



This is a repository copy of *Linkage mapping of a polymorphic plumage locus associated with intermorph incompatibility in the Gouldian finch (Erythrura gouldiae)*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/106558/>

Version: Accepted Version

Article:

Kim, K-W., Griffith, S.C. and Burke, T. orcid.org/0000-0003-3848-1244 (2016) Linkage mapping of a polymorphic plumage locus associated with intermorph incompatibility in the Gouldian finch (*Erythrura gouldiae*). *Heredity*, 116 (4). pp. 409-416. ISSN 0018-067X

<https://doi.org/10.1038/hdy.2015.114>

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Published in Kim et al. 2016. *Heredity* 116, 409-416

(DOI: 10.1038/hdy.2015.114)

Author Accepted Manuscript (there may be slight differences with published one)

**Linkage mapping of a polymorphic plumage locus associated with inter-morph
incompatibility in the Gouldian finch (*Erythrura gouldiae*)**

Kang-Wook Kim¹, Simon C. Griffith² and Terry Burke¹

¹Department of Animal and Plant Sciences, University of Sheffield, Sheffield, United Kingdom;

²Department of Biological Sciences, Macquarie University, Sydney, Australia.

Correspondence: Kang-Wook Kim

Department of Animal & Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

Tel +44 (0)114 222 0113

Fax +44 (0)114 222 0002

Email: kangwook.kim.ego@gmail.com

Abstract

Colour polymorphism is known to facilitate speciation but the genetic basis of animal pigmentation and how colour polymorphisms contribute to speciation is poorly understood. Restricted recombination may promote linkage disequilibrium between the colour locus and incompatibility genes. Genomic rearrangement and the position of relevant loci within a chromosome are important factors that influence the frequency of recombination. Therefore, it is important to know the position of the colour locus, gene order and recombination landscape of the chromosome to understand the mechanism that generates incompatibilities between morphs.

Recent studies showed remarkable pre- and post-zygotic incompatibilities between sympatric colour morphs of the Gouldian finch (*Erythrura gouldiae*), in which head feather colour is genetically determined by a single sex-linked locus, Red. We constructed a genetic map for the Z chromosome of the Gouldian finch (male-specific map distance = 131 cM), using 618 captive-bred birds and 34 microsatellite markers, to investigate the extent of inter- and intra-specific genomic rearrangements and variation in recombination rate within the Z chromosome. We refined the location of the Red locus to a ~7.2-cM interval in a region with a moderate recombination rate but outside the least-recombining, putative centromeric region. There was no evidence of chromosome-wide genomic rearrangements between the chromosomes carrying the red or black alleles with the current marker resolution. This work will contribute to identifying the causal gene, which will in turn enable alternative explanations for the association between incompatibility and colouration, such as fine-scale linkage disequilibrium, genomic rearrangements and pleiotropy, to be tested.

Keywords: sex linkage, colour polymorphism, genetic incompatibility, linkage map.

Introduction

Colour polymorphisms are often associated with other physiological and behavioural traits (McKinnon and Pierotti, 2010), which may generate incompatibilities between morphs, so facilitating speciation (Hugall and Stuart-Fox, 2012; Poelstra et al, 2014). One of the explanations for these associations is pleiotropy, in which the colour polymorphism is controlled by transcription factors or hormones that influence multiple physically unlinked genes (McKinnon and Pierotti, 2010). An alternative hypothesis for the association is that there are linkage disequilibria (LD) among the loci that control these traits. Recombination is a key process that counteracts the formation of LD between loci. To date, various mechanisms that might alter the rate of recombination have been proposed, including inversion (Rieseberg, 2001), the centromere effect (Carneiro et al, 2009) and recombination hot-spots (Singhal et al, 2015), which collectively generate highly variable recombination landscapes between and within species (Backström et al, 2010a; Auton et al, 2012).

For example, chromosome-wide LD caused by complex rearrangements and reduced recombination in the second chromosome of white-throated sparrows (*Zonotrichia albicollis*) generates the association between plumage polymorphism and social behaviour (Thornycroft, 1966; Tuttle, 2003; Huynh et al, 2011). Similarly, the three alternative, phenotypically complex and distinctly plumaged morphs of the ruff *Philomachus pugnax* are determined by a large inversion and the multiple genes combined within such an inversion may act together as alleles of a supergene (Küpper et al, 2015; Lamichhaney et al, 2015).

In the zebra finch (*Taeniopygia guttata*; hereafter ZF), the macrochromosomes show greatly reduced recombination in the central part of each chromosome and recombination is concentrated towards the distal parts of the chromosomes, but these extreme patterns are not found in the microchromosomes (Stapley et al, 2010). In contrast, the orthologous macrochromosomes in chicken and flycatcher show moderate recombination rates in the more central regions, although there is again a tendency for recombination rates to be elevated towards the chromosome ends (Kawakami et al, 2014; Singhal et al, 2015). If a colour locus resides in a recombination desert, such as the one that spans ~50 Mb in the central region of the Z chromosome of ZF (Stapley et al, 2010), incompatibility loci in the same region might potentially be coupled with the colour locus. Thus, in order to

understand the mechanism underlying any incompatibility between morphs, it is necessary to study the genomic organization and the recombination landscape of the respective chromosome, as well as locating the position of the gene that causes the polymorphism.

The Gouldian finch (*Erythrura gouldiae*; hereafter GF) is an ideal species in which to study colour polymorphism in relation to incompatibility between morphs. The GF belongs to the Estrildidae, the same family as ZF, and is estimated to have separated from ZF 12.8 million years ago (Jetz et al, 2012). This was one of the first avian species in which a plumage colour polymorphism was studied using classical genetics (Southern, 1945). A single sex-linked Mendelian locus (hereafter, the Red locus) with dominant red (Z^R) and recessive black (Z^r) alleles determines the expression of head colour (Southern, 1945). In females, the heterogametic sex in birds, the phenotype matches the hemizygous genotype (Z^R , red; Z^r , black), whereas males can be homozygous Z^rZ^r (black), Z^RZ^R (red) or heterozygous Z^RZ^r (red). This polymorphic trait attracts special attention because the sex chromosome is expected to have a disproportionately large influence on the speciation process, either by selection or drift (Qvarnström and Bailey, 2008; Mank et al, 2010), and colour polymorphisms are often subject to sexual selection (Wellenreuther et al, 2014). Moreover, the Red locus has been shown to be associated with both pre- and post-zygotic incompatibilities between morphs in a captive population (Pryke and Griffith, 2007; Pryke and Griffith, 2009b; Pryke, 2010). However, it has not been confirmed whether incompatibility occurs in the wild population and a lack of knowledge on the genetic architecture of the polymorphism hampers our understanding of the underlying cause of the reported incompatibilities.

In the present study, we examined two large-scale factors that might potentially cause the genetic incompatibility associated with head colour by constructing a genetic map for the Z chromosome of the GF and localizing the Red locus on the map. First, we examined whether the Red locus is located in a genomic region with reduced recombination, such as the centromere or a large recombination desert as is found in the ZF. Second, we compared the gene order between Z chromosomes containing the alternative alleles at the Red locus to test for genomic rearrangements between chromosomes determining the two morphs.

Materials and Methods

Samples

Blood samples (n = 618) were collected from a captive population that was originally established with wild-type, but domesticated, birds from 23 aviculturists throughout Australia (Pryke and Griffith, 2006) and stored in 100% ethanol at room temperature prior to DNA extraction using the ammonium-acetate precipitation method (Bruford et al, 1998).

Marker development and genotyping

Previously, we developed a large number of microsatellite markers for GF in various ways (Kim et al, 2015), including Sanger and 454 sequencing of microsatellite-enriched libraries, in silico primer design based on ZF microsatellite sequences found in the reference genome, *Taeniopygia_guttata*-3.2.4 (Warren et al, 2010), and testing microsatellite primers designed in other species (Supplementary Table S1). In total, 39 sex-linked markers were selected and multiplex sets were developed and tested for genotyping. The flanking sequences of the cross-species and library-derived microsatellites were used to search for the physical positions of markers in the ZF genome assembly with the **blastn** option (e-value < e-10) in a locally installed BLAST program.

DNA (~20 ng) was dried down in PCR plates in preparation for the PCR reaction. Two-microlitre PCR reactions were set up containing 1 µl of QIAGEN Multiplex PCR Kit (QIAGEN, UK) and 1 µl of diluted primers (0.2 µM each). PCR amplification was performed using DNA Engine Tetrad PTC-225 Peltier thermal cyclers (MJ Research, BIO-RAD, Hemel Hempstead). PCR products were separated on an ABI 3730 DNA Analyser and genotypes were scored using GENEMAPPER v.3.7 (Applied Biosystems).

For the sequencing, PCR products were first cleaned with ExoSAP-IT (USB). After the sequencing reaction was completed using a BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems), and the reaction products precipitated using the ethanol/EDTA method, the products were separated on an ABI 3730 DNA Analyser. Sequences were aligned and edited using CodonCode Aligner v. 2.0.6 (CodonCode Corporation).

Genotype error correction

A Perl program (CriMaker; Dryad doi: @@@@/@@@@) was developed to compile the genotypic data into CRI-MAP format and to identify inconsistencies in the Mendelian inheritance of markers. Initial error checking revealed 1804 genotypic mismatches between parents and offspring, but these were caused mostly by a few individuals, who were concluded to have false pedigree information (the birds had been bred in aviaries containing social groups, so some misidentification of parents was possible), and some error-prone markers. By removing 27 individuals that caused errors across the markers, and four markers that caused errors across families, the number of errors dropped to 245. These errors consisted of inherited null alleles, 1-bp binning mismatches or rare mutations, and all could be corrected manually by comparing the pedigree and genotypes. The remaining 34 loci were used in the subsequent analysis (Supplementary Table S1).

Construction of Z-chromosome genetic map and linkage mapping of the Red locus

Two- or three-generation half-sib families in which 73 fathers each mated with 1–4 females were used to construct the genetic linkage map of the Z chromosome. If a parent's DNA was not available then genotypes were, whenever possible, inferred from the offspring's genotypes. The genotype of the mother (ZW) could be inferred with certainty because a maternal allele on the Z chromosome should be seen in all sons but not in daughters, whereas the paternal alleles could be transmitted to both sons and daughters. The sex of each individual was determined with two sex-linked markers, Z-002D (Dawson, 2007) and Z37B (Dawson et al, 2015), that each amplified alleles on both the Z and W chromosomes. The results agreed perfectly with other markers that showed only a single allele in females and one or two alleles in homozygous and heterozygous males, respectively. Pedigree information and the head colour after moulting were recorded, but due to a high mortality rate in mixed-morph breeding of offspring from embryos up to the age at which the trait is expressed in six to nine month old juveniles (Pryke and Griffith, 2009b), only 65.4% of samples had phenotypic data available (404/618). The mother's genotype at the colour locus could be inferred directly from their phenotype, whereas the genotype of red-headed males had to be inferred from the pedigree

information. For example, a black male (always Z^rZ^r) breeding with a red female (Z^R) will produce only red heterozygous male offspring, while a heterozygous red male (Z^RZ^r) breeding with a black or red female (Z^r) can produce both red and black offspring.

To determine the marker order and intervals between markers, a maximum-likelihood linkage analysis was performed using CRI-MAP (Green et al, 1990) as improved by Evans and Maddox (ver. 2.503, <http://www.animalgenome.org/tools/share/crimap/>). First, a framework marker map was created using the option **build** and the positions of the remaining markers were determined with the option **all**, so adding one marker at a time. If this option gave two or more positions with non-significant difference in the \log_{10} likelihood (< 3) between positions, the order in the ZF reference genome was considered to be the most likely one. This order was then tested with options **flips2**, **flips3** and **flips4** to see if there was a better order, and an increase in \log_{10} likelihood of three or more was taken as evidence of a significantly more likely map (Morton, 1955). None of these tests produced a better order than the original. Undetected errors that introduce less-likely double recombinants were examined with the option **chrompic** and the map was recreated after correction, as described above.

In birds, recombination on the sex chromosome (other than in the pseudoautosomal regions) occurs only in the male, so the effective recombination rate on the Z chromosome should be 2/3 of the male-specific estimate (Bennett and Oertel, 1965). Caution is required when using CRI-MAP as it calculates the sex-averaged recombination rates on the sex chromosome as being half of the male-specific rate. We report male-specific maps with Kosambi map distances to account for meiosis only in the father and to compare with previous literature that used male-specific rates, but these values can be easily converted to sex-averaged rates by multiplying by 2/3. The maps were drawn using MapChart v2.2 (Voorrips, 2002).

In total, 11 families were informative for segregation at the Red locus and used to obtain two-point LOD scores between the colour locus and genetic markers. These families were either fully informative crosses between a heterozygous red male (Z^RZ^r) and a black female (Z^r), or partially informative crosses between a heterozygous red male and a red female (Z^R), where only female offspring are informative because all male offspring will have the red phenotype. The best position for

the Red locus was searched for by comparing the likelihoods of alternative map positions of the Red locus using the **all** function of CRI-MAP and an increase of three or more in \log_{10} likelihoods was considered as evidence of a significantly more likely map position.

Intraspecific genomic rearrangement

In order to test for intraspecific genomic rearrangements between morphs, the above-described mapping families were subdivided into three subsets according to the father's genotype at the colour locus based on pedigree information across several generations and their phenotypes: first, homozygous black fathers ($n=31$) that had two 'black chromosomes' (the chromosome that has the recessive black allele Z^f); second, homozygous red fathers ($n=15$) that had two 'red chromosomes' (the chromosome that has the dominant red allele Z^R); and third, heterozygous red fathers ($n=17$) that had both black and red chromosomes. Genotypes of ten fathers in the mapping families could not be determined due to the limited pedigree information we had for them and these were excluded from this analysis. Using these subsets, three separate maps were constructed and the marker orders were compared with each other.

Results

Linkage map for the Gouldian finch Z chromosome and intra-chromosomal rearrangements between species

A total of 73 families were genotyped with 34 microsatellite markers (Supplementary Table S1), giving 582 informative crossovers. First, we created a framework marker map composed of 28 markers and determined the positions of six remaining markers to obtain the best-order map; this had a mean inter-marker spacing of 4 cM (0–31.9 cM) and a total male-specific map length of 131 cM (Supplementary Figure S1, Figure 1a). This map corresponded to 97.2% of the physical map length of the ZF Z chromosome, assuming the physical size of the Z chromosome is conserved in both species because the same terminal markers in GF (TguE103 and TguE157) were also located at the ends of the chromosome in ZF (Figure 1a).

Genetic map positions of the GF microsatellite markers were compared with the physical map position on the ZF reference genome assembly (Warren et al, 2010). Various degrees of intra-chromosomal rearrangement were observed between the two species and these resulted from changes in position of five chromosomal fragments that shared the same marker order between the two species (P1–P5) (Figure 1b). First, a major pericentric inversion including the region between markers Ego001 and TguE144 interchanged the positions of both telomeric regions (P1 and P5) in relation to the P4 region. Second, a small fraction of the chromosome, including only one marker (Ego010), showed a pattern of intrachromosomal transposition (Figure 1b). To verify and define the range of an intra-chromosomal translocation of this locus, primers for eleven introns of the genes flanking marker Ego010, between markers Z37B and Ego011, were tested for sequencing. Informative SNPs that were obtained from 24 individuals in 5 families using four amplifying primers showed that a chromosomal fragment with a size of up to 3.5 Mb had changed its position in GF compared to ZF (Supplementary Table S2, Supplementary Figure S2). Third, the positions of chromosomal sections P2 and P3 were different in the two species (ZF and GF). This could not be explained by a simple inversion because the gene order within each fragment was the same as in ZF, and the orientations of the fragments were

unchanged (Figure 1b). However, the marker order in these regions was not resolved well because of the small number of informative crossovers.

Despite large-scale genomic rearrangements between species, marker orders within chromosomal fragments were conserved between the two species. For example, the P4 region, the largest chromosomal fragment in the q-arm spanning 35.8Mb in ZF between Ego005 and TguE144, showed almost the same order of markers, except for Ego010 (Figure 1a).

The rate of recombination

The overall recombination rate of the GF Z chromosome was estimated by comparing the genetic distance between the endmost markers in the GF Z chromosome with the physical distance of the corresponding positions of the ZF Z chromosome. Assuming a similar physical size of the Z chromosome in the two species, the estimated physical distance between the endmost markers (TguE157 and TguE103) in GF was about 70.9 Mb and the overall recombination rate was 1.85 (131/70.9) cM/Mb.

Linear regression of GF linkage map position against the ZF physical map position was performed for each chromosomal fragment to obtain the slopes as estimates of regional recombination rates. There was considerable local variation in the rate of recombination among five chromosomal fragments (Figure 2). High recombination rates were found in the terminal sections: 3.4 cM/Mb and 11.8 cM/Mb in P1 and P5, respectively. In particular, the P5 region, spanning just 3.5 Mb in ZF, occupied almost a quarter of the genetic distance of the GF genetic map.

In contrast, the P2 and P3 regions showed low rates of recombination. These regions, spanning 14.7 Mb of the p-arm in ZF, corresponded to a small genetic distance (4.4 cM) in GF, resulting in low recombination rates: 0.7 and 0.2 cM/Mb in P2 and P3, respectively. This, in turn, suggests that these regions may be close to the centromere. The region spanning the greatest physical length in ZF, P4, showed an intermediate recombination rate in GF (1 cM/Mb, excluding a translocated marker, Ego010, Figure 2).

Linkage mapping of the Red locus

Linkage between the colour locus and genetic markers was examined with two-point LOD scores (Figure 3). All markers on the Z chromosome showed significant linkage with the colour locus (LOD >3.3). However, markers (Ego008~TguE78) within the region between Z-013 and Z-040, which spanned 7.2 cM (~8 Mb in ZF), showed perfect co-segregation with the phenotype (i.e. the recombination fractions between these markers and the colour locus were zero). These markers showed the highest LOD score for linkage (two-point LOD = 26.2, Figure 3). The likelihood of the map position of the Red locus was equally high across this range, while any other map position for the Red locus resulted in a significant decrease (>4.5) in the \log_{10} likelihood of the map. The Red locus was therefore mapped to the centre of the P4 region, and this candidate region showed an intermediate recombination rate (0.9 cM/Mb, Figure 2).

Intra-chromosomal rearrangement between morphs

To examine if there were genomic rearrangements between chromosomes carrying each colour allele, linkage maps for three subsets of families that each had fathers with one of the three combinations of Black- and Red-determining chromosomes (Black x Black, Black x Red and Red x Red) were constructed and the sizes of the linkage maps were 131.7, 128.1 and 123.7 cM, respectively (Supplementary Figure S3, Figure 4). If a chromosomal polymorphism were present, then it would be expected to produce different genetic maps from fathers with Black x Black and Red x Red chromosomes. In addition, the map from Black x Red would be expected to show reduced recombination in the rearranged region. The size of the linkage map was smallest in the Red x Red chromosome combination, but it was not possible to conclude that this is a property of this combination because the number of families used for the mapping of this type was smallest, allowing few observations of crossovers. Overall, there was no evidence of changes in marker order, suggesting no detectable intraspecific genomic rearrangements around the Red locus at the resolution provided by the current marker density.

Discussion

Genomic rearrangements between species

Information on genomic organization is essential for mapping a focal trait on a chromosome and for studying the effect of genomic location on the trait. Many avian species share highly conserved synteny, but with various degrees of intra-chromosomal rearrangement (Kawakami et al, 2014; Zhang et al, 2014). For example, no genomic rearrangements were found in the sex chromosomes between closely related *Ficedula* flycatcher species (Backström et al, 2010b), whereas there are karyotypic polymorphisms on the Z chromosome in ZF (Itoh et al, 2011) and an inversion is associated with alternative reproductive morphs in the ruff (Küpper et al, 2015; Lamichhaney et al, 2015). Estrildid finches have experienced several genomic rearrangements in the process of speciation. An estimation based on karyotypic banding patterns over part of this phylogenetic branch suggest that an inversion has been fixed every 2.6 My in the estrildid finches (Hooper and Price, 2015). Within the genus *Erythrura*, for example, the blue-faced parrotfinch (*Erythrura trichroa*) is separated from GF by two autosomal inversions (Christidis, 1986b; Hooper and Price, 2015). Christidis (1986a) described karyotypic polymorphism in the ZF Z chromosome based on early cytological and banding techniques. Recently, Itoh et al. (2011) rediscovered these within-species genomic rearrangements and mapped the relative positions of some genes by using fluorescent in situ hybridization (FISH) mapping (Itoh et al, 2006; Itoh et al, 2011), which showed that the current ZF reference genome sequence was based on a submetacentric Z chromosome (Zsm), whereas an alternative metacentric Z chromosome (Zm) has been created by an inversion.

The Gouldian finch is known to have a submetacentric Z chromosome (Christidis, 1986b). Although it is not possible to determine which marker should be placed at the top of the map using a genetic map alone, our illustration (Figure 1) shows the most likely orientation of the submetacentric Z chromosome of GF under the conventional definition of chromosome arms, with the P5 region belonging to the short (p) arm and P1 region being in the long (q) arm, for the following reasons. First, although the precise position of the centromere of GF needs verification, we can predict the approximate position of the centromere based on the recombination rate. The order of fragments P2 and P3 in the GF best-order map appears to have been changed compared to ZF, but this cannot be

explained by a simple inversion because the orientation of the markers inside the fragments was unchanged. It is likely that the marker order in this region has not been fully resolved because only a few informative crossovers were available due to the low recombination rate, which in turn suggests that the GF centromere may be close to this region. Moreover, the predicted physical size of P4 plus P1 is greater than the combined size of the P5, P3 and P2 regions, suggesting that the GF q arm may consist of the P4 and P1 regions. Interestingly, when the relative position of the centromere in Zsm of ZF, which was shown to be at approximately 40% of the chromosomal length from the p-terminus (Itoh et al, 2011), was applied to the reference sequence of the ZF Z chromosome, the location of the centromere in ZF was predicted to be between TguE129 and Ego005 in ZF (Figure 1b). This location is close to the P3 region, which we found to have the lowest recombination rate in GF (Figure 2).

A comparison between ZF and GF shows that a pericentric inversion has changed the relative locations of the terminal regions of the Z chromosome. To examine the changes in gene order in relation to the major inversion between the GF and ZF Z chromosome karyotypes, the positions of five genes that were used in the previous study (Itoh et al, 2011) were compared with the relative positions of the microsatellite markers used in this study (Figure 1b). Assuming no gene order changes within the five pre-defined fragments, the chromosomal fragment that contains GHR (growth hormone receptor) and DMRT1 (doublesex and mab-3 related transcription factor 1) was found on the long arm in GF and in both chromosomal variants of ZF. In ZF, SMAD2 (mothers against DPP homologue 2), HINT1 (histidine triad nucleotide binding protein 1) and SPIN (spindlin) were found in the short arm of Zsm, whereas SMAD2 and the major part of an HINT1-containing BAC probe were mapped onto the long arm, and SPIN was found in the short arm in Zm (Itoh et al, 2011). In GF, however, the chromosomal fragment P1 (between TguE103 and TguE118) that may contain HINT1, SPIN and possibly SMAD2 (distance between SMAD2 and TguE103 is 122 kb in ZF), was mapped onto the long arm. Therefore, neither of the chromosomal polymorphisms in ZF had the same gene order as GF, suggesting that the inversions in the two species had different origins.

We also found evidence for an intrachromosomal translocation that has occurred within a genomic region around Ego010 (maximum 3.5Mb), although this could have been derived from more complex genomic rearrangements (Supplementary Table S2, Supplementary Figure S2). This region

provides a good model for the future study of the relationship between genomic rearrangements and sequence evolution because protein-coding sequences may evolve faster in chromosomes that have undergone structural rearrangements (Navarro and Barton (2003). Currently, the position of Ego010 within the GF genome is not supported very well due to the sparse marker density between the P1 and P4 regions. It is possible that the change in genomic position may also have influenced the recombination rate in the translocated fragment because it was moved closer to the end of the chromosome, where the recombination rate is elevated. Interestingly, the translocated region contained a putative pigmentation gene, proteinase-activated receptor 2 precursor (PAR-2 or F2RL1), that up-regulates melanin and melanosome ingestion by increasing keratinocyte phagocytosis (Sharlow et al, 2000). Therefore, further studies in this region may provide an insight into the relative roles of selection and genomic rearrangements in the evolution of pigmentation genes.

Variation in recombination rate

Some researchers assume similar recombination landscapes across confamilial species (Hooper and Price, 2015). However, our results indicate that finch species in the same family can have very different recombination landscapes. In the ZF, an exceptionally biased recombination landscape is observed, in which the majority of recombination occurs in the distal regions, whereas large recombination deserts are found in the central parts of the macrochromosomes (Kawakami et al, 2014; Singhal et al, 2015). For example, the three largest chromosomes (TGU1–TGU3) have large recombination deserts spanning over 100 Mb, with a mean recombination rate of 0.1 cM/Mb, and the Z chromosome shows an extremely low recombination rate in the central part, causing significant LD across a 50-Mb region (Backström et al, 2010a; Stapley et al, 2010).

In the present study, we examined if this pattern was also found in GF. Indeed, a highly heterogeneous recombination rate was again found within the Z chromosome but the pattern was different to that of ZF. Based on the recombination rate and a comparison with ZF, we predicted that the GF centromere might be located close to the P3 (0.2cM/Mb) region. However, the reduced recombination did not extend beyond this region. In contrast, both termini of the chromosome showed greatly increased recombination rates that account for a large fraction of the entire genetic map,

consistent with other avian species (Kawakami et al, 2014). The P4 region that has the largest collinear gene order compared with ZF showed an intermediate level of recombination.

The mechanism that causes this variation in recombination rate is not clearly understood, but chromosome size, GC content and sequence motif have been proposed to be important factors (Kawakami et al, 2014; Singhal et al, 2015). In this study, direct comparison of genomic sequences and recombination rates between these species was not possible due to the non-availability of a genome sequence for the GF. An alternative mechanism that might cause the differences in the rate of recombination might be the described genomic rearrangement, because such a major pericentric inversion as that seen between ZF and GF would have changed the relative positions of many loci within the chromosome.

Colour-gene mapping

The extent of genomic rearrangements and the rate of recombination provide crucial information for gene mapping. The conserved marker order in the P4 section of the Z chromosome, when compared with ZF, allows us to infer that gene order is well conserved in this area. If a candidate locus were found in this region, then it should be possible to refer to the well-annotated ZF genome databases to obtain a list of candidate genes and predict their function. However, as can be seen in the transposition of locus Ego010, it is also possible that some small regions have changed their position, such that they cannot be detected with the current density of markers.

The striking differences in the rate of recombination between distal and central regions of the chromosome indicate that different strategies might be needed in different regions of the chromosome in order to identify the location of a target gene. First, more markers should be added in the distal regions to increase the resolution of recombination mapping. Second, more informative mapping families are required to resolve the high LD between markers in the central part of the Z chromosome. However, these options are not always available and are sometimes impractical. We concluded that the Red locus is in the P4 region, where the local recombination rate is about 0.9 cM/Mb. Thus, thousands of meioses might be required to observe enough informative recombinant events in order to map the locus more finely to a < 1Mb region. In the GF it is also likely to be difficult to obtain

enough samples that include adequate phenotypic information, because informative crosses between heterozygous red males and black females produce offspring of low viability in this captive population (Pryke and Griffith, 2009a; Pryke and Griffith, 2009b). Therefore, it would be difficult to obtain finer resolution than the currently identified candidate region from linkage mapping; a genome-wide association study using many more markers is required to help narrow down the candidate region.

Colour locus and mating incompatibilities

In the captive GF population also studied here, the Red locus was shown to be associated with pre- and post-zygotic reproductive incompatibilities (Pryke and Griffith, 2007; Pryke and Griffith, 2009b). In theory, the observed association may be caused either by LD between the Red locus and incompatibility genes or a pleiotropic effect of the Red locus.

In the present study, we mainly focused on two large-scale factors that might generate high LD between the Red locus and incompatibility genes. First, we examined local recombination around the Red locus. Usually, the recombination rate in the centromeric region is significantly reduced, which often results in faster genetic differentiation between diverging species (Stump et al, 2005; Carneiro et al, 2009). We identified a putative centromeric region where the recombination rate was greatly reduced below that of any other region but this region was much smaller than the recombination desert in the Z chromosome in ZF (Backström et al, 2010a; Singhal et al, 2015). The Red locus was mapped to a region with a moderate recombination rate, far from the putative centromere region, suggesting that the Red locus was not influenced by a centromere effect.

Second, we examined the possibility of the inter-morph rearrangement, such as an inversion, which may lead to genetic differentiation by reducing the fitness of heterozygotes or by reducing gene flow through the suppression of recombination (Rieseberg, 2001). We could not find any evidence of a large-scale genomic rearrangement between morphs and the marker order was highly conserved among the three maps that were constructed from different combinations of chromosomes. However, this result does not necessarily exclude the possibility of genomic rearrangement at a finer scale. For example, if a small-scale inversion were present in the region of low recombination, or in the regions

between the currently typed markers, then it would have been difficult to detect by linkage mapping without a much larger sample size.

Although not tested in this study, LD may be higher in the captive population than wild populations due to reduced effective population size and inbreeding. In particular, it is worth noting that there might have been artificial selection in captive populations in favour of high-quality head colouration (Evans and Fidler, 2005), or to produce particular colour morphs, by breeding them separately, and thus driving genetic structuring in the domesticated population. Therefore, we need to verify whether the incompatibility is a captive population-specific phenomenon, with the increased LD coupling the Red locus with incompatibility genes.

Alternatively, pleiotropy may have played an important role in the association of colour locus with incompatibility. In vertebrates, several genes involved in melanin-based colouration (e.g. Proopiomelanocortin, POMC) have been reported to be associated with sexual behaviour, aggressiveness, resistance to stressors, immune response and metabolic rate through pleiotropic effects (Ducrest et al, 2008). The red/black colour polymorphism in the Gouldian finch is mainly due to a difference in the pigments that are deposited in the feathers (Brush and Seifried, 1968). The black morphs deposit melanin, whereas red morphs deposit carotenoid in the feathers of the head. The gene(s) involved in pigment production or deposition may regulate very heterogeneous processes, which may in turn generate the incompatibility between morphs.

To conclude, there was no evidence of a large-scale reduction in recombination around the Red locus. In addition, we showed pronounced between- and within-species variation in the recombination rate within the Z chromosome. However, the Red locus was mapped to a region with a moderate degree of recombination that is greater than in the putative centromeric region, which showed the lowest recombination rate. Although there were various degrees of chromosomal rearrangement in comparison with ZF, no genomic rearrangements between the chromosomes that contain polymorphic colour alleles were detected with the current marker density. To investigate the cause of incompatibilities in the Gouldian finch, including the possibilities of a locally reduced recombination rate, fine-scale genomic rearrangements and pleiotropy, it will be necessary to identify the causal gene for the colour polymorphism using a higher-density marker approach. Our

identification of the position of the Red locus will assist in targeting such efforts and revealing the genetic basis of the colour polymorphism.

Acknowledgements

This study was supported by funding from the University of Sheffield for a studentship to K.-W.K., and from the Australian Research Council to S.C.G. and T.B. We are grateful to Sarah Pryke and the Save the Gouldian Fund for providing access to samples.

Compliance with Ethical Standards

Samples: Blood sampling of birds was conducted under the approval of the Animal Ethics Committees at the University of New South Wales and Macquarie University.

Conflict of Interest

The authors declare that they have no conflict of interest.

Data Archiving

CriMaker and Microsatellite genotypes: Dryad doi: <http://dx.doi.org/10.5061/dryad.fg7g4>.

Supporting Information

Supplementary file S1: supplementary figures and tables.

Figures

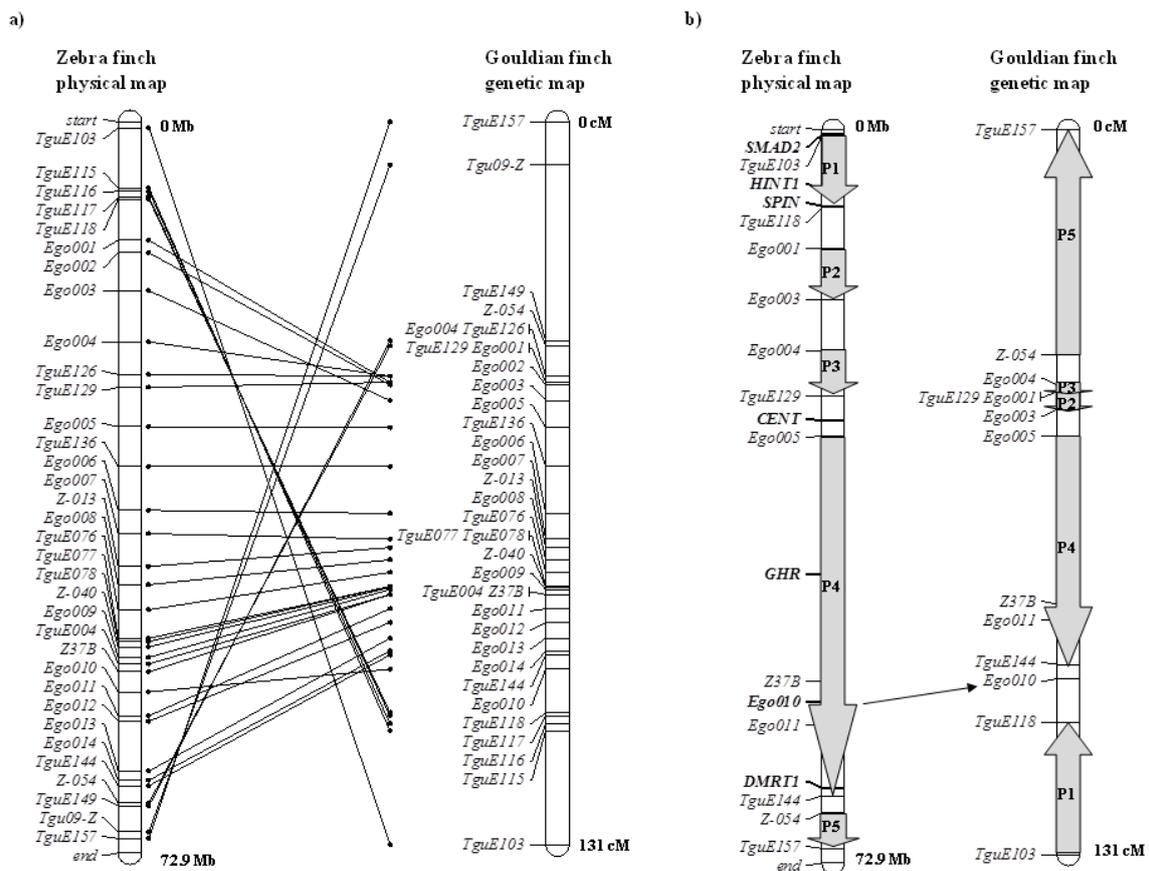


Figure 1 Inter-chromosomal rearrangements between zebra finch and the Gouldian finch. (a) The best-order genetic map positions of 34 microsatellite markers in the GF Z chromosome are compared to the physical map positions in ZF, with straight lines connecting identical marker positions. (b) Schematic overview of the inter-chromosomal rearrangements between the two species. The orientations of five chromosomal regions (P1~P5) that share an identical marker order are indicated by arrows. The positions of the putative centromere (CENT) and five genes in ZF are based on Itoh et al. (2011). The GF map is orientated to illustrate the collinear relationship between the two species in the P4 region, which has the biggest predicted physical length. In addition, it is more likely that our illustration represents the order in the submetacentric karyotype of GF, given the relative size of each region and the predicted position of the centromere in GF.

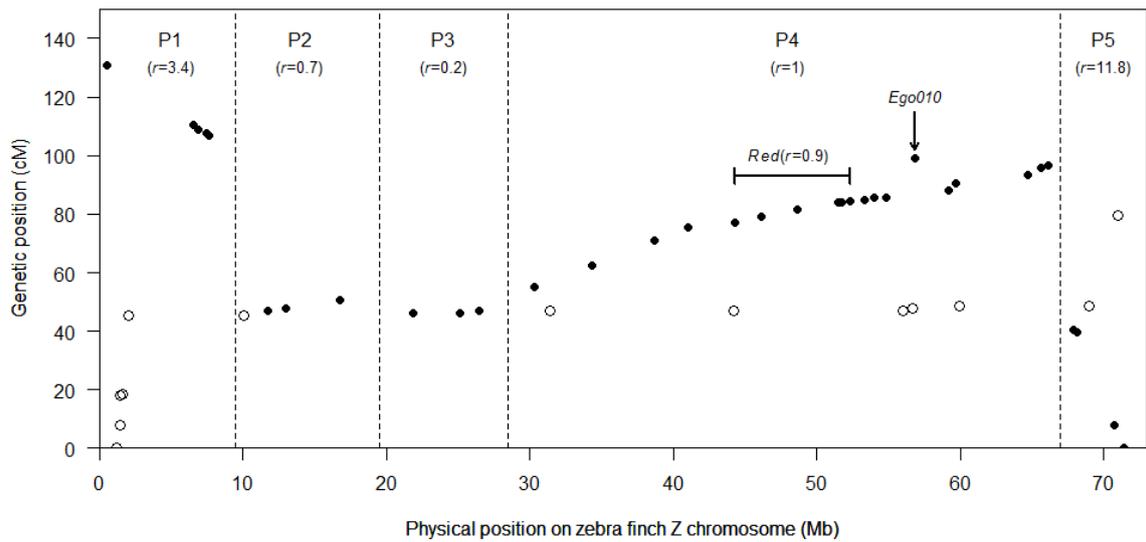


Figure 2 Linkage map position (cM) plotted against the predicted position in the zebra finch (Mb) for the Gouldian finch (filled circles) and zebra finch (open circles: data courtesy of Stapley (2010)). Vertical dotted lines divide the chromosome into five sections (P1 ~ P5; see text for details), each of which has similarity with the respective chromosomal section in the zebra finch, with the most likely order in GF being P5–P3–P2–P4–P1 (Figure 1). The horizontal bar (Z-013 ~ Z-040) shows the candidate region that contains the Red locus (see Figure 3). r , the absolute linear regression slope within each fragment represents the regional recombination rate (cM/Mb).

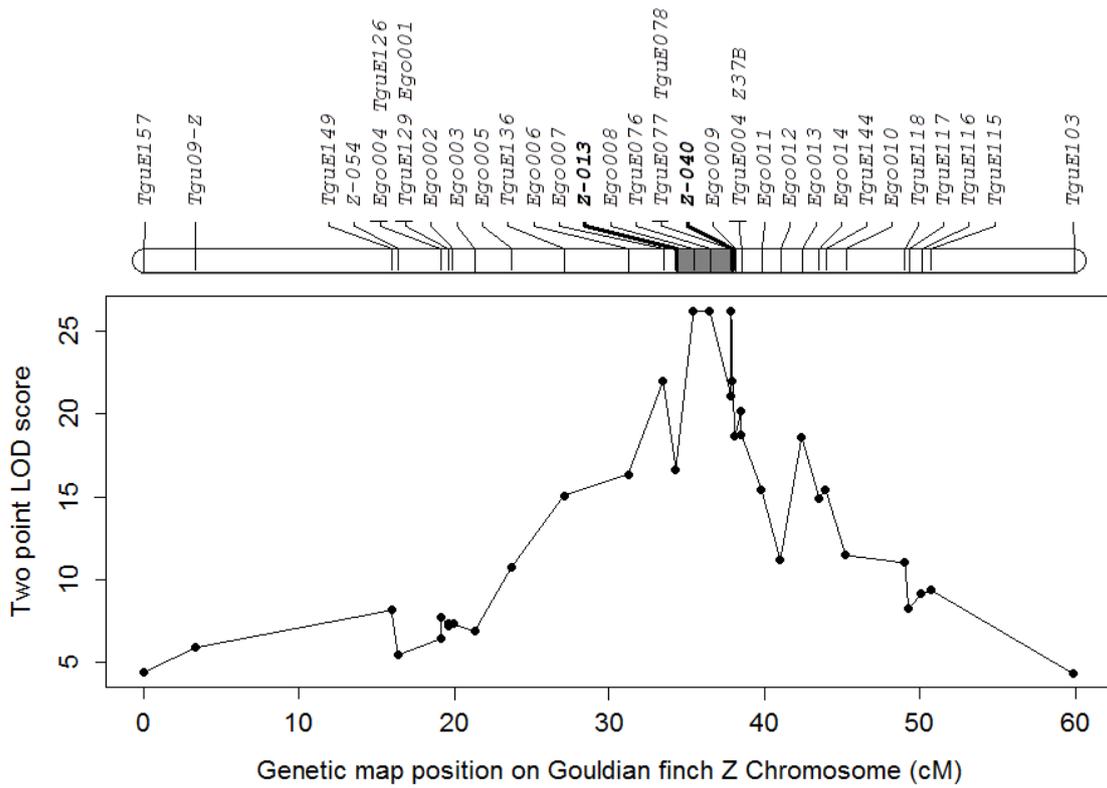


Figure 3 Male-specific two-point LOD score profiles for linkage between the Red locus and microsatellite markers plotted against Z chromosome genetic map position in the Gouldian finch. The interval that is most likely to contain the Red locus is shaded in grey between Z-013 and Z-040.

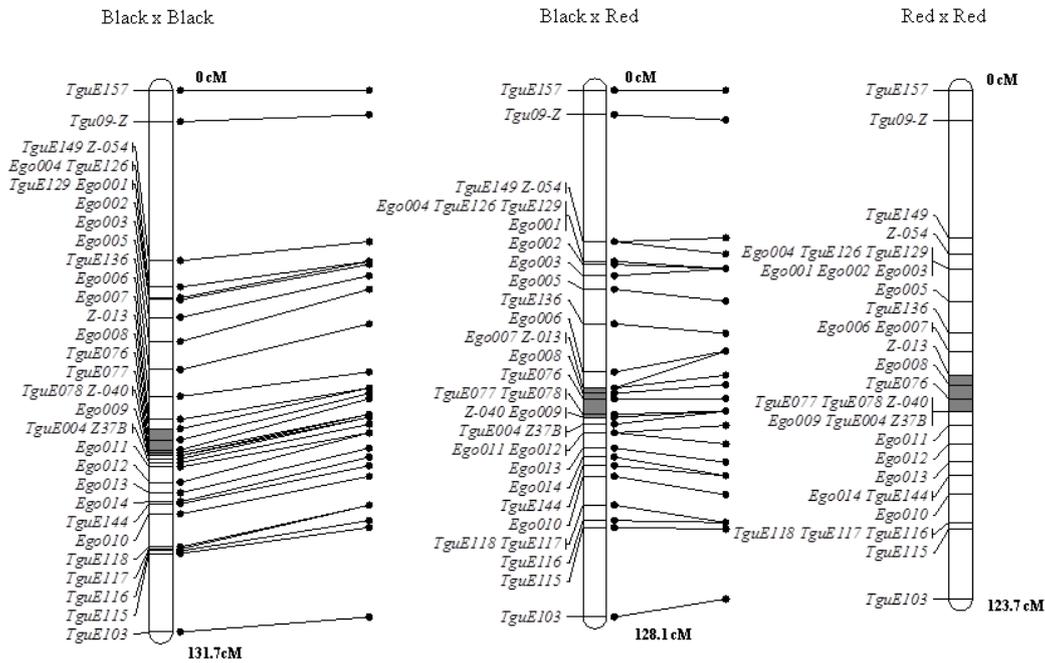


Figure 4 Comparative maps of the markers from three types of father, who each had different combinations of Black and Red chromosomes, in relation to the genotype at the Red locus: Black x Black, Red x Red, and Black x Red, where the Black and Red chromosomes denote the chromosomes that have a recessive black allele (Z^I) or the red allele (Z^R), respectively. The father's genotypes were inferred from the pedigree and the phenotypes of their offspring. The shaded areas in the bar show the most likely map position of the Red locus in the Z chromosome.

References

Auton A, Fledel-Alon A, Pfeifer S, Venn O, Ségurel L, Street T et al. (2012). A fine-scale chimpanzee genetic map from population sequencing. *Science* **336**: 193-198.

Backström N, Forstmeier W, Schielzeth H, Mellenius H, Nam K, Bolund E et al. (2010a). The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research* **20**: 485-495.

Backström N, Palkopoulou E, Qvarnström A, Ellegren H (2010b). No evidence for Z-chromosome rearrangements between the pied flycatcher and the collared flycatcher as judged by gene-based comparative genetic maps. *Molecular Ecology* **19**: 3394-3405.

Bennett JH, Oertel CR (1965). The approach to a random association of genotypes with random mating. *Journal of Theoretical Biology* **9**: 67-76.

Bruford MW, Hanotte O, Brookfield JFY, Burke T (1998). Multilocus and single-locus DNA fingerprinting. In: Hoelzel AR (ed) *Molecular genetic analysis of populations: A practical approach*. IRL Press: Oxford, U.K., pp 225-269.

Brush AH, Seifried H (1968). Pigmentation and feather structure in genetic variants of the Gouldian finch, *Poephila gouldiae*. *The Auk* **85**: 416-430.

Carneiro M, Ferrand N, Nachman MW (2009). Recombination and speciation: Loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics* **181**: 593-606.

Christidis L (1986a). Chromosomal evolution within the family Estrildidae (Aves) I. The Poephilae. *Genetica* **71**: 81-97.

Christidis L (1986b). Chromosomal evolution within the family Estrildidae (Aves) II. The Lonchurae. *Genetica* **71**: 99-113.

Dawson D, Brekke P, Dos Remedios N, Horsburgh G (2015). A marker suitable for sex-typing birds from degraded samples. *Conservation Genet Resour* **7**: 337-343.

Dawson DA (2007). *Genomic analysis of passerine birds using conserved microsatellite loci*. PhD Thesis thesis, University of Sheffield, UK.

Ducrest A-L, Keller L, Roulin A (2008). Pleiotropy in the melanocortin system, coloration and behavioural syndromes. *Trends in Ecology and Evolution* **23**: 502-510.

Evans S, Fidler M (2005). *The Gouldian Finch*. Indruss Productions: Queensland.

Green P, Falls K, Crooks S. (1990). Washington University, St. Louis.

Hooper DM, Price TD (2015). Rates of karyotypic evolution in Estrildid finches differ between island and continental clades. *Evolution* **69**: 890-903.

Hugall AF, Stuart-Fox D (2012). Accelerated speciation in colour-polymorphic birds. *Nature* **485**: 631-634.

Huynh LY, Maney DL, Thomas JW (2011). Chromosome-wide linkage disequilibrium caused by an inversion polymorphism in the white-throated sparrow (*Zonotrichia albicollis*). *Heredity* **106**: 537-546.

Itoh Y, Kampf K, Arnold A (2006). Comparison of the chicken and zebra finch Z chromosomes shows evolutionary rearrangements. *Chromosome Research* **14**: 805-815.

Itoh Y, Kampf K, Balakrishnan CN, Arnold AP (2011). Karyotypic polymorphism of the zebra finch Z chromosome. *Chromosoma* **120**: 255-264.

Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO (2012). The global diversity of birds in space and time. *Nature* **491**: 444-448.

Kawakami T, Smeds L, Backström N, Husby A, Qvarnström A, Mugal CF et al. (2014). A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. *Molecular Ecology* **23**: 4035-4058.

Kim K-W, Zhang H, Horsburgh G, Bird S, Dawson D, Griffith S et al. (2015). Four-way development of microsatellite markers for the Gouldian finch (*Erythrura gouldiae*). *Conservation Genet Resour* **7**: 899-907.

Küpper C, Stocks M, Risse JE, dos Remedios N, Farrell LL, McRae SB et al. (2015). A supergene determines highly divergent male reproductive morphs in the ruff. *Nature Genetics* **advance online publication doi: 10.1038/ng.3443**.

Lamichhaney S, Fan G, Widemo F, Gunnarsson U, Thalmann DS, Hoepfner MP et al. (2015). Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nature Genetics* **advance online publication doi: 10.1038/ng.3430**.

Mank JE, Nam K, Ellegren H (2010). Faster-Z evolution is predominantly due to genetic drift *Molecular Biology and Evolution* **27**: 661-670.

McKinnon JS, Pierotti MER (2010). Colour polymorphism and correlated characters: genetic mechanisms and evolution. *Molecular Ecology* **19**: 5101-5125.

Morton NE (1955). Sequential tests for the detection of linkage. *American Journal of Human Genetics* **7**: 277-318.

Navarro A, Barton NH (2003). Chromosomal speciation and molecular divergence-Accelerated evolution in rearranged chromosomes. *Science* **300**: 321-324.

Poelstra JW, Vijay N, Bossu CM, Lantz H, Ryll B, Muller I et al. (2014). The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **344**: 1410-1414.

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

Pryke SR, Griffith SC (2006). Red dominates black: agonistic signalling among head morphs in the colour polymorphic Gouldian finch. *Proceedings of the Royal Society B* **273**: 949-957.

Pryke SR, Griffith SC (2007). The relative role of male vs. female mate choice in maintaining assortative pairing among discrete colour morphs. *Journal of Evolutionary Biology* **20**: 1512-1521.

Pryke SR, Griffith SC (2009a). Genetic incompatibility drives sex allocation and maternal investment in a polymorphic finch. *Science* **323**: 1605-1607.

Pryke SR, Griffith SC (2009b). Postzygotic genetic incompatibility between sympatric color morphs. *Evolution* **63**: 793-798.

Qvarnström A, Bailey RI (2008). Speciation through evolution of sex-linked genes. *Heredity* **102**: 4-15.

Rieseberg LH (2001). Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution* **16**: 351-358.

Sharlow ER, Paine CS, Babiarez L, Eisinger M, Shapiro S, Seiberg M (2000). The protease-activated receptor-2 upregulates keratinocyte phagocytosis. *Journal of Cell Science* **113**: 3093-3101.

Singhal S, Leffler EM, Sannareddy K, Turner I, Venn O, Hooper DM et al. (2015). Stable recombination hotspots in birds. *Science* **350**: 928-932.

Southern HN (1945). Polymorphism in *Poephila gouldiae* Gould. *Journal of Genetics* **47**: 51-57.

Stapley J, Birkhead TR, Burke T, Slate J (2010). Pronounced inter- and intrachromosomal variation in linkage disequilibrium across the zebra finch genome. *Genome Research* **20**: 496-502.

Stump AD, Shoener JA, Costantini C, Sagnon NF, Besansky NJ (2005). Sex-linked differentiation between incipient species of *Anopheles gambiae*. *Genetics* **169**: 1509-1519.

Thornycroft HB (1966). Chromosomal polymorphism in the white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Science* **154**: 1571-1572.

Tuttle EM (2003). Alternative reproductive strategies in the white-throated sparrow: behavioral and genetic evidence. *Behavioral Ecology* **14**: 425-432.

Voorrips RE (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity* **93**: 77-78.

Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Künstner A et al. (2010). The genome of a songbird. *Nature* **464**: 757-762.

Wellenreuther M, Svensson EI, Hansson B (2014). Sexual selection and genetic colour polymorphisms in animals. *Molecular Ecology* **23**: 5398-5414.

Zhang G, Li C, Li Q, Li B, Larkin DM, Lee C et al. (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**: 1311-1320.

Supplementary File S1

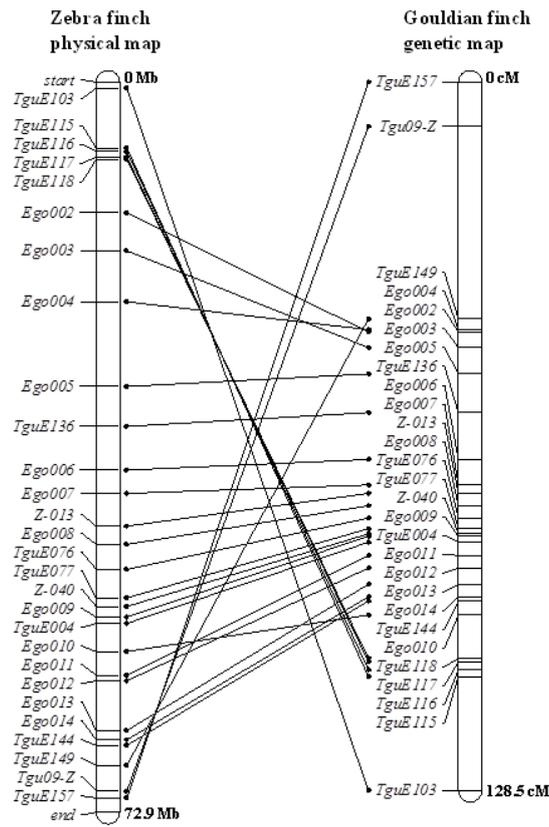
Supplementary Table S1 Summary of 34 microsatellite markers used to map the Gouldian finch Z chromosome.

Locus	Position*	Reference
TguE103	569652	Kim et al, 2015
TguE115	6565637	Kim et al, 2015
TguE116	6910485	Kim et al, 2015
TguE117	7474079	Kim et al, 2015
TguE118	7679813	Kim et al, 2015
Ego001	11788525	Kim et al, 2015
Ego002	13029255	Kim et al, 2015
Ego003	16805901	Kim et al, 2015
Ego004	21903548	Kim et al, 2015
TguE126	25169498	Kim et al, 2015
TguE129	26442253	Kim et al, 2015
Ego005	30359933	Kim et al, 2015
TguE136	34338648	Kim et al, 2015
Ego006	38701383	Kim et al, 2015
Ego007	41041863	Kim et al, 2015
Z-013	44295478	Dawson et al, 2015b
Ego008	46132393	Kim et al, 2015
TguE076	48610858	Kim et al, 2015
TguE077	51465521	Kim et al, 2015
TguE078	51783687	Kim et al, 2015
Z-040	52307826	Dawson et al, 2015b
Ego009	53372288	Kim et al, 2015
TguE004	53998897	Kim et al, 2015
Z37B	54802039	Dawson et al, 2015a
Ego010	56811660	Kim et al, 2015
Ego011	59183392	Kim et al, 2015
Ego012	59730427	Kim et al, 2015
Ego013	64693362	Kim et al, 2015
Ego014	65618011	Kim et al, 2015
TguE144	66170871	Kim et al, 2015
Z-054	67880083	Dawson et al, 2015b
TguE149	68168281	Kim et al, 2015
Tgu09-Z	70750947	Kim et al, 2015; Slate et al, 2007
TguE157	71420647	Kim et al, 2015

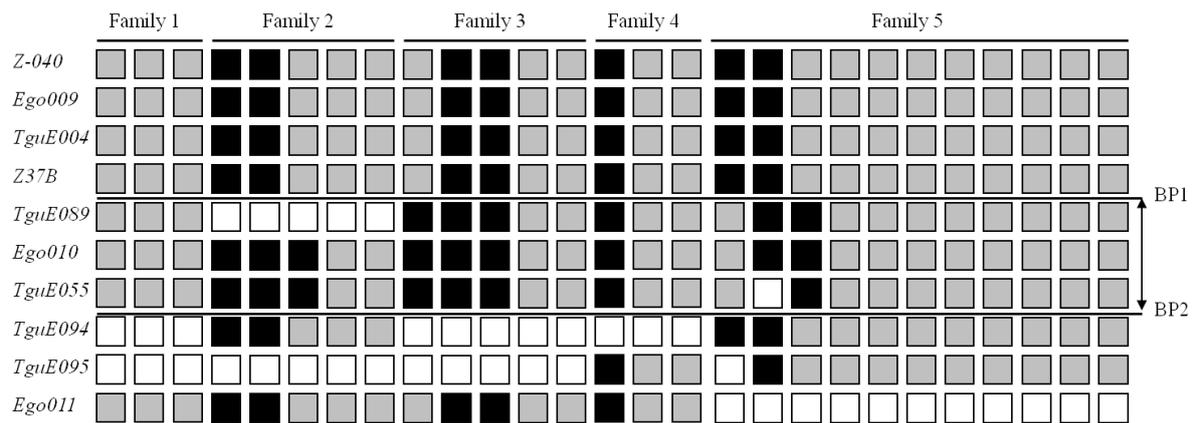
* 5' position of PCR products in the zebra finch reference genomic sequence *Taeniopygia_guttata*-3.2.4.

Supplementary Table S2 Summary of the primers used to identify informative SNPs by sequencing around locus Ego010. The end positions of the PCR amplicons are based on the ZF reference Z chromosome sequence. Microsatellite markers (SSR) are shaded (see **Supplementary Table S1**).

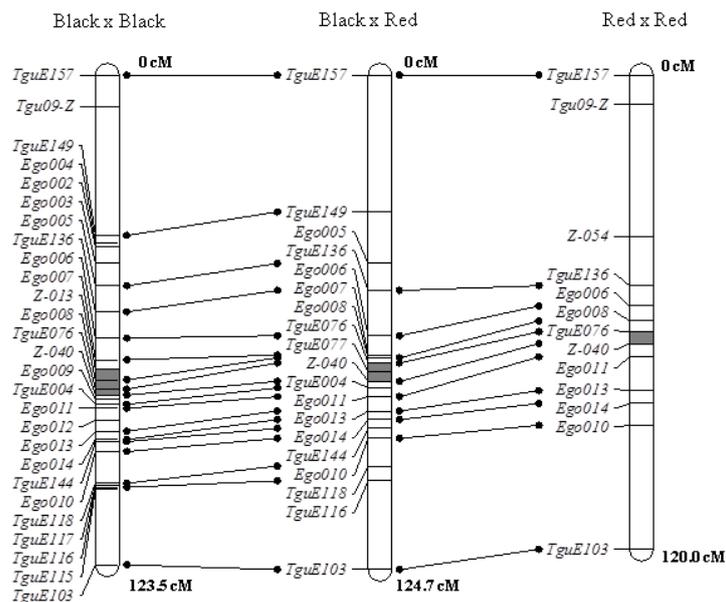
Locus	Gene	start	end	Forward primers (5' →3')	Reverse primers (5' →3')
Z37B	SSR	54802039	54802142	Dawson et al, 2015a	
TguE089	JMY	55843177	55844463	CCTGGGTTAGTCAGGAACGA	CCAGGCGTAGTCTGGTTGAT
Ego010	SSR	56811661	56811940	Kim et al, 2015	
TguE055	PAR2	56982536	56983699	AGTTAGGAGCACCTGCCTCA	ATCGCTTGTGCAATTTAGCC
TguE094	ADFP	58333832	58334807	CACACGTGCCTACCAACAAG	ACGCAGATCATCACCATCAT
TguE095	SH3GL2	58840618	58841761	CCCAC TTCAGAATCTCCATGA	CGTCGACCTTCCATTTTCTT
Ego011	SSR	59183392	59183539	Kim et al, 2015	



Supplementary Figure S1 Genetic map positions of 28 microsatellite markers in the framework marker map of the Gouldian finch Z chromosome, compared with the physical map positions in the zebra finch.



Supplementary Figure S2 Breakpoint analysis for locus Ego010, which was shown to have translocated into another part of the Z chromosome (Figure 1) by using five Gouldian finch families. Each column represents a single offspring chromosome inherited from the father and each row represents a marker locus, ordered according to position in the zebra finch genome. Black and grey squares represent two alternative alleles whose phases are inferred from the pedigree. White squares indicate that the markers were not informative or that the individuals were not genotyped. Two putative break points (BP1 and BP2) show the pattern of double crossing-over, ranging from a minimum of 1.1 Mb (TguE089 ~ TguE055) to a maximum of 3.5 Mb (Z37B ~ TguE094).



Supplementary Figure S3 Comparison of the framework maps from three types of father, who each had different combinations of Black and Red chromosomes, in relation to the genotype at the Red locus: Black x Black, Red x Red, and Black x Red, where Black and Red denote the chromosomes that have the recessive black allele (Z^f) or the red allele (Z^R), respectively. The father's genotypes were inferred from the pedigree and the phenotypes of their offspring. The shaded areas in the bar show the most likely map position of the Red locus in the Z chromosome.

References

Dawson D, Brekke P, Dos Remedios N, Horsburgh G (2015a). A marker suitable for sex-typing birds from degraded samples. *Conservation Genetics Resources* **7**: 337–343.

Dawson DA, Bird S, Horsburgh GJ, Ball AD (2015b). Autosomal and Z-linked microsatellite markers enhanced for cross-species utility and assessed in a range of birds, including species of conservation concern. *Conservation Genetics Resources* **7**: 881-886.

Kim K-W, Zhang H, Horsburgh GJ, Bird S, Dawson DA, Griffith SC et al. (2015). Four-way development of microsatellite markers for the Gouldian finch (*Erythrura gouldiae*). *Conservation Genetics Resources* **7**: 899-907.

Slate J, Hale M, Birkhead T (2007). Simple sequence repeats in zebra finch (*Taeniopygia guttata*) expressed sequence tags: a new resource for evolutionary genetic studies of passerines. *BMC Genomics* **8**: 52.