenreuther & Bernatchez 2018). The maintenance of these polymorphisms over time can be considered an evolutionary puzzle (Hoffmann *et al.* 2004; Hoffmann & Rieseberg 2008; Kirkpatrick 2010; Wellenreuther & Bernatchez 2018), because the offspring of individuals that are heterozygous for the inversion suffer from increased mortality caused by unbalanced cross-overs in the inverted region (Roberts 1967; Knief *et al.* 2016).

Inversion polymorphisms could be maintained through various forms of balancing selection. (1) Heterosis, i.e. heterozygous individuals have higher fitness than homozygous carriers. As an extreme example, in *Drosophila tropicalis* all individuals that are homozygous for an inversion on chromosome 2 die before becoming adults (Dobzhansky & Pavlovsky 1955). (2) Antagonistic pleiotropy, where genes show opposing effects on different traits. For example, in the seaweed fly *Coelopa frigida*, where a chromosomal inversion polymorphism causes opposite effects on female fecundity and offspring survival (Mérot *et al.* 2020). (3) Inter-chromosomal incompatibilities between loci (epistasis), which is difficult to detect due to the complexity of multi-locus interactions. For example, in *Drosophila melanogaster*, the co-occurrence of two inversions on two different chromosomes is non-random, where there was a significant excess of individuals that were dual homozygous for the ancestral types or dual heterozygous for both inversions (Singh & Das 1991).

Regardless of the mechanism maintaining them, chromosomal inversions can have large phenotypic effects. In a variety of organisms, inversion polymorphisms result in distinct morphs or alternative life-history strategies (e.g. in ruffs *Philomachus pugnax* (Küpper *et al.* 2015; Lamichhaney *et al.* 2015), white-throated sparrows *Zonotrichia albicollis* (Tuttle 2003; Tuttle *et al.* 2016), fire ants *Solenopsis invicta* (Pracana *et al.* 2017), butterflies *Heliconius numata* (Jay *et al.* 2018) and stick insects *Timema cristinae* (Lindtke *et al.* 2017)). Inversion polymorphisms are also known to play a role in local adaptation. In this case, they typically manifest as changes in the frequency of the inverted allele along

an environmental gradient (e.g. in yellow monkeyflowers *Mimulus guttatus* (Lowry & Willis 2010; Lee *et al.* 2016) and fruit flies *Drosophila melanogaster* (Kennington *et al.* 2007; Durmaz *et al.* 2018)). Other inversion polymorphisms appear to have less obvious phenotypic effects, but they affect multiple life-history traits. For instance, in the seaweed fly *Coelopa frigida*, an inversion on chromosome I shows opposite effects on fitness and survival (Betran *et al.* 1998; Mérot *et al.* 2020), and in the zebra finch, *Taeniopygia guttata*, an inversion on chromosome Z increases the mortality of embryos sired by heterozygous males (Knief *et al.* 2016), but these males also have faster-swimming sperm and have higher siring success (Kim *et al.* 2017; Knief *et al.* 2017).

Overall, the mechanism behind the maintenance of many inversion polymorphisms remains unclear (Wellenreuther & Bernatchez 2018), and this is particularly true for inversions that have no obvious phenotypic effects. Yet, the above examples have in common that inversions capture many loci, which implies that there is at least potential for antagonistic pleiotropic effects on fitness.

Zebra finches have at least four large (>250 genes) chromosomal inversions that all segregate at around 50% allele frequencies, both in the wild and in captivity (Christidis 1986; Itoh *et al.* 2011; Knief *et al.* 2016). A previous study investigated the effects of these inversions on a range of morphological traits (Knief *et al.* 2016). This study reported that half of the combinations of inversions and morphological traits (20 out of 40) showed effects beyond sampling noise (9 of which were statistically significant). This illustrates that when combining the allelic effects of about 250 genes into a haplotype (i.e. as an inversion), this will affect many of the studied morphological traits, which are thought to have a highly polygenic genetic architecture (Knief *et al.* 2016). Fitness as one of the most polygenic traits, is also thought to be a particularly large target to be influenced by *de novo* mutations, so we expect that inversions may often capture loci that have effects on fitness-related traits.

Among the four chromosomal inversions of the zebra finch, only the largest one, located on the sex chromosome *TguZ*, appears to be maintained by an overdominance effect (Kim *et al.* 2017; Knief *et al.* 2017, 2019). The mechanism maintaining the other three inversion polymorphisms remains unclear (Knief *et al.* 2016). Although, the derived (inverted) type on chromosome *Tgu11* showed beneficial additive effects on female fecundity and male siring success in captive zebra finches (Knief *et al.* 2016), selective forces that would balance the system remain unclear. The inversion on *Tgu11* captures 265 genes and spans 57% of the assembled chromosome (12Mb) (Knief *et al.* 2016), such that it is indeed likely to influence multiple fitness-related traits (i.e. pleiotropy).

In this study, we first quantified the fitness effects of the inversion. We measured individual survival and reproductive performance in relation to *Tgu11* inversion genotype using 14 years of breeding data with more than ten thousand informative embryos from four captive populations of zebra finches. We then used individual-based simulations to investigate the evolution of the *Tgu11* inversion polymorphism. We simulated populations containing the inversion and used the estimated effect sizes on fitness to assess frequency changes of the inversion over time.

## Methods

### Study populations

We studied four captive populations of Australian zebra finches that were maintained at the Max Planck Institute for Ornithology, Seewiesen, Germany: (1) ‘Seewiesen’ (population #18 in (Forstmeier, Segelbacher, Mueller, & Kempenaers, 2007)), (2) ‘Bielefeld’ (population #19 in (Forstmeier et al., 2007)), (3) ‘Melbourne’ (Jerónimo et al., 2018) and (4) ‘Krakow’ (hybrids between a zebra finch population from Krakow (#11 in Forstmeier et al. 2007) and the ‘Seewiesen’ population). These populations were studied over 3 to 14 generations (see Table S1). Populations ‘Seewiesen’ and ‘Krakow’ were domesticated while ‘Bielefeld’ and ‘Melbourne’ were recently wild-derived. Wild-derived birds were smaller in body size (~11 g) and shyer compared to the domesticated (~15 g) birds (Pei (裴一凡) *et al.* 2020).

### Tag SNP genotyping

Inversion types of all adult birds were taken from (Knief *et al.* 2016). Note that we here refer to the ancestral genotype as allele A and to the derived, inverted genotype as allele B (opposite to Knief et al., 2016, who labelled the ancestral genotype as allele B because its allele frequency was below 50%). For genotyping all other individuals (dead embryos and nestlings), we selected one tag SNP (WZF00031805, see Table S2 for details) from Knief et al. (2016) based on the highest  $r^2$ -value (i.e. variance explained of the inversion genotype of the individual,  $r^2 = 0.994$  based on 948 individuals from the wild) and a call rate  $>0.99$ . To study the effects of the inversion genotype on offspring survival, we genotyped 3,022 dead embryos and 1,522 dead nestlings with the Roche LightCycler Instrument following the manufacturer’s guide. Genotypes were called by the default parameters on the Roche LightCycler Software. Eight out of 11,458 cases (total number of genotyped individuals) failed the inheritance check and were removed from analysis. Failed cases (0.07%) were either embryos (N = 7)

nestlings that had died at a young age ( $N = 1$ ), and inheritance errors were likely due to aneuploidy or to falsely assigning an individual as homozygous. The allele frequency of the tag SNP in wild birds and among the captive populations is shown in Supplementary Table S1.

#### *Fecundity, fertility and viability measurements*

Data on reproductive performance, lifespan, and nestling body mass at eight days of age were taken from (Pei (裴一凡) *et al.* 2020; Pei *et al.* 2020). Reproductive performance traits had been measured either (1) in ‘cages’ containing a single breeding pair or (2) in ‘aviaries’ with a group of males and females (typically 12 birds with equal sex ratios, range: 10-42 birds and 0.4-0.6 in the proportion of males) where breeding pairs were formed freely. During the breeding season, nests were checked daily to record hatching dates. Details on rearing conditions, timing of the breeding seasons, and all measured reproductive traits are described in Pei *et al.* (2020b).

In brief, we studied four female reproductive performance traits: (1) clutch size in cages ( $N = 562$  females), (2) clutch size in aviaries ( $N = 703$  females), (3) the total number of eggs laid over the entire breeding season in aviaries where offspring rearing was not allowed and eggs were replaced by plastic eggs ( $N = 285$  females), and (4) the number of independent young a female produced (genetic mother, determined using genotyping with microsatellite markers) within a breeding season in aviaries where offspring rearing was allowed, ( $N = 438$  females). For males, we studied (5) egg fertility in cages, i.e. whether an egg laid in a cage was fertilized ( $N = 504$  males), (6) egg siring success in aviaries, i.e. whether an egg laid by the male’s social partner was fertilized by him ( $N = 512$  males; determined using genotyping with microsatellite markers), (7) the total number of eggs sired in an aviary, including both within- and extra-pair ( $N = 724$  males), and (8) the total number of independent young sired in an aviary ( $N = 432$  males).

We also estimated the effect of the *Tgu11* inversion on individual (offspring) survival. Here, we included (9) the effect of the genotype of the genetic mother on embryo hatching success ( $N = 940$  females), because Pei *et al.* (2020b) showed that the identity of the genetic mother significantly predicted embryo survival, and the effect of the individual’s own genotype on (10) hatching success (yes/no;  $N = 7335$  embryos), (11) body mass at an age of 8 days ( $N = 4622$  nestlings), (12) survival

until independence (i.e. day 35, yes/no; N = 3819 nestlings), and (13) lifespan (N = 2904 birds; also see Table S3).

We explored the effects of the inversion on offspring survival in more detail, by dividing different periods of mortality. Embryonic age at death was recorded by one observer (K. Martin) in days (following Mak et al., 2015). Most embryos died in the first few days after incubation started, with a second, smaller mortality peak a few days before hatching (Fig. S1A). Based on the age that showed the lowest embryonic death rate (i.e. 8 days after incubation started), we divided the data in two classes: early ( $\leq 7$  days) versus late embryo mortality ( $\geq 8$  days; Fig. S1A). Similarly, we also categorized offspring mortality, into early nestling mortality ( $\leq 7$  days), late offspring mortality (death between days 8 and 35) and mortality after day 35 until recruitment as adults (day 100; Fig. S1B).

All adult birds were genotyped for the *Tgull* inversion throughout the study period (2004-2017), whereas dead embryos and nestlings were collected and genotyped only from 2012 onwards. Hence, when estimating the effect of the offspring's genotype on mortality, we only used the 2012-2017 breeding seasons.

### *Statistical analyses*

We fitted mixed-effect models using the function `lmer` in the package 'lme4' (Bates et al., 2018) in R V3.6.1. We adapted the model structures used in (Pei et al., 2019) on reproductive performance, offspring survival and lifespan (Supplementary Tables S3 and S4 show sample sizes and all fixed effects, while Table S5 lists all random effects). To compare and meta-summarize the effects of inversion genotype on different traits, we Z-scaled each response variables (i.e. the 13 traits listed above), assuming a normal distribution (Knief & Forstmeier 2018). We fitted the inversion genotype as a fixed effect with three levels (AA: homozygous ancestral type, AB: heterozygous, BB: homozygous derived) and show the result relative to the AA ancestral baseline (Table S3). We controlled for the following Z-scaled covariates in all models: inbreeding (i.e. pedigree-based inbreeding coefficient  $F$ ), age of reproduction (i.e. age at laying for male infertility, embryo and nestling mortality; age at the start of the clutch for models of clutch size; and age at the start of the breeding season for male siring success, female fecundity and the number of seasonal recruits), or – for nestling mass – age at measurement (most nestlings were weighed at 8 days of age, but some weighed at day 7, 9 or 10), the number of days the focal bird participated in a breeding season and sex

ratio (i.e. the proportion of males in an aviary). Whenever applicable, we also controlled for laying order (eggs within a clutch), hatching order (nestlings within a brood), clutch order (within a breeding season), whether an egg was laid while nestlings from a previous clutch were still present (yes/no), pair formation type (i.e. forced pairs in cages or free-choice in aviaries), cross-fostering regime (i.e. no cross-fostering, cross fostered to parents from the same population, domesticated nestlings being cross-fostered to wild-derived parents and wild-derived nestlings being cross-fostered to domesticated parents), and for population (Table S4). To account for non-independence between observations, we also included female, male, pair, clutch and breeding season identity as random effects, where applicable (Table S5). To check the robustness of the effects of the inversion genotype, we fitted all models again without the covariates, but including the random effects.

We meta-summarized the effects of the inversion genotype (i.e. AB and BB) in two weighted ‘lm’ models using the R package ‘stats’, one for the reproductive performance traits (1-8) and one for the survival traits (9-13) listed above). As response variables, we used the estimated genotype effects, from the mixed-models, weighted by the multiplicative inverse of its standard error to account for its level of uncertainty. We removed the intercept and fitted the genotype as the only fixed effect (2 levels).

#### *Simulating the evolution of the Tgu11 inversion via antagonistic pleiotropic effects*

To simulate the evolution of the inversion polymorphism, i.e. the allele frequency changes, over time, we re-estimated the effects of the inversion on fitness traits on their original scale. We focused on key phases of reproduction and survival per generation, covering the entire reproductive cycle in aviary setups (phases P1a, P1b, P2 and G1-G4 in Supplementary Fig. S2 and Table S6). The number of eggs laid by females (P1a in Fig. S2 and Table S6) and sired by males (P2 in Fig. S2 and Table S6) within a breeding season followed a zero-inflated Poisson distribution. Hence, we calculated the proportion of zeros and the mean after removing zeros for individuals with genotypes AA, AB and BB, respectively. As the absolute number of eggs produced in a breeding season depends on the experimental setup, we extracted measurements of female fecundity and male siring success from a single breeding season (i.e. the aviary experiment ‘SelectionLines S3’ in (Pei (裴一凡) *et al.* 2020)) that provided standardized estimates for 219 females and 217 males (Wang *et al.* 2017, 2020). Note that the fertility probability of an egg is mainly a male-specific trait (Pei (裴一凡) *et al.* 2020), but in aviary setups it is complicated by within- and extra-pair paternities. Therefore, to complete the full reproduction cycle for simple simulations, we additionally estimated the female fertility probability

for each genotype (AA, AB and BB) in aviaries experiments, i.e. if an egg laid by the female was fertilized or not (P1b in Fig. S2 and Table S6), regardless of it being fertilized a within-pair or extra-pair male ( $N = 536$  females). Then, the survival of an embryo to recruitment was simulated from binomial distributions, with the probability of survival based on the genotype combination of the embryo and its mother (G1-G4 in Table S6). To estimate early embryo survival (G1), nestling survival to fledging (G3) and fledgling survival to recruitment (G4) as a function of the offspring's own genotype and to estimate late embryo survival as a function of the mother's genotype (G2), we fitted the same models as described in *Statistical analyses* but without scaling the response variables and we removed the intercept. To simplify the simulation of the genotype effects on survival, we then calculated the overall survival probability from embryo to recruitment (G1-4 in Table S6) as the product  $G1 \times G2 \times G3 \times G4$  based on the genotypes of the embryo and its mother.

To assess whether the *Tgu11* inversion can spread or whether the inversion polymorphism can be maintained by antagonistic effects on reproductive performance and survival, we simulated frequency changes in the B allele (i.e. the derived inversion genotype) using our estimated effect sizes on fitness-related traits in individual-based models. For simplicity, we assumed a single locus following Mendelian inheritance, segregating in a population with equal sex ratio, with discrete generations, without mutations and with random mating. We examined B allele frequency changes in 50 independent simulations over 1000 generations each.

To evaluate the initial spreading of the derived B allele, we ran a simulation starting with an initial population of 400 adult birds with a B allele frequency of 0.125% in Hardy-Weinberg equilibrium (i.e. 399 birds of type AA and 1 of type AB; with sex randomly assigned). To evaluate the maintenance of the B allele, we ran a simulation starting with an initial population of 400 adult birds with a B allele frequency of 50% in Hardy-Weinberg equilibrium (i.e. 100 birds of type AA, 200 of type AB and 100 of type BB; with sex randomly assigned). We repeated this with an initial population of 5000 adult birds.

Each generation started with the production of embryos (phases P1a, P1b and P2 in Supplementary Table S6, Fig. S2). First, we simulated the number of eggs laid by each female based on her genotype. Second, for each egg, we simulated its chance of being fertilized based on the fertility rate of its maternal genotype. Third, we simulated the number of embryos sired by each male based on his

genotype. Fourth, we randomly paired the previously simulated genotypes between males and females to form a “fertilized” embryo (because we randomly generated egg numbers for males and females separately, we needed to “discard” excess embryos that had only one parent assigned, which was done randomly with regard to genotype). Then, the genotype of each embryo is simulated by randomly drawing one allele from each genetic parent, following Mendel’s law of inheritance.

In the next step, the survival of each embryo is determined by drawing from a binomial distribution, based on the survival probability estimated by the combination of genotypes of the embryo and its mother (Supplementary Table S6 and Fig. S2).

Finally, the frequency of the derived B allele is calculated from all surviving recruits. Then, 400 (or 5000) birds were drawn randomly from those recruits as the starting point for the next simulated generation.

We ran the simulation described above in four scenarios to assess (1) the initial spreading of the B allele with full antagonistic pleiotropy effects (reproduction and survival effects), (2) the maintenance of B with the full antagonistic pleiotropy effects, (3) the effect of drift by fixing all trait values of genotypes AB and BB to equal those of AA, (4) additive positive effects on reproductive performance by using only the estimated trait values of female fecundity and male siring success for each genotype, and (5) recessive negative effects on survival by using only the estimated trait values of survival for each genotype (Table S6, Fig. S2). We compared the number of cases of fixation versus loss of the B allele between different scenarios with the `t.test` function in the R package ‘stats’.

All data and all scripts will be uploaded to the Open Science Framework or available upon request.

## Results

### *Population frequency of the inverted allele*

In the four captive populations, the allele frequency of the derived allele B ranged from 0.34 to 0.55 (Supplementary Table S1), which is not far from the frequency of 0.53 that has been estimated in the wild (Knief et al. 2016).

*Additive beneficial effects of the inverted allele on reproductive performance*

In comparison to the ancestral genotype AA, the inverted type B significantly increased male and female reproductive performance in an additive manner (Figs. 1A and 1C; meta-summarized standardized effect sizes of genotype AB:  $b = 0.094 \pm 0.024$  SE,  $P = 0.002$  and genotype BB:  $b = 0.157 \pm 0.026$  SE,  $P < 0.0001$ ). Female zebra finches that carried at least one copy of the B allele laid significantly more eggs per clutch (in aviaries and in cages), while male zebra finches with the B allele secured more paternity within their clutch and sired significantly more eggs in communal aviaries (Fig. 1A; also see Supplementary Figure S3 for estimates from models without covariates).

*Recessive deleterious effects of the inverted allele on survival*

Inversion type B reduced the probability of offspring survival recessively (Figs. 1B and 1C; meta-summarized standardized effect sizes, AB:  $b = -0.003 \pm 0.005$  SE,  $P = 0.53$ ; BB:  $b = -0.047 \pm 0.005$  SE,  $P < 0.0001$ ; see Tables S3 for each estimated effect size). Embryos were less likely to survive if the embryo itself or the genetic mother of the embryo were homozygous for the B allele. Similarly, nestlings with a BB genotype were on average lighter at 8 days of age and less likely to survive to independence and they overall had a shorter lifespan (Fig. 1B). Note, however, that none of these detrimental effects reached statistical significance on their own.

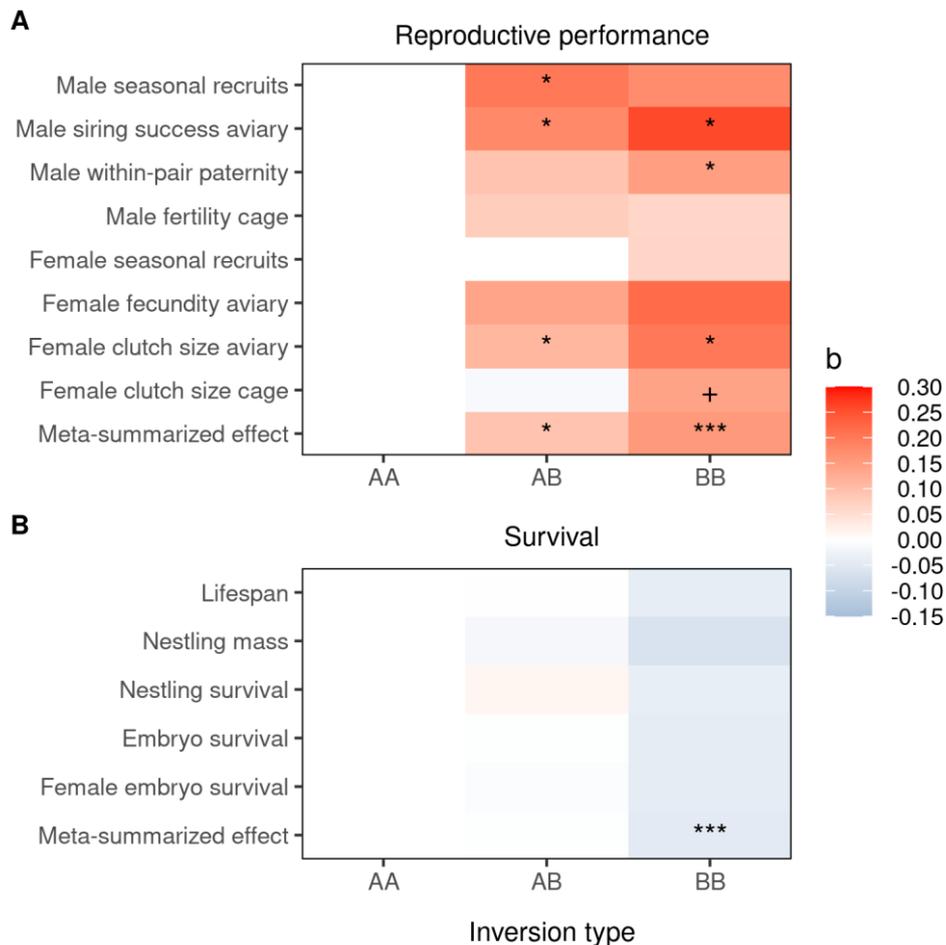
Partitioning embryo and nestling survival into different stages (Supplementary Fig. S1) showed that the probability of early embryo survival tended to be lower if the embryo was homozygous for the B allele ( $b = -0.069 \pm 0.038$ ,  $P = 0.07$ ), but was not influenced by the genotype of its mother. In contrast, the probability of late embryo survival tended to be lower if its genetic mother was BB ( $b = -0.062 \pm 0.037$ ,  $P = 0.098$ ), but did not depend on its own *Tgu11* inversion genotype (Supplementary Table S3). Early nestling survival was not affected by the nestlings' genotype, whereas late nestling survival was slightly reduced if the nestling carried the B allele, although this was not significant (Supplementary Table S3).

The overall standardized effect sizes of the inverted *Tgu11* allele on all survival traits were small ( $b$  of BB ranges from -0.064 to -0.039; Supplementary Table S3).

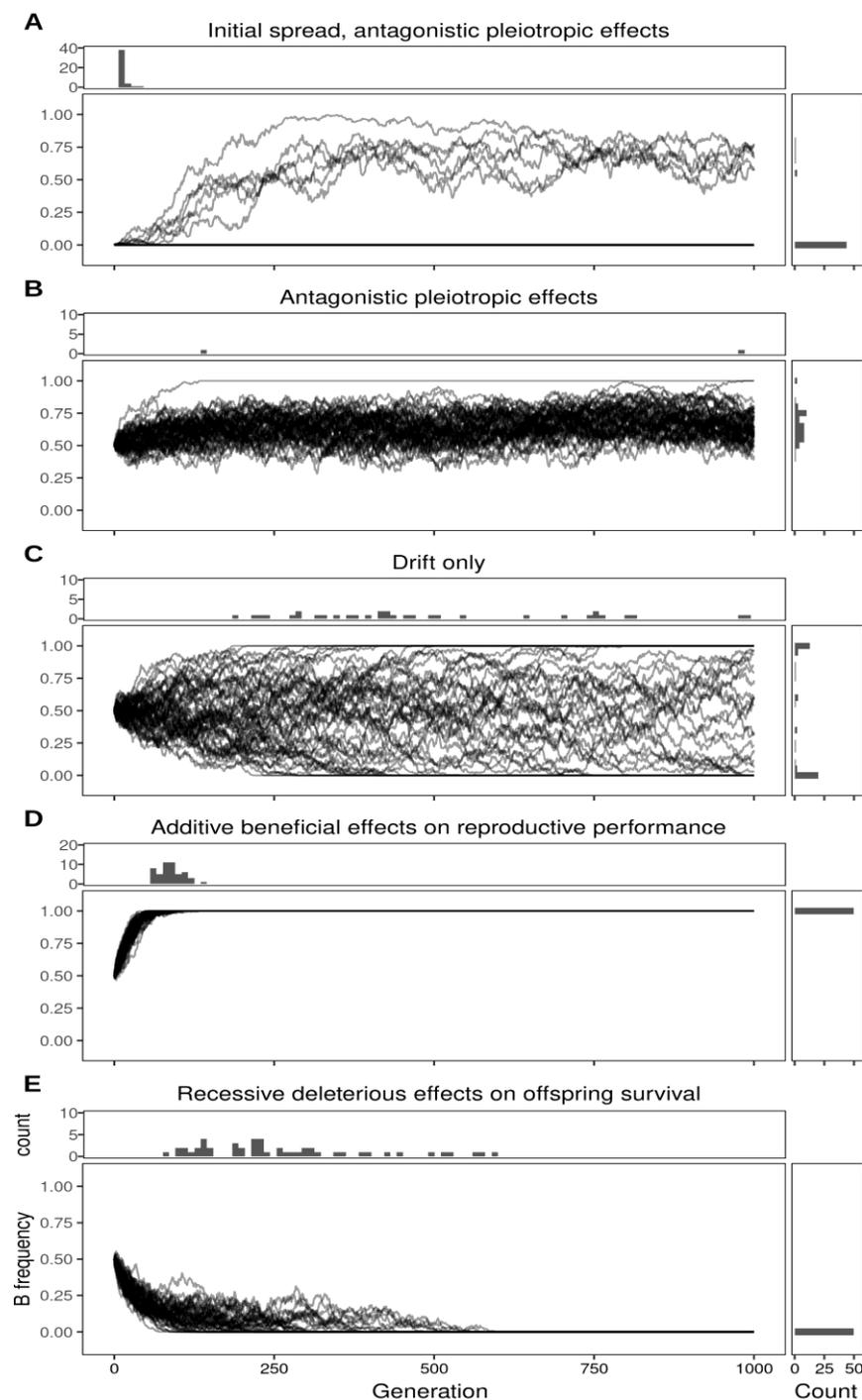
*Initial spreading of the inverted allele and maintenance of both alleles via antagonistic pleiotropy*

The B allele was able to spread in a population from a single mutational event in 12% of 50 simulated cases (Fig. 2A). When starting from an initial population of 400 adults and a 50% B allele frequency, the inverted allele was maintained for more than 1000 generations in 96% of all simulations with the antagonistic pleiotropic effects (Fig. 2B). With an initial population of 5000 adults, this increased to 100% of simulations; Fig. S4). In contrast, under scenarios with only drift or with either only the positive effects on reproductive performance or the negative effects on offspring survival, the polymorphism was less likely to be maintained (in 34% cases with population size of 400 and only drift, Fig. 2C, see also Figs. 2D and 2E).

**Figure 1.** Estimated standardized effect sizes of the *Tgull1* inversion genotypes (AB and BB) relative to the homozygous ancestral genotype (AA) representing the baseline phenotype (indicated in white). Effects on male and female reproductive performance traits (A) and on offspring survival (B), with their meta-summarized effects. Red indicates positive effects on fitness-related traits whereas blue indicates negative effects. For each genotype, ‘+’ indicates  $P < 0.1$ , ‘\*’ indicates  $P < 0.05$ , ‘\*\*\*’ indicates  $P < 0.0001$ ; no symbol indicates  $P \geq 0.1$ . All estimates can be found in Supplementary Tables S3 and S4. See also Supplementary Fig. S3 for estimates in models excluding all covariates.



**Figure 2.** Changes in the simulated frequency of the inverted B allele in a population with 400 adult birds over 1000 generations using estimated total fitness measures (Table S6; combination of female fecundity corrected for fertility in aviaries, male siring success, and embryo survival until recruitment; Fig. S2). We also simulated random genetic drift by randomly drawing 400 recruits from the pool of all produced recruits of the previous generation (336 to 652 recruits; see methods for details). Shown are five simulation scenarios illustrating the initial spread of allele B (A), the maintenance of allele B via antagonistic pleiotropic effects on both reproductive performance and offspring survival (B), drift only (C), additive positive effects of allele B on reproductive performance (D), and recessive deleterious effects of allele B on offspring survival (E). Histograms at the top of each panel (A–E) show time points of allele fixation; histograms on the right of each panel show B-allele frequency at generation 1000. Note that the B allele was maintained at a frequency of 0.65 (SD = 0.12) via effects of antagonistic pleiotropy in populations with 400 birds (B; see also Fig. S4 for simulations with populations of 5000 birds).



## Discussion

Our analyses suggest that the inversion polymorphism on chromosome *Tgu11* could be maintained by antagonistic pleiotropic effects on fitness-related traits. Expanding the results of Knief *et al.* (2016), who found an effect of the inversion allele on male siring success and female fecundity, we show that the derived inversion allele should be able to spread initially by its beneficial additive effects on reproductive success (Fig. 1A,C and Fig. 2B). Yet, the fixation of the derived allele may be slowed down or potentially prevented (Fig. 2) by its negative recessive effects in BB homozygotes on offspring survival (Fig. 1B, C).

Our results thus confirm theoretical studies that have predicted that genetic polymorphisms are likely maintained via antagonistic pleiotropic effects on multiple life-history traits, imposing life-history trade-offs (Rose 1985; Zajitschek & Connallon 2018). Because chromosome inversions typically link many loci into fixed haplotypes, the probability increases that they capture multiple allelic effects on fitness into highly pleiotropic fitness phenotypes. A similar life-history trade-off presumably maintains a large inversion polymorphism in the seaweed fly: the (presumably) inverted alpha allele (Mérot *et al.* 2021) reduced larvae survival but increased female fecundity and male competitive ability (Mérot *et al.* 2020).

In the case of *Tgu11*, simulations including the observed antagonistic pleiotropic effects, show that the inversion polymorphism is maintained at an average B-allele frequency of 0.64 (excluding the two cases where the B allele was fixed; Fig. 2B and Fig. S4B). This simulated frequency is higher than the observed frequency of 0.526 from the wild (see Supplementary Table S1). The discrepancy suggests that either the detrimental effects on offspring survival (Fig. 1B) were underestimated in our data from a captive population, or that the beneficial effects on reproductive performance (Fig. 1A) were overestimated. Such differences in costs or benefits are not unlikely, because the estimated effect sizes are based on captive populations of zebra finches, where individuals bred with *ad libitum* food and water supply. In harsh natural environments (Zann 1996), the genetic effects of *Tgu11* on fitness-related traits, e.g. nestling survival and individual lifespan, might be stronger, or the positive effects on reproductive performance weaker, e.g. due to more environmental noise. Similarly, in the seaweed fly, no advantage of the heterozygous genotype on offspring developmental time was found when larval density was low (Butlin *et al.* 1984; Mérot *et al.* 2020). Alternatively, but less likely, the fixation process of the derived B allele could still be ongoing. In the wild, the zebra finch has a very large effective population size (Balakrishnan & Edwards 2009). Thus, the inversion polymorphism is likely

more stable than in our simulations with a population size of only 400 individuals (Fig. 2B). Indeed, simulations with 5000 individuals showed that the B allele frequency was already markedly more stable (compare Fig. 2B with Fig. S4B).

The maintenance of large inversion polymorphisms has mostly been studied in systems with large effects on phenotypic traits (e.g. Jay et al., 2018; Küpper et al., 2015; Lamichhaney et al., 2015; Tuttle et al., 2016), or with a clear heterozygote advantage (e.g. Kim et al., 2017; Knief et al., 2017). Examples of inversions showing antagonistic pleiotropic effects on fitness-related traits are less common and mostly stem from insects (Betran *et al.* 1998; Troth *et al.* 2018; Mérot *et al.* 2020). In our study, the estimated standardized effect sizes of the *Tgu11* inversion were quite small (Fig. 1A,B, also see Fig. S3), indicating that large sample sizes are required for their detection. Hence, many segregating inversion polymorphisms might be maintained by small antagonistic effects on fitness-related traits, which may be difficult to quantify (e.g. Knief et al. 2016) or being practically unidentifiable (due to small effect sizes), especially in the wild (Wellenreuther & Bernatchez 2018).

## Conclusion and outlook

Our results highlight a potential mechanism on how a large chromosomal inversion, derived from a single mutational event, could have spread and be maintained in populations over many generations via antagonistic pleiotropic effects on different fitness-related traits. Although the tight physical linkage of a large number of allelic variants at multiple loci makes it virtually impossible to identify the gene(s) responsible for the fitness effects, it is the linkage of several small effects that allows to study these genetic effects on fitness. Inversions are sometimes portrayed as “supergenes” for highly polygenic traits; in analogy, inversions may also bind together a whole range of relatively minor fitness-related phenotypic effects. The sum of these effects of the entire inversion haplotype might be large enough to be detectable given a sufficient sample size. We observed small additive beneficial effects on reproductive success (a supposedly rare phenomenon that makes the initial spread of the derived allele more likely), as well as small recessive deleterious effects on survival, which are thought to be more common (Charlesworth & Willis 2009), because selection against recessive deleterious alleles is ineffective as long as the allele is rare. Our findings thus suggest that large chromosomal inversions may link beneficial and deleterious mutations together that synergistically prevent the inversion from going extinct or from spreading to fixation. Our study calls for future work that estimates the effect of the *Tgu11* inversion on fitness-related traits in the wild.

### **Data and code accessibility**

Reproductive performance and offspring survival data can be found at <https://osf.io/tgsz8/>; all other supporting data and R scripts will be uploaded to the Open Science Framework OSF or upon request.

### **Competing interests**

We have no competing interests.

### **Funding**

This research was supported by the Max Planck Society (to BK). YP was a member of the International Max Planck Research School for Organismal Biology.

### **Acknowledgements**

We thank M. Schneider for molecular work, K. Martin, T. Aronson, E. Bolund, M. Ihle, S. Janker, H. Schielzeth, J. Schreiber, and D. Wang for help with data collection, and S. Bauer, E. Bodendorfer, J. Didsbury, A. Grötsch, A. Kortner, F. Lehmann, P. Neubauer, K. Piehler, F. Weigel and B. Wörle for animal care and help with zebra finch breeding.

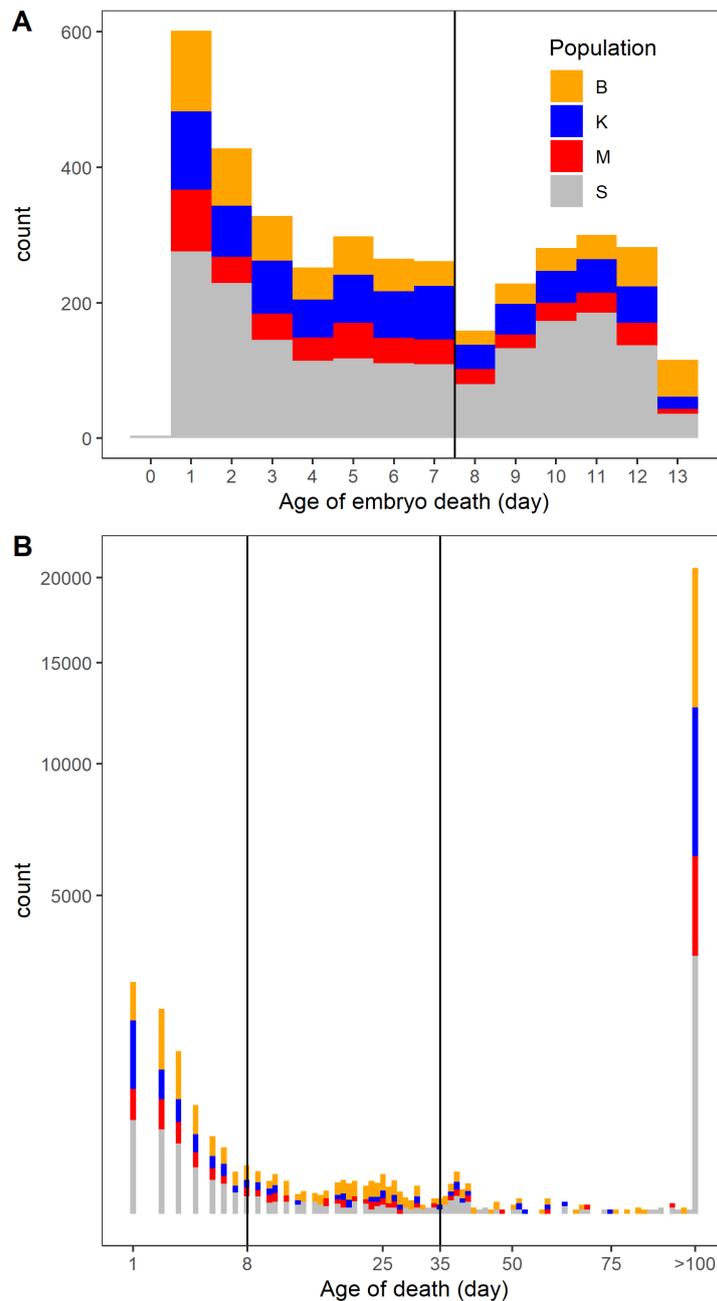
## References

- Balakrishnan, C.N. & Edwards, S. V. (2009). Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra Finch (*Taeniopygia guttata*). *Genetics*, 181, 645–660.
- Betran, E., Santos, M. & Ruiz, A. (1998). Antagonistic Pleiotropic Effect of Second-Chromosome Inversions on Body Size and Early Life-History Traits in *Drosophila buzzatii*. *Evolution (N. Y.)*, 52, 144.
- Butlin, R.K., Collins, P.M. & Day, T.H. (1984). The effect of larval density on an inversion polymorphism in the seaweed fly *Coelopa frigida*. *Heredity (Edinb.)*, 52, 415–423.
- Charlesworth, D. & Willis, J.H. (2009). The genetics of inbreeding depression. *Nat. Rev. Genet.*, 10, 783–796.
- Christidis, L. (1986). *Chromosomal evolution within the family Estrildidae (Aves) I. The Poephilae. Genetica.*
- Dobzhansky, T. & Pavlovsky, O. (1955). AN EXTREME CASE OF HETEROSIS IN A CENTRAL AMERICAN POPULATION OF *DROSOPHILA TROPICALIS*. *Proc. Natl. Acad. Sci.*, 41, 289–295.
- Durmaz, E., Benson, C., Kapun, M., Schmidt, P. & Flatt, T. (2018). An inversion supergene in *Drosophila* underpins latitudinal clines in survival traits. *J. Evol. Biol.*, 31, 1354–1364.
- Hoffmann, A.A. & Rieseberg, L.H. (2008). Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? *Annu. Rev. Ecol. Evol. Syst.*
- Hoffmann, A.A., Sgrò, C.M. & Weeks, A.R. (2004). Chromosomal inversion polymorphisms and adaptation. *Trends Ecol. Evol.*, 19, 482–488.
- Itoh, Y., Kampf, K., Balakrishnan, C.N. & Arnold, A.P. (2011). Karyotypic polymorphism of the zebra finch Z chromosome. *Chromosoma*, 120, 255–264.
- Jay, P., Whibley, A., Frézal, L., Rodríguez de Cara, M.Á., Nowell, R.W., Mallet, J., *et al.* (2018). Supergene Evolution Triggered by the Introgression of a Chromosomal Inversion. *Curr. Biol.*, 28, 1839-1845.e3.
- Kennington, W.J., Hoffmann, A.A. & Partridge, L. (2007). Mapping regions within cosmopolitan inversion In(3R)Payne associated with natural variation in body size in *Drosophila melanogaster*. *Genetics*.
- Kim, K.-W., Bennison, C., Hemmings, N., Brookes, L., Hurley, L.L., Griffith, S.C., *et al.* (2017). A sex-linked supergene controls sperm morphology and swimming speed in a songbird. *Nat. Ecol. Evol.*, 1, 1168–1176.
- Kirkpatrick, M. (2010). How and why chromosome inversions evolve. *PLoS Biol.*, 8.
- Knief, U. & Forstmeier, W. (2018). Violating the normality assumption may be the lesser of two evils. *bioRxiv*.
- Knief, U., Forstmeier, W., Kempnaers, B. & Wolf, J.B.W. (2019). A sex chromosome inversion is associated with copy number variation of mitochondrial DNA in zebra finch sperm. *bioRxiv*.
- Knief, U., Forstmeier, W., Pei, Y., Ihle, M., Wang, D., Martin, K., *et al.* (2017). A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nat. Ecol. Evol.*, 1, 1177–1184.
- Knief, U., Hemmrich-Stanisak, G., Wittig, M., Franke, A., Griffith, S.C., Kempnaers, B., *et al.* (2016). Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biol.*, 17, 199.
- Küpper, C., Stocks, M., Risse, J.E., Dos Remedios, N., Farrell, L.L., McRae, S.B., *et al.* (2015). A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.*, 48, 79–83.
- Lamichhaney, S., Fan, G., Widemo, F., Gunnarsson, U., Thalmann, D.S., Hoepfner, M.P., *et al.* (2015). Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nat. Genet.*, 48, 84–88.
- Lee, Y.W., Fishman, L., Kelly, J.K. & Willis, J.H. (2016). A segregating inversion generates fitness variation in yellow monkeyflower (*Mimulus guttatus*). *Genetics*, 202, 1473–1484.
- Lindtke, D., Lucek, K., Soria-Carrasco, V., Villoutreix, R., Farkas, T.E., Riesch, R., *et al.* (2017). Long-term

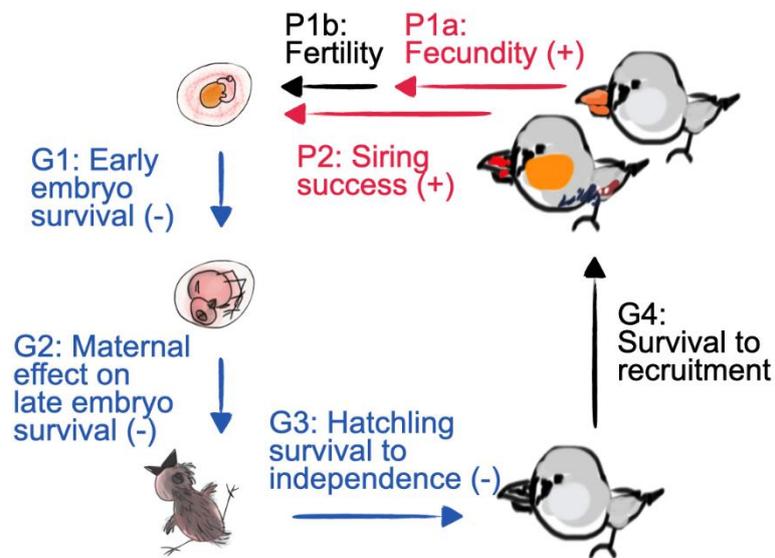
- balancing selection on chromosomal variants associated with crypsis in a stick insect. *Mol. Ecol.*
- Lowry, D.B. & Willis, J.H. (2010). A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.*, 8.
- Mérot, C., Berdan, E.L., Cayuela, H., Djambazian, H., Ferchaud, A.-L., Laporte, M., *et al.* (2021). Locally Adaptive Inversions Modulate Genetic Variation at Different Geographic Scales in a Seaweed Fly. *Mol. Biol. Evol.*
- Mérot, C., Llaurens, V., Normandeau, E., Bernatchez, L. & Wellenreuther, M. (2020). Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. *Nat. Commun.*, 11, 670.
- Pei (裴一凡), Y., Forstmeier, W., Wang (王代平), D., Martin, K., Rutkowska, J. & Kempnaers, B. (2020). Proximate Causes of Infertility and Embryo Mortality in Captive Zebra Finches. *Am. Nat.*, 196, 577–596.
- Pei, Y., Forstmeier, W. & Kempnaers, B. (2020). Offspring performance is well buffered against stress experienced by ancestors. *Evolution (N. Y.)*, 74, 1525–1539.
- Pracana, R., Levantis, I., Martínez-Ruiz, C., Stolle, E., Priyam, A. & Wurm, Y. (2017). Fire ant social chromosomes: Differences in number, sequence and expression of odorant binding proteins. *Evol. Lett.*
- Roberts, P.A. (1967). A positive correlation between crossing over within heterozygous pericentric inversions and reduced egg hatch of *Drosophila* females. *Genetics*.
- Rose, M.R. (1985). Life history evolution with antagonistic pleiotropy and overlapping generations. *Theor. Popul. Biol.*, 28, 342–358.
- Singh, B. & Das, A. (1991). Epistatic interaction between unlinked inversions in Indian natural populations of *Drosophila melanogaster*. *Genet. Sel. Evol.*, 23, 371.
- Troth, A., Puzey, J.R., Kim, R.S., Willis, J.H. & Kelly, J.K. (2018). Selective trade-offs maintain alleles underpinning complex trait variation in plants. *Science (80-. )*, 361, 475–478.
- Tuttle, E.M. (2003). Alternative reproductive strategies in the white-throated sparrow: Behavioral and genetic evidence. *Behav. Ecol.*, 14, 425–432.
- Tuttle, E.M., Bergland, A.O., Korody, M.L., Brewer, M.S., Newhouse, D.J., Minx, P., *et al.* (2016). Divergence and functional degradation of a sex chromosome-like supergene. *Curr. Biol.*, 26, 344–350.
- Wang, D., Forstmeier, W., Martin, K., Wilson, A. & Kempnaers, B. (2020). The role of genetic constraints and social environment in explaining female extra-pair mating. *Evolution (N. Y.)*, 74, 544–558.
- Wang, D., Kempnaers, N., Kempnaers, B. & Forstmeier, W. (2017). Male zebra finches have limited ability to identify high-fecundity females. *Behav. Ecol.*, 28, 784–792.
- Wellenreuther, M. & Bernatchez, L. (2018). Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends Ecol. Evol.*, 33, 427–440.
- Zajitschek, F. & Connallon, T. (2018). Antagonistic pleiotropy in species with separate sexes, and the maintenance of genetic variation in life-history traits and fitness. *Evolution (N. Y.)*, 72, 1306–1316.
- Zann, R. (1996). *The Zebra Finch: A synthesis of field and laboratory Studies*. Oxford University Press, New York, USA.

## Supplementary Materials

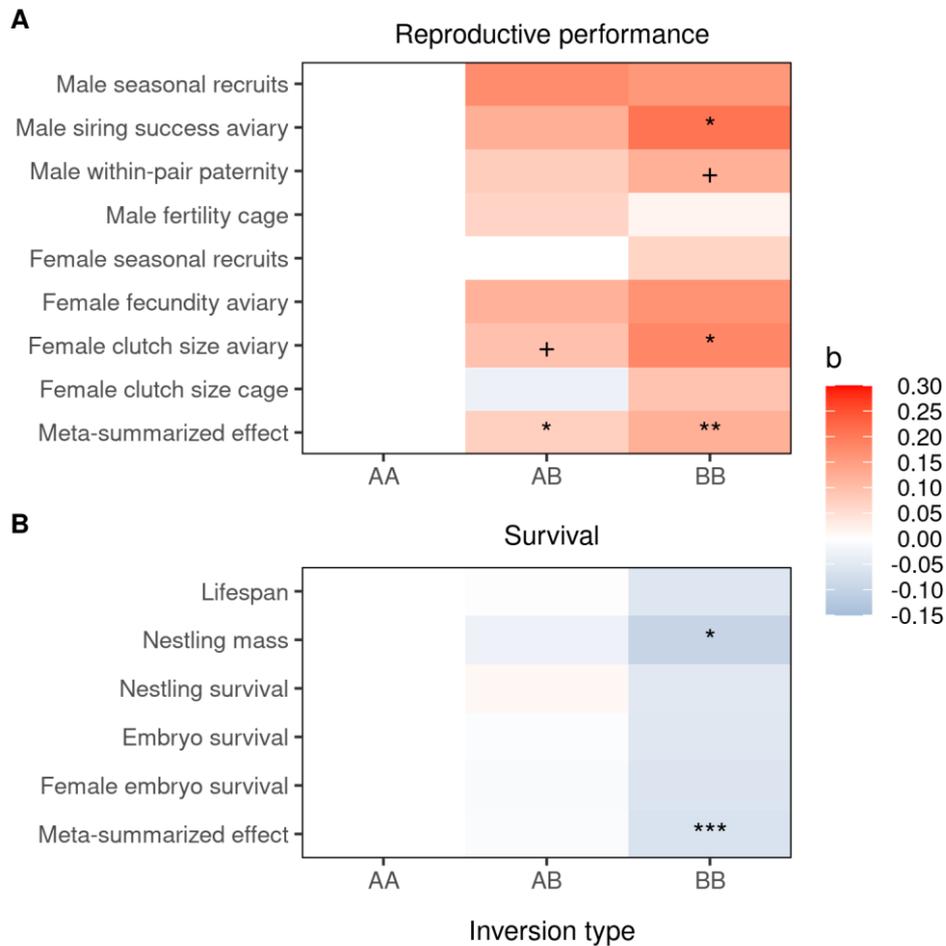
**Figure S1.** Stacked histogram of the age of embryo death (top) and the age of death after hatching (bottom) for populations Bielefeld ('B', orange), Krakow ('K', blue), Melbourne ('M', red), and Seewiesen ('S', grey). Embryo survival is divided into early (0-7 days of incubation, left of black line) and late (8-13 days of incubation) stages. Offspring survival is divided into early (1-8 days), late (9-35 days), and post-independence (35-100 days post hatching) stages. Note that in Fig. S1B the x and y axes were square-root transformed for visualization.



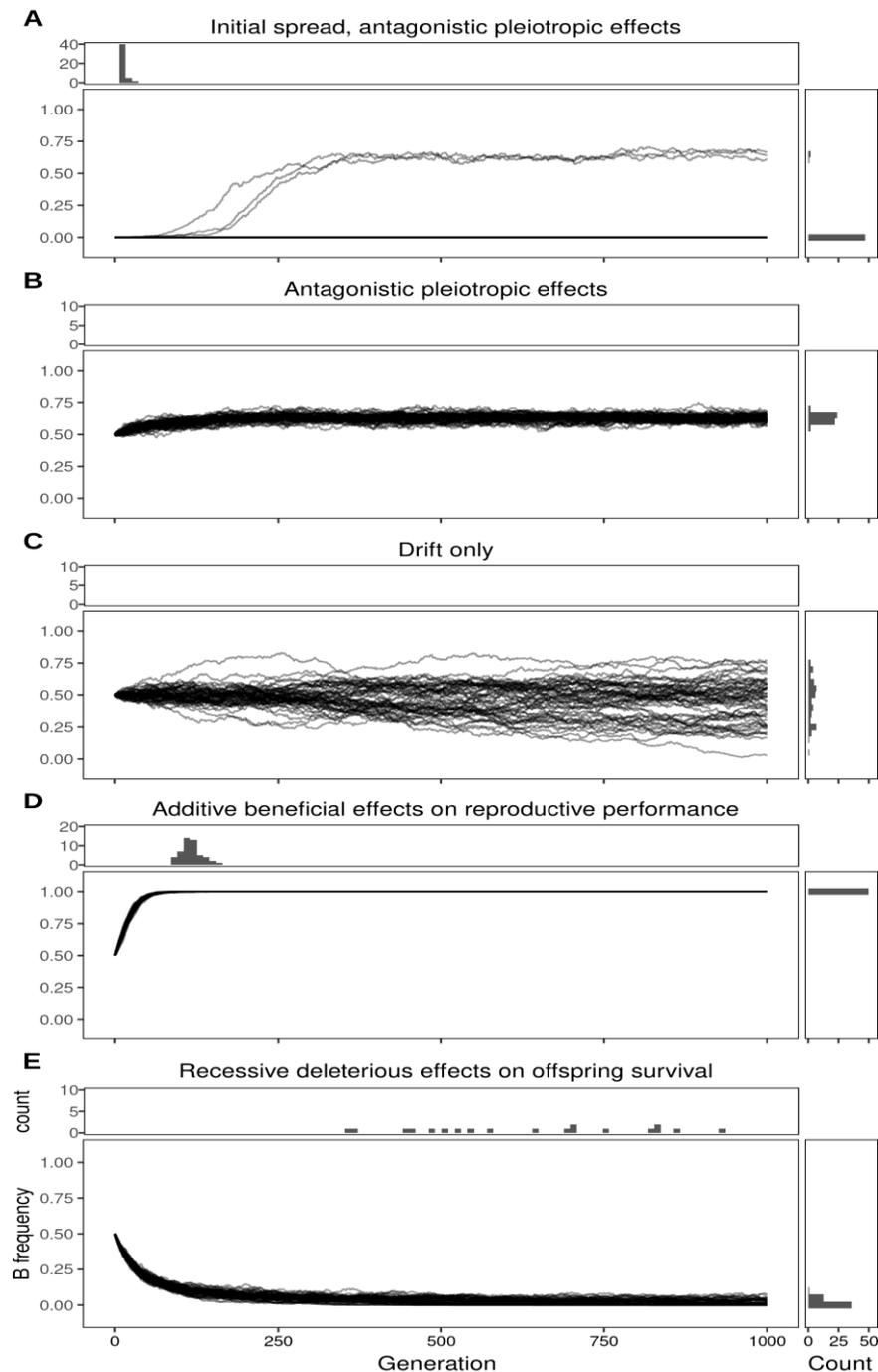
**Figure S2.** Zebra finch life stages influenced by the derived *Tgu11* inversion B-allele. B alleles increase female fecundity (number of eggs, P1a, adjusted by fertility in aviaries, P1b, i.e. the probability of an egg to be fertilized), and male siring success (P2) additively (red), but reduce early embryo survival (G1), maternal effects on late embryo survival (G2), and hatchling survival (G3) in a recessive way (blue). Survival to recruitment (G4) was not influenced by *Tgu11* genotype (black), but we kept the estimates in the simulation for completeness (full generation cycle). See Fig S1 for the details of categorization of early and late stages of embryo death. Note that estimates of female fecundity and male siring success comprise effects of adult survival during the breeding season. Estimates of genotype effects on each phase used for simulations are listed in Table S6.



**Figure S3.** Estimated standardized effect sizes of the Tgu11 inversion genotypes (AB and BB) relative to the homozygous ancestral genotype (AA) representing the baseline phenotype (indicated in white), from models excluding all covariates. Effects on female and male reproductive performance traits (A) and on offspring survival (B), with their meta-summarized effects. Red indicates positive effects on fitness-related traits whereas blue indicates negative effects. For each genotype, ‘+’ indicates  $P < 0.1$ , ‘\*’ indicates  $P < 0.05$ , ‘\*\*’ indicates  $P < 0.001$ , ‘\*\*\*’ indicates  $P < 0.0001$ ; no symbol indicates  $P \geq 0.1$ . See Fig 1 (Table S3) for estimates from full models including covariates.



**Figure S4.** Changes in the simulated frequency of the inverted B allele in a population with 5000 adult birds over 1000 generations using estimated total fitness measures (Table S6; combination of female fecundity corrected for fertility in aviaries, male siring success, and embryo survival until recruitment; Fig. S2). We also simulated random genetic drift by randomly drawing 5000 recruits from the pool of all produced recruits of the previous generation (5000 to 7280 recruits; see methods for details). Shown are five simulation scenarios illustrating the initial spread of allele B (A), the maintenance of allele B via antagonistic pleiotropic effects on both reproductive performance and offspring survival (B), drift only (C), additive positive effects of allele B on reproductive performance (D), and recessive deleterious effects of allele B on offspring survival (E). Histograms at the top of each panel (A–E) show time points of allele fixation; histograms on the right of each panel show B-allele frequency at generation 1000. Note that the B allele was maintained at a frequency of 0.62 (SD = 0.04) via effects of antagonistic pleiotropy in populations with 5000 birds (B; see also Fig. 2 for simulations on populations of 400 birds).



**Table S1.** Number of zebra finches that survived at least to 8 days of age with Tgu11 inversion genotypes AA, AB, and BB, and B-allele frequency in populations Bielefeld, Krakow, Melbourne, Seewiesen, and in the wild (the latter is from Knief et al 2016). Type A is the ancestral type (referred to as type B in Knief et al 2016) whereas type B is the derived inversion type (i.e. type A in Knief et al 2016), judged from analysis of heterozygosity and diversity in A and B inversion types among 948 wild zebra finches (Knief et al 2016).

<b>Population</b>	<b>Type AA</b>	<b>Type AB</b>	<b>Type BB</b>	<b>B allele frequency</b>	<b>N generations maintained at Seewiesen until May 2018</b>
Bielefeld	552	633	128	0.339	7
Krakow	273	585	286	0.506	6
Melbourne	138	265	191	0.545	3
Seewiesen	1034	1682	918	0.484	14
<i>Wild</i>	<i>218</i>	<i>463</i>	<i>267</i>	<i>0.526</i>	-

**Table S2.** Tag SNP of Tgu11 inversion types. Allele G represents the ancestral type A (referred to as type B in Knief et al. 2016) and allele A is linked to the derived inversion type B. Chromosomal position is based on genome built TaeGut1.

SNP Name	Chromosome	Position	Ancestral A	Inverted B
WZF00031805	11	12289339	G	A

**Table S3.** Estimated standardized effect sizes for Tgu11 inversion types AB and BB in comparison to AA (homozygous ancestral) on reproductive performance and survival traits. Nobs refers to the number of observations (e.g. eggs, clutches, individual breeding seasons, or individuals).

This table contains 43 rows.

**Table S4.** Effect sizes of confounding fixed effects in models on reproductive performance and survival traits.

This table contains 341 rows.

**Table S5.** Random effect estimates for all mixed-effects models.

This table contains 95 rows.

Tables S3-S5 will be uploaded with the publication

**Table S6.** Estimated trait values for Tgu11 inversion types AA, AB and BB used in the simulation of the evolution of B-allele frequency. Details of life cycle stages are described in Fig S2. Female fecundity and male siring success (i.e. reproductive performance traits) followed a zero-inflated Poisson distribution, therefore we estimated the proportion of zeros and their Poisson mean (i.e. mean excluding zeros) from the raw data, only using the Experiment "SelectionLines S3" that had the most observations (Pei et al. 2020b). For simplicity in simulation, a single survival rate from embryo to recruitment was calculated (G1-4). This was calculated as the product of early embryo survival as a function of embryo genotype (G1), late embryo survival as a function of the mother's genotype (G2), hatchling survival (G3), and survival to recruitment (G4) as a function of offspring genotype (i.e. embryo to recruitment survival rate), for each combination of female and offspring genotype. All embryo to recruitment survival rates were estimated from mixed-effect models using the raw trait values as response variable while controlling for all possible fixed and random effects.

Trait	Life cycle stage (in Fig S2)	Genotype	Genotype of	Trait type	Estimate
Female fecundity aviary proportion of zeros	P1a	AA	Female	Reproductive performance	0.179
Female fecundity aviary proportion of zeros	P1a	AB	Female	Reproductive performance	0.085
Female fecundity aviary proportion of zeros	P1a	BB	Female	Reproductive performance	0.083
Female fecundity aviary mean without zeros	P1a	AA	Female	Reproductive performance	10.663
Female fecundity aviary mean without zeros	P1a	AB	Female	Reproductive performance	10.081
Female fecundity aviary mean without zeros	P1a	BB	Female	Reproductive performance	10.339
Female fertility aviary	P1b	AA	Female	Reproductive performance	0.834
Female fertility aviary	P1b	AB	Female	Reproductive performance	0.826
Female fertility aviary	P1b	BB	Female	Reproductive performance	0.815
Male siring success aviary proportion of zeros	P2	AA	Male	Reproductive performance	0.163
Male siring success aviary proportion of zeros	P2	AB	Male	Reproductive performance	0.181
Male siring success aviary proportion of zeros	P2	BB	Male	Reproductive performance	0.106
Male siring success aviary mean without zeros	P2	AA	Male	Reproductive performance	8.832
Male siring success aviary mean without zeros	P2	AB	Male	Reproductive performance	9.234

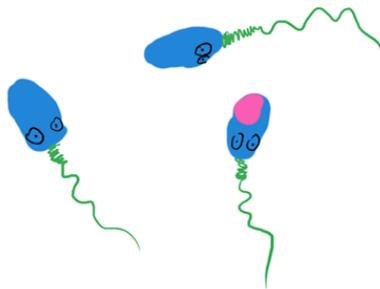
Male siring success aviary mean without zeros	P2	BB	Male	Reproductive performance	9.570
Embryo to recruitment survival rate	G1-4	AA_AA	Female and embryo	Survival (calculated as the product of estimated survival rates in G1-G4)	0.346
Embryo to recruitment survival rate	G1-4	AA_AB	Female and embryo	Survival (calculated as the product of estimated survival rates in G1-G4)	0.348
Embryo to recruitment survival rate	G1-4	AB_AA	Female and embryo	Survival (calculated as the product of estimated survival rates in G1-G4)	0.346
Embryo to recruitment survival rate	G1-4	AB_AB	Female and embryo	Survival (calculated as the product of estimated survival rates in G1-G4)	0.348
Embryo to recruitment survival rate	G1-4	AB_BB	Female and embryo	Survival (calculated as the product of estimated survival rates in G1-G4)	0.322
Embryo to recruitment survival rate	G1-4	BB_AB	Female and embryo	Survival (calculated as the product of estimated survival rates in G1-G4)	0.339
Embryo to recruitment survival rate	G1-4	BB_BB	Female and embryo	Survival (calculated as the product of estimated survival rates in G1-G4)	0.314
Early embryo survival	G1	AA	Embryo	Detailed embryo to recruitment survival rate	0.747
Early embryo survival	G1	AB	Embryo	Detailed embryo to recruitment survival rate	0.744
Early embryo survival	G1	BB	Embryo	Detailed embryo to recruitment survival rate	0.718
Maternal late embryo survival	G2	AA	Female	Detailed embryo to recruitment survival rate	0.835
Maternal late embryo survival	G2	AB	Female	Detailed embryo to recruitment survival rate	0.835
Maternal late embryo survival	G2	BB	Female	Detailed embryo to recruitment survival rate	0.812
Nestling survival	G3	AA	Nestling	Detailed embryo to recruitment survival rate	0.568
Nestling survival	G3	AB	Nestling	Detailed embryo to recruitment survival rate	0.575
Nestling survival	G3	BB	Nestling	Detailed embryo to recruitment survival rate	0.549
Fledgling survival	G4	AA	Fledgling	Detailed embryo to recruitment survival rate	0.977
Fledgling survival	G4	AB	Fledgling	Detailed embryo to recruitment survival rate	0.975
Fledgling survival	G4	BB	Fledgling	Detailed embryo to recruitment survival rate	0.979



## Chapter 5

### Germline-restricted chromosome

"Hey, look, that sperm has one extra chromosome that we don't have!"



## Chapter 5

### Occasional paternal inheritance of the germline restricted chromosome

Yifan Pei, Wolfgang Forstmeier, Francisco J. Ruiz-Ruano, Jakob C. Mueller, Josefa Cabrero, Juan Pedro M. Camacho, Juan D. Alché, Andre Franke, Marc Hoepfner, Stefan Börno, Ivana Gessara, Moritz Hertel, Kim Teltscher, Ulrich Knief, Alexander Suh, Bart Kempenaers

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 can also be paternally inherited. Confocal microscopy using GRC-specific FISH probes indicated that a considerable fraction of sperm heads (1-19%) 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, analysing GRC haplotypes across nine *castanotis* matriline, estimated to have diverged for up to 250,000 years, 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 event 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 outcompeting other GRC haplotypes that were limited to strict maternal inheritance, even if this was partly detrimental to organismal fitness.

Prepared as:

Pei, Y., Forstmeier, W., Ruiz-Ruano, F. J., Mueller, J. C., Cabrero, J., Camacho, J. P. M., ... & Kempenaers, B. (2021). Occasional paternal inheritance of the germline-restricted chromosome in songbirds. *bioRxiv*.

## 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 ability of paternal inheritance differs between families. We further show that the genetic diversity of the GRC is extremely reduced compared to the highly divergent mtDNA, suggesting that a single GRC haplotype has spread through the Australian zebra finch population relatively recently, via the additional route of paternal inheritance. Our study therefore suggests that the GRC is prone to evolve in a selfish manner that could result in intra-genomic conflict.

## Main

In sexually reproducing eukaryotes, the stable inheritance of the nuclear DNA typically requires 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 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 some genetic elements successfully manipulate their mode of inheritance to their own advantage (so called 'selfish genetic elements' (10–12)). Hence, as a key step towards understanding the evolution and the function of the GRC, it appears essential to fully understand how the GRC is inherited.

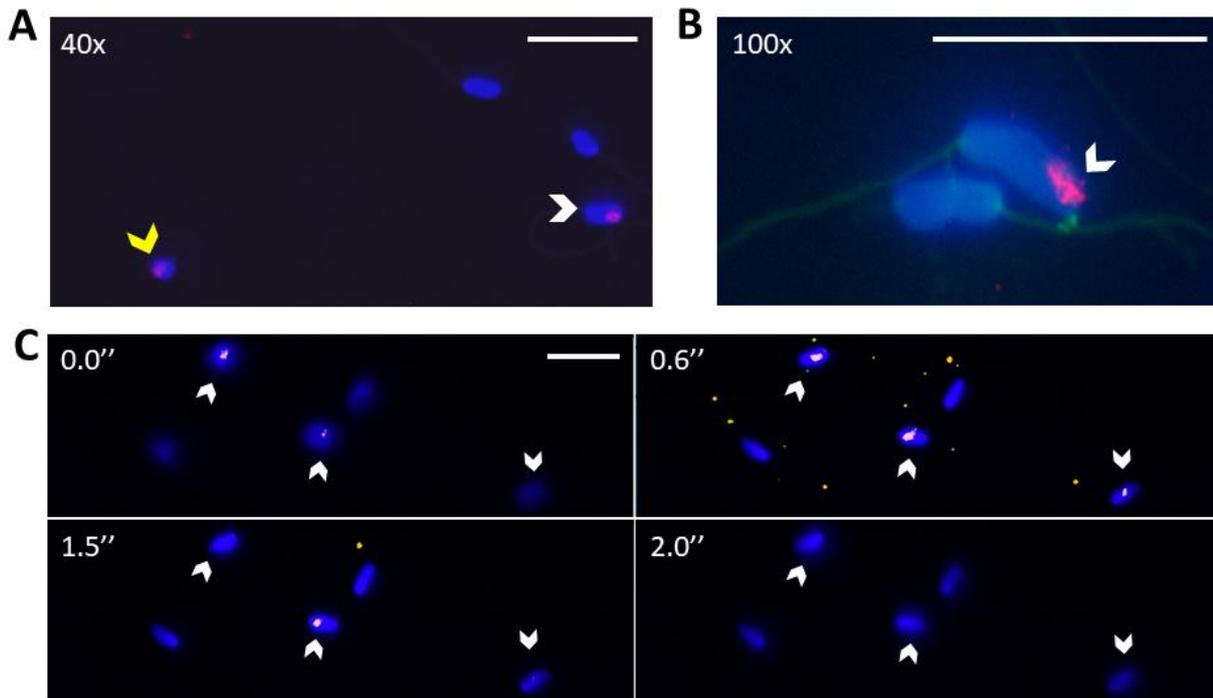
In most studies to date, the GRC was observed in two copies in female (primary) oocytes (5, 7, 9, 13), and as a single chromosome in male spermatogonia (1, 4–6, 9, 13). Cytogenetic investigations in males (predominantly *T. g. castanotis*) found that this single-copy GRC is eliminated from nuclei during 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. Yet to date, the mode of inheritance of the GRC has not been tested with genetic markers, implying that suggestions about the evolutionary significance of the GRC (2, 4, 9) remain speculative. In this study, we used the two zebra finch subspecies *T. g. castanotis* of Australia

(hereafter *castanotis*) and *T. g. guttata* of the Indonesian islands such as Timor (hereafter *guttata*) to address the issue of inheritance. Specifically, we combined cytogenetic and genomic data to study (a) the elimination efficiency of the GRC during spermatogenesis, (b) the strictness of the proposed matrilineal inheritance (i.e. expected co-inheritance with the mitochondrial genome, ‘mtDNA’) and (c) the genetic variation of the GRC and the co-evolutionary 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: (i) expelled free-floating GRC micronuclei (1, 15), (ii) small, digested DNA fragments (14, 15) and (iii) non-expelled 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 7 natural ejaculate samples from 5 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: 1-19%) and only a few free-floating micronuclei (mean = 1 micronucleus per hundred sperm heads, range: 0-2; **Fig. 1; SI Appendix, Table 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; SI Appendix, Supplementary video**). These results show that, in the zebra finch, the GRC is not completely eliminated during spermatogenesis, and imply that a non-negligible number of spermatozoa can potentially transmit the GRC.

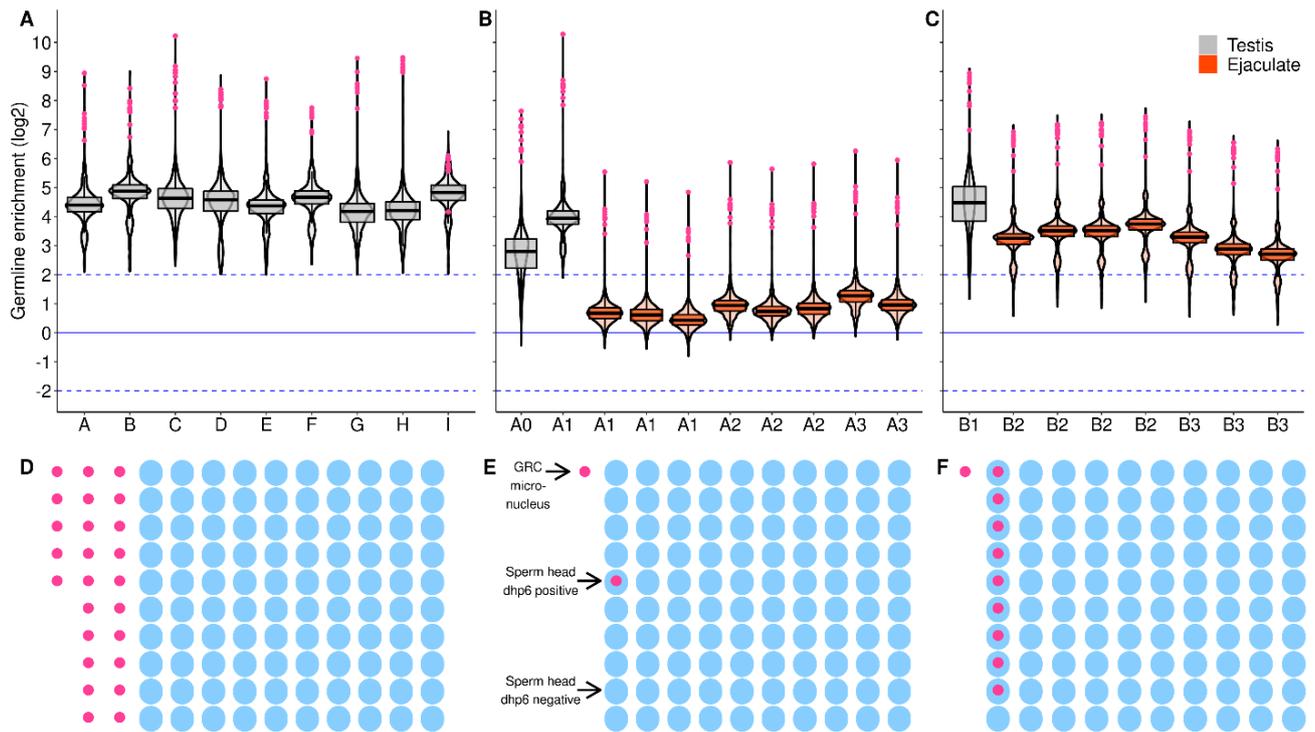


**Fig 1. Cytogenetic evidence for the presence of the GRC in the nucleus of zebra finch *Taeniopygia guttata castanotis* sperm.** The GRC-amplified probe *dph6* (see Methods) indicates the presence of the GRC (red) inside some sperm heads (white arrow in (A-C)) and in free-floating micronuclei (yellow arrow in (A)). Blue DAPI stain without red indicates sperm heads without GRC. Green autofluorescence shows the sperm flagellum (in (B)). (A) 40X magnification. (B) 100X 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 the **SI Appendix, video**. The video consisted of 24 sections representing a total of 6.0  $\mu\text{m}$  depth. Scale bars are 20  $\mu\text{m}$ .

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. To do this, we used Illumina PCR-free libraries of pairs of germline and its corresponding soma samples, and quantified the GRC content as coverage enrichment in the germline samples (as  $\log_2$  germline-to-soma coverage ratio; for list of all samples and library pairs see **SI Appendix, Table S2** and for details see **Materials and Methods**). We first defined a set of high-copy-number sequences on the GRC from nine *castanotis* testes from nine different matriline (i.e. mtDNA haplotypes), and found 1,742 windows of 1kb length with a cut-off of  $\log_2$  testis-to-soma coverage ratio larger than 2 (**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. S1, S2**). In these selected windows, we then examined the GRC content of 15 ejaculates from five additional *castanotis* males stemming from two families (i.e. matriline A and B), including one male per family for which we had obtained cytogenetic data (as presented above; ejaculates in **Fig. 2B,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,C**).

Based on median coverage ratios for each sample, ejaculates from the same individual male zebra finch were remarkably repeatable in their GRC content (**Fig. 2B,C**;  $R_{\text{male}} = 0.98$ ,  $P < 0.001$ ; **SI Appendix, Table S3**). The majority of the repeatable variation was due to a between-matriline difference ( $R_{\text{matriline}} = 0.96$ ,  $P = 0.012$ ; **SI Appendix, Table S3**). Ejaculates from males of matriline B had significantly higher amounts of GRC (**Fig. 2B**) than those from males of matriline A (**Fig. 2C**;  $b_{\text{matrilineB}} = 2.4$ ,  $SE = 0.28$ ,  $P < 0.001$ ; see also **SI Appendix, Figs. S3C,E, S4** and **Table S3**), qualitatively confirming the results from the cytogenetic analysis (1% versus 9% of sperm heads were GRC-positive, respectively; **Fig. 2D,E**).

These results indicate that certain GRC haplotypes (e.g. those in matriline B) are more likely transmitted via sperm than others (e.g. those in matriline A), and hence more likely to potentially spread in a ‘selfish’ manner. The high individual repeatability and consistency within matriline 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 the GRC haplotype itself (including epigenetic marks such as histone or DNA modifications) or due to epistatic interactions between the GRC and the A-chromosomal (i.e. autosomal and sex-chromosomal) genome.

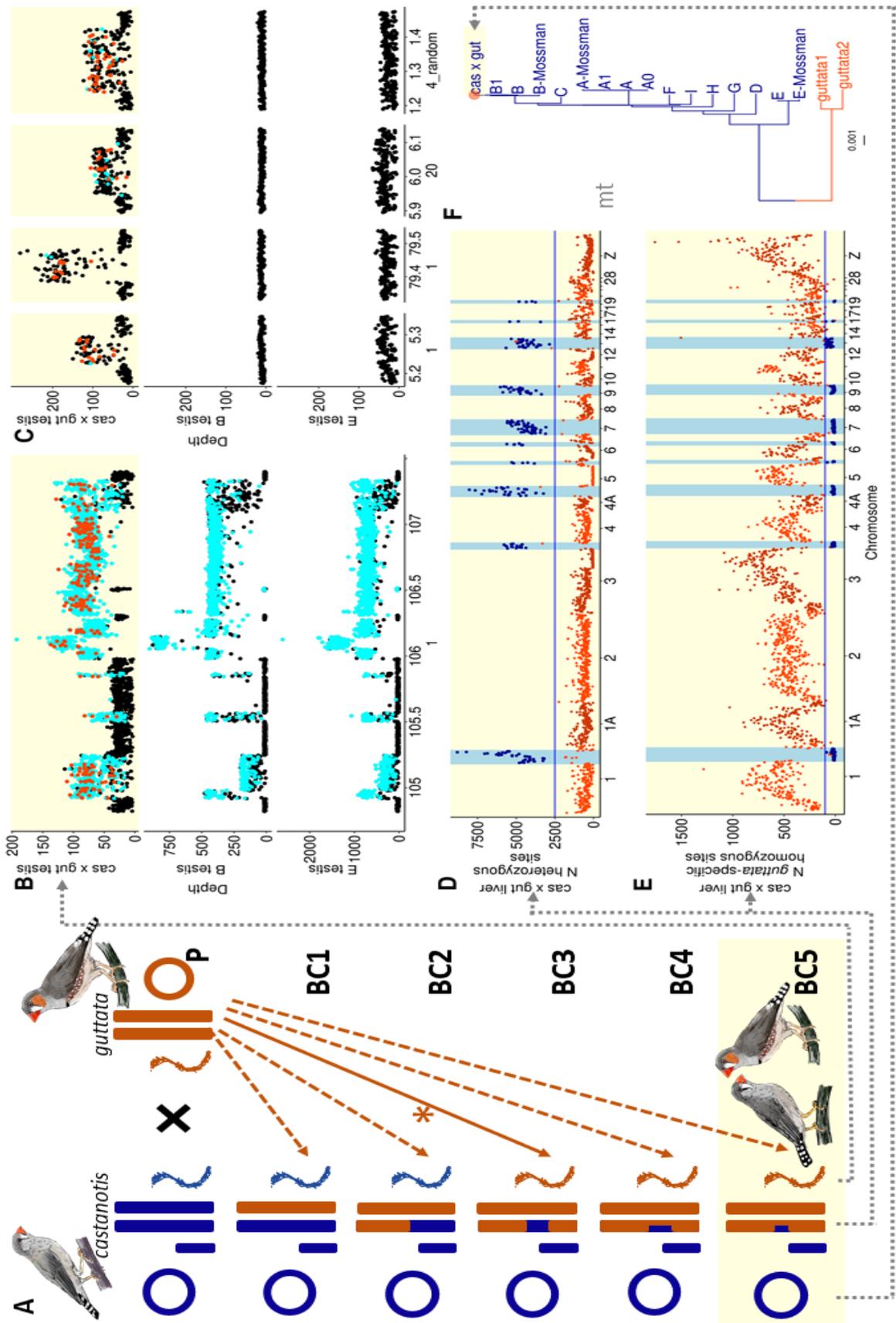


**Fig 2. The elimination efficiency of the GRC differs between *castanotis* matriline A and B.** (A-C) Comparison of sequencing coverage of GRC-containing (testis, indicated in grey, and ejaculate, orange) and GRC-free tissue (liver) identifies sequences that are GRC-linked in high copy number (4). Male IDs are shown on the X-axis. The solid blue line refers to a  $\log_2$  germline-to-soma coverage ratio = 0, i.e. no germline enrichment; the dashed blue line refers to a 4-fold 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 4-fold ( $\log_2 > 2$ ) 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-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:  $b = -2.56$ ,  $SE = 0.29$ ,  $P < 0.001$ ; **SI Appendix, Table S3**), as expected from previous work (14, 15). (C) Coverage ratios of the selected windows in matriline B (three brothers B1-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:  $b = -1.21$ ,  $SE = 0.39$ ,  $p = 0.02$ ; **SI Appendix, Table S3**). (D-F) Illustration of the expected (D) and observed (E, 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, see 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 micro-nuclei 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; also see **SI Appendix, Table S1**).

### Paternally inherited GRC in a hybrid individual

Given the observed occurrence of the GRC in spermatozoa, we tested for paternal inheritance of the typically maternally-transmitted GRC (1, 7, 9, 13). First, we identified polymorphic markers that reliably distinguish different GRC haplotypes. Then, we looked for cases in which the maternally transmitted mtDNA and the GRC markers were not co-inherited (i.e. informative paternal transmission of the GRC). This search rests on the assumption of strict maternal inheritance of mtDNA haplotypes, which is highly likely because (a) paternal inheritance of the mtDNA is known to be rare (16) and was usually detected in individuals with mitochondrial diseases and heteroplasmy (17, 18), and (b) the avian W chromosome and mtDNA are in high linkage disequilibrium (16). We used data on all 12 *castanotis* males stemming from nine matriline from which testes and soma samples had been sequenced (in **Fig. 2A-C**), and we additionally sequenced one pair of testis and liver (**SI Appendix, Table S2**) of one *castanotis* x *guttata* hybrid descendant male with *guttata* phenotype but *castanotis*-introgressed genotype (i.e. *castanotis* matriline B and 5% of *castanotis* introgressed tracts on an otherwise *guttata* A-chromosomal genome), using 10X linked-read technology (**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 novel GRC haplotype that uniquely diverged from all the 12 GRC haplotypes examined so far (**Fig. 3B,C**; **SI Appendix, Fig. S5**). 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 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. S5**), 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. S5**; including 13 genes with GRC paralogs that were not found previously (4), see **SI Appendix, Table S4**; also see **SI Appendix, Table S5**). These results suggest that the GRC was inherited from at least one of the potential *guttata* males during back-crossing (**Fig. 3A**). Note that the *castanotis* x *guttata* testis sample had extremely low coverage on the GRC, presumably due to underdeveloped testes (**Materials and Methods**), and that no individuals from a previous generation could be retrieved (also see **SI Appendix, Behind-the-paper**). We therefore assume that the *guttata* GRC coexisted with, recombined with or replaced the *castanotis* GRC, and stably co-inherited with the *castanotis* mitochondria thereafter (**Fig. 3**).



**Fig 3. A case of paternal inheritance of the GRC in a captive-bred hybrid *castanotis* x *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 (53) 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-C) Raw read depth of selected genome regions of testis libraries (mapped to somatic-reference taeGut1 (36)) for the *castanotis* x *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  $\geq 3$  germline-specific SNPs. Each orange-brown dot shows a 1-kb window that contains private testis-specific SNPs. Note that only the hybrid has individual testis-specific SNPs (B-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. S5** for genome-wide plots and more individuals. (D-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* x *guttata* hybrid compared to the pooled DNA from 100 wild-caught *castanotis* zebra finches (43). The horizontal blue line indicates the cut-off of 2500 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 (51) and SRR3208120 (52) respectively; orange-brown). The scale bar shows the number of substitutions per site.

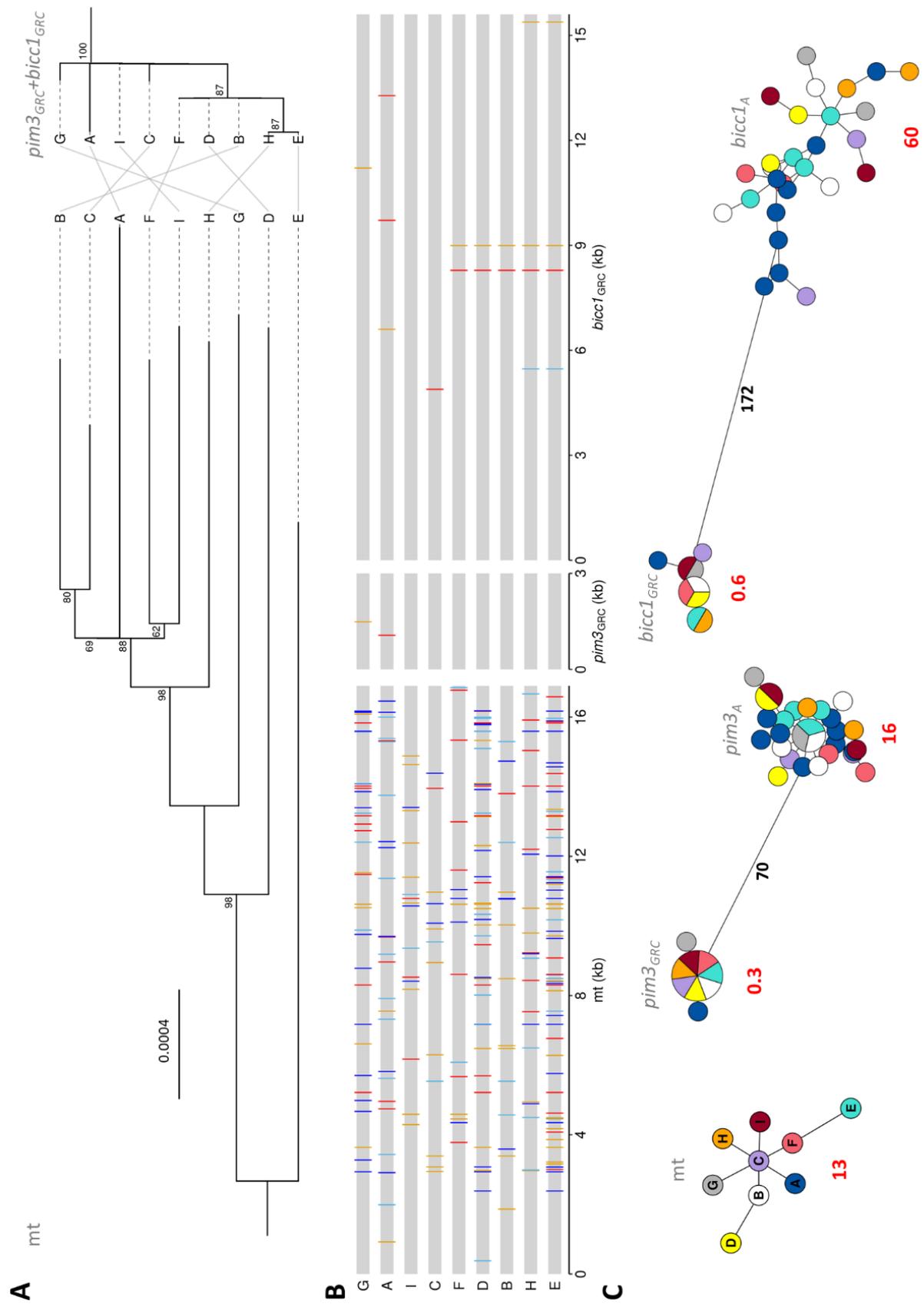
## Discordant evolutionary histories between the mtDNA and the associated GRC haplotypes

We found surprisingly little variation among the 12 *castanotis* GRCs from nine matriline, some of which were old matriline as judged from mtDNA divergence (**Fig. 4; SI Appendix, Fig. S1**). In sharp contrast to the overall highly diverged *guttata* GRC, which harboured hundreds of private testis-specific SNPs (top row in **Fig. 3B-C, SI Appendix, Fig. S5A and Table S5**), we did not find a single high-confidence GRC-specific SNP that was private to one of the 12 *castanotis* individuals (middle and bottom rows in **Fig. 3B-C and SI Appendix, Fig. S5B-M**). The lack of detection of such SNPs may partly be due to the sensitivity of the subtractive method (each high-confidence allele required 3 reads support, see **Materials and Methods**) and to the highly diverged A-chromosomal paralogs (i.e. alleles that are shared between GRC haplotypes and the soma will be missed). Note, that in our full *castanotis* dataset of 12 GRCs, most of the males that share the same mtDNA are brothers that can be expected to also share the same GRC-haplotype (only one male, B, is not closely related; **SI Appendix, Table S2**). Indeed, for the two single-copy GRC genes, the *castanotis* males that have the same mtDNA haplotype also have the same GRC haplotype (3 males of matriline A, and 4 males of matriline B; see **SI Appendix, Fig. S6**).

Therefore, we studied the within-population variation of the GRC (in *castanotis*) from the nine males (A-I) representing nine independent matriline (**SI Appendix, Figs. S1,S2**). We compared the evolutionary histories between the GRC haplotypes and the associated mitogenomes, by focusing on two single-copy GRC genes, *pim3<sub>GRC</sub>* and *biccl<sub>GRC</sub>* (4), that were highly diverged from their A-chromosomal paralogs (for consensus sequence construction and additional details see **Materials and Methods and SI Appendix, Results**). The GRC paralogs of *pim3* and *biccl* were clearly in single copy, because all GRC mappings were homozygous (also see **SI Appendix, Fig. S6A,B**) and showed 39% read coverage compared to the genomic background of the corresponding sample (**SI Appendix, Fig. S7A,B**). 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).

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 GRC genes *pim3<sub>GRC</sub>* and *biccl<sub>GRC</sub>*), we found no positive correlation between the two pairwise-distance-matrices

(**Fig. 4A**;  $r = -0.03$ ,  $SE = 0.06$ ,  $P=0.66$ ; **SI Appendix, Table S6**). Strikingly, multiple males from clearly diverged mtDNA lineages shared the same GRC haplotype (**Fig. 4**, **SI Appendix, Figs. S6 and S8**). For instance, we estimated that the mtDNA haplotypes E and H diverged for about 250,000 years, based on 0.5% substitutions per site, and assuming a molecular clock with 2% substitutions per site per million years (19). However, males E and H carried the same GRC haplotype (characterized by two diagnostic derived mutations; **Fig. 4B**). Hence, we observed considerable genetic diversity in mtDNA (13 SNPs/kb; **Fig. 4C**; **SI Appendix, Fig. S8A and Table S7**), in striking contrast to the two single-copy GRC-linked genes showing hardly any genetic diversity (*pim3<sub>GRC</sub>*: 0.3 SNPs/kb, *biccl1<sub>GRC</sub>*: 0.6 SNPs/kb; **Fig. 4C**; **SI Appendix, Fig. S8B,C**; also see **SI Appendix, Results and Fig. S8D** for the double-copy GRC gene *elavel4<sub>GRC</sub>* haplotype 1: 0.1 SNPs/kb and haplotype 2: 1.7 SNPs/kb). This lack of diversity in the GRC even across deeply diverged matriline suggests that older GRC haplotypes were probably driven to extinction by a more recent type that had the ability to cross 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.



**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 9 haplotypes (A-I) of the *castanotis* mitogenome (mt, left) and the 9 associated GRC haplotypes (concatenated sequence of the two single-copy genes *pim3<sub>GRC</sub>* and *bicc1<sub>GRC</sub>*; right). The mtDNA tree was rooted by the mitogenome assemblies of two published *guttata* datasets (SRA accession numbers SRR2299402 (51) and SRR3208120 (52) (not shown here, see Fig. 3F), whereas the tree of GRC haplotypes was rooted by the concatenated sequences of the two A-chromosomal paralogs (not shown here, see (C) and SI Appendix, Fig. S10). Node support bootstrap values are shown if >60 (based on 1000 bootstraps). The scale bar indicates 0.004 substitutions per site. (B) The alignments of the mitogenomes and the two single-copy genes on the GRC (4) (*pim3<sub>GRC</sub>* 3,286 bp and *bicc1<sub>GRC</sub>* 13,822 bp) from the nine individuals used in (A). Grey indicates consensus among the 9 haplotypes. Orange, red, blue and dark blue indicate a mutation towards 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 (mt, 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 (SI Appendix, Table 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 kilobase for each cluster of haplotypes; black numbers refer to the number of mutations per kilobase 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 *elav1/4<sub>GRC</sub>*. Note the highly reduced genetic diversity (short branch length in (A) and little variation in (B,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 (also see SI Appendix, Table S7).

## 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, Figs. S6, S7**), suggesting that one GRC is removed (either always the same one, i.e. an efficient driver, or with a 50:50 probability). In either case, the GRC that escapes 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), which sets the stage for selfish evolution of DNA. The fact that we see low diversity on the *castanotis* GRCs even across highly diverged matrilineages (**Fig. 4**), suggests that such a selfish GRC may have spread relatively recently, while the discordance in tree topologies suggests that the crossing of matriline boundaries is still ongoing. The highly diverged *guttata* GRC (**Fig. 3**) argues against the alternative explanation that the GRC has an unusually low mutation rate, which is also contradicted by the observed dynamic GRC evolution in general (4, 5). Hence, the lack of GRC variation in *castanotis* is more suggestive of a recent spreading event.

Selfish evolution of the GRC might be widespread across songbirds, because we find evidence for paternal inheritance in both zebra finch subspecies (**Figs. 1-2** and **4** for *castanotis* GRC, and **Fig. 3** for *guttata* GRC), and because of the observed similarity of meiotic behaviour (i.e. elimination during male spermatogenesis and recombination in female oocytes) of the GRC across all songbirds examined to date (6, 9, 13). Interestingly, Malinovskaya et al. (9) reported a GRC copy number mosaicism in spermatogonia and pachytene spermatocytes in males of the pale martin (*Riparia diluta*). We hypothesize that the spermatogonia and spermatocytes that possess an extra copy of the GRC might result in GRC-carrying sperm. The GRC-carrying sperm thereby brings 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).

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 evolutionary 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 ability to spread in both zebra finch subspecies (**Fig. 2** and **Fig. 3**). After a certain GRC haplotype has successfully spread to the entire population (going to fixation), its ability to spread through this second (i.e. paternal transmission) route may lose its adaptive value, because there is no alternative haplotype to compete with. 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 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 a highly dynamic nature of the content of the GRC. Across species, there is excessive variation in 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 of the GRC in sexual reproduction. The genetic diversity of the two examined single-copy genes on the GRC is 56-fold lower than the diversity of their A-chromosomal paralogs (**Fig. 4; SI Appendix, Table S7**; also see **SI Appendix, Results, Figs. S8, S9** and (23)). Other non-recombining sex-specific chromosomes also show highly reduced genetic diversity compared to their autosomes, possibly due to strong sexual selection (e.g. the human Y-chromosome shows a 23-41-fold reduction in the mean number of pairwise differences per site, i.e.  $\pi$  (24)), or due to Hill-Robertson interference with the mitogenome (e.g. the avian W chromosome shows a 46-104-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 non-recombining 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 co-inherited. With paternal spill-over 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 – such that this new combination can rise in frequency. 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 the **SI Appendix, Table S2** (also see **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 (#18 in (28)). Using a dummy female, we collected natural ejaculates from 8 brothers from two families (mtDNA haplotype A,  $N_{\text{ejaculates}} = 9$ , from males A1-A3; mtDNA haplotype B,  $N_{\text{ejaculates}} = 11$ , from males B2-B6). Males were 100-1036 days 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 the **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* x *guttata* hybrid (see **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-A3, and males B1-B6 were brothers; **SI Appendix, Figs. S1,S2**; see **SI Appendix, Materials and Methods** for mtDNA sequencing and assembly). Note that ejaculates of males B4-B6 were only sampled for cytogenetic analysis (see **SI Appendix, Table S1** and **Behind-the-paper**). Matriline A,

B and E were first described in (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' #19 in (28) and 'Melbourne' (30)) and three domesticated populations ('Krakow' #11 in (28), 'Seewiesen' and 'Spain' (4)). The sampled individuals were 238-1882 days old (**SI Appendix, Table S2**). All 11 captive *castanotis* zebra finch testes were large (longest diameter: 3-5 mm) compared to the testis of the *castanotis* x *guttata* zebra finch (longest diameter: ~ 1mm), suggesting that the *castanotis* x *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. S3F,S5A,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 **SI Appendix, Tables S1, S2**), following a protocol modified from (31). In brief, we collected fresh ejaculate in 10  $\mu$ l of PBS with a 20  $\mu$ l pipette, and osmotic-shocked the sample by adding 250  $\mu$ l (~ 20-fold volume of the sample) of 1% sodium-citrate solution for 20 minutes. 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 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-ACGTCTTTGCCTGACCCTTTCAGA, R- TGCATAGAGTTCTCCATCAGACAGACA, taken from (4), and then labelled it with tetramethylrhodamine-5-dUTP via nick translation (32). FISH was then performed on the ejaculate preparations. The hybridization mix consisted of 12  $\mu$ l formamide, 6  $\mu$ l dextran sulfate, 1.5  $\mu$ l 20 $\times$ SSC, 0.5  $\mu$ l salmon sperm, 0.5  $\mu$ l SDS, 3  $\mu$ l *dph6* probe, and 6.5  $\mu$ 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-UV (405 nm) and green (550 nm) sequential excitation, obtained with a pE-4000 (CooLED) 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 8 fields (2 rows x 4 columns) from each FISH preparation where there were more than 20 non-overlapping 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 40x magnification. We took photos with a Zeiss Axiocam 512 color camera using ZEN blue 3.1 software. We then counted the total number of spermatocytes, the number of GRC-carrying spermatocytes and the number of expelled free-floating GRC micronuclei for all sampled fields. Additional images were generated under Leica DM 6000/HX PC APO 100x-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 **SI Appendix, Table S2**. Note that the raw sequencing reads of *castanotis* males A (SR00100), B (Spain\_1) and F (Spain\_2) were taken from (4) (SRA accession numbers SRX6431677-SRX6431681, SRX6431686 and SRX6431688-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 x 150 bp, paired-end) at the Institute of Clinical Molecular Biology (IKMB) 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 of 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 co-evolutionary history of the GRC and its associated mtDNA, we sequenced (PCR-free) DNA from 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-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 Kinsella et al. (4)) from one captive *guttatta* x *castanotis* hybrid with *guttata* phenotype (see **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 **SI Appendix, Table S2**) as paired-end ( $2 \times 150$  bp) using the Illumina NovaSeq 6000 S4 platform at SciLifeLab Stockholm.

### *Raw reads processing*

Raw reads from all WGS 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 adaptor 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 analysed coverage and called SNPs (see below) for downstream analysis. For a detailed pipeline and the corresponding scripts see **Code 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 Kinsella et al. (4). As GRC-linked sequences are often difficult to distinguish from their ancestral A-chromosomal paralogs (as the latter vary substantially between individuals), GRC content is most easily quantified by focussing on sequences that reside on the GRC in numerous copies (compared to just two A-chromosomal copies) (4). To do this, we calculated 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 non-overlapping adjacent windows of 1 kb 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 germline and its corresponding soma library and  $\log_2$ -transformed the values, after correcting for variation in library size, i.e. by dividing the coverage-per-window by the total number of base pairs sequenced for that library.

To quantify testis enrichment in coverage (compared to soma), we first removed windows with too low coverage, i.e. those where both soma and germline samples had <3-fold coverage. Second, we calculated mean and standard deviation (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 quantification of copy number enrichment in testis difficult. Third, we centred the  $\log_2$ -transformed germline-to-soma coverage-ratios of the high-quality windows on the median of the above-selected 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-I, see Samples and **Fig. 2A**). Then, we calculated the median  $\log_2$  ejaculate-to-soma coverage ratios of the selected windows for the 15 ejaculate samples and the 3 testis samples from seven additional males (A0-A3 and B1-B3 in **Fig. 2B,C**) for statistical analysis.

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; see **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 above selected windows) for each germline sample as 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 **SI Appendix, Table S3**.

### *SNP analysis*

We used the SAMtools v1.6 (39) mpileup and bcftools v1.9 call (42) tools to call SNPs, and a customized R script to filter for high-confidence SNPs of interest (see **Code accessibility**), as follows. To study the overall between-individual variation in GRC haplotypes, we called SNPs for all

testis/soma pairs (one hybrid *castanotis* x *guttata* and 12 *castanotis* males) simultaneously and selected high-confidence sites (see **SI Appendix, Material 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 that were identified above, we focused on regions that contained multiple private testis-specific SNPs, defined as 10-kb non-overlapping adjacent windows with at least three such private SNPs (**SI Appendix, Fig. S5 and Fig. 3B,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,C**, the four blood samples of their parents and one pair of testis and liver of one uncle of family A (A0 in **Fig. 2B**). We filtered for high-quality, germline-specific alleles, following the same procedure as described above. We then identified the 1-kb non-overlapping adjacent windows that contained at least 15 germline-specific SNPs (**SI Appendix, Fig. 4**).

To study the extent of A-chromosomal introgression of *castanotis* DNA into the captive population of *castanotis* x *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 non-overlapping 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 non-introgressed 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

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 ( $N > 50$ ) and high density ( $> 3$  per kb) of germline-specific SNPs, and identified five such genes (*pim3*, *elavl4*, *bicc1*, *trim71* and *cpeb1*). The genes *trim71* and *cpeb1* were dropped because we were unable to assemble a GRC contig that was longer than two kilobases (see below) and gene *elavl4* was dropped because it turned out to have two copies on the GRC (see **SI Appendix, Results** and **Fig. S7C,D**). For the remaining two genes, we first generated a consensus for each GRC paralog, and then created haplotypes for each sample, as described below.

To generate a consensus of each GRC paralog, first, we used SAMtools v1.6 (39) view for each germline and soma library to subset reads that mapped on the A-chromosomal paralog (i.e. *taeGut1* (36)) of the focal gene. Second, we used MEGAHIT v1.2.6 (44) to assemble those mapped reads into contigs *de novo*. We then used MAFFT v7.429 (45) to align the assembled contigs that were longer than 2 kb. To build a consensus of the GRC paralogs, we manually selected the contigs that were absent from all soma libraries but present in more than half of the germline libraries, whereby we generated the consensus using the most common allele (for sites that contained SNPs). The consensus of *bicc1<sub>GRC</sub>* and *pim3<sub>GRC</sub>* were constructed from a single *de novo* assembled contig.

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 of that gene. 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 SAMtools v1.6 (39) mpileup to call SNPs for each library of those mapped reads. For the two single-copy GRC genes, *bicc1<sub>GRC</sub>* and *pim3<sub>GRC</sub>* (**Fig. 4** and **SI Appendix, Figs. S6-S9**), 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 **Code accessibility**). Unfortunately, the coverage of GRC-linked reads was too low for the 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.

We used the A-chromosomal paralogs as the outgroup for the GRC genes. Using read aware phasing of SHAPEIT v2.r904 (46), we constructed the two A-chromosomal haplotypes for each somatic library for the GRC genes (**Fig. 4**; see **Code accessibility**).

To compare the genetic diversity between different GRC genes and their A-chromosomal paralogs, we first used MAFFT v7.429 (45) to align the above-constructed haplotypes of GRC and A-chromosomal paralogs for each gene. We then used BMGE v1.12 (47) to trim positions with > 20% “missingness” (i.e. gaps between GRC and A-chromosomal paralogs) in each alignment (see **Code accessibility**). Then, for phylogenetic analysis, we concatenated the two alignments (i.e. *bicc1* and *pim3*) to represent GRC haplotypes. Finally, we used DnaSP v6.12.01 (48) to calculate the mean number of pairwise differences per site ( $\pi$ ) for each GRC gene and each A-chromosomal paralog in the gap-free alignments, and used this value as an estimate of genetic diversity. For each set of genes (GRC and A-chromosomal paralogs), we analyzed and plotted a haplotype network using the ‘haplotype’ function from the R package pegas v0.14 (49) (see **Code accessibility**).

### *Phylogenetic analysis*

Chapter 1 All phylogenetic trees were built using RAxML-NG v1.0.2 (50) assuming a general time-reversible model and discrete GAMMA model of rate heterogeneity with 100 randomized parsimony starting trees and 1000 bootstrap replicates (for details see **Code accessibility**). To demonstrate that the matriline of the hybrid *castanotis x 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**).

Chapter 2 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 (51) and SRR3208120 (52)), whereas the GRC haplotype tree was rooted by 28 constructed A-chromosomal paralogs from all somatic libraries (**Fig. 4A**; for a haplotype network analysis see **SI Appendix, Fig. S8**).

Chapter 3 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 ID as two factors to control for overall position of each sample in the matrices

(SI Appendix, Table S9). We standardized the response variable and the covariate to account for the drastic difference in units between the two types of sequences.

#### *Data accessibility*

All NGS data have been deposited in the Sequence Read Archive (accession number PRJNA741250). All alignments have been deposited in figshare (<https://figshare.com/s/d4a1cfa1c6c126fc15b2>). Additional supporting data have been deposited in the Open Science Framework <https://osf.io/n9x2g/>.

#### *Code accessibility*

All supporting pipelines and scripts have been deposited in the Open Science Framework <https://osf.io/n9x2g/>.

### **Funding**

This research was supported by the Max Planck Society (to B.K.), the Swedish Research Council Formas (2017-01597 and 2020-04436 to A.S.), and the Swedish Research Council Vetenskapsrådet (2016-05139 to A.S.). Y.P. was part of the International Max Planck Research School for Organismal Biology. F.J.R.R. was supported by a postdoctoral fellowship from Sven och Lilly Lawskis fond and a Marie Curie Individual Fellowship (875732).

### **Acknowledgements**

We thank Melanie Schneider and Christine Baumgartner for support with molecular work, Martin Irestedt for help with DNA extractions for the 10X samples, Shouwen Ma for discussion on microscopic image processing, Keren Sadanandan for discussion on tanglegram analysis, and Katrin Martin, Isabel Schmelcher, Claudia Scheicher, Sonja Bauer, Edith Bodendorfer, Jane Didsbury, Annemarie Grötsch, Andrea Kortner, Petra Neubauer, Frances Weigel and Barbara Wörle for animal care and help with breeding zebra finches. We thank Leo Joseph and the Australian National Wildlife Collection for providing testis and liver samples from a wild *T. g. castanotis* individual. We thank Leo Joseph, Julie Blommaert, Octavio Palacios, Simone Fouché and three anonymous reviewers for comments on the manuscript. Some of the computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) through the Uppsala Multidisciplinary Center

for Advanced Computational Science (UPPMAX). The authors acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory, the Knut and Alice Wallenberg Foundation and the Swedish Research Council.

## References

1. 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*. *Chromosom. Res.* **6**, 105–113 (1998).
2. B. Hansson, On the origin and evolution of germline chromosomes in songbirds. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 11570–11572 (2019).
3. J. J. Smith, V. A. Timoshevskiy, C. Saraceno, Programmed DNA Elimination in Vertebrates. *Annu. Rev. Anim. Biosci.* **9**, 1–29 (2021).
4. C. M. Kinsella, *et al.*, Programmed DNA elimination of germline development genes in songbirds. *Nat. Commun.* **10**, 5468 (2019).
5. A. A. Torgasheva, *et al.*, Germline-restricted chromosome (GRC) is widespread among songbirds. *Proc. Natl. Acad. Sci.* **116**, 201817373 (2019).
6. 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).
7. 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).
8. 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).
9. 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**, 1–10 (2020).
10. J. H. Werren, Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 10863–10870 (2011).
11. Austin, B. U. R. T., R. Trivers, A. Burt, *Genes in conflict: the biology of selfish genetic elements* (Harvard University Press, 2009).
12. L. D. Hurst, A. Atlan, B. O. Bengtsson, Genetic conflicts. *Q. Rev. Biol.* **71**, 317–364 (1996).
13. 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).
14. 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).
15. 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).
16. L. Smeds, *et al.*, Evolutionary analysis of the female-specific avian W chromosome. *Nat. Commun.* **6**, 7330 (2015).
17. S. Luo, *et al.*, Biparental inheritance of mitochondrial DNA in humans. *Proc. Natl. Acad. Sci.* **115**, 13039–13044 (2018).
18. 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).
19. A. Arcones, R. Ponti, D. R. Vieites, Mitochondrial substitution rates estimation for divergence time analyses in modern birds based on full mitochondrial genomes. *Ibis (Lond. 1859)*, ibi.12965 (2021).
20. H. Ellegren, Evolutionary stasis: the stable chromosomes of birds. *Trends Ecol. Evol.* (2010) <https://doi.org/10.1016/j.tree.2009.12.004>.
21. A. Kapusta, A. Suh, Evolution of bird genomes—a transposon’s-eye view. *Ann. N. Y. Acad. Sci.* **1389**, 164–185 (2017).
22. T. M. Degrandi, *et al.*, Introducing the bird chromosome database: an overview of cytogenetic studies

- in birds. *Cytogenet. Genome Res.* **160**, 199–205 (2020).
23. 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).
  24. M. A. Wilson Sayres, K. E. Lohmueller, R. Nielsen, Natural selection reduced diversity on human Y chromosomes. *PLoS Genet.* **10**, e1004064 (2014).
  25. S. Berlin, H. Ellegren, Chicken W: A genetically uniform chromosome in a highly variable genome. *Proc. Natl. Acad. Sci.* **101**, 15967–15969 (2004).
  26. W. G. Hill, A. Robertson, The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).
  27. J. M. Comeron, A. Williford, R. M. Kliman, The Hill-Robertson effect: Evolutionary consequences of weak selection and linkage in finite populations. *Heredity (Edinb.)* **100**, 19–31 (2008).
  28. 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).
  29. 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).
  30. S. Jerónimo, *et al.*, Plumage color manipulation has no effect on social dominance or fitness in zebra finches. *Behav. Ecol.* **29**, 459–467 (2018).
  31. R. Meredith, A simple method for preparing meiotic chromosomes from mammalian testis. *Chromosoma* **26**, 254–258 (1969).
  32. J. Camacho, J. Cabrero, M. López-León, D. Cabral-de-Mello, F. Ruiz-Ruano, “Grasshoppers (Orthoptera)” in *Protocols for Cytogenetic Mapping of Arthropod Genomes*, Igor V. Sharakhov, Ed. (CRC Press, 2014), pp. 381–438.
  33. J. Schindelin, *et al.*, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* (2012) <https://doi.org/10.1038/nmeth.2019>.
  34. V. der A. GA, *et al.*, From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Curr. Protoc. Bioinforma.* (2013).
  35. B. Bushnell, BBMap. <https://sourceforge.net/projects/bbmap/> (2015).
  36. W. C. Warren, *et al.*, The genome of a songbird. *Nature* **464**, 757–762 (2010).
  37. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv Prepr. arXiv* (2013).
  38. Broad Institute, Picard Tools. *Broad Institute, GitHub Repos.* (2019).
  39. H. Li, *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* (2009) <https://doi.org/10.1093/bioinformatics/btp352>.
  40. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
  41. R Core Team, R: A Language and Environment for Statistical Computing (2020).
  42. 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).
  43. U. Knief, *et al.*, Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biol.* **17**, 199 (2016).
  44. D. Li, C. M. Liu, R. Luo, K. Sadakane, T. W. Lam, MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* (2015) <https://doi.org/10.1093/bioinformatics/btv033>.
  45. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in

- performance and usability. *Mol. Biol. Evol.* (2013) <https://doi.org/10.1093/molbev/mst010>.
46. O. Delaneau, B. Howie, A. J. Cox, J. F. Zagury, J. Marchini, Haplotype estimation using sequencing reads. *Am. J. Hum. Genet.* (2013) <https://doi.org/10.1016/j.ajhg.2013.09.002>.
  47. A. Criscuolo, S. Gribaldo, BMGE (Block Mapping and Gathering with Entropy): A new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* (2010) <https://doi.org/10.1186/1471-2148-10-210>.
  48. J. Rozas, *et al.*, DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* (2017) <https://doi.org/10.1093/molbev/msx248>.
  49. E. Paradis, Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* (2010) <https://doi.org/10.1093/bioinformatics/btp696>.
  50. 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).
  51. S. Singhal, *et al.*, Stable recombination hotspots in birds. *Science* (80-. ). **350**, 928–932 (2015).
  52. J. H. Davidson, C. N. Balakrishnan, Gene Regulatory Evolution During Speciation in a Songbird. *Genes/Genomes/Genetics* **6**, 1357–1364 (2016).
  53. U. Knief, W. Forstmeier, Y. Pei, J. Wolf, B. Kempnaers, A test for meiotic drive in hybrids between Australian and Timor zebra finches. *Ecol. Evol.*, ece3.6951 (2020).

## Supplementary Materials

### Supplementary text

#### Materials and Methods

##### *Ejaculate collection*

We collected ejaculates using a method modified from (1). The focal male was placed in a 240 cm long cage, i.e. a row of four connected single cages of 60 cm × 40 cm × 45 cm (length × width × height). Initially, the cage was divided into two equally-sized compartments by a wooden divider. In one side of the cage, we kept the focal male with a female, and provided them with two nest boxes and coconut fibers as nesting material. In the other side, we kept two “stimulus” males. Focal males typically start nest building and courting the female within a day or two after entering the cage, and after 7 days the nest would typically be built. Seven days after the initial setup, we replaced the wooden divider by a metal mesh and we also separated the focal male from his female using another metal mesh divider. In this way, the focal male could stay in visual and vocal contact both with his female and with the two males.

From this point on, every other day for two weeks, we placed a dummy female mounted in a copulation solicitation posture in the focal male’s compartment. The female’s dummy cloaca contained a piece of fresh silicon tube filled with PBS buffer.

Within 3-5 minutes after introducing the dummy female, the focal male would either show no response, or briefly court the dummy female and then mount it, making contact with the false cloaca.

Immediately after a copulation with the dummy female, an observer quickly separated the male from the dummy female using a non-transparent divider, gently removed the dummy female, and checked for the presence of an ejaculate in or near the false cloaca. An ejaculate was visible as a white or beige coloured viscous cloud in the PBS buffer, or as a dot of brown or a small patch of white liquid either on the hard part of the body around the false cloaca or on the feathers. Each ejaculate that was detected in the false cloaca dissolved in PBS was collected using a pipet (20 µl). Each ejaculate that was found outside the cloaca was collected using a pipet (20 µl) that contained 5-10 µl PBS.

##### *DNA extraction from ejaculates, testis and liver*

Except for the samples used for 10X Chromium library preparation (which were extracted with a KingFisher Duo robot using the KingFisher Cell and Tissue DNA Kit following the manufacturer's recommendation; see also Kinsella et al. (2)), we extracted DNA from each ejaculate, testis and liver sample using two modified versions of the phenol-chloroform method described in (3).

#### A. Protocol for the ejaculate samples

1. 160 µl TNE buffer  
20 µl SDS  
10 µl Proteinase K  
10 µl 1% dithiothreitol (DTT)  
Incubate at 56°C with gentle shaking overnight.
2. Add 200 µl phenol, gently invert 15 times, centrifuge at 8000 g for 10 minutes, pipette off the upper aqueous layer and transfer to a new tube (discard the rest).
3. Add 200 µl phenol plus 100 µl chloroform, gently invert 15 times, centrifuge at 8000 g for 10 minutes, pipette off the upper aqueous layer and transfer to a new tube (discard the rest).

4. Add 200  $\mu$ l phenol/chloroform/isoamyl alcohol (25:24:1), gently invert tubes 15 times. Centrifuge at 8000 g for 10 minutes, pipette off the upper aqueous layer and transfer to a new tube (discard the rest).
5. Add 200  $\mu$ l chloroform/isoamyl alcohol (24:1) to the supernatant, gently invert tubes 15 times. Centrifuge at 8000 g for 10 minutes, pipette off the upper aqueous layer and transfer to a new tube (discard the rest).
6. Add 200  $\mu$ l isopropanol plus 20  $\mu$ l 3M NaOAc (4°C) to the supernatant, gently invert tubes 15 times, incubate at -80°C for one hour. Thaw briefly at room temperature then centrifuge at 13000 g for 30 minutes at 4°C. Carefully remove the supernatant.
7. Wash the pellet with 800  $\mu$ l -20°C 100% ethanol. Centrifuge at 13000 g for 30 minutes at 4°C. Carefully remove the supernatant so as not to disturb the pellet, leave to dry approximately 15 minutes at room temperature.
8. Resolve DNA pellet in 50  $\mu$ l distilled water.

#### B. Protocol for the testis and liver samples.

1. 485  $\mu$ l TNE buffer  
20  $\mu$ l SDS  
15  $\mu$ l Proteinase K  
Incubate at 56°C with gentle shaking overnight.
2. Add 450  $\mu$ l phenol/chloroform/isoamyl alcohol (25:24:1), gently invert tubes 15 times. Centrifuge at 8000 g for 10 minutes. Pipette off the upper aqueous layer and transfer to a new tube (discard the rest).
3. Repeat step 2 twice.
4. Add 450  $\mu$ l chloroform/isoamyl alcohol (24:1) to the supernatant, gently invert tubes 15 times. Centrifuge at 8000 g for 10 minutes. Pipette off the upper aqueous layer and transfer to a new tube (discard the rest).
5. Add 45  $\mu$ l 3M NaOAc plus 500  $\mu$ l cold isopropanol (-20 °C) to the supernatant. Incubate at -80°C for at least one hour. Thaw briefly at room temperature then centrifuge at 13000 g for 30 minutes at 4°C. Carefully remove the supernatant.
6. Wash the pellet with 1 ml -20°C 100% ethanol plus 10  $\mu$ l NaOAc. Centrifuge at 13000 g for 30 minutes at 4°C. Carefully remove the supernatant.
7. Wash the pellet with 600  $\mu$ l -20°C 70% ethanol. Centrifuge at 13000 g for 15 minutes at 4°C. Carefully remove the supernatant so as not to disturb the pellet, leave to dry approximately 15 minutes at room temperature.

8. Resolve DNA pellet in 70  $\mu$ l distilled water.

#### *Amplicon-sequencing of mtDNA and mitogenome analysis*

We used the following approach to identify the matriline in the four captive populations of the Australian zebra finches *Taeniopygia guttata castanotis*, i.e. Seewiesen, Krakow, Bielefeld and Melbourne, and in the captive population of *T. g. castanotis* x *T. g. guttata* hybrids, with Timor zebra finch *Taeniopygia guttata guttata* phenotype (thereafter *castanotis* x *guttata*). First, for the founder females of each population (or from one of her offspring; N = 141), we extracted DNA from muscle tissue (or embryonic tissue) using the DNeasy Blood & Tissue Kit (Qiagen), or from blood using the NucleoSpin Blood QuickPure Kit (Macherey-Nagel) following the manufacturers' instructions. Second, we amplified the DNA using four overlapping long-range PCRs that cover the entire zebra finch mitochondrial genome (4) (for primer sequences and additional details of PCR conditions see **SI Appendix, Table S8**). Third, we pooled equal volumes of the four PCR products for each individual, ran a small aliquot of this pooled product on an agarose gel, and visually checked for bands with the expected size (indicating that at least one of the long-range PCRs worked). Fourth, pooled PCR products were cleaned using 70% ethanol and Agencourt AMPure XP beads (Beckman Coulter) following the manufacturer's instruction. Fifth, cleaned PCR products were multiplexed by individual identity and pool-sequenced using HiSeq4000 for Illumina paired-end short reads (length 75 bp) sequencing at the Sequencing Core Facility, Max Planck Institute for Molecular Genetics in Berlin, Germany.

To investigate the phylogenetic relationships of the mitogenomes between our captive matriline, wild Australian zebra finches *T. g. castanotis* and related species, we additionally sequenced and assembled the mitogenomes from 55 wild-caught zebra finches *T. g. castanotis* (that were longer than 13 kb); one diamond finch *Stagonopleura guttata*, three long-tailed finches *Poephila acuticauda* (two *P. a. hecki* and one *P. a. acuticauda*), one painted finch *Emblema pictum*, one double-banded finch *Taeniopygia bichenovii*, one masked finch *Poephila personata* (kindly provided by Frank Rößler); two Gouldian finches *Erythrura gouldiae* (one black-headed and one yellow-headed; tissue samples kindly provided by Dr. Sue Anne Zollinger); one red-browed finch *Neochmia temporalis*, one plum-headed finch *Neochmia modesta* and one crimson finch *Neochmia phaeton* (kindly provided by the Stuttgart Zoo), as described above. All DNA was extracted from blood samples unless otherwise stated.

To increase the sample size, we included published whole-genome sequencing data from 19 wild *T. g. castanotis* (five of the 19 samples were dropped due to low amount of mitochondrial genome-linked reads), 20 wild long-tailed finches (10 *P. a. hecki* and 10 *P. a. acuticauda*) and one double-banded finch *T. bichenovii* (5). Additionally, we included data from two published pure *T. g. guttata* samples (SRA accession numbers SRR2299402 (5) and SRR3208120 (6)).

We also assembled the mitogenomes from the whole-genome sequencing data for *T. g. castanotis* males (A-I, A0-A3 and B1-B3) and the *castanotis* x *guttata* hybrid (male cas x gut) to confirm their matriline (i.e. for the GRC-individuals used in this study; **Fig. 3F**, see also Materials and Methods).

The whole-genome sequencing data generated from blood samples were low in amount of mtDNA per cell (hence contained a high fraction of the nuclear copy of the mitochondria sequences, i.e. NUMT (7)), whereas other samples (liver, muscle, testis and ejaculate) were strongly enriched for mtDNA. Therefore, we first assembled *de novo* a NUMT haplotype from one sample using MEGAHIT v1.2.6 (8), then mapped reads of all samples against both the true mitochondrial genome (4) and our assembled NUMT haplotype to filter the NUMT sequences from true mtDNA. Then true mtDNA-derived reads were selected and *de novo* assembled using MEGAHIT v1.2.6 (8) using the default settings. For each sample, we selected the longest assembled contig, and aligned it to the reference

using minimap2 (9). We then visually checked the aligned contigs and grouped them into haplotypes. The mtDNA haplotypes of all wild-caught *T. g. castanotis* were distinct and the mtDNA haplotypes of the captive populations of *T. g. castanotis* and the *castanotis* x *guttata* hybrid were grouped into 16 matriline. We used mitogenome assemblies from the two pure *T. g. guttata* as an outgroup for *T. g. castanotis*.

Then, we built two phylogenetic trees using the assembled mtDNA described above (>13 kb) using RAxML-NG v1.0.2 (10) with 100 tree searches and 1000 bootstrap replicates: one tree for all GRC-individuals (*T. g. castanotis* and the *castanotis* x *guttata* hybrid), the 55 wild *T. g. castanotis*, the two true *T. g. guttata*, all other related species and three published mitochondria haplotypes that were also found in our founders (11) (N = 120; **SI Appendix, Fig. S1**), and one tree for the captive *T. g. castanotis* founders, the two true *T. g. guttata* and five published mitochondrial genomes(11) (N = 29; **SI Appendix, Fig. S2**).

### Genotyping of a W-chromosome specific sequence

To identify a W-chromosome-specific marker that can distinguish the two matriline A and B, we used the W-specific intronic sequence CHD1Wi16 (12) as a reference. First, we identified one SNP that distinguishes W-chromosome haplotypes by examining the mapped reads from the sequenced blood samples of the two domesticated *T. g. castanotis* zebra finch females with mitochondria haplotypes A and B (i.e. SR12333 and SR15062 in **SI Appendix, Table S2**). Second, we designed the primer sequences F- GTAAGAATTTTGCTAGTAATAGTCAAG and R-GAGATTGAATGATACAGTTAAAAGG to amplify a short W-specific sequence containing the informative SNP. We then amplified the W-specific sequence for all founder females of the *castanotis* x *guttata* hybrids (N = 2), using DNA extracted from blood samples. The PCR products were Sanger-sequenced and compared to the haplotypes of the two *T. g. castanotis* females. The two *castanotis* x *guttata* hybrid females had the same haplotype as the *T. g. castanotis* female zebra finch with mitochondrial haplotype B.

### Haplotype network analysis

To estimate the genetic diversity among GRC haplotypes, we analyzed the haplotype networks on two single-copy GRC genes (*pim3<sub>GRC</sub>* and *bicc1<sub>GRC</sub>*) and one double-copy GRC gene (*elavl4<sub>GRC</sub>*) and their associated mtDNA haplotypes.

We first constructed the consensus sequence of *elavl4<sub>GRC</sub>* and phased the two A-chromosomal haplotypes (*elavl4<sub>A</sub>*) for each somatic library. Then, we used the haplotype analysis methods described in the main text (see **Materials and Methods**). The consensus of the *elavl4<sub>GRC</sub>* consisted of five *de novo* assembled contigs that were scaffolded with multi-N nucleotides in between. We artificially phased the reference allele of the consensus sequence into haplotype 1 and the alternative allele into haplotype 2.

To maximize the sample size, we constructed a GRC haplotype for each of the testis samples of the 12 *castanotis* males, and for the pooled ejaculate samples of the 2 males from matriline B (males B2 and B3 in **Fig. 2C** in the main text).

To compare the genetic diversity between different GRC genes and their A-chromosomal paralogs, we aligned and then gap-trimmed the above-constructed haplotypes of GRC and A-chromosomal paralogs for each gene (for details see **Materials and Methods**). Finally, we used DnaSP v6.12.01(13) to calculate the mean number of pairwise differences per site ( $\pi$ ) for each GRC gene and each A-chromosomal paralog in the gap-free alignments, and used this value as an estimate of genetic diversity.

For each set of genes (GRC and A-chromosomal paralogs), we analyzed and plotted a haplotype network using the ‘haplotype’ function from the R package *pegas* v0.14 (14) (see **Code accessibility**).

## Results

### *Detection of a castanotis x guttata hybrid*

Analysis of the mtDNA tree of all zebra finches *T. g. castanotis* (thereafter *castanotis*) and the putative *T. g. guttata* (thereafter *guttata*) included in this study (N = 55 wild, N = 13 captive individuals) as well as published sequences (5) (N = 15 wild individuals) showed that our captive population of the subspecies *guttata* (“Timor finch”; based on one male and two female founders for which we sequenced the mtDNA) did not cluster with the highly diverged published sequences of *guttata* (5), but rather showed a mtDNA sequence identical to the *castanotis* haplotype B, which is abundant among domesticated birds from Europe but rare in the wild (**Fig. 3F**; see also **SI Appendix, Figs. S1, S2**). This suggests that our population of putative *guttata* birds (bought from a local amateur breeder) was not recently wild-derived pure *guttata* birds, but had been domesticated locally by crossing *guttata* males with an already domesticated *castanotis* female founder. We confirmed this putative scenario by genotyping a W-chromosome specific sequence, which is exclusively maternally inherited, and showing that our putative hybrids indeed carried the same sequence as found in the domesticated *castanotis* matriline B (**SI Appendix, Materials and Methods**).

To quantify the degree of A-chromosomal introgression of *castanotis* genes into this population of a *guttata* phenotype, we used whole-genome sequencing data from one founder male of the alleged *guttata* population and compared it with a pool of 100 wild *castanotis* zebra finch genomes (15). We identified a total of ten large segments (mean length 7.4 Mb or 16.4 cM) that showed a striking excess in the number of heterozygous sites (**Fig. 3D**) and a deficiency in the number of fixed *guttata*-specific alleles (**Fig. 3E**); see also **SI Appendix, Table S8**). This implies that only 5% of the diploid (somatic) genome of this hybrid male was of heterozygous *castanotis* origin (**Fig. 3**), most likely coming from the domesticated founder female. The low proportion of *castanotis* DNA suggests that hybrid females had been back-crossed with *guttata* males for approximately five generations (**Fig. 3A**), and explains why the birds were phenotypically indistinguishable from pure *guttata* birds (see also (16)).

### *Reduced genetic diversity in low-copy-number genes on the GRC*

For the GRC paralog of *elavl4*, all germline samples shared the same heterozygous sites and mostly had 94% in read coverage compared to the genomic background of the same sample (**SI Appendix, Figs. S6C, S7C**), suggesting that the GRC contains two copies of *elavl4* (about 40% of read coverage is expected for single-copy GRC genes, see **Fig. S7**). Therefore, we artificially phased two GRC haplotypes for each germline sample by keeping the reference allele for haplotype 1 while using the alternative allele for haplotype 2. This introduces a bias towards higher genetic diversity of haplotype 2, but this was not problematic for our purpose.

We compared the genetic diversity both among the GRC haplotypes, and between the GRC and their A-chromosomal paralogs using the gap-free alignments of the three GRC-linked genes, *bicc1*, *pim3* and *elavl4*. The A-chromosomal paralogs of the three genes represented the range of genome-wide autosomal genetic diversity found in *castanotis* in the wild (15), with *pim3<sub>A</sub>* representing low levels of diversity (16 SNPs per kb), *elavl4<sub>A</sub>* representing average levels (37 SNPs/kb), and *bicc1<sub>A</sub>* representing high levels (60 SNPs/kb; **SI Appendix, Figs. S8B-D right, S9 and Table S7**). We also found considerable genetic diversity in mitochondrial DNA (13 SNPs/kb; **SI Appendix, Fig. S8A and Table S7**). In contrast, the GRC-linked genes showed low genetic diversity (*pim3<sub>GRC</sub>*: 0.3 SNPs/kb, *bicc1<sub>GRC</sub>*:

0.6 SNPs/kb, *elavel4<sub>GRC</sub>* haplotype 1: 0.1 SNPs/kb and haplotype 2: 1.7 SNPs/kb; **SI Appendix, Fig. S8, Table S7**; also see **Fig. 4**).

## Behind-the-paper

Our findings on occasional paternal inheritance of the GRC resulted from (re)analyses of seven sources of data collected during 2017-2020. Most of these data had been generated for purposes unrelated to the study of the germline-restricted chromosome (GRC).

The Illumina PCR-free libraries from 15 ejaculates of brothers from families A and B and from the corresponding four blood samples of their parents came from data generated in 2017 and 2018 (data source #1; see ejaculates in **Fig. 2B,C**). The initial purpose of these data was to test for meiotic drive in ejaculates of males that were bred to be heterozygous for a putative meiotic driver on chromosome *Tgu2* (17). Their parents had been sequenced to facilitate phasing the two alleles in the ejaculate samples. No evidence for meiotic drive in ejaculates was found. However, during quality control we observed areas in the genome (particularly on chromosomes *Tgu1* and *Tgu10*) with unusually high sequencing coverage that appeared to contain no SNPs in all ejaculate libraries (after filtering), even though the two corresponding parental genotypes suggested that the male offspring should be heterozygous. Additionally, we detected apparently heterozygous sites (with minor allele frequency < 10%) for some regions that should have been homozygous, because the minor allele was absent in the parental libraries. Eventually, we figured out that these unexpected signals in the ejaculate samples must have come from the GRC, which we had started to investigate around the same time (see (2)). Then, we reanalyzed the ejaculate sequence data systematically and found that they all contained sequences from the GRC (**SI Appendix, Fig. S4**). Interestingly, we found that the signal of the GRC in ejaculates from family A looked consistently different from those of family B. Specifically, in the Manhattan plots of  $\log_2$  ejaculate-to-soma coverage ratios, ejaculates from family A missed the prominent high-copy-number peaks of reads mapping to regions on chromosomes *Tgu1*, *Tgu3*, *Tgu4* and *Tgu10* (in terms of ejaculate enrichment in coverage comparing to the soma library; **SI Appendix, Fig. S4**; also see **Fig. 2B,C** and (2)).

Given that our sequenced ejaculates contained GRC sequences and given the observed variation between the two families, we wanted to find out (a) where in the ejaculate these sequences can be found, and (b) whether the observed difference of GRC signal between the two families results from differences in elimination patterns, perhaps due to different GRC haplotypes. To this end, we conducted the FISH study on the ejaculates using a GRC-specific probe (2) (data source #2, collected in 2019), and we found that the GRC signals in ejaculates were mainly coming from the sperm heads, rather than from the eliminated free-floating GRC micro-nuclei (**Fig. 1**). Meanwhile, we sequenced (PCR-free) three testis samples with their corresponding soma samples from the two families to study the between-family differences in GRC haplotypes in terms of the presence/absence of high-copy-number peaks (data source #3; testis samples in **Fig. 2B,C**). Contrary to the initial idea of there being clearly different GRC haplotypes, we found that the GRC haplotypes of the two families were surprisingly similar (**SI Appendix, Fig. S5**), even in comparison to the already published GRCs from three *castanotis* individuals A, B and F (2) in **Fig. 2A** (data source #4). Only later, during manuscript preparation in 2020, did we understand that the family-difference in the efficiency of elimination of the GRC from sperm heads (**Figs. 1, 2E,F** and **Table S1**) was sufficient to explain the apparently missing high-copy number peaks (on *Tgu1*, 3, 4, and 10) in ejaculates of family A (**Fig. 2B,C**, and **SI Appendix, Fig. S4**). The high repeatability within families and the significant difference between families in the proportion of GRC-carrying sperm implied that the (potential) occurrence of paternal inheritance may vary between families, or between GRC haplotypes.

In 2019, we launched another project in which we sequenced mtDNA haplotypes of all our captive populations of zebra finches to assess mtDNA variation in zebra finches and to compare variation in captive versus wild-derived individuals (data source #5). This study was also done for a different purpose, namely to examine possible mito-nuclear incompatibilities that could explain substantial variation in infertility and embryo mortality between populations (see (18)). We sequenced four PCR

amplicons that spanned the whole mitogenome from 55 wild zebra finches and all female founders of the captive populations (or one of their offspring; **SI Appendix, Figs. S1 and S2**).

To be able to root the mtDNA tree, we also included a couple of related species as well as samples from two female zebra finches of the *guttata* subspecies. These latter samples came from a small population of alleged *guttata* birds (showing a *guttata* plumage phenotype) maintained by an amateur breeder. Back in 2013, we had bought two males and two females from this breeder to study meiotic drive in F1 hybrids and backcrosses between the two zebra finch subspecies *guttata* and *castanotis*. The results of that study were published recently (16). However, from sequencing the mtDNA of two such “*guttata*” individuals, we discovered, rather unexpectedly, that they carried the *castanotis*-B mtDNA haplotype (**Fig. 3F** and **SI Appendix, Figs. S1, S2**, also see (16)), rather than clustering with mtDNA sequences of *guttata*, which we extracted from published raw data (SRA accession numbers SRR2299402 (5) and SRR3208120 (6)). This suggested a history of hybridization in the past, possibly in captivity in Europe, where the mtDNA-haplotype B is widespread. In a subsequent analysis on the somatic DNA of one *guttata* male from this population (part of data source #6), we found that this alleged *guttata* male actually carried 5% of the *castanotis* A-chromosomal genome (**Fig. 3D,E**) and the same B-*castanotis* mtDNA haplotype as the two “*guttata*” females (**Fig. 3F**). This led us to deduce a (hypothetical) history of initial hybridization followed by backcrossing to restore the *guttata* phenotype. As we made these discoveries, we acquired all the remaining birds of that hybrid population from the same amateur breeder (7 males, 4 females) for future follow-up work.

As part of our ongoing work, now specifically aimed at studying the GRC of the zebra finch (see (2)), we sequenced the testis and liver samples of three zebra finches using linked reads, one wild-caught *castanotis* male, one *castanotis* male from an old but common matriline in captivity, and one “*guttata*” male founder from the meiotic-drive study (data source #6, collected in 2018–2019). We sequenced the testes of that one “*guttata*” male thinking it was a pure *guttata* and thus suitable as an outgroup for the *castanotis* GRCs. Later, we found out that this individual’s A-chromosomes contained 5% *castanotis* sequences, as well as a *castanotis* mtDNA. Based on (2), we assumed that the GRC is maternally inherited and hence co-inherited with the mtDNA (19–21). Thus, given that our hybrid *guttata*  $\times$  *castanotis* male had a *castanotis* mtDNA haplotype B, we expected to find a typical *castanotis* GRC. To our surprise, when checking the testis library of this hybrid male in 2020 (part of data source #6), we found a completely novel GRC-haplotype that differed strikingly from all other *castanotis* haplotypes (**Fig. 3B,C; SI Appendix, Fig. S5**). We therefore hypothesize that the novel haplotype represents the GRC of the subspecies *guttata*.

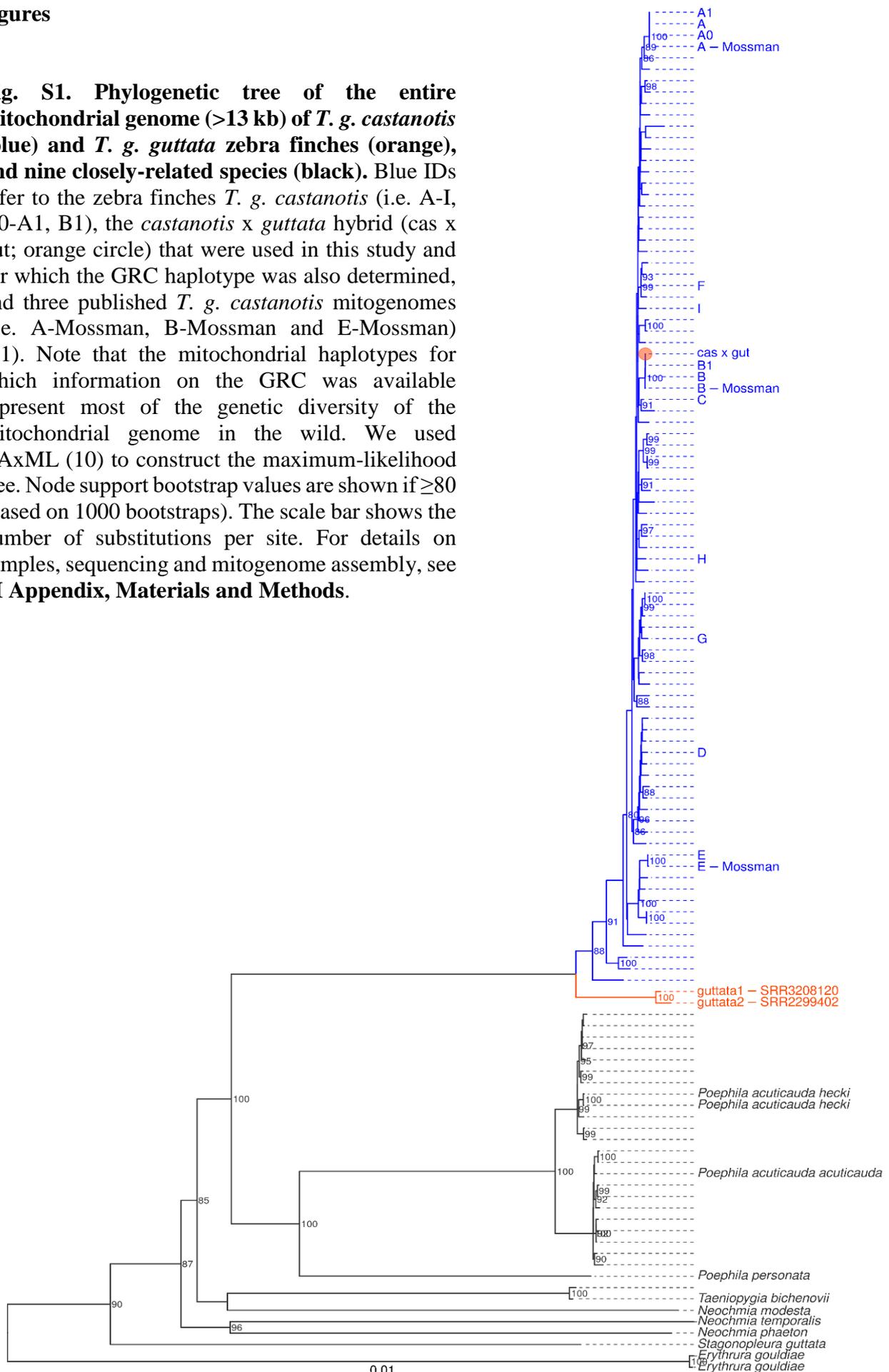
Meanwhile, we had put together a comprehensive dataset to study variation in GRC haplotypes across a total of nine *castanotis* matriline A-I (**Fig. 2A and 4**), which we considered to broadly represent the entire tree of mtDNA (**SI Appendix, Fig. S1**, which includes many mtDNA haplotypes from samples in the wild). The dataset included the three published GRCs (2) (data source #4), two *castanotis* zebra finches that were sequenced by linked reads (data source #6) and libraries (PCR-free) from testis and liver samples from four individuals representing additional matriline that dominate our captive populations “Krakow”, “Bielefeld” and “Melbourne” (see **SI Appendix, Fig. S2**), two of which are recently wild-derived (data source #7; also see **Materials and Methods**).

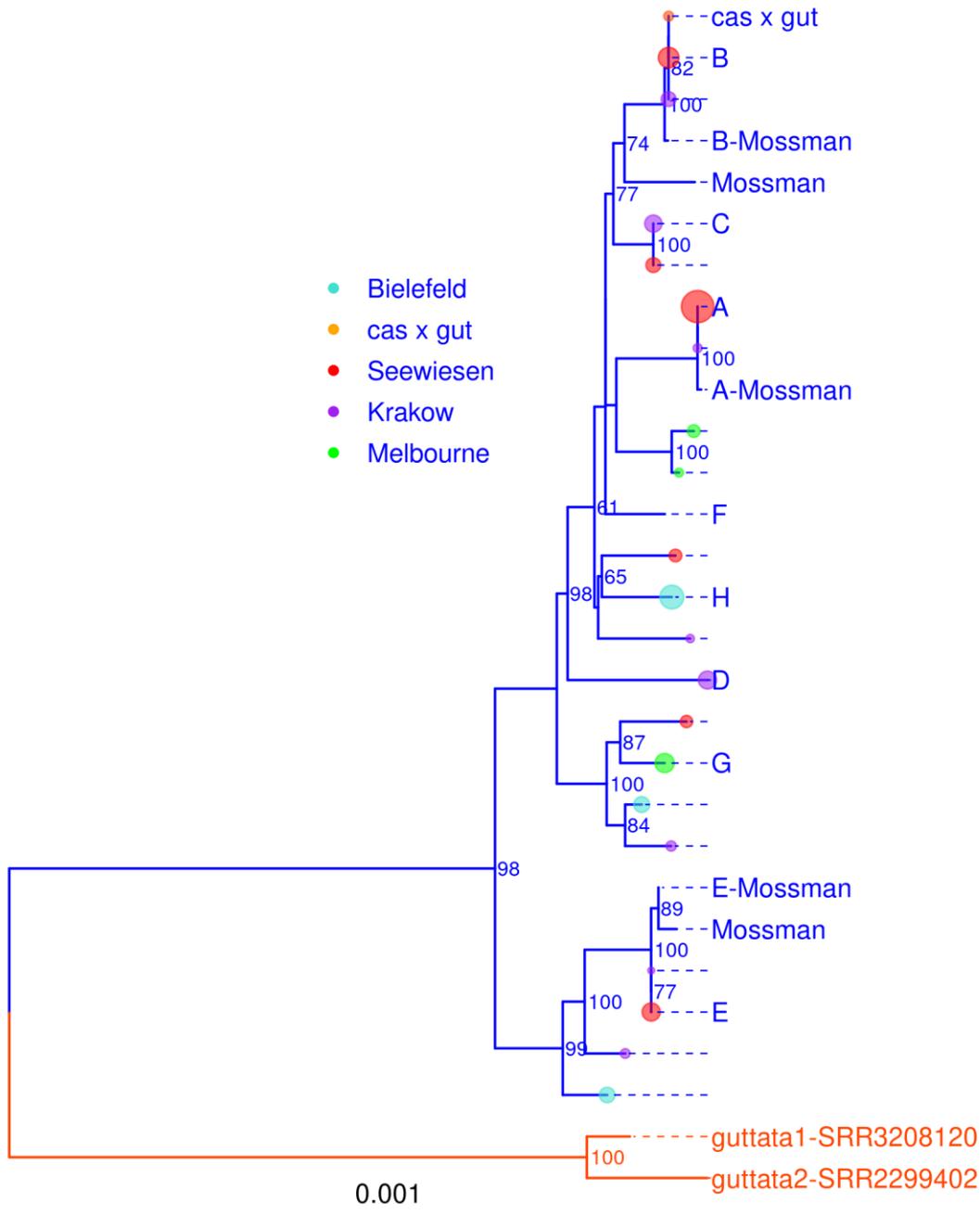
During the examination of all evidence for paternal inheritance of the GRC in the winter of 2020, we noticed that all examined *castanotis* GRCs are extremely similar, even though they came from independent populations, in clear contrast to the highly diverged mtDNA haplotypes of the same set of males (**Figs. 4 and S8**). To explain this, we hypothesize that a single GRC haplotype expanded in the recent evolutionary history of the *castanotis* population, crossing matriline boundaries via paternal inheritance.

At that stage, we put together the manuscript, aiming to transparently present all the available data from the aforementioned ongoing projects. Taken together, we conclude that occasional paternal inheritance (“spillover”, **Fig. 3**) of the GRC via sperm (**Figs. 1, 2**) is essential to understand the currently existing low genetic diversity among the GRC haplotypes (**Fig. 4**).

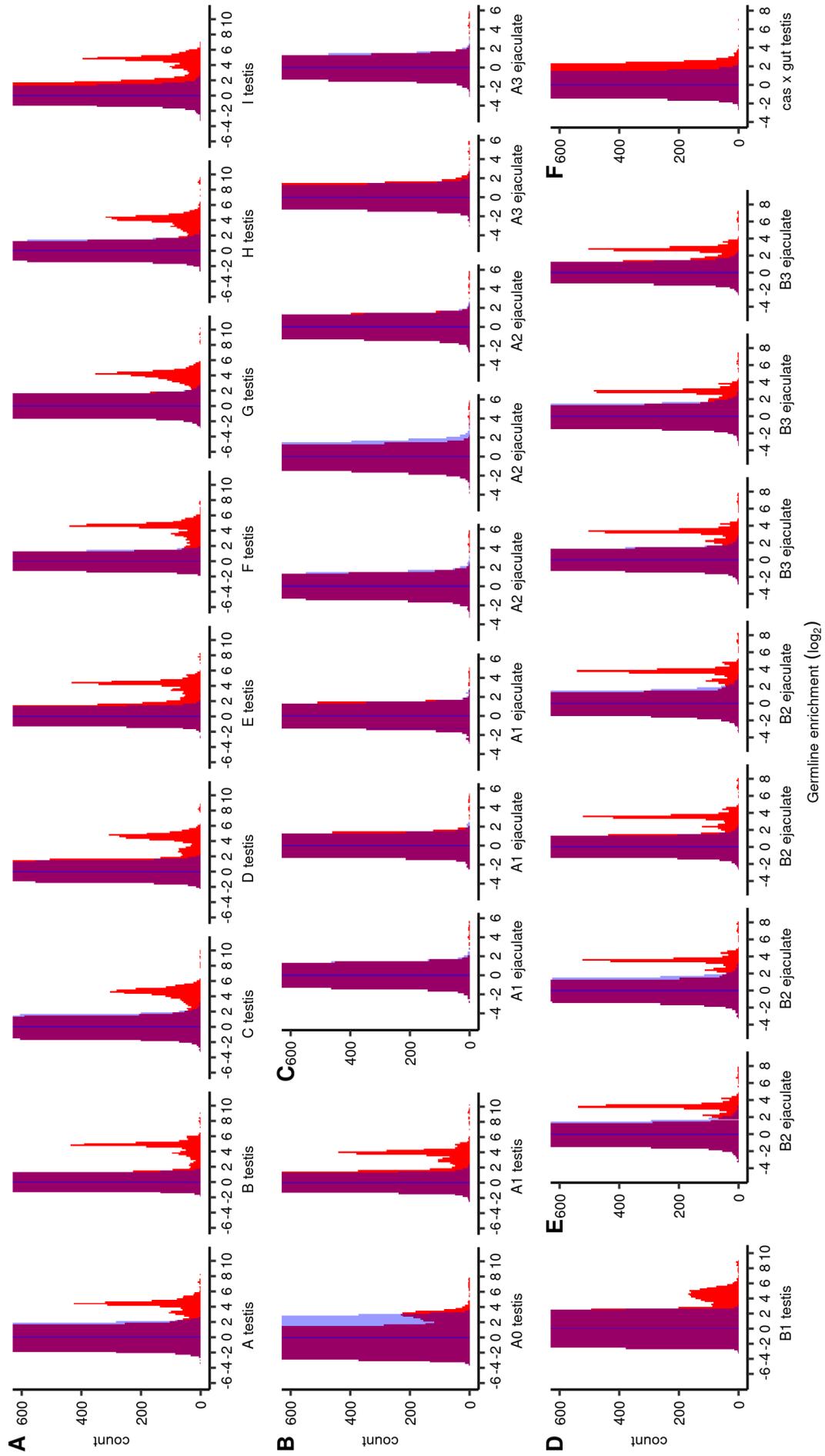
## Figures

**Fig. S1. Phylogenetic tree of the entire mitochondrial genome (>13 kb) of *T. g. castanotis* (blue) and *T. g. guttata* zebra finches (orange), and nine closely-related species (black).** Blue IDs refer to the zebra finches *T. g. castanotis* (i.e. A-I, A0-A1, B1), the *castanotis* x *guttata* hybrid (cas x gut; orange circle) that were used in this study and for which the GRC haplotype was also determined, and three published *T. g. castanotis* mitogenomes (i.e. A-Mossman, B-Mossman and E-Mossman) (11). Note that the mitochondrial haplotypes for which information on the GRC was available represent most of the genetic diversity of the mitochondrial genome in the wild. We used RAxML (10) to construct the maximum-likelihood tree. Node support bootstrap values are shown if  $\geq 80$  (based on 1000 bootstraps). The scale bar shows the number of substitutions per site. For details on samples, sequencing and mitogenome assembly, see **SI Appendix, Materials and Methods**.

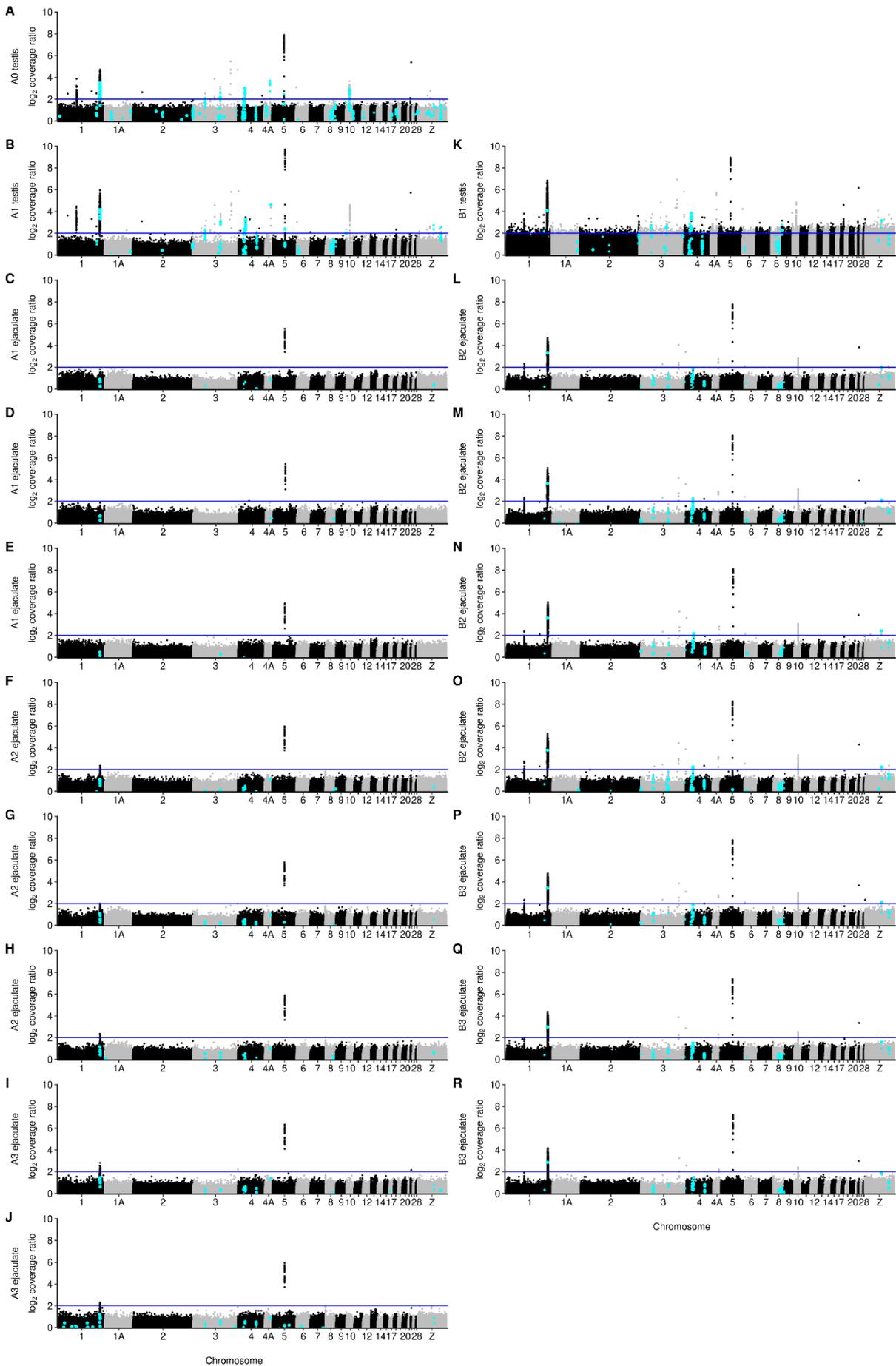




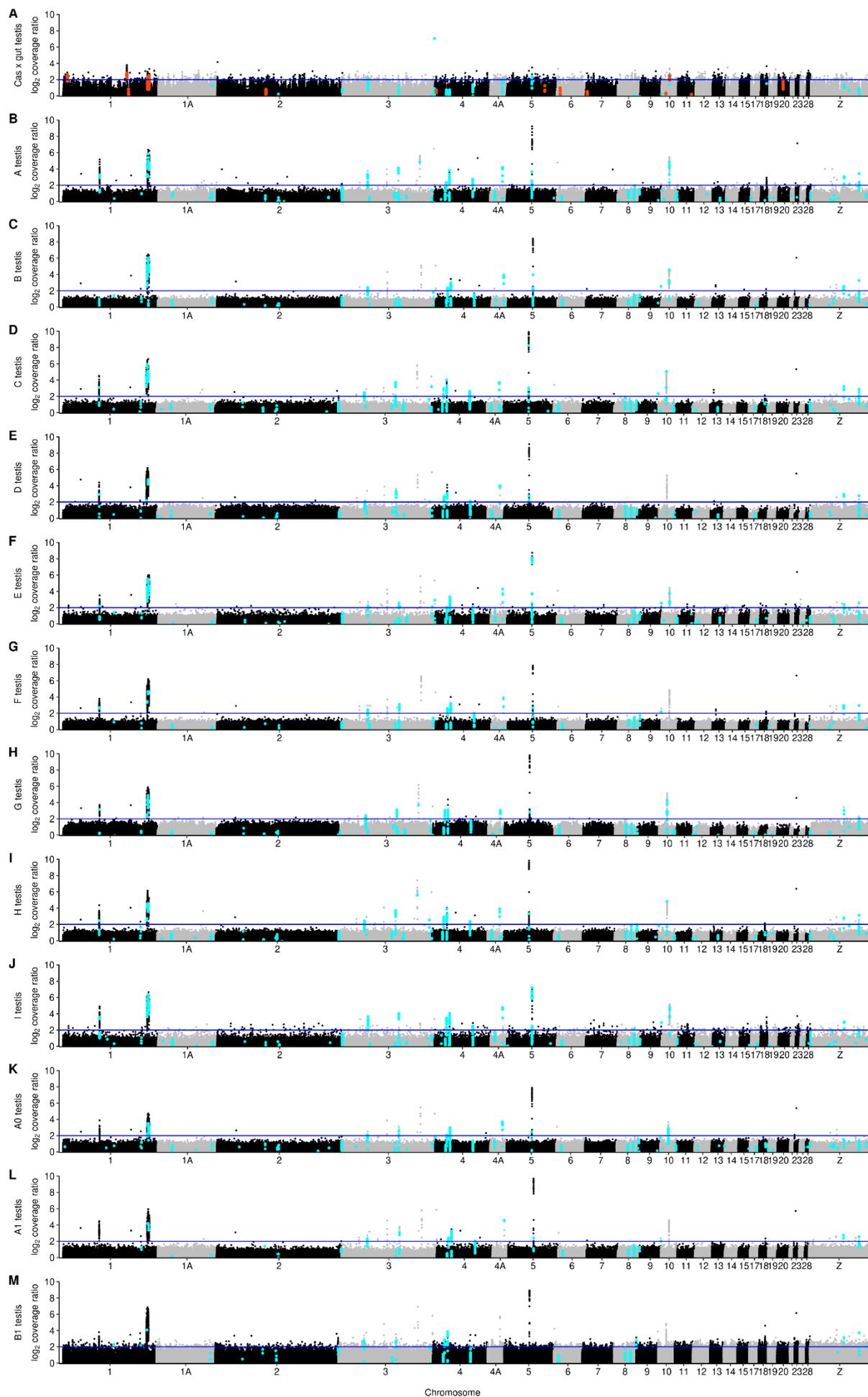
**Fig. S2. Phylogenetic tree of the mitochondrial genome of our four captive populations of *T. g. castanotis* zebra finches (matrilines A-H), our captive *castanotis x guttata* hybrid zebra finches (see Supplementary results), and five published mitochondrial haplotypes (blue IDs labeled with Mossman; from (11)).** The blue labels A to H and cas x gut refer to matriline of different populations in which at least one male was sequenced for its GRC (testis and soma) in this study (details see **Materials and Methods**; also see **Fig. 3F**). The true *guttata* mitochondrial assemblies (in orange) were included as the outgroup. Node support bootstrap values are shown if >60 (based on 1000 bootstraps). Circle size indicates the number of birds (range: 5 to 2245) that carried that haplotype in our captive population. Circle color shows the population in which that haplotype was found (see **Materials and Methods** in the main text for population background). The scale bar shows the number of substitutions per site. For additional details see **SI Appendix, Materials and Methods**.



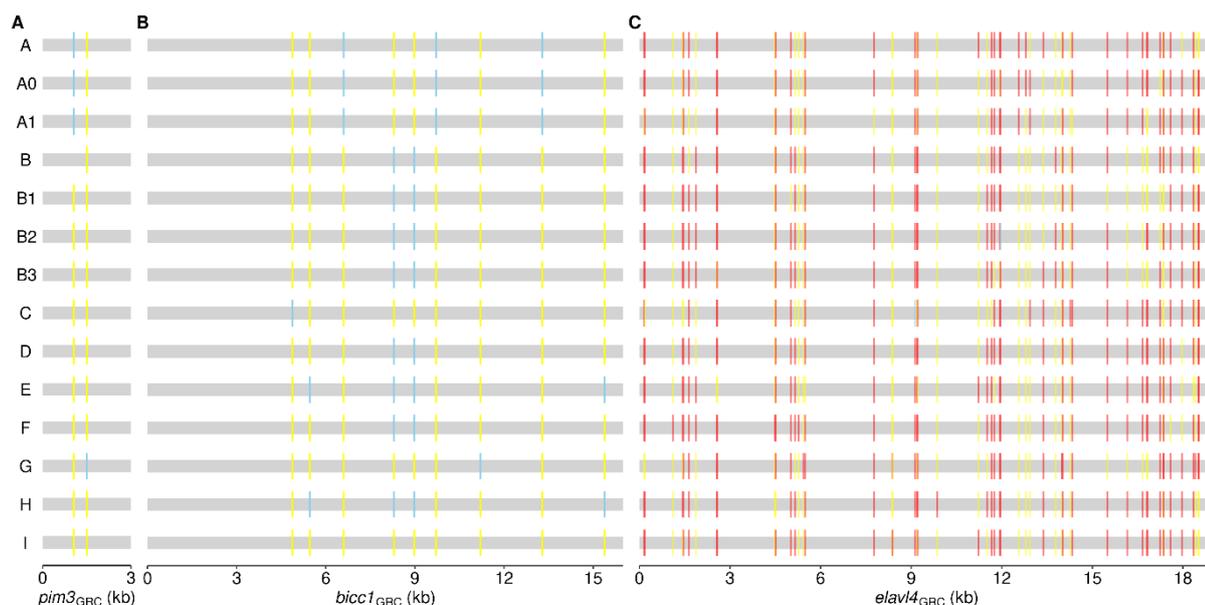
**Fig. S3. Histograms of the log<sub>2</sub> germline-to-soma coverage ratio for all testis and ejaculate samples of *T. g. castanotis* zebra finches (A-E) and for one *castanotis* x *guttata* hybrid zebra finch (F) used in this study.** Red shows the germline enrichment in coverage (thereafter germline enrichment) in testis and ejaculate compared to the soma samples, in 1-kb windows of the A-chromosomal paralogs. Blue depicts the symmetrical distribution of the log<sub>2</sub> germline-to-soma coverage ratios (in 1-kb windows of the A-chromosomal paralogs) if there was no GRC-paralog in the germline samples. (A) Germline enrichment (red) in testis samples for nine *T. g. castanotis* males from different matriline (A to I; i.e. mitochondrial haplotypes; **SI Appendix, Table S2**; for details of population background also see **Materials and Methods** in main text). (B) Germline enrichment (red) in testis samples for male A1 and his uncle (A0) representing mitochondrial haplotype A. (C) Germline enrichment (red) in ejaculate samples for three brothers (A1-A3). (D) Germline enrichment (red) in testis sample for male B1 representing mitochondrial haplotype B. (E) Germline enrichment (red) in ejaculate samples for the two brothers (B2 and B3) of male B1. (F) Germline enrichment (red) in testis sample for the *castanotis* x *guttata* hybrid zebra finch (cas x gut). The low germline enrichment in coverage ratio of the cas x gut testis is due to the difference in sampling (smaller testes, see **Materials and Methods** in main text and **SI Appendix, Table S2**). Note the remarkable reduction of germline enrichment in ejaculates from family A (C) compared to family B (E) and the remarkable similarity in germline enrichment in all testis samples of *castanotis* zebra finches (A,B,D).



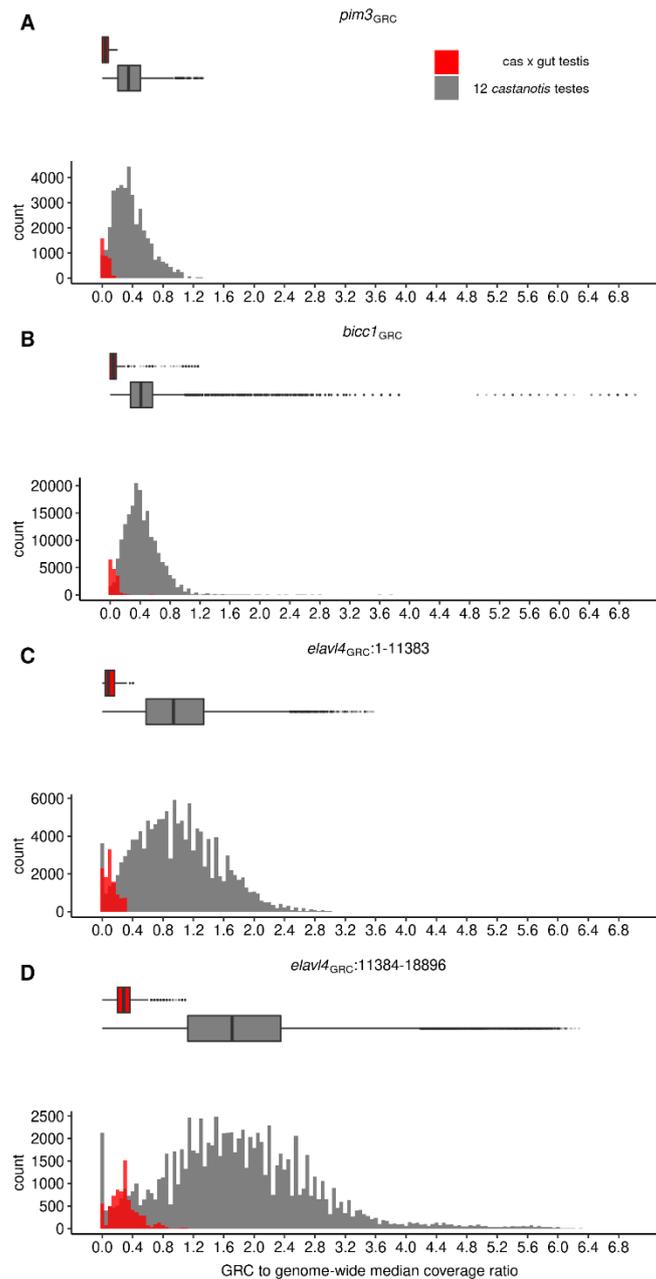
**Fig. S4. Manhattan plots of  $\log_2$  testis-to-soma coverage ratios ((A-B) and (K)) and  $\log_2$  ejaculate-to-soma coverage ratios ((C-J) and (L-R)) across the major A chromosomes (larger than 5 Mb) for *T. g. castanotis* males (brothers) from families of mitochondrial haplotypes A (A-J) and B (right; (K-R)).** Each dot shows a 1-kb window. Cyan depicts windows that have >15 high confidence testis- or ejaculate-specific SNPs. The horizontal blue line indicates  $\log_2$  testis-to-soma and ejaculate-to-soma coverage ratios = 2. Note that all ejaculate samples from family A (C-J) have fewer cyan dots (i.e. windows that contain ejaculate-specific SNPs) and fewer dots above the blue line (i.e. windows in which the read coverages are enriched for ejaculate sample comparing to the soma, e.g. on chromosomes 1, 3, 4 and 10) compared to ejaculate samples from family B (L-R) and all testis samples ((A-B) and (K)), indicating a lower GRC amount in ejaculate samples from family A compared to family B (for coverage enrichment see **Fig. 2B-C** and **SI Appendix, Fig. 3B-E**).



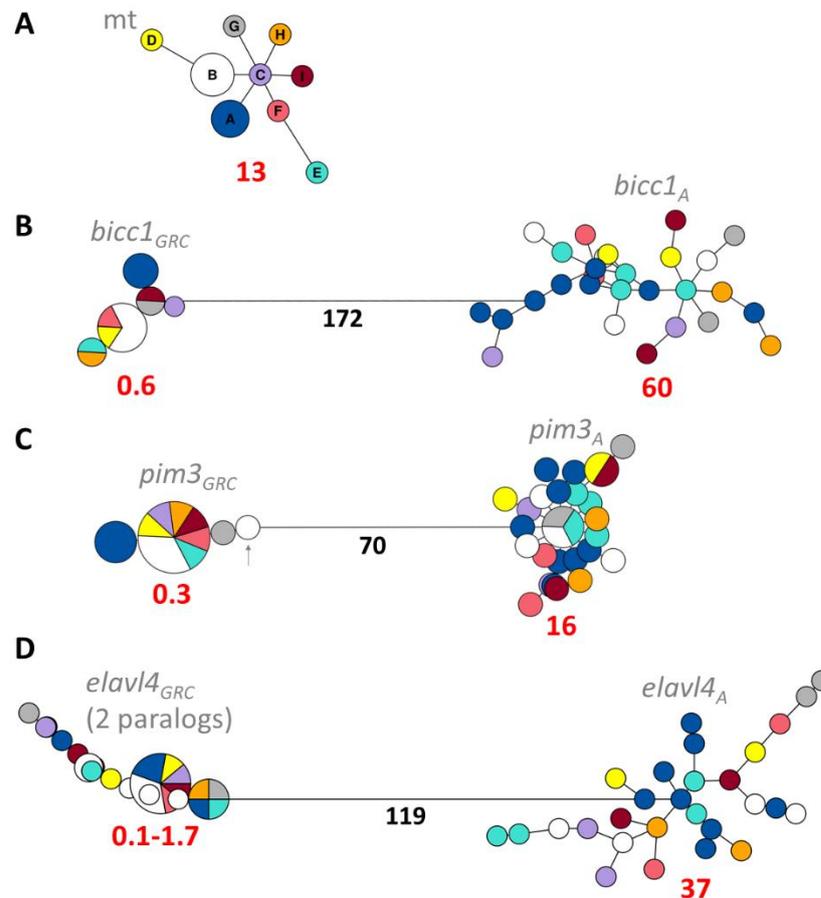
**Fig. S5. Manhattan plots of  $\log_2$  testis-to-soma coverage ratios in 1-kb windows across the major A chromosomes of the captive *castanotis* x *guttata* hybrid (A; i.e. male cas x gut) in comparison to *T. g. castanotis* zebra finches (B-M; i.e. males A to I, A0, A1, and B1, respectively) with nine different mitochondrial haplotypes A to I. Males A, A0 and A1 had the same mitochondrial haplotype A, whereas males B and B1 had the same mitochondrial haplotype B (for details see **Table S2** and figure legend of **SI Appendix, Fig. S3**). Each dot shows a 1-kb window. Red depicts windows that contain high-confidence private testis-specific SNPs (i.e. the testis-specific allele is private to that male and absent from all other testis or soma libraries; see **Fig. 3B-C** for high confidence regions of example samples; for details see **Methods**). Cyan depicts windows that contain >15 high confidence testis-specific SNPs (for coverage enrichment of testis samples see **Fig. 2A-C** and **SI Appendix, Fig. S3A,B,D,F**). The horizontal blue line indicates  $\log_2$  testis-to-soma coverage ratio = 2. Note that only the *castanotis* x *guttata* hybrid shows clustered windows with private testis-specific SNPs (A). Also note the similarity of the distribution of the windows that are enriched for testis-specific SNPs (cyan dots) and the regions with testis-enriched coverage (dots that are above the blue horizontal line) of all *T. g. castanotis* samples (B-M), despite their different mitochondrial haplotypes.**



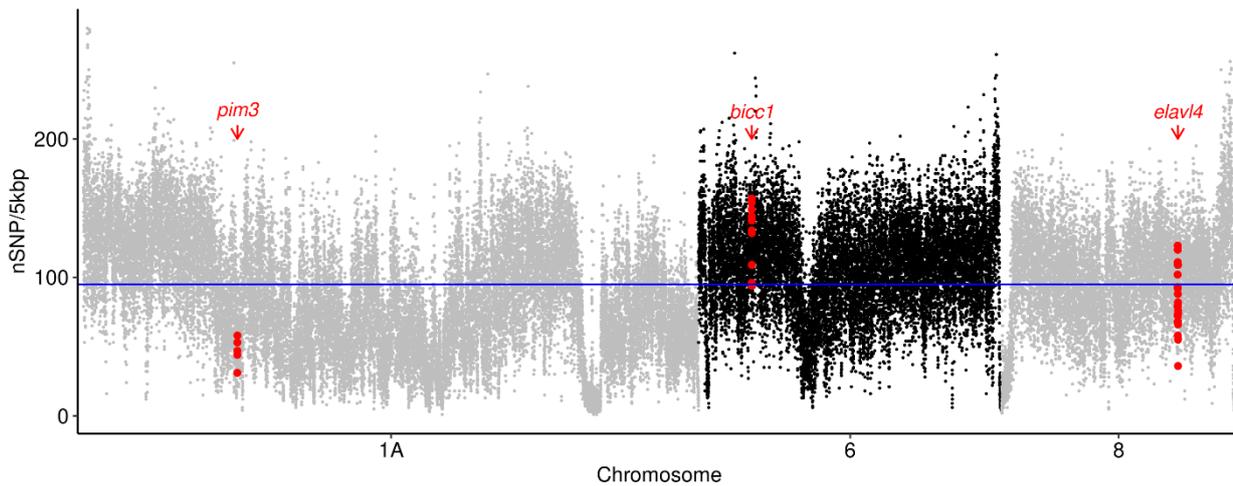
**Fig. S6. GRC haplotypes of all examined *T. g. castanotis* zebra finches based on the genotype calls of GRC paralogs *pim3<sub>GRC</sub>* ((A);  $N_{SNPs} = 2$ ), *bicc1<sub>GRC</sub>* ((B);  $N_{SNPs} = 9$ ) and *elavl4<sub>GRC</sub>* ((C);  $N_{SNPs} = 52$ ).** Each row represents one individual. Grey indicates consensus sequence among all samples. Each vertical bar represents a SNP, whereby the genotype is depicted as either homozygous for the reference allele (yellow), homozygous for the alternative allele (blue) or heterozygous (red). Note that some vertical bars appear orange due to overlapping of red and yellow bars. All males are homozygous for SNP genotypes on GRC paralogs *pim3<sub>GRC</sub>* and *bicc1<sub>GRC</sub>* (i.e. all bars are blue or yellow in (A-B)), indicating that *pim3<sub>GRC</sub>* and *bicc1<sub>GRC</sub>* are in single-copy on the GRC (also see **SI Appendix, Figs. S7A,B**). All males have heterozygous genotypes on *elavl4<sub>GRC</sub>* (i.e. red in (C)), indicating that the *elavl4<sub>GRC</sub>* is present in at least two copies on the GRC (also see **SI Appendix, Figs. S7C,D**). Note that for the two single-copy GRC paralogs *pim3<sub>GRC</sub>* and *bicc1<sub>GRC</sub>*, males with the same mitochondrial haplotype (i.e. males A, A0 and A1 representing mitochondrial haplotype A, and males B, B1 to B3 representing mitochondrial haplotype B) have the same genotype. Additionally, males with different mitochondrial haplotypes can also share the same genotype (e.g. males B1, D and F; males E and H in (A,B)). For details on the population background of the males see **Materials and Methods** and **SI Appendix, Table S2**.



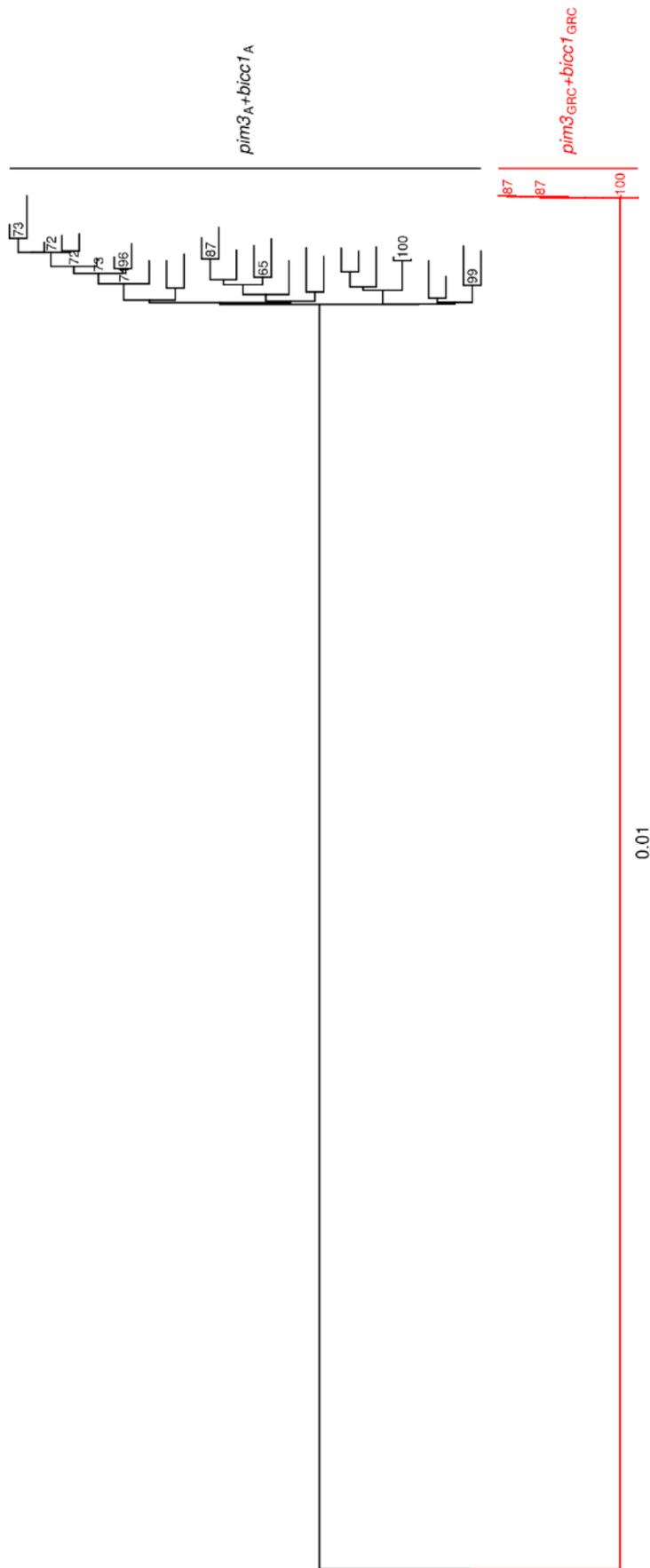
**Fig. S7. Distributions of the coverage ratio between the read depth of each base position of the filtered GRC paralogs (free from somatic reads) to the genome-wide median read depth for 12 *castanotis* testis samples (grey) and the *castanotis x guttata* testis sample (red) for GRC paralogs *pim3<sub>GRC</sub>* ((A); 3,349 bp; median coverage ratio = 0.35), *bicc1<sub>GRC</sub>* ((B); 15,619 bp; median coverage ratio = 0.41) and *elavl4<sub>GRC</sub>* ((C-D); 18,896 bp, median coverage ratio = 1.17). Thick horizontal lines show the median, and boxes indicate the 25th and 75th percentiles. The assembled *elavl4<sub>GRC</sub>* consisted of five contigs; (C) shows the coverage ratio for the first four contigs of *elavl4<sub>GRC</sub>* (base positions < 11384; median coverage ratio = 0.94) while (D) shows the last contig (i.e, positions > 11383; median coverage ratio = 1.71). Paralogs *pim3<sub>GRC</sub>* and *bicc1<sub>GRC</sub>* are clearly in single copy on the GRC (also see **SI Appendix, Fig. S6**). The majority of *elavl4<sub>GRC</sub>* is present in dual copies (C) whereas part of it is present in four copies on the GRC (D); also see **SI Appendix, Fig. S6**). Note that the *castanotis x guttata* testis sample had extremely low coverage on the GRC (median coverage ratio = 0.04 for the two single-copy GRC genes (A,B); i.e. median depth = 1 read, with lower and upper 95% quantiles of 0 and 4), presumably due to underdeveloped testes.**



**Fig. S8. Haplotype network showing the genetic diversity in GRC-linked genes ( $gene_{GRC}$ ) compared to their A-chromosomal paralogs ( $gene_A$ ) and the mtDNA (mt) in *castanotis* zebra finches.** Shown are gap-free alignments of mitogenomes (A), two single-copy genes on the GRC (2) (B, 13,822 bp; C, 3,286 bp) and a low-copy-number gene (presumably two paralogs) on the GRC (2) (D, 17,492 bp) with the two haplotypes of the respective single-copy A-chromosome paralogs from all *castanotis* germline samples used in this study ( $N = 14$  individuals). Note that all haplotypes of the A-chromosome paralogs were constructed from all available *castanotis* somatic libraries (SI Appendix, Table S2). Colors represent the different mitogenome haplotypes. The size of each circle indicates the number of individuals 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 kilobase for each cluster of haplotypes; black numbers refer to the number of mutations per kilobase between the GRC-linked and A-chromosomal paralogs (B-D). Note the highly reduced genetic diversity (red numbers) in the GRC genes (which corresponds to a mean number of pairwise differences per site  $\pi = 0.00015$ - $0.00027$ ; SI Appendix, Table S7) in comparison to their A-chromosomal paralogs ( $\pi = 0.0013$ - $0.011$ ; SI Appendix, Table S7 and Fig. S9; (B-D)). Further note that different mitogenome haplotypes may share the same GRC haplotypes (i.e. the circles contain multiple colors in GRC genes in (B-D); also see Fig. 4), indicating that mtDNA and GRC are not always inherited together. Repeated samples of haplotypes A (blue) and B (white) mostly consist of close relatives, explaining why they share the same mtDNA, GRC, and similar A-chromosomal DNA. The one haplotype of  $pim3_{GRC}$  (arrow in (C)) that deviated from the major haplotype of  $pim3_{GRC}$  (pie-chart in (C)) was due to a missing data on one SNP.



**Fig. S9. Manhattan plot of the number of segregating sites per 5-kb window in the pool of 100 wild-caught *T. g. castanotis* zebra finches.** The blue line shows the median number of SNPs (median = 94) per 5-kb in overlapping sliding windows, based on pool sequencing of blood samples. Note that the three A-chromosomal paralogs of the GRC-linked genes *pim3*, *bicc1* and *elavl4* (shown in red) represent low, high and intermediate levels of genetic diversity (also see **SI Appendix, Fig. S8**) in comparison to the average autosomal genetic diversity (blue line).



**Fig. S10. Phylogenetic tree of the 9 GRC haplotypes (red; i.e. concatenated sequences of the two single-copy genes *pim3*<sub>GRC</sub> and *bicc1*<sub>GRC</sub>; also see Fig. 4A) from *T. g. castanotis* males A-I rooted by the concatenated sequences of the two A-chromosomal paralogs from all *castanotis* somatic libraries (black; *pim3*<sub>A</sub> + *bicc1*<sub>A</sub>; SI Appendix, Table S2). The tree was built using gap-free alignment. Node support bootstrap values are shown if >60 (based on 1000 bootstraps). The scale bar shows the number of substitutions per site. For additional details see Fig. 4.**

**Legend for supplementary video.****Video of the z-sections showing the presence of the GRC inside the nucleus of zebra finch *Taeniopygia guttata castanotis* sperm.**

The GRC-linked multi-copy probe *dph6* (see **Materials and Methods**) shows the presence of the GRC (red) inside some sperm heads. Blue DAPI stain without red shows sperm heads without GRC. Individual z-sections under a confocal microscope show the sequential appearance and disappearance of the *dph6* signal (red) along consecutive sections, indicating the location of the GRC within the nucleus of the spermatozoa. For additional details see **Fig. 1**. Scale bar is 20  $\mu\text{m}$ .

(The Video will be uploaded with the publication.)

## SI References

1. S. Immler, T. I. M. R. Birkhead, Short communication A non-invasive method for obtaining spermatozoa from birds. *Ibis (Lond. 1859)*, 827–830 (2005).
2. C. M. Kinsella, *et al.*, Programmed DNA elimination of germline development genes in songbirds. *Nat. Commun.* **10**, 5468 (2019).
3. J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual, 2nd ed., Vols. 1, 2 and 3* (1989).
4. W. C. Warren, *et al.*, The genome of a songbird. *Nature* **464**, 757–762 (2010).
5. S. Singhal, *et al.*, Stable recombination hotspots in birds. *Science (80-. )*. **350**, 928–932 (2015).
6. J. H. Davidson, C. N. Balakrishnan, Gene Regulatory Evolution During Speciation in a Songbird. *Genes/Genomes/Genetics* **6**, 1357–1364 (2016).
7. B. Liang, N. Wang, N. Li, R. T. Kimball, E. L. Braun, Comparative genomics reveals a burst of homoplasmy-free numt insertions. *Mol. Biol. Evol.* (2018) <https://doi.org/10.1093/molbev/msy112>.
8. D. Li, C. M. Liu, R. Luo, K. Sadakane, T. W. Lam, MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* (2015) <https://doi.org/10.1093/bioinformatics/btv033>.
9. H. Li, Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* (2018) <https://doi.org/10.1093/bioinformatics/bty191>.
10. 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).
11. 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).
12. A. Suh, J. O. Kriegs, J. Brosius, J. Schmitz, Retroposon insertions and the chronology of avian sex chromosome evolution. *Mol. Biol. Evol.* **28**, 2993–2997 (2011).
13. J. Rozas, *et al.*, DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* (2017) <https://doi.org/10.1093/molbev/msx248>.
14. E. Paradis, Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics* (2010) <https://doi.org/10.1093/bioinformatics/btp696>.
15. U. Knief, *et al.*, Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biol.* **17**, 199 (2016).
16. U. Knief, W. Forstmeier, Y. Pei, J. Wolf, B. Kempnaers, A test for meiotic drive in hybrids between Australian and Timor zebra finches. *Ecol. Evol.*, ece3.6951 (2020).
17. U. Knief, H. Schielzeth, H. Ellegren, B. Kempnaers, W. Forstmeier, A prezygotic transmission distorter acting equally in female and male zebra finches *Taeniopygia guttata*. *Mol. Ecol.* **24**, 3846–3859 (2015).
18. Y. Pei (裴一凡), *et al.*, Proximate Causes of Infertility and Embryo Mortality in Captive Zebra Finches. *Am. Nat.* **196**, 577–596 (2020).
19. 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).
20. 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*. *Chromosom. Res.* **6**, 105–113 (1998).
21. 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).

## Tables

**Table S1.** Estimated amounts of GRC-carrying sperm heads (i.e. dph6 probe positive) and GRC micronuclei from fluorescent in situ hybridization (FISH) preparations of *Taeniopygia guttata castanotis* ejaculates. For details of collection of natural ejaculates see SI Appendix, Materials and Methods and for details of FISH see Materials and Methods. On average, ejaculates of matriline A and matriline B showed 1% and 9% of GRC-positive sperm heads, respectively (also see Fig. 2).

FISH slide ID	Sperm ID	Male ID	Matriline	N sperm heads	N free-floating GRC micronuclei	N GRC-positive sperm heads	Proportion of GRC-positive sperm head	N GRC micronuclei per 100 sperm heads
ZFSP86-06082019	ZFSperm86	A3*	A	677	4	6	0.01	1
ZFSP87-10092019	ZFSperm87	B2*	B	273	0	35	0.13	0
ZFSP88-10092019	ZFSperm88	B4	B	196	4	18	0.09	2
ZFSP89-10092019	ZFSperm89	B5	B	277	6	16	0.06	2
ZFSP89-06082019	ZFSperm89	B5	B	271	2	19	0.07	1
ZFSP89-12082019	ZFSperm89	B5	B	270	5	14	0.05	2
ZFSP90-10092019	ZFSperm90	B6	B	266	1	51	0.19	0

\*Additional ejaculates of males A3 and B2 were sequenced (SI Appendix, Table S2; sequencing results see Fig. 2 and SI Appendix, Figs. S3-S4).

**Table S2.** Description of all germline (testis and ejaculate) and their corresponding soma (liver, muscle or blood of the two parents) samples used in this study. Note that germline and their soma samples were always processed using the same library preparation and sequencing methods.

Sample name	Sex	Male ID	Matri-line* ID	Age (days)	Tissue type	Individual ID in pedigree or database	Population†	Mother ID	Father ID	Library preparation‡	Reference
taeGutGut-testis	M	cas x gut	B	>1828	Testis	TR13008	European captive	-	-	10X	This study
taeCasSA-testis	M	I	I	Unknown	Testis	B56513	Chimooli Dam, 8.3 km ENE Wyndham, Western Australia wild	-	-	10X	This study
MR1904-2-LT	M	G	G	279	Left testis	MR19042	Melbourne recently wild-derived	-	-	PCR-free	This study
BR19117-LT	M	H	H	238	Left testis	BR19117	Bielefeld recently wild-derived	-	-	PCR-free	This study
CR19001-LT	M	C	C	328	Left testis	CR19001	Krakow domesticated	-	-	PCR-free	This study
CR19103-LT	M	D	D	304	Left testis	CR19103	Krakow domesticated	-	-	PCR-free	This study
taeCasSE-testis	M	E	E	1882	Left testis	SR12451	Seewiesen domesticated	-	-	10X	This study
SpainM2-testis	M	F	F	Unknown	Testis	Spain_2	Spain domesticated	-	-	PCR-free	Kinsella et al. 2019 (1)
SpainM1-testis	M	B	B	Unknown	Testis	Spain_1	Spain domesticated	-	-	PCR-free	Kinsella et al. 2019 (1)
SR00100-testis	M	A	A	1371	Testis	SR00100	Seewiesen domesticated	-	-	10X	Kinsella et al. 2019 (1)
SR12483-LT	M	A0	A	1810	Left testis	SR12483 §	Seewiesen domesticated	-	-	PCR-free	This study
SR16288-LT	M	A1	A	880	Left testis	SR16288	Seewiesen domesticated	SR12333 §	SR15064	PCR-free	This study
ZFSperm 29	M	A1	A	192	Ejaculate	SR16288	Seewiesen domesticated	SR12333 §	SR15064	PCR-free	This study
ZFSperm 34	M	A1	A	194	Ejaculate	SR16288	Seewiesen domesticated	SR12333 §	SR15064	PCR-free	This study
ZFSperm 37	M	A1	A	196	Ejaculate	SR16288	Seewiesen domesticated	SR12333 §	SR15064	PCR-free	This study
ZFSperm 24	M	A2	A	186	Ejaculate	SR16289	Seewiesen domesticated	SR12333 §	SR15064	PCR-free	This study
ZFSperm 27	M	A2	A	188	Ejaculate	SR16289	Seewiesen domesticated	SR12333 §	SR15064	PCR-free	This study
ZFSperm 30	M	A2	A	190	Ejaculate	SR16289	Seewiesen domesticated	SR12333 §	SR15064	PCR-free	This study
ZFSperm 23	M	A3	A	100	Ejaculate	SR16298	Seewiesen domesticated	SR12333 §	SR15064	PCR-free	This study
ZFSperm 33	M	A3	A	106	Ejaculate	SR16298	Seewiesen domesticated	SR12333 §	SR15064	PCR-free	This study
SR16281-LT	M	B1	B	969	Left testis	SR16281	Seewiesen domesticated	SR15062	SR12047	PCR-free	This study
ZFSperm 22	M	B2	B	167	Ejaculate	SR16294	Seewiesen domesticated	SR15062	SR12047	PCR-free	This study
ZFSperm 26	M	B2	B	169	Ejaculate	SR16294	Seewiesen domesticated	SR15062	SR12047	PCR-free	This study
ZFSperm 32	M	B2	B	173	Ejaculate	SR16294	Seewiesen domesticated	SR15062	SR12047	PCR-free	This study
ZFSperm 40	M	B2	B	177	Ejaculate	SR16294	Seewiesen domesticated	SR15062	SR12047	PCR-free	This study
ZFSperm 21	M	B3	B	163	Ejaculate	SR16297	Seewiesen domesticated	SR15062	SR12047	PCR-free	This study
ZFSperm 31	M	B3	B	169	Ejaculate	SR16297	Seewiesen domesticated	SR15062	SR12047	PCR-free	This study
ZFSperm 35	M	B3	B	171	Ejaculate	SR16297	Seewiesen domesticated	SR15062	SR12047	PCR-free	This study
taeGutGut-soma	M	cas x gut	B	>1828	Liver	TR13008	European captive	-	-	10X	This study

taeCasSA-soma	M	I	I	Unkown	Liver	B56513	Chimooli Dam, 8.3 km ENE Wyndham, Western Australia wild	-	-	10X	This study
MR19042	M	G	G	279	Liver	MR19042	Melbourne recently wild-derived	-	-	PCR-free	This study
BR19117	M	H	H	238	Liver	BR19117	Bielefeld recently wild-derived	-	-	PCR-free	This study
CR19001	M	C	C	328	Liver	CR19001	Krakow domesticated	-	-	PCR-free	This study
CR19103	M	D	D	304	Liver	CR19103	Krakow domesticated	-	-	PCR-free	This study
taeCasSE-soma	M	E	E	1882	Liver	SR12451	Seewiesen domesticated	-	-	10X	This study
SpainM2-soma	M	F	F	Unkown	Muscle	Spain_2	Spain domesticated	-	-	PCR-free	Kinsella et al. 2019 (1)
SpainM1-soma	M	B	B	Unkown	Muscle	Spain_1	Spain domesticated	-	-	PCR-free	Kinsella et al. 2019 (1)
SR00100-soma	M	A	A	1371	Liver	SR00100	Seewiesen domesticated	-	-	10X	Kinsella et al. 2019 (1)
SR12483 §	M	A0	A	1810	Liver	SR12483 §	Seewiesen domesticated	SR11156	SR11451	PCR-free	This study
SR12333 §	F	-	A	-	Blood	SR12333 §	Seewiesen domesticated	SR11156	SR11451	PCR-free	This study
SR15064	M	-	A	-	Blood	SR15064	Seewiesen domesticated	SR14018	SR14024	PCR-free	This study
SR15062	F	-	B	-	Blood	SR15062	Seewiesen domesticated	SR14009	SR14027	PCR-free	This study
SR12047	M	-	E	-	Blood	SR12047	Seewiesen domesticated	SR11068	SR11600	PCR-free	This study
ZFSperm 86	M	A3	A	833	Ejaculate	SR16298	Seewiesen domesticated	SR12333 §	SR15064	-	This study
ZFSperm 87	M	B2	B	872	Ejaculate	SR16294	Seewiesen domesticated	SR15062	SR12047	-	This study
ZFSperm 88	M	B4	B	906	Ejaculate	SR16292	Seewiesen domesticated	SR15062	SR12047	-	This study
ZFSperm 89	M	B5	B	901	Ejaculate	SR16295	Seewiesen domesticated	SR15062	SR12047	-	This study
ZFSperm 90	M	B6	B	912	Ejaculate	SR16291	Seewiesen domesticated	SR15062	SR12047	-	This study

\*Matrilines (i.e. mitochondria haplotypes) A, B and D were first described in Mossman et al. 2006 (2).

†For details of population background, see Samples in Methods.

‡Details of library preparation see Methods.

§Male A0 (i.e. SR12483) was a brother of female SR12333.

#### References:

1. C. M. Kinsella, et al., Programmed DNA elimination of germline development genes in songbirds. *Nat. Commun.* 10, 5468 (2019).
2. 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).

**Table S3.** Estimated fixed and random effects from mixed-effect models (lmer) for (median) log<sub>2</sub> ejaculate-to-soma coverage ratio (of the selected 1,742 1-kb windows that were enriched for GRC-linked reads, for details see Materials and Methods and also legend of Fig. 2) of ejaculates from males of matriline A and B. Estimate refers to the fixed effect estimates from linear models (lm) for the median of log<sub>2</sub> germline to soma coverage ratio (of the selected 1,742 1-kb windows) for matriline A and B. Individual repeatability,  $R_{\text{male}} = 0.98$  for the medians of 15 ejaculates (and  $R_{\text{male}} = 0.82$  for the values per window among 26,023 data points), was calculated as the estimated variance of male identity divided by the sum of estimated variances for all random effects and the residual variance. Matriline repeatability,  $R_{\text{matriline}} = 0.96$  for the medians of 15 ejaculates (and  $R_{\text{matriline}} = 0.86$  for the values per window among 26,023 data points), was calculated as the estimated variance of male identity divided by the sum of estimated variances for all random effects and the residual variance in that model. For results of linear models also see Fig. 2.

Model purpose and response variable	Effects	N	Estimate	SE	Z	P§
Individual repeatability in median log <sub>2</sub> ejaculate to soma coverage ratio, lmer (N=15 ejaculates)	Var(Male identity)	5*	1.78	-	-	<0.001
	Var(Residual)	-	0.04	-	-	-
Matriline repeatability in median log <sub>2</sub> ejaculate to soma coverage ratio, lmer (N=15 ejaculates)	Var(Male identity)	5*	0.08	-	-	0.012
	Var(Matriline ID)	2*	2.85	-	-	0.017
	Var(Residual)	-	0.04	-	-	-
Between matriline difference in median log <sub>2</sub> ejaculate to soma coverage ratio, lmer (N=15 ejaculates)	Var(Male identity)	5*	0.08	-	-	0.05
	Var(Residual)	-	0.04	-	-	-
	Intercept (matriline A)	8†	0.84	0.18	4.59	<0.001
	matriline B	7†	2.40	0.28	8.44	<0.001
Individual repeatability in log <sub>2</sub> ejaculate to soma coverage ratio, lmer (N=26,023 windows)	Var(Window identity)	1742*	0.14	-	-	<0.001
	Var(Ejaculate identity)	15*	0.02	-	-	<0.001
	Var(Male identity)	5*	0.92	-	-	<0.001
	Var(Residual)	-	0.04	-	-	-
Matriline repeatability in log <sub>2</sub> ejaculate to soma coverage ratio, lmer (N=26,023 windows)	Var(Window identity)	1742*	0.14	-	-	<0.001
	Var(Ejaculate identity)	15*	0.02	-	-	<0.001
	Var(Male identity)	5*	0.05	-	-	<0.001
	Var(Matriline identity)	2*	1.47	-	-	0.013
	Var(Residual)	-	0.04	-	-	-
Between matriline difference in log <sub>2</sub> ejaculate to soma coverage ratio, lmer (N=26,023 windows)	Var(Window identity)	1742*	0.14	-	-	<0.001
	Var(Ejaculate identity)	15*	0.02	-	-	<0.001
	Var(Male identity)	5*	0.05	-	-	0.047
	Var(Residual)	-	0.04	-	-	-
	Intercept (matriline A)	8†	-0.81	0.13	-6.02	<0.001
matriline B	7†	1.73	0.21	8.27	<0.001	
Family A, reduction of median log <sub>2</sub> germline to soma coverage ratio in ejaculates comparing to testis, lm (N=15 ejaculates)	Intercept (tissue testis)	2‡	3.37	0.26	12.83	<0.001
	tissue ejaculate	8‡	-2.56	0.29	-8.73	<0.001
Family B, reduction of median log <sub>2</sub> germline to soma coverage ratio in ejaculates comparing to testis, lm (N=15 ejaculates)	Intercept (tissue testis)	1‡	4.48	0.37	12.19	<0.001
	tissue ejaculate	7‡	-1.21	0.39	-3.07	0.022

\*N levels for the random effect.

†N ejaculates in the matriline.

‡N samples in the tissue type.

§P values in mixed-effect models were estimated via ANOVA comparison (using ML method) between models with and without the focal effect. In linear models, P values were estimated in the model.

**Table S4.** Identified genes with GRC paralogs that were private to the GRC of the *castanotis x guttata* hybrid, but not found previously in Kinsella et al. 2019 (1). EnsGene table was extracted from UCSC table browser. We used the four identified *guttata* GRC-specific regions from SI Appendix, Table S5 to search against taeGut1 (2) database. Gene names were identified from ensemblToGeneName table from UCSC table browser.

Gene name	N bins	Ensembl ID	Chr	Transcription start (bp)	Transcription end (bp)	Coding region start (bp)	Coding region end (bp)	N exons	Score
	625	ENSTGUT00000018323.1	1	5260708	5260806	5260806	5260806	1	0
<i>mir221</i>	625	ENSTGUT00000018545.1	1	5261212	5261311	5261311	5261311	1	0
	1190	ENSTGUT00000013332.1	1	79374313	79384810	79374313	79384810	4	0
<i>mtmr2</i>	148	ENSTGUT00000013336.1	1	79425809	79464732	79425809	79464732	15	0
	1191	ENSTGUT00000013338.1	1	79468408	79484293	79468408	79484293	12	0
<i>pleg1</i>	630	ENSTGUT00000004995.1	20	5914462	5958118	5914462	5958118	39	0
<i>zhx3</i>	630	ENSTGUT00000005004.1	20	5966798	5969733	5966798	5969733	4	0
<i>lpin3</i>	630	ENSTGUT00000005061.1	20	6018567	6026350	6018567	6026350	26	0
<i>emilin3</i>	631	ENSTGUT00000005067.1	20	6033348	6042229	6033348	6042229	4	0
<i>emilin3</i>	631	ENSTGUT00000005078.1	20	6033381	6039014	6033381	6039014	7	0
<i>rpn2</i>	631	ENSTGUT00000005113.1	20	6088659	6130880	6088659	6130880	15	0
	594	ENSTGUT00000015806.1	4_random	1300114	1302349	1300114	1302349	4	0
	595	ENSTGUT00000015805.1	4_random	1326408	1327522	1327522	1327522	2	0

References:

1. C. M. Kinsella, et al., Programmed DNA elimination of germline development genes in songbirds. *Nat. Commun.* 10, 5468 (2019).
2. W. C. Warren, et al., The genome of a songbird. *Nature* 464, 757–762 (2010).

**Table S5.** Identified *guttata* GRC content where the putative *guttata* GRC (of the *castanotis* x *guttata* hybrid; see Supplementary Results) was highly diverged from the 12 *castanotis* GRCs (i.e. *guttata-castanotis* shared), or private to the putative *guttata* GRC (Fig. 3). Because there is no GRC reference assembly available, the genomic regions were based on A chromosomes of *taeGut1* (1). Note that due to the small testis of the *castanotis* x *guttata* hybrid (hence low amount of the GRCs, see also Materials and Methods and SI Appendix, Table S2), we only focused on GRC-amplified regions (i.e. with coverage support).

Category	Chr	Start and end (bp) including 1-kb flanking regions	Evidence	N <sup>high confidence</sup> SNPs private to the <i>guttata</i> testis	N <sup>high confidence</sup> testis-specific SNPs shared between the <i>guttata</i> testis and the pool of 12 <i>castanotis testes</i>	Start and end of the testis-specific SNP (bp)	Log <sub>2</sub> testis-to-soma coverage-ratio cutoff	length (bp)
<i>guttata</i> GRC specific	1	5224000-5298000	Coverage and SNP	23	40	5225598-5294718	1	74,000
<i>guttata</i> GRC specific	1	79381000-79490000	Coverage and SNP	12	14	79384699-79478740	2	109,000
<i>guttata</i> GRC specific	20	5944974-6091000	Coverage and SNP	29	74	5945974-6096940	0.5	146,026
<i>guttata</i> GRC specific	4_ran dom	1239000-1425000	Coverage and SNP	55	86	1240765-1423755	1	186,000
<i>guttata-castanotis</i> shared	1	104919000-107339000	Coverage and SNP	312	5535	104920645-107309548	-	2,420,000

\*Testis specific SNPs were identified with  $\geq 3$  reads support (see Materials and Methods for details).

Reference:

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

**Table S6.** Estimated fixed and random effects from mixed-effect models (lmer) for pairwise distance (number of substitutions per site) matrices between the mtDNA haplotype tree and the associated GRC haplotype tree (concatenated sequence of the two single-copy genes *pim3GRC* and *bicc1GRC*) from *castanotis* males A to I. Fixed effect and the response variable were standardized for better model performance. Note that the pairwise distance matrices of the mtDNA and the GRC haplotypes are not correlated, indicating different evolutionary histories. For additional details on the phylogenetic trees see Materials and Methods and also legend of Fig. 4.

Effects	Estimate	SE	Z	P§
Var(Matriline ID in column)	1.78	-	-	<0.001
Var(Matriline ID in row)	0.04	-	-	<0.001
Var(Residual)	0.08	-	-	-
Intercept (Distance mtDNA)	-0.31	0.33	-0.95	-
Distance GRC haplotype	-0.03	0.06	-0.45	0.66

**Table S7.** Estimated genetic diversity and number of mutational steps (i.e. N SNPs) of the mtDNA, the three low-copy-number GRC paralogs (gene<sub>GRC</sub>), their A-chromosomal paralogs (gene<sub>A</sub>), using the alignments that excluded gaps (for details see Materials and Methods and SI Appendix, Materials and Methods).

No.	Alignment content	Alignment content subgroup 1	Alignment content subgroup 2	N sequences subgroup 1	N sequences subgroup 2	Alignment length (bp)	N SNPs *	N SNPs/kb*	$\pi$ †
1	mtDNA	mtDNA	-	14	-	16629	224	13.5	0.00281
2	<i>bicc1<sub>GRC</sub></i>	<i>bicc1<sub>GRC</sub></i>	-	14	-	13822	8	0.6	0.0002
3	<i>pim3<sub>GRC</sub></i>	<i>pim3<sub>GRC</sub></i>	-	14	-	3286	1	0.3	0.00004
4	<i>elavl4<sub>GRC</sub></i> haplotype 1	<i>elavl4<sub>GRC</sub></i> haplotype 1	-	14	-	17492	1	0.1	0.00003
5	<i>elavl4<sub>GRC</sub></i> haplotype 2	<i>elavl4<sub>GRC</sub></i> haplotype 2	-	14	-	17492	30	1.7	0.00042
6	<i>bicc1<sub>A</sub></i>	<i>bicc1<sub>A</sub></i>	-	28	-	13822	833	60.3	0.01128
7	<i>pim3<sub>A</sub></i>	<i>pim3<sub>A</sub></i>	-	28	-	3286	51	15.5	0.00223
8	<i>elavl4<sub>A</sub></i>	<i>elavl4<sub>A</sub></i>	-	28	-	17492	642	36.7	0.0064
9	<i>bicc1<sub>GRC</sub></i> and <i>bicc1<sub>A</sub></i>	<i>bicc1<sub>GRC</sub></i>	<i>bicc1<sub>A</sub></i>	14	28	13822	2378	172.0	-
10	<i>pim3<sub>GRC</sub></i> and <i>pim3<sub>A</sub></i>	<i>pim3<sub>GRC</sub></i>	<i>pim3<sub>A</sub></i>	14	28	3286	229	69.7	-
11	<i>elavl4<sub>GRC</sub></i> haplotypes 1, 2 and <i>elavl4<sub>A</sub></i>	<i>elavl4<sub>GRC</sub></i> haplotypes 1 and 2	<i>elavl4<sub>A</sub></i>	28	28	17492	2082	119.0	-
12	<i>elavl4<sub>GRC</sub></i> haplotypes 1 and 2	<i>elavl4<sub>GRC</sub></i> haplotype 1	<i>elavl4<sub>GRC</sub></i> haplotype 2	14	14	17492	39	2.2	-

\*N SNPs and N SNPs/kb of alignments No. 9-12 were the estimated number of mutational steps between the cluster of sequences in subgroup 1 and subgroup 2 (see SI Appendix, Fig. S8).

†Mean number of pairwise differences per site ( $\pi$ ) was calculated in DnaSP v6.12.01.

**Table S8.** Primer sequences of the four overlapping long-range PCR that covered the whole zebra finch mitogenome (1). Long-range PCR was done using Phusion High Fidelity Taq polymerase (ThermoScientific) following the manufacturer's instruction.

Primer ID	Primer sequence	Expected PCR product size (bp)	tm (eurofin)	Primer pair
tguM2_4F	TGATCCTAACCTCCGCAATC	4632	57.3	tguM2
tguM2_8R	RCCTTGGAATGTGCTTTCTYG	4632	57.9	tguM2
tguM3_8F	AGCCTTYCCCCTATGACTTG	4596	58.3	tguM3
tguM3_12R	ATRTCCTGTTCGCCGTTTAG	4596	56.3	tguM3
tguM4_12F	AGCCTTCCTCCACATCTCAA	4920	57.3	tguM4
tguM4_0R	GGCTCGATTGTCCCCTTTA	4920	57.3	tguM4
tguM5_0F	CGAGAACTACGAGCACTAAC	3982	57.3	tguM5
tguM5_4R	GTGGGAGATGGATGAGAAG	3982	56.7	tguM5

Reference:

1. 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).

**Table S9.** Identified introgressed *castanotis* tracks on the *guttata* A chromosomes of the captive *castanotis* x *guttata* hybrid (i.e. male cas x gut). For details see Fig. 3 and SI Appendix, Results. Chromosomal positions were based on taeGut1 (1) and genetic positions (in centimorgan, cM) were estimated from Backström et al. 2010 (2).

Chromosome	Start (bp)	End (bp)	Start (cM)	End (cM)	Length (Mb)	Length (cM)
1	87500000	108500000	28.2	46.1	21.0	17.9
13	500000	16962381	1.6	19.4	16.5	17.8
15	12500000	14000000	44.4	47.9	1.5	3.5
19	9000000	11000000	46.1	61.4	2.0	15.4
26_random	500000	1000000	-	-	0.5	-
26	1500000	3000000	15.3	30.6	1.5	15.3
27	0	3000000	0	32.5	3.0	32.5
4	0	8500000	0	11.8	8.5	11.8
4_random	500000	3000000	-	-	2.5	-
5	1000000	16500000	4.86	33.15	15.5	28.3
5	57500000	60000000	49.60	67.46	2.5	17.9
5_random	0	2000000	-	-	2.0	-
6	27000000	30500000	49.0	52.7	3.5	3.7
7	10500000	34000000	16.1	24.9	23.5	8.7
8_random	4000000	4500000	NA	NA	0.5	-
9	11500000	25500000	27.2	51.3	14.0	24.1
total	-	-	-	-	118.5	196.9
mean	-	-	-	-	7.4	16.4

References:

1. W. C. Warren, et al., The genome of a songbird. *Nature* 464, 757–762 (2010).
2. N. Backström, et al., The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Res.* 20, 485–95 (2010).



## General discussion

### Searching for causes of reproductive failure

My dissertation attempted to find the cause of reproductive failure in the model system of captive zebra finches. One might typically think that reproductive failure (e.g. infertility and embryo mortality) is due to inbreeding or due to environmental factors, such as pollutants, virus and parasite infection, e.g. (Jackson et al. 2011). However, inbreeding explained only little of the variance in reproductive success (**Chapter 1**), and environmental factors cannot easily explain the striking differences in rates of reproductive failure among group-living individuals in a controlled environment. Infectious diseases should easily spread in the closed environment and pollution should also influence the group as a whole. Genetic differences in disease resistance or pollution tolerance should lead to rapid adaptation as selection on reproductive performance is strong in every generation. Although our captive zebra finches were housed under the same conditions (e.g. as a group of individuals in the same aviary), certain individuals repeatedly produce infertile eggs or dead embryos while the others managed to raise all the eggs to become recruits (**Chapter 1**). Then one might ask, are certain poor conditions experienced by an individual (Amos et al. 2001; Pemberton et al. 2009; Kraft et al. 2019) or its ancestors (i.e. transgenerational effect (Bonduriansky and Crean 2018; Engqvist and Reinhold 2018)), affecting its reproductive failure? I first ruled out transgenerational effects, which are absent or completely negligible in this species (**Chapter 2**), as has also been argued to be the case in humans (Horsthemke 2018). I found some proximate factors (**Chapter 1**) that were significantly correlated with the rates of infertility and embryo mortality (mean  $r$  ranges from -0.06 to -0.12). These factors include aging, low body mass at eight days old and inbreeding. But together, they explained maximally 3% of the observed variation in reproductive performance (**Chapter 1**). Next, one would suggest to search for genetic causes of the repeated reproductive failure. It is puzzling to blame the lack of genetic adaptation to the captive environment, because selection should quickly fix beneficial alleles, particularly those alleles that directly affect reproductive output. Lack of genetic variability could hardly be a reason for the lack of ongoing adaptation, since most captive populations carry typically 10 or more alleles across most of the genome. Inspired by this puzzle, my **Chapters 1, 3 & 4** tried to study reproductive failure from a genetic perspective.

### The genetic architecture of fitness-related traits in the zebra finch

### *Additive genetic effects*

Fitness describes the ability of an individual to propagate its genes. Fitness is often estimated through a wide range of complex traits that relate to reproductive performance, including infertility and embryo mortality. To study the genetic basis of complex traits, in my **Chapter 1**, I first estimated the heritability of all fitness-related traits from captive zebra finches, using a quantitative genetics approach. The fundamental theorem of natural selection (Fisher 1930) suggests that traits close to fitness (e.g. reproductive performance) show low heritability due to the rapid fixation of beneficial alleles. I found that all fitness-related traits shown low heritability (median 7% of phenotypic variation), confirming the theorem. Similarly low heritability of fitness has previously been reported from collared flycatchers *Ficedula albicollis* from the wild (Merilä and Sheldon 2000). My findings hence suggest some small additive genetic component behind zebra finch reproductive performance (**Chapter 1**). My **Chapter 1** additionally found that male-specific traits, especially infertility, tend to be negatively correlated with female-specific and offspring traits (including embryo mortality) at the additive genetic level. These suggest that some of the standing additive-genetic variation in fitness-related traits can be maintained via antagonistic pleiotropy between the sexes (sexual antagonism). Besides the small amount of additive genetic variation, the majority (median 71%) of the individually repeatable variation in reproductive performance remains unexplained, suggesting potentially an important role of segregating genetic incompatibility.

### *Effects of supergenes- chromosomal inversions*

The maintenance of genetic variation typically involves certain forms of balancing selection, including heterosis and antagonistic pleiotropic effects. These balancing effects may influence different life-history traits, including infertility and embryo mortality, casting a potentially unavoidable link between genetic variation and the variation in reproductive failure. In search for such genetic elements, my **Chapters 3 & 4** focused on the how the genotypic effects of the segregating supergenes – large chromosomal inversions – influences the fitness-related traits.

Up until now, inversions have been identified as an important element in explaining some striking intra-specific phenotypic variants (Lowry and Willis 2010; Küpper et al. 2015; Lamichhaney et al. 2015; Tuttle et al. 2016; Mérot et al. 2020). However, the mechanisms that maintain the many other inversion polymorphisms are still unknown (Wellenreuther and Bernatchez 2018). In the zebra finch, our group previously found four large chromosomal inversions, mostly on macrochromosomes (larger than 20 Mb), that segregate both in the wild and in captivity (Knief et al. 2016), also see (Itoh and

Arnold 2005; Itoh et al. 2011). Among the four inversions, only *TguZ* inversion polymorphism was found to show heterosis in males, whereby heterozygous males had increased siring success (Kim et al. 2017; Knief et al. 2017). Yet, the fitness consequence of the remaining three large inversions (on *Tgu5*, *Tgu11* and *Tgu13*) as well as the presence and absence of inversions on zebra finch microchromosomes were still unclear.

In my **Chapter 3**, I combined linked-reads and conventional short-read sequencing data to study the microchromosomes as well as the evolution of chromosomal inversions in the zebra finch. I found that two microchromosomes, *Tgu26* and *Tgu27* contain chromosomal inversions that both span about half (~3 Mb) of the chromosome. Additionally, I found that the inversion on *Tgu27* had significant heterotic effects, where heterozygous individuals performed significantly better in all fitness-related traits (**Chapter 3**). In my **Chapter 4**, I studied the previously detected inversion polymorphism on *Tgu11* (Knief et al. 2016). I found that it had opposing effects of different fitness components, where individuals that carried the derived allele had higher siring success and fecundity but individuals that are homozygous for the derived inversion had lower survival rate. Interestingly, the derived allele on *Tgu11* causes a rare (additive) beneficial effect, and hence the allele frequency rises (**Chapter 4**). But as often, the derived allele that links 3-12 Mb of DNA seems to contain recessive (mildly) deleterious effects, which prevent the inversion from going to fixation (**Chapters 3 & 4**). My findings suggest that chromosomal inversions (e.g. 3 Mb) are large mutational target that can easily contain loci with effects on complex traits, e.g. beneficial effects on reproductive performance and negative effects on individual survival.

### *Summary*

In sum, my findings in **Chapters 3 & 4** suggest that, segregating inversion polymorphisms may be maintained by a net heterosis effect on fitness-related traits. Perhaps it is the typical situation that the segregating inversions do not have large phenotypic effects but rather show only small effects on fitness-related traits. However, the small genotypic effects (**Chapters 1, 3 & 4**) suggest that the key genetic cause for reproductive failure might be due to epistasis (e.g. incompatibility between different loci) or sit somewhere outside of the regular chromosomes (mitochondrial genome, autosomes and sex-chromosomes).

Notably, in my **Chapter 1** I found that infertility is a male-specific problem, whereas embryo mortality is influenced by both the genetic mother and the particular combination of the two genetic parents. And these findings suggest that: (a) the genetic effects on infertility and embryo mortality might

happen at three places, i.e. the male testes, the female ovary and the developing embryo. (b) The death of embryo might involve genetic incompatibility (Orr 1996). Genetic incompatibility could be evolutionary stable if certain combinations of genotypes perform better than others, and such incompatibility may arise between nuclear loci or between the nuclear and the mitochondrial genomes (Zeh and Zeh 2005). Using the currently available data, I couldn't identify any such combination of loci in my dissertation. Therefore, I moved on to search for genetic elements that sit outside of the regular chromosomes.

### **The germline-restricted chromosome**

Interestingly, the zebra finch does have an odd genetic element that I studied as part of my search for genetic causes of infertility and embryo mortality, namely the germline-restricted chromosome (GRC). Only recently, the GRC was found to be an accessory and probably indispensable chromosome across all songbirds examined to date, that is only present in the germline cells but is absent from all somatic tissue (Pigozzi and Solari 1998; Torgasheva et al. 2019). Hence, the genes on the GRC may only be expressed in male testis, female ovary (Kinsella et al. 2019), and in the primordial germ cells in the developing embryo.

Besides the cytogenetic observations (Del Priore and Pigozzi 2014; Torgasheva et al. 2019; Malinovskaya et al. 2020) and the characterization of the gene content of the zebra finch GRC (Biederman et al. 2018; Kinsella et al. 2019), there is little knowledge regarding the elimination efficiency of GRC during spermatogenesis, the strictness of the proposed matrilineal inheritance and its intra-specific genetic variation. My **Chapter 5** offered the first study to address those issues in the zebra finch.

#### *Occasional paternal inheritance of the GRC*

Using both cytogenetic and sequencing methods, I found GRCs in some of the sperm heads in zebra finch ejaculates (**Chapter 5**). Intriguingly, the proportion of GRC-carrying sperm in the ejaculate was significantly repeatable for individual males, and differed repeatedly between males from different families (1% versus 9% of the sperm; **Chapter 5**). Additionally, when comparing the GRC-haplotypes between different individuals, we found that one hybrid male (that carried *castanotis* mitochondria, with 5% *castanotis* autosomal DNA and 95% *guttata* autosomal and sex-chromosomal DNA). Following the common belief that the GRC is only transmitted from the mothers (Goday and Pigozzi 2010; Schoenmakers et al. 2010; Del Priore and Pigozzi 2014), one would expect to find a *castanotis*

GRC in this hybrid male. Unexpectedly, we found that this *castanotis* x *guttata* hybrid carried a novel GRC haplotype that was different from all other pure *castanotis* GRCs (**Chapter 5**). Hence, I concluded that probably this is the unknown *guttata* GRC, that was inherited from a pure *guttata* father. These findings suggest that at least some GRC haplotypes can occasionally be transmitted from the fathers, and such occasional paternal transmission ability might differ between different GRC haplotypes (or at least between the families that we studied). In general, GRC haplotypes that can both be transmitted maternally and occasionally paternally may convey a selfish advantage (Werren et al. 1988) compared to the strictly maternally transmitted GRC haplotypes. This means that they could potentially spread in a population, even if they had detrimental effects on individual fitness.

### *Intra-specific variation*

To study the intra-specific variation of the zebra finch GRC, in **Chapter 5**, I sequenced and compared 12 GRCs from nine different matriline, including one sample from the wild. I found that despite of high divergence of the mitochondrial haplotypes, the GRC haplotypes (17 Kb GRC-linked sequence, in single copy) were very similar among each other. Interestingly, I found that a few diagnostic SNPs were shared between GRC-haplotypes that came from very diverged matriline (mitochondrial haplotypes; **Chapter 5**), which further confirms the idea of at least occasional paternal inheritance (lack of linkage to matriline). The extremely low genetic diversity of GRC haplotypes suggests that, possibly, one GRC haplotype had spread relatively recently across the whole zebra finch population crossing matrilineal boundaries, via the ability of occasional paternal spill-over. My findings in **Chapter 5** open up new avenues, not only for the growing field of research on the songbird germline-restricted chromosome, but also for the study of songbird biology, including the potential linkage between the GRC and reproductive failure.

### **Outlook**

My **Chapters 1 - 4** ruled out the proximate factors, additive genetic effects and supergenes to be the key cause of infertility and embryo mortality. My findings from **Chapter 5** suggest that the germline-restricted chromosome (GRC) haplotypes that can occasionally be transmitted from the fathers would have the potential to evolve as a selfish genetic element. The selfish GRC haplotypes may raise in allele frequency while imposing detrimental effects on individual fitness. Interestingly the GRCs are only present in the germline tissues, suggesting them to be a prospective candidate of the cause of the unavoidable reproductive failure. My dissertation also opens many more directions beyond the study of zebra finch reproductive failure. For instance, **Chapter 5** only studied the elimination efficiency of

males from two families. Future follow-ups on examining the population-level variation in GRC elimination during spermatogenesis as well as the genetic (or epigenetic) basis would bring more insights into the inheritance pattern of the GRC. One could also predict that paternally spreading haplotypes convey fitness costs to the organism (e.g. reduced fertility), otherwise the system would be unstable and quickly reach an equilibrium with little variance in fitness. Therefore, studying the fitness consequence between males with different GRC-elimination efficiencies or different GRC haplotype would help with explaining the observed phenotypic and genotypic variations. Additionally, studies on the complete GRC assemblies for both zebra-finch subspecies may shed more light on GRC evolution.

### **Beyond the zebra finch**

#### *Meta-summarization of effect sizes - turning multiple testing into a strength*

Throughout my dissertation, I worked with many statistical tests (up to 1000 tests in **Chapter 2**) for various combinations of traits and focal fixed effects. My dissertation shows that, multiple testing can be turned into a strength. That is, start from clear, one-tailed hypotheses (Ruxton and Neuhäuser 2010); then meta-summarize all effects (e.g. from a certain focal factor on all fitness-related traits; **Chapters 1 - 4**) to obtain an average effect size estimate (with very narrow confidence interval) and lastly -if possible- verify the average effect in an independent dataset (**Chapter 2**).

#### *Avian microchromosomes matter*

Within a species, chromosomal inversions are typically identified through (a) cytogenetic observations such as ‘pericentric inversions’ that change the position of the centromere, e.g. (Bailey et al. 1996; Itoh and Arnold 2005; Itoh et al. 2011; Hooper and Price 2017), (b) large phenotypic effects when studying the genetic basis for remarkable morphological variants, e.g. (Küpper et al. 2015; Lamichhaney et al. 2015; Tuttle et al. 2016) and more recently (c) bioinformatic analysis such as genome-wide scans using a large number of SNPs and a large number of individuals, e.g. (Knief et al. 2016; Mérot et al. 2021). In the study of the chromosomal inversions, avian microchromosomes were traditionally omitted, presumably due to their small size and being acrocentric (Burt 2002). My **Chapter 3** shows that, avian microchromosomes may also contain large chromosomal inversions that may be maintained by heterosis.

*The more we know, the more we know we don't know*

Given the rich body of literature, are there any interesting questions that haven't been answered already in a model species? Before starting my dissertation, I had this thought in mind. I guess it is normal to think that model species such as the zebra finch are relatively boring to study. However, my own dissertation proved me to be wrong. In **Chapter 3**, using linked-read and short-read sequencing data, I found two microchromosomal inversions in the zebra finch and one of them showed consistent heterosis for nearly all fitness-related traits. In **Chapter 5**, I found, for the first time, that the special accessory germline-restricted chromosome can be present in some of the sperm heads, and can be inherited occasionally from the father as well. It is the fast advancement of sequencing technology and the open science community that made these new discoveries possible. Given the new discoveries, new questions arrive consequently (see **Outlook**).

### **Significance**

I here present a comprehensive dissertation on the evolutionary genetics of reproductive performance in captive zebra finches. My dissertation highlights the meta-summarization framework to be readily applicable for statistical testing in biology, which turns multiple testing from a burden into a strength. Along with the identified genetic factors such as chromosomal inversions that correlate with infertility and embryo mortality, my findings also suggest the evolutionary importance of the previously understudied microchromosomes. Last but not least, I present the peculiar occasional paternal transmission ability of the germline-restricted chromosome in the zebra finch, which opens new possibilities in songbird research and in the field of the evolution of selfish genetic elements.

## References

- Amos, W., J. Worthington Wilmer, K. Fullard, T. M. Burg, J. P. Croxall, D. Bloch, and T. Coulson. 2001. The influence of parental relatedness on reproductive success. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268:2021–2027.
- Bailey, S. M., J. Meyne, M. N. Cornforth, T. S. McConnell, and E. H. Goodwin. 1996. A new method for detecting pericentric inversions using COD-FISH. *Cytogenetic and Genome Research* 75:248–253.
- Biederman, M. K., M. M. Nelson, K. C. Asalone, A. L. Pedersen, C. J. Saldanha, and J. R. Bracht. 2018. Discovery of the first germline-restricted gene by subtractive transcriptomic analysis in the zebra finch, *Taeniopygia guttata*. *Current Biology* 28:1620–1627.e5.
- Bonduriansky, R., and A. J. Crean. 2018. What are parental condition-transfer effects and how can they be detected? *Methods in Ecology and Evolution* 9:450–456.
- Burt, D. W. 2002. Origin and evolution of avian microchromosomes. *Cytogenetic and Genome Research* 96:97–112.
- Del Priore, L., and M. I. Pigozzi. 2014. Histone modifications related to chromosome silencing and elimination during male meiosis in Bengalese finch. *Chromosoma* 123:293–302.
- Engqvist, L., and K. Reinhold. 2018. Adaptive parental effects and how to estimate them: A comment to Bonduriansky and Crean. *Methods in Ecology and Evolution* 9:457–459.
- Fisher, R. A. 1930. *The genetical theory of natural selection*. Clarendon Press, Oxford.
- Goday, C., and M. I. Pigozzi. 2010. Heterochromatin and histone modifications in the germline-restricted chromosome of the zebra finch undergoing elimination during spermatogenesis. *Chromosoma* 119:325–336.
- Hooper, D. M., and T. D. Price. 2017. Chromosomal inversion differences correlate with range overlap in passerine birds. *Nature Ecology & Evolution* 1:1526–1534.
- Horsthemke, B. 2018. A critical view on transgenerational epigenetic inheritance in humans. *Nature Communications* 9:2973.
- Itoh, Y., and A. P. Arnold. 2005. Chromosomal polymorphism and comparative painting analysis in the zebra finch. Pages 47–56 *in* *Chromosome Research* (Vol. 13).
- Itoh, Y., K. Kampf, C. N. Balakrishnan, and A. P. Arnold. 2011. Karyotypic polymorphism of the zebra finch Z chromosome. *Chromosoma* 120:255–264.
- Jackson, A. K., D. C. Evers, M. A. Etterson, A. M. Condon, S. B. Folsom, J. Detweiler, J. Schmerfeld, et al. 2011. Mercury exposure affects the reproductive success of a free-living terrestrial songbird, the Carolina Wren (*Thryothorus ludovicianus*). *The Auk* 128:759–769.
- Kim, K.-W., C. Bennison, N. Hemmings, L. Brookes, L. L. Hurley, S. C. Griffith, T. Burke, et al. 2017. A sex-linked supergene controls sperm morphology and swimming speed in a songbird. *Nature Ecology & Evolution* 1:1168–1176.
- Kinsella, C. M., F. J. Ruiz-Ruano, A.-M. Dion-Côté, A. J. Charles, T. I. Gossmann, J. Cabrero, D. Kappei, et al. 2019. Programmed DNA elimination of germline development genes in songbirds. *Nature Communications* 10:5468.
- Knief, U., W. Forstmeier, Y. Pei, M. Ihle, D. Wang, K. Martin, P. Opatová, et al. 2017. A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nature Ecology & Evolution* 1:1177–1184.
- Knief, U., G. Hemmrich-Stanisak, M. Wittig, A. Franke, S. C. Griffith, B. Kempenaers, and W. Forstmeier. 2016. Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biology* 17:199.
- Kraft, F.-L. O. H., S. C. Driscoll, K. L. Buchanan, and O. L. Crino. 2019. Developmental stress reduces body condition across avian life-history stages: A comparison of quantitative magnetic resonance data and condition indices. *General and Comparative Endocrinology* 272:33–41.

- Küpper, C., M. Stocks, J. E. Risse, N. Dos Remedios, L. L. Farrell, S. B. McRae, T. C. Morgan, et al. 2015. A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics* 48:79–83.
- Lamichhaney, S., G. Fan, F. Widemo, U. Gunnarsson, D. S. Thalmann, M. P. Hoepfner, S. Kerje, et al. 2015. Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nature Genetics* 48:84–88.
- Lowry, D. B., and J. H. Willis. 2010. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biology* 8.
- Malinovskaya, L. P., K. S. Zadesenets, T. V. Karamysheva, E. A. Akberdina, E. A. Kizilova, M. V. Romanenko, E. P. Shnaider, et al. 2020. Germline-restricted chromosome (GRC) in the sand martin and the pale martin (*Hirundinidae*, Aves): synapsis, recombination and copy number variation. *Scientific Reports* 10:1–10.
- Merilä, J., and B. C. Sheldon. 2000. Lifetime Reproductive Success and Heritability in Nature. *The American Naturalist* 155:301–310.
- Mérot, C., E. L. Berdan, H. Cayuela, H. Djambazian, A.-L. Ferchaud, M. Laporte, E. Normandeau, et al. 2021. Locally Adaptive Inversions Modulate Genetic Variation at Different Geographic Scales in a Seaweed Fly. (D. Charlesworth, ed.) *Molecular Biology and Evolution*.
- Mérot, C., V. Llaurens, E. Normandeau, L. Bernatchez, and M. Wellenreuther. 2020. Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. *Nature Communications* 11:670.
- Orr, H. A. 1996. Dobzhansky, Bateson, and the Genetics of Speciation. *Genetics* 144:1331–1335.
- Pemberton, J. M., L. E. B. Kruuk, A. Morris, T. H. Clutton-Brock, M. N. Clements, and D. H. Nussey. 2009. Inter- and intrasexual variation in aging patterns across reproductive traits in a wild red deer population. *The American Naturalist* 174:342–357.
- Pigozzi, M. I., and A. J. Solari. 1998. Germ cell restriction and regular transmission of an accessory chromosome that mimics a sex body in the zebra finch, *Taeniopygia guttata*. *Chromosome Research* 6:105–113.
- Ruxton, G. D., and M. Neuhäuser. 2010. When should we use one-tailed hypothesis testing? *Methods in Ecology and Evolution* 1:114–117.
- Schoenmakers, S., E. Wassenaar, J. S. E. Laven, J. A. Grootegeod, and W. M. Baarends. 2010. Meiotic silencing and fragmentation of the male germline restricted chromosome in zebra finch. *Chromosoma* 119:311–324.
- Torgasheva, A. A., L. P. Malinovskaya, K. S. Zadesenets, T. V. Karamysheva, E. A. Kizilova, E. A. Akberdina, I. E. Pristyazhnyuk, et al. 2019. Germline-restricted chromosome (GRC) is widespread among songbirds. *Proceedings of the National Academy of Sciences* 116:201817373.
- Tuttle, E. M., A. O. Bergland, M. L. Korody, M. S. Brewer, D. J. Newhouse, P. Minx, M. Stager, et al. 2016. Divergence and functional degradation of a sex chromosome-like supergene. *Current Biology* 26:344–350.
- Wellenreuther, M., and L. Bernatchez. 2018. Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends in Ecology and Evolution* 33:427–440.
- Werren, J. H., U. Nur, and C. I. Wu. 1988. Selfish genetic elements. *Trends in Ecology and Evolution* 3:297–302.
- Zeh, J. A., and D. W. Zeh. 2005. Maternal inheritance, sexual conflict and the maladapted male. *Trends in Genetics* 21:281–286.

The more you know, the more you know you don't know.

- Aristotle



## Addresses of co-authors

### **Juan D. Alché**

Department of Biochemistry, Cell & Molecular Biology of Plants, Estación Experimental del Zaidín, Spanish National Research Council (CSIC), Profesor Albareda 1, E-18008 Granada, Spain  
Email: [juandedios.alche@eez.csic.es](mailto:juandedios.alche@eez.csic.es)

### **Stefan Börno**

Sequencing Core Facility, Max Planck Institute for Molecular Genetics, Ihnestraße 63-73, 14195 Berlin, Germany  
Email: [boerno@molgen.mpg.de](mailto:boerno@molgen.mpg.de)

### **Josefa Cabrero**

Department of Genetics, University of Granada, E-18071, Granada, Spain  
Email: [jcabrero@ugr.es](mailto:jcabrero@ugr.es)

### **Juan Pedro M. Camacho**

Department of Genetics, University of Granada, E-18071, Granada, Spain  
Email: [jpmcamac@ugr.es](mailto:jpmcamac@ugr.es)

### **Anne-Marie Dion-Côté**

Département de Biologie, Université de Moncton, New Brunswick, Canada  
Email: [anne-marie.dion-cote@umoncton.ca](mailto:anne-marie.dion-cote@umoncton.ca)

### **Wolfgang Forstmeier**

Department of Behavioural Ecology & Evolutionary Genetics, Max Planck Institute for Ornithology Seewiesen, Germany  
Email: [forstmeier@orn.mpg.de](mailto:forstmeier@orn.mpg.de)

### **Andre Franke**

Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany  
Email: [a.franke@mucosa.de](mailto:a.franke@mucosa.de)

### **Ivana Gessara**

Department of Behavioural Neurobiology, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany  
Email: [igessara@orn.mpg.de](mailto:igessara@orn.mpg.de)

### **Moritz Hertel**

Department of Behavioural Neurobiology, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany  
Email: [hertel@orn.mpg.de](mailto:hertel@orn.mpg.de)

### **Marc Hoepfner**

Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany  
Email: [m.hoepfner@ikmb.uni-kiel.de](mailto:m.hoepfner@ikmb.uni-kiel.de)

**Bart Kempnaers**

Department of Behavioural Ecology & Evolutionary Genetics, Max Planck Institute for Ornithology  
Seewiesen, Germany

Email: [b.kempnaers@orn.mpg.de](mailto:b.kempnaers@orn.mpg.de)

**Ulrich Knief**

Division of Evolutionary Biology, Faculty of Biology, Ludwig Maximilian University of Munich,  
Großhaderner Straße 2, D-82152 Planegg-Martinsried, Germany

Email: [knief@biologie.uni-muenchen.de](mailto:knief@biologie.uni-muenchen.de)

**Katrin Martin**

Department of Behavioural Ecology & Evolutionary Genetics, Max Planck Institute for Ornithology  
Seewiesen, Germany

Email: [k.martin@orn.mpg.de](mailto:k.martin@orn.mpg.de)

**Jakob C. Mueller**

Department of Behavioural Ecology & Evolutionary Genetics, Max Planck Institute for Ornithology  
Seewiesen, Germany

Email: [mueller@orn.mpg.de](mailto:mueller@orn.mpg.de)

**Francisco J. Ruiz-Ruano**

Department of Organismal Biology, Evolutionary Biology Centre (EBC), Science for Life  
Laboratory, Uppsala University, SE-752 36, Uppsala, Sweden

Email: [francisco.ruiz-ruano@ebc.uu.se](mailto:francisco.ruiz-ruano@ebc.uu.se)

**Joanna Rutkowska**

Institute of Environmental Sciences, Faculty of Biology, Jagiellonian University, 30387 Krakow,  
Poland

Email: [joanna.rutkowska@uj.edu.pl](mailto:joanna.rutkowska@uj.edu.pl)

**Alexander Suh**

School of Biological Sciences, University of East Anglia, Norwich NR4 7TU, UK &  
Department of Organismal Biology, Evolutionary Biology Centre (EBC), Science for Life  
Laboratory, Uppsala University, SE-752 36, Uppsala, Sweden

Email: [alexander.suh@ebc.uu.se](mailto:alexander.suh@ebc.uu.se)

**Kim Teltscher**

Department of Behavioural Ecology & Evolutionary Genetics, Max Planck Institute for Ornithology  
Seewiesen, Germany

Email: [teltscher@orn.mpg.de](mailto:teltscher@orn.mpg.de)

**Daiping Wang**

Department of Behavioural Ecology & Evolutionary Genetics, Max Planck Institute for Ornithology  
Seewiesen, Germany

Email: [dwang@orn.mpg.de](mailto:dwang@orn.mpg.de)

**Jochen Wolf**

Division of Evolutionary Biology, Faculty of Biology, Ludwig Maximilian University of Munich,  
Großhaderner Straße 2, D-82152 Planegg-Martinsried, Germany

Email: [j.wolf@biologie.uni-muenchen.de](mailto:j.wolf@biologie.uni-muenchen.de)

## Author contributions

**Chapter 1:** W.F. and B.K. designed and planned the study. W.F., D.W., and K.M. collected reproductive performance data. Y.P. and W.F. analysed the data, with input from J.R. Y.P., W.F., and B.K. interpreted the results and wrote the manuscript, with input from J.R. All authors contributed to the final manuscript.

**Chapter 2:** W.F. and B.K. designed the study. W.F. collected the morphological data. Y.P. and W.F. analysed the data and interpreted the results with input from B.K. Y.P. and W.F. wrote the manuscript with help from B.K.

**Chapter 3:** Y.P., W.F. and B.K. conceived the study. U.K. designed the tag SNPs. Y.P. analysed the data with inputs from W.F., J.W., U.K., A.S. and A.M.D.C. Y.P. interpreted the results with inputs from W.F., J.W., U.K., A.S. and B.K. Y.P. wrote the manuscript with inputs from W.F., U.K. and B.K.

**Chapter 4:** Y.P., W.F. and B.K. conceived the study. Y.P. analysed the data with inputs from W.F. and U.K. Y.P. and W.F. interpreted the results with inputs from U.K. and B.K. Y.P. and W.F. wrote the manuscript with inputs from U.K. and B.K.

**Chapter 5:** Conceptualization: Y.P., W.F., F.J.R.R., A.S. and B.K. Ejaculates sampling and sequencing design: Y.P., U.K. and W.F. with inputs from J.C.M. Testis and liver sampling and sequencing design: W.F., A.S. and Y.P. Ejaculates collection: Y.P. Tissue dissection: Y.P., K.T. and M.H. DNA extraction: K.T. PCR-free sequencing: A.F. and M.H. Mitogenome amplicon sequencing design: Y.P. and W.F. with inputs from J.C.M. and U.K. Primer design: Y.P. Mitogenome sequencing and de-multiplexing: S.B. Sequencing data analysis: Y.P. with help from F.J.R.R. and inputs from W.F., A.S. and J.C.M. FISH: Y.P. with help from F.J.R.R. and inputs from J.C. and J.P.M.C. FISH counting: I.G. Confocal microscopy: J.P.M.C. and J.D.A. Cytogenetic interpretation: Y.P., W.F. with help from F.J.R.R., J.C., J.P.M.C. and M.H. Results interpretation: Y.P., W.F. and

A.S. with inputs from F.J.R.R. Funding acquisition: B.K. and A.S. Illustration: Y.P. with input from B.K. First draft of manuscript: Y.P. Writing: Y.P., W.F., A.S. and B.K., with inputs from all authors.

.....

Bart Kempenaers, Doktorvater

.....

Yifan Pei, Doktorand

## Acknowledgement

Firstly, I would like to acknowledge my PhD supervisors, Wolfgang Forstmeier and Bart Kempnaers.

My deepest thank goes to you, Wolfgang, for offering me this fascinating system to work with, for guiding me to be careful, efficient and independent in research, for your patient, critical and intellectual discussions with me, for not losing faith in me and for mentoring me throughout my PhD. I'm mostly inspired by your enthusiasm and wide yet in detailed knowledge as a researcher and a naturalist.

Bart, I'm very grateful to you, for allowing me to join this fantastic group, for your critical mind that helped me making decisions, for your great intellectual edits and comments to all my manuscripts, for trusting me and for always being there to support me.

Secondly, I thank my PhD advisory committee members, Jochen Wolf and Ulrich Knief, for your time traveling to Seewiesen, for your constructive discussions and feedbacks on my projects.

My thank also goes to my collaborators, Joanna Rutkowska, Anne-Marie Dion-Côté, Josefa Cabrero, Juan Pedro. M. Camacho, Juan D. Alché, Stefan Börno, especially, Alexander Suh, Francisco J. Ruiz-Ruano, Moritz Hertel and Jakob C. Mueller for your help and discussions whenever I was in doubt.

I would like to thank people behind the data collection, without whom my PhD would not be possible: Melanie Schneider, Christine Baumgartner, Kim Teltscher for the molecular work, Katrin Martin, Holger Schielzeth, Elisabeth Bolund, Malika Ihle, Ulrich Knief, Daiping Wang, Claudia Burger, Sanja Janker, Johannes Schreiber, Tibaud Aronson, and Anais Edme for help with data collection, and Sonja Bauer, Edith Bodendorfer, Jane Didsbury, Annemarie Grötsch, Johann Hacker, Andrea Kortner, Jenny Minshull, Petra Neubauer, Claudia Scheicher, Frances Weigel, and Barbara Wörle for animal care and help with breeding, Tim Birkhead, Tobias Krause, John Lesku, and Kylie Roberts for providing birds. Especially, I thank Katrin and Melanie, whom I indebt the most for the great amount of work that you have been doing behind my dissertation.

My PhD wouldn't be that smooth without the amazing members in the department and at the Institute that helped me with solving issues in one way or another: Mihai, Lotte, Emmi, Agnes, the Peters (1-

3), Sylvia, Christine, Alex, Melanie, Katrin for being the most amazing office mate, Carmen and Uschi for helping me with paper works, Mäggi and Corinna for helping me with IMPRS-related issues.

Additionally, I would like to thank those people that built up all the light-hearted colourful memories during my PhD. Yaoyao's friends: Keren and Gabriel for inviting me to the 'Ghetto Mensa', Eunbi, Mekhla, Ale and Esha for the international meals and you accompany while walking to bus stops, Qiaoyi for sharing with me your passion of plants and making delicious food, Gianina for introducing me your mom's 'holy oil' and jam, Annie and Rowan for the board games, Pame, Elena, Giulia, Pietro and Safari for catsitting Yaoyao. Ivy and Kim for checking with me the shining cells in the darkroom. Hannes for the infinite number of car-rides and with Maggie, Carol, Peter, P3 and the dogs (Nuna and Pippi) for the BBQs and the flickering charcoals in the summer evenings. Kristina for 'hosting' my broken bike in Maising for about half a year. Esteban and Nils for being the first to encourage me to try scientific illustration. Maggie (Ko) and Luisana for the many last-minute reminders. Glenn, Neetash, Thejasvi, Julia, Daniel, Shouwen, Amanda ... for chats here and there. Elisavet for our adventures in Athens and Poland. My friends since the old time: Junyi, Bei and your sister for hosting me in Beijing and Shanghai and for visiting me here, Xue (Luo) for cheering me up. Also, the cat, Yaoyao, for catching invisible 'creatures' to protect me and for brightening up my days.

Lastly, I owe a great debt of thanks to my parents, Jianzhang Pei and Jing Fan, for their unconditional love and support to me and for them being openminded to all my decisions in life.

Yifan Pei  
August 2021  
Seewiesen

## Curriculum vitae

**Yifan Pei** (裴一凡)

### EDUCATION

- 2016.07 – present      **PhD candidate, International Max Planck Research School for Organismal Biology (IMPRS), Max-Planck-Institut für Ornithologie-Seewiesen, Germany**  
Thesis: Evolutionary genetics of reproductive performance in the zebra finch  
Supervisors: Dr. Wolfgang Forstmeier & Prof. Dr. Bart Kempenaers
- 2014.10 – 2016.10      **MSc, Evolution, Ecology and Systematics (EES), Ludwig-Maximilians-Universität München (LMU), Germany**  
Thesis: Interference competition pressure predicts the number of avian predators that shifted their timing of activity  
Supervisors: Dr. Mihai Valcu & Prof. Dr. Bart Kempenaers
- 2009.09 - 2013.07      **BSc, Marine Science, Shanghai Ocean University, China**  
Thesis: Laboratory observations on the foraging behaviour of the Japanese common squid, *Todarodes pacificus*  
Supervisors: Prof. Dr. Xinjun Chen & Prof. Dr. Yasunori Sakurai:

### PUBLICATIONS

- Pei, Y.**; Forstmeier, W.; Wang, D.; Martin, K.; Rutkowska, J.; Kempenaers, B.: Proximate causes of infertility and embryo mortality in captive zebra finches. *American Naturalist* 196 (5), pp. 577 - 596 (2020)
- Pei, Y.**; Forstmeier, W.; Kempenaers, B.: Offspring performance is well buffered against stress experienced by ancestors. *Evolution: International journal of organic evolution* 74 (7), pp. 1525 - 1539 (2020)
- Pei, Y.**; Valcu, M.; Kempenaers, B.: Interference competition pressure predicts the number of avian predators that shifted their timing of activity. *Proceedings of the Royal Society B: Biological Sciences* 285 (1880), 20180744 (2018)
- Knief, U.; Forstmeier, W.; **Pei, Y.**; Wolf, J.; Kempenaers, B.: A test for meiotic drive in hybrids between Australian and Timor zebra finches. *Ecology and Evolution* 10 (23), pp. 13464 - 13475 (2020)
- Knief, U.; Forstmeier, W.; **Pei, Y.** et al.: A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nature Ecology & Evolution* 1 (8), pp. 1177 - 1184 (2017)

### PREPRINT

- Pei, Y.**; Forstmeier, W.; Francisco J. R. R.; et al.: Occasional paternal inheritance of the germline-restricted chromosome in songbirds. *bioRxiv* <https://doi.org/10.1101/2021.01.28.428604>

## DATA & SCRIPTS

**Pei, Y.:** Supporting alignments for “Occasional paternal inheritance of the germline-restricted chromosome in songbirds”, available on figshare at <https://figshare.com/s/d4a1cfa1c6c126fc15b2> (2021)

**Pei, Y.:** Supporting information for "Offspring performance is well buffered against stress experienced by ancestors", DOI 10.17605/OSF.IO/WJG3Q, available on Open Science Framework at <https://osf.io/wjg3q/> (2019)

**Pei, Y.:** Supporting information for "Proximate causes of infertility and embryo mortality in captive zebra finches", DOI 10.17605/OSF.IO/TGSZ8, available on Open Science Framework at <https://osf.io/tgsz8/> (2019)

**Pei, Y.:** Supporting information for "Interference competition pressure predicts the number of avian predators that shifted their timing of activity" <https://osf.io/2rfm7/> (2018)

## PRESENTATIONS

Lightning talk and poster, Virtual symposium on within-individual genome variation and germline/soma distinction (SMBE), 2020

Talk, Workshop on causes and consequences of inclusive inheritance, Plön, Germany, 2019

Poster, European society for evolutionary biology (ESEB), Turku, Finland, 2019

Poster, European society for evolutionary biology (ESEB), Groningen, the Netherlands, 2017

## SCHOLARSHIPS AND AWARDS

- |                |  |
|----------------|--|
| 2020           | <b>Best Student Paper Award</b> , issued from the IMPRS for organismal biology                             |
| 2012 -<br>2013 | <b>JASSO Gakushu-Shoreihi Scholarship</b> , issued from the Japan Student Services Organization            |
| 2010 -<br>2011 | <b>National Scholarship (top 2 in cohort within A.Y.)</b> , issued from the Ministry of Education of China |

## Statutory declaration and statement

### Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt ist.

München, den 30.08.2021

.....

Yifan Pei

### Erklärung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist. Des Weiteren erkläre ich, dass ich mich nicht anderweitig einer Doktorprüfung ohne Erfolg unterzogen habe oder ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich einer Doktorprüfung zu unterziehen.

München, den 30.08.2021

.....

Yifan Pei