Quantitative and molecular genetics of phenotypic variation in the zebra finch

Dissertation
Fakultät für Biologie
Ludwig-Maximilians-Universität München

durchgeführt am
Max-Planck-Institut für Ornithologie
Seewiesen

vorgelegt von
Johann Ulrich Knief
August 2015

Erstgutachter: Prof. Dr. **Bart Kempenaers**
Zweitgutachter: Prof. Dr. **John Parsch**

Eingereicht am: 1.9.2015
Tag der mündlichen Prüfung: 12.1.2016

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

Contents

General introduction

"Genetics is a field of the future"
— Nelson *et al.* 2013

Evolutionary genetics in the age of genomics

Sequencing and genotyping costs decreased rapidly over the last 15 years (Wetterstrand 2015), making now the golden age for evolutionary genetics, in which we can address old questions with new sequencing and genotyping tools in non-model organisms (Nadeau & Jiggins 2010). As a consequence, we can now analyze the distribution and effect sizes of individual genetic variants contributing to phenotypic variation in quantitative traits (genetic architecture), we can study inbreeding effects at the genomic level, or localize regions under selection within genomes.

Most of these questions had been addressed in a quantitative genetic framework, which was built over the past 100 years mostly in the absence of directly observable genotypic data (Fisher 1930). Nevertheless, the theory was successfully employed to understand the sources of continuous phenotypic variation and to predict the nature of evolutionary change in wild and captive populations (Falconer & Mackay 1996). Using specific breeding designs (such as line-crosses or inbreeding experiments), it has even been possible to crudely characterize the genetic architecture of continuous phenotypic traits (Lynch & Walsh 1998). However, shifting from these statistical descriptions of populations in terms of variances and covariances to an actual understanding of the functional molecular genetic basis of phenotypic variation remained elusive (Erickson *et al.* 2004), but is essential if we want to understand the influence of genetic architecture on phenotypic evolution (Lee *et al.* 2014).

Two complementary research strategies exist for the study of genetic architecture. The forward genetics approach uses the phenotype to identify the underlying genetic basis (for example via QTL mapping or association studies, **Chapters 1, 5 and 6**), whereas the reverse genetics approach first identifies outlier regions in a genome (for example outliers in terms of Tajima's D or diversity π, **Chapter 4, 5 and 7**) and then searches for the effects of these regions on the organism. Both strategies are not new, but have been revolutionized by the progress in sequencing and genotyping possibilities, which ultimately has led to their union in the field of population genomics (reviewed in Pardo-Diaz *et al.* 2015; Stinchcombe & Hoekstra 2008).

In population genomics, locus-specific effects (such as selection, mutation or recombination) are separated from genome-wide effects (demographic events like bottlenecks or inbreeding) by genome-wide sampling schemes, which should lead to an

improved understanding of the molecular basis of phenotypic variation and ultimately evolution (Black IV *et al.* 2001; Luikart *et al.* 2003).

Population genomics approaches are by definition unbiased and hypothesis-free, in the sense that genome-wide approaches do not require prior knowledge about candidate genes or the mode of gene action (van Helden 2013). However, they often suffer from high false-positive rates due to an immense multiple testing burden (Kraft *et al.* 2009; Pavlidis *et al.* 2012). Thus, confirmatory sampling has become standard in genomic studies of model organisms like humans or *Drosophila sp.* (NCI-NHGRI working group on replication in association studies *et al.* 2007), but has not been pervasively introduced in the population genomics field of non-model organisms, not to the least because results can be published without replicates or controls (M Zody 2012, pers. comm.).

My PhD thesis bridges the gap between unbiased whole-genome methods (**Chapters 1, 2, 4 and 7**) and hypothesis-driven confirmatory sampling (**Chapters 4, 5, 6 and 7**) to study the genetic architecture of phenotypic variation in an emerging model species for evolutionary genetics, the zebra finch (*Taeniopygia guttata*).

The nature of genetic variation

In order to understand the genetic basis of complex phenotypic traits one needs to identify the number of genes affecting them, the relative effects and the mode of action of individual genes, which in total make up the genetic architecture (Erickson *et al.* 2004). To achieve these goals one needs to integrate quantitative and molecular genetics (Mackay 2001) to first decompose the phenotypic variation into its environmental and genetic component and then further describe the genetic component quantitatively (quantitative genetics) and qualitatively (molecular genetics).

Genetic variation as variance components

Research in quantitative genetics is generally concerned with separating phenotypic variation into an environmental and a genetic component and to further split the genetic component into additive, dominance and epistatic effects (Erickson *et al.* 2004). The resemblance between relatives and consequently the response to selection depends on the additive genetic variance component, because additive genetic effects are independent of interactions between alleles (dominance) and loci (epistasis; Lynch & Walsh 1998). Consequently, researchers are mostly interested in partitioning the additive genetic variance component from all the non-additive variance (Roff 1997), and this is usually done by assessing the phenotypic resemblance between relatives in families or pedigrees. While it was for a long time necessary to follow specific breeding designs to separate the different sources of variation, the introduction of the statistical mixed-modelling framework into quantitative genetics (the so called “animal model”) made it possible to partition variance

components in outbred populations with an arbitrary pedigree structure (Kruuk 2004). In my PhD thesis, I make use of the "animal model" and extensions thereof (e.g. Almasy & Blangero 2010) in **Chapters 1, 6 and 7** to partition the phenotypic variance in several morphological traits into their different variance components with a special emphasize on additive genetic variance.

Deviations from strict additivity will introduce dominance variation (e.g. Nietlisbach & Hadfield 2015), which is the interaction of the two alleles at a locus. In my PhD thesis, I was concerned with dominance variation in light of (1) inbreeding depression, the decline of a trait's value due to the mating of related individuals (**Chapters 2 and 3**), which only occurs if there is some degree of dominance (Roff 1997) and (2) heterosis, the superiority of heterozygous individuals over homozygotes in terms of fitness (**Chapter 7**).

Genetic variation as molecular genetic diversity

Assuming that the infinitesimal additive model of gene action is a good approximation to the genetic architecture of a complex trait, an understanding of how selection will affect the evolution of these traits can be achieved via variance partitioning (Rockman 2012; Roff 1997). However, there is also much to be gained by studying the underlying molecular genetic variants, because this could potentially lead to an understanding of the selective forces shaping and maintaining genetic variation in more detail (for examples see Gratten *et al.* 2008; Johnston *et al.* 2013).

Molecular genetic variation ranges from single nucleotides (SNPs) to large-scale chromosome- or genome-wide differences, such as deletions, duplications, inversions or translocations (Conrad & Hurles 2007; White 1977). To date, most research on intraspecific genetic variation is focused on single nucleotides because SNPs are the most abundant genetic variants within a genome, they can be identified easily and genotyped in large populations, which make them the ideal variants to cover and characterize the whole genome of an organism (Frazer *et al.* 2009). For these reasons I used genome-wide SNP data throughout my PhD thesis (**Chapters 1, 2, 4, 6 and 7**). Moreover, SNPs can be used as proxies for larger structural variants because generally some of them will be in linkage disequilibrium with the structural variants (McCarroll *et al.* 2008). For the first time in a bird species, I made use of this to detect and study the phenotypic effects of large inversion polymorphisms in the zebra finch genome (**Chapter 7**).

Genetic mapping is the forward genetics approach to relate molecular genetic variation with phenotypic diversity (Lander & Schork 1994). Using individual SNPs or haplotypes composed of multiple SNPs as markers for a specific genomic region, one can relate the inheritance patterns of a complex phenotypic trait with these regions (linkage mapping; Morton 1956). Without pedigree information one can exploit linkage disequilibrium stemming from the demographic history of a population to identify genomic regions associated with a complex

phenotype (association mapping; Risch & Merikangas 1996). Linkage mapping is powerful but suffers from a low spatial resolution, whereas in association studies a trade-off between power and resolution exists depending on the amount of linkage disequilibrium within a population (De La Chapelle & Wright 1998). In any case, both methods generally do not identify the causal variants underlying phenotypic variation, but rather genomic intervals or haplotypes which are linked to the causal variants. In my PhD thesis, I used linkage mapping in a captive population of zebra finches to identify genomic regions associated with phenotypic variation in beak morphology (**Chapter 1**) and to localize centromeres on chromosomes (**Chapter 5**), as well as association mapping in a wild population of zebra finches for fine-mapping phenotypic variation (**Chapter 6**).

Genome scans are the reverse genetics approach to connect genotypic with phenotypic variation (Pardo-Diaz *et al.* 2015). Using genome-wide data on genetic variants (e.g. SNPs), the whole genome is screened for outliers in terms of diversity, heterozygosity or differentiation, which could potentially be footprints of selection (both at the intra- or interspecific level). Regions of low diversity and heterozygosity could be under positive selection (Cutter & Payseur 2013; Rubin *et al.* 2010), whereas regions of high heterozygosity and/or diversity may indicate balancing selection (Fijarczyk & Babik 2015). Once these genomic regions have been identified, their phenotypic effects are assessed. In my PhD thesis, I conducted several genome scans: (1) I screened the genome for regions which are transmitted to the next generation more often than expected under Mendelian segregation (transmission distortion, **Chapter 4**). (2) After sequencing a pooled sample of DNA from 100 wild-caught zebra finches to an average genome-coverage of 247.5 x, I screened the genome for outliers in diversity (Figure 1) and heterozygosity (**Chapter 7**). Regions of low diversity seemed to reside predominantly at centromeres, which let me to map the positions of centromeres on all chromosomes in the zebra finch genome (**Chapter 5**).

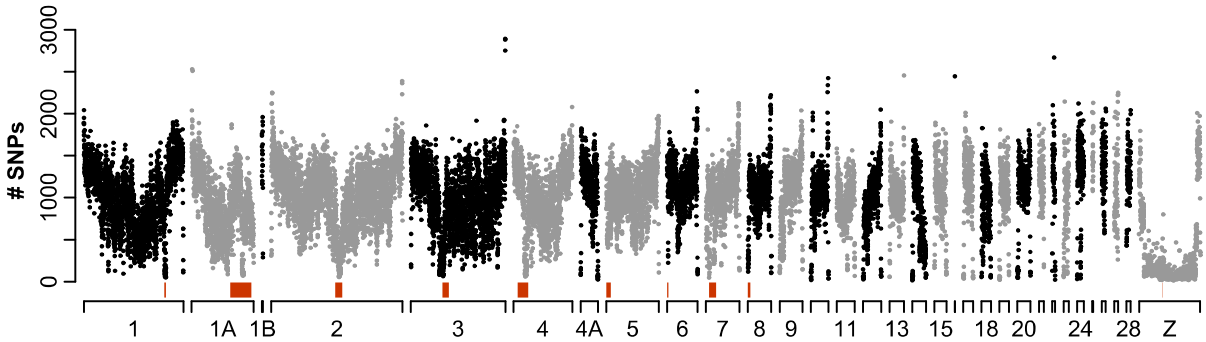


Figure 1: *Genome scan for diversity (number of SNPs in 50 kb non-overlapping windows) along the zebra finch genome. SNPs were called as part of a 247.5 x coverage pooled-sequencing experiment using DNA of 100 wild-caught zebra finches. The positions of previously known centromeres are marked in red (mapped via fluorescence in situ hybridization; Warren et al. 2010).*

The false-positive rates of both genetic mapping and genome scans are generally high, meaning that associations or outlier regions often appear by chance events alone (Pardo-Diaz *et al.* 2015). Consequently, I followed-up all the results stemming from the unbiased forward and reverse genetics genome-wide approaches with detailed confirmatory analyses, in which I characterized the outlier loci in new samples (if possible even from different populations to reduce effects stemming from the genetic background).

The zebra finch as a model in evolutionary genetics

The zebra finch has a long history of being a model species in neuroscience, behavioral ecology and physiology, but only recently gained popularity in evolutionary and molecular genetics research (Griffith & Buchanan 2010). Its success as a model species resulted from the fact that it can be kept in large quantities in the lab and that it readily breeds in captivity with no apparent seasonality (Figure 1; Zann 1996).

Figure 2: *A male zebra finch (left) and a female. Photo taken by Wolfgang Forstmeier.*

By the year 2010, the zebra finch genome was only the second bird and the first songbird genome to be sequenced (Warren *et al.* 2010). In contrast to the more recently sequenced bird genomes (Jarvis *et al.* 2014; Zhang *et al.* 2014), the zebra finch genome was built using a combination of shot-gun sequencing, Sanger sequencing of bacterial artificial chromosomes and a linkage map (Stapley *et al.* 2008; Warren *et al.* 2010), which overall guarantees a high-quality genome alignment, which is necessary to conduct genome scans efficiently and allows studying chromosome structure and structural variants.

Bird genomes are karyotypically stable across species with a high degree of intrachromosomal synteny (Romanov *et al.* 2014; Skinner & Griffin 2012). Generally they consist of a few macro- and several microchromosomes, which are differentiated according to physical size. The current zebra finch genome assembly (WUSTL v3.2.4) consists of 32

chromosomes, which means that all macrochromosomes are included and that eight microchromosomes are missing (Itoh & Arnold 2005; Pigozzi & Solari 1998).

In every chapter of my PhD thesis I use the physical genome assembly and often also a linkage map developed with birds from our captive population (Backström *et al.* 2010). Building these genomic resources from scratch would have been a severe hurdle and thus the zebra finch is an ideal model organism to address evolutionary genetics and genomics questions.

Thesis outline

The overall aim of my PhD thesis was to study the quantitative and molecular genetic variation in the zebra finch genome and how this variation contributes to complex phenotypic traits. In general, I screened the genome for additive effects on morphology (**Chapters 1, 6 and 7**), dominance effects (**Chapters 2 and 3**), genomic structures (**Chapters 5 and 7**) and selfish genetic elements (**Chapter 4**).

Specifically, in **Chapter 1**, I performed quantitative genetic analyses and linkage mapping for the three beak dimensions (beak length, beak depth and beak width) in a captive population of zebra finches. Beak shape has been extensively studied in Darwin's finches and has become one of the textbook examples for natural selection acting on a phenotypic trait (Grant & Grant 2002). Recently, beak morphology of Darwin's finches has been the subject of transcriptomic (Abzhanov *et al.* 2006; Abzhanov *et al.* 2004; Mallarino *et al.* 2011) and also genomic (Lamichhaney *et al.* 2015) studies, which gave me strong candidate genes to study in zebra finches. Thus, after describing the beak in a multivariate quantitative genetic analysis, I conducted a genome-wide linkage analysis and tested whether the known candidate genes from Darwin's finches overlapped with the identified QTL regions.

To get closer to the actual causal variants, linkage disequilibrium should span only short genomic regions (De La Chapelle & Wright 1998). To get this low linkage disequilibrium in captive populations, it would require unrealistically large pedigrees, but wild Australian zebra finches had been found in previous studies to exhibit rapid decay of linkage disequilibrium over short genomic distances (Balakrishnan & Edwards 2009), which made them well-suited for my fine-mapping efforts. Thus, I sampled zebra finches in the wild, conducted whole-genome next-generation sequencing of pooled DNA samples, called SNPs and performed genome scans for outlier loci in terms of diversity (**Chapter 5**) and heterozygosity (**Chapter 7**). I subsequently developed a SNP-genotyping-array which covered candidate genes in previously identified QTL regions, outlier regions from the genome scans and the whole genome (described in **Chapters 2, 6 and 7**). After genotyping almost 1,000 wild-caught zebra finches with this SNP-array, I first used the genotype data to

exemplify how inbreeding can be quantified from dense SNP panels without the need for an extended pedigree (**Chapter 2**).

I followed up on this idea in **Chapter 3**. Here, I used gene-dropping simulations to see how the method described in **Chapter 2** compares to traditional marker- and pedigree-based inbreeding estimates in terms of precision and accuracy and how it can be used in heterozygosity-fitness correlations.

In **Chapter 4** I conducted the last genome-wide scan in my PhD thesis and combined it with confirmatory analyses. In our captive population of zebra finches, I first screened the genome for regions exhibiting segregation distortion, which means regions that do not get passed on in a Mendelian fashion. These regions could either harbor recessive deleterious mutations or be selfish, exploiting the meiotic machinery to their own advantage (Burt & Trivers 2006). In any case, deviations from fair Mendelian segregation can only result from some form of selection: In case of viability effects it would be selection at the organismal level, whereas in case of selfish genetic elements selection would act at the level of the gene itself (and could potentially even be harmful to the organism; Traulsen & Reed 2012). Selfish genetic elements often consist of multiple genes in high linkage disequilibrium. Thus, they have the potential to fix deleterious mutations due to Hill-Robertson interference and shape the recombination and diversity landscape in a genome (Dyer *et al.* 2007), thereby influencing genetic variation markedly.

In **Chapter 5** I followed up on the genome-wide scan for diversity presented in Figure 1. The genomic positions of ten centromeres had been mapped to the zebra finch genome previously via fluorescence in situ hybridization (Warren *et al.* 2010) and their positions overlapped with regions of highly reduced diversity (an around 10-fold reduction compared to the genome-wide average), which could be due to background (purifying) or positive selection. All microchromosomes in the zebra finch genome are known to be acrocentric from cytogenetic studies (Pigozzi 2008) and most of the microchromosomes showed regions of reduced diversity at both chromosomal ends (Figure 1). In order to map the centromeres to one of the chromosomal ends, I used a specific linkage mapping technique which makes use of naturally occurring triploid individuals (triploidy mapping; Chakravarti *et al.* 1989).

In **Chapter 6** I followed up on the genome-wide linkage mapping results for beak morphology (**Chapter 1**) and other linkage scans, which had been conducted in our captive population of zebra finches (Schielzeth *et al.* 2012). The aim of the study was to fine-map the causal variants in previously identified QTL regions. Therefore, I first performed association mapping between those SNPs that had been genotyped in candidate genes (in QTL regions) and their respective phenotypes in the wild sample of zebra finches. After I had identified the strongest associations in the wild, which should be close to the causal variants or be indeed the causal variants, I genotyped those SNPs in additional four captive

populations to (1) verify the associations and (2) test whether these SNPs are causal or linked to the causal variants.

In **Chapter 7** I combined results from the genome scan for heterozygosity and the genotyping in the wild sample of birds. The genome scan showed sharp peaks in heterozygosity on several chromosomes and I used the individual SNP-genotyping in the wild birds to clarify what kind of genetic structures had let to these peaks. They appeared to be breakpoints of inversion polymorphisms and I followed up on these inversions by genotyping tagging SNPs, which are in perfect linkage disequilibrium with the inversions, in four captive populations. Then I tested the inversions for their fitness effects and phenotypic associations. Overall, this is the first study of large-scale intraspecific inversion polymorphisms in a bird genome in terms of their molecular and quantitative genetic effects.

References

Abzhanov A, Kuo WP, Hartmann C*, et al.* (2006) The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* **442**, 563–567.

Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004) *Bmp4* and morphological variation of beaks in Darwin's finches. *Science* **305**, 1462–1465.

Almasy L, Blangero J (2010) Variance component methods for analysis of complex phenotypes. *Cold Spring Harb Protoc* **2010**, pdb.top77.

Backström N, Forstmeier W, Schielzeth H*, et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research* **20**, 485–495.

Balakrishnan CN, Edwards SV (2009) Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* **181**, 645–660.

Black IV WC, Baer CF, Antolin MF, DuTeau NM (2001) Population genomics: genome-wide sampling of insect populations. *Annual Review of Entomology* **46**, 441–469.

Burt A, Trivers R (2006) *Genes in conflict: the biology of selfish genetic elements*. Belknap Press, Harvard University Press, Cambridge, Massachusetts, London, England.

Chakravarti A, Majumder PP, Slaugenhaupt SA*, et al.* (1989) Gene-centromere mapping and the study of non-disjunction in autosomal trisomies and ovarian teratomas. *Prog Clin Biol Res* **311**, 45–79.

Conrad DF, Hurles ME (2007) The population genetics of structural variation. *Nature Genetics* **39**, S30–S36.

Cutter AD, Payseur BA (2013) Genomic signatures of selection at linked sites: unifying the disparity among species. *Nature Reviews Genetics* **14**, 262–274.

De La Chapelle A, Wright FA (1998) Linkage disequilibrium mapping in isolated populations: the example of Finland revisited. *Proceedings of the National Academy of Sciences of the USA* **95**, 12416–12423.

Dyer KA, Charlesworth B, Jaenike J (2007) Chromosome-wide linkage disequilibrium as a consequence of meiotic drive. *Proceedings of the National Academy of Sciences of the USA* **104**, 1587–1592.

Erickson DL, Fenster CB, Stenoien HK, Price D (2004) Quantitative trait locus analyses and the study of evolutionary process. *Molecular Ecology* **13**, 2505–2522.

Falconer D, Mackay T (1996) *Introduction to quantitative genetics*, 4th edn. Longmann, Harlow, UK.

Fijarczyk A, Babik W (2015) Detecting balancing selection in genomes: limits and prospects. *Molecular Ecology* **24**, 3529–3545.

Fisher RA (1930) *The genetical theory of natural selection*. Oxford University Press, Oxford, UK.

Frazer KA, Murray SS, Schork NJ, Topol EJ (2009) Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics* **10**, 241–251.

Grant PR, Grant BR (2002) Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**, 707–711.

Gratten J, Wilson AJ, McRae AF*, et al.* (2008) A localized negative genetic correlation constrains microevolution of coat color in wild sheep. *Science* **319**, 318–320.

Griffith SC, Buchanan KL (2010) The Zebra Finch: the ultimate Australian supermodel. *Emu* **110**, v–xii.

Itoh Y, Arnold AP (2005) Chromosomal polymorphism and comparative painting analysis in the zebra finch. *Chromosome Research* **13**, 47–56.

Jarvis ED, Mirarab S, Aberer AJ*, et al.* (2014) Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346**, 1320–1331.

Johnston SE, Gratten J, Berenos C*, et al.* (2013) Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* **502**, 93–95.

Kraft P, Zeggini E, Ioannidis JPA (2009) Replication in genome-wide association studies. *Statistical Science* **24**, 561–573.

Kruuk LEB (2004) Estimating genetic parameters in natural populations using the 'animal model'. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **359**, 873–890.

Lamichhaney S, Berglund J, Almén MS*, et al.* (2015) Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375.

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. *Science* **265**, 2037–2048.

Lee YW, Gould BA, Stinchcombe JR (2014) Identifying the genes underlying quantitative traits: a rationale for the QTN programme. *AoB Plants* **6**, plu004.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics* **4**, 981–994.

Lynch M, Walsh B (1998) *Genetics and analysis of quantitative traits*. Sinauer, Sunderland, MA.

Mackay TFC (2001) The genetic architecture of quantitative traits. *Annual Review of Genetics* **35**, 303–339.

Mallarino R, Grant PR, Grant BR*, et al.* (2011) Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proceedings of the National Academy of Sciences USA* **108**, 4057–4062.

McCarroll SA, Kuruvilla FG, Korn JM*, et al.* (2008) Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nature Genetics* **40**, 1166–1174.

Morton NE (1956) The detection and estimation of linkage between the genes for elliptocytosis and the Rh-blood type. *American Journal of Human Genetics* **8**, 80–96.

Nadeau NJ, Jiggins CD (2010) A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics* **26**, 484–492.

NCI-NHGRI working group on replication in association studies, Chanock SJ, Manolio T*, et al.* (2007) Replicating genotype-phenotype associations. *Nature* **447**, 655–660.

Nelson RM, Pettersson ME, Carlborg O (2013) A century after Fisher: time for a new paradigm in quantitative genetics. *Trends in Genetics* **29**, 669–676.

Nietlisbach P, Hadfield JD (2015) Heritability of heterozygosity offers a new way of understanding why dominant gene action contributes to additive genetic variance. *Evolution* **69**, 1948–1952.

Pardo-Diaz C, Salazar C, Jiggins CD (2015) Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution* **6**, 445–464.

Pavlidis P, Jensen JD, Stephan W, Stamatakis A (2012) A critical assessment of storytelling: gene ontology categories and the importance of validating genomic scans. *Molecular Biology and Evolution* **29**, 3237–3248.

Pigozzi MI (2008) Relationship between physical and genetic distances along the zebra finch Z chromosome. *Chromosome Research* **16**, 839–849.

Pigozzi MI, Solari AJ (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.

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. *Science* **273**, 1516–1517.

Rockman MV (2012) The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**, 1–17.

Roff D (1997) *Evolutionary quantitative genetics*. Chapman & Hall, London, UK.

Romanov MN, Farré M, Lithgow PE*, et al.* (2014) Reconstruction of gross avian genome structure, organization and evolution suggests that the chicken lineage most closely resembles the dinosaur avian ancestor. *Bmc Genomics* **15**, Artn 1060.

Rubin CJ, Zody MC, Eriksson J*, et al.* (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* **464**, 587–593.

Schielzeth H, Forstmeier W, Kempenaers B, Ellegren H (2012) QTL linkage mapping of wing length in zebra finch using genome-wide single nucleotide polymorphisms markers. *Molecular Ecology* **21**, 329–339.

Skinner BM, Griffin DK (2012) Intrachromosomal rearrangements in avian genome evolution: evidence for regions prone to breakpoints. *Heredity* **108**, 37–41.

Stapley J, Birkhead TR, Burke T, Slate J (2008) A linkage map of the zebra finch *Taeniopygia guttata* provides new insights into avian genome evolution. *Genetics* **179**, 651–667.

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity (Edinb)* **100**, 158–170.

Traulsen A, Reed FA (2012) From genes to games: cooperation and cyclic dominance in meiotic drive. *Journal of Theoretical Biology* **299**, 120–125.

van Helden P (2013) Data-driven hypotheses. *EMBO Rep* **14**, 104.

Warren WC, Clayton DF, Ellegren H*, et al.* (2010) The genome of a songbird. *Nature* **464**, 757–762.

Wetterstrand KA (2015) DNA sequencing costs: data from the NHGRI genome sequencing program (GSP), www.genome.gov/sequencingcosts.

White MJD (1977) *Animal cytology and evolution*. Cambridge University Press, UK, Cambridge.

Zann RA (1996) *The zebra finch: a synthesis of field and laboratory studies*. Oxford University Press, USA.

Zhang GJ, Li C, Li QY*, et al.* (2014) Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320.

Chapter 1: QTL and quantitative genetic analysis of beak morphology reveals patterns of standing genetic variation in an Estrildid finch

Abstract

The intra- and interspecific diversity of avian beak morphologies is one of the most compelling examples for the power of natural selection acting on a morphological trait. The development and diversification of the beak have also become a textbook example for evolutionary developmental biology, and variation in expression levels of several genes is known to causally affect beak shape. However, until now, no genomic polymorphisms have been identified, which are related to beak morphology in birds. QTL mapping does reveal the location of causal polymorphisms, albeit with poor spatial resolution. Here, we estimate heritability and genetic correlations for beak length, depth and width and perform a QTL linkage analysis for these traits based on 1,404 informative single-nucleotide polymorphisms genotyped in a four-generation pedigree of 992 captive zebra finches (Taeniopygia guttata). Beak size, relative to body size, was sexually dimorphic (larger in males). Heritability estimates ranged from 0.47 for beak length to 0.74 for beak width. QTL mapping revealed four to five regions of significant or suggestive genome-wide linkage for each of the three beak dimensions (nine different regions in total). Eight out of 11 genes known to influence beak morphology are located in these nine peak regions. Five QTL do not cover known candidates demonstrating that yet unknown genes or regulatory elements may influence beak morphology in the zebra finch.

Published as: Knief U, H Schielzeth, B Kempenaers, H Ellegren, W Forstmeier 2012: QTL and quantitative genetic analysis of beak morphology reveals patterns of standing genetic variation in an Estrildid finch. Molecular Ecology 21: 3704–3717.

MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 3704–3717 doi: 10.1111/j.1365-294X.2012.05661.x

QTL and quantitative genetic analysis of beak morphology reveals patterns of standing genetic variation in an Estrildid finch

ULRICH KNIEF,* HOLGER SCHIELZETH,†‡ BART KEMPENAERS,* HANS ELLEGREN† and WOLFGANG FORSTMEIER*

**Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Eberhard-Gwinner-Str., 82319 Seewiesen, Germany, †Department of Evolutionary Biology, Uppsala University, Norbyvägen 18D, 752 36 Uppsala, Sweden*

Abstract

The intra- and interspecific diversity of avian beak morphologies is one of the most compelling examples for the power of natural selection acting on a morphological trait. The development and diversification of the beak have also become a textbook example for evolutionary developmental biology, and variation in expression levels of several genes is known to causally affect beak shape. However, until now, no genomic polymorphisms have been identified, which are related to beak morphology in birds. QTL mapping does reveal the location of causal polymorphisms, albeit with poor spatial resolution. Here, we estimate heritability and genetic correlations for beak length, depth and width and perform a QTL linkage analysis for these traits based on 1404 informative single-nucleotide polymorphisms genotyped in a four-generation pedigree of 992 captive zebra finches (*Taeniopygia guttata*). Beak size, relative to body size, was sexually dimorphic (larger in males). Heritability estimates ranged from 0.47 for beak length to 0.74 for beak width. QTL mapping revealed four to five regions of significant or suggestive genome-wide linkage for each of the three beak dimensions (nine different regions in total). Eight out of 11 genes known to influence beak morphology are located in these nine peak regions. Five QTL do not cover known candidates demonstrating that yet unknown genes or regulatory elements may influence beak morphology in the zebra finch.

Keywords: candidate genes, G matrices, genetic correlations, QTL mapping, quantitative genetics, sexual dimorphism, *Taeniopygia guttata*, zebra finch

Received 8 February 2012; revision received 9 April 2012; accepted 23 April 2012

Introduction

Ever since Darwin's 'On the Origin of Species' (1859), the intra- and interspecific diversity of avian beak morphologies has been one of the most compelling examples for the power of natural selection acting on a morphological trait (Grant 1999; Grant & Grant 2002). Beak shape is of major ecological importance (Grant & Grant 2002; Herrel *et al.* 2010): it is almost axiomatic in ornithology that granivorous birds tend to have relative deep and short beaks, whereas insectivorous species tend to have narrow and long ones (Gosler 1987).

Intraspecific variation in beak shape has been related to prey size and hardness, but also to the patterns of parental care and to mating systems. For example, in great tits (*Parus major*), birds with deeper beaks caught bigger and harder prey (Gosler 1987), and in polygynous dusky warblers (*Phylloscopus fuscatus*), females with deeper beaks more often mated as secondary females and received less male assistance in feeding their offspring (Forstmeier *et al.* 2001).

Variation in beak shape may also constrain song features in passerines both within and across species

Correspondence: Ulrich Knief, Fax: +49 8157 932400; E-mail: uknief@orn.mpg.de
‡Current address: Department of Evolutionary Biology, Bielefeld University, Morgenbreede 45, 33615 Bielefeld, Germany.

(Podos 2001). As the beak is part of the vocal tract, birds with larger beaks usually produce songs with lower rates of syllable repetition, lower vocal frequencies and narrower frequency bandwidths (reviewed in Podos & Nowicki 2004).

Recently, beak morphology has become a model for developmental processes and modularity (Badyaev 2010). The molecular genetics of beak morphology variation has been extensively studied in a comparative framework using Darwin's finches (*Geospiza* sp.), chicken (*Gallus gallus*), quails (*Coturnix coturnix*) and ducks (*Anas platyrhynchos*). Abzhanov *et al.* (2004), (2006) and Mallarino *et al.* (2011) found that beak shape in Darwin's finches and chicken is regulated by variation in the expression of *bone morphogenetic protein 4* (*BMP4*), *calmodulin 1* (*CALM1*), *dickkopf 3* (*DKK3*), *transforming growth factor beta receptor II* (*TGFBR2*) and *β-catenin* (*CTNNB1*). Higher expression of *BMP4* is associated with wider and deeper beaks, whereas higher expression of *CALM1* is associated with longer beaks. Acting at a later stage of development and in a different tissue, higher expression of *DKK3*, *TGFBR2* and *CTNNB1* is causally connected to longer and deeper beaks. Wu *et al.* (2004), (2006) further showed that higher expression of *BMP4* also increased beak length in chicken and ducks. *Sonic hedgehog* (*SHH*) and *fibroblast growth factor 8* (*FGF8*) expressions synergistically promote frontonasal process outgrowth and may thus also influence beak length (Schneider *et al.* 2001; Hu *et al.* 2003; Abzhanov & Tabin 2004). Other studies point at *BMP2*, *BMP7* and *retinaldehyd dehydrogenase 2* and *3* (*RALDH2*, *RALDH3*) expression as factors correlated with variation in beak morphology (Lee *et al.* 2001; Schneider *et al.* 2001; Song *et al.* 2004).

All studies so far have used expression level analysis or expression level manipulation to detect genes associated with variation in beak shape. However, the causal polymorphisms have not yet been identified for any of those genes and it is also possible that additional (possibly regulatory) polymorphisms contribute to variation in beak morphology. It is therefore unclear whether the expression differences originate from regulatory polymorphisms that act in *cis* or *trans* to the affected genes. In principal, QTL mapping can answer this question; yet, the low resolution (i.e. the large confidence intervals of QTL linkage peaks) does not allow drawing a definite conclusion.

Here, we focus on standing additive genetic variation in beak size and shape in a captive population of zebra finches (*Taeniopygia guttata*), in which we do not expect strong recent directional selection on beak morphology. Both the absence of directional selection and the reduced environmental variation in captivity promote the detection of genetic polymorphisms contributing to the phenotype (Falconer & Mackay 1996). Because our captive population went through a bottleneck in population size (Forstmeier *et al.* 2007), the genetic polymorphisms influencing beak morphology might be reduced in comparison with wild zebra finches. On the one hand, this increases the power to detect such polymorphisms because individual QTL are likely to explain a larger fraction of the additive genetic variance (Flint & Mackay 2009). On the other hand, we might miss some polymorphisms because of the loss of variation by drift. Our population might be comparable to island or otherwise isolated populations that sustain only small effective population sizes.

Here, we use extensive single-nucleotide polymorphism (SNP) genotyping combined with pedigree information and a linkage map (Backström *et al.* 2010) to map quantitative trait loci for beak morphology. This population has previously been used to map QTL for beak colour and wing morphology (Schielzeth *et al.* 2012a,b). We then utilized the availability of the zebra finch genome (Warren *et al.* 2010) to examine whether the known candidate genes for beak morphology are located in the genomic regions associated with phenotypic variation in beak shape.

Methods

Study population and pedigree

Data on beak morphology were collected for 1090 birds (532 females and 558 males) from four generations of a captive population at the Max Planck Institute for Ornithology in Seewiesen. For 992 of those birds (510 males and 482 females), SNP data were available (Backström *et al.* 2010). Because the parents of our first-generation birds are also known, the pedigree data cover five generations. Individuals from the second to fourth generation (83% of all birds) were cross-fostered to disentangle genetic from early environmental effects.

Phenotypic data

Beak length, width and depth were measured with a digital calliper to the nearest 0.01 mm (all measurements were taken by the same person: UK). We measured the length from the proximal end of the rhampotheca, which is the horny part of the beak, to the tip of the culmen (mean ± SD: 10.42 ± 0.50, $n = 990$), the depth as the maximum at the base of the beak (mean ± SD: 8.34 ± 0.29, $n = 990$) and the width as the maximum of the upper mandible (mean ± SD: 6.62 ± 0.22, $n = 990$) (Svensson 1992). Measurements were taken on dead ($n = 642$) and alive birds ($n = 350$) that were between 112 and 3256 days of age. When a

bird died, it was immediately frozen to −20 °C and defrosted for measuring. Repeated measures of beak morphology were taken in direct sequence (length, depth, width, length, depth, width) without releasing the birds between measurements. For 17 birds, beak length and depth were measured four times (twice in sequence while alive and twice when dead). For three of those individuals, beak width was measured three times (twice in sequence when alive and once when dead). All other individuals were measured twice in sequence (beak length: $n = 871$; beak depth: $n = 1072$; beak width: $n = 278$) or only once (beak length: $n = 200$; beak width: $n = 808$). All measurements were included in a permanent environment model in VCE (see below) to estimate individual repeatability. This estimate is likely to be slightly inflated because most measurements from the same individual were taken in direct sequence and are therefore not completely independent. For QTL analyses (and also for the assessment of fixed effects), we used only one aggregated measurement per individual. Whenever we had measurements from a bird as dead and as alive, we took the measurement from the dead bird. For beak width (mostly measured once), we always used the first measurement. For beak length and depth, the average of two measurements was taken.

Beak measurements partly reflect variation in the size and shape of the beak (relative to body size), but also to some extent variation in overall body size. In an effort to distinguish between these two sources of variation, we carried out two types of analyses, one using the raw beak measurements and one controlling for three measures of overall body size as covariates (tarsus length, wing length and mass, all three fitted simultaneously). The comparison between the two types of models helps separating the sources of variation, but one should bear in mind that this commonly used approach (of controlling for size covariates) does only remove part of the overall variation in body size (Forstmeier 2011). As the measures of body size, we took single measurements of tarsus length to the nearest 0.1 mm (mean ± SD: 17.2 ± 0.6 mm, age: 34–367 days), wing length to the nearest 0.5 mm (mean ± SD: 58.4 ± 1.6 mm, age: 32–367 days) and mass to the nearest 0.1 g (mean ± SD: 16.9 ± 1.8 g, age: 87–367). Mass measurements were cube-root transformed to account for the difference in dimensionality (single beak dimensions are linear, mass is volumetric).

Quantitative genetic analysis

We used pedigree-based animal models in REML-VCE 6.0.2 for estimating narrow-sense heritability of the three beak dimensions and the genetic correlations between them (Groeneveld *et al.* 2008). VCE implements REML estimation to decompose phenotypic variance (V_P) into additive genetic variance (V_A), general maternal environment variance (V_M, random effect of mother identity not linked to the pedigree), early environment variance (V_{EE}, brood identity), permanent environment variance (V_{PE}, individual identity) and residual variance (V_R). Because most repeated measurements were taken immediately after each other, the 'permanent environment variance' might be slightly inflated (on expense of the residual variance), because the two measurements of the same individual were not entirely independent. This, however, does not affect the additive genetic variance V_A that is of primary interest in this article.

The three beak dimensions were fitted simultaneously as response variables and sex, status (dead or alive), age, and inbreeding coefficient based on our five-generation pedigree were fitted as fixed effects. In this three-trait model, a full variance–covariance matrix was fitted for each of the causal components of variance. Because V_M and V_{EE} were very small (Table S1, Supporting information), the corresponding covariances were difficult to estimate, leading to problems with model convergence (VCE output 'status 3'). Hence, we fitted a simpler model including only V_A, V_{PE} and V_R, which converged well (VCE output 'status 1') and yielded nearly identical estimates for those three variance–covariance matrices. In the above described models, all 1090 phenotyped birds and all repeated measures were included. We calculated narrow-sense heritabilities and coefficients of additive genetic variation CV_A (Houle 1992). Following Hansen & Houle (2008), we calculated the conditional additive genetic variance (V_{A*}) of each beak dimension conditional on the other two beak dimensions. This estimates the amount of additive genetic variance of each beak dimension that is independent of the additive genetic variance of the other two dimensions. From this, we calculated the conditional heritability as V_{A*}/V_P.

To check the reliability of this model (including V_A, V_{PE} and V_R), we also fitted models using ASReml 3 (Gilmour *et al.* 2009). All genetic correlations and heritability estimates were identical between VCE and ASReml to the forth decimal, but standard errors were slightly larger in ASReml (on average by 8.7% of the standard error estimate; range, 3.7–14.0%). In the Results section, we report the estimates from VCE (for consistency with Table S1, Supporting information).

SNP genotyping and linkage map

Out of the 1090 birds with data on beak morphology, 992 had previously been genotyped for 1920 SNPs using the Illumina GoldenGate Assay (Fan *et al.* 2003)

at the SNP Technology Platform in Uppsala, Uppsala University. Thousand four hundred and four informative markers were included in a best-order linkage map that covered 32 linkage groups (equivalent to chromosomes 1–28 without 22, i.e., *c.* 92% of the annotated genome) and 1479 cM (Backström *et al.* 2010). Average (±SD) marker spacing was 1.1 ± 3.4 cM, but genetic distances were much shorter in the centre of chromosomes and greater in telomeric regions (Backström *et al.* 2010).

Identity-by-descent probabilities

We used the linkage disequilibrium and linkage analysis (LDLA) module in GridQTL, a web-based QTL mapping software for estimating identity-by-descent (IBD) probabilities of genomic regions of interest and for fitting QTL models (Hernandez-Sanchez *et al.* 2009). IBD probabilities between all pairs of individuals were estimated based solely on our known pedigree using a recursive deterministic method (Pong-Wong *et al.* 2001). We ignored the option provided by GridQTL to model the generations of population founders to pedigree founders (by setting the number of generations from population founders to pedigree founders to zero), which makes the software perform a standard linkage analysis.

QTL model

GridQTL (with its underlying ASReml software) estimates the variance components using restricted maximum likelihood (REML). We used likelihood ratio tests (LRTs) to test whether variances are significantly greater than zero. The LRT compares a polygenic linear mixed model (without QTL effect, model 1) with a QTL linear mixed model (model 2). The two models are as follows:

$$\mathrm{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{A}\boldsymbol{\gamma}_1 + \varepsilon \quad (1)$$

$$\mathrm{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{A}\boldsymbol{\gamma}_1 + \mathbf{Q}\boldsymbol{\gamma}_2 + \varepsilon \quad (2)$$

where $\mathbf{y}$ is a vector of phenotypes, $\boldsymbol{\beta}$ is a vector of fixed effects, γ_1 is vector of additive nonlocalized polygenic effects, γ_2 is a vector of additive QTL effects at a specific location, $\boldsymbol{\varepsilon}$ is a vector of residuals, $\mathbf{X}$ is a design matrix relating fixed effects to observations, $\mathbf{A}$ is the additive genetic relatedness matrix and $\mathbf{Q}$ is the identity-by-decent matrix at the locus of interest.

Single-trait mixed models were fitted for all three beak dimensions (length, depth and width). We included the identity of the mother and the identity of the nest of rearing as random effects to account for maternal and early environment effects. These effects were small and are therefore only presented in the Supplement (Table S2, Supporting information). Whether the bird was dead or alive at measuring, its age at measuring in days (or age when it died), its inbreeding coefficient based on our five-generation pedigree and its sex were included as fixed effects to account for freezing, age-dependent changes, inbreeding depression and sexual dimorphism. We fitted models with and without correction for overall body size, as described above. We performed whole genome scans in 5-cM intervals and subsequently refined scans to 1 cM for chromosomes of interest, that is, regions that showed linkage in the initial scan.

Significance testing

The LRT statistic was used to assess the significance of a QTL effect

$$D = 2^{*} \ln(\mathrm{L_{QTL}}/(\mathrm{L_{poly}})$$

where D is the likelihood ratio (the test statistic of the LRT), $\mathrm{L_{QTL}}$ is the likelihood of the data given the fitted QTL model (model 2) and $\mathrm{L_{poly}}$ is the likelihood of the data given the fitted polygenic model without QTL effect (model 1). Because tests were performed at the boundary of possible values for variance components, we approximated *P*-values by comparing D to a 50 : 50 mixture of a χ^2_1 and a χ^2_0 distribution (George *et al.* 2000).

We used the formula provided by Lander & Kruglyak (1995) to calculate the significance threshold for the LRT assuming a map length of 1479.4 cM covering 32 chromosomes. The approximated *P*-values for significant genome-wide linkage (expecting one false-positive in 20 genome scans) and suggestive genome-wide linkage (expecting one false-positive in every genome scan) were $1.28{*}10^{-4}$ (D = 13.37) and $4.69{*}10^{-3}$ (D = 6.75), respectively.

Support intervals were calculated around linkage peaks to estimate the location of a trait locus. We defined them as the region around linkage peaks where $-\log_{10}(\mathrm{P})$ drops by one unit compared to the maximum value on that chromosome, that is, a more than 90% confidence region (Lander & Botstein 1989; Dupuis & Siegmund 1999; Pavlicev *et al.* 2008). Additionally, we defined approximately 95% confidence intervals by a $-1.5{*}\log_{10}(\mathrm{P})$ drop. D can be converted to LOD scores by using the formula LOD = D/(2*ln(10)) (Ziegler & König 2006).

Conversion of genetic to physical map and gene search

QTL effects were estimated between flanking markers. Although the physical location of every marker is

known, the physical location of QTL could only be estimated by linear interpolation from the genetic and physical position of the flanking markers. Whenever flanking markers had the same genetic position but different physical locations, we used the physical location of the more distant marker for calculating the physical location of the QTL effect (resulting in larger and hence more conservative support intervals). Furthermore, whenever genetic and physical marker orders were inconsistent, we checked whether all the markers within the QTL interval (based on the linkage map) were located inside the confidence interval on the physical map. If not, we expanded the interval to cover these markers. This was the case on chromosome *Tgu2*, but even if intervals were not expanded, the same candidate genes were inside the QTL peak regions. Thus, physical positions should be treated as approximations rather than exact locations.

To find genes that are located in the confidence intervals, we used the zebra finch genome (taeGut 3.2.4) of Ensembl 61 accessed via BioMart (http://www.biomart.org). We included all predicted and annotated genes.

Other statistics

Statistical analyses were performed with SPSS (SPSS Inc. 2009) and R 2.12.2 (R Development Core Team 2011). All tests described below were two-tailed, and *P*-values smaller than 0.05 were considered significant. We used linear models to investigate the relationships between beak dimensions (response) and five fixed effect predictors: sex (two levels), status (two levels, dead or alive), inbreeding coefficient (continuous) and age (continuous, measured in years). In an alternative model, we additionally fitted tarsus and wing length and the cube root of mass as covariates to control for overall body size. We converted estimates to standardized effect sizes (*d* for categorical and *r* for continuous predictors) (Nakagawa & Cuthill 2007).

Results

Phenotypic variation

All three beak dimensions displayed sexual dimorphism independent of whether measurements were taken on dead or alive birds (estimates were similar when splitting the data by status; details not shown). Females had significantly shorter, flatter and narrower beaks than males, with beak depth showing the strongest sexual dimorphism (Table 1, Fig. S1, Supporting information). The sexual dimorphism became even more pronounced when controlling for variation in body size (Table S3, Supporting information).

The status of the bird at the time of measurement, that is, whether it was dead or alive, had a significant effect on the measurements of all three beak dimensions. Possibly due to drying or freezing of the beak, dead birds had significantly longer but flatter and narrower beaks with small-to-moderate effect sizes (Table 1, Table S3, Supporting information). Inbred birds were significantly smaller in all three beak dimensions (Table 1), but this effect largely disappeared when controlling for overall body size (Table S3, Supporting

Table 1 Linear models with beak dimensions as dependent variables (measured in mm), sex and status (dead or alive when measured) as independent factors, and age (measured in years) and inbreeding coefficient (*f*) as covariates. All interactions between predictors were nonsignificant ($P > 0.25$). Standardized effect sizes were calculated according to Nakagawa & Cuthill (2007). *b* = slope, *d* = standardized effect size, *r* = partial correlation coefficient.

Dimension	Parameter	*b* ± SE	*t*-value	*P* value	Effect size
Beak length	Intercept	10.046 ± 0.056			
	$Sex^{female = 0}$	0.165 ± 0.030	5.5	$5*10^{-8}$	$d = 0.33$
	$Status^{alive = 0}$	0.075 ± 0.033	2.2	0.025	$d = 0.14$
	f	−1.062 ± 0.310	−3.4	$6*10^{-4}$	$r = -0.10$
	Age	0.065 ± 0.010	6.4	$2*10^{-10}$	$r = 0.19$
Beak depth	Intercept	8.275 ± 0.031			
	$Sex^{female = 0}$	0.186 ± 0.017	11.3	$5*10^{-28}$	$d = 0.69$
	$Status^{alive = 0}$	−0.096 ± 0.018	−5.2	$3*10^{-7}$	$d = -0.32$
	f	−0.448 ± 0.171	−2.6	$9*10^{-3}$	$r = -0.08$
	Age	0.009 ± 0.006	1.7	0.093	$r = 0.05$
Beak width	Intercept	6.589 ± 0.025			
	$Sex^{female = 0}$	0.083 ± 0.013	6.3	$3*10^{-10}$	$d = 0.39$
	$Status^{alive = 0}$	−0.077 ± 0.015	−5.2	$2*10^{-7}$	$d = -0.34$
	f	−0.603 ± 0.136	−4.4	$1*10^{-5}$	$r = -0.13$
	Age	0.010 ± 0.004	2.2	0.031	$r = 0.07$

information). Hence, inbreeding seems to affect absolute body size rather than beak size itself.

Beak length increased markedly with age (age at death or age at measuring for alive birds), while for depth and width, the age effect was only marginally significant (Table 1). Controlling for body size further reduced the age effects on beak depth and width (Table S3; $P = 0.32$ and $P = 0.23$, respectively, Supporting information) but had almost no impact on the age effect on beak length (Table S3; $P = 3*10^{-9}$, Supporting information).

In cross-sectional analyses like this, an increase in beak length with age could either reflect within-individual growth (longitudinally) or differential mortality. To separate these possibilities, a subset of 61 live birds was measured on two occasions that were on average 287 days apart. This confirmed that the elongation was indeed because of within-individual growth (elongation by 0.13 ± 0.03 mm, paired *t*-test: $t_{60} = 4.4$, $P = 4*10^{-5}$).

Quantitative genetics

Heritability estimates and genetic and phenotypic correlations for all three beak dimensions were estimated in three-trait animal models (Table 2, Table S1, Supporting information). For beak length, the heritability estimate was substantially lower than the estimates for depth and width; yet, coefficients of additive genetic variation were similar in all three dimensions (beak length: $CV_A = 3.216$; beak depth: $CV_A = 2.674$; beak width: $CV_A = 2.778$). The difference in heritability estimates was mainly due to higher permanent environment variation in beak length (beak length: $CV_{PE} = 3.372$; beak depth: $CV_{PE} = 1.852$; beak width: $CV_{PE} = 1.639$). Maybe as a consequence of this, genetic and phenotypic correlations between beak length and the other two dimensions were lower than the correlations between depth and width. Genetic correlations tended to be slightly stronger than the phenotypic correlations (Table 2). The conditional heritability of each of the three dimensions conditional on the additive genetic variance of the other two beak dimensions was similar (beak length: 0.351; beak depth: 0.322; beak width: 0.367).

Table 2 Phenotypic and genetic correlation matrix, narrow-sense heritability and additive genetic (co)variance (V_A) estimates for the three beak dimensions. In the ratios, part above the diagonal are pair-wise phenotypic Pearson correlation coefficients ± SE (all $P < 10^{-25}$), below the diagonal are pair-wise genetic correlations ± SE and in bold on the diagonal are the narrow-sense heritability estimates ± SE (all $P < 10^{-8}$ using likelihood ratio tests). Sample sizes for the different dimensions ranged from 1085 to 1088 individuals.

	Dimension	Length	Depth	Width
Ratios	Length	**0.470 ± 0.058**	0.379 ± 0.028	0.307 ± 0.029
	Depth	0.465 ± 0.079	**0.648 ± 0.046**	0.612 ± 0.024
	Width	0.462 ± 0.075	0.690 ± 0.048	**0.736 ± 0.041**
V_A	Length	**0.113 ± 0.018**		
	Depth	0.035 ± 0.008	**0.050 ± 0.005**	
	Width	0.028 ± 0.006	0.028 ± 0.004	**0.034 ± 0.003**

QTL signals

We found four regions with suggestive genome-wide linkage (chromosomes *Tgu2*, *Tgu5*, *Tgu6* and *Tgu20*) for beak length controlled for body size (Fig. 1 and Fig. S2, Supporting information). The regions on chromosomes *Tgu2*, *Tgu5* and *Tgu6* also showed suggestive linkage for absolute beak length, and additionally, a region on chromosome *Tgu4* displayed suggestive linkage when not controlling for size. The four suggestive linkage peaks on chromosomes *Tgu2*, *Tgu5*, *Tgu6* and *Tgu20* together explained 34% of the total phenotypic variance when controlling for body size and measured at the point of highest significance. This corresponds to 106% of the heritability and is hence evidently an overestimate (see Discussion). QTL peaks covered large genomic regions spanning 20–64 cM on the genetic map in the 1.5-LOD drop intervals and 19–50 cM when considering the 1-LOD drop intervals. In total, this corresponded to 228 Mb covering 2649 genes and 225 Mb covering 2601 genes, respectively (Table 3).

Five regions showed suggestive genome-wide linkage (chromosomes *Tgu1*, *Tgu2*, *Tgu5*, *Tgu9* and *Tgu17*) for beak depth, absolute or controlled for overall body size (Fig. 2 and Fig. S3, Supporting information). The variance explained by these five regions amounts to 46% of the phenotypic variance and 70% of the heritability estimate when controlling for body size. Both the 1.5-LOD drop intervals and the 1-LOD drop intervals for the QTL peaks spanned large genomic regions, ranging from 15–56 cM to 11–44 cM, respectively. Accordingly, the 1.5-LOD and 1-LOD drop confidence intervals covered 319 Mb comprising 3394 genes and 299 Mb comprising 3180 genes (Table 3).

For beak width, two regions showed significant genome-wide linkage (chromosomes *Tgu2* and *Tgu4*) and three regions suggestive linkage (chromosomes *Tgu1*, *Tgu9* and *Tgu10*) (Fig. 3 and Fig. S4, Supporting information). Mapping absolute beak width gave significant linkage peaks on chromosomes *Tgu1* and *Tgu2* and suggestive ones on chromosomes *Tgu4* and *Tgu10*. Together, the five peak regions for relative beak width explained 46% of the phenotypic variance when controlling for body size, which corresponded to 63% of the heritability estimate. The 1.5-LOD drop and the

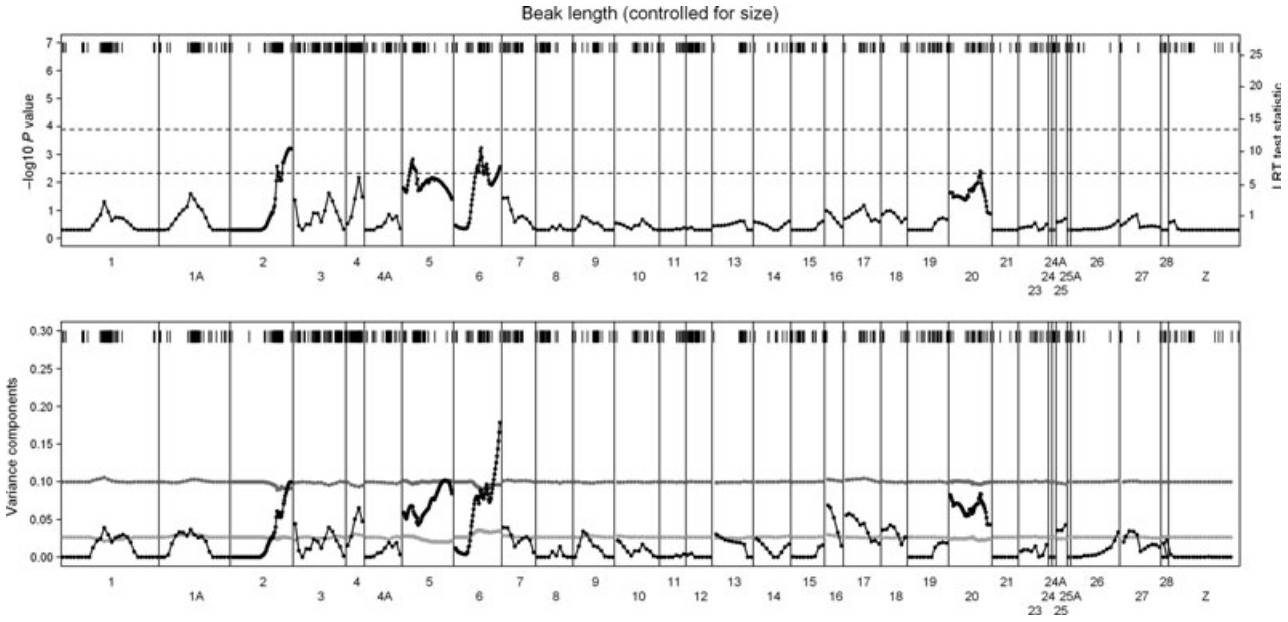


Fig. 1 Genome-wide linkage scan for additive genetic variance in beak length controlled for body size scanned in 5 cM intervals (and 1 cM intervals on chromosomes showing suggestive genome-wide linkage). The upper graph shows *D* and *P* values, the lower graph shows variance components explained by each interval, i.e. effect sizes. Chromosome identities are given on the abscissa. Marker positions are given by the vertical lines in the upper part of each graph. The upper and lower dotted lines represent significant and suggestive genome-wide linkage thresholds, respectively. Additive genetic variances are shown in black, maternal effects in dark grey and early rearing effects in light grey.

1-LOD drop confidence intervals covered large genomic regions (12–56 cM corresponding to 321 Mb including 3262 genes and 11–31 cM corresponding to 275 Mb and 2798 genes, respectively) (Table 3).

We identified genes from the literature, which are known to influence at least one beak dimension, and are supported by either gain- or loss-of-function experiments. All genes known to have an effect on beak morphology lay at least within the suggestive QTL regions except *BMP2*, *DKK3* and *RALDH3* (but see discussion for *DKK3* residing on *Tgu5_random*; Table 4). With the exception of *CTNNB1* (beak length) and *RALDH2* (beak width), all are located within the 1-LOD drop intervals. *CTNNB1*, *TGFBR2*, *BMP4*, *CALM1* and *FGF8* were present in suggestive QTL regions of those beak dimensions, which they are known to influence. Although there is evidence that *BMP4* expression influences all beak dimensions, it is only present in the beak length suggestive QTL peak region. *SHH*, *RALDH2* and *BMP7* are located in suggestive QTL regions of at least one beak dimension, but not for those dimensions that they have been shown to influence.

Five peak regions (1 for depth, 1 for width and 2 for depth and width) do not host any of the candidates. One of them is the significant QTL peak region for beak width on chromosome *Tgu4*, the other three are suggestive QTL peaks on chromosomes *Tgu1*, *Tgu9* and *Tgu17*. The other significant QTL peak region for beak width on chromosome *Tgu2* harbours three candidate genes (*CTNNB1*, *SHH* and *TGFBR2*), but none of them is known to influence beak width.

Discussion

Sexual dimorphism in beak size

The sexual dimorphism in beak size has to our knowledge never been described in zebra finches before, but it is a well-studied phenomenon in other species of birds and is usually interpreted as an adaptation to ecological niche separation between the sexes (Selander 1966; Gosler 1987; Radford & Du Plessis 2003). However, in zebra finches, niche separation is unlikely because pairs stay together throughout the year (including the nonbreeding season) and they typically forage together on the same kinds of food items (Zann 1996). However, beak colour may be a sexually selected trait in the zebra finch (reviewed in Collins & ten Cate 1996), and sexual selection on colour might also have led to sexual selection on beak size because of sensory exploitation (Ryan 1998). In our captive population, we do not find female preferences for males with redder (Forstmeier & Birkhead 2004) or larger beaks (unpublished data) and we discuss the cur-

Table 3 Significant and suggestive additive QTL peak regions and their 1.5-LOD and 1-LOD drop intervals for all three beak dimensions. The two significant QTL peaks are in bold type.

					Δ1.5 LOD				Δ1 LOD			
Dimension	Chromosome	*D* value	Significance ($-\log_{10}P$)	Effect size (% phenotypic variance explained) ± SE	Genetic position (cM)	Physical position (Mb)	Width (Mb)	# genes	Genetic position (cM)	Physical position (Mb)	Width (Mb)	# genes
Length	*Tgu2*	10.47	3.22	9.92 ± 4.17	59–79	29.99–155.29	125.30	1020	60–79	29.99–155.29	125.30	1020
	Tgu5	8.82	2.83	6.67 ± 3.15	0–64	3.01–61.64	58.63	890	5–55	4.46–60.39	55.93	854
	Tgu6	10.54	3.23	8.97 ± 3.56	24–59	3.91–35.16	31.25	450	27–59	4.20–35.16	30.96	447
	Tgu20	7.06	2.41	8.40 ± 3.89	0–51	0.86–14.17	13.31	289	0–46	0.86–13.43	12.57	280
Depth	*Tgu1*	9.34	2.95	8.48 ± 3.48	45–101	1.21–106.72	105.51	1026	50–94	1.73–104.15	102.42	1021
	Tgu2	13.30	3.88	8.67 ± 3.45	19–70	1.05–146.92	145.87	1252	30–67	2.15–142.03	139.88	1198
	Tgu5	10.90	3.32	8.42 ± 3.61	0–24	3.01–53.80	50.79	768	0–16	3.01–45.42	42.41	660
	Tgu9	11.84	3.54	8.13 ± 3.20	26–41	13.30–24.11	10.81	183	27–38	15.68–23.35	7.67	140
	Tgu17	8.28	2.70	12.04 ± 5.23	0–34	1.65–8.15	6.50	165	0–31	1.65–7.85	6.20	161
Width	*Tgu1*	8.08	2.65	5.61 ± 2.71	38–94	1.73–110.31	108.58	1042	53–84	2.47–100.26	97.79	966
	Tgu2	**23.34**	**6.17**	**11.83 ± 3.67**	**27–64**	**1.76–118.07**	**116.31**	**996**	**34–61**	**2.66–93.54**	**90.88**	**767**
	Tgu4	**18.12**	**4.98**	**9.43 ± 3.18**	**8–20**	**4.62–63.17**	**58.55**	**583**	**9–20**	**7.25–63.17**	**55.92**	**558**
	Tgu9	7.34	2.47	7.15 ± 3.19	13–51	3.86–25.78	21.92	374	19–47	5.73–25.54	19.81	299
	Tgu10	8.32	2.71	12.02 ± 5.37	0–37	1.08–17.12	16.04	267	0–28	1.08–11.32	10.24	208

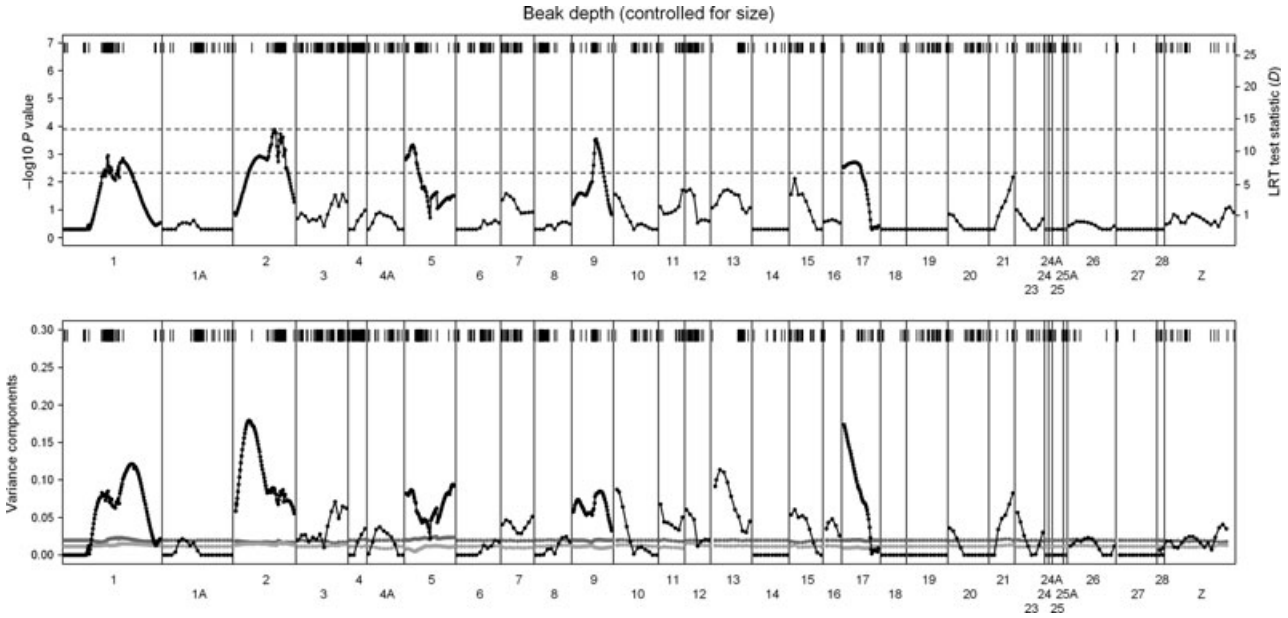


Fig. 2 Genome-wide linkage scan for additive genetic variance in beak depth controlled for size scanned in 5 cM intervals (and 1 cM intervals on chromosomes showing suggestive genome-wide linkage). For detailed information see Fig. 1.

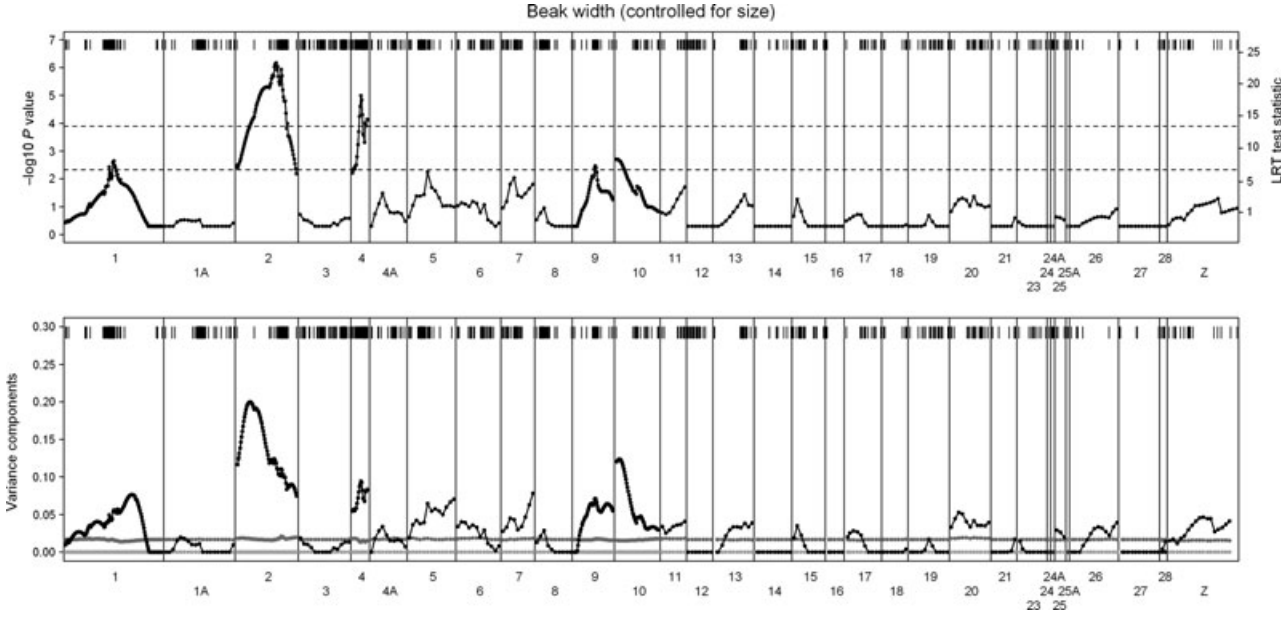


Fig. 3 Genome-wide linkage scan for additive genetic variance in beak width controlled for size scanned in 5 cM intervals (and 1 cM intervals on chromosomes showing significant and suggestive genome-wide linkage). For detailed information see Fig. 1.

rent evidence for sexual selection on beak colour elsewhere (Schielzeth *et al*. 2012b). Still it is possible that beak size has been subject to sexual selection in the wild, where both beak colour (Zann 1996) and beak size (own unpublished data) have been found to be sexually dimorphic.

Examples of beak size dimorphism in species with otherwise equally sized sexes are rare, and there is even less evidence for sexual selection acting on beak size in any species. An interesting exception, however, is the emerald toucanet (*Aulacorhynchus prasinus*), a frugivorous species with elaborate beak coloration. Male toucanets have 20% longer beaks than females, while there is no difference in body size. As in the zebra finch, the feeding ecology and the roles in nesting behaviour are similar in both sexes, suggesting that sexual selection

Table 4 Genes known to be functionally related to beak dimensions and their presence in QTL support intervals in our study. Bold-type x are genes located in 1-LOD drop support intervals, normal-type x are genes that reside in 1.5-LOD drop intervals. (x) are genes in support intervals for a beak dimension other than the known one. Length = beak length, depth = beak depth, width = beak width.

			Evidence for functional relation			In QTL peaks		
Candidate gene	Chromosome	Physical position (Mb)	Length	Depth	Width	Length	Depth	Width
CTNNB1[*]	*Tgu2*	64.38–64.39	x	x		x	**x**	**(x)**
SHH[†,‡]	*Tgu2*	8.96–8.97	x				**(x)**	**(x)**
TGFBR2[*]	*Tgu2*	60.12–60.18	x	x		**x**	**x**	**(x)**
BMP2[§]	*Tgu3*	25.96–25.96		x	x			
BMP4[§,¶,‡]	*Tgu5*	59.39–59.39	x	x	x	**x**		
CALM1[**]	*Tgu5*	44.69–44.70	x			**x**	**(x)**	
DKK3[*]	*Tgu5*	1.62–1.65	x	x				
FGF8[†]	*Tgu6*	22.05–22.15	x			**x**		
RALDH2[††]	*Tgu10*	6.79–6.84	x					(x)
RALDH3[††]	*Tgu10*	17.98–18.01	x					
BMP7[§]	*Tgu20*	12.99–13.03		x	x	**(x)**		

[*]Mallarino *et al.* (2011); [†]Abzhanov & Tabin (2004); [‡]Wu *et al.* (2006); [§]Abzhanov *et al.* (2004); [¶]Wu *et al.* (2004); [**]Abzhanov *et al.* (2006); [††]Song *et al.* (2004).

was involved in creating the sexual dimorphism (Riley & Smith 1992).

Beak size and growth

Because beaks are worn off when used, the rhampotheca needs to grow constantly to compensate abrasion (e.g. Matthysen 1989). Usually, fledglings have smaller beaks than their parents and beak growth might be observed until birds reach sexual maturity (Smith *et al.* 1986). Beak size changes in adults have rarely been described, perhaps because of small effect sizes and the resulting difficulty of detection. Price & Grant (1984) showed that in Darwin's medium ground finches (*Geospiza fortis*), beak length and width increased slightly between 1 and 3 years of age. Our result suggests a lifelong increase in zebra finch beak length, although this might be specific to the captive environment, because of reduced abrasion in cages and aviaries (Fox 1952).

Quantitative genetics of beak morphology

All three beak dimensions show considerable additive genetic variance. The narrow-sense heritability estimates for beak depth and width are among the highest estimates for morphological traits in general and certainly for beak dimensions in avian species (Keller *et al.* 2001; Åkesson *et al.* 2008; Tschirren & Postma 2010). This might reflect the reduced environmental variation in the captive environment (but see the review of Weigensberg & Roff 1996 for sometimes higher heritability estimates in the wild as compared to the laboratory). Heritability estimates for beak length are considerably lower, but estimates of CV_A are similar indicating a similar potential for proportional change. Beaks are increasing in length with age in our population, and in a few individuals (approximately 2% of the population), the upper mandible is growing too long and has to be trimmed from time to time to avoid malformations and starvation. Birds with recently trimmed beaks were excluded from the analysis, but the data might include some birds with earlier beak trimming. This treatment causes an increase in the variance because of environmental effects and a corresponding decrease in the heritability estimate.

Genetic correlations between the three measures of beak size were high. This may be due to pleiotropic effects of the same gene(s) or to coincidental linkage disequilibrium between genes that affect more than one trait. Unfortunately, we cannot separate the two possibilities, because in our captive population, linkage disequilibrium is high in the centre of the macrochromosomes and spans large genomic regions (Backström *et al.* 2010). In accordance with the strong genetic correlation between beak depth and width, we found more overlap of QTL peak regions between these two beak dimensions than between them and beak length. Nevertheless, all beak dimensions have exclusive linkage peaks (2 for beak length and width and 1 for beak depth), which means that there may be additive genetic variance for each dimension independent of the other beak dimensions. In accordance with this, the conditional heritability estimates are still quite high, indicating the potential for all beak dimen-

sions to evolve independently of the other dimensions. Genetic and phenotypic correlations were approximately similar in strength, which is expected if the heritability is high and genetic influences therefore also dominate the phenotypic correlation (Hadfield *et al.* 2007).

QTL signals

QTL maps show four regions of suggestive genome-wide linkage for beak length, five regions for beak depth and three for beak width. Two additional regions reached genome-wide significant linkage for beak width. Effect sizes are similar in all peak regions (ranging from 5.6% to 12.1%) and together explain 63%, 70% and 106% of the heritability estimate for beak width, depth and length, respectively. Effect size estimates are known to be inflated in regions where the power of detecting a QTL is low (Beavis 1998). Because this power is not constant in a genome scan, we defined QTL regions by means of the significance peaks rather than the effect size peaks. Subsequently, we calculated effect sizes point-wise at significance peaks. However, effect size estimates are still inflated, which reflects the general problem of estimating both the genomic regions influencing a trait and their effect sizes using the same data set. As these estimates are not independent, higher significance tends to coincide with higher effect size estimates (Göring *et al.* 2001), which is also true for linear models in general (Forstmeier & Schielzeth 2011). The variance explained by QTLs should therefore be considered an upper limit of the true effect size.

Five and three of the 11 known candidate genes are present in suggestive and significant QTL peak regions, respectively. From those three genes that are not included in the QTL regions (*BMP2*, *DKK3* and *RALDH3*), *DKK3* is located on chromosome *Tgu5* but is not covered by our linkage map (it is located on chromosome *Tgu5_random*, that is, it is known to be on chromosome *Tgu5* but its exact position could not be determined). DKK3 acts downstream of TGFBR2 and CTNNB1 (Mallarino *et al.* 2011), such that the association between *DKK3* expression levels and beak morphology may be more parsimoniously explained by polymorphisms in *TGFBR2* or *CTNNB1*.

CTNNB1, *CALM1*, *FGF8* and *TGFBR2* are found in the QTL peak regions of those beak dimensions for which they were previously described to be influential (Abzhanov & Tabin 2004; Abzhanov *et al.* 2006; Mallarino *et al.* 2011). *CTNNB1* and *TGFBR2* are found in the peak regions of all three beak dimensions on chromosome *Tgu2*. Because these peaks show the highest significance values in all three beak dimensions, *CTNNB1* and *TGFBR2* are strong candidates for influencing beak morphology in our population. *BMP4* and *BMP7* are found in QTL regions showing suggestive genome-wide linkage to variation in beak length. Abzhanov *et al.* (2004) showed that these two genes are connected to variation in beak depth and width, while Wu *et al.* (2004, 2006) showed that also beak length varies with changes in *BMP4* expression. *SHH* is found in the QTL peak region of beak depth and width and *RALDH2* is found in the peak region of beak width, although previous studies suggest that both genes contribute to variation in beak length.

All the above genes known to affect beak morphology were identified using expression level analysis or expression level manipulations. These expression differences could stem from polymorphisms acting in *cis* (e.g. variation in promoter regions) or in *trans* (e.g. transcription factors) to the affected genes. Even though in total 5355 genes are covered by our 95% confidence intervals, that is, 29% of all known and predicted genes in the zebra finch genome (Ensembl 61), our QTL peak regions match rather well with the locations of those genes known to affect beak morphology. Thus, it seems likely that the causal polymorphisms reside predominantly in *cis*-regulatory elements of these genes rather than in *trans*-acting factors, which is in line with the general assumption of a large contribution of *cis*-regulatory mutations to morphological evolution (Wray 2007). However, it is difficult to test the significance of overlap between the candidate gene locations and QTL peak regions. The probability of drawing 8 or more out of 11 possible successes when 29% of all genes are located in our peak regions is only $P = 0.003$; yet, this is certainly anticonservative because the 11 candidate genes are not independently distributed across the genome.

Furthermore, *cis*-acting polymorphisms are known to have larger effect sizes than *trans*-acting factors (Veyrieras *et al.* 2008; Cheung & Spielman 2009). Because *cis*-acting variants are in close proximity to each other, their effect sizes add up and appear as one QTL with large effect, whereas *trans*-acting factors are more dispersed and thus their effects cannot accumulate. This may also explain our result of a predominance of causal *cis*-regulatory polymorphisms. Currently, there is a growing body of evidence that *trans*-acting factors play an important role in gene expression regulation too (Cheung *et al.* 2010). Also, methods for decomposing expression variation into *cis*- and *trans*-regulatory components usually neglect the possibility of *cis*-*trans* compensation, thereby spuriously exaggerating the contribution of *cis*-acting polymorphisms (Takahasi *et al.* 2011). Thus, the contribution of *cis*-acting variants to gene expression variation is probably still overestimated.

Finally, three suggestive and one significant QTL do not include any of the known candidate genes. This means that there may be still unknown genetic variants that influence beak morphology in the zebra finch. These regions might include coding genes and/or transcription factors.

Acknowledgements

We thank S. Bauer, E. Bodendorfer, A. Grötsch, J. Hacker, M. Halser, K. Martin, J. Minshull, P. Neubauer, F. Preininger, M. Ruhdofer and A. Türk for animal care and help with breeding; M. Schneider for laboratory work; N. Backström and H. Mellenius for kindly providing the linkage map. U.K. is part of the International Max Planck Research School for Organismal Biology. This study was funded by financial support from the Max Planck Society (B.K.), the Swedish Research Council (H.E.), the Knut and Alice Wallenberg Foundation (H.E.), the German Research Foundation grants FO340/1-1, FO340/1-2 and FO340/1-3 (W.F.) and the European Union grant HPMF-CT-2002-01871 (W.F.).

References

Abzhanov A, Tabin CJ (2004) *Shh* and *Fgf8* act synergistically to drive cartilage outgrowth during cranial development. *Developmental Biology*, **273**, 134–148.

Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004) *Bmp4* and morphological variation of beaks in Darwin's finches. *Science*, **305**, 1462–1465.

Abzhanov A, Kuo WP, Hartmann C *et al.* (2006) The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature*, **442**, 563–567.

Åkesson M, Bensch S, Hasselquist D, Tarka M, Hansson B (2008) Estimating heritabilities and genetic correlations: comparing the 'animal model' with parent-offspring regression using data from a natural population. *PLoS One*, **3**, e1739.

Backström N, Forstmeier W, Schielzeth H *et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research*, **20**, 485–495.

Badyaev AV (2010) The beak of the other finch: coevolution of genetic covariance structure and developmental modularity during adaptive evolution. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **365**, 1111–1126.

Beavis WD (1998) QTL analysis: power, precision, and accuracy. In: *Molecular Dissection of Complex Traits* (ed. Paterson AH), pp. 145–162. CRC Press, Boca Raton, Florida.

Cheung VG, Spielman RS (2009) Genetics of human gene expression: mapping DNA variants that influence gene expression. *Nature Reviews Genetics*, **10**, 595–604.

Cheung VG, Nayak RR, Wang IXR *et al.* (2010) Polymorphic *cis*- and *trans*-regulation of human gene expression. *Plos Biology*, **8**, e1000480.

Collins SA, ten Cate C (1996) Does beak colour affect female preference in zebra finches? *Animal Behaviour*, **52**, 105–112.

Darwin C (1859) *On the Origin of Species by Means of Natural Selection*. John Murray, London.

Dupuis J, Siegmund D (1999) Statistical methods for mapping quantitative trait loci from a dense set of markers. *Genetics*, **151**, 373–386.

Falconer D, Mackay T (1996) *Introduction to Quantitative Genetics*. Longmann, Harlow, UK.

Fan JB, Oliphant A, Shen R *et al.* (2003) Highly parallel SNP genotyping. *Cold Spring Harbor Symposia on Quantitative Biology*, **68**, 69–78.

Flint J, Mackay TFC (2009) Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Research*, **19**, 723–733.

Forstmeier W (2011) Women have relatively larger brains than men: a comment on the misuse of general linear models in the study of sexual dimorphism. *Anatomical Record*, **294**, 1856–1863.

Forstmeier W, Birkhead TR (2004) Repeatability of mate choice in the zebra finch: consistency within and between females. *Animal Behaviour*, **68**, 1017–1028.

Forstmeier W, Schielzeth H (2011) Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology*, **65**, 47–55.

Forstmeier W, Leisler B, Kempenaers B (2001) Bill morphology reflects female independence from male parental help. *Proceedings of the Royal Society B-Biological Science*, **268**, 1583–1588.

Forstmeier W, Segelbacher G, Mueller JC, Kempenaers B (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular Ecology*, **16**, 4039–4050.

Fox W (1952) Behavioral and evolutionary significance of the abnormal growth of beaks of birds. *Condor*, **54**, 160–162.

George AW, Visscher PM, Haley CS (2000) Mapping quantitative trait loci in complex pedigrees: a two-step variance component approach. *Genetics*, **156**, 2081–2092.

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009) *ASReml User Guide*. VSN International, Hemel Hempstead, UK.

Göring HHH, Terwilliger JD, Blangero J (2001) Large upward bias in estimation of locus-specific effects from genomewide scans. *American Journal of Human Genetics*, **69**, 1357–1369.

Gosler AG (1987) Pattern and process in the bill morphology of the great tit *Parus major*. *Ibis*, **129**, 451–476.

Grant PR (1999) *Ecology and Evolution of Darwin's Finches*. Princeton University Press, Princeton, New Jersey.

Grant PR, Grant BR (2002) Unpredictable evolution in a 30-year study of Darwin's finches. *Science*, **296**, 707–711.

Groeneveld E, Kovač M, Mielenz N (2008) *VCE User's Guide and Reference Manual. Version 6.0*. Institute of Farm Animal Genetics, Friedrich Loeffler Institute (FLI), Mariensee, Germany.

Hadfield JD, Nutall A, Osorio D, Owens IPF (2007) Testing the phenotypic gambit: phenotypic, genetic and environmental correlations of colour. *Journal of Evolutionary Biology*, **20**, 549–557.

Hansen TF, Houle D (2008) Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology*, **21**, 1201–1219.

Hernandez-Sanchez J, Grunchec JA, Knott S (2009) A web application to perform linkage disequilibrium and linkage analyses on a computational grid. *Bioinformatics*, **25**, 1377–1383.

Herrel A, Soons J, Aerts P *et al.* (2010) Adaptation and function of the bills of Darwin's finches: divergence by feeding type and sex. *Emu*, **110**, 39–47.

Houle D (1992) Comparing evolvability and variability of quantitative traits. *Genetics*, **130**, 195–204.

Hu D, Marcucio RS, Helms JA (2003) A zone of frontonasal ectoderm regulates patterning and growth in the face. *Development*, **130**, 1749–1758.

Keller LF, Grant PR, Grant BR, Petren K (2001) Heritability of morphological traits in Darwin's Finches: misidentified paternity and maternal effects. *Heredity*, **87**, 325–336.

Lander ES, Botstein D (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121**, 185–199.

Lander E, Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics*, **11**, 241–247.

Lee SH, Fu KK, Hui JN, Richman JM (2001) Noggin and retinoic acid transform the identity of avian facial prominences. *Nature*, **414**, 909–912.

Mallarino R, Grant PR, Grant BR *et al.* (2011) Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proceedings of the National Academy of Sciences USA*, **108**, 4057–4062.

Matthysen E (1989) Seasonal variation in bill morphology of nuthatches *Sitta europaea* – dietary adaptations or consequences? *Ardea*, **77**, 117–125.

Nakagawa S, Cuthill IC (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*, **82**, 591–605.

Pavlicev M, Kenney-Hunt JP, Norgard EA *et al.* (2008) Genetic variation in pleiotropy: differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. *Evolution*, **62**, 199–213.

Podos J (2001) Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature*, **409**, 185–188.

Podos J, Nowicki S (2004) Beaks, adaptation, and vocal evolution in Darwin's finches. *BioScience*, **54**, 501–510.

Pong-Wong R, George AW, Woolliams JA, Haley CS (2001) A simple and rapid method for calculating identity-by-descent matrices using multiple markers. *Genetics Selection Evolution*, **33**, 453–471.

Price TD, Grant PR (1984) Life history traits and natural selection for small body size in a population of Darwin's finches. *Evolution*, **38**, 483–494.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Radford AN, Du Plessis MA (2003) Bill dimorphism and foraging niche partitioning in the green woodhoopoe. *Journal of Animal Ecology*, **72**, 258–269.

Riley CM, Smith KG (1992) Sexual dimorphism and foraging behavior of emerald toucanets *Aulacorhynchus prasinus* in Costa Rica. *Ornis Scandinavica*, **23**, 459–466.

Ryan MJ (1998) Sexual selection, receiver biases, and the evolution of sex differences. *Science*, **281**, 1999–2003.

Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2011) Data from: QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. *Dryad Digital Repository*, doi:10.5061/dryad.r044b.

Schielzeth H, Forstmeier W, Kempenaers B, Ellegren H (2012a) QTL linkage mapping of wing length in zebra finch using genome-wide single nucleotide polymorphisms markers. *Molecular Ecology*, **21**, 329–339.

Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2012b) QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. *Evolution*, **66**, 18–30.

Schneider RA, Hu D, Rubenstein JL, Maden M, Helms JA (2001) Local retinoid signaling coordinates forebrain and facial morphogenesis by maintaining FGF8 and SHH. *Development*, **128**, 2755–2767.

Selander RK (1966) Sexual dimorphism and differential niche utilization in birds. *Condor*, **68**, 113–151.

Smith JNM, Arcese P, Schluter D (1986) Song sparrows grow and shrink with age. *Auk*, **103**, 210–212.

Song Y, Hui JN, Fu KK, Richman JM (2004) Control of retinoic acid synthesis and FGF expression in the nasal pit is required to pattern the craniofacial skeleton. *Developmental Biology*, **276**, 313–329.

SPSS Inc. (2009) *PASW Statistics for Windows, Rel. 18.0.0*. SPSS Inc., Chicago, Illinois.

Svensson L (1992) *Identification Guide to European Passerines*, 4th edn. L. Svensson, Stockholm.

Takahasi KR, Matsuo T, Takano-Shimizu-Kouno T (2011) Two types of *cis-trans* compensation in the evolution of transcriptional regulation. *Proceedings of the National Academy of Sciences USA*, **108**, 15276–15281.

Tschirren B, Postma E (2010) Quantitative genetics research in zebra finches: where we are and where to go. *Emu*, **110**, 268–278.

Veyrieras JB, Kudaravalli S, Kim SY *et al.* (2008) High-resolution mapping of expression-QTLs yields insight into human gene regulation. *Plos Genetics*, **4**, e1000214.

Warren WC, Clayton DF, Ellegren H *et al.* (2010) The genome of a songbird. *Nature*, **464**, 757–762.

Weigensberg I, Roff DA (1996) Natural heritabilities: Can they be reliably estimated in the laboratory? *Evolution*, **50**, 2149–2157.

Wray GA (2007) The evolutionary significance of *cis*-regulatory mutations. *Nature Reviews Genetics*, **8**, 206–216.

Wu P, Jiang TX, Suksaweang S, Widelitz RB, Chuong CM (2004) Molecular shaping of the beak. *Science*, **305**, 1465–1466.

Wu P, Jiang TX, Shen JY, Widelitz RB, Chuong CM (2006) Morphoregulation of avian beaks: comparative mapping of growth zone activities and morphological evolution. *Developmental Dynamics*, **235**, 1400–1412.

Zann RA (1996) *The Zebra Finch: A Synthesis of Field and Laboratory Studies*. Oxford University Press, USA.

Ziegler A, König IR (2006) *A Statistical Approach to Genetic Epidemiology: Concepts and Applications*. John Wiley & Sons, Incorporated, Weinheim, Germany.

U.K. is an evolutionary geneticist interested in intragenomic conflicts and the genetic architecture of quantitative traits. H.S. is an evolutionary biologist and is interested in the genetic architecture of quantitative traits. B.K. is a behavioural ecologist interested in sexual selection and the evolution of mating behaviour. H.E. interest is the integration of genomics and evolutionary biology. W.F. is an evolutionary geneticist and is working on the genetics of sexual behaviour in the zebra finch.

Data accessibility

Pedigree and genotype data as well as a linkage map: DRYAD entry doi: 10.5061/dryad.r044b.

Phenotypes (Schielzeth *et al*. 2011): DRYAD entry doi:10.5061/dryad.9vq45n5s.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Age related changes in beak dimensions.

Fig. S2 Fine-mapping linkage scan for additive genetic variation in beak length controlled for body size scanned in 1 cM intervals.

Fig. S3 Fine-mapping linkage scan for additive genetic variation in beak depth controlled for size scanned in 1 cM intervals.

Fig. S4 Fine-mapping linkage scan for additive genetic variation in beak width controlled for size scanned in 1 cM intervals.

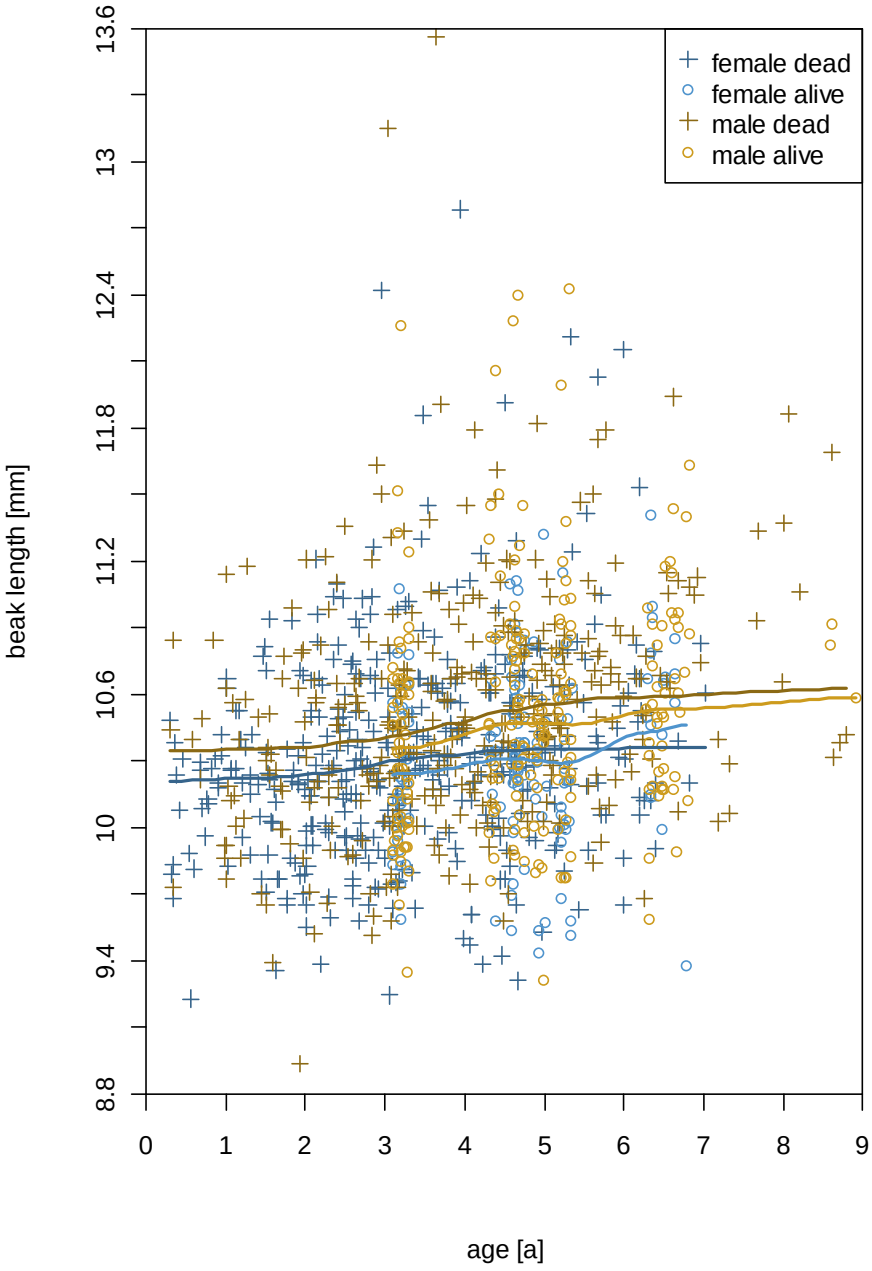
Table S1 Variance components (standardized by the phenotypic variance) and correlation estimates along with their standard errors for all three beak dimensions.

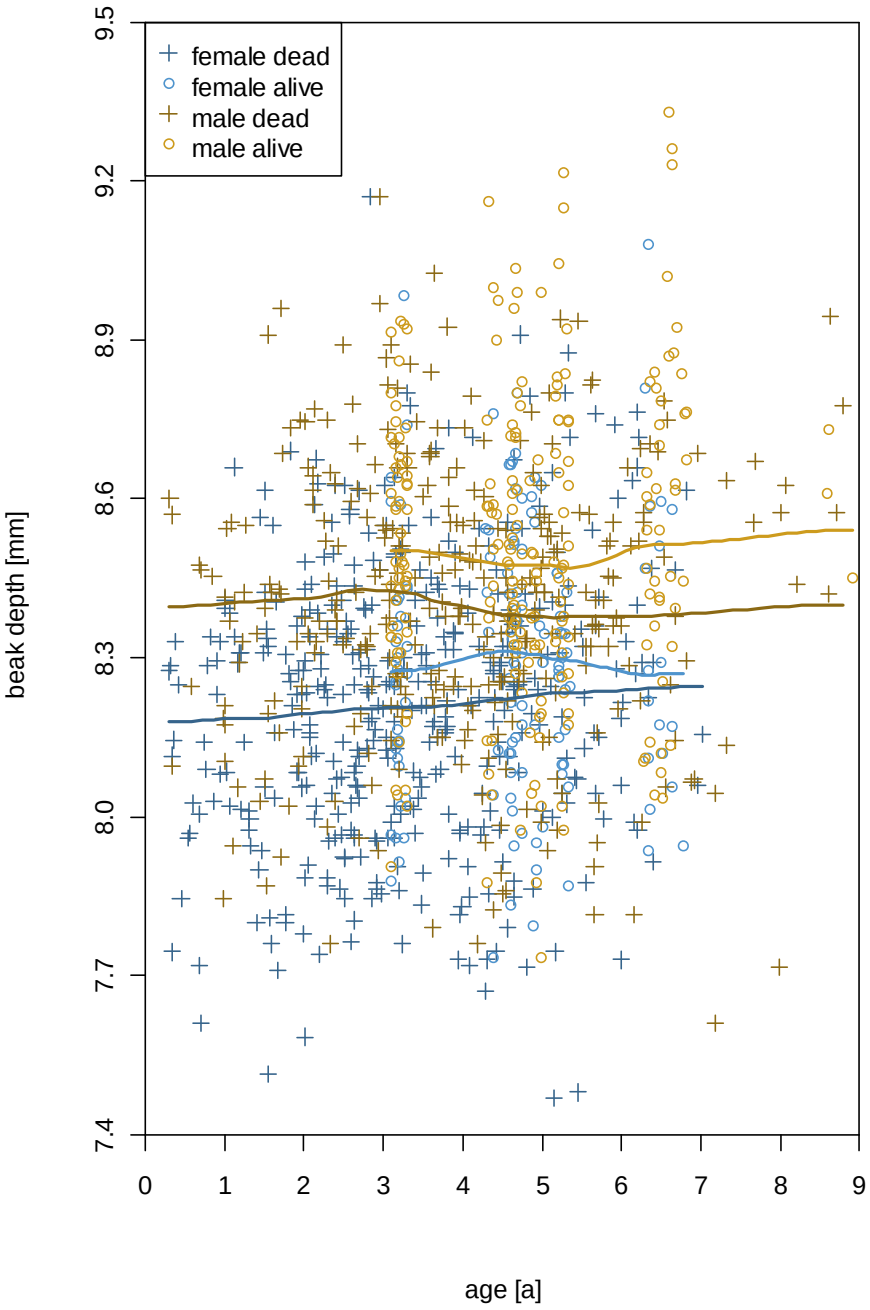
Table S2 Maternal and early rearing effects estimated in GridQTL by fitting mother identity and brood identity as random effects.

Table S3 Linear models with beak dimensions as dependent variables (measured in mm), sex and status as factorial predictors (two levels each), age (measured in years), inbreeding coefficient (*f*) and cube root of mass (measured in g), tarsus length and wing length (measured in mm) as covariates.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Supplement


female dead
female alive
male dead
male alive
beak length [mm]
8.8
9.4
10
10.6
11.2
11.8
12.4
13
13.6
0
1
2
3
4
5
6
7
8
9
age [a]


female dead
female alive
male dead
male alive
beak depth [mm]
9.5
9.2
8.9
8.6
8.3
8.0
7.7
7.4
0
1
2
3
4
5
6
7
8
9
age [a]

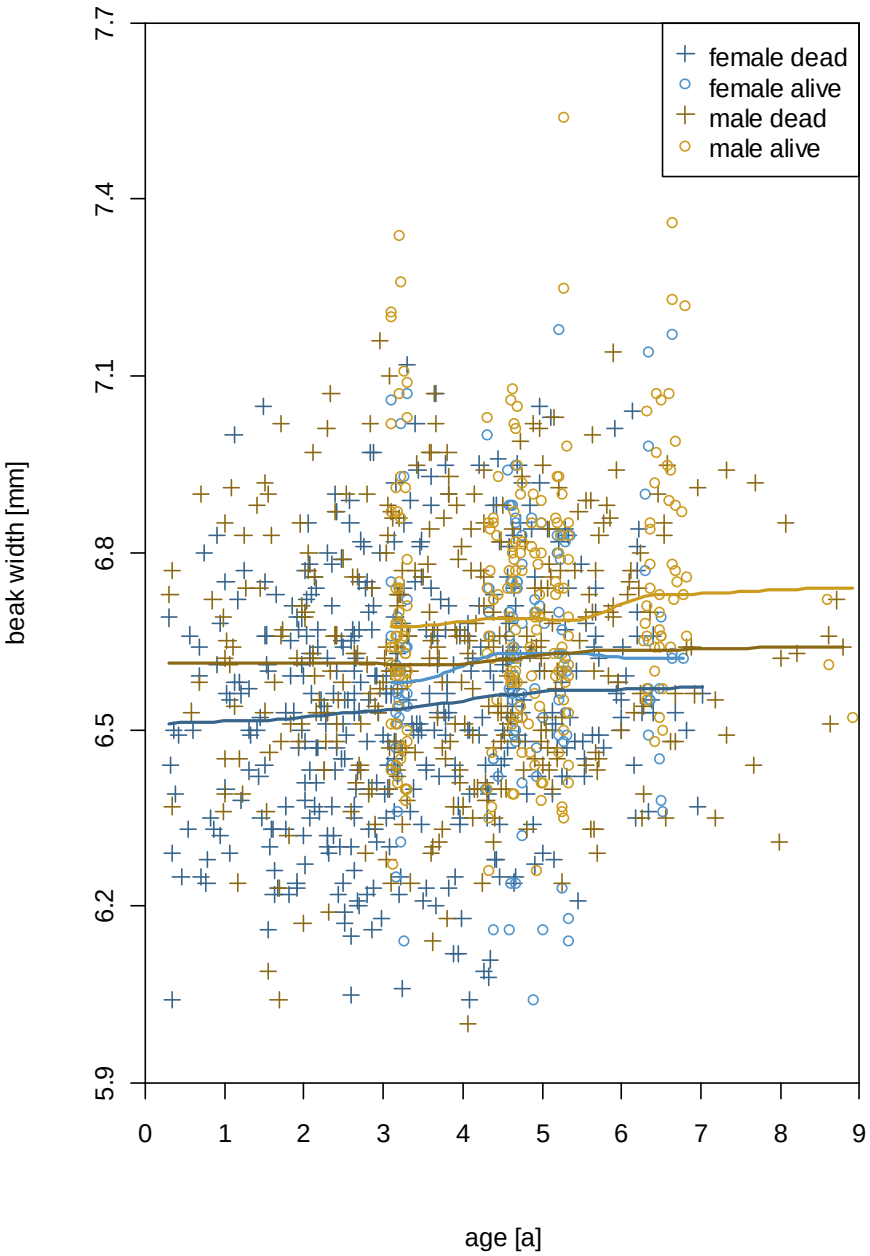


Figure S1: *Age related changes in beak dimensions. Data are separated by sex and whether birds were alive or dead. To illustrate potential curvature we show local polynomial regression lines with 0 degree of local polynomial used (using the loess function implemented in R). For each beak dimension we give coefficients of determination (R^2), P values and slopes of linear regression lines with 1 degree of freedom. (A) beak length (female dead birds: $R^2 = 0.043$, $P = 3*10^{-5}$, slope ± SE = 0.066 ± 0.016; female alive birds: $R^2 = 0.031$, P = 0.04, slope ± SE = 0.070 ± 0.034; male dead birds: $R^2 = 0.073$, $P = 6*10^{-7}$, slope ± SE = 0.082 ± 0.016; male alive birds: $R^2 = 0.027$, P = 0.01, slope ± SE = 0.069 ± 0.028); (B) beak depth (female dead birds: $R^2 = 0.020$, $P = 5*10^{-3}$, slope ± SE = 0.025 ± $8.66*10^{-3}$; female alive birds: $R^2 = 8.52*10^{-4}$, P = 0.74, slope ± SE = 7.13*10-3 ± 0.021; male dead birds: $R^2 = 4.32*10^{-5}$, P = 0.91, slope ± SE = $9.93*10^{-4}$ ± $8.30*10^{-3}$; male alive birds: $R^2 = 8.86*10^{-3}$, P = 0.16, slope ± SE = $2.19*10^{-2}$ ± 0.016); (C) beak width (female dead birds: $R^2 = 0.020$, $P = 4*10^{-3}$, slope ± SE =*

*0.020 ± 6.99*10^{-3}; female alive birds: R^2 = 8.59*10^{-3}, P = 0.29, slope ± SE = 0.020 ± 0.019; male dead birds: R^2 = 4.41*10^{-3}, P = 0.23, slope ± SE = 8.04*10^{-3} ± 6.64*10^{-3}; male alive birds: R^2 = 0.010, P = 0.13, slope ± SE = 0.018 ± 0.012).*

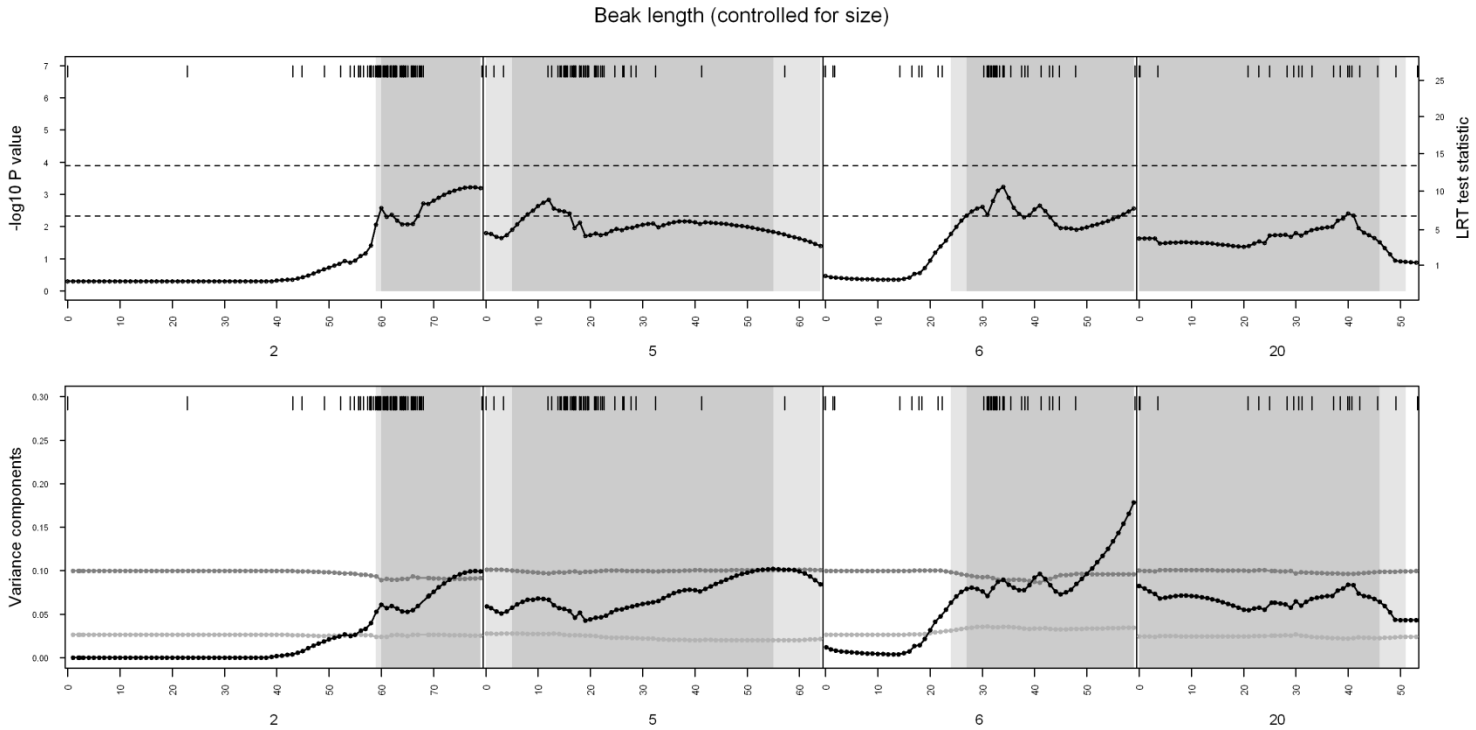


Figure S2: *Fine-mapping linkage scan for additive genetic variation in beak length controlled for body size scanned in 1 cM intervals. Included are those chromosomes that showed at least suggestive linkage in the initial 5 cM scan. The upper graph shows D and P values, the lower graph shows variance components explained by each interval, i.e. effect sizes. Marker positions are given by the vertical lines in the upper part of each graph. The upper and lower dotted lines represent significant and suggestive genome-wide linkage thresholds, respectively. Additive genetic variances are shown in black, maternal effects in dark grey and early rearing effects in light grey. Shaded dark grey regions: 1 LOD drop intervals; shaded light grey regions: 1.5 LOD drop intervals. The abscissa is labelled in cM and shows the chromosome identity.*

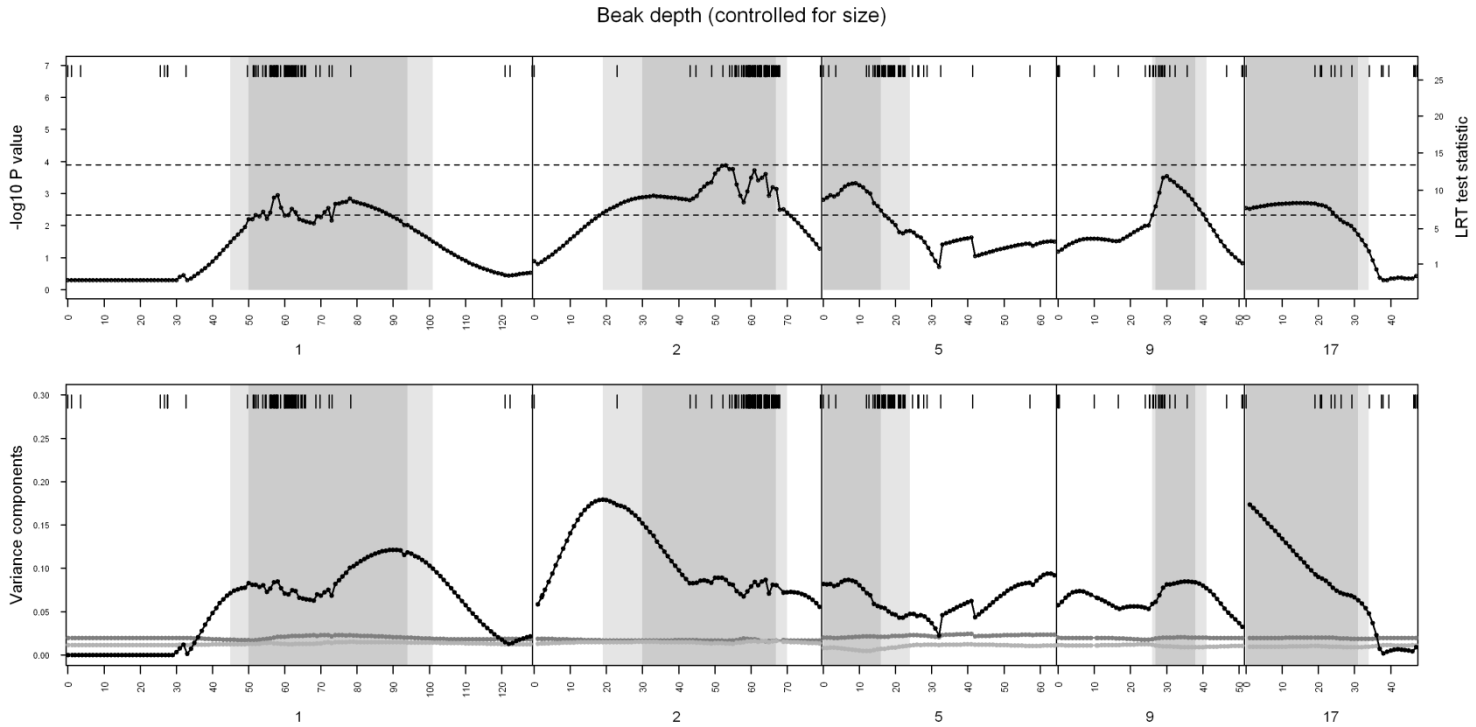


Figure S3: *Fine-mapping linkage scan for additive genetic variation in beak depth controlled for size scanned in 1 cM intervals. Included are those chromosomes that showed at least suggestive linkage in the initial 5 cM scan. The upper graph shows D and P values, the lower graph shows variance components explained by each interval, i.e. effect sizes. Marker positions are given by the vertical lines in the upper part of each graph. The upper and lower dotted lines represent significant and suggestive genome-wide linkage thresholds, respectively. Additive genetic variances are shown in black, maternal effects in dark grey and early rearing effects in light grey. Shaded dark grey regions: 1 LOD drop intervals; shaded light grey regions: 1.5 LOD drop intervals. The abscissa is labelled in cM and shows the chromosome identity.*

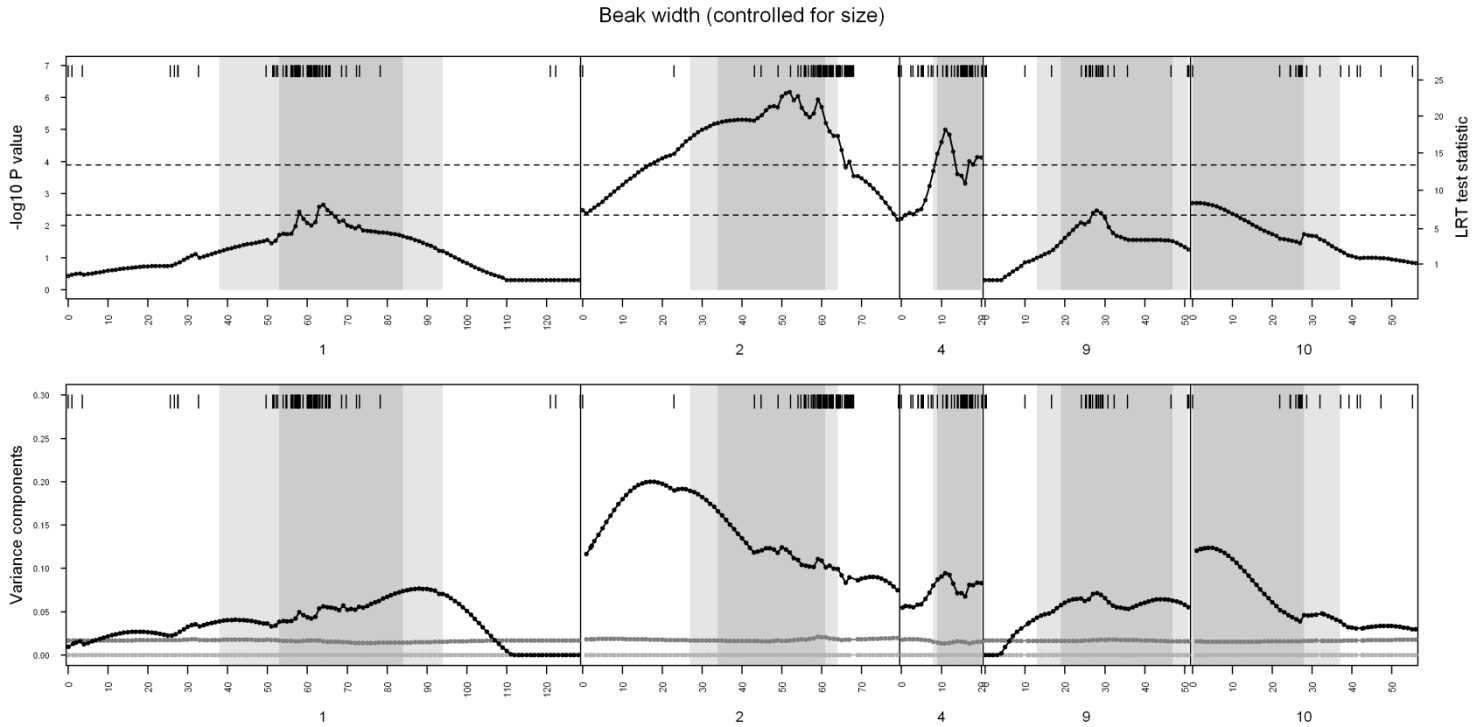


Figure S4: *Fine-mapping linkage scan for additive genetic variation in beak width controlled for size scanned in 1 cM intervals. Included are those chromosomes that showed at least suggestive linkage in the initial 5 cM scan. The upper graph shows D and P values, the lower graph shows variance components explained by each interval, i.e. effect sizes. Marker positions are given by the vertical lines in the upper part of each graph. The upper and lower dotted lines represent significant and suggestive genome-wide linkage thresholds, respectively. Additive genetic variances are shown in black, maternal effects in dark grey and early rearing effects in light grey. Shaded dark grey regions: 1 LOD drop intervals; shaded light grey regions: 1.5 LOD drop intervals. The abscissa is labelled in cM and shows the chromosome identity.*

Table S1: *Variance components (standardized by the phenotypic variance) and correlation estimates along with their standard errors for all three beak dimensions. Since the model on which these estimates are based had problems with convergence, standard errors should be treated with caution.*

Effect	Dimension	Length	Depth	Width
additive genetic	length	0.420 ± 0.069	0.469 ± 0.084	0.466 ± 0.096
	depth		0.648 ± 0.046	0.700 ± 0.045
	width			0.731 ± 0.047
maternal	length	0.057 ± 0.027	0.510 ± 0.654	-0.141 ± 0.400
	depth		0.004 ± 0.008	0.780 ± 0.432
	width			0.034 ± 0.017
early environment	length	0.038 ± 0.017	0.994 ± 0.052	0.576 ± 0.329
	depth		0.018 ± 0.012	0.662 ± 0.362
	width			0.028 ± 0.022
permanent environment	length	0.473 ± 0.067	0.179 ± 0.088	0.056 ± 0.107
	depth		0.290 ± 0.045	0.298 ± 0.093
	width			0.200 ± 0.046
residual	length	0.013 ± 0.001	0.157 ± 0.031	0.006 ± 0.048
	depth		0.041 ± 0.003	0.100 ± 0.042
	width			0.007 ± 0.001

Table S2: *Maternal and early rearing effects estimated in GridQTL by fitting mother identity and brood identity as random effects.*

Effect	Dimension	Effect ± SE	LRT value	P value
maternal	length	0.100 ± 0.033	13.34	$1.3*10^{-4}$
	depth	0.020 ± 0.020	1.16	0.14
	width	0.017 ± 0.018	1.04	0.15
early environment	length	0.026 ± 0.028	0.89	0.17
	depth	0.012 ± 0.022	0.3	0.29
	width	$1.1*10^{-11}$ ± $2.2*10^{-12}$	<0.01	>0.46

Table S3: *Linear models with beak dimensions as dependent variables (measured in mm), sex and status as factorial predictors (2 levels each), age (measured in years), inbreeding coefficient (f) and cube root of mass (measured in g), tarsus length and wing length (measured in mm) as covariates. Standardized effect sizes were calculated according to Nakagawa & Cuthill (2007). b = slope, d = standardized effect size, r = partial correlation coefficient.*

Dimension	Parameter	b ± SE	t value	P value	Effect size
beak length	intercept	5.423 ± 0.634			
	sex $^{female=0}$	0.200 ± 0.030	6.7	$3*10^{-11}$	d = 0.41
	status $^{alive=0}$	0.068 ± 0.032	2.1	0.034	d = 0.13
	f	-0.656 ± 0.304	-2.2	0.031	r = -0.07
	age	0.059 ± 0.010	6.0	$3*10^{-9}$	r = 0.18
	mass	1.265 ± 0.173	7.3	$5*10^{-13}$	r = 0.22
	tarsus	0.045 ± 0.029	1.6	0.12	r = 0.05
	wing	0.011 ± 0.010	1.1	0.27	r = 0.03
beak depth	intercept	5.206 ± 0.348			
	sex $^{female=0}$	0.195 ± 0.016	11.9	$1*10^{-30}$	d = 0.72
	status $^{alive=0}$	-0.101 ± 0.018	-5.7	$2*10^{-8}$	d = -0.35
	f	-0.173 ± 0.167	-1.0	0.30	r = -0.03
	age	0.005 ± 0.005	1.0	0.32	r = 0.03
	mass	0.510 ± 0.095	5.4	$9*10^{-8}$	r = 0.16
	tarsus	0.064 ± 0.016	4.0	$6*10^{-5}$	r = 0.12
	wing	0.012 ± 0.005	2.2	0.027	r = 0.07
beak width	intercept	2.885 ± 0.261			
	sex $^{female=0}$	0.095 ± 0.012	7.7	$3*10^{-14}$	d = 0.47
	status $^{alive=0}$	-0.083 ± 0.013	-6.2	$6*10^{-10}$	d = -0.38
	f	-0.273 ± 0.125	-2.2	0.030	r = -0.07
	age	0.005 ± 0.004	1.2	0.23	r = 0.04
	mass	0.635 ± 0.071	8.9	$2*10^{-18}$	r = 0.26
	tarsus	0.072 ± 0.012	6.1	$2*10^{-9}$	r = 0.18
	wing	0.014 ± 0.004	3.7	$3*10^{-4}$	r = 0.11

Chapter 2: Quantifying realized inbreeding in wild and captive animal populations

Abstract

Most molecular measures of inbreeding do not measure inbreeding at the scale that is most relevant for understanding inbreeding depression—namely the proportion of the genome that is identical-by-descent (IBD). The inbreeding coefficient F_{Ped} obtained from pedigrees is a valuable estimator of IBD, but pedigrees are not always available, and cannot capture inbreeding loops that reach back in time further than the pedigree. We here propose a molecular approach to quantify the realized proportion of the genome that is IBD (propIBD), and we apply this method to a wild and a captive population of zebra finches (Taeniopygia guttata). In each of 948 wild and 1,057 captive individuals we analyzed available single-nucleotide polymorphism (SNP) data (260 SNPs) spread over four different genomic regions in each population. This allowed us to determine whether any of these four regions was completely homozygous within an individual, which indicates IBD with high confidence. In the highly nomadic wild population, we did not find a single case of IBD, implying that inbreeding must be extremely rare (propIBD = 0–0.00094, 95% CI). In the captive population, a five-generation pedigree strongly underestimated the average amount of realized inbreeding (F_{Ped} = 0.013 < propIBD = 0.064), as expected given that pedigree founders were already related. We suggest that this SNP-based technique is generally useful for quantifying inbreeding at the individual or population level, and we show analytically that it can capture inbreeding loops that reach back up to a few hundred generations.

Published as: Knief U, G Hemmrich-Stanisak, M Wittig, A Franke, SC Griffith, B Kempenaers, W Forstmeier 2015: Quantifying realized inbreeding in wild and captive animal populations. Heredity 114: 397–403.

Heredity (2015) 114, 397–403

www.nature.com/hdy

ORIGINAL ARTICLE

Quantifying realized inbreeding in wild and captive animal populations

U Knief[1], G Hemmrich-Stanisak[2], M Wittig[2], A Franke[2], SC Griffith[3,4], B Kempenaers[1] and W Forstmeier[1]

Most molecular measures of inbreeding do not measure inbreeding at the scale that is most relevant for understanding inbreeding depression—namely the proportion of the genome that is identical-by-descent (IBD). The inbreeding coefficient F_{Ped} obtained from pedigrees is a valuable estimator of IBD, but pedigrees are not always available, and cannot capture inbreeding loops that reach back in time further than the pedigree. We here propose a molecular approach to quantify the realized proportion of the genome that is IBD (propIBD), and we apply this method to a wild and a captive population of zebra finches (*Taeniopygia guttata*). In each of 948 wild and 1057 captive individuals we analyzed available single-nucleotide polymorphism (SNP) data (260 SNPs) spread over four different genomic regions in each population. This allowed us to determine whether any of these four regions was completely homozygous within an individual, which indicates IBD with high confidence. In the highly nomadic wild population, we did not find a single case of IBD, implying that inbreeding must be extremely rare (propIBD = 0–0.00094, 95% CI). In the captive population, a five-generation pedigree strongly underestimated the average amount of realized inbreeding ($F_{Ped} = 0.013 < propIBD = 0.064$), as expected given that pedigree founders were already related. We suggest that this SNP-based technique is generally useful for quantifying inbreeding at the individual or population level, and we show analytically that it can capture inbreeding loops that reach back up to a few hundred generations.
Heredity (2015) **114,** 397–403; doi:10.1038/hdy.2014.116; published online 14 January 2015

INTRODUCTION

Inbreeding, the mating of genetically related individuals, leads to genomic regions inherited identical-by-descent (IBD) from both parents to their offspring. Thereby recessive deleterious mutations within those regions become homozygous and fully express their deleterious effects, known as inbreeding depression (Keller and Waller, 2002). Thus, inbred individuals are more often affected by recessive diseases (Campbell *et al.*, 2007), show poorer phenotypic traits and reduced fitness (reviewed by Chapman *et al.*, 2009).

The concept of identity-by-descent is important for understanding inbreeding depression, but it may also be confusing because, in any individual, most base pairs in its genome are homozygous due to common ancestry, often reaching back millions of generations. So, in a sense, most of the genome is 'IBD' (Powell *et al.*, 2010). However, such old coancestry is unproblematic, because it unlikely concerns a recessive deleterious mutation. Most recessive deleterious mutations persist in a population at low allele frequencies (because at higher frequencies they get selected against), and they do not persist for very long times (Li, 1975; Kiezun *et al.*, 2013), because rare alleles are frequently lost from a population even by genetic drift alone. Hence inbreeding depression does result from identity-by-descent, but only from the fraction of coancestry that is more recent than the respective deleterious mutation (Powell *et al.*, 2010). So how can we quantify this recent fraction of inbreeding in the complete absence of knowledge about deleterious mutations? Traditionally, this is done with pedigree information.

The pedigree-based inbreeding coefficient (F_{Ped}) has often been used to quantify the amount of inbreeding in an individual. F_{Ped} is defined as the probability that the two alleles at any autosomal locus are IBD, or equivalently, as an individual's average proportion of the autosomal genome being IBD (Wright, 1922). Here, IBD is defined with respect to a base population (the pedigree founders) in which all individuals are assumed to be unrelated (Powell *et al.*, 2010). F_{Ped} usefully estimates inbreeding as the proportion of the genome being IBD, but it only captures the most recent inbreeding loops that are included in the pedigree, not the ones from when the common ancestor lived before the pedigree founders. Because of this, F_{Ped} systematically underestimates the amount of IBD that leads to inbreeding depression (Broman and Weber, 1999), since it will miss the causal recessive deleterious mutations inherited twice from a common ancestor that lived before the start of the pedigree (but still after the average deleterious mutation arose).

Several estimates of inbreeding have been proposed that do not rely on pedigree data, but use genetic markers instead (reviewed by Coltman and Slate, 2003). In animal studies, microsatellites have usually been used as molecular markers to quantify individual heterozygosity, for example, as the percentage of loci within an individual that is heterozygous. Likewise, for populations, one can

[1]Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Seewiesen, Germany; [2]Institute of Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany; [3]Department of Biological Sciences, Macquarie University, Sydney, New South Wales, Australia and [4]School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales, Australia
Correspondence: U Knief, Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Eberhard-Gwinner-Strasse, Starnberg (Seewiesen) 82319, Germany.
E-mail: uknief@orn.mpg.de
Received 13 March 2014; revised 10 November 2014; accepted 14 November 2014; published online 14 January 2015

estimate the percentage of individuals that are heterozygous at any given locus or across a range of loci (multilocus heterozygosity). For instance, a heterozygosity of 80% means that 20% of the individuals are homozygous at a given locus. The disadvantage of this approach is that it remains unclear how many of these 20% are homozygous because of recent inbreeding and how many carry two different copies that are not closely related phylogenetically (that is, allozygous) but only happen to be the same by chance. In other words, a locus can be identical-by-state (IBS), but this does not necessarily mean it is IBD. Hence, IBS gives an approximate upper limit of the proportion of the genome being IBD (here 20%). To distinguish true IBD from IBS occurring due to chance, we need to inspect the information content of the flanking regions that surround a polymorphic site (Broman and Weber, 1999). If a homozygous marker is surrounded by other markers, all of which are homozygous too, then this strongly indicates IBD, because the combined probability of all markers being homozygous by chance becomes very small.

In dense marker panels, genetic regions that are IBD stand out as tracts in which all markers are homozygous, so-called 'runs of homozygosity' (ROH; McQuillan *et al.*, 2008) and these can be used to quantify realized inbreeding (Broman and Weber, 1999). Because microsatellites are usually not found at high density in a genome, ROH can be better detected with readily available single-nucleotide polymorphisms (SNPs), even though their allelic richness is lower and they are thus individually less informative (Broman and Weber, 1999). Studies on humans suggest that when more than 50 neighboring SNPs are homozygous, one can safely infer IBD because IBS by chance is too unlikely to ever occur (Powell *et al.*, 2010). Across many generations, recombination between the neighboring markers will lead to the breaking up of haplotypes, making runs of homozygosity shorter, the longer back the common ancestor was (McQuillan *et al.*, 2008). As we will show in this paper analytically, such runs of homozygosity allow us to detect IBD arising from a shared ancestor up to a few hundred generations back, much longer than any pedigree information. We propose that this yields a better measure of inbreeding than F_{Ped}, because it captures more of the relevant inbreeding events, while arguably still being on the safe side in that the majority of recessive deleterious mutations are older than those rather long inbreeding loops that are captured (Li, 1975; Fu *et al.*, 2013).

The aim of the present study is to demonstrate that the ROH-based method can be practically useful for quantifying the realized proportion of the genome that is IBD (propIBD) in both wild and captive animal populations. In the wild, when a study species is highly mobile, it is often impossible to compile a pedigree, so the amount of IBD can only be assessed molecularly. Also, in captivity, it is typically unknown how closely related the pedigree founders were, and by how much F_{Ped} underestimates the levels of IBD that are responsible for inbreeding depression (Ruiz-López *et al.*, 2009).

First, we demonstrate the utility of this molecular method, using two available SNP data sets that had been designed for other purposes.

(1) For a wild population of highly nomadic zebra finches, where no pedigree can be compiled, we use SNP data from an association study (ongoing research by the authors), where 18 candidate genes are being examined. Unfortunately, only four of these genes contained enough SNPs ($n = 56$–75) to confidently infer the presence or absence of IBD at a locus within an individual and to exclude that IBS occurred by chance alone. Thus, this data set allows us to take four snapshots for every genotyped individual to assess how frequently IBD occurs in this large and presumably panmictic population (Balakrishnan and Edwards, 2009), whose inbreeding levels are hitherto unknown (Zann, 1996).

(2) For a captive population of zebra finches with a five-generation pedigree, we have data on 1395 SNPs spread widely across the genome that were genotyped for the purpose of quantitative trait locus mapping (Backström *et al.*, 2010). From this data set, we selected four genomic regions with a matching set of 56–75 SNPs, that allow us to infer IBD with confidence. In this case, however, the SNPs are spread over much larger genetic distances than in the candidate gene data set (on average 4 cM vs 0.07 cM in the wild population), such that recombination will break up the haplotypes after a few generations. While this prevents us from detecting IBD *via* long inbreeding loops, the outcome is still sufficiently striking to show the utility of the method.

Because both exemplary data sets comprise only four loci (or genomic regions) per individual, we can only assess the population-wide level of inbreeding, which is relevant, for example, for studies in conservation genetics (Jamieson *et al.*, 2003). However, we note that our method is also suitable to study between-individual variation in inbreeding, provided a sufficient number of genomic regions per individual has been genotyped.

Second, we estimate analytically for how many generations haplotypes persist before being broken up by recombination or altered by mutation. In other words, we estimate the average length of inbreeding loops that can be detected with our method.

MATERIALS AND METHODS

Study populations and sample collection

We collected blood samples from 948 wild adult zebra finches (480 females, 468 males) at Fowler's Gap, NSW, Australia, in two contiguous places (S 30°57′ E 141° 46′ and S 31°04′ E 141°50′) between October and December 2010 and in April/May 2011. More details on the study sites and catching procedure using a walk-in trap at feeders are given in Griffith *et al* (2008) and Mariette and Griffith (2012).

For a comparison of inbreeding estimates based on the ROH approach and the pedigree, we used data from 1057 individuals from a captive population held at the Max Planck Institute for Ornithology in Seewiesen, Germany. Pedigree information available for this study covered five generations (Backström *et al.*, 2010; Schielzeth *et al.*, 2011, 2012). We calculated Wright's inbreeding coefficients using Pedigree Viewer v6.5 (Kinghorn and Kinghorn, 2010) and averaged them to get an average F_{Ped_5gen}. Founders (hatched in 2001) were known to be related to each other, showing an average F_{Ped_18gen} of 0.030 (calculated using Pedigree Viewer v5.1, for details see Forstmeier *et al.*, 2004), estimated from another 18-generation pedigree for the years 1985–2001 (Forstmeier *et al.*, 2004), which however is not owned by the authors. In a small and closed population, the increase in inbreeding coefficients during the first generations is almost linear (Falconer and Mackay, 1996) so the two average inbreeding coefficients F_{Ped_5gen} and F_{Ped_18gen} can be summed to obtain a rough estimate of the inbreeding coefficient from a 23-generation pedigree.

SNP genotyping and quality assessment

For the wild population, an initial dense SNP panel ($>$ 20 million SNPs) was discovered by sequencing a pooled non-barcoded sample of equal amounts of whole genomic DNA of 100 from the 948 individuals caught at Fowler's Gap with the Illumina HiSeq 2000 platform. The whole SNP discovery pipeline is described in the Supplementary Material.

In the course of an association study that will be described elsewhere, we genotyped all 948 wild-caught individuals at 685 SNPs located in 18 genes (10–75 SNPs per gene; Supplementary Table S2 and Supplementary Material) with an Illumina Infinium iSelect HD Custom BeadChip (Illumina Inc., Eindhoven, The Netherlands) on the Illumina iScan platform. Genotype quality was checked for each SNP (clustering of genotype calls, Hardy-Weinberg tests, the occurrence of heterozygous deletions (Ziegler *et al.*, 2010; Gogarten *et al.*, 2012)) and we assessed the possible impact of genotyping errors on our results (for details see the Supplementary Material).

In the captive population, all individuals had previously been genotyped for 1395 SNPs using the Illumina GoldenGate Assay (Fan *et al.*, 2003; Backström *et al.*, 2010). Since this SNP panel was originally designed to cover the whole

genome for quantitative trait locus mapping, the average physical marker spacing was much larger than in the wild birds (mean distance between neighboring SNPs ± s.d. = 701.5 ± 1117.5 kb vs 1.3 ± 4.6 kb, respectively). The genotype quality of these SNP calls had been checked previously and not a single inheritance error in our five-generation pedigree had been found (Backström *et al.*, 2010). Thus, here we assume no genotyping errors in these samples.

ROH-based estimation of population-level inbreeding

A large enough number of markers will, by chance alone, practically never be homozygous at the same time (Broman and Weber, 1999). Because our SNP data are not phased and haplotype frequencies cannot be established with confidence, we used a simplifying approach to get an estimate of the realized population-level of inbreeding (propIBD).

In the wild population, we first selected genes that were covered with enough SNPs so that IBS would not occur by chance alone. We identified four genes with 56–75 SNPs *via* a selection procedure described in the Supplementary Material. In brief, the procedure estimates the probability that all markers in a gene would be IBS by chance, and we then picked the four genes with the lowest probabilities ($P = 6 \times 10^{-6}$–7×10^{-5}, translating into $\mu = 0.01$–0.07 individuals expected to be IBS by chance alone). Thus, we expected fewer than one individual to be IBS by chance alone (see Supplementary Material for details). The alternative approach is simply to pick all genes covered with more than 55 SNPs, which yielded the same four genes. Please note that this selection was only needed because our genetic data stem from work that was not specifically designed for estimating population-level inbreeding. If a study is designed specifically for estimating inbreeding, then we recommend genotyping at least 75 SNPs per genomic region (which is the maximum number of SNPs per region used in this study, Supplementary Table S2). Yet the precise number of SNPs needed depends on linkage between SNPs, their allele frequencies and genotyping failure rates.

For each of the four selected regions, we calculated propIBD as the proportion of individuals that were IBS (and hence likely IBD). The population-wide propIBD could in principle be estimated from a single genomic region—provided that a large sample of individuals is used—, because each region should theoretically have the same probability of becoming IBD. Empirically, however, there is variation in IBD among regions in a genome (Weir *et al.*, 2006). Because of that several regions should be used to estimate propIBD to minimize the impact of this variation (which could, for example, stem from a region being located within an inversion polymorphism that is the target of disassortative mating; Thorneycroft, 1975). We calculated confidence limits for our estimated propIBD using Blaker's exact confidence interval (CI) (Blaker, 2000) for a binomial proportion with 0 successes (success = all homozygotes) and 3792 trials (4 regions × 948 individuals = 3792 trials) as implemented in Scherer (2013).

In the captive population, windows containing an equal number of genotyped SNPs as in the selected four candidate genes and spanning the smallest genetic length were selected on chromosomes *Tgu1* (67 SNPs, spanning 4.6 cM and 655 genes), *Tgu1A* (56 SNPs, spanning 4.0 cM and 476 genes), *Tgu2* (75 SNPs, spanning 3.8 cM and 653 genes) and *Tgu4* (62 SNPs, spanning 4.5 cM and 531 genes). Since these regions span around 4 cM, they will be broken up by cross-over events in roughly 4% of the meioses (1 cM is defined as an expected number of 0.01 cross-over events per meiosis). Hence, the true extent of IBD will be somewhat underestimated because cross-over leads to the loss of IBS for the entire genomic region even when a part of it is still IBD.

To obtain a 95% CI for our estimate of propIBD in captivity, we fitted a generalized linear model with the glm function in R (v2.15.3; R Core Team, 2013). We used counts of completely homozygous individuals vs not completely homozygous individuals for each of the four regions (bound with the cbind function in R) as the response variable, and the intercept as the sole predictor. We specified a quasibinomial error distribution and a logit-link function, because the data were overdispersed. Since we used the logit-link function, we back-transformed (inverse-logit) the intercept and the 95% CI (estimated using the confint function in R), to obtain the estimate of propIBD and its 95% CI (which is equivalent to the average proportion of regions being IBS weighted by their sample sizes; Crawley, 2007).

Persistence of ROH

To get an idea about the persistence of IBD segments over the course of many generations, we estimated the recombination rate per region in both populations from Backström *et al* (2010) and used the average of all four selected regions for our calculations (0.068 cM per 41.6 kb and 4.23 cM per 59.56 Mb for wild and captive populations, respectively). Because the zebra finch exhibits very low recombination rates in the center of its macrochromosomes (0.12 cM/Mb; Backström *et al.*, 2010), which is not representative of most other species, we also estimated the persistence of IBD segments for a more typical example, namely the chicken (*Gallus gallus*). We considered a hypothetical locus that is 65 kb long, such that 65 SNPs can easily be found in that region considering the observed diversity in the chicken genome (Wong *et al.*, 2004). This locus then spans 0.20 cM, given the genome-wide average recombination rate of 3.11 cM/Mb (Groenen *et al.*, 2009). The probability of persistence is then given as $(1 - L/100)^{2 \times G}$ where G is the number of generations and L is the length of a region in cM (see also Hayes *et al.*, 2003). We further assumed a mutation rate of 1.2×10^{-8} per nucleotide per generation (based on studies on humans; Kong *et al.*, 2012). The probability of persistence was then calculated as $(1 - 1.2 \times 10^{-8})^{nSNP*2*G}$ where G is the number of generations and nSNP is the average number of SNPs genotyped within one region. The calculations show that the mutation rate had almost no effect on the overall persistence of IBD segments: in the absence of recombination, half of the IBD segments would persist for around 440 000 generations. Hence, uncertainty about the mutation rate in the species of interest is likely to be relatively uncritical.

RESULTS

Estimates of probIBD in our study populations

Figure 1 depicts the average heterozygosity calculated across 56–75 SNPs (depending on the region) for $n = 948$ wild and $n = 1057$ captive zebra finches. These averages are approximately normally distributed, and the left tail of the bell-shaped curve is sufficiently far from zero, indicating that IBS is not expected to occur by chance alone. Hence, individuals that are completely homozygous for a gene region strongly indicate IBD. The proportion of completely homozygous individuals (propIBD) is highlighted for each gene region in Figure 1.

Across the four genes from the wild population, we did not observe a single case that would indicate IBD, suggesting a complete absence of inbreeding (Supplementary Table S2, Figure 1, left panels). It is unlikely that genotyping errors were the cause of the absence of IBD regions, for two reasons. First, all four genes had at least two SNPs that were heterozygous in each individual. Second, those individuals with the least number of heterozygous SNPs per gene had ratios of allelic intensities for the heterozygous SNPs that were in the range of the heterozygous SNPs of the whole population (Supplementary Figure S2). On the basis of these four genes, the estimated population level of propIBD was practically 0 (upper 95% confidence value = 0.00094).

For the captive population, the average pedigree-based inbreeding coefficient was $F_{Ped\_5gen} = 0.013$. However, pedigree founders were already related by an average F_{Ped_18gen} of 0.030 (Forstmeier *et al.*, 2004). Thus, the birds from our captive population had an F_{Ped_23gen} of approximately 0.030+0.013 = 0.043. On the basis of the SNPs from four selected genomic regions, the estimated realized propIBD for this population was 0.064 (95% CI = 0.036–0.102) (Supplementary Table S3, Figure 1 right panels).

Persistence of ROH

Our calculations show that recombination events plus *de novo* mutations occur at such a low frequency that it should take 508 generations to break up 50% of the haplotypes that we assessed for IBD in the wild zebra finch population (markers spread over an average genetic length of 0.068 cM). For the captive zebra finch

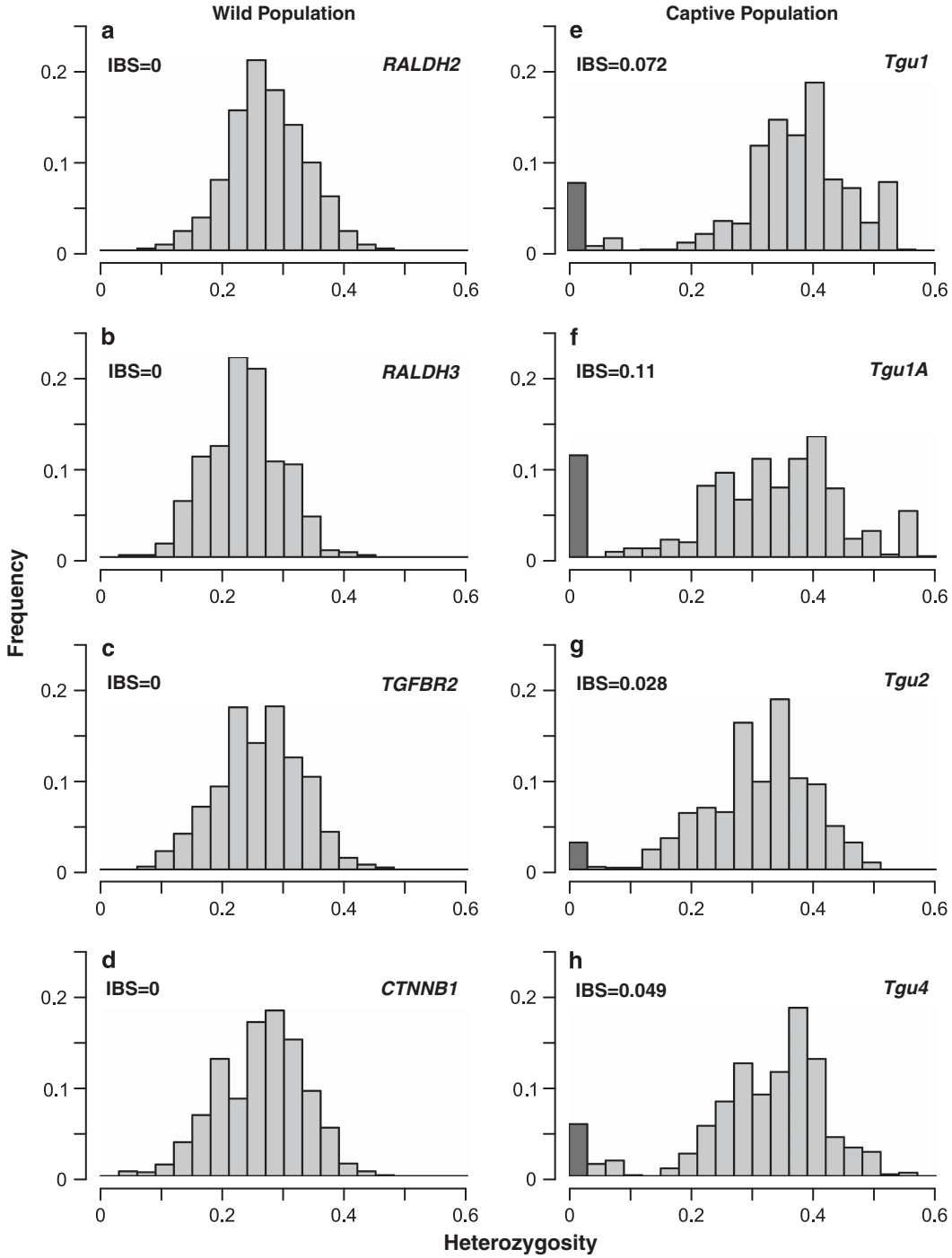


Figure 1 Each histogram shows the mean heterozygosity of every individual in a population for a particular gene region (four regions per population). The wild population ($n=948$ individuals, **a**–**d**) contains no individual that is fully homozygous in any of the four gene regions (identical-by-state, IBS = 0). In contrast, between 2.8 and 11.0% of the 1057 individuals of the captive population (**e**–**h**) are fully homozygous in a given gene region indicating homozygosity by descent and hence inbreeding (dark bars). Gene regions are not the same between the two populations, but were matched for the number of SNPs genotyped per region. $n=67$ SNPs for (**a**) *RALDH2* and (**e**) *Tgu1*, $n=56$ SNPs for (**b**) *RALDH3* and (**f**) *Tgu1A*, $n=75$ SNPs for (**c**) *TGFBR2* and (**g**) *Tgu2*, $n=62$ SNPs for (**d**) *CTNNB1* and (**h**) *Tgu4*.

population, however, where our markers were spread over much larger genetic distances (about 4.23 cM), we estimate that 50% of the studied genomic regions would persist for only eight generations. We also estimated the persistence of a hypothetical region in the chicken genome, in which the recombination rate is considerably higher than in the zebra finch (Backström *et al.*, 2010). In such a—potentially more broadly applicable—hypothetical avian genome, 50% of the 65 kb haplotypes spanning 0.20 cM should persist for 171 generations.

DISCUSSION

We here propose a novel method to estimate inbreeding using ROH of molecular markers without the need for pedigree information, thus

avoiding problems stemming from incomplete pedigree information and relatedness of pedigree founders. The method should not be confused with existing ROH methods that rely on sliding-window approaches to find stretches of markers that are IBS and use miscellaneous methods to discern IBS from IBD (Howrigan *et al.*, 2011). Because these previous methods try to identify every IBD segment within a single genome, they are influenced by variation in linkage disequilibrium and minor allele frequencies of SNPs across windows and tend to overestimate inbreeding when markers are not linkage disequilibrium-pruned before analysis (Polašek *et al.*, 2010). However, this is not the case for our method, because we focus on selected regions that are densely covered with SNPs and hence practically never become homozygous by chance alone (in all four examined regions in the wild zebra finch population together we expect less than 0.2 cases of IBS by chance alone among the $4 \times 948 = 3792$ cases and in the captive population less than 0.7 cases of IBS among the 3958 cases; Supplementary Tables S2 and S3). Thus, theoretically each such region in a genome becomes representative for the population-mean inbreeding and could be used interchangeably to estimate propIBD, if we assume no selection against homozygotes in a region or other special cases like inversion polymorphisms or targets of mate choice. To mitigate errors in population comparisons, normally, the same regions should be used to estimate propIBD in the different populations, which, however, was not possible in the present study because we utilized available SNP data rather than designing the genotyping for our purpose. Yet, reassuringly, all four comparisons (Figure 1) lead to the same conclusion. In the present study, the physical and genetic lengths of the regions were very different between the populations (42 kb and 0.07 cM in the wild vs 60 Mb and 4.2 cM in captivity). As a consequence of using SNPs that were widely spread over long distances, we could only capture rather short inbreeding loops in the captive population, because recombination will have broken up some of the regions studied, leading to an underestimate of propIBD. In that sense, estimates of propIBD are not quite comparable between our two data sets, but the conclusion that inbreeding is much more frequent in the captive than in the wild population only confirms the obvious, and the difference in estimated inbreeding coefficients should be highly conservative.

In the wild population, only four out of 18 genes analyzed could be used to reliably distinguish IBD from cases of IBS occurring by chance alone. These were the four genes with the most SNPs genotyped, emphasizing the need for a dense marker set to reliably infer IBD. Other factors that influence the reliable discrimination between IBD and IBS are population diversity, allele frequencies and linkage between SNPs (Gibson *et al.*, 2006). Australian mainland zebra finches exhibit exceptionally high levels of nucleotide diversity, rapid decay of linkage disequilibrium and high population recombination rates (Balakrishnan and Edwards, 2009), making 56–75 SNPs sufficiently powerful to distinguish IBD from IBS. Although human population demography has been quite different from that of the zebra finch Howrigan *et al* (2011) suggested that similar marker densities were sufficient for IBD detection in humans.

In our wild study population, not a single individual was completely homozygous in any of the four selected genes, indicating that inbreeding in wild zebra finches is an extremely rare event. With such a low rate of inbreeding, recessive deleterious mutations are not effectively purged and are expected to accumulate in this large panmictic population (Bataillon and Kirkpatrick, 2000). The severe inbreeding depression that has been observed in captive populations (Bolund *et al.*, 2010; Forstmeier *et al.*, 2012; Hemmings *et al.*, 2012) is in line with such an accumulation of recessive deleterious mutations.

In the captive population, the estimated realized propIBD was 0.064. This value still underestimates the true realized inbreeding because cross-over will have broken up some of the tracts of homozygosity (50% decay after 8 generations). Figure 1 (right panels, especially e and h) shows a few odd cases with heterozygosity < 0.1 but larger than zero (*Tgu1*: $n = 21$, *Tgu1A*: $n = 13$, *Tgu2*: $n = 7$, *Tgu4*: $n = 40$). These might represent IBD segments where just a few of the SNPs had recombined. Consistent with this interpretation, the heterozygous SNPs in those specific cases were not distributed randomly across the examined regions but were concentrated at one of the ends of the regions (data not shown).

In the captive population, there was more variation between regions in the percentage of individuals being IBD than expected by chance (we had to specify a quasi-binomial error distribution in our GLM). Specifically, fewer individuals than expected were IBD for chromosome *Tgu2*. A lack of homozygous individuals for chromosome *Tgu2* had been shown previously in our population (Forstmeier *et al.*, 2007). This could result from non-random mating or be indicative of positive or negative selection. In any case, this emphasizes the need for using multiple regions to estimate population-level inbreeding to ensure against variation in IBD among regions due to evolutionary forces (for example, selection) or structural variants (for example, inversions). In particular, it also illustrates that comparisons of inbreeding levels between populations should normally be based on the same regions in a genome (Weir *et al.*, 2006).

PropIBD was substantially higher than F_{Ped} from a 5-generation pedigree ($F_{Ped\_5gen} = 0.013$). Even when accounting for relatedness from another 18-generation pedigree, $F_{Ped\_23gen} \approx 0.030+0.013 = 0.043$ was still lower than the estimated propIBD. This might be surprising because we estimated that 50% of the studied genomic regions would persist for only eight generations. However, both the pedigree and the propIBD estimate may be biased. On the one hand, some individuals have been introduced into the 23-generation pedigree in a later generation, which then are treated as founders, making the pedigree actually shorter and consequently biasing F_{Ped} downwards (that is, 23 generations are the maximum, but not the average length of the pedigree). Furthermore, the founders of the 23-generation pedigree (maintained in a laboratory since 1985 and originating from the population of domesticated birds maintained by aviculturists in the United Kingdom for about one hundred years before that) must have been related to each other to some extent. Indeed, this seems inevitable when founding a captive population from other captive populations. Consequently, F_{Ped} again underestimates the true levels of inbreeding. On the other hand, recombination within a studied region does not always break up the homozygous stretch of SNPs; if the cross-over happens at one end of the studied region the allelic state of the few affected SNPs might not change. Because of such cases, runs of homozygosity may persist for longer than the estimated eight generations. Finally, it should be noted that F_{Ped_23gen} and propIBD were not significantly different (95% CI overlap).

Our calculations on the persistence of haplotypes assessed for IBD confirmed that our method is able to detect inbreeding loops that reach back in time much further than typical pedigree information obtained from wild animal populations. Even in organisms with high recombination rates like the chicken it should be possible to detect inbreeding loops over more than 100 generations with a sufficiently dense marker panel. From calculations of the mean age of a recessive deleterious allele in a population of constant size (Li, 1975), it is reasonable to assume that in species with a sufficiently large effective

population size the majority of recessive deleterious mutations is much older than 100 generations. We mention this because if most such mutations had arisen only recently, this would undermine the utility of quantifying long inbreeding loops. Instead, this suggests that such long inbreeding loops that reach far back into the past are of importance to study the full extent of inbreeding depression. Thus, our method may be a useful tool in conservation genetics to assess the amount of population-level inbreeding in wild animal populations, even when pedigree information is available. We here show the utility of our method for a large, outbred wild population as well as for a captive population with moderate levels of inbreeding, which could serve as an example for a bottlenecked population under conservation efforts. For future empirical or modelling studies, it would be interesting to assess the utility of our method for populations with high levels of inbreeding, in which the background heterozygosity might not be normally distributed anymore.

Our study suggests that inbreeding can be reliably quantified in a population using ROH based on high-density SNP genotyping without the need for pedigree data. It should be noted that individual variation in inbreeding could also be measured with our method, for example, by genotyping 80 regions each covered with approximately 75 SNPs (that is, a 6k SNP array, yet the exact number of regions and SNPs might depend on the species, research question and marker characteristics). Among others, the false-positive rate of our method decreases with the number of SNPs assessed for a ROH, whereas the false-negative rate increases with the genetic distance covered by the SNPs. Consequently, each region should span only a short genetic distance to detect all relevant stretches of IBD that may cause inbreeding depression.

Only F_{Ped} and ROH-based methods measure inbreeding at the scale that is most relevant for understanding inbreeding depression—namely the proportion of the genome that is IBD. Even if pedigree data are available, the proposed method can identify cases of inbreeding that reach back many more generations than are typically covered by pedigree information. This may be of particular interest because recessive deleterious mutations persist in a population over many more generations than covered by the available pedigrees. High-density SNP genotype data from a large number of individuals are necessary, but these are increasingly becoming available in wild animal populations, for example, through candidate-gene based association studies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Christa Beckmann, Aliza Sager and Mylene Mariette for assistance with sampling the birds. The wild birds were sampled and banded under approval of the Macquarie University Animal Ethics Committee, the Australian Bird and Bat Banding Scheme, and a Scientific License from NSW National Parks and Wildlife Service. We further thank Melanie Schneider for laboratory work in Seewiesen and Markus Schilhabel, the Next-Generation Sequencing team and the Genotyping team at the IKMB in Kiel for laboratory work. UK is part of the International Max Planck Research School for Organismal Biology. This study was funded by the Max Planck Society (BK), with the zebra finch study at Fowler's Gap funded by support to SCG from the Australian Research Council. Genotype data of wild zebra finches are accessible through the Dryad Digital Repository: http://doi.org/10.5061/dryad.j678b.

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

Balakrishnan CN, Edwards SV (2009). Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* **181**: 645–660.

Bataillon T, Kirkpatrick M (2000). Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. *Genet Res* **75**: 75–81.

Blaker H (2000). Confidence curves and improved exact confidence intervals for discrete distributions. *Can J Stat* **28**: 783–798.

Bolund E, Martin K, Kempenaers B, Forstmeier W (2010). Inbreeding depression of sexually selected traits and attractiveness in the zebra finch. *Anim Behav* **79**: 947–955.

Broman KW, Weber JL (1999). Long homozygous chromosomal segments in reference families from the centre d'étude du polymorphisme humain. *Am J Hum Genet* **65**: 1493–1500.

Campbell H, Carothers AD, Rudan I, Hayward C, Biloglav Z, Barac L *et al.* (2007). Effects of genome-wide heterozygosity on a range of biomedically relevant human quantitative traits. *Hum Mol Genet* **16**: 233–241.

Chapman JR, Nakagawa S, Coltman DW, Slate J, Sheldon BC (2009). A quantitative review of heterozygosity-fitness correlations in animal populations. *Mol Ecol* **18**: 2746–2765.

Coltman DW, Slate J (2003). Microsatellite measures of inbreeding: a meta-analysis. *Evolution* **57**: 971–983.

Crawley MJ (2007). *The R Book*. John Wiley & Sons Ltd: Chichester, England.

Falconer DS, Mackay TFC (1996). *Introduction to Quantitative Genetics*. Longman: Harlow, Essex, UK.

Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F, Gunderson KL *et al.* (2003). Highly parallel SNP genotyping. *Cold Spring Harb Symp Quant Biol* **68**: 69–78.

Forstmeier W, Coltman DW, Birkhead TR (2004). Maternal effects influence the sexual behavior of sons and daughters in the zebra finch. *Evolution* **58**: 2574–2583.

Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B (2012). Heterozygosity-fitness correlations in zebra finches: microsatellite markers can be better than their reputation. *Mol Ecol* **21**: 3237–3249.

Forstmeier W, Schielzeth H, Schneider M, Kempenaers B (2007). Development of polymorphic microsatellite markers for the zebra finch (*Taeniopygia guttata*). *Mol Ecol Notes* **7**: 1026–1028.

Fu WQ, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM *et al.* (2013). Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* **493**: 216–220.

Gibson J, Morton NE, Collins A (2006). Extended tracts of homozygosity in outbred human populations. *Hum Mol Genet* **15**: 789–795.

Gogarten SM, Bhangale T, Conomos MP, Laurie CA, McHugh CP, Painter I *et al.* (2012). GWASTools: an R/Bioconductor package for quality control and analysis of genome-wide association studies. *Bioinformatics* **28**: 3329–3331.

Griffith SC, Pryke SR, Mariette M (2008). Use of nest-boxes by the Zebra Finch (*Taeniopygia guttata*): implications for reproductive success and research. *EMU* **108**: 311–319.

Groenen MAM, Wahlberg P, Foglio M, Cheng HH, Megens HJ, Crooijmans RPMA *et al.* (2009). A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Res* **19**: 510–519.

Hayes BJ, Visscher PM, McPartlan HC, Goddard ME (2003). Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Res* **13**: 635–643.

Hemmings NL, Slate J, Birkhead TR (2012). Inbreeding causes early death in a passerine bird. *Nat Commun* **3**: 863.

Howrigan DP, Simonson MA, Keller MC (2011). Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC Genomics* **12**: 460.

Jamieson IG, Roy MS, Lettink M (2003). Sex-specific consequences of recent inbreeding in an ancestrally inbred population of New Zealand Takahe. *Conserv Biol* **17**: 708–716.

Keller LF, Waller DM (2002). Inbreeding effects in wild populations. *Trends Ecol Evol* **17**: 230–241.

Kiezun A, Pulit SL, Francioli LC, van Dijk F, Swertz M, Boomsma DI *et al.* (2013). Deleterious alleles in the human genome are on average younger than neutral alleles of the same frequency. *PLoS Genet* **9**: e1003301.

Kinghorn BP, Kinghorn AJ (2010). *Pedigree Viewer 6.5*. University of New England: Armidale, Australia.

Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G *et al.* (2012). Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**: 471–475.

Li WH (1975). The first arrival time and mean age of a deleterious mutant gene in a finite population. *Am J Hum Genet* **27**: 274–286.

Mariette MM, Griffith SC (2012). Conspecific attraction and nest site selection in a nomadic species, the zebra finch. *Oikos* **121**: 823–834.

McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, Barac-Lauc L *et al.* (2008). Runs of homozygosity in European populations. *Am J Hum Genet* **83**: 359–372.

Polašek O, Hayward C, Bellenguez C, Vitart V, Kolčic I, McQuillan R *et al.* (2010). Comparative assessment of methods for estimating individual genome-wide homozygosity-by-descent from human genomic data. *BMC Genomics* **11**: 139.

Powell JE, Visscher PM, Goddard ME (2010). Reconciling the analysis of IBD and IBS in complex trait studies. *Nat Rev Genet* **11**: 800–805.

R Core Team. (2013). *R Foundation for Statistical Computing*. Vienna, Austria.
Ruiz-López MJ, Roldán ERS, Espeso G, Gomendio M (2009). Pedigrees and microsatellites among endangered ungulates: what do they tell us? *Mol Ecol* **18**: 1352–1364.
Scherer R (2013). PropCIs: Various confidence interval methods for proportions. *R package version 0.2-4*. http://CRAN.R-project.org/package=PropCIs.
Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2012). QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. *Evolution* **66**: 18–30.
Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2011). Data from: QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. Dryad Data Repository.
Thorneycroft HB (1975). A cytogenetic study of the white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Evolution* **29**: 611–621.
Weir BS, Anderson AD, Hepler AB (2006). Genetic relatedness analysis: modern data and new challenges. *Nat Rev Genet* **7**: 771–780.
Wong GKS, Liu B, Wang J, Zhang Y, Yang X, Zhang ZJ *et al*. (2004). A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* **432**: 717–722.
Wright S (1922). Coefficients of inbreeding and relationship. *Am Nat* **56**: 330–338.
Zann RA (1996). *The Zebra Finch: A Synthesis of Field and Laboratory Studies* Vol 5, Oxford University Press: Oxford, UK.
Ziegler A, König IR, Pahlke F (2010). *A Statistical Approach to Genetic Epidemiology: Concepts and Applications*. Wiley-VCH: Weinheim, Germany.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)

Supplement

Methods

SNP discovery pipeline

SNPs were discovered by paired-end sequencing a pooled non-barcoded sample of equal amounts of whole genomic DNA of 100 from the 948 individuals caught at Fowler's Gap with the Illumina HiSeq 2000 platform (1 flow cell = 8 lanes). We mapped the reads to the zebra finch reference genome (version July 2008, WUSTL v3.2.4 assembly; Warren *et al*, 2010) with BWA (v0.5.9; Li and Durbin, 2009); settings: bwa aln -n 4 -q 20 -l 5000; meaning that:

- 4 mismatches were allowed per read (-n 4),
- reads were cut when quality dropped below QS=20 (-q 20),
- and that seeding was disabled which is slower but performs better (-l 5000)).

This yielded an average genome coverage of 247.5 x (calculated using BEDTools (v2.17.0; Quinlan and Hall, 2010) coverageBed function after removing alignment gaps). Coverage was evenly distributed across regions. SNPs were called using GATKs (v2.1-11-g13c0244; McKenna *et al*, 2010) UnifiedGenotyper function (settings: -stand_call_conf 50.0 -stand_emit_conf 10.0 -dcov 1000 -mbq 10 -mmq 10 -glm BOTH; meaning that:

- a minimum phred-scaled confidence threshold of 50 was used at which variants were called (-stand_call_conf 50.0),
- a minimum phred-scaled confidence threshold of 10 was used at which variants were emitted (-stand_emit_conf 10.0),
- a downsampling threshold of 1000 reads was used (-dcov 1000),
- a minimum base quality of 10 was required to consider a base for calling (-mbq 10),
- a minimum read mapping quality of 10 was required to consider a read for calling (-mmq 10),
- and that both indels and SNPs were called (-glm BOTH)).

SNPs were filtered with the VariantFiltration function (settings: -filterExpression "(AF > 1.00)" -filterName "filter_AF" -filterExpression "(DP < 10.0 || DP > 600.0)" -filterName "filter_DP" -filterExpression "(HRun > 10.0)" -filterName "filter_HRun" -filterExpression "(MQ0 > 50.0 || ((MQ0/(1.0*DP)) > 0.50))" -filterName "filter_MQ0" -filterExpression "(QD < 0.04)" -filterName "filter_QDlow" -filterExpression "(QD > 39.0)" -filterName "filter_QDhigh" -filterExpression "(MQ < 10.0)" -filterName "filter_MQ" -filterExpression "(Dels > 40.0)" -filterName "filter_Dels" -clusterWindowSize 0 -mask InDels -maskName "InDel" -maskExtension 0; meaning that SNPs were removed if:

- all individuals were fixed for the non-reference allele (-filterExpression "(AF > 1.00)" -filterName "filter_AF"),
- the depth of coverage at the given position was <10 or >600 (-filterExpression "(DP < 10.0 || DP > 600.0)" -filterName "filter_DP"),
- the longest continuous homopolymer run of the variant allele was longer than 10 (-filterExpression "(HRun > 10.0)" -filterName "filter_HRun"),
- the number of reads with a mapping quality of 0 at a locus was >50 (-filterExpression "(MQ0 > 50.0 || ((MQ0/(1.0*DP)) > 0.50))" -filterName "filter_MQ0"),

- the SNP quality score divided by the unfiltered depth of all non-reference samples was <0.04 or >39 (-filterExpression "(QD < 0.04)" -filterName "filter_QDlow" -filterExpression "(QD > 39.0)" -filterName "filter_QDhigh")
- the root mean square mapping quality of all reads was <10 (-filterExpression "(MQ < 10.0)" -filterName "filter_MQ"),
- the percentage of reads with deletions spanning the position was >40 (-filterExpression "(Dels > 40.0)" -filterName "filter_Dels"),
- indels called in the previous SNP calling were spanning the position with no extension (-mask InDels -maskName "InDel" -maskExtension 0). The clustered SNP filter was disabled (-clusterWindowSize 0)).

The exact description of the resulting SNP set is beyond the scope of this study. The minor allele counts from the pooled sequencing of the 648 SNPs studied here were highly correlated with their minor allele frequency estimated from the Illumina iSelect genotyping of the 948 individuals (Pearson's r = 0.95).

Description of candidate genes

The 18 genes studied in the wild population are part of a candidate-gene based association study for morphological traits which will be described elsewhere. Details on each gene are provided in Table S1.

Quality control of SNP genotyping

Genotype quality of SNP calls was only checked in the SNP panel used for the wild-caught zebra finches (n = 685 SNPs). In the captive birds, where we used an existing SNP data set, genotype quality had been checked previously and not a single inheritance error in our five-generation pedigree had been found Backström *et al.* 2010.

For each of the 685 SNPs used in the wild birds, cluster plots, which separate homozygous from heterozygous individuals, were automatically analyzed using Illuminas GenomeStudio software (v2011.1, genotyping module 1.9.4) and manually inspected using R v2.15.3; R Core Team 2013. The R package GWASTools v1.6.2; Gogarten *et al.* 2012 was used for further quality checks like the occurrence of heterozygous deletions and to test for Hardy-Weinberg-disequilibrium (HWD). HWD is generally taken as a sign for genotyping errors but could also indicate selection, potentially acting against homozygotes which could bias our results Ziegler *et al.* 2010. 37 of the 685 SNPs were in significant HWD after Bonferroni correction, distributed randomly among genes (chi-square test for homogeneity $\chi^2 = 24.10$, df = 17, P = 0.12). For all but one SNP heterozygotes were missing, which is indicative of genotyping errors Ziegler *et al.* 2010 and suggests that there was no selection acting against homozygotes. All 37 SNPs were removed from further analyses. From the remaining 648 SNPs 99.65% were successfully called, totaling 612184 genotypes. Due to ascertainment bias in the SNP-detection pipeline, SNPs with a minor allele frequency (MAF) below 0.1 were rarer than expected (Figure S1). This should increase the power to detect true ROH (lower

false-negative rate) Broman & Weber 1999 but increase the false-positive rate because the probability of having a hidden heterozygous SNP of lower minor allele frequency within the ROH increases Powell *et al.* 2010.

Genotyping errors could either increase (heterozygotes called as homozygotes) or decrease (homozygotes called as heterozygotes) the number of ROH. Since we found not a single ROH in the wild population (see Results) we were only concerned with the latter case. Illumina Infinium iSelect HD Custom BeadChips use a fluorescence signal to distinguish the two alleles (alleles A and B) at each locus Peiffer *et al.* 2006. In heterozygous individuals the signal of both alleles is emitted in approximately equal intensities. We checked for genotyping errors only in those individuals with the least number of heterozygous SNPs per gene since these individuals had the highest chance of being mistakenly assigned as non-homozygous. For each of these individuals and at each heterozygous SNP we calculated the ratio of allelic intensities by dividing the intensity of allele A by the sum of the intensities of allele A and B and compared it to the same ratio of heterozygous individuals from the whole population (shown in Figure S2). Whenever the allelic ratio of the individual and the SNP under consideration lay between the 2.5 and 97.5 percentile of the distribution of the allelic intensity of the whole population we considered that SNP correctly called as heterozygous.

Selection of informative regions for IBD detection

In the wild population, we first had to select genes that were covered with enough SNPs so that IBS would not occur by chance alone. This was only needed because our genetic data was not specifically designed for estimating population-level inbreeding and some genes did not contain sufficient information for inferring IBD with confidence. Once a threshold for including a gene into our study was found (a rather post hoc procedure), we used the same threshold also in the captive population.

Specifically, for each individual we calculated its mean heterozygosity in each of the 18 genes listed in Table S2 as its number of heterozygous SNPs in that region divided by the total number of SNPs in that region. For each gene this gave us as many estimates of individual-mean heterozygosities as there are individuals (Figure 1). An individual-mean heterozygosity of 0 indicates that in this specific individual this region is IBS and hence potentially IBD. For each region, we then calculated the population-mean heterozygosity ($\bar{x}$) and its (between-individual) standard deviation (SD) as the mean and standard deviation of the individual-mean heterozygosities. These two properties characterize the bell-shaped curves shown in Figure 1, and our aim was to estimate how far from zero (in terms of multiples of SD) the mean heterozygosity is. We did this by calculating Z-scores for each region at a mean heterozygosity of zero as $Z = (0 - \bar{x}) / SD$ Zar 2010, which yields a Z-score for being IBS at all SNPs in that region by chance alone. Assuming that the population-mean heterozygosity within each region follows a normal distribution which is a valid assumption because of the central limit theorem; see Nei & Roychoud 1974 and Figure 1, we

transformed our Z-scores to P-values for each region using the pnorm function in R (v2.15.3). The expected number of homozygous individuals (μ) in the absence of inbreeding was then estimated as P*nInd, with nInd being the number of individuals studied. P and μ should be seen only as a rough guideline for identifying the most informative regions, given that the true haplotype frequencies were not known and for some of the regions the assumption that mean heterozygosity was normally distributed was violated. Only for those regions where we expected less than one individual to be homozygous by chance alone (μ < 0.5), probIBD was calculated. The other regions were not considered for our estimate of probIBD since IBD could not be reliably distinguished from IBS. Only four out of the 18 genes in the wild population were covered with enough markers to confidently identify cases of IBD (μ < 0.5 individuals per gene).

For comparison, we calculated the population-mean heterozygosity and its between-individual standard deviation also for each selected region in the captive population after removing all completely homozygous individuals. This should yield the baseline heterozygosity in that region in the population in the absence of inbreeding. Z-scores, P-vales and the expected number of homozygous individuals (μ) in the absence of inbreeding were calculated as described above (Table S3).

References

Backström N, Forstmeier W, Schielzeth H*, et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research* **20**, 485–495.

Broman KW, Weber JL (1999) Long homozygous chromosomal segments in reference families from the centre d'étude du polymorphisme humain. *American Journal of Human Genetics* **65**, 1493–1500.

Gogarten SM, Bhangale T, Conomos MP*, et al.* (2012) GWASTools: an R/Bioconductor package for quality control and analysis of genome-wide association studies. *Bioinformatics* **28**, 3329–3331.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760.

McKenna A, Hanna M, Banks E*, et al.* (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297–1303.

Nei M, Roychoud AK (1974) Sampling variances of heterozygosity and genetic distance. *Genetics* **76**, 379–390.

Peiffer DA, Le JM, Steemers FJ*, et al.* (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Research* **16**, 1136–1148.

Powell JE, Visscher PM, Goddard ME (2010) Reconciling the analysis of IBD and IBS in complex trait studies. *Nature Reviews Genetics* **11**, 800–805.

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842.

R Core Team (2013) R: a language and environment for statistical computing. *3.0.2*.

Warren WC, Clayton DF, Ellegren H*, et al.* (2010) The genome of a songbird. *Nature* **464**, 757–762.

Zar JH (2010) *Biostatistical analysis*, 5 edn. Prentice-Hall/Pearson, New Jersey.

Ziegler A, König IR, Pahlke F (2010) *A Statistical Approach to Genetic Epidemiology: Concepts and Applications*, 2nd edn. Wiley-VCH, Weinheim, Germany.

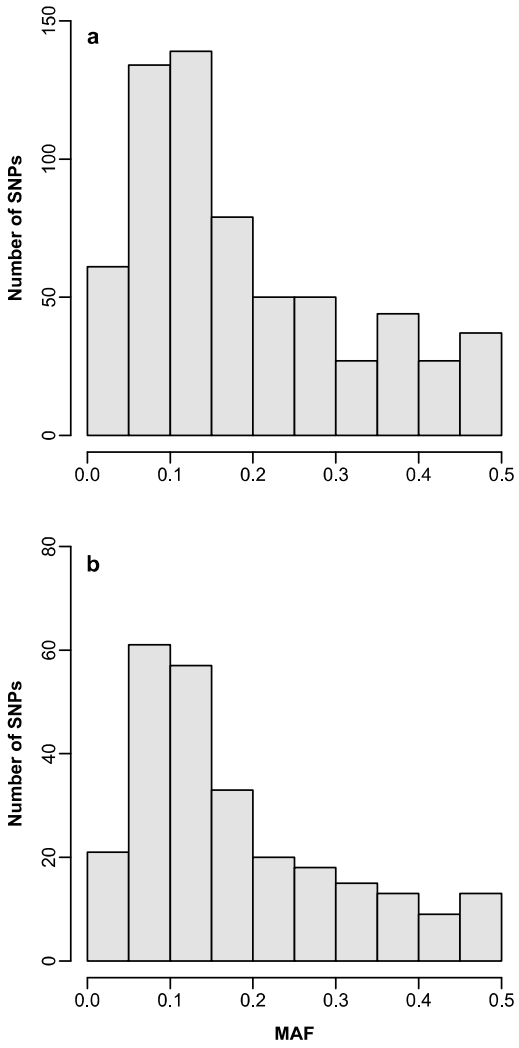


Figure S1: *Minor allele frequency spectra of (a) the 648 SNPs located in all 18 genes and (b) the 260 SNPs located in the 4 genes RALDH2, RALDH3, TGFBR2 and CTNNB1.*

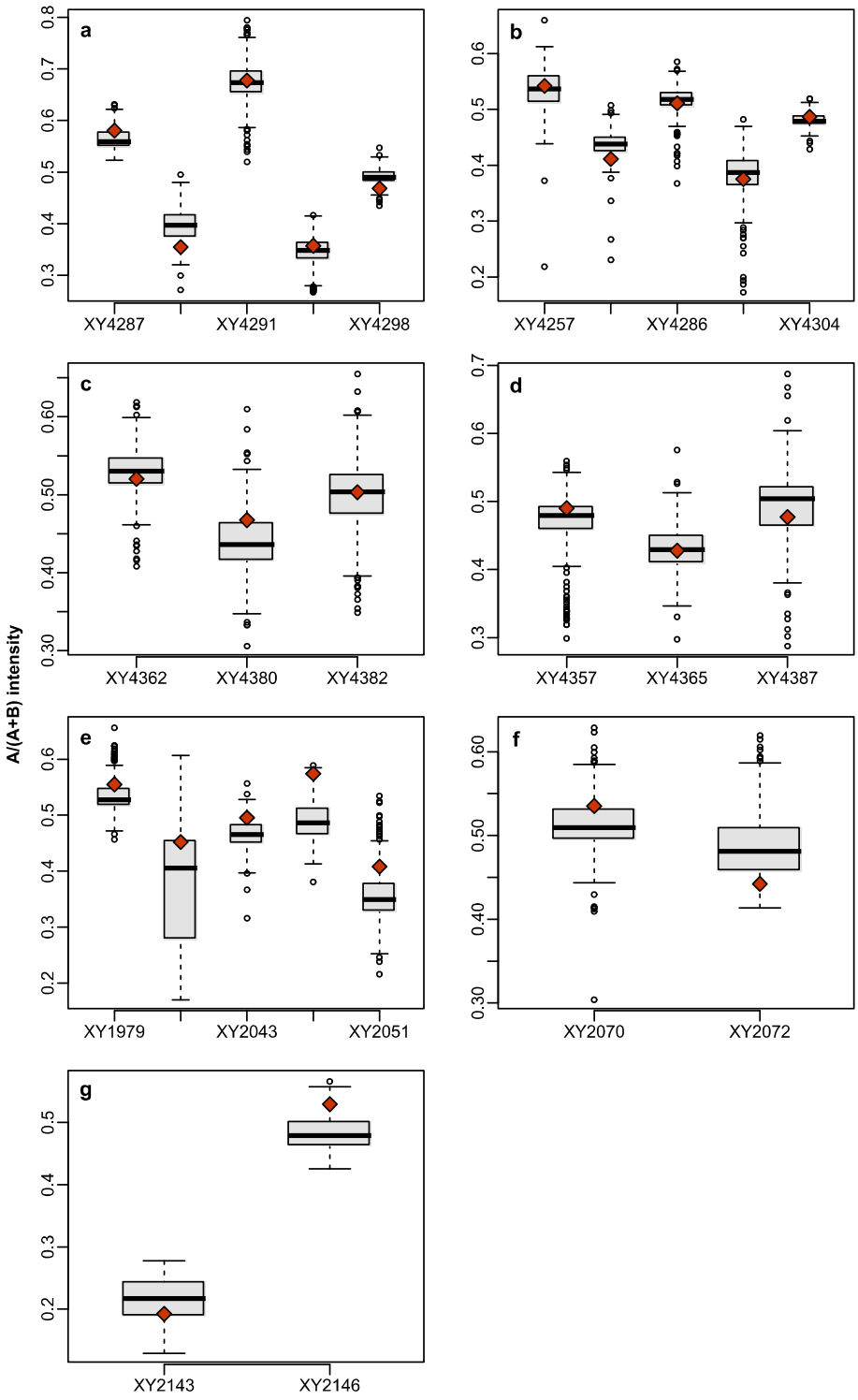


Figure S2: *For those individuals with the least number of heterozygous SNPs in the top four genes the ratios of allelic intensities for their heterozygous SNPs are shown. Each plot represents one individual. Boxplots are ratio estimates from each heterozygous individual in the population, diamonds represent the call ratios for those critical cases that could be genotyping errors. Yet, 24 out of the 25 cases lie within the 95% percentile of call ratios of the other individuals and the one outside the 95% percentile is located in (e), where all other 4 are located inside the 95% CI. (a, b) 2 individuals heterozygous at 5 SNPs in RALDH2, (c, d)*

2 individuals heterozygous at 3 SNPs in RALDH3, (e) 1 individual heterozygous at 5 SNPs in TGFBR2, (f, g) 2 individuals heterozygous at 2 SNPs in CTNNB1. Internal SNP names are given on the abscissa. Boxes represent the interquartile range (IQR), the thick line the median and whiskers extend 1.5 times the IQR. Values further than 1.5 times the IQR are represented by dots.

Table S1: *Overview of the 18 genes used in this study. For each gene the morphological trait which it may influence in the zebra finch and the gene ontology (GO) annotation is given.*

Gene	Trait	Molecular function (GO)[1]
RALDH2	Beak morphology	retinal dehydrogenase activity, 3-chloroallyl aldehyde dehydrogenase activity, oxidoreductase activity, oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor, retinal binding
RALDH3	Beak morphology	aldehyde dehydrogenase (NAD) activity, aldehyde dehydrogenase [NAD(P)+] activity, oxidoreductase activity, oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor, protein homodimerization activity, thyroid hormone binding, NAD+ binding
TGFBR2	Beak morphology	protein kinase activity, protein serine/threonine kinase activity, transmembrane receptor protein serine/threonine kinase activity, receptor signaling protein serine/threonine kinase activity, protein tyrosine kinase activity, receptor activity, transforming growth factor beta-activated receptor activity, transforming growth factor beta receptor activity, type II, contributes to protein binding, ATP binding, glycosaminoglycan binding, transferase activity, transferring phosphorus-containing groups, activin-activated receptor activity, mitogen-activated protein kinase kinase kinase binding, type I transforming growth factor beta receptor binding, type III transforming growth factor beta receptor binding, SMAD binding, metal ion binding, transforming growth factor beta binding
CTNNB1	Beak morphology	RNA polymerase II activating transcription factor binding, DNA binding, chromatin binding, double-stranded DNA binding, sequence-specific DNA binding transcription factor activity, transcription coactivator activity, signal transducer activity, structural molecule activity, binding, protein binding, protein C-terminus binding, transcription factor binding, enzyme binding, kinase binding, protein kinase binding, protein phosphatase binding, estrogen receptor binding, protein complex binding, ionotropic glutamate receptor binding, nuclear hormone receptor binding, transcription regulatory region DNA binding, ion channel binding, alpha-catenin binding, cadherin binding, SMAD binding, androgen receptor binding, I-SMAD binding, R-SMAD binding, repressing transcription factor binding
BMP7	Beak morphology	cytokine activity, contributes to protein binding, growth factor activity
LIN28B	Digit ratio	nucleic acid binding, DNA binding, RNA binding, protein binding, zinc ion binding

WNT5A	Wing length	sequence-specific DNA binding transcription factor activity, receptor binding, frizzled binding, frizzled-2 binding, receptor tyrosine kinase-like orphan receptor binding, cytokine activity, protein binding, protein domain specific binding, transcription regulatory region DNA binding, receptor agonist activity
RB1	Tarsus length	core promoter binding, RNA polymerase II activating transcription factor binding, molecular_function, DNA binding, sequence-specific DNA binding transcription factor activity, transcription coactivator activity, protein binding, transcription factor binding, enzyme binding, kinase binding, ubiquitin protein ligase binding, identical protein binding, androgen receptor binding, phosphoprotein binding
FOXO1A	Tarsus length, Body size	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in negative regulation of transcription, DNA binding, chromatin binding, double-stranded DNA binding, sequence-specific DNA binding transcription factor activity, RNA polymerase II distal enhancer sequence-specific DNA binding transcription factor activity, protein binding, transcription factor binding, DNA binding, bending, protein kinase binding, sequence-specific DNA binding, protein phosphatase 2A binding
INTS6	Tarsus length, Body size	transmembrane signaling receptor activity, protein binding
BMP2	Beak morphology	retinol dehydrogenase activity, receptor binding, cytokine activity, protein binding, growth factor activity, phosphatase activator activity, SMAD binding, protein heterodimerization activity, BMP receptor binding
WNT6	Wing length	receptor binding, frizzled binding
DKK3	Beak morphology	
CALM1	Beak morphology	calcium ion binding, protein binding, kinase activity, protein domain specific binding, titin binding, thioesterase binding, N-terminal myristoylation domain binding, phospholipase binding, phosphatidylinositol 3-kinase binding, ion channel binding, calcium-dependent protein binding, protein phosphatase activator activity
RNASEH2B	Tarsus length, Body size	
SHH	Beak morphology	glycoprotein binding, signal transducer activity, patched binding, calcium ion binding, protein binding, glycosaminoglycan binding, peptidase activity, zinc ion binding, morphogen activity, laminin-1 binding

ESR1	Digit ratio	core promoter sequence-specific DNA binding, DNA binding, chromatin binding, sequence-specific DNA binding transcription factor activity, steroid hormone receptor activity, ligand-activated sequence-specific DNA binding RNA polymerase II transcription factor activity, steroid binding, protein binding, beta-catenin binding, transcription factor binding, zinc ion binding, enzyme binding, nitric-oxide synthase regulator activity, estrogen receptor activity, estrogen response element binding, estrogen-activated sequence-specific DNA binding RNA polymerase II transcription factor activity, identical protein binding, sequence-specific DNA binding
KPNA3	Tarsus length, Body size	binding, protein binding, protein C-terminus binding, nuclear localization sequence binding, protein transporter activity

[1] From www.genecards.org, accessed 05/16/2014

Table S2: *Summary statistics on SNPs and heterozygosity in the 18 genes used for this study on wild zebra finches. In our analysis we focus on the first four genes, for which the expected number of homozygotes in the population in the absence of inbreeding (μ) is less than 1 out of the n = 948 individuals studied. Note that for less informative loci (below the line) μ is typically smaller than the observed numbers of IBS individuals, since the former is just a rough approximation. This is especially prominent in genes covered with only a few SNPs (WNT6, CALM1, SHH).*

Gene	Chromosome	Length (kb)	N SNPs	Call rate	Mean ($\bar{x}$) heterozygosity	SD heterozygosity	Z-score	P	μ	N individuals 100% IBS*	Min N heterozygous SNPs
RALDH2	*Tgu10*	62.3	67	99.63	0.28	0.06	4.39	$5.6*10^{-6}$	0.01	0	5
RALDH3	*Tgu10*	42.5	56	99.71	0.24	0.06	3.96	$3.8*10^{-5}$	0.04	0	3
TGFBR2	*Tgu2*	33.7	75	99.49	0.26	0.07	3.83	$6.4*10^{-5}$	0.06	0	5
CTNNB1	*Tgu2*	27.9	62	99.64	0.26	0.07	3.79	$7.4*10^{-5}$	0.07	0	2
BMP7	*Tgu20*	39.6	35	99.39	0.20	0.06	3.18	0.00074	0.70	0	1
LIN28B	*Tgu3*	89.9	55	99.57	0.29	0.10	2.75	0.0030	2.84	6	0
WNT5A	*Tgu12*	30.0	19	99.35	0.23	0.09	2.52	0.0059	5.64	9	0
RB1	*Tgu1*	61.9	36	99.65	0.33	0.13	2.46	0.0070	6.64	16	0
FOXO1A	*Tgu1*	39.8	25	99.87	0.24	0.12	2.10	0.018	16.92	15	0
INTS6	*Tgu1*	38.3	31	99.77	0.28	0.14	2.08	0.019	17.78	37	0
BMP2	*Tgu3*	24.9	15	99.85	0.27	0.13	2.07	0.019	18.07	26	0
WNT6	*Tgu7*	13.8	11	98.88	0.26	0.13	2.01	0.022	20.95	31	0
DKK3	*Tgu5_random*	22.0	42	99.83	0.33	0.17	1.98	0.024	22.59	2	0
CALM1	*Tgu5*	7.0	13	99.61	0.25	0.14	1.80	0.036	33.71	70	0
RNASEH2B	*Tgu1*	37.5	41	99.84	0.29	0.18	1.61	0.053	50.48	55	0
SHH	*Tgu2*	15.7	10	99.60	0.17	0.12	1.45	0.074	70.15	149	0
ESR1	*Tgu3*	185.2	36	99.96	0.23	0.20	1.15	0.12	118.01	110	0
KPNA3	*Tgu1*	18.6	19	99.88	0.23	0.24	0.92	0.18	169.13	166	0

* Identical-By-State.

Table S3: *Summary statistics on SNPs and heterozygosity in the 4 genomic regions used for this study on captive zebra finches. We removed all individuals that were completely homozygous at a particular region (yielding the values in N inds) for calculating the mean heterozygosity and its standard deviation to get the baseline heterozygosity in the absence of inbreeding. In all four regions the expected number of homozygotes in the population in the absence of inbreeding (μ) was less than 1.*

Chromosome	Length (kb)	Genes covered	N SNPs	Call rate	Mean heterozygosity	SD heterozygosity	Z-score	P	μ	N inds	N individuals 100% IBS*
Tgu1	73513.3	655	67	98.01	0.38	0.09	4.29	$9.0*10^{-6}$	0.009	981	76
Tgu2	76077.2	653	75	98.73	0.31	0.08	3.82	$6.7*10^{-5}$	0.069	1027	30
Tgu4	55193.2	531	62	97.43	0.33	0.09	3.55	$2.0*10^{-4}$	0.196	1005	52
Tgu1A	33472.4	476	56	98.74	0.34	0.10	3.36	$3.8*10^{-4}$	0.363	945	112

* Identical-By-State.

Chapter 3: Comparing precision and accuracy of inbreeding estimates derived from pedigrees versus molecular markers

Abstract

For studies of inbreeding it is important to know the proportion of an individual's genome that is identical-by-descent (IBD). This can be calculated from pedigrees (inbreeding coefficient F) or estimated from molecular markers, whereby both estimators contain an error component. Pedigree F is inaccurate at the individual level due to chance events during chromosome segregation (Mendelian noise). Molecular estimates suffer from sampling error of markers plus the error that occurs when a marker is homozygous (identical-by-state, IBS) without reflecting common ancestry (IBD-IBS discrepancy). Here, we quantify these three sources of error via simulations for a captive population of zebra finches, and compare them with estimates from studies on humans. In the zebra finch, where the genome contains large blocks that are rarely broken up by recombination, the Mendelian noise was remarkably large (nearly twofold larger standard deviations compared to humans). On the other hand, the IBD-IBS correspondence for microsatellite markers was relatively high, explaining why a limited number of informative markers can yield useful estimates of genome-wide IBD. Our simulations show how many markers are needed to reach the same precision as pedigree-based predictions, and they illustrate how the various noise components introduce systematic and random errors into inbreeding-fitness relationships.

Submitted to Molecular Ecology Resources as: Knief U, B Kempenaers, W Forstmeier: Comparing precision and accuracy of inbreeding estimates derived from pedigrees versus molecular markers.

Introduction

Offspring of genetically related individuals are inbred, meaning that they harbor genomic segments homozygous due to common ancestry (identical-by-decent, IBD; Wright 1922). The proportion of an individual's genome that is IBD (genome-wide IBD, GWIBD) is arguably the best predictor of the homozygous mutation load of an individual because all recessive deleterious mutations that become IBD contribute to inbreeding depression (Keller *et al.* 2011).

Studies on inbreeding often aim at quantifying the amount of variance in fitness explained by inbreeding or the inbreeding load within a population. These two aims are not the same and we should attempt to reach both of them (Szulkin *et al.* 2010). For the first, we need to find a way to quantify all relevant inbreeding in order to minimize bias and reduce the random sampling error in our GWIBD estimator. For the second, an accurate and precise estimate of the regression slope of a fitness trait over a measure of inbreeding within a population is required. Both objectives can be addressed by using information from pedigrees or from molecular markers, whereby each has its own limitations in terms of precision and accuracy. Throughout this manuscript we use the term 'precision' to refer to random errors around a true value, and the term 'accuracy' to refer to systematic errors (bias) away from the true value (Sokal & Rohlf 1995).

In pedigrees, the expected GWIBD (Pedigree F) of a diploid individual whose parents are *k*th generation linear descendants of a common ancestor is traditionally quantified using Wright's path method as the inbreeding coefficient $F = 2^{-(2*k+2)}$ and can be extended to also incorporate complex inbreeding loops (Malécot 1948; Wright 1922). As a consequence of limited recombination in meiosis, however, genomes do not get transmitted as independent basepairs but rather in segments of DNA, leading to linkage between adjacent segments and variation around the expected GWIBD (Fisher 1949). In the following, we will call this random error the Mendelian sampling noise (Figure 1). Interestingly, the magnitude of this error does not only change with the degree of inbreeding, but also with the genomic architecture of the species under consideration (Franklin 1977; Hill & Weir 2011; Rasmuson 1993; Stam 1980), as we will explain below.

The variation in GWIBD for a given inbreeding constellation will be smaller, the more segments in a genome segregate independently (law of large numbers; Visscher 2009). Consequently, since chromosomes get inherited as independent units in meiosis the Mendelian noise for a given inbreeding constellation will be smaller in species with more chromosomes (Hill & Weir 2011). Likewise, genomic segments on a given chromosome will be broken up by cross-overs during meiosis, which reduces linkage and consequently increases the number of independently segregating units in a genome. Thus, the longer the genetic map of a genome, which by definition is the expected number of cross-overs in each meiosis, the smaller the Mendelian sampling noise (Hill & Weir 2011). To our knowledge, all

analytical analyses so far have assumed a uniform distribution of cross-overs along chromosomes (e.g. Franklin 1977; Hill & Weir 2011; Stam 1980; but see Suarez *et al.* (1979) and Libiger & Schork (2007) for Monte Carlo simulations on relatedness). Although this assumption holds more or less for the human genome (Matise *et al.* 2007), linkage maps from other species have shown that recombination along chromosomes can be highly skewed (Backström *et al.* 2010; Gore *et al.* 2009). This will result in a more block-like inheritance pattern of genomic segments on a given chromosome, which in turn will increase the Mendelian noise (Forstmeier *et al.* 2012; Guo 1995; Risch & Lange 1979).

Another aspect of the use of pedigrees is that pedigree F reflects IBD only within the pedigree (descent from the founders). In other words, only when all founders are unrelated and outbred, pedigree F accurately reflects all inbreeding. Throughout this manuscript we will focus on such an idealized scenario, where all inbreeding is fully defined and captured by the pedigree information. Thus, we further define IBD with reference to the founder population. To emphasize this point we indicate the number of generations of the pedigree as a subscript (e.g. F_7, IBD_7, $GWIBD_7$ refer to values from a pedigree that is seven generations long). In real populations, pedigree founders will also be related to some extent, and inbreeding loops that are longer than the pedigree will also contribute to inbreeding depression. Such long inbreeding loops can be captured by molecular information (but not by pedigree information), which generally means that pedigrees underestimate the full amount of inbreeding. Further discussion of this problem lies outside the scope of this study and is treated elsewhere (e.g. Knief *et al.* 2015; Powell *et al.* 2010; Speed & Balding 2015; Thompson 2013).

As an alternative to pedigrees one can use molecular markers such as microsatellites or single nucleotide polymorphisms (SNPs) to estimate GWIBD (Broman & Weber 1999). If we could sequence the whole genome of all individuals we are interested in, we could estimate GWIBD more or less exactly (Prado-Martinez *et al.* 2013). Usually, however, only parts of a genome are covered by molecular markers and IBD within these parts (Marker IBD) is used as a proxy for GWIBD (Powell *et al.* 2010). This will introduce variation into our estimates of GWIBD, which we call the marker sampling noise (Figure 1). Unlike the Mendelian sampling noise which is an inherent result of the meiotic process, the marker sampling noise is only a consequence of the limited number of markers used; hence, the precision of GWIBD estimates increases when more molecular markers are sampled.

Another problem arises when using molecular markers. Marker homozygosity only equals Marker IBD when each monophyletic haplotype in a population is identifiable by a unique allele of the molecular marker in use (referred to as “ideal marker”). Studies on heterozygosity-fitness correlations combine homozygosity at multiple unlinked microsatellites into a measure of GWIBD, e.g. as the percentage of markers being homozygous. Yet microsatellite homozygosity does not translate directly into IBD for two

reasons. (1) Two homologous DNA segments may have been inherited from a shared ancestor (IBD), but the microsatellite marker located in this segment may have changed due to mutation since the common ancestor. Then the marker will not be identical-by-state (IBS) even though the segment is IBD. However, if we define IBD with regard to common ancestors that lived rather recently, such cases will be exceedingly rare, because microsatellite mutation rates are too low (Goldstein & Schlötterer 1999). Hence, for simplicity, in the following we ignore this first possibility assuming no mutations. (2) Because a microsatellite marker can adopt only a finite number of states (alleles), two phylogenetically unrelated haplotypes (not IBD) may carry the same allele (IBS) by chance alone. We will call the combination of Marker IBD with this chance homozygosity the Marker identity-by-state (Marker IBS).

Depending on how well Marker IBS reflects Marker IBD this may introduce severe error into the GWIBD estimate, which we call the IBD-IBS discrepancy (Thompson 1976; Figure 1). Generally, this discrepancy becomes smaller by increasing the allelic richness of a molecular marker, because more haplotypes can be distinguished uniquely. Clearly, multiallelic microsatellite markers contain more information on GWIBD than biallelic SNPs. Because genomes get transmitted in segments rather than by individual basepairs, we can consider several closely linked markers jointly as a single marker in order to further decrease the IBD-IBS discrepancy (Kong *et al.* 2008). By doing this, we generate more diverse and consequently more informative markers for GWIBD estimation (practically ideal markers), because the combined probability of all markers being homozygous just by chance becomes very small (see Knief *et al.* (2015) for an example of constructing such ideal markers from dense SNP panels).

Although neither pedigrees nor molecular markers estimate GWIBD precisely, a multitude of studies have interpreted a weak correlation between Pedigree F and Marker IBS as a sign of weakness of molecular markers in predicting GWIBD, culminating in the demand for more and better pedigrees (see Forstmeier *et al.* 2012 for a critical discussion on the subject). Here, we argue that the optimal method depends on the goal of the study. If we are interested in estimating the inbreeding load, Pedigree F may be indeed – as we will show – superior to molecular markers. However, if the aim is to quantify individual GWIBD, molecular markers may be the better choice because both the marker sampling noise and the IBD-IBS discrepancy can be reduced by using a larger number of highly informative markers and no assumptions about the relatedness of the founders have to be made.

Here we illustrate the properties of Pedigree F and molecular markers and address the following questions. (1) How does the amount of random noise introduced into estimates of GWIBD by Mendelian segregation compare to the amount of marker sampling noise and IBD-IBS discrepancy stemming from a limited number of molecular markers? (2) How many molecular markers are needed to become a superior method to pedigree information in

approximating GWIBD (Forstmeier *et al.* 2012)? (3) How do molecular markers compare to pedigree-based estimates of inbreeding in heterozygosity-fitness correlations/regressions, when we are interested in precise and accurate estimates of the inbreeding load (Szulkin *et al.* 2010)?

Because GWIBD can be assessed precisely only by sequencing the whole genome of an individual and because for our purpose we also need to look at a large population of individuals, we use a Monte Carlo gene-dropping simulation to answer these questions for two genomes that we expect to show high (zebra finch, *Taeniopygia guttata*) and low (human) amounts of Mendelian noise. Our expectation is based on the fact that although the zebra finch genome consist of 39 autosomes compared to only 22 in humans, about half of its autosomal genome is made up of only four chromosomes (*Tgu1*, *Tgu1A*, *Tgu2* and *Tgu3*) and shows extremely low recombination rates in the centers of these chromosomes (Backström *et al.* 2010) whereas in humans recombination is distributed quite uniformly on the megabase scale along chromosomes (Matise *et al.* 2007). Here, we follow the zebra finch genomes of 159 diploid pedigree founders, who we assume to be unrelated, through an empirical seven-generations pedigree of 3,404 individuals. We estimated the Mendelian sampling noise in Pedigree F_7, ignoring inbreeding stemming from relatedness between pedigree founders. To quantify the amount of marker sampling noise we estimated $GWIBD_7$ from subsets of the simulated genomes (Marker IBD_7). We used empirical data from 11 microsatellites genotyped in the seventh generation of our pedigree to estimate the IBD-IBS discrepancy and introduced this noise component into our simulations (Marker IBS_7). This allowed us to compare the precision of Pedigree F_7, Marker IBD_7 and Marker IBS_7 in predicting $GWIBD_7$. Finally, we demonstrate the effects of the three noise components (Figure 1) on estimates of heterozygosity-fitness correlations/regressions.

Methods

Empirical linkage and physical maps

For the zebra finch, we used the sex-averaged linkage map described in Backström *et al.* (2010) covering 33 chromosomes and based on 1,395 SNPs (data accessible from Schielzeth *et al.* 2011, 2012). We excluded the sex chromosome and removed 10 microchromosomes covered by less than 10 markers to include only those 22 autosomes containing enough information to fit a smoothed line (see below). However, because the variance in relatedness and inbreeding depends on the number of chromosomes and because the zebra finch genome consists of 39 autosomes (Pigozzi & Solari 1998), we added 17 “artificial” chromosomes to our linkage map.

These 17 chromosomes were constructed using both the physical and the genetic length of known chromosomes to be as close to reality as possible. The physical lengths of known chromosomes were taken from the WUSTL v3.2.4 assembly (Warren *et al.* 2010). First, we

used the complete published physical genome length (1,222,864,721 bp), subtracted the length of the sex chromosome (72,861,351 bp) and rounded the result up to the nearest 100 kb (11,501 x 100 kb). Then we subtracted the length of the 22 well-characterized autosomes from the linkage map (9,146 x 100 kb), leaving 2,355 x 100 kb.

Calderón & Pigozzi (2006) estimated the total genetic length of the zebra finch genome (containing all 39 autosomes) to be 2,272.5 cM (from MLH1 foci mapping). We subtracted the estimated total genetic length of the 22 well-characterized autosomes (1,391.711 cM, based on our smoothed line fitting function; see below) from the total genetic length of the genome, leaving 880.789 cM. In the zebra finch genome, recombination along the microchromosomes follows a more or less uniform distribution (Backström *et al.* 2010; Calderón & Pigozzi 2006). Thus, we constructed the 17 additional chromosomes as being 2,355 x 100 kb / 17 = 138.5294 x 100 kb and 880.789 cM / 17 = 51.81114 cM long with a uniform distribution of recombination along the chromosomes (Figure S1). Assuming a uniform distribution of recombination essentially reduces the amount of Mendelian noise stemming from these chromosomes to a minimum (Risch & Lange 1979).

For humans, we used the sex-averaged Rutgers Map v.2 (Matise *et al.* 2007) covering 24,168 markers and all 22 autosomes and the X chromosome (Figure S2). We removed the sex chromosome from all further analyses. The physical lengths of chromosomes were taken from the hg19/GRCh37 assembly (Collins *et al.* 2004).

Linkage map smoothing

In order to get a monotonously increasing linkage map for each chromosome we followed Roesti *et al.* (2013) and fitted a local polynomial regression (loess function in R v3.0.2 (R Core Team 2013) with span parameter = 10 / number of markers on the chromosome, degree = 1, control = loess.control(surface = "direct")) with the genetic positions of the markers as the independent variable and the physical positions of the markers as the predictor. Since we based our simulations on physical 100 kb intervals, we predicted the genetic positions every 100 kb from the loess-smoothing and sorted them in ascending order (Figure S1 and S2).

Finding the optimal interference distance

In meiosis, accurate chromosomal segregation requires at least one cross-over per chromosome (i.e. the minimal length of a chromosome is 50 cM; Petronczki *et al.* 2003). Because we wanted our simulations to be as realistic as possible, we also required one cross-over per chromosome in each meiosis in our gene-dropping simulations (see below). Yet adding up random and forced cross-overs would lead to an overestimation of the genetic length of the chromosomes. We solved this problem by introducing an interference distance for each chromosome, which suppresses additional cross-overs in a given physical interval surrounding an initial cross-over. We define the optimal interference distance (in

kb) as the distance which returns the distribution of recombination events as close to the actual linkage map as possible. In order to find the optimal interference distance we ran our gene-dropping simulations for each chromosome 10,000 times for interference distances ranging from 0 to maximally 50 Mb (in 100 kb steps) and selected for each chromosome the one with the minimal sum of squared residuals, a residual being the difference between the simulation result and the actual smoothed linkage map (Figure S1 and S2). The optimal interference distances for each chromosome were subsequently used in our gene-dropping simulations. It should be noted that interference actually happens in meiosis (Muller 1916; Sturtevant 1913) and is thus not an artefact introduced by our simulation procedure. We also developed a simulation which does not require interference, which did not change the results (data not shown, but the script is available upon request).

Gene-dropping simulations

Our gene-dropping simulations extend on those described in Libiger & Schork (2007). We first specified a genome by the number and size (both genetically and physically) of chromosomes. Then we split each chromosome in predefined physical segments (we used 100 kb segments) and calculated the recombination probability between segments using a smoothed linkage map (see above) and the Kosambi map function (Kosambi 1943). Finally, we followed each segment through a specified genealogy by using these recombination probabilities and assuming that founders of the pedigree were unrelated.

(1) For each founder we simulated a diploid chromosome set (two unique haplotypes without inbreeding for each chromosome).

(2) The simulation proceeded by creating offspring of the pedigree founders. We tried to simulate meiosis as realistic as possible and implemented the following steps. (a) Prior to meiosis I the homologous chromosomes (2N2C = 2 homologous chromosomes, 2 chromatids) in both the mother and the father duplicate to form two sister chromatids (2N4C) that are identical. (b) In meiosis I (which leads to 1N2C), cross-overs between chromatids of the homologous chromosomes occur with probabilities as defined by the linkage map. Cross-overs may occur between both chromatids of the homologous chromosomes but not between the two sister chromatids of a single chromosome (remember also that the two sister chromatids are identical). Thus, also unrecombined chromatids may get inherited. (c) One of the four chromatids (i.e. 1N1C) in both the mother and the father was chosen randomly to create the offspring which is then 2N2C again.

(3) Within each offspring the total length of all autozygous stretches was determined as homozygosity for a founder haplotype ($GWIBD_7$). The end of an autozygous stretch was placed at a randomly chosen base pair between the flanking autozygous and non-autozygous segment.

(4) To show how our simulations of inbreeding can be extended to also quantify relatedness among individuals we did the following. When all individuals in the pedigree had been simulated, the genetic similarity between two individuals was calculated as the proportion of sharing of founder haplotypes between individuals (the pedigree-based equivalent is the additive genetic relatedness, which is twice the coefficient of coancestry). We focus on the genetic similarity matrix because it is commonly used when estimating variance components in animal models (Speed & Balding 2015; Results are presented in the Supplement, Figure S3).

Simulated pedigrees

We ran our gene-dropping simulations 10,000 times on two pedigrees. (1) A designed pedigree comprised of full-sibs and their offspring (full-sib mating; FSM), first-cousins and their offspring (first-cousin mating; FCM) and second-cousins and their offspring (second-cousin mating; SCM; Figure 2). (2) An empirical pedigree from our captive population of zebra finches held at the Max Planck Institute for Ornithology in Seewiesen, Germany, comprised of n = 159 founders and n = 3,404 individuals in total (Figure S4). The pedigree spans seven generations: in the first three generations the aim was to produce outbred individuals, the fourth generation contains offspring of full-sib matings and in the last three generations selection lines were produced (six lines in total), which increased the overall level of inbreeding. We focus our analysis on the last (seventh) generation of the pedigree (n = 681 individuals) since here we have the most precise information about coancestry (Figure S5; Balloux *et al.* 2004). Although we focus our analyses on a single pedigree, results are qualitatively transferable to other pedigrees, yet results will change quantitatively (e.g. the number of markers needed to reach the same precision as Pedigree F) when considering pedigrees with more or less variance in inbreeding (i.e. different levels of identity disequilibrium; Miller & Coltman 2014). However, the simulation script is available for download and can be applied to any other pedigree. Running the simulation once on the 3,404-individuals pedigree with a zebra finch linkage map takes around 45 minutes to complete on a single computer core (Intel® Core™ i7-2600, 3.4 GHz and 16 Gb RAM), which adds up to a runtime of around 312 days for 10,000 simulation runs.

Estimating the IBD-IBS discrepancy of microsatellites

In the following we use a simplified approach to quantify how often two phylogenetically independent haplotypes (*i.e.* they are not IBD) carry the same allele by chance alone (IBD-IBS discrepancy). A more sophisticated empirical approach to this problem is presented elsewhere (Knief *et al.* 2015), but for the purpose of our simulations the way we do it here (illustrated in Figure S6D) will suffice.

We genotyped the seventh generation of our pedigree with 11 microsatellite markers spread across 11 chromosomes (*Tgu1A*, *Tgu2*, *Tgu3*, *Tgu5*, *Tgu6*, *Tgu11*, *Tgu14*, *Tgu15*, *Tgu22*, *Tgu26*, *Tgu27*). Details on each microsatellite and the PCR protocol are given in Table

S1. For each of the 681 individuals we calculated its average homozygosity across the 11 microsatellites (Marker IBS), which is equivalent to 1 - multilocus heterozygosity. Regressing Marker IBS on the pedigree-based inbreeding coefficient (Pedigree F_7) follows a Berkson error model (Berkson 1950), which means that regression slopes are unbiased. Because of that the regression line will intersect the point [F = 1, Marker IBS = 1], which represents the point of maximal values of the independent and dependent variable, respectively. It follows that the slope β of the regression line is

$$\beta = 1 - \alpha$$

with α being the intercept. We will use the intercept as a measure of the baseline homozygosity in the absence of IBD (IBD-IBS discrepancy). Since regression lines intersect the mean of the data points [average Pedigree F ($\bar{x}$), average Marker IBS ($\bar{y}$)], substituting β with 1 - α in the standard regression equation yields

$$\bar{y} = (1 - \alpha) * \bar{x} + \alpha$$

Thus, the noise due to incomplete information about IBD added to an average microsatellite is

$$\alpha = \text{IBD-IBS discrepancy} = (\bar{y} - \bar{x}) / (1 - \bar{x})$$

Comparing molecular estimates of inbreeding with pedigree-based estimates

To compare the precision of pedigree-based with molecular marker-based estimates of inbreeding we ran our gene-dropping simulations 1,000 times on the empirical seven-generations pedigree for both the zebra finch and the human genome and recorded $GWIBD_7$ and an estimate of inbreeding based on n = 5, 10, 20, 40, 80 and 160 randomly chosen segments in the genome (Marker IBD_7). Ignoring mutations, all markers that were IBD were designated as IBS. We further incorporated the IBD-IBS discrepancy by calling a marker IBS with the above described probability α, irrespective of whether it was IBD or not.

We took the coefficient of determination (r^2) between both the pedigree-based estimate of inbreeding (Pedigree F_7) and $GWIBD_7$ and between $GWIBD_7$ and the marker-based estimates (Marker IBD_7 and Marker IBS_7) in the seventh generation of our pedigree as a measure of precision in predicting the dependent variables (Zar 2010). We also recorded the slopes of OLS regressions with $GWIBD_7$ as the dependent variable and Pedigree F_7, Marker IBD_7 and Marker IBS_7 as predictors to get an estimate of the accuracy of the prediction.

Fitness and heterozygosity-fitness correlations

To illustrate how the various measures of inbreeding are expected to relate to fitness, we used a simplifying approach that is only meant as a proof of principle. For each individual in

the seventh generation of our pedigree we calculated a measure of fitness depending on the individuals' inbreeding level. First, we sampled the number of offspring for each individual from a Poisson distribution with λ slightly larger than two to mimic a stable population (λ would be two if there is no effect of inbreeding, but must be adjusted upwards to account for negative inbreeding effects). Then we introduced a mutational load of one lethal equivalent per haploid genome (Morton *et al.* 1956) by rendering individuals infertile (assigning a fitness of zero offspring) if they happen to be homozygous (IBD) for a recessive deleterious mutation with probability (S) given its inbreeding level as

$$S = 1 - e^{-GWIBD7}$$

We refrained from simulating inbreeding depression in a more complex way (localized effects of varying effect size and allele frequency) because this would not have altered any of our general conclusions. The absolute amount of noise would have been different (see Visscher (2009) for an analogous situation), but not the qualitative answer whether a slope of regression is biased or unbiased.

We used relative fitness (standardized to a mean of one) as the independent variable in heterozygosity-fitness OLS regressions with Pedigree F_7, $GWIBD_7$, Marker IBD_7 or Marker IBS_7 as predictors. As suggested by Szulkin *et al.* (2010) we report both the slopes of the OLS regressions (β), which are a measure of the inbreeding load, and the coefficients of determination (r^2), which quantify the amount of variance in fitness explained by inbreeding.

Results

Mendelian noise in GWIBD and a comparison to analytical results

As has been shown analytically before (Hill & Weir 2011), for any class of individuals with the same inbreeding history and hence same pedigree F, there is considerable variation among individuals in the proportion of the genome that is inherited IBD (Figure 3). This variation – caused by Mendelian sampling noise – is markedly larger in the zebra finch (FSM SD = 0.0838, FCM SD = 0.0461, SCM SD = 0.0231) compared to humans (FSM SD = 0.0454, FCM SD = 0.0251, SCM SD = 0.0115). Although the standard deviation of GWIBD decreases with more distant inbreeding levels, the coefficient of variation (CV), which can be interpreted as a measure of the relative standard deviation, increases in both the zebra finch (FSM CV = 0.335, FCM CV = 0.732, SCM CV = 1.466) and the human (FSM CV = 0.181, FCM CV = 0.402, SCM CV = 0.737) genome, which has also been shown analytically (Hill & Weir 2011).

By constructing an artificial human linkage map with a strict uniform distribution of cross-overs we were able to compare our simulation-based estimates of the standard deviation in

GWIBD in humans with their analytical expectations (calculated with the formulas provided in Franklin 1977; Hill & Weir 2011). The simulations yielded slightly larger standard deviations of GWIBD than expected analytically, at maximum a deviation of 3.3% (FSM SD = 0.0412 vs 0.0420, FCM SD = 0.0226 vs 0.0234, SCM SD = 0.0104 vs 0.0107). The deviation is probably caused by the use of an infinitesimal model in the analytical approach (Franklin 1977; Hill & Weir 2011), whereas we simulated 100 kb segments (for computational feasibility) which slightly increased the standard deviation in IBD sharing. In line with this interpretation, analytical models yield larger standard deviations in IBD sharing between relatives when they use a localized distribution of cross-overs instead of an infinitesimal model (Risch & Lange 1979; Suarez *et al.* 1979; Visscher 2009).

Precision and accuracy of Pedigree F in predicting GWIBD

As shown above, the standard deviation of GWIBD resulting from FSM, FCM and SCM is almost twice as large in zebra finches as in humans because of the difference in their genomic architectures. Likewise, Pedigree F in the seventh generation of the empirical pedigree was more precise in predicting GWIBD when simulating a human linkage map (r^2 = 0.82, 95% quantile range (QR) = 0.79–0.85; Figure 4B) than when simulating a zebra finch linkage map (r^2 = 0.57, 95% QR = 0.49–0.63; Figure 4A). In preceding generations with lower inbreeding levels (where many individuals have F = 0, SD = 0), the pedigree-based inbreeding estimates appeared more precise on average (Table 1), because Mendelian noise can only contribute to variation in GWIBD whenever F > 0.

Regressing $GWIBD_7$ over Pedigree F_7 from 1,000 simulations yielded an unbiased mean slope of β = 1.00 in both zebra finches (95% QR = 0.88–1.12; Figure 4A) and humans (95% QR = 0.93–1.07; Figure 4B), as is expected for a direct cause-effect relationship and a Berkson error model.

Precision and accuracy of Marker IBD and IBS in predicting GWIBD

In the following, we first consider ideal markers which are never homozygous by chance alone (no IBS without IBD). Ideal markers were more precise and more accurate in predicting GWIBD when simulating a zebra finch linkage map than when simulating a human linkage map (compare the red lines in Figures 5A and 5C with 5B and 5D). For instance, 20 markers yielded a precision of r^2 = 0.63 ± 0.025 SE in the zebra finch, but only r^2 = 0.53 ± 0.016 SE in humans. Thus, the markers are more reliable in the species with the less reliable pedigree-based prediction. We then ask: how many markers are needed to obtain higher precision than given by Pedigree F?

In the seventh generation of our empirical pedigree, around 15 randomly distributed segments in the zebra finch genome (out of 11509 autosomal 100kb segments = 0.13% of the autosomal genome) gave the same precision as the pedigree-based estimate (Figure 5A). In the human genome, 80 randomly distributed segments (out of 28801 autosomal

100kb segments = 0.28% of the autosomal genome) were needed (Figure 5B). However, the OLS regression slopes of $GWIBD_7$ over Marker IBD_7 estimated from 15 and 80 randomly distributed segments in the zebra finch and human genome, respectively, were biased downwards and only when using around 160 segments to estimate Marker IBD_7 the slopes became almost unbiased (Figure 5C, D).

We now consider non-ideal microsatellite markers, which can be IBS without being IBD. We empirically estimated the IBD-IBS discrepancy (mean from 11 microsatellites) as 13.3%. After incorporating this into our simulations around 40 randomly distributed segments in the zebra finch genome (out of 11509 autosomal 100kb segments = 0.35% of the autosomal genome) are needed to obtain an as precise estimate of $GWIBD_7$ as the pedigree-based estimate in the seventh generation of our pedigree (Figure 5A). By assuming the same IBD-IBS discrepancy in humans we found that around 160 randomly distributed segments (out of 28801 autosomal 100kb segments = 0.56% of the autosomal genome) were needed (Figure 5B). Regressing $GWIBD_7$ over Marker IBS_7 yielded slopes that were biased downwards. Using around 80 and 160 segments for the marker-based inbreeding estimates in zebra finches and humans yields comparable slopes between Marker IBD_7 and Marker IBS_7, which is due to the fact that the standard deviation of Marker IBS_7 is decreasing more rapidly than the standard deviation of Marker IBD_7, probably because the distribution of Marker IBS_7 is less skewed (the slope of an OLS regression line with independent variable x and dependent variable y is calculated as $r_{xy} * SD_y / SD_x$; Figure 5C, D). However, one should keep in mind that the precision of Marker IBS_7 is lower than the precision of Marker IBD_7.

Precision and accuracy in heterozygosity-fitness correlations

The best predictor in terms of precision and accuracy of the inbreeding load was $GWIBD_7$, because we introduced inbreeding depression into our simulations by using $GWIBD_7$, and hence all other estimators should be compared to this standard (r^2 = 0.013, 95% QR = 0.00089–0.032 and β = -0.94, 95% QR = -1.60–-0.27 in zebra finches and r^2 = 0.010, 95% QR = 0.00022–0.028 and β = -0.98, 95% QR = -1.78–-0.14 in humans; Figure 6). Interestingly, $GWIBD_7$ explains more of the variance in fitness in zebra finches than in humans (elevation of yellow line in Figure 6A versus 6B), presumably because there is a greater realized variance in $GWIBD_7$ in zebra finches than in humans (compare scatter in Figure 4A versus 4B) and consequently the coefficient of determination is larger (King 1986). Generally, the large amounts of scatter in the relationships with fitness (low r^2 in Figures 6A and 6B) follow from chance events in drawing a lethal equivalent that reduces fitness to zero. This scatter could have been reduced by simulating a larger number of recessive deleterious mutations, each with a smaller negative effect on fitness, but this is not critical for the following assessment of precision and accuracy.

Pedigree F_7 does not capture the Mendelian noise (of which there is more in zebra finches than in humans) and accordingly the precision of Pedigree F_7 in predicting fitness was lower,

which was more pronounced in case of the zebra finch linkage map than in case of the human linkage map (r^2 = 0.0080, 95% QR = 0.00011–0.024 and r^2 = 0.0086, 95% QR = 0.00018–0.025 for zebra finches and humans, respectively; compare black to yellow lines in Figures 6A and 6B). However, although Pedigree F_7 did not account for the Mendelian noise in each individual, the regression slope of fitness over Pedigree F_7 was unbiased when using both the zebra finch (β = -0.95, 95% QR = -1.80–-0.083) and human linkage map (β = -0.99, 95% QR = -1.85–-0.094; compare black to yellow lines in Figures 6C and 6D). This finding of unbiased slopes is in line with expectations for cause-effect relationships (Pedigree F is causal to fitness, Figure 1) and Berkson error models.

The precision of Marker IBD_7 and Marker IBS_7 in predicting the inbreeding load increased when more markers were used to estimate the inbreeding level of an individual. Equivalent to the $GWIBD_7$ predictions, about 15 ideal markers (Marker IBD_7) and 40 non-ideal markers (Marker IBS_7) were needed with a zebra finch linkage map to reach the same precision as Pedigree F_7, and about 80 ideal markers (Marker IBD_7) and more than 160 non-ideal markers (Marker IBS_7) in case of a human linkage map. Using molecular markers as predictors of fitness in an OLS regression yielded shallower slopes than when using $GWIBD_7$ or Pedigree F_7 as the predictors but became more accurate when increasing the number of markers. These downward biased slopes are expected when using OLS regression (Marker IBD and IBS are not causal to fitness but related indirectly, see Figure 1).

The expectations and findings for error components and slopes are summarized in Table S2. Illustrative examples of scatter plots showing these relationships are presented in Figure S6.

Discussion

Mendelian noise in GWIBD

Both the genetic map length and the number of chromosomes under consideration are known to influence the variation in IBD (Franklin 1977; Hill & Weir 2011; Stam 1980). Similarly, it has been predicted that cross-over interference and the distribution of cross-overs along chromosomes will influence the amount of variance in IBD (e.g. Forstmeier *et al.* 2012; Guo 1995; Rasmuson 1993; Risch & Lange 1979), but to our knowledge it has never received attention in a modelling framework. Here we show that the variance in IBD is much larger in zebra finches than in humans, because almost half of the genome is inherited in only four segments (i.e. the interiors of chromosomes *Tgu1*, *Tgu1A*, *Tgu2* and *Tgu3*) which only rarely break up by cross-overs (Backström *et al.* 2010). Within birds, the zebra finch linkage map appears to be quite special in its distribution of recombination (Kawakami *et al.* 2014, but see Calderón & Pigozzi 2006), yet it should be noted that even more extreme examples of a skewed distribution of recombination can be found (e.g. corn (*Zea mays*); Gore *et al.* 2009).

Comparison of pedigree- and molecular marker-based estimates of inbreeding

Due to differences in Mendelian sampling noise the precision of pedigree-based estimates of GWIBD is higher in humans than in zebra finches. In our empirical pedigree, assessing around 15 and 80 genomic regions for their IBD status (Marker IBD_7) in zebra finches and humans yields as precise estimates of $GWIBD_7$ as a complete seven-generations pedigree. The surprisingly small number of genomic regions needed in the zebra finch genome is in good agreement with an earlier empirical estimate for our captive population which suggested that 11 microsatellites reflect an individual's realized inbreeding coefficient equally well as the pedigree (Forstmeier *et al.* 2012). Microsatellites used in that study were spread across nine chromosomes, including the macrochromosomes *Tgu1*, *Tgu1A*, *Tgu2*, *Tgu3*, *Tgu5*, *Tgu6* and *Tgu9*, which together sum up to half the physical zebra finch genome and rarely break up in meiosis (Backström *et al.* 2010). Thus, they are potentially more informative than a random set of genomic segments (as considered here). Due to a limited number of segregating haplotypes in our captive population, being IBS for a single microsatellite reflects IBD quite well (Forstmeier *et al.* 2012), but this may not be the case in large and panmictic populations in the wild (Knief *et al.* 2015). To our knowledge, empirical field studies have rarely assessed the extent to which IBS of microsatellite markers reflects IBD of the surrounding genomic region, a question that now can be addressed by either using dense SNP panels or several microsatellites located within small genomic regions (e.g. 100kb, see Knief *et al.* 2015). As expected, incorporating the IBD-IBS discrepancy into our simulations decreased the precision of the molecular markers and consequently more markers were needed to get as precise estimates of $GWIBD_7$ as with Pedigree F_7 (40 and 160 markers with IBD-IBS discrepancy = 13.3% in zebra finches and humans, respectively).

The lower number of markers needed in zebra finches as compared to humans to reach the same precision as Pedigree F can largely be explained by the lower precision of Pedigree F in the former. Yet each single marker in the zebra finch genome also contributes more to an increase in precision of Marker IBD or Marker IBS than in the human genome, which is reflected in the steeper increase in precision with an increasing number of markers (see Figure 5A versus 5B). Whenever a marker in the zebra finch genome is located in the center of a macrochromosome, it will add additional information on the inbreeding level of an individual, but whenever it is located more towards the telomeres it will be less informative. In humans, the pronounced block-like inheritance of genomic regions is absent and consequently each marker adds approximately the same but on average less information, which is also evident in the smaller standard errors in humans compared to zebra finches in Figure 5.

Implications for heterozygosity-fitness correlations and regression slopes

In studies of heterozygosity-fitness correlations (in the wider sense) we are generally interested in both the slope of an OLS regression of a fitness-related trait over an estimate of inbreeding and the coefficient of determination (Szulkin *et al.* 2010). Both Pedigree F and

GWIBD gave unbiased slopes when used as predictors of fitness because both are causal to variation in fitness (see causal arrows in Figure 1), and (in our simulations) both are known without error. Thus, both provide an unbiased estimate of the inbreeding load, which is defined as the reduction in mean fitness within a population due to inbreeding (Szulkin *et al.* 2010). This interpretation is consistent with a 'controlled' experiment as described in Berkson (1950). Since Pedigree F is the expected value of GWIBD, which we quantify without error, both will yield unbiased slopes in an OLS regression. Yet, when using Pedigree F instead of GWIBD as a predictor of fitness, we are one step further away from the dependent variable (F affects GWIBD which in turn affects fitness), so the relationship contains the component of Mendelian noise and hence the coefficient of determination (r^2) is lower.

In contrast, using molecular markers as predictors of fitness yields slopes that are biased downwards because the relationship is indirect (GWIBD affects both the predictor and the dependent variable, see Figure 1) and the predictor is measured with error (including marker sampling noise and the IBD-IBS discrepancy; Szulkin *et al.* 2010). Hence, the assumption of OLS regression that all noise lies in the dependent variable rather than in the predictor is violated, leading to shallower regression slopes (Forstmeier 2011). The bias becomes smaller the more markers are used, but can only be completely eliminated if the whole genome would be sampled, which is equivalent to assessing GWIBD. Nevertheless, with a sufficient number of markers the coefficient of determination in heterozygosity-fitness correlations becomes higher using molecular markers than pedigree information, because marker sampling noise is reduced whereas Mendelian noise stays constant. Consequently, molecular markers explain more of the variance in fitness caused by inbreeding.

When comparing pedigree versus marker-based information, it should be noted that we simulated a perfect pedigree. In reality, pedigrees are usually incomplete for two reasons. (1) Missing or wrongly assigned parentage will introduce (random) errors in Pedigree F and will consequently bias the slopes of relationships with fitness downwards and also lead to a lower coefficient of determination. (2) Pedigree-based estimates of inbreeding assume that the founders are unrelated which is normally not the case. Because of that Pedigree F is biased downwards but this systematic bias has no effect on the slope of the heterozygosity-fitness regression as long as the background relatedness does not become too high (the relationship between inbreeding and fitness is not linear; Morton *et al.* 1956). However, the bias in Pedigree F will lead to a downward bias in the coefficient of determination in the heterozygosity-fitness correlation.

Conclusion

Our results suggest that it is advantageous to use either molecular markers or pedigrees to estimate inbreeding, depending on the relative importance of accurately and precisely

predicting individual inbreeding coefficients or of accurately quantifying the slopes in heterozygosity-fitness regressions.

Mendelian noise is species specific (Rasmuson 1993) and consequently cannot be changed within a study organism. Because of that Pedigree F will always underestimate the effect size of GWIBD on fitness. This is also true for molecular markers but already a limited number of informative markers will result in higher precision than Pedigree F. It remains to be tested empirically whether it is more effective to reduce the marker sampling noise (by increasing the number of randomly distributed markers) or the IBD-IBS discrepancy (by clustering markers to get more reliable information about IBD).

Provided that the utilized pedigree comes without parentage errors and covers a sufficient number of generations, Pedigree F will yield unbiased slopes in predictions of fitness, which are often used to quantify the inbreeding load in a population. In contrast, the slopes obtained by using molecular markers as predictors of fitness will be biased downwards. Increasing the number of markers or reducing the IBD-IBS discrepancy will lessen the bias but it is only when the whole genome is sampled (which is essentially GWIBD) that they will yield unbiased estimates of the regression slopes.

Acknowledgements

We are grateful to S. Bauer, E. Bodendorfer, A. Grötsch, A. Kortner, K. Martin, P. Neubauer, F. Weigel, and B. Wörle for animal care and help with breeding and to M. Schneider for help with genotyping. Y.G. Araya-Ajoy and S. van Dongen gave valuable comments on an earlier version of the manuscript. L. Keller brought the Berkson error model to our attention. This study was funded by the Max Planck Society. U.K. is part of the International Max Planck Research School for Organismal Biology.

References

Backström N, Forstmeier W, Schielzeth H*, et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research* **20**, 485–495.

Balloux F, Amos W, Coulson T (2004) Does heterozygosity estimate inbreeding in real populations? *Molecular Ecology* **13**, 3021–3031.

Berkson J (1950) Are there two regressions? *Journal of the American Statistical Association* **45**, 164–180.

Broman KW, Weber JL (1999) Long homozygous chromosomal segments in reference families from the centre d'étude du polymorphisme humain. *American Journal of Human Genetics* **65**, 1493–1500.

Calderón PL, Pigozzi MI (2006) MLH1-focus mapping in birds shows equal recombination between sexes and diversity of crossover patterns. *Chromosome Research* **14**, 605–612.

Collins FS, Lander ES, Rogers J, Waterston RH, Conso IHGS (2004) Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945.

Fisher RA (1949) *The theory of inbreeding* Oliver & Boyd, Edinburgh, UK.

Forstmeier W (2011) Women have relatively larger brains than men: a comment on the misuse of general linear models in the study of sexual dimorphism. *Anatomical Record* **294**, 1856–1863.

Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B (2012) Heterozygosity-fitness correlations in zebra finches: microsatellite markers can be better than their reputation. *Molecular Ecology* **21**, 3237–3249.

Franklin IR (1977) The distribution of the proportion of the genome which is homozygous by descent in inbred individuals. *Theoretical Population Biology* **11**, 60–80.

Goldstein DB, Schlötterer C (1999) *Microsatellites: evolution and applications* Oxford University Press.

Gore MA, Chia JM, Elshire RJ*, et al.* (2009) A first-generation haplotype map of maize. *Science* **326**, 1115–1117.

Guo SW (1995) Proportion of genome shared identical by descent by relatives: concept, computation, and applications. *American Journal of Human Genetics* **56**, 1468–1476.

Hill WG, Weir BS (2011) Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genetics Research* **93**, 47–64.

Kawakami T, Smeds L, Backström N*, 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.

Keller MC, Visscher PM, Goddard ME (2011) Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics* **189**, 237–249.

King G (1986) How not to lie with statistics: avoiding common mistakes in quantitative political science. *American Journal of Political Science* **30**, 666–687.

Knief U, Hemmrich-Stanisak G, Wittig M*, et al.* (2015) Quantifying realized inbreeding in wild and captive animal populations. *Heredity*.

Kong A, Masson G, Frigge ML*, et al.* (2008) Detection of sharing by descent, long-range phasing and haplotype imputation. *Nature Genetics* **40**, 1068–1075.

Kosambi DD (1943) The estimation of map distances from recombination values. *Annals of Eugenics* **12**, 172–175.

Libiger O, Schork NJ (2007) A simulation-based analysis of chromosome segment sharing among a group of arbitrarily related individuals. *European Journal of Human Genetics* **15**, 1260–1268.

Malécot G (1948) *Les mathématiques de l'hérédité* Masson, Paris.

Matise TC, Chen F, Chen WW*, et al.* (2007) A second-generation combined linkage–physical map of the human genome. *Genome Research* **17**, 1783–1786.

Miller JM, Coltman DW (2014) Assessment of identity disequilibrium and its relation to empirical heterozygosity fitness correlations: a meta-analysis. *Molecular Ecology* **23**, 1899–1909.

Morton NE, Crow JF, Muller HJ (1956) An estimate of the mutational damage in man from data on consanguineous marriages. *Proceedings of the National Academy of Sciences of the United States of America* **42**, 855–863.

Muff S, Riebler A, Held L, Rue H, Saner P (2015) Bayesian analysis of measurement error models using integrated nested Laplace approximations. *Journal of the Royal Statistical Society Series C-Applied Statistics* **64**, 231–252.

Muller HJ (1916) The mechanism of crossing-over. *American Naturalist* **50**, 193–221.

Petronczki M, Siomos MF, Nasmyth K (2003) Un ménage à quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112**, 423–440.

Pigozzi MI, Solari AJ (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.

Powell JE, Visscher PM, Goddard ME (2010) Reconciling the analysis of IBD and IBS in complex trait studies. *Nature Reviews Genetics* **11**, 800–805.

Prado-Martinez J, Hernando-Herraez I, Lorente-Galdos B*, et al.* (2013) The genome sequencing of an albino Western lowland gorilla reveals inbreeding in the wild. *Bmc Genomics* **14**.

R Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rasmuson M (1993) Variation in genetic identity within kinships. *Heredity* **70**, 266–268.

Risch N, Lange K (1979) Application of a recombination model in calculating the variance of sib pair genetic identity. *Annals of Human Genetics* **43**, 177–186.

Roesti M, Moser D, Berner D (2013) Recombination in the threespine stickleback genome—patterns and consequences. *Molecular Ecology* **22**, 3014-3027.

Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2011) Data from: QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. *Dryad Digital Repository*.

Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2012) QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. *Evolution* **66**, 18–30.

Sokal RR, Rohlf FJ (1995) *Biometry* W. H. Freeman, New York.

Speed D, Balding DJ (2015) Relatedness in the post-genomic era: is it still useful? *Nature Reviews Genetics* **16**, 33–44.

Stam P (1980) The distribution of the fraction of the genome identical by descent in finite random mating populations. *Genetical Research* **35**, 131–155.

Sturtevant AH (1913) A third group of linked genes in *Drosophila ampelophila*. *Science* **37**, 990–992.

Suarez BK, Reich T, Fishman PM (1979) Variability in sib pair genetic identity. *Human Heredity* **29**, 37–41.

Szulkin M, Bierne N, David P (2010) Heterozygosity-fitness correlations: a time for reappraisal. *Evolution* **64**, 1202–1217.

Thompson EA (1976) Population correlation and population kinship. *Theoretical Population Biology* **10**, 205–226.

Thompson EA (2013) Identity by descent: variation in meiosis, across genomes, and in populations. *Genetics* **194**, 301–326.

Visscher PM (2009) Whole genome approaches to quantitative genetics. *Genetica* **136**, 351–358.

Warren WC, Clayton DF, Ellegren H*, et al.* (2010) The genome of a songbird. *Nature* **464**, 757–762.

Wright S (1922) Coefficients of inbreeding and relationship. *American Naturalist* **56**, 330–338.

Zar JH (2010) *Biostatistical analysis*, 5 edn. Prentice-Hall/Pearson, New Jersey.

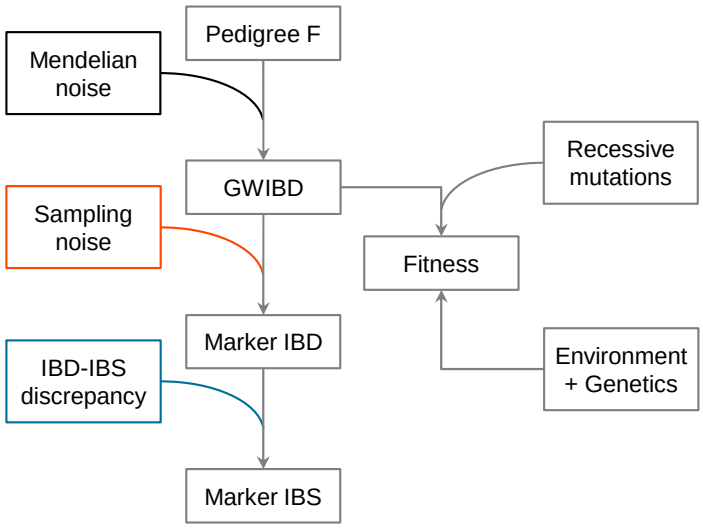


Figure 1: *Conceptualization of the causality in inbreeding and heterozygosity-fitness correlations/regressions. Arrowheads represent the causal direction (see also Szulkin et al. 2010). Mating patterns as represented by pedigrees are causal to genome-wide identity-by-descent (GWIBD) and Mendelian segregation is adding noise to this effect. We are generally interested in estimating GWIBD because it is most directly related to the homozygous mutational load and inbreeding depression. The heterogeneous distribution of recessive deleterious mutations is introducing noise into that relation. The environment and other genetic factors like epistatic interactions further add noise to fitness. Importantly, both Pedigree F and GWIBD are error-free predictors in fitness-over-inbreeding regressions and consequently the regression slopes are unbiased. Berkson (1950) explains this in terms of a 'controlled' experiment, in which case both an underlying variable measured accurately (GWIBD) and its expected value (Pedigree F) give unbiased regression slopes (see also Muff et al. 2015). Frequently, molecular markers are used as predictors of GWIBD and fitness. Yet the direction of causality is reversed, namely GWIBD is causing Marker IBD which in turn is affecting Marker IBS. Both dependencies are affected by noise components, which will introduce error in the predictors (Marker IBD or Marker IBS) in heterozygosity-fitness regressions. Consequently, the regression slopes will be biased downwards.*

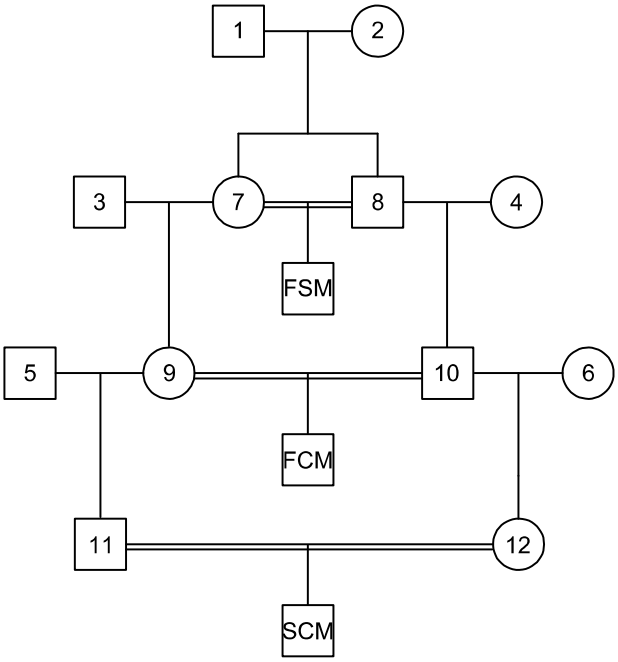


Figure 2: *The designed pedigree used to illustrate the variance in relatedness and inbreeding for full-sibs (individuals 7, 8) and their offspring (full-sib mating, FSM), for first-cousins (individuals 9, 10) and their offspring (first cousin mating, FCM) and for second-cousins (individuals 11, 12) and their offspring (second cousin mating, SCM). Circles represent females and squares males (note that we simulate identical genomes for both sexes). Lines connect parents and offspring and double lines represent inbreeding.*

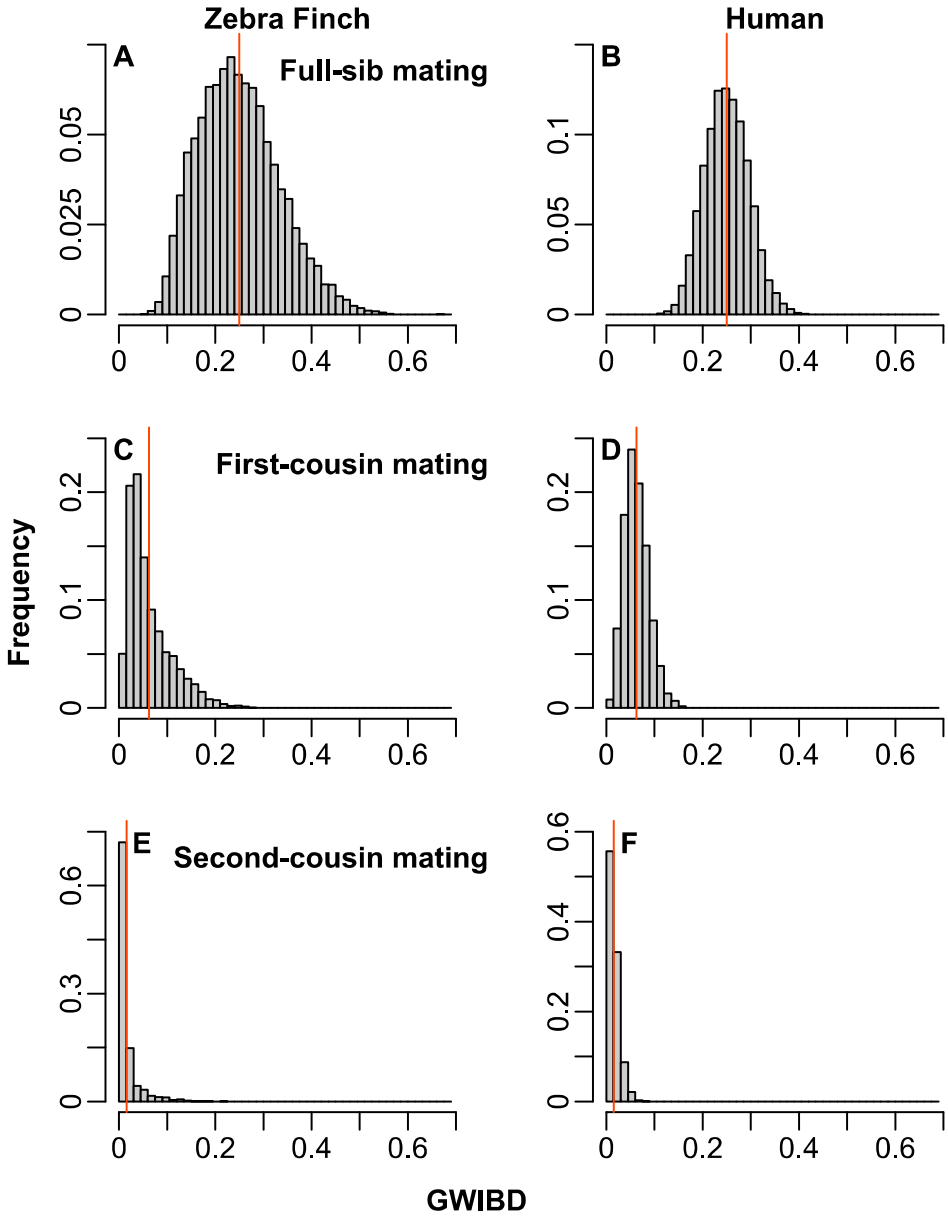


Figure 3: *Variance in inbreeding (realized genome-wide identity-by-descent GWIBD) for offspring of a full-sib mating (A, B), first-cousin mating (C, D) and second-cousin mating (E, F). Left panels show the results for zebra finches, panels on the right the results for humans. The red lines indicate the expected inbreeding coefficients calculated with Wright's path method (1/4, 1/16 and 1/64, respectively).*

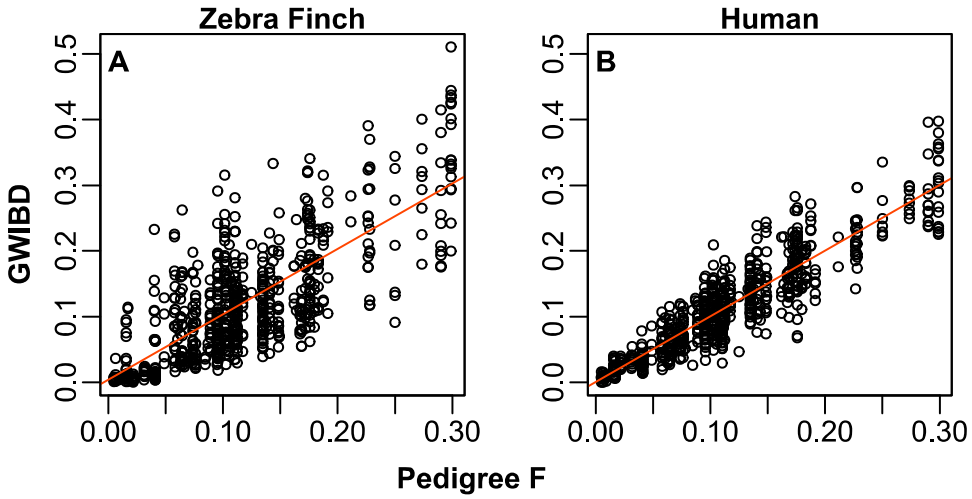


Figure 4: *An exemplary simulation of the realized genome-wide identity-by-descent (GWIBD) in 681 individuals of the seventh generation of our pedigree over the expected values (Pedigree F). Simulations based on the linkage maps of (A) the zebra finch and (B) the human genome. Shown are the most representative simulation runs (out of 1,000 each) where regression slopes (β = 1.00, 95% QR = 0.88–1.12 and β = 1.00, 95% QR = 0.93–1.07) and coefficients of determination (r^2 = 0.57, 95% QR = 0.49–0.63 and r^2 = 0.82, 95% QR = 0.79–0.85) were closest to the mean values from 1,000 runs.*

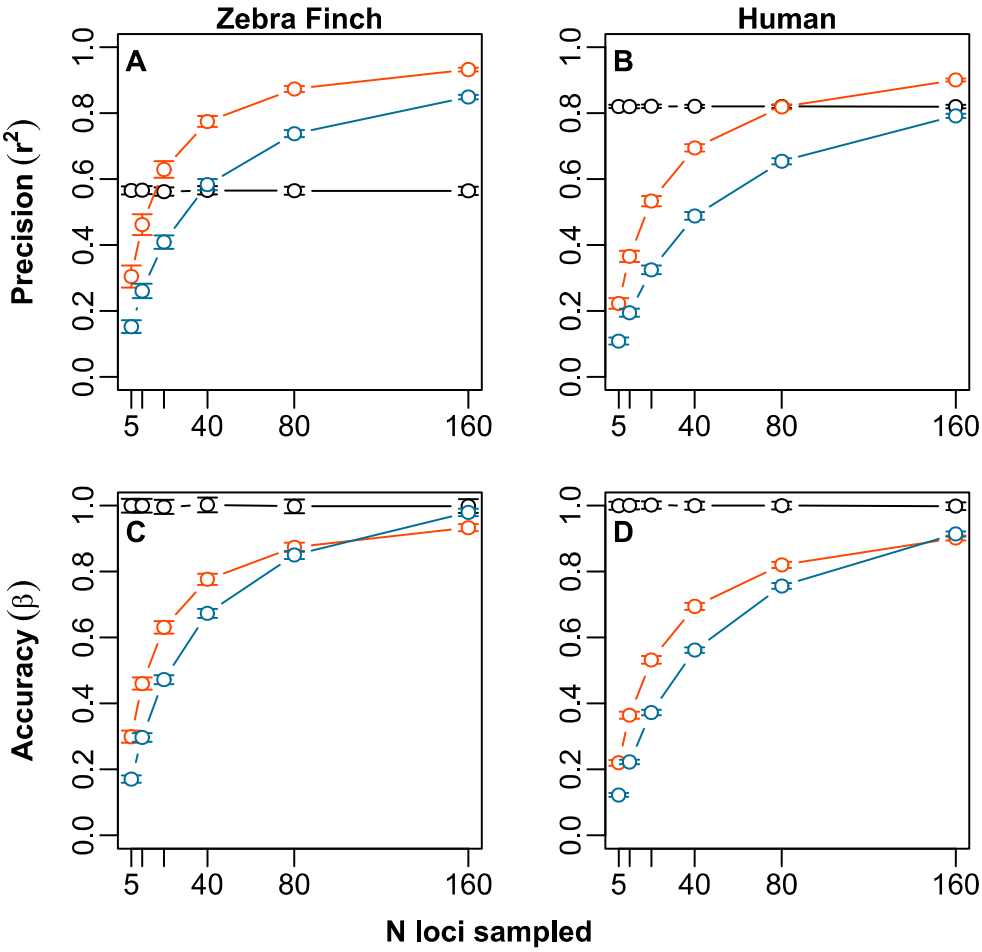


Figure 5: *Comparison of the precision (A, B) and accuracy (C, D) when predicting GWIBD by Pedigree F_7 (black), Marker IBD_7 (red), and Marker IBS_7 (blue). The left and right panels are estimates from 1,000 simulation runs in zebra finches and humans, respectively. The black line indicates the average precision and accuracy of pedigree-based estimates (Pedigree F_7) of inbreeding (± 1 SE) which is not influenced by the number of markers. The red and blue lines indicate the average precision and accuracy (± 1 SE) of Marker IBD_7 and Marker IBS_7, respectively, for varying numbers of markers (which are 100 kb genomic segments) used for predicting the inbreeding level of individuals ($GWIBD_7$) in the seventh generation of the empirical pedigree.*

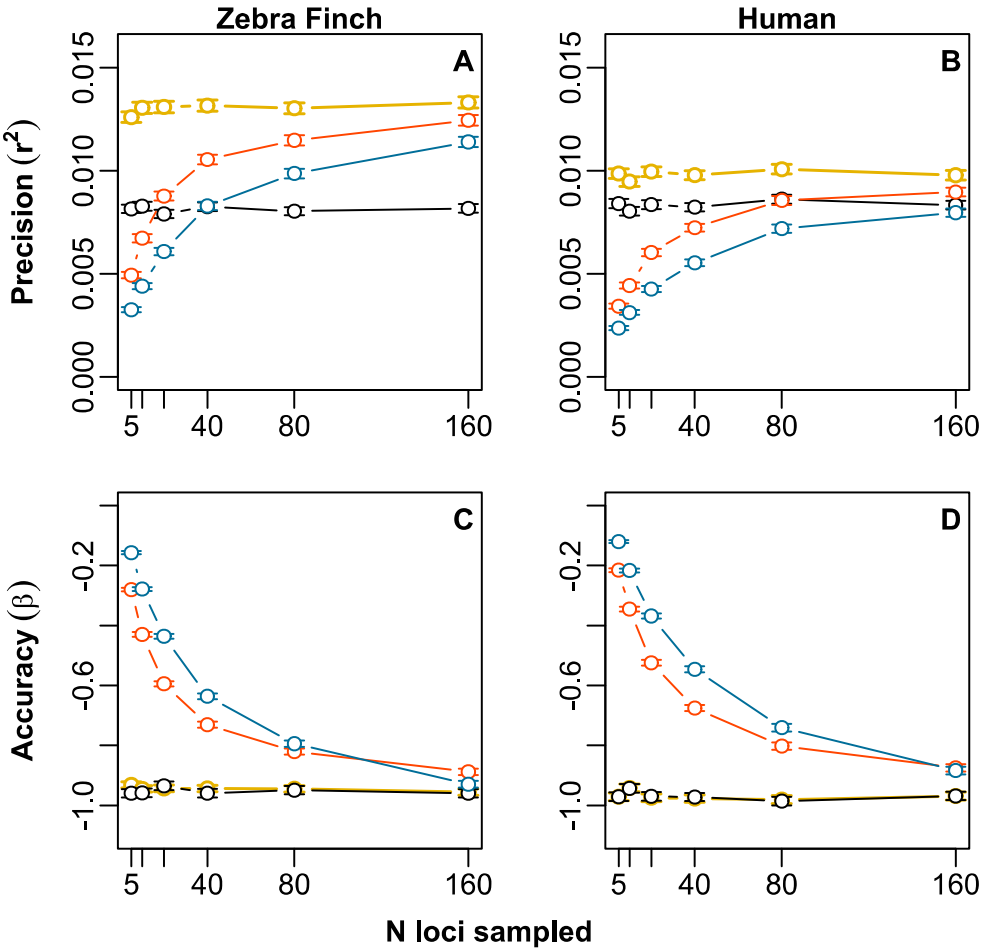


Figure 6: *Comparison of the precision (A, B) and accuracy (C, D) of pedigree- and molecular marker-based estimates of inbreeding depression. The left and right panels are estimates from 1,000 simulation runs in zebra finches and humans, respectively. The yellow line indicates the average precision and accuracy of $GWIBD_7$ in estimating the inbreeding load. $GWIBD_7$ is the best estimator of inbreeding depression and the other estimators should be compared to it. The black line indicates the average precision and accuracy of Pedigree F_7 (± 1 SE) which is not influenced by the numbers of markers. The red and blue line indicate the average precision and accuracy (± 1 SE) of Marker IBD_7 and Marker IBS_7, respectively, for varying numbers of markers (which are 100 kb genomic segments) used for predicting the inbreeding load of individuals in the seventh generation of the empirical pedigree.*

Table 1: *Precision of Pedigree F in predicting GWIBD within each generation of the empirical zebra finch pedigree. In the first three generations close inbreeding was avoided, the fourth generation contains offspring of full-sib matings and in the last three generations six selection lines were bred which increased the inbreeding coefficient.*

Generation	Mean Pedigree F	SD Pedigree F	% individuals with Pedigree F=0	N individuals	Precision of Pedigree F (r^2)
1	0	0	100%	231	-
2	0	0	100%	309	-
3	0.00190	0.00131	97.15%	526	0.718
4	0.08783	0.03576	39.87%	153	0.826
5	0.01877	0.01367	57.89%	710	0.751
6	0.06787	0.04422	1.89%	635	0.476
7	0.11813	0.05697	0%	681	0.562
All	0.04663	0.02550	47.30%	3245	0.755

Supplement

Results

Variation in relatedness and comparison to empirical estimates

As has been previously shown analytically (Hill & Weir 2011) there is considerable variation around the expected additive genetic relatedness between individuals in both the zebra finch (full-sibs SD = 0.072, first-cousins SD = 0.043, second-cousins SD = 0.023) and the human (full-sibs SD = 0.042, first-cousins SD = 0.023, second-cousins SD = 0.013) genome (Figure S3). Although the variance in relatedness estimates decreases in more distant relationships, the coefficient of variation (CV), which can be interpreted as a measure of the relative standard deviation, increases in both the zebra finch (full-sibs CV = 0.14, first-cousins CV = 0.34, second-cousins CV = 0.74) and the human (full-sibs CV = 0.084, first-cousins CV = 0.18, second-cousins CV = 0.42; inbreeding: FSM CV = 0.18, FCM CV = 0.40, SCM CV = 0.77) genome, which has also been shown analytically (Hill & Weir 2011).

The comparison between the estimated variance in IBD sharing between full-sibs in humans from our gene-dropping simulation using the Rutgers Map v.2 and empirical estimates indicated a slightly larger standard deviation resulting from our simulations (SD = 0.0420 vs 0.036 (Visscher *et al.* 2007; Visscher *et al.* 2006 and 0.0395 Gagnon *et al.* 2005). All empirical estimates of IBD sharing had been performed with the Merlin package (Abecasis *et al.* 2002), which could potentially buffer some of the actual variance in IBD sharing (yet Gagnon *et al.* 2005 account for this). More importantly, Merlin's estimates of IBD sharing are based solely on the genetic distances between markers and thus do not take variation in recombination rate into account. Consequently, the empirical estimates of the variance in IBD are similar to the analytical expectations (Hill & Weir 2011), yet for most questions we should be more interested in the actual physical genome shared IBD. Visscher *et al.* (2006) also note that the empirical variance in IBD sharing (SD of IBD sharing between full-sibs = 0.036) which they estimated is likely to be an underestimate due to imperfect marker information. Furthermore, genotyping errors will decrease the variance in IBD sharing, which could bias the empirical estimates downwards but only a rather unrealistic error rate of 5% would reduce the variance in a comparable magnitude (Gagnon *et al.* 2005).

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nature Genetics* **30**, 97–101.

Backström N, Forstmeier W, Schielzeth H*, et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research* **20**, 485–495.

Forstmeier W, Schielzeth H, Schneider M, Kempenaers B (2007) Development of polymorphic microsatellite markers for the zebra finch (*Taeniopygia guttata*). *Molecular Ecology Notes* **7**, 1026–1028.

Gagnon A, Beise J, Vaupel JW (2005) Genome-wide identity-by-descent sharing among CEPH siblings. *Genetic Epidemiology* **29**, 215–224.

Hill WG, Weir BS (2011) Variation in actual relationship as a consequence of Mendelian sampling and linkage. *Genetics Research* **93**, 47–64.

Kinghorn B, Kinghorn A (2010) Pedigree Viewer 6.5. *University of New England, Armidale, Australia*.

Matise TC, Chen F, Chen WW*, et al.* (2007) A second-generation combined linkage–physical map of the human genome. *Genome Research* **17**, 1783–1786.

Visscher PM, Macgregor S, Benyamin B*, et al.* (2007) Genome partitioning of genetic variation for height from 11,214 sibling pairs. *American Journal of Human Genetics* **81**, 1104–1110.

Visscher PM, Medland SE, Ferreira MAR*, et al.* (2006) Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *Plos Genetics* **2**, 316–325.

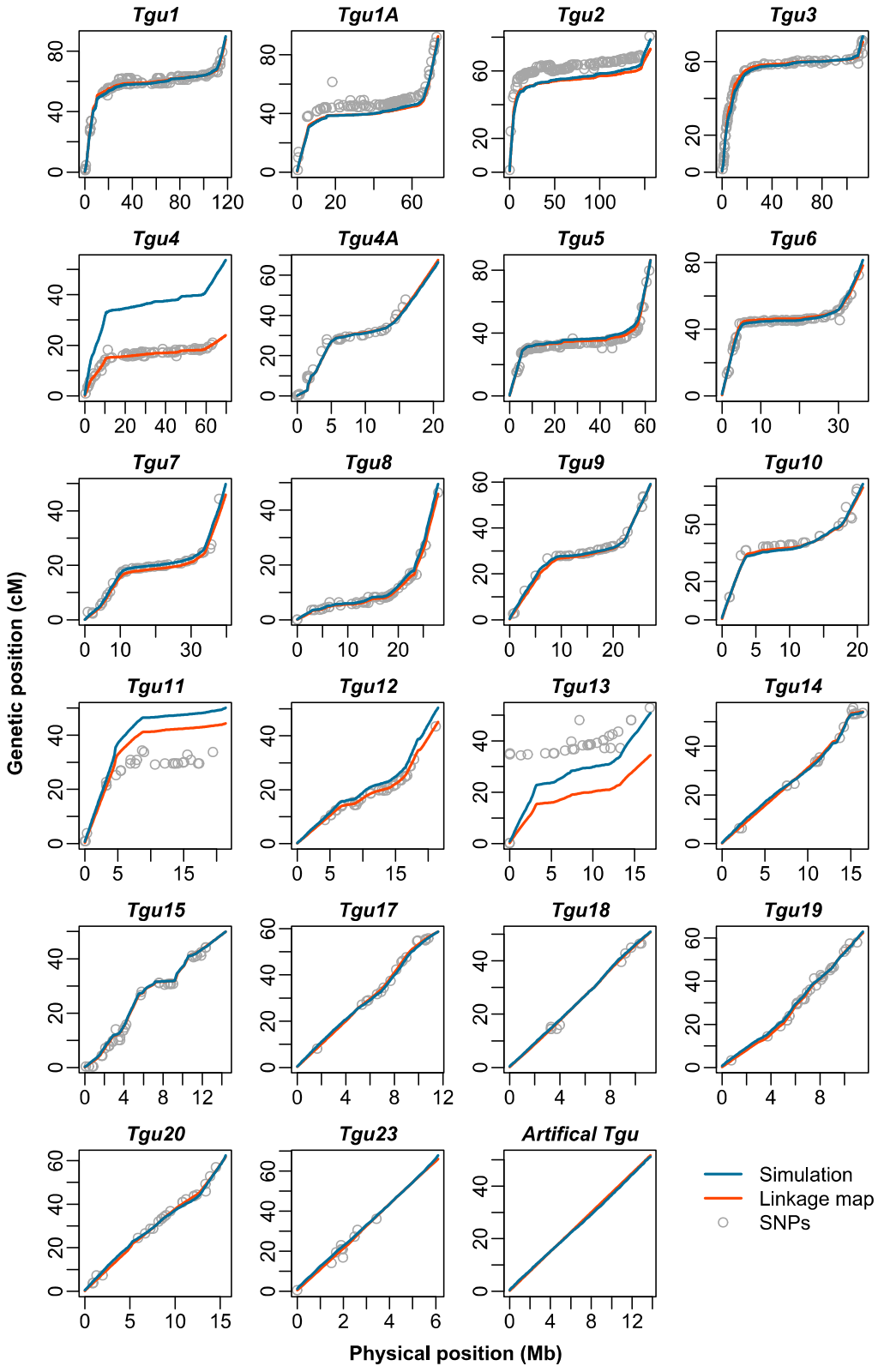


Figure S1: *Actual marker positions (gray dots) and loess smoothed linkage maps (red lines; Backström et al. 2010) for each chromosome in the zebra finch genome (WUSTL v3.2.4 assembly) which were used in the gene-dropping simulations. The "Artificial Tgu" was duplicated 17 times in order to have 39 autosomes. Overlaid are the distributions of cross-overs from n = 10,000 simulation runs using the optimal interference distance (blue lines).*

Lines were shifted to originate from [0, 0]. Remember that the minimal genetic length of a chromosome is 50 cM and that our simulations actually force each chromosome to be at least of that length.

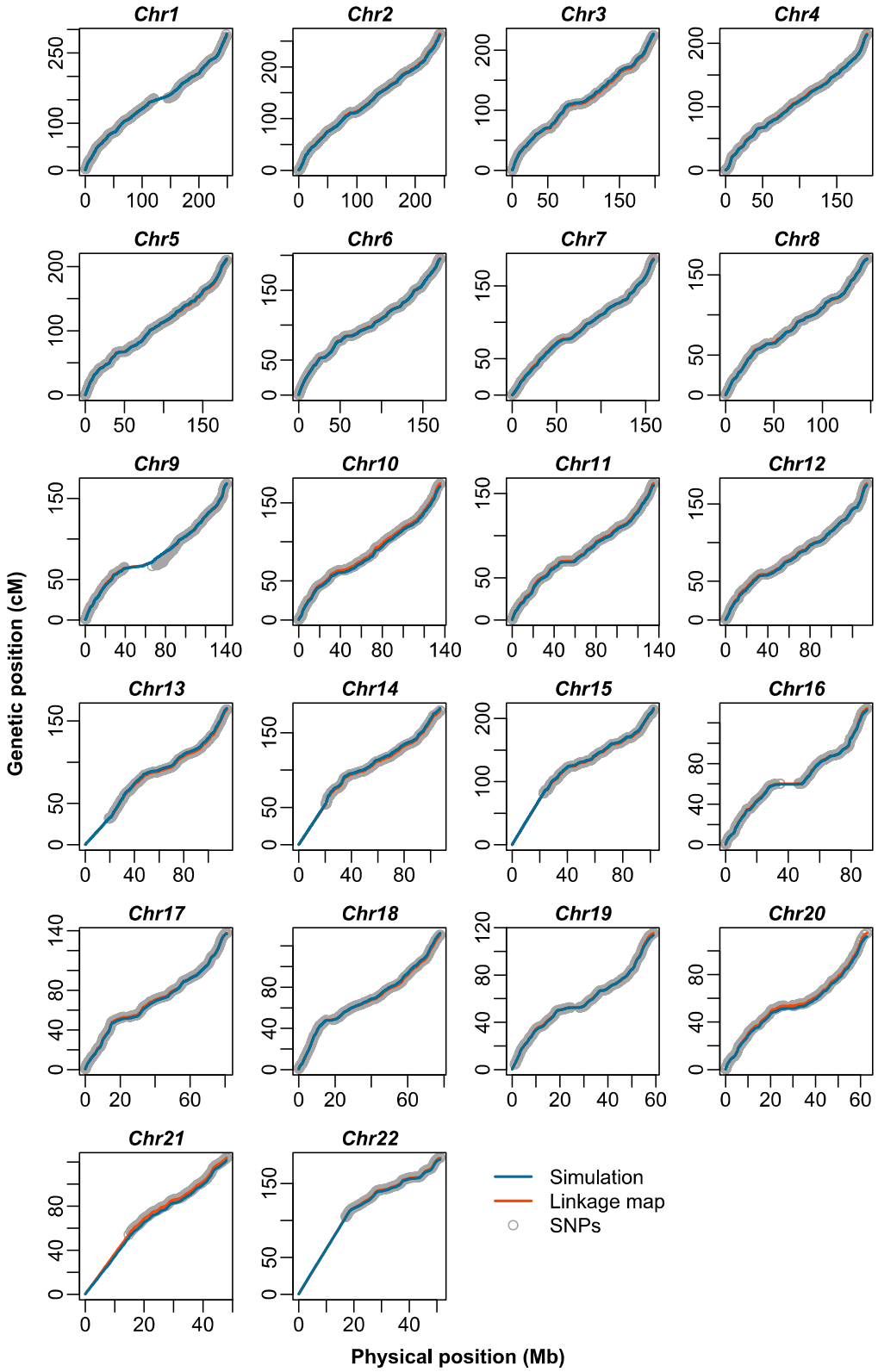


Figure S2: *Actual marker positions (gray dots) and loess smoothed linkage maps (red lines; Matise et al. 2007) for each chromosome in the human genome (hg19/GRCh37 assembly) which were used in the gene-dropping simulations. Overlaid are the distributions of cross-overs from n = 10,000 simulation runs using the optimal interference distance (blue lines). Lines were shifted to originate from [0, 0].*

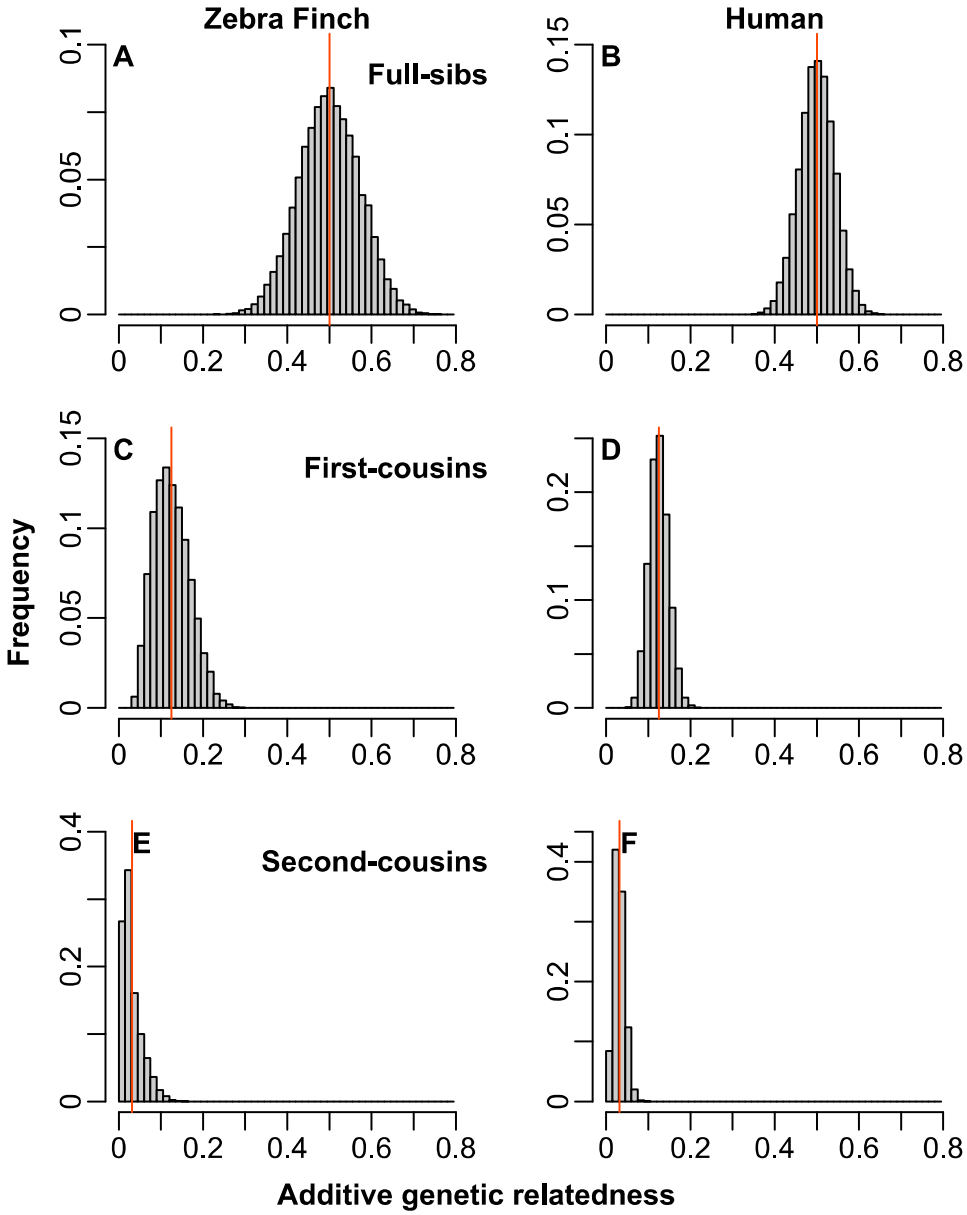


Figure S3: *Variance in in additive genetic relatedness for (A, B) full sibs, (C, D) first cousins and (E, F) second cousins. Left panels show the results for zebra finches, panels on the right the results for humans. The red lines indicate the expected additive genetic relatedness calculated with Wright's path method (1/2, 1/8 and 1/32, respectively).*

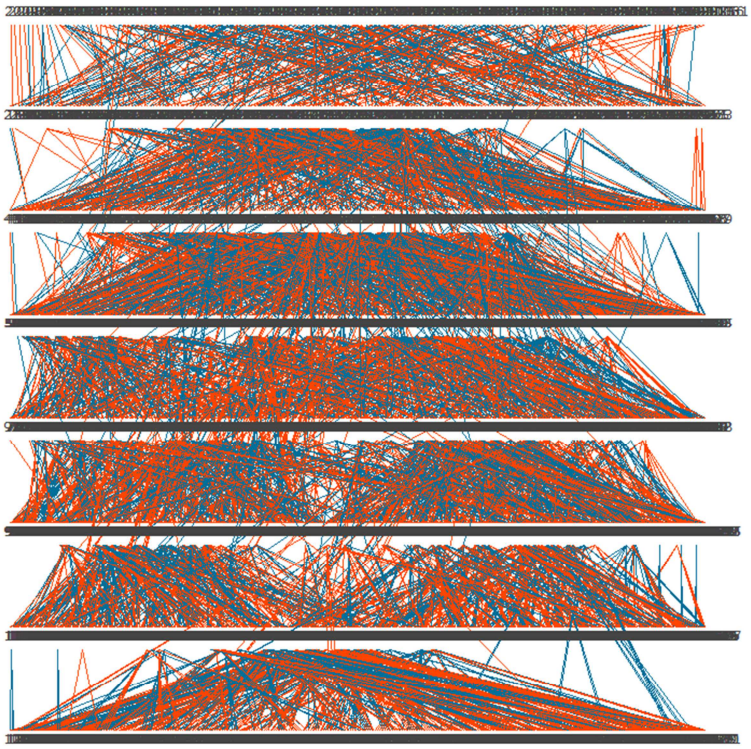


Figure S4: *Our empirical seven-generations pedigree of zebra finches held at the Max Planck Institute for Ornithology in Seewiesen, Germany. Red lines indicate maternal, blue lines paternal connections between parents and their offspring (pedigree drawn using Pedigree Viewer v6.5b; Kinghorn & Kinghorn 2010).*

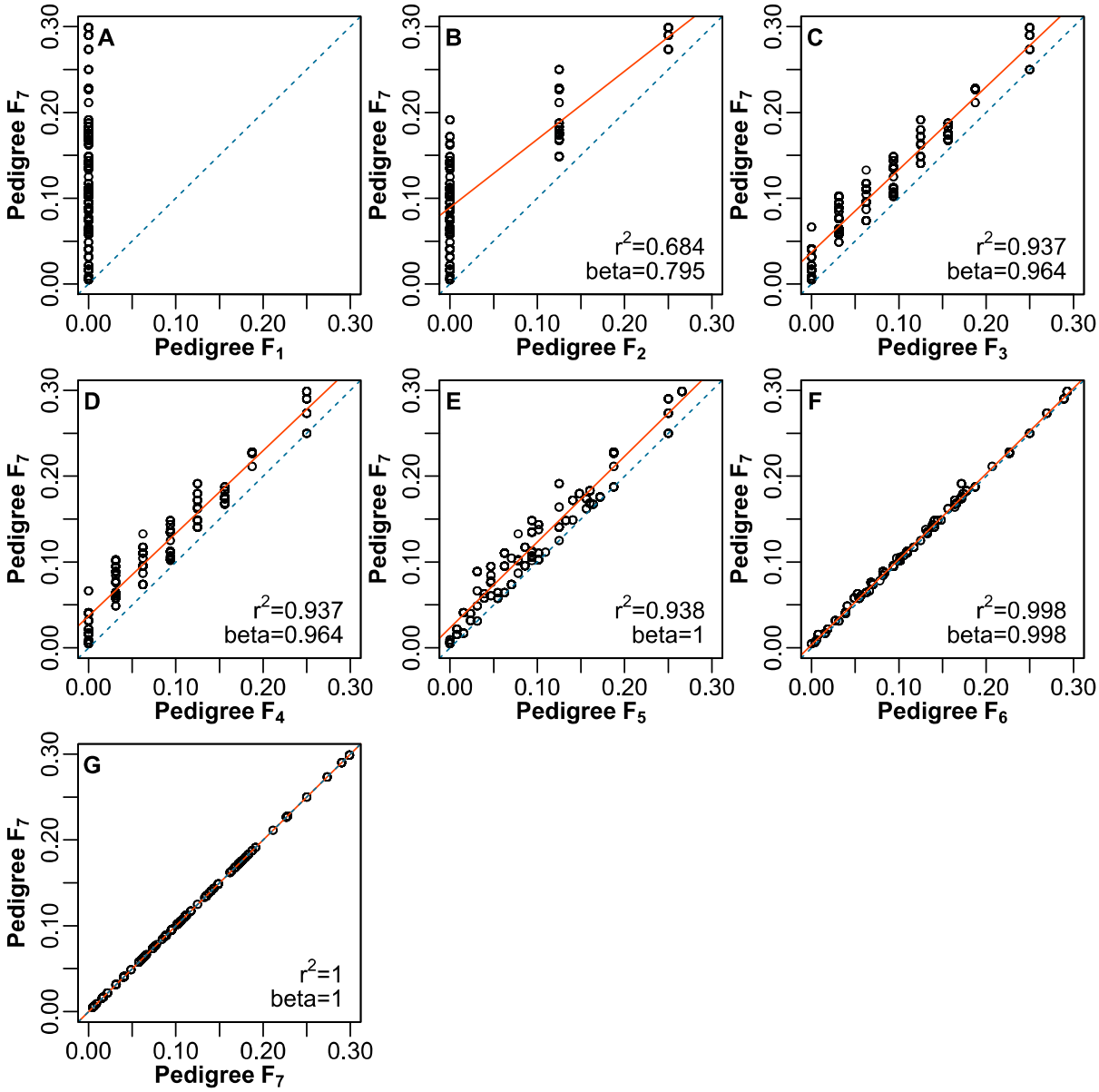


Figure S5: *Precision and accuracy of Pedigree F using increasingly remote generations as the founder generation. On the ordinate is the pedigree-based inbreeding estimate in the seventh generation using the full seven-generations pedigree and on the abscissa the pedigree-based inbreeding estimate using the (A) sixth generation, (B) fifth generation, (C) fourth generation, (D) third generation, (E) second generation, (F) first generation and (G) the original founding generation as founders. (C) and (D) are the same by coincidence. Dashed blue are the identity lines and solid red the OLS regression lines. Beta is the regression slope and r^2 the coefficient of determination.*

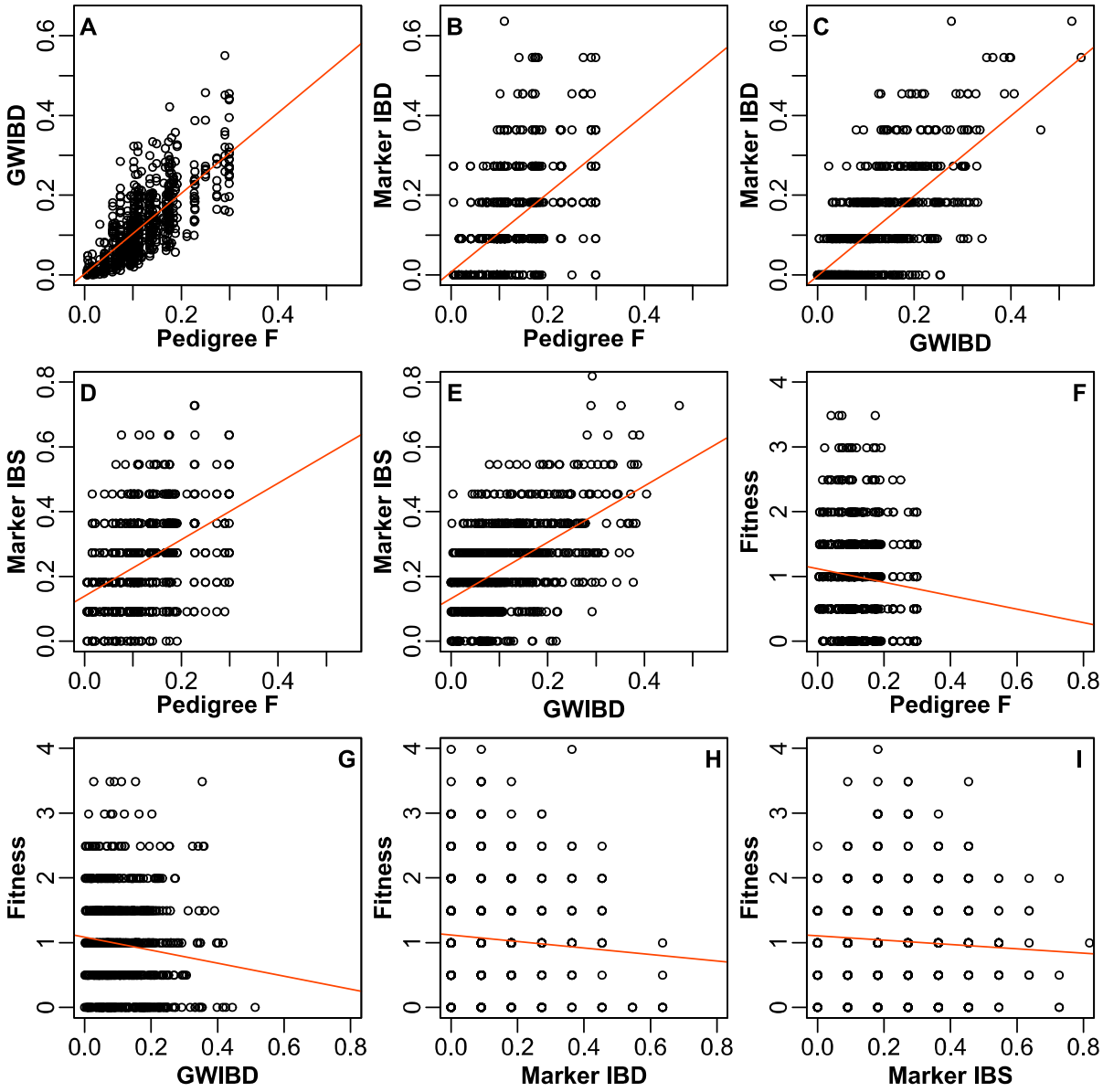


Figure S6: *A selection of exemplary simulations of zebra finch genomes closest to the mean regression slope and coefficient of determination in the seventh generation of our pedigree. Regressions of (A) GWIBD over Pedigree F, (B) Marker IBD over Pedigree F, (C) Marker IBD over GWIBD, (D) Marker IBS over Pedigree F, (E) Marker IBS over GWIBD, (F) Fitness over Pedigree F, (G) Fitness over GWIBD, (H) Fitness over Marker IBD, (I) Fitness over Marker IBS.*

Table S1: *Summary of the microsatellites and primers used for estimating the IBD-IBS discrepancy. All multiplex PCRs were done by using the Quiagen Type-it Microsatellite PCR Kit.*

Chromosome	Position (bp)	Primer Name	sequence	Fluorescence label	Annealing temperature (°C)	Number of cycles	Multiplex
Tgu1A	30,057,585	chr1A_39-F	GGCTCCTTAAAAGCCCAGCTC	NED	60	24	1
	30,057,585	chr1A_39-R	CTCTGCTGGACCCTCTCTAG		60	24	1
Tgu2	51,246,375	Tgu8-F*	GGGAGAGATAAAAGGTATTTTCAGG	6FAM	57	24	2
	51,246,375	Tgu8-R*	GAAAGGCATGGCAATAGTGAAG		57	24	2
Tgu3	58,288,063	chr3_58-F	CCTGATTCACCATGCCCAGT	PET	60	24	1
	58,288,063	chr3_58-R	AAAGGGCAGAAGGTAGACCATGA		60	24	1
Tgu5	34,270,862	chr5_34-F	GCAACTGCTGCTCTGAAGGA	PET	59	26	3
	34,270,862	chr5_34-R	AGCTGCACATGGGGAAGCTA		59	26	3
Tgu6	16,264,885	chr6_16-F	TCTGCCGTGTGTGTTTCTGG	VIC	59	26	3
	16,264,885	chr6_16-R	TAGCCATCTGGGCTCCTCAA		59	26	3
Tgu11	8,246,501	chr11_8-F	TTGCAGGCAGGTTCAGTGTG	NED	59	26	3
	8,246,501	chr11_8-R	TGGTTGCCTGGAGAAGATGG		59	26	3
Tgu14	8,898,802	chr14_9-F	GATGGAAAGGCTCTGGCACC	NED	60	24	4
	8,898,802	chr14_9-R	CTGAGTGGGTCGCAGGTGAT		60	24	4
Tgu15	6,303,408	chr15_6-F	AGCCGAGGGCCTAAAGATGA	6FAM	60	24	1
	6,303,408	chr15_6-R	GAGCCAGGATGAAAGGAGGT		60	24	1
Tgu22	2,509,697	chr22_3-F	TGGCCTTGCTGACTTCTGCT	VIC	60	24	1
	2,509,697	chr22_3-R	AGCAGGTTGTGAGGGCTTGT		60	24	1
Tgu26	2,541,304	chr26_3-F	GAAAGGACCTCTGGGCTCTG	6FAM	60	24	4
	2,541,304	chr26_3-R	AGCTTGCACCGTGAGGTAGC		60	24	4
Tgu27	1,136,439	chr27_1-F	GATCTGGAAATACCCTGGAGC	6FAM	59	26	3
	1,136,439	chr27_1-R	TGAAGCATTTCCCTCTGGAGTC		59	26	3

* Tgu8 was first described in Forstmeier *et al.* (2007).

Table S2: *Expectations and findings (estimate and 95% QR) for error components and slopes of OLS regressions between the different estimates of inbreeding and fitness. Marker based estimates of inbreeding (Marker IBD and IBS) were based on 11 markers with an IBD-IBS Discrepancy of 13.3%. α = OLS intercept, z_2 and z_{23} = bias introduced by random error in the predictors. Error 1 = Mendelian noise, error 2 = Marker sampling noise, error 3 = IBD-IBS Discrepancy, error 4 = heterogeneous distribution of recessive deleterious mutations and error 5 = additional environmental and genetic noise in fitness.*

Variables		Causality	Expectation			Simulation	
Y	**X**		**OLS slope**	**Error in Y**	**Error in X**	**Correlation (r)**	**Slope (β)**
GWIBD	Pedigree F	Regression	1	1		0.75 (0.70–0.79)	1.00 (0.89–1.11)
Marker IBD	Pedigree F	Regression	1	1, 2		0.52 (0.42–0.61)	0.99 (0.79–1.16)
Marker IBD	GWIBD	Regression	1	2		0.70 (0.59–0.82)	1.00 (0.78–1.23)
Marker IBS	Pedigree F	Regression	1-α	1, 2, 3		0.40 (0.30–0.46)	0.87 (0.66–1.03)
Marker IBS	GWIBD	Regression	1-α	2, 3		0.53 (0.44–0.64)	0.88 (0.66–1.07)
Fitness	Pedigree F	Regression	b	1, 4, 5		-0.09 (-0.16–-0.02)	-1.05 (-1.87–-0.23)
Fitness	GWIBD	Regression	b	4, 5		-0.12 (-0.18–-0.04)	-1.01 (-1.60–-0.39)
Fitness	Marker IBD	Correlation	$b-z_2$	4, 5	2	-0.082 (-0.15–-0.01)	-0.50 (-0.94–-0.08)
Fitness	Marker IBS	Correlation	$(b-z_{23})/(1-\alpha)$	4, 5	2, 3	-0.064 (-0.14–0.00)	-0.34 (-0.77–0.03)

Chapter 4: A prezygotic transmission distorter acting equally in female and male zebra finches *Taeniopygia guttata*

Abstract

The two parental alleles at a specific locus are usually inherited with equal probability to the offspring. However, at least three processes can lead to an apparent departure from fair segregation: early viability selection, biased gene conversion and various kinds of segregation distortion. Here, we conduct a genome-wide scan for transmission distortion in a captive population of zebra finches (Taeniopygia guttata) using 1,302 single-nucleotide polymorphisms (SNPs) followed by confirmatory analyses on independent samples from the same population. In the initial genome-wide scan, we found significant distortion at three linked loci on chromosome Tgu2 and we were able to replicate this finding in each of two follow-up data sets [overall transmission ratio = 0.567 (95% CI = 0.536–0.600), based on 1,101 informative meioses]. Although the driving allele was preferentially transmitted by both heterozygous females [ratio = 0.560 (95% CI = 0.519–0.603)] and heterozygous males [ratio = 0.575 (95% CI = 0.531–0.623)], we could rule out postzygotic viability selection and biased gene conversion as possible mechanisms. Early postzygotic viability selection is unlikely, because it would result in eggs with no visible embryo and hence no opportunity for genotyping, and we confirmed that both females and males heterozygous for the driving allele did not produce a larger proportion of such eggs than homozygous birds. Biased gene conversion is expected to be rather localized, while we could trace transmission distortion in haplotypes of several megabases in a recombination desert. Thus, we here report the rare case of a prezygotically active transmission distorter operating equally effectively in female and male meioses.

Published as: Knief U, H Schielzeth, H Ellegren, B Kempenaers, W Forstmeier 2015: A prezygotic transmission distorter acting equally in female and male zebra finches *Taeniopygia guttata*. Molecular Ecology 24: 3846–3859.

MOLECULAR ECOLOGY

Molecular Ecology (2015) **24**, 3846–3859 doi: 10.1111/mec.13281

A prezygotic transmission distorter acting equally in female and male zebra finches *Taeniopygia guttata*

ULRICH KNIEF,* HOLGER SCHIELZETH,† HANS ELLEGREN,‡ BART KEMPENAERS* and WOLFGANG FORSTMEIER*

*Department of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, Eberhard-Gwinner-Str., 82319 Seewiesen, Germany, †Department of Evolutionary Biology, Bielefeld University, Morgenbreede 45, 33615 Bielefeld, Germany, ‡Department of Evolutionary Biology, Uppsala University, Norbyvägen 18D, 752 36 Uppsala, Sweden

Abstract

The two parental alleles at a specific locus are usually inherited with equal probability to the offspring. However, at least three processes can lead to an apparent departure from fair segregation: early viability selection, biased gene conversion and various kinds of segregation distortion. Here, we conduct a genome-wide scan for transmission distortion in a captive population of zebra finches (*Taeniopygia guttata*) using 1302 single-nucleotide polymorphisms (SNPs) followed by confirmatory analyses on independent samples from the same population. In the initial genome-wide scan, we found significant distortion at three linked loci on chromosome *Tgu2* and we were able to replicate this finding in each of two follow-up data sets [overall transmission ratio = 0.567 (95% CI = 0.536–0.600), based on 1101 informative meioses]. Although the driving allele was preferentially transmitted by both heterozygous females [ratio = 0.560 (95% CI = 0.519–0.603)] and heterozygous males [ratio = 0.575 (95% CI = 0.531–0.623)], we could rule out postzygotic viability selection and biased gene conversion as possible mechanisms. Early postzygotic viability selection is unlikely, because it would result in eggs with no visible embryo and hence no opportunity for genotyping, and we confirmed that both females and males heterozygous for the driving allele did not produce a larger proportion of such eggs than homozygous birds. Biased gene conversion is expected to be rather localized, while we could trace transmission distortion in haplotypes of several megabases in a recombination desert. Thus, we here report the rare case of a prezygotically active transmission distorter operating equally effectively in female and male meioses.

Keywords: genic drive, meiotic drive, segregation distortion, selfish DNA

Received 18 March 2015; revision received 13 June 2015; accepted 17 June 2015

Introduction

According to Mendel's law of segregation, at each locus of a diploid organism, the two alleles are separated during gamete formation and are transmitted to the next generation with equal probability, that is each allele has a 50% chance of being inherited to each offspring (Mendel 1865). Although such Mendelian segregation is ubiquitously found in nature (Úbeda & Haig 2005), three distinct biological processes may lead to an apparent departure from fair transmission of alleles, which we collectively refer to as 'transmission distortion': (i) postzygotic viability selection, (ii) biased gene conversion and (iii) several kinds of prezygotic 'segregation distortion' (Burt & Trivers 2006; Duret & Galtier 2009; Meyer *et al.* 2012). In the following, we put aside other forms of bias by transposons and other 'jumping genes' that are not bound to a specific genomic location (Burt & Trivers 2006).

Correspondence: Ulrich Knief, Fax: +49 8157 932 400; E-mail: uknief@orn.mpg.de

1 Postzygotic viability selection can lead to transmission distortion when one of the alleles reduces the

viability of the progeny at an age prior to the assessment of transmission ratios (Meyer *et al.* 2012). Selection will lead to an overrepresentation of the beneficial allele at that locus in surviving individuals. Usually, both sexes contribute equally to this type of transmission distortion, because there is no universal reason for a deleterious or beneficial allele to have a differential effect when inherited from the mother or from the father (with the exception of imprinted genes, but these seem to be absent in the avian lineage, Frésard *et al.* 2014). Zöllner *et al.* (2004) suggested that viability selection is a common phenomenon, found across most chromosomes in the human genome, leading to full-siblings sharing more than the expected 50% of their genomes identical by descent. However, a large-scale study could not replicate the upward bias in allele sharing (Visscher *et al.* 2006).

2 Gene conversion is a process that occurs during meiotic recombination, where double-strand breaks can initiate crossing-over (Duret & Galtier 2009). In this process, information from an intact chromosome is used to repair a double-strand break in the homologous chromosome. Whenever the two chromosomes carry different alleles, this leads to a replacement of one allele by the other. Such gene conversion usually occurs from the uncut to the cut chromosome. Thus, if one chromosome happens to initiate double-strand breaks more often than the other, biased gene conversion may occur (Jeffreys & Neumann 2002). Alternatively, the repair enzymes may favour one allele over the other (usually GC over AT motifs; Birdsell 2002). Weber *et al.* (2014) suggested that biased gene conversion has shaped the GC content across the avian lineage, and GC content is also positively correlated with recombination rate in the zebra finch genome (Backström *et al.* 2010). Biased gene conversion should be equally common in males and in females, but at least in humans, there seems to be a male bias (Webster *et al.* 2005; Dreszer *et al.* 2007). Yet the effect of biased gene conversion is typically very localized, with about 50–1000 bp getting converted from the uncut to the cut strand in humans (Jeffreys & May 2004), and their locations are usually associated with recombination hotspots (Williams *et al.* 2015).

3 Segregation distortion refers to all cases in which the functional products of meiosis carry a specific allele at a given locus in more than 50% of the cases (Meyer *et al.* 2012). These deviations can result from (i) '(true) meiotic drive' (=chromosomal drive; Lyttle 1993) acting in the asymmetric meioses of females, or (ii) from 'postmeiotic competition among gametes' (=genic drive; Lyttle 1993) acting after the symmetric meioses in males, or from (iii) a variety of other prezygotic viability effects (Bernasconi *et al.* 2004).

Meiotic drive is acting primarily in females (Pardo-Manuel de Villena & Sapienza 2001; Meyer *et al.* 2012). Their asymmetric meiosis leads to the formation of one oocyte and three polar bodies from a single oogonium. The two homologous chromosomes are expected to compete for inclusion into the egg cell nucleus rather than ending up in the polar body, which represents an evolutionary dead end (Axelrod & Hamilton 1981). Length and sequence motif polymorphisms at or near centromeres are supposed to influence meiotic drive, because the microtubules of the spindle apparatus bind to the centromeres during meiosis (Henikoff *et al.* 2001). Chromosomes orient themselves before spindle attachment, and telomere-associated movement is proposed as a further process leading to meiotic drive (Axelsson *et al.* 2010; Tsai & McKee 2011). Four loci exhibiting meiotic drive in females were found in a genome-wide scan in chicken, and all were located close to centromeres or telomeres (Axelsson *et al.* 2010), but so far, none of these findings has been verified in follow-up studies.

In contrast, males show two symmetric cell divisions during spermatogenesis, which give rise to four functional spermatozoa from a single primary spermatocyte. Segregation distorters in males can act postmeiotically by disrupting or otherwise outperforming other sperm not carrying the driving allele (Pardo-Manuel de Villena & Sapienza 2001). The *t*-complex in mice (*Mus musculus*) and the *SD* locus in *Drosophila melanogaster* are well-known examples of this kind of drive (Lyttle 1991).

The few known examples of segregation distortion show strong effects, with more than 90% of all offspring from heterozygous parents carrying the driving allele (Lyttle 1993). If there is no compensating reduction in the fitness of the carriers, the driving locus would rapidly spread to fixation, and no stable polymorphism could evolve (Traulsen & Reed 2012). Due to the rapid fixation of drivers, such events should be difficult to detect, but stable polymorphisms can be maintained for long times if the driving allele carries a deleterious load. For example, male *D. melanogaster* homozygous for the driving allele at the *SD* locus have severely reduced fertility (Hartl 1969) and male mice homozygous for the driving allele at the *t*-complex are sterile (Lyon 2003). Female meiotic drive may not lead to a drop in the number of eggs produced but may lead to nondisjunction of homologous chromosomes in males that are heterozygous for the driving allele (Henikoff *et al.* 2001). Alternatively, nondisjunction may happen in females that are homozygous for the driving allele, resulting in aneuploidy and higher embryo mortality (Axelrod & Hamilton 1981; Day & Taylor 1998).

Results of tests for transmission distortion in several species from diverse phyla suggest that it is more common than previously thought (e.g. Fishman & Saunders 2008; Meyer *et al.* 2012). However, to our knowledge, so far explicit tests have been performed in only two bird species, the lesser kestrel (*Falco naumanni*; Aparicio *et al.* 2010) and domestic chicken (*Gallus gallus domesticus*; Axelsson *et al.* 2010). Ellegren *et al.* (2012) recently suggested that meiotic drive may have played an important role during speciation in flycatchers *Ficedula* sp. and therefore recommended testing for meiotic drive in pedigreed populations.

Here, we test for transmission distortion in the zebra finch (*Taeniopygia guttata*), a colonial-breeding passerine bird native to Australia that has been bred in captivity for many generations. It has become a model species in evolutionary biology with extensive genetic and genomic resources (e.g. Balakrishnan *et al.* 2010; Griffith & Buchanan 2010; Warren *et al.* 2010) and is thus well suited for this type of study. First, we make use of 1302 single-nucleotide polymorphisms (SNPs) that have been genotyped in a pedigree of 1057 individuals to screen the entire genome for any signs of transmission distortion. The initial scan suggested transmission distortion at two loci in the genome, but did not allow us to disentangle viability selection from biased gene conversion and segregation distortion, because we only genotyped birds that survived until maturity. Therefore, we conducted a large-scale follow-up analysis (follow-up 1) to examine whether viability selection was responsible for the transmission bias using an additional 1751 adult birds, 715 embryos that died during incubation and 985 embryos that were DNA sampled before hatching. To replicate and complement findings of follow-up study 1, we set up additional breeding pairs from birds heterozygous for the driving alleles and produced another 494 embryos that were mainly sampled before hatching (follow-up 2). This enabled us (i) to replicate the results of the initial scan and (ii) to distinguish between the different mechanisms leading to transmission distortion. If viability selection was acting, we expected the nondriving allele to be overrepresented among the embryos that died naturally during incubation. However, in the case of biased gene conversion, meiotic or genic drive, we expected transmission distortion in favour of the driving allele also in the dead embryos. The 985 embryos that were collected before hatching should show a preferential transmission of the driving allele in the case of biased gene conversion, meiotic or genic drive and fair Mendelian segregation in the case of viability selection.

Methods

Initial genome-wide scan

Study population. In a first step, we screened 1057 birds from 258 families for transmission distortion at 1302 SNPs (see below). These birds were taken from a four-generation pedigree of a captive population held at the Max Planck Institute for Ornithology in Seewiesen, Germany. We refer to these four successive generations as P, F1, F2 and F3. Recently, inbreeding coefficients in this population had been estimated with molecular markers and were rather low (estimated F = 0.064; Knief *et al.* 2015).

The birds used in this study originate from a population held at the University of Sheffield since 1985 (Forstmeier *et al.* 2004). European captive zebra finches are from the Australian (sub)species *Taeniopygia guttata castanotis*, which forms one large panmictic population across the Australian continent (Balakrishnan & Edwards 2009). Although the Lesser Sunda islands north of Australia harbour the (sub)species *T. g. guttata* (Timor zebra finch), and both subspecies interbreed successfully in captivity (Zann 1996), an admixture between the two subspecies prior to 1985 in captivity is unlikely because Timor zebra finches were only recently imported to Europe in small quantities and interbreeding is strongly discouraged among aviculturists.

SNP genotyping. One thousand and sixty-seven birds from the P, F1, F2 and F3 generations had previously been genotyped for 1920 SNPs using the Illumina GoldenGate assay (Fan *et al.* 2003) at the SNP Technology Platform of Uppsala University (Backström *et al.* 2010; Schielzeth *et al.* 2011, 2012). Thousand three-hundred and ninety-five of those SNPs were informative and contained no Mendelian errors (1358 autosomal, 37 on sex chromosome *TguZ*). They had been previously assembled into 32 linkage groups (Backström *et al.* 2010), which are equivalent to chromosomes *Tgu1–Tgu28* and *TguZ* (excluding *Tgu22*, because marker selection and genotyping was performed before the genome assembly was published; Warren *et al.* 2010) of genome built WUSTL v3.2.4. In total, SNPs covered about 92% of the annotated zebra finch genome (Backström *et al.* 2010). Ten individuals for which more than half of the SNPs could not be genotyped were removed from the current analysis. As transmission distortion tests are strongly affected by genotyping errors (Mitchell *et al.* 2003), we checked the missing call rate for each SNP. Genotyping failure might indicate problems with genotype scoring in general. Therefore, we excluded 93 SNPs with missing call rates >0.1 from further analyses (leaving 1265 autosomal SNPs + 37 SNPs on chromo-

some *TguZ* = 1302 SNPs). The genome scan based on the remaining 1302 SNPs for both sexes combined did not show signs of poor genotype calls judged by a log quantile-quantile *P*-value plot (e.g. Balding 2006).

Expected and realized inheritance. We used the method described in Axelsson *et al.* (2010), which makes use of the exact binomial test to assess whether there is transmission distortion at a specific locus. We combined information from all families in the analysis as families are usually small and power to detect transmission distortion at the family-level would thus be prohibitively low. Specifically, at every diallelic locus, we counted the number of offspring from two heterozygous parents (N_{Aa}), the number of offspring from a heterozygous parent that had a partner homozygous for the major allele A (N_{AA}) and the number of offspring from a heterozygous parent paired to a bird homozygous for the minor allele a (N_{aa}). We defined major and minor alleles based on allele frequencies across all generations. Furthermore, we define N_A for each locus as the sum of all major alleles inherited among the $N_{Aa} + N_{AA} + N_{aa}$ offspring. Offspring from a heterozygous parent that was paired to a partner homozygous for the major allele (in total N_{AA} offspring) always inherited one major allele from the homozygous parent. Among informative meioses, the realized number of inheritances of the major allele (n_A) is thus given by:

$$n_A = N_A - N_{AA}$$

The number of informative meioses (n) is then:

$$n = 2 \times N_{Aa} + N_{AA} + N_{aa}$$

In the absence of distortion, the major and minor allele are each inherited in half of the informative meioses, such that the expected number of inheritance events of the major allele [$E(n_A)$] and of the minor allele [$E(n_a)$] are equal:

$$E(n_A) = E(n_a) = n/2$$

The difference between $E(n_A)$ and n_A gives evidence for transmission distortion and can be tested against H_0: $E(n_A) = n_A$ by the exact binomial test.

Combined and sex-specific genome scans. Following Axelsson *et al.* (2010), we made the above calculations first considering both heterozygous parents at the same time ('combined-sexes genome scan') and second taking only heterozygous males or females (including cases where both parents are heterozygous). These latter sex-specific scans are still a mixture of informative meioses from both sexes, but with one sex contributing more than the other. We third also performed the sex-specific scans excluding cases where both parents were heterozygous, but the power in these scans was reduced due to a limited number of informative meioses and the results are therefore only shown in the Supplement.

If transmission distortion resulted from either viability selection or biased gene conversion, we expected the effect to show up in all of the genome scans. In the case of true meiotic drive, the effect should show up in the female-specific scan and to a lesser extent in the combined-sexes scan. In the case of genic drive, the effect should show up in the male-specific scan and to a lesser extend in the combined-sexes scan.

We tested the alleles on chromosome *TguZ* only against other *TguZ* alleles in males (which is the homogametic sex in birds), because there is no evidence for sex-ratio distortion in our population.

To correct for multiple testing, we used strict Bonferroni correction (Foulkes 2009) such that *P*-values smaller than 3.8×10^{-5} (=0.05/1302) in male-specific tests and *P*-values smaller than 4.0×10^{-5} (=0.05/1265, excluding SNPs on chromosome *TguZ*) in female-specific and combined-sexes tests were considered significant. The Bonferroni correction was also used to adjust the confidence intervals (CI), such that adjusted 95% CIs in female-specific and in the combined-sexes scans were calculated for a confidence level of $1 - (0.05/1265)$ and in the male-specific scan for a confidence level of $1 - (0.05/1302)$. The Bonferroni adjustment for multiple comparisons is conservative in the context of this study because SNPs are in linkage disequilibrium with each other (Foulkes 2009).

Tests for weak genome-wide distortion. Transmission distortion could be weak but widespread in the genome, for example as a consequence of viability selection. Making no assumptions about whether major or minor alleles are generally transmitted more frequently, if many loci exhibited weak transmission distortion, then we would predict that the standard deviation around the 50% transmission rate would be higher than expected from sampling noise. To test this prediction, we simulated 10 000 data sets containing the same number of SNPs with the same number of informative meioses per SNP as in the empirical data set. The number of transmissions was sampled from a binomial distribution with Mendelian segregation. Linkage between SNPs on the same chromosome could potentially skew the distribution of transmission ratios. To account for this nonindependence, we sampled from each of these data sets only one SNP per chromosome and calculated the standard deviation of the transmission ratios of these SNPs. We then sampled 10 000 times from the empirical data set one SNP from each chromosome and calculated the standard deviation of the transmission ratios. The distributions of standard deviations were right-skewed in both the simulated and

the empirical data sets. Thus, we checked where the median of the empirical data set was located in the simulated data set and considered a position in the top 5% of the simulated values as significant.

All statistical analyses were performed in R 3.0.2 (R Core Team 2013). Power analyses were performed in R 3.0.2 using the pwr.p.test() function from the PWR-package (version 1.1.1; Champely 2012).

Follow-up 1: confirmation in embryos and adults

The genome-wide scan identified two genomic regions that deviated from random segregation, so we designed a follow-up study to first verify those cases and second to distinguish between viability selection, biased gene conversion and segregation distortion.

Study population. To replicate significant effects from the initial genome-wide scan, we used 1392 additional birds from 315 families of the next two generations (labelled F4 and F5) of the same population. We also included 359 birds from the P, F1, F2 and F3 generations that had not previously been genotyped for the SNPs. In addition, we genotyped 715 embryos from the F2, F3, F4 and F5 generations that died naturally at any age during incubation and 985 embryos from the same four generations that were either collected for DNA analysis before hatching (typically four days after onset of incubation, $n = 944$) or that died because the eggs were broken by their parents or while handling ($n = 41$).

Microsatellite genotyping. In the follow-up studies, we used microsatellites instead of SNPs to track the segregation of alleles because this could be performed more conveniently in our own laboratory. We used microsatellite alleles to impute the genotype of the two SNPs showing the strongest segregation distortion in the genome-wide scan (one locus on chromosome *Tgu5* and one on *Tgu2*), as described below. We designed primers for two microsatellites that were physically close to the significant locus on chromosome *Tgu5* (Tgu5_SD and Tgu5_SD4) and two microsatellites on chromosome *Tgu2* (Tgu2_SD44 and Tgu2_SD60; see Table S1, Supporting information for exact location and primer sequences). In addition, we used information from microsatellite Tgu8 on chromosome *Tgu2* (Forstmeier *et al.* 2007). Genotyping was carried out as described in Forstmeier *et al.* (2007).

For the locus on chromosome *Tgu2*, we later confirmed the reliability of our imputation by SNP genotyping a subset of 2511 individuals (including all 744 individuals that served as parents) within a larger genotyping project using the Sequenom MassARRAY iPLEX platform (Gabriel *et al.* 2009). This showed that only 7 individuals (0.3%) had been imputed incorrectly (four samples were imputed to be heterozygous although they were homozygous for the major allele and three samples *vice versa*). We consider these errors unproblematic, as they should not introduce any larger systematic error into our estimates of segregation ratios.

We genotyped all individuals used in this study ($n = 1057 + 359 + 1392 + 715 + 985 = 4508$) for all five microsatellite markers. The two microsatellites on chromosome *Tgu5* indicated that for this region, there were at least 17 different haplotypes segregating in our population, 10 of which contained the major allele C at the significant SNP (rs82439270) and the other 7 haplotypes carried the minor allele T. To avoid problems of multiple testing, we a priori decided to analyse only those two groups of haplotypes. The microsatellites on *Tgu2* identified a total of 12 haplotypes, 10 of which contained the major allele C at the significant SNP (rs82439854) and the other 2 haplotypes carried the minor allele G. Again, we present only the analyses contrasting those two groups of haplotypes.

Follow-up 2: selective breeding of the driving allele on chromosome Tgu2

For locus rs82439854 on chromosome *Tgu2*, a second follow-up study was necessary, because the low allele frequency of the undertransmitted allele meant that the number of informative meioses was low, rendering follow-up study 1 inconclusive.

Study population and microsatellite genotyping. We estimated the combined effect of the driving locus on chromosome *Tgu2* from the initial genome-wide scan and the follow-up study 1 to be close to 0.57. To design follow-up study 2, we performed a power analysis to determine that the number of informative meioses needed to reach at least 80% power was $n = 313$. At the time of setting up the experiment, only 20 individuals (12 females, 8 males) were alive that carried a copy of the undertransmitted allele G at SNP rs82439854. These individuals were given partners that were homozygous for the winning allele C. Of those 20 pairs, 16 produced eggs yielding information on segregation distortion for 10 heterozygous females and 6 heterozygous males in a total of 494 informative meioses. Eggs were mainly collected at day four of incubation but some ($n = 55$) were allowed to hatch, to maintain the under-represented allele in our population. Microsatellite genotyping and analyses were performed in the same way as in follow-up study 1.

Further statistical analyses. To get a combined estimate of the initial genome-wide study and the follow-up studies, it is misleading to just sum up all informative meioses as the point estimates may be biased upwards in the initial

genome scan due to selective focus on the most significant loci. Thus, we used the meta.summaries() function of the R package rmeta (version 2.16; Lumley 2012) to combine the estimates from the initial genome-wide scans (with Bonferroni corrected CIs) with the estimates from the follow-up 1 study (all birds that hatched, embryos that died during development and embryos collected from incubated eggs pooled together) and the follow-up study 2 (for chromosome *Tgu2*).

Generalized linear mixed models were fitted using the glmer() function of the R package lme4 (version 1.1-6; Bates *et al.* 2014). All covariates were mean-centred and scaled by twice their standard deviation to aid interpretability (Schielzeth 2010).

Results

Initial genome scan

The power analysis for the genome scan with combined sexes suggested good power for detecting distortion down to a transmission rate of the driving allele of 0.7 (80% power at 145 informative meioses which was reached by 89% of all SNPs; Fig. 1). In the sex-specific scans (including pairs in which both parents are heterozygous), power was almost as high (80% power down to a transmission rate of 0.7 was reached by 79% and 80% of all SNPs for the female and male-specific tests, respectively). Power diminishes rapidly below transmission ratios of 0.7, such that only 44% of the SNPs in the combined-sexes scan and only 18% and 17% of the SNPs in the female and male-specific scans could detect a transmission distortion of 0.6 with 80% power. Excluding pairs in which both parents are heterozygous in the sex-specific scans resulted in a severe drop in power (Fig. S1, Supporting information).

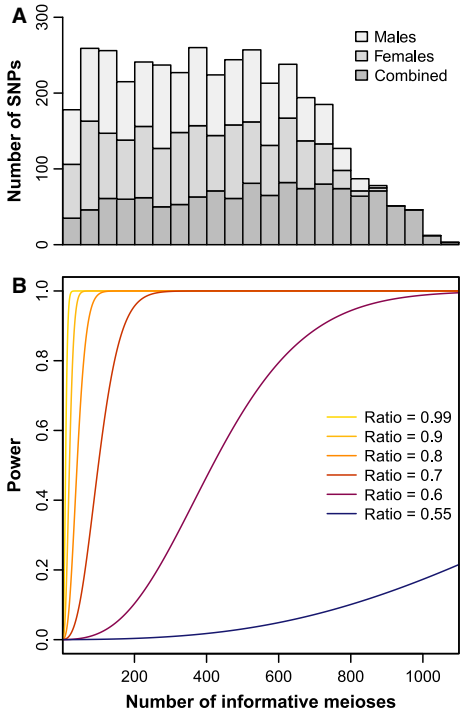


Fig. 1 (A) Histogram of the number of informative meioses available for each SNP in the initial whole-genome scans and (B) power analysis for the number of informative meioses for different transmission ratios considering a *P*-value of 4.0×10^{-5} as significant.

The initial genome scan for transmission distortion in both sexes combined yielded one SNP on chromosome *Tgu5* (rs82439270) that was significant after Bonferroni correction (Figs 2A and S2A, Supporting information, Table 1). At this locus, the major allele was preferentially transmitted (minor allele frequency, MAF = 0.202).

The genome scan for females yielded three SNPs that were significant after Bonferroni correction (Figs 2B and S2B, Supporting information, Table 1). All three markers were located within a 1.22 Mb region on chromosome *Tgu2* (rs82439854, rs82439741, rs82478694) and were in strong linkage disequilibrium with each other. A fourth SNP close to the other three markers on chromosome *Tgu2* (rs82478683) almost reached the significance threshold (Table 1). At all significant loci, the major allele was preferentially transmitted (for MAFs see Table 1). The genome scan that excluded those pairs in which both parents were heterozygous did not show any significant transmission distortion, presumably due to reduced power (Fig. S3A, Supporting information).

The genome scan in males did not show significant transmission distortion after Bonferroni correction (Figs 2C and Fig. S2C, Supporting information). Excluding pairs in which both parents were heterozygous also gave no significant transmission distortion (Fig. S3B, Supporting information).

At a genome-wide scale transmission was fair and simulations showed no indications of any excess variance beyond what is expected from Mendelian segregation [empirical vs. simulated medians of SD of transmission ratios: sexes combined 0.031 vs. 0.027 (95% quantile range (QR) = 0.019–0.047), $P = 0.26$; female-specific 0.035 vs. 0.036 (95% QR = 0.024–0.093), $P = 0.52$; male-specific 0.041 vs. 0.036 (95% QR = 0.023–0.094), $P = 0.29$].

Follow-up 1

The first follow-up study focused on the two most strongly driving SNPs found in the genome-wide scan, namely rs82439270 on chromosome *Tgu5* and rs82439854 on chromosome *Tgu2* (all four SNPs showing transmission distortion on chromosome *Tgu2* seemed to be linked to the same driving haplotype).

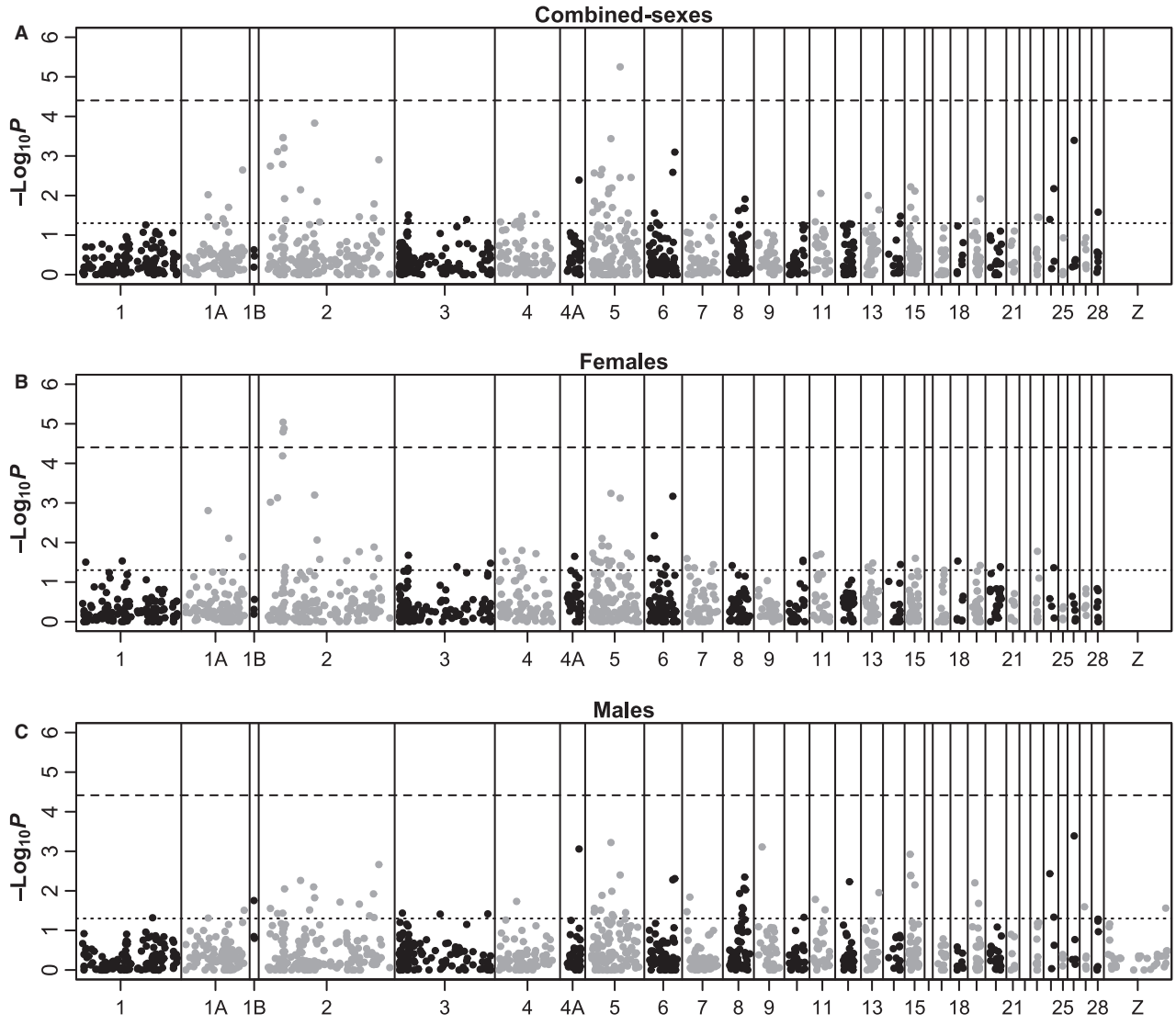


Fig. 2 Whole-genome scans for transmission distortion in (A) both heterozygous sexes combined, (B) heterozygous female parents only and (C) heterozygous male parents only. On the abscissa are the chromosomes (missing chromosome names are shown in Figure S3) and on the ordinate are the $-\log_{10}$ (*P*-values). The dashed line indicates the significance threshold after genome-wide Bonferroni correction and the dotted line before genome-wide Bonferroni correction. Chromosome *TguZ* was tested only in males because only males carry two copies of *TguZ*.

SNP rs82439270 on chromosome *Tgu5* did not show significant transmission distortion in any of the samples in the follow-up data set, neither when both heterozygous parents were combined (birds that hatched: 509 major allele vs. 520 minor allele transmissions, ratio = 0.495, $P = 0.76$; dead embryos: 186 major allele vs. 181 minor allele transmissions, ratio = 0.507, $P = 0.83$; collected embryos: 315 major allele vs. 279 minor allele transmissions, ratio = 0.530, $P = 0.15$; Fig. 3A) nor when female and male parents were analysed separately (Table S2, Supporting information). When combining all samples from the follow-up data set, there was no significant transmission distortion either (1010 major allele vs. 980 minor allele transmissions, ratio = 0.508, $P = 0.52$; Fig. 3A). Finally, also all samples from the initial genome-wide scan and the follow-up study combined showed no significant deviation from fair Mendelian segregation [1310 major allele vs. 1178 minor allele transmissions, ratio = 0.514 (95% CI of the transmission ratio = 0.493–0.536); Fig. 3A].

In contrast, SNP rs82439854 on chromosome *Tgu2* showed significant transmission distortion in favour of the major allele (252 major allele vs. 195 minor allele transmissions, ratio = 0.564, $P = 0.008$) when analysing

Table 1 Significant and almost significant ($P_{after} < 0.1$) SNPs of the initial whole-genome scan for both sexes combined and for heterozygous females only. 95% CI_{before} and P_{before} give 95% confidence intervals and *P*-values of a binomial test before Bonferroni correction, 95% CI_{after} and P_{after} after Bonferroni correction. MAF is the minor allele frequency and was calculated from the SNP data of the initial scan. $E(n_A)$ and n_A are the expected and observed numbers of inheritance events of the major allele, respectively, and were calculated as described in the Methods

Scan	SNP	Position	Major allele	Minor allele	MAF	Informative meioses	$E(n_A)$	n_A	Transmission ratio	95% CI_{before}	95% CI_{after}	P_{before}	P_{after}
Combined	rs82439270	*Tgu5*: 37 818 654	C	T	0.202	498	249	300	0.602	0.558, 0.646	0.510, 0.690	5.60×10^{-6}	0.0071
Female	rs82439854	*Tgu2*: 24 794 801	C	G	0.042	106	53	76	0.717	0.621, 0.800	0.516, 0.871	9.10×10^{-6}	0.012
	rs82478694	*Tgu2*: 26 015 177	C	T	0.043	105	52.5	75	0.714	0.618, 0.798	0.512, 0.870	1.32×10^{-5}	0.017
	rs82439741	*Tgu2*: 24 794 975	A	T	0.042	107	53.5	76	0.710	0.615, 0.794	0.510, 0.866	1.60×10^{-5}	0.020
	rs82478683	*Tgu2*: 24 218 493	G	A	0.042	93	46.5	66	0.710	0.606, 0.799	0.494, 0.874	6.47×10^{-5}	0.082

all samples in the follow-up data set combined (Fig. 3B). This drive was significant in males (ratio = 0.580, $P = 0.021$) and nonsignificant in females (ratio = 0.538, $P = 0.22$), but overall, the number of informative meioses was low due to the high major allele frequency of the driving allele (Tables 1, 2, and S3, Supporting information). Overall, these findings regarding SNP rs82439854 on chromosome *Tgu2* were inconclusive with respect to the mechanism of transmission distortion because the initial genome-wide scan suggested transmission bias in heterozygous females while the follow-up study 1 showed a transmission bias in heterozygous males. We therefore performed another replicate study (follow-up study 2) to clarify this issue for this locus.

Follow-up 2

The combined-sexes analysis of follow-up study 2 yielded the same effect as the combined-sexes analysis in follow-up study 1 (278 major allele vs. 216 minor allele transmissions, ratio = 0.563, $P = 0.0060$). This effect was about equally strong in females (ratio = 0.557, $P = 0.073$) as in males (ratio = 0.569, $P = 0.042$; Table S3, Supporting information).

Summary across data sets

A meta-analytic summary of transmission ratios from the initial genome-wide scan (with Bonferroni corrected CIs), follow-up study 1 and follow-up study 2 for locus rs82439854 on chromosome *Tgu2* yielded a weighted transmission bias of 0.567 for the two sexes combined [633 major allele vs. 468 minor allele transmissions (95% CI of the transmission ratio = 0.536–0.600); Fig. 3B]. Note that, this value is more conservative than the observed raw average (ratio = 0.575, $P = 7 \times 10^{-7}$; Table S3, Supporting information), which could still be somewhat inflated by a detection bias in the initial scan. When considering the sexes separately, meta-analytic summaries of transmission ratios gave very similar values for females [ratio = 0.560 (95% CI = 0.519–0.603)] and males [ratio = 0.575 (95% CI = 0.531–0.623)].

Ruling out postzygotic viability selection

As we found the preferential inheritance of the major allele of SNP rs82439854 in all subsamples of the follow-up studies (including dead embryos and embryos sampled before hatching), viability selection as a mechanism for the transmission bias seems unlikely. However, there is also a considerable fraction of eggs (about 30% of all eggs laid) that show no visible embryo development and that we therefore judge as

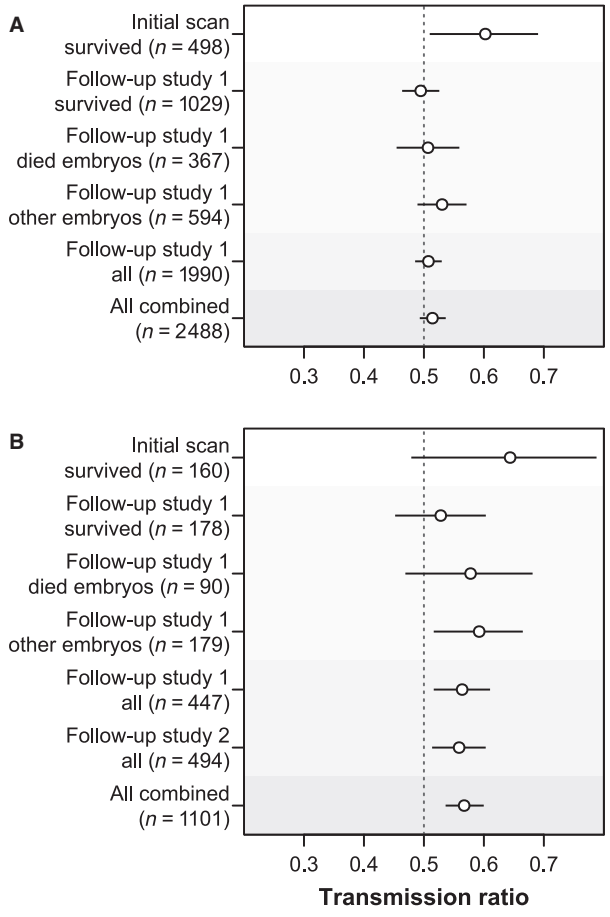


Fig. 3 Transmission ratios of the major allele ±95% CIs for the initial genome scans (combined sexes, initial scan) and their follow-ups (combined sexes, follow-up study 1 and follow-up study 2) for the regions showing transmission distortion in the initial genome scan on chromosomes (A) *Tgu5* and (B) *Tgu2*. For the initial scan, Bonferroni corrected CIs are shown. For the combined analysis of the initial genome scan with the follow-up studies 1 and 2, a weighted average is shown. For each group, *n* refers to the number of informative meioses.

infertile. Problematically, this pool of apparently infertile eggs may also comprise at least some eggs containing zygotes that died before any visible embryo development. As these eggs could not be sampled for DNA, they could potentially contain the missing fraction of the minor allele.

To test for this possibility, we fitted binomial generalized linear mixed models using 11 617 eggs that had been laid by single pairs in isolated cages as the binary response variable (fertile vs. infertile, scored as the presence vs. absence of a visible embryo) and the genotypes (heterozygous for the driving allele vs. homozygous) of each of the parents as two predictors. We controlled for the age and the pedigree-based inbreeding coefficient of the parents, their pairing duration at egg laying and the laying order by fitting appropriate predictors as covariates, and we added the clutch, pair, mother, father and experiment identity as random effects (covariates: all $P < 0.01$, Table 3). Neither the mother's nor the father's genotype was significantly associated with the fraction of infertile eggs with a trend in the wrong direction (Table 3). Pairs in which both partners were homozygous produced undeveloped eggs at a rate of 31.64% (based on 9339 eggs). Given this baseline, we would have expected heterozygous females and males to produce undeveloped eggs at a rate of 39.72% if those would contain in addition the missing fraction of the undertransmitted minor allele (31.64% baseline infertile, 8.08% undetected minor alleles, 26.10% detected minor alleles and 34.18% detected major alleles). However, heterozygous females paired to homozygous males laid only 25.75% of undeveloped eggs (based on 1231 eggs) and heterozygous males paired to homozygous females produced only 23.88% undeveloped eggs (based on 917 eggs). This indicates that the missing fraction of the minor allele is not contained in the (apparently) infertile eggs and argues against postzygotic viability selection as the cause of the transmission bias.

Discussion

After performing an initial genome-wide scan for transmission distortion and two follow-up studies, one locus on chromosome *Tgu2* showed a consistent preferential inheritance of 56.7% of a particular allele from both heterozygous females and males to their offspring. Although both sexes seemed to contribute equally to the transmission distortion, we could rule out postzygotic viability selection as the causal mechanism because distortion was equally strong in embryos that died during development and because heterozygous individuals did not produce a larger proportion of undeveloped eggs than homozygous birds. We detected the transmission bias with SNPs and with microsatellites that were separated by several megabases, which excludes biased gene conversion as a causal mechanism. The fact that the signal was seen with several different markers also strongly argues against genotyping errors as causing the distortion.

The initial genome-wide scan using both sexes combined gave one significant locus on chromosome *Tgu5* after Bonferroni correction ($P = 0.007$), at which the major allele was preferentially transmitted. However, the four times larger follow-up study showed fair segregation at this locus, leaving three possible explanations for the discrepancy. (i) The drive disappeared due to an interaction between genotype and an environmental factor that changed between early and later generations

Table 2 Transmission ratio of the major allele at locus rs82439854 on chromosome *Tgu2* in follow-up study 1 for both heterozygous females and heterozygous males separately (including cases in which both parents are heterozygous). The follow-up data set consisted of (i) a replicate of the initial genome scan which contained only birds that hatched (survived), (ii) embryos that died naturally during incubation (died embryos) and (iii) embryos that were collected for DNA analysis before hatching (other embryos). The observed number of inheritance events of the major allele is shown under n_A

Sex	Sample	Informative meioses	n_A	Transmission ratio	95% CI	*P*-value
Females	Survived	121	59	0.488	0.396, 0.580	0.86
	Died embryos	60	33	0.550	0.416, 0.679	0.52
	Other embryos	109	64	0.587	0.489, 0.681	0.084
	All	290	156	0.538	0.479, 0.596	0.22
Males	Survived	85	48	0.565	0.453, 0.672	0.28
	Died embryos	42	24	0.571	0.410, 0.723	0.44
	Other embryos	92	55	0.598	0.490, 0.699	0.076
	All	219	127	0.580	0.512, 0.646	0.021

Table 3 Parameter estimates of a binomial generalized linear mixed model using fertility status (based on visible embryo development) of 11 617 eggs as the binary response variable (fertile = 0, infertile = 1; with 30.22% being infertile). All eggs were laid by pairs in individual cages. We fitted the clutch [2783 levels, variance component = 3.50 (95% CI = 2.98–4.10)], pair [875 levels, variance component = 2.07 (95% CI = 1.24–3.09)], mother [440 levels, variance component = 0.44 (95% CI = 0–1.14)], father [436 levels, variance component = 2.44 (95% CI = 1.65–3.36)] and experiment [21 levels, variance component = 0.078 (95% CI = 0–0.35)] as random effects. Covariates were mean-centred and scaled by twice their standard deviation. Parameter estimates are presented on the logit scale, and confidence intervals were obtained by the profile likelihood method. F_{Ped} = pedigree-based inbreeding coefficient

Parameter	Estimate	95% CI	*P*-value
Intercept	−1.55	−1.89, −1.25	$<2 \times 10^{-16}$
Mother genotype (1=heterozygous)	−0.072	−0.70, 0.56	0.82
Father genotype (1=heterozygous)	−0.34	−1.23, 0.56	0.46
F_{Ped} of mother	0.88	0.49, 1.27	1×10^{-5}
F_{Ped} of father	0.89	0.43, 1.36	2×10^{-4}
Mother age	0.66	0.16, 1.16	1×10^{-2}
Father age	0.89	0.37, 1.42	8×10^{-4}
Pairing duration	−0.44	−0.70, −0.18	8×10^{-4}
Laying position	−0.32	−0.45, −0.18	3×10^{-6}

(generations P–F3 were used for the initial scan, while most samples in the follow-up were from generations F4–F5). However, in those individuals from the P–F3 generation that were missed in the initial scan and were genotyped in the follow-up study, the minor allele rather than the major allele was preferentially transmitted (22 major allele vs. 36 minor allele transmissions, ratio = 0.379, $P = 0.087$), casting doubt on such a genotype by environment interaction. (ii) The SNP rs82439270 may exhibit biased gene conversion, which was not captured when imputing the genotype at this location from surrounding microsatellite markers. Genomic regions prone to gene conversion typically colocalize with regions of high recombination rate (Williams *et al.* 2015), but the cumulative genetic map length in the region surrounding SNP rs82439270 is not increasing markedly (flanking markers rs82439433 and rs82480084 lie at 18.6 cM/37 609 078 bp and 19 cM/38 292 102 bp, respectively), meaning that there is no strong recombination hotspot in between them and that SNP rs82439270 resides within a recombination desert (maximally 0.4 cM; see also Backström *et al.* 2010). (iii) The initial finding may have been a type 1 error, and from the above reasoning, we consider this the most likely explanation.

After Bonferroni correction, the genome-wide scan in females showed significant preferential transmission of the major allele at three loci on chromosome *Tgu2*. Thus, at first, a meiotic drive system in which the two homologous chromosomes compete for inclusion into the egg cell nucleus appeared the most likely mechanism. However, the significant loci on chromosome *Tgu2* are around 50 Mb away from the centromere and almost 25 Mb away from the nearest telomere, which is where the main driving loci in case of meiotic drive are expected (Henikoff *et al.* 2001; Axelsson *et al.* 2010; Tsai & McKee 2011). Indeed, in follow-up study 1, predictions for a meiotic drive system were not fulfilled, because the transmission bias was equally strong in females and males and apparent in both embryos and surviving birds. Follow-up study 2 confirmed this finding. This leads to four conclusions.

1 The mechanism leading to the transmission bias is acting both in females and males.
2 We can rule out viability selection as the underlying mechanism because the driving allele is also

over-represented in embryos that died during development. We can also exclude very early (post-zygotic) embryo mortality, because heterozygous individuals did not produce a larger proportion of (apparently) infertile eggs than homozygous individuals. The high number of apparently infertile eggs (30%) that cannot be analysed may appear somewhat worrying in this context. We believe that most of these eggs are truly infertile due to birds failing to copulate. First, 'infertile' eggs are most frequent (around 30%) when single pairs breed in isolated cages, and a considerable fraction produced several repeated clutches where none of the eggs showed any sign of fertility. 'Infertile' eggs are less frequent (10–20%) when pairs breed in communal aviaries (W. Forstmeier, unpublished data). Second, unpublished data from extensive video recording of pairs breeding in such communal aviaries shows that the probability of laying 'infertile' eggs strongly declines with the number of copulations performed.

3 Biased gene conversion is unlikely to have such a strong effect at this genomic location, because in the zebra finch, recombination rates drop rapidly with increasing distance from the telomeres (Backström *et al.* 2010) and the driving SNP is located more than 24 Mb from each chromosome end. Thus, although the linkage map for chromosome *Tgu2* was difficult to build, the driving SNP was certainly located in a recombination desert with flanking markers rs82439650 and rs82438778 at 58 cM/24 057 156 bp and 58.1 cM/26 288 644 bp, respectively (Backström *et al.* 2010). Furthermore, the effects of biased gene conversion are typically localized (Jeffreys & May 2004; Duret & Galtier 2009), but we observed the transmission distortion both at several linked SNPs distributed over more than 1 Mb and at microsatellites several Mb away from the SNPs. Thus, we can exclude biased gene conversion as the underlying mechanism leading to the transmission distortion.

4 Consequently, the transmission bias on chromosome *Tgu2* must arise from a transmission distorter, which is acting prezygotically in females and males with similar efficiency. An argument against prezygotic selection may be that females heterozygous for the driving allele do not show longer laying gaps between consecutive eggs within a clutch than homozygous females (based on 10 688 measures of laying intervals fitted as the response variable in a Poisson generalized linear mixed model, data not shown), which may be expected in case of haploid selection acting against the undertransmitted allele.

In crosses between differentiated populations or subspecies segregation distortion is often observed either as a result of Dobzhansky–Muller incompatibilities, leading to viability selection, or because of the reactivation of segregation distorters that were inactivated within the founding populations by suppressor genes (Presgraves 2010). Segregation distorters could also have evolved independently on the same chromosome in the parental populations, once in males and once in females, and only appear to be a single segregation distorter acting in both sexes in the hybrid population (yet under this scenario direction and strength of the distortion do not need to be the same in the two sexes). However, in our population of zebra finches, both mechanisms would lead to between-individual differences in the strength of transmission distortion because they require interacting genotypes (Presgraves 2010), and we do not have a defined genetic background. Thus, in some individuals, the segregation distorter would be associated with the unlinked suppressor and not active, while in other individuals, the segregation distorter would be in a naïve genetic background showing transmission bias. However, there is no between-individual variation in the strength of segregation distortion in males and only very little in females (analyses not shown). Also, because Australian zebra finches form one large panmictic population (Balakrishnan & Edwards 2009), it seems rather unlikely that two differentiated populations mixed at some point in the recent past.

A cytogenetic mechanism to explain the observed patterns of transmission bias is not obvious, because both gametes and meioses are fundamentally different in males and females. In the common shrew (*Sorex araneus*), metacentric chromosomes resulting from a Robertsonian translocation were found preferentially transmitted over their acrocentric homologues in both male and female heterozygous parents (Wyttenbach *et al.* 1998; Fedyk & Chetnicki 2007). The effect was more pronounced in males, which showed a transmission distortion of around 0.60–0.71 in favour of the metacentric chromosomes.

Paracentric inversions are another type of structural variant that could potentially bias the segregation of alleles in meiosis in both oogenesis and spermatogenesis. A crossover within such an inversion in a heterozygous carrier leads to the formation of a dicentric bridge and an acentric chromosomal fragment (reviewed in Burnham 1962). Meiotic outcomes in the gametes differ between inversions and between species (Koehler *et al.* 2002) and could potentially lead to segregation distortion by selectively eliminating or arresting one of the abnormal meiotic products.

Whether any such structural variant exists in our population is unknown and needs further cytogenetic research. The linkage map for chromosome *Tgu2* in our population showed discrepancies in gene order to the

zebra finch genome assembly (Backström *et al.* 2010) and patterns of linkage disequilibrium are consistent with the idea that there is a structural polymorphism for chromosome *Tgu2* in our population.

Conclusion and outlook

We here describe a segregation distortion allele which is passed on in 56.7% of the meioses from both heterozygous female and male zebra finches. A structural variant most likely contributes to the transmission bias and future cytogenetic work should clarify the nature of any such structural polymorphism and how it leads to transmission distortion, while additional genomic work should reveal the evolutionary history of the competing alleles. In our population, at least 13 of the 153 pedigree founders carried the undertransmitted allele, demonstrating that it is not a particularly recent *de novo* mutation within our population. In line with this, the undertransmitted allele is associated with microsatellite alleles that identify two different variants of the haplotype. In the absence of fitness benefits, the undertransmitted allele would have gone rapidly to extinction, so it may be interesting to study the fitness consequences for individuals that carry these alleles.

Until now, localized transmission distortion has been searched for in three bird species and was found in two of them (Axelsson *et al.* 2010, current study). Although publication bias always works against the null hypothesis, it still seems possible that transmission distortion is a more common phenomenon in avian genomes and studies in other pedigreed bird populations are needed to draw general conclusions.

Acknowledgements

We thank S. Bauer, E. Bodendorfer, A. Grötsch, J. Hacker, M. Halser, K. Martin, J. Minshull, P. Neubauer, F. Preininger, M. Ruhdofer and A. Türk for animal care and help with breeding; M. Schneider and M. Thomas for laboratory work; N. Backström and H. Mellenius for kindly providing the linkage map. U.K. is part of the International Max Planck Research School for Organismal Biology. H.S. is supported by an Emmy-Noether fellowship of the German Research Foundation (SCHI 1188/1-1). This study was funded by financial support from the Max Planck Society (B.K.), the Swedish Research Council (H.E.), the Knut and Alice Wallenberg Foundation (H.E.), the German Research Foundation grants FO340/1-1, FO340/1-2 and FO340/1-3 (W.F.) and the European Union grant HPMF-CT-2002-01871 (W.F.).

References

Aparicio JM, Ortego J, Calabuig G, Cordero PJ (2010) Evidence of subtle departures from Mendelian segregation in a wild lesser kestrel (*Falco naumanni*) population. *Heredity*, **105**, 213–219.

Axelrod R, Hamilton WD (1981) The evolution of cooperation. *Science*, **211**, 1390–1396.

Axelsson E, Albrechtsen A, van AP *et al.* (2010) Segregation distortion in chicken and the evolutionary consequences of female meiotic drive in birds. *Heredity*, **105**, 290–298.

Backström N, Forstmeier W, Schielzeth H *et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research*, **20**, 485–495.

Balakrishnan CN, Edwards SV (2009) Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics*, **181**, 645–660.

Balakrishnan CN, Edwards SV, Clayton DF (2010) The Zebra Finch genome and avian genomics in the wild. *EMU*, **110**, 233–241.

Balding DJ (2006) A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*, **7**, 781–791.

Bates D, Maechler M, Bolker B, Walker S (2014) lme4: Linear mixed-effects models using Eigen and S4. *R package version 1.1-6*.

Bernasconi G, Ashman TL, Birkhead TR *et al.* (2004) Evolutionary ecology of the prezygotic stage. *Science*, **303**, 971–975.

Birdsell JA (2002) Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Molecular Biology and Evolution*, **19**, 1181–1197.

Burnham CR (1962) *Discussions in Cytogenetics*, 1st edn. Burgess Publishing Company, Minneapolis, USA.

Burt A, Trivers R (2006) *Genes in Conflict: The Biology of Selfish Genetic Elements*. Belknap Press, Harvard University Press, Cambridge, Massachusetts, London, UK.

Champely S (2012) pwr: Basic functions for power analysis. *R package version 1.1.1*.

Day T, Taylor PD (1998) Chromosomal drive and the evolution of meiotic nondisjunction and trisomy in humans. *Proceedings of the National Academy of Sciences, USA*, **95**, 2361–2365.

Dreszer TR, Wall GD, Haussler D, Pollard KS (2007) Biased clustered substitutions in the human genome: the footprints of male-driven biased gene conversion. *Genome Research*, **17**, 1420–1430.

Duret L, Galtier N (2009) Biased gene conversion and the evolution of mammalian genomic landscapes. *Annual Review of Genomics and Human Genetics*, **10**, 285–311.

Ellegren H, Smeds L, Burri R *et al.* (2012) The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature*, **491**, 756–760.

Fan JB, Oliphant A, Shen R *et al.* (2003) Highly parallel SNP genotyping. *Cold Spring Harbor Symposia on Quantitative Biology*, **68**, 69–78.

Fedyk S, Chetnicki W (2007) Preferential segregation of metacentric chromosomes in simple Robertsonian heterozygotes of *Sorex araneus*. *Heredity*, **99**, 545–552.

Fishman L, Saunders A (2008) Centromere-associated female meiotic drive entails male fitness costs in monkeyflowers. *Science*, **322**, 1559–1562.

Forstmeier W, Coltman DW, Birkhead TR (2004) Maternal effects influence the sexual behavior of sons and daughters in the zebra finch. *Evolution*, **58**, 2574–2583.

Forstmeier W, Schielzeth H, Schneider M, Kempenaers B (2007) Development of polymorphic microsatellite markers for the zebra finch (*Taeniopygia guttata*). *Molecular Ecology Notes*, **7**, 1026–1028.

Foulkes AS (2009) *Applied Statistical Genetics with R: For Population-based Association Studies*. Springer, Dordrecht, Heidelberg, London, New York.

Frésard L, Leroux S, Servin B *et al.* (2014) Transcriptome-wide investigation of genomic imprinting in chicken. *Nucleic Acids Research*, **42**, 3768–3782.

Gabriel S, Ziaugra L, Tabbaa D (2009) SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Current Protocols in Human Genetics* **Chapter 2**, Unit 2 12.

Griffith SC, Buchanan KL (2010) The Zebra Finch: the ultimate Australian supermodel. *EMU*, **110**, v–xii.

Hartl DL (1969) Dysfunctional sperm production in *Drosophila melanogaster* males homozygous for segregation distorter elements. *Proceedings of the National Academy of Sciences, USA*, **63**, 782–789.

Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. *Science*, **293**, 1098–1102.

Jeffreys AJ, May CA (2004) Intense and highly localized gene conversion activity in human meiotic crossover hot spots. *Nature Genetics*, **36**, 151–156.

Jeffreys AJ, Neumann R (2002) Reciprocal crossover asymmetry and meiotic drive in a human recombination hot spot. *Nature Genetics*, **31**, 267–271.

Knief U, Hemmrich-Stanisak G, Wittig M *et al.* (2015) Quantifying realized inbreeding in wild and captive animal populations. *Heredity*, **114**, 397–403.

Koehler KE, Millie EA, Cherry JP *et al.* (2002) Sex-specific differences in meiotic chromosome segregation revealed by dicentric bridge resolution in mice. *Genetics*, **162**, 1367–1379.

Lumley T (2012) rmeta: Meta-analysis. *R package version 2.16*.

Lyon MF (2003) Transmission ratio distortion in mice. *Annual Review of Genetics*, **37**, 393–408.

Lyttle TW (1991) Segregation distorters. *Annual Review of Genetics*, **25**, 511–557.

Lyttle TW (1993) Cheaters sometimes prosper: distortion of mendelian segregation by meiotic drive. *Trends in Genetics*, **9**, 205–210.

Mendel G (1865) Versuche über Pflanzen-Hybriden. *Verhandlungen des naturforschenden Vereins in Brünn*, **4**, 3–47.

Meyer WK, Arbeithuber B, Ober C *et al.* (2012) Evaluating the evidence for transmission distortion in human pedigrees. *Genetics*, **191**, 215–232.

Mitchell AA, Cutler DJ, Chakravarti A (2003) Undetected genotyping errors cause apparent overtransmission of common alleles in the transmission/disequilibrium test. *American Journal of Human Genetics*, **72**, 598–610.

Pardo-Manuel de Villena F, Sapienza C (2001) Nonrandom segregation during meiosis: the unfairness of females. *Mammalian Genome*, **12**, 331–339.

Presgraves DC (2010) The molecular evolutionary basis of species formation. *Nature Reviews Genetics*, **11**, 175–180.

R Core Team (2013) R: a language and environment for statistical computing. *version 3.0.2*.

Schielzeth H (2010) Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, **1**, 103–113.

Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2011) Data from: QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. *Dryad Digital Repository*, doi: 10.5061/dryad.r044b.

Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2012) QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. *Evolution*, **66**, 18–30.

Traulsen A, Reed FA (2012) From genes to games: cooperation and cyclic dominance in meiotic drive. *Journal of Theoretical Biology*, **299**, 120–125.

Tsai JH, McKee BD (2011) Homologous pairing and the role of pairing centers in meiosis. *Journal of Cell Science*, **124**, 1955–1963.

Úbeda F, Haig D (2005) On the evolutionary stability of Mendelian segregation. *Genetics*, **170**, 1345–1357.

Visscher PM, Medland SE, Ferreira MAR *et al.* (2006) Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *Plos Genetics*, **2**, 316–325.

Warren WC, Clayton DF, Ellegren H *et al.* (2010) The genome of a songbird. *Nature*, **464**, 757–762.

Weber CC, Boussau B, Romiguier J, Jarvis ED, Ellegren H (2014) Evidence for GC-biased gene conversion as a driver of between-lineage differences in avian base composition. *Genome Biology*, **15**, 549.

Webster MT, Smith NGC, Hultin-Rosenberg L, Arndt PF, Ellegren H (2005) Male-driven biased gene conversion governs the evolution of base composition in human Alu repeats. *Molecular Biology and Evolution*, **22**, 1468–1474.

Williams AL, Geneovese G, Dyer T *et al.* (2015) Non-crossover gene conversions show strong GC bias and unexpected clustering in humans. *eLife*, **4**, e04637.

Wyttenbach A, Borodin P, Hausser J (1998) Meiotic drive favors Robertsonian metacentric chromosomes in the common shrew (*Sorex araneus*, Insectivora, Mammalia). *Cytogenetics and Cell Genetics*, **83**, 199–206.

Zann RA (1996) *The Zebra Finch: A Synthesis of Field and Laboratory Studies*. Oxford University Press, New York, USA.

Zöllner S, Wen X, Hanchard NA *et al.* (2004) Evidence for extensive transmission distortion in the human genome. *American Journal of Human Genetics*, **74**, 62–72.

U.K., H.S., H.E., B.K. and W.F. jointly conceived of the study; U.K. analysed the data with input from H.S. and W.F.; and U.K. wrote the manuscript with input from all authors.

Data accessibility

For the initial genome-wide scan: Pedigree and genotype data as well as a linkage map: DRYAD entry doi: 10.5061/dryad.r044b. Results from binomial tests for transmission distortion for every SNP: DRYAD entry doi: 10.5061/dryad.s181n. For the test for genome-wide transmission distortion: Simulation script and input file:

DRYAD entry doi: 10.5061/dryad.s181n. For the two follow-up studies: Pedigree and microsatellite and Sequenom genotype data: DRYAD entry doi: 10.5061/dryad.s181n.

Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1. Methods and results of a test for genome-wide weak transmission distortion in favour of major alleles.

Table S1 Primer sequences, positions in the genome and melting temperatures used in the PCR.

Table S2 Transmission ratio of locus rs82439270 on chromosome *Tgu5* in the follow-up analysis for both heterozygous males and females separately.

Table S3 Summary of all transmission ratios for the driving locus on chromosome *Tgu2* in heterozygous females + both parents heterozygous, males + both parents heterozygous, combined sexes, only heterozygous females (excluding cases where both parents are heterozygous) and only heterozygous males (excluding cases where both parents are heterozygous).

Fig. S1 (A) Histogram of the number of informative meioses available for each SNP in the initial whole genome scans. For the sex-specific scans we here exclude all cases where both parents are heterozygous. (B) Power analysis for the number of informative meioses for different transmission ratios considering a *P*-value of 4.0×10^{-5} as significant.

Fig. S2 Whole-genome transmission ratios in (A) both heterozygous sexes combined, (B) heterozygous female parents only and (C) heterozygous male parents only.

Fig. S3 Whole-genome scans for transmission distortion in (A) heterozygous female parents only, excluding those pairs in which both female and male are heterozygous and (B) heterozygous male parents only, excluding those pairs in which both male and female are heterozygous.

Supplement

Methods and results of a test for genome-wide weak transmission distortion in favor of major alleles

Methods

Aparicio *et al.* (2010) suggested that weak transmission distortion may more commonly favor the major alleles due to recessive deleterious mutations, which are more likely to be associated with the minor alleles (though founder effects in our population are expected to add noise to this relationship). To get an estimate of the genome-wide transmission ratio of the major allele, we first sampled one SNP from each chromosome, then summed up all inheritance events of the major and minor alleles of these SNPs and calculated the transmission ratio. We repeated this 10,000 times and calculated the 95% quantile range (QR).

Results

Summing up the informative inheritance events did not show a consistent transmission bias in favor of the major alleles across all loci (combined-sexes transmission ratio [95% QR] = 0.501 [0.493–0.509], female-specific transmission ratio = 0.501 [0.493–0.510], male-specific transmission ratio = 0.501 [0.492–0.511]).

Discussion

Neither in the combined-sexes nor in the sex-specific genome scans did we detect any sign of weak genome-wide transmission bias. A small effect in favor of major alleles may be expected due to recessive deleterious mutations, which are more likely to be associated with the minor alleles and which was found by Aparicio *et al.* (2010) in a wild population of lesser kestrels. However, those minor alleles that are indeed linked to recessive deleterious mutations may be at especially low frequencies and their effect would be masked by SNPs with higher allele frequencies (which are more likely not linked to recessive deleterious mutations) and founder effects could additionally change allele frequencies in captivity. Also, a weak transmission bias towards major alleles might be expected if genotyping errors were common (Mitchell *et al.* 2003). In any case, major and minor alleles were inherited with similar probabilities in our population.

Although examples for segregation distorters are sparse and probably prone to detection bias, classical distorters often bias transmission by more than 90% (Lyttle 1993) and often have low stable equilibrium frequencies in a population (e.g. Presgraves *et al.* 2009). We found 12 and 17 segregating haplotypes on chromosomes *Tgu2* and *Tgu5*, respectively, and a previous study identified 17 haplotypes across the *ESR1* gene on chromosome *Tgu3* (Forstmeier *et al.* 2012), which suggests that we were unable to tag all low frequency haplotypes within our population with a unique marker since on average we had 39 SNPs per chromosome (median 27 SNPs). However, more than 79% of the SNPs would detect

transmission rates of 0.7 with 80% power, such that transmission distortion of SNPs which tag multiple non-driving haplotypes along with the driving haplotype could be detected. Nonetheless, in the genome-wide scans we did not identify any such strong distorters but they could potentially be still present at low frequency or on the eight chromosomes missing in the current genome assembly (Pigozzi & Solari 1998; Warren *et al.* 2010).

Studies on transmission distortion in other bird species were based on comparable numbers of informative meioses as our initial scan and also showed only subtle departures from Mendelian segregation (Aparicio *et al.* 2010; Axelsson *et al.* 2010). Since their results were not replicated in independent sets of birds, conclusions should be treated carefully, as illustrated by our data on chromosome *Tgu5*. In a recent paper, Ellegren *et al.* (2012) suggested to test for meiotic drive by extensive genotyping in pedigrees because they found genomic divergence peaks between two flycatcher species close to potential centromeres and telomeres and invoke a meiotic drive model of speciation. Such studies could easily be conducted with microsatellite markers used for paternity analysis in wild populations, although genome-wide coverage is difficult to achieve. However, as replication is the gold standard for validation of any genetic association study (NCI-NHGRI Working Group on Replication in Association Studies 2007), following guidelines for replication should also be crucial when testing for transmission distortion.

References

Aparicio JM, Ortego J, Calabuig G, Cordero PJ (2010) Evidence of subtle departures from Mendelian segregation in a wild lesser kestrel (*Falco naumanni*) population. *Heredity* **105**, 213–219.

Axelsson E, Albrechtsen A, Van AP*, et al.* (2010) Segregation distortion in chicken and the evolutionary consequences of female meiotic drive in birds. *Heredity* **105**, 290–298.

Ellegren H, Smeds L, Burri R*, et al.* (2012) The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* **491**, 756–760.

Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B (2012) Heterozygosity-fitness correlations in zebra finches: microsatellite markers can be better than their reputation. *Molecular Ecology* **21**, 3237–3249.

Lyttle TW (1993) Cheaters sometimes prosper: distortion of mendelian segregation by meiotic drive. *Trends in Genetics* **9**, 205–210.

Mitchell AA, Cutler DJ, Chakravarti A (2003) Undetected genotyping errors cause apparent overtransmission of common alleles in the transmission/disequilibrium test. *American Journal of Human Genetics* **72**, 598–610.

NCI-NHGRI Working Group on Replication in Association Studies, Chanock SJ, Manolio T*, et al.* (2007) Replicating genotype-phenotype associations. *Nature* **447**, 655–660.

Pigozzi MI, Solari AJ (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.

Presgraves DC, Gerard PR, Cherukuri A, Lyttle TW (2009) Large-scale selective sweep among segregation distorter chromosomes in African populations of *Drosophila melanogaster*. *Plos Genetics* **5**, e1000463.

Warren WC, Clayton DF, Ellegren H*, et al.* (2010) The genome of a songbird. *Nature* **464**, 757–762.

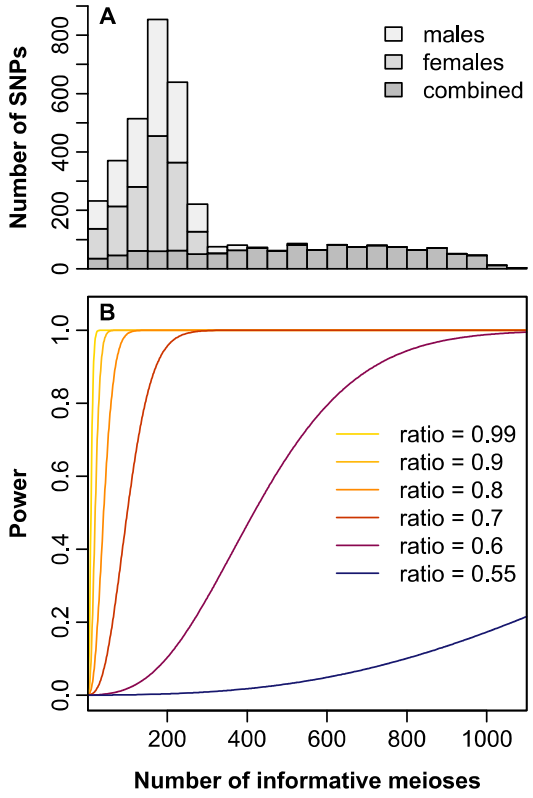


Figure S1: *(A) Histogram of the number of informative meioses available for each SNP in the initial whole genome scans. For the sex-specific scans we here exclude all cases where both parents are heterozygous. (B) Power analysis for the number of informative meioses for different transmission ratios considering a P-value of 4.0 x* 10^{-5} *as significant.*

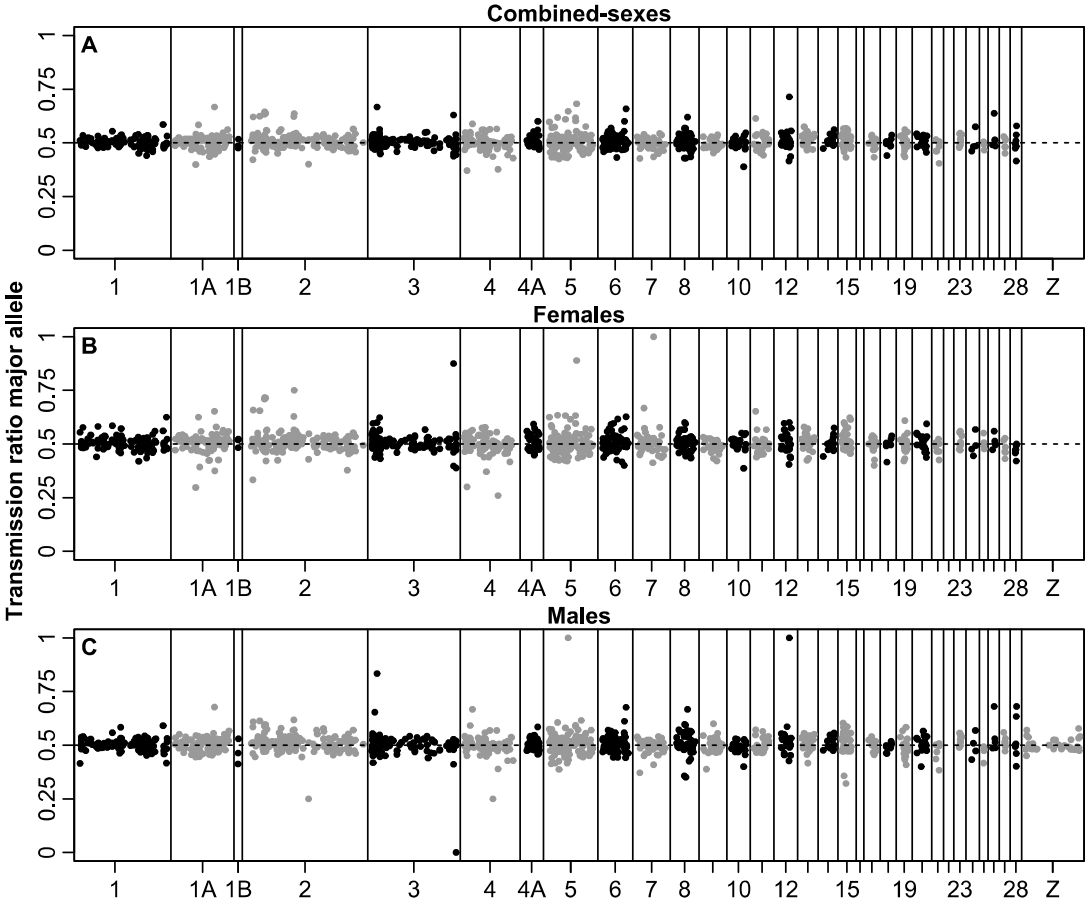


Figure S2: *Whole-genome transmission ratios in (A) both heterozygous sexes combined, (B) heterozygous female parents only and (C) heterozygous male parents only. On the abscissa are the chromosomes and on the ordinate are transmission ratios of the major allele for each SNP. The dashed line indicates fair segregation of 0.5. Chromosome TguZ was tested only in males because only males carry two copies of TguZ. Those SNPs with transmission ratios above 0.8 or below 0.2 had less than 10 informative meioses.*

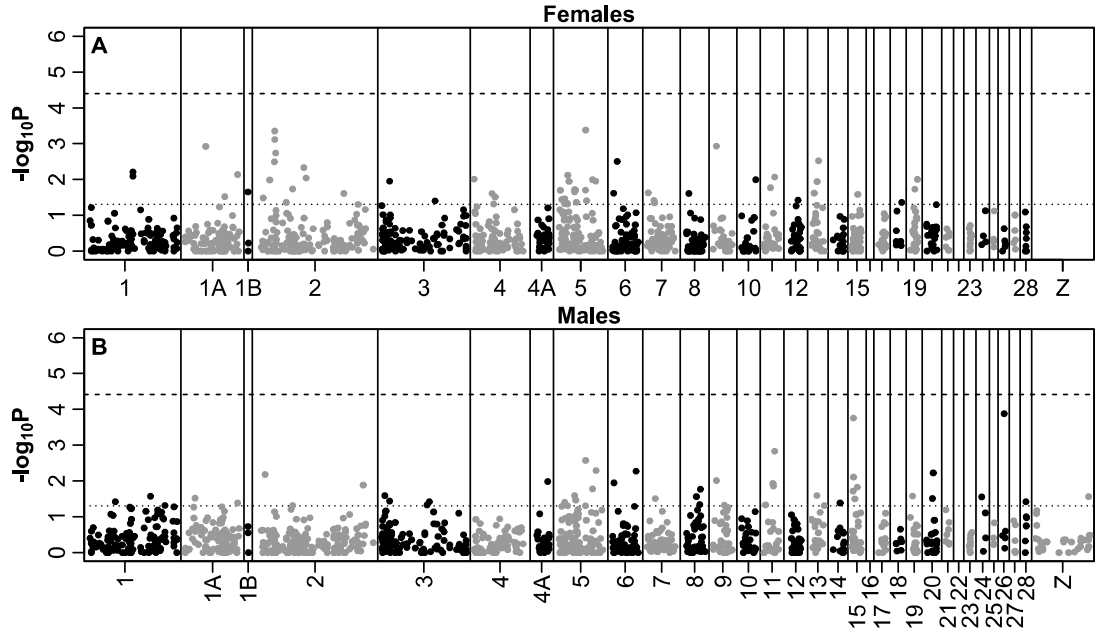


Figure S3: *Whole-genome scans for transmission distortion in (A) heterozygous female parents only, excluding those pairs in which both female and male are heterozygous and (B) heterozygous male parents only, excluding those pairs in which both male and female are heterozygous. On the abscissa are the chromosomes and on the ordinate are the -log$_{10}$(P-values). The dashed line indicates the significance threshold after genome-wide Bonferroni correction and the dotted line before genome-wide Bonferroni correction. Chromosome TguZ was tested only in males because only males carry two copies of TguZ.*

Table S1: *Primer sequences, positions in the genome and melting temperatures used in the PCR.*

Microsatellite	Start	End	Distance to SNP (Mb)	Motif	Primer name	Sequence	Tm (°C)
Tgu5_SD	35,669,252	35,669,308	2.15	TG	-F	CTACAGTCAGTGAAACCGTTC	57
					-R	GCATGGAACTGCATGCCTTA	57
Tgu5_SD4	38,137,613	38,137,648	0.32	CA	-F	CCCTTGGGGCTTCTCATCAT	58
					-R	TGCACCATCCCACTGAACTG	58
Tgu2_SD44	43,803,808	43,803,904	19.01	TG	-F	TGGAAGTGGCAAGGACAACA	57
					-R	TCCCTGCTCCCTATCTGTAT	57
Tgu2_SD60	60,167,999	60,168,042	35.37	CA	-F	CGTCCCAAAACACCAATCGT	57
					-R	CCTCACAACACGAAGCAGAT	57

Table S2: *Transmission ratio of locus rs82439270 on chromosome Tgu5 in the follow-up analysis for both heterozygous males and females separately. The follow-up analysis consisted of (1) a replicate of the initial genome-wide scan which contained only birds that hatched (survived), (2) embryos that died naturally during incubation (died embryos) and (3) embryos that were collected for DNA analysis before hatching or whose egg shells broke and whose fate is thus unknown (other embryos). The observed number of inheritance events of the major allele is shown under* n_A.

Sex	Sample	Informative meioses	n_A	Transmission ratio	95% CI	P-value
females	survived	640	311	0.486	0.447, 0.525	0.50
	died embryos	245	123	0.502	0.438, 0.566	1.00
	other embryos	422	223	0.528	0.480, 0.577	0.26
	all	1307	657	0.503	0.475, 0.530	0.87
males	survived	679	332	0.489	0.451, 0.527	0.59
	died embryos	220	109	0.495	0.428, 0.563	0.95
	other embryos	338	186	0.55	0.496, 0.604	0.073
	all	1237	627	0.507	0.479, 0.535	0.65

Table S3: *Summary of all transmission ratios for the driving locus on chromosome Tgu2 in heterozygous females + both parents heterozygous, males + both parents heterozygous, combined sexes, only heterozygous females (excluding cases where both parents are heterozygous) and only heterozygous males (excluding cases where both parents are heterozygous). The observed number of inheritance events of the major allele is shown under* n_A.

Sex	Sample	Informative meioses	n_A	Transmission ratio	95% CI	P-value
females	genome scan	106	76	0.717	0.621, 0.800	9×10^{-6}
	follow-up 1 survived	121	59	0.488	0.396, 0.580	0.86
	follow-up 1 died embryos	60	33	0.55	0.416, 0.679	0.52
	follow-up 1 other embryos	109	64	0.587	0.489, 0.681	0.084
	follow-up 1 all	290	156	0.538	0.479, 0.596	0.22
	follow-up 2 all	262	146	0.557	0.495, 0.618	0.073
	all	658	378	0.574	0.536, 0.613	0.00015
males	genome scan	90	53	0.589	0.480, 0.692	0.11
	follow-up 1 survived	85	48	0.565	0.453, 0.672	0.28
	follow-up 1 died embryos	42	24	0.571	0.410, 0.723	0.44
	follow-up 1 other embryos	92	55	0.598	0.490, 0.699	0.076
	follow-up 1 all	219	127	0.58	0.512, 0.646	0.021
	follow-up 2 all	232	132	0.569	0.503, 0.634	0.042
	all	541	312	0.577	0.534, 0.619	0.00041
combined	genome scan	160	103	0.644	0.564, 0.718	0.00034
	follow-up 1 survived	178	94	0.528	0.452, 0.603	0.50
	follow-up 1 died embryos	90	52	0.578	0.469, 0.681	0.17
	follow-up 1 other embryos	179	106	0.592	0.516, 0.665	0.017
	follow-up 1 all	447	252	0.564	0.516, 0.610	0.0080
	follow-up 2 all	494	278	0.563	0.518, 0.607	0.0060
	all	1101	633	0.575	0.545, 0.604	7×10^{-7}
females only	genome scan	70	50	0.714	0.594, 0.816	0.00044
	follow-up 1 survived	93	46	0.495	0.389, 0.600	1.00
	follow-up 1 died embryos	48	28	0.583	0.432, 0.724	0.31
	follow-up 1 other embryos	87	51	0.586	0.476, 0.691	0.13
	follow-up 1 all	228	125	0.548	0.481, 0.614	0.16
	follow-up 2 all	262	146	0.557	0.495, 0.618	0.073
	all	560	321	0.573	0.531, 0.615	0.00061
males only	genome scan	54	27	0.5	0.361, 0.639	1.00
	follow-up 1 survived	57	35	0.614	0.476, 0.740	0.11
	follow-up 1 died embryos	30	19	0.633	0.439, 0.801	0.20
	follow-up 1 other embryos	70	42	0.6	0.476, 0.715	0.12
	follow-up 1 all	157	96	0.611	0.531, 0.688	0.0065
	follow-up 2 all	232	132	0.569	0.503, 0.634	0.042
	all	443	255	0.576	0.528, 0.622	0.0017

Chapter 5: Triploidy mapping of centromeres of microchromosomes in the zebra finch (*Taeniopygia guttata*)

Abstract

Centromeres are the attachment site of the spindle microtubules and are essential for the proper segregation of chromosomes in mitosis and meiosis. They usually consist of hundreds of kilobases of repetitive sequence and are often missing from assembled genomes. Because of that their location on physical chromosome maps has to be inferred from flanking sequences via fluorescence in situ hybridization (FISH) or by linkage analysis of half-tetrads (resulting from the accidental inheritance of two chromatids from a single meiosis). Here we genotype 37 zebra finches (Taeniopygia guttata) that were triploid or tetraploid due to inheritance errors (and mostly died as embryos) together with their parents at 64 microsatellite markers (at least two per chromosome) to (1) determine the parental origin of the supernumerary haploid chromosome set, (2) determine the stage in meiosis in which the non-disjunction occurred and (3) map the location of centromeres on all assembled chromosomes in the zebra finch reference genome. For the purpose of centromere mapping paternally derived triploidies are not suitable, since they are often caused by dispermy such that a half-tetrad cannot be recovered. The majority of triploidies had a maternal origin (n = 22; including two tetraploids that were both maternally and paternally derived) and arose equally frequent from errors in the first and second meiotic division (n = 12 and n = 10, respectively). For the ten largest chromosomes the centromere positions were already known from FISH mapping and here we map the centromeres of an additional 19 acrocentric microchromosomes in reference to the current genome assembly.

Prepared as: Knief U, W Forstmeier: Triploidy mapping of centromeres of microchromosomes in the zebra finch (*Taeniopygia guttata*).

Introduction

Centromeres are the attachment site of the spindle microtubules and are essential for the proper segregation of chromosomes in mitosis and meiosis. The location of centromeres can be readily identified by means of cytogenetic methods, yet integrating the cytogenetic with the linkage/physical map can be quite difficult, because centromeres are not defined by a specific sequence motive, they usually consist of hundreds of kilobases of repetitive sequence and because of that they are often missing from assembled genomes (Krasikova *et al.* 2006; Shang *et al.* 2010). However, it has been recently suggested that centromeric/telomeric DNA prominently contributes to species divergence in birds and mammals (Carneiro *et al.* 2009; Ellegren *et al.* 2012) and because of that more knowledge about the position of these cytogenetic features is needed for any reference genome.

The zebra finch (*Taeniopygia guttata*) was the second avian species whose genome was sequenced and it was assembled into 32 chromosomes using bacterial artificial chromosomes (BACs) and a linkage map (Warren *et al.* 2010). With its genomic resources at hand it is arguably qualitatively the second best annotated avian genome after the chicken (*Gallus gallus*), even though karyotypically the genome consists of n = 40 chromosomes (Pigozzi & Solari 1998), meaning that eight chromosomes have not been assembled yet. Throughout this paper we will use the chromosome nomenclature introduced by (Itoh & Arnold 2005; Warren *et al.* 2010).

Bird genomes consist of a few large macro- and several smaller microchromosomes; the exact definition of them being rather loose. It is generally accepted that the zebra finch genome consists of seven macrochromosomes (*Tgu1*–*Tgu5*, *Tgu1A* and *TguZ*; Itoh & Arnold 2005) with an assembled size range of 62–156 Mb and 33 microchromosomes ranging from 9 kb to 40 Mb.

Notwithstanding the amount of genomic and molecular tools available for the zebra finch, only for the ten largest of the 32 assembled chromosomes the location of the centromere is known in reference to the physical map (WUSTL v3.2.4; Warren *et al.* 2010). These positions were inferred from flanking sequences via fluorescence *in situ* hybridization (FISH). Among those ten chromosomes seven are submetacentric (*Tgu1*, *Tgu1A*, *Tgu2*, *Tgu3*, *Tgu4*, *Tgu7*, *TguZ*), meaning that the centromere is located slightly off the middle of the chromosome, and the remaining more or less acrocentric (*Tgu5*, *Tgu6*, *Tgu8*), meaning that the centromere is located on either end of the chromosome (Pigozzi & Solari 1998). Chromosome *Tgu5* is known to be polymorphic for a pericentric inversion which changes the chromosome to be submetacentric (Christidis 1986; Itoh & Arnold 2005). All other chromosomes that are smaller than chromosome *Tgu8*, including those that are not yet assembled, are known to be acrocentric from cytogenetic studies (Pigozzi 2008). However, their centromeric ends have not been distinguished from their distal ends in reference to the physical map.

To fill this gap we here apply centromere-marker-mapping using linkage analysis in half-tetrads, which requires that at least two chromatids of a single meiosis are recovered together (Mather 1938). In a normal meiosis, homologous chromosomes are separated in the first meiotic division (meiosis I) and sister-chromatids in the second meiotic division (meiosis II), giving rise to four haploid gametes. In the female meiosis, three of these haploid cells degenerate (the so called polar bodies) and a single oocyte survives. Centromeres are the attachment sites for the spindle microtubules that mediate the separation of chromosomes in meiosis I and II. Accordingly, in the first meiotic division the centromeres of homologous chromosomes get separated and molecular markers located close to the centromere tend to be reduced. Specifically, this means if the mother is heterozygous at a marker close to the centromere the two alleles will separate in meiosis I and the two sister chromatids within each daughter cell will be homozygous. Whenever an uneven number of cross-overs between the centromere and the molecular marker occurs, the two alleles separate in the second meiotic division (Johnson *et al.* 1996).

Triploid individuals may carry two chromatids of a single meiosis (a so called half-tetrad; Zhao & Speed 1998). The supernumerary haploid chromosome set originates either from the mother (digyny) or the father (diandry) and may arise from non-disjunction of homologous chromosomes at meiosis I or by non-disjunction of sister-chromatids at meiosis II. Diandric triploidies can also result from dispermy, the fertilization of a single egg by two sperm cells (Jacobs & Morton 1977). Since dispermy is not caused by a meiotic failure, a half-tetrad cannot be recovered and hence those triploids are not suitable for centromere mapping.

In order to distinguish between non-disjunction at the first or second meiotic division a molecular marker close to the centromere is needed (Chakravarti & Slaugenhaupt 1987). Whenever the parent contributing the third chromosome set is heterozygous for that marker, the triploid offspring inherits both alleles (i.e. both homologous chromosomes) in case of a meiosis I error and two copies of the same allele (i.e. both sister-chromatids) in case of an error in the second meiotic division. A molecular marker at the distal end may convey additional information, if exactly one or an uneven number of cross-overs happens between itself and the centromere, because then an error in the second meiotic division would always lead to the inheritance of both alleles (because of the cross-over the sister-chromatids carry parts of both homologous chromosomes at that position; Côté & Edwards 1975).

In order to obtain several half-tetrads per chromosome we use naturally occurring triploid zebra finches, which usually die at the embryo stage (Forstmeier & Ellegren 2010) but occasionally survive to adulthood (Girndt *et al.* 2014). By genotyping each triploid individual at one microsatellite close to the ten known centromeres and one close to a distal end on

each of these chromosomes we were able to distinguish digynic from diandric triploidies and subsequently identify whether the digynic meiotic failures occurred in the first or second meiotic division. Since all remaining chromosomes with an unknown physical centromere position are acrocentric (Pigozzi 2008), we then designed primers for microsatellites on both ends of each chromosome. Since one of the two markers was close to the centromere, we were able to orient the physical/linkage map in respect to the centromere for almost all assembled microchromosomes in the zebra finch genome.

Material and Methods

Individuals and populations

When using microsatellites for identifying cases of triploidy not all markers are expected to show three alleles because they (1) could be blind due to homozygosity or mother and father sharing the same alleles or (2) are located at a physical position along the chromosome that gets inherited twice from the same homolog (see Introduction). During regular paternity analysis of 4,993 alive birds (we consider them as alive birds if they hatch) and 2,999 embryos (including cases where the egg shell broke or the egg was opened before the due date of hatching) 6 birds (3 of which survived to adulthood) and 28 embryos had been identified as being trisomic for at least three chromosomes (range: 3–16) and we assumed that these 34 individuals were triploid. In previous studies using single nucleotide polymorphism (SNP) markers spread across the whole genome a subset of these 34 individuals, namely 8 embryos (n=1,395 SNPs; Forstmeier & Ellegren 2010) and 2 adult birds (n=2,417 SNPs; Girndt *et al.* 2014) were confirmed as being triploid (trisomic for all 32 chromosomes in the WUSTL v3.2.4 assembly). An additional 3 embryos were found to be triploid by genotyping the same SNP set as in Girndt *et al.* (2014) in 115 embryos that had died from natural causes. Thus, in total we had 37 individuals that were triploid. To determine whether the supernumerary haploid chromosome was inherited from the mother or the father we included all the parents of the triploid individuals in our study.

The 37 triploid individuals stemmed from three different populations: (1) Our main population held at the Max-Planck-Institute for Ornithology in Seewiesen (n=19; study population 18 in Forstmeier *et al.* 2007), (2) a recently wild-derived population held at the Max-Planck-Institute for Ornithology in Seewiesen (n=13; originating from study population 4 in Forstmeier *et al.* 2007), (3) a population that was produced by crossing individuals from a captive population held in Cracow (study population 11 in Forstmeier *et al.* 2007) with our main population (n=5). Since we used differing microsatellite sets for trisomy detection within and between each of the three populations, detection probabilities varied and a comparison of the rate of triploidy between populations is not meaningful. The only unbiased estimate of the rate of triploidy can be obtained from the 115 dead embryos genotyped with 2,417 SNPs (Girndt *et al.* 2014), which yielded three triploids (2.6%) among

naturally dying embryos (with about 25–30% of all embryos dying naturally during development).

Genetic markers

For each of the ten chromosomes with a known centromere location we designed primers to amplify two microsatellites, one of them located close to the known centromere and the other at the most distant chromosome end. On chromosome *Tgu5* and *Tgu6* the FISH probes mapping closest to the centromere are located on sequences, whose positions within the chromosomes are not known (chromosomes *Tgu5_random* and *Tgu6_random*; Warren *et al.* 2010). Thus, we designed primers for microsatellites that are positioned on the same Contig as the FISH probes. Yet on chromosome *Tgu6_random* the marker appears to be quite far from the centromere so we designed an additional primer pair for a microsatellite on chromosome *Tgu6* which should be located close to the centromere. For the 22 microchromosomes with an unknown centromere location we designed primers for two microsatellites, one at the start and one at the end of each chromosome (excluding the difficult-to-assemble chromosome *Tgu16* which is only 9.9 kb in the current genome assembly but known to be several hundred times larger; Ekblom *et al.* 2011; Pichugin *et al.* 2001). Since all chromosomes with an unknown centromere position are acrocentric (Pigozzi 2008), one microsatellite should be located close to the centromere and the other one close to the distal end (see Supplementary Table S1 for detailed information for each primer pair). However, if parts of the chromosome are missing from the assembly, markers could be further away from the centromere or from the distal end (see Discussion).

We used the primer pair 3007/3112 for sexing all embryos, which amplifies an intron in the *CHD1* gene differing in length on chromosome *TguZ* and chromosome *TguW* (Ellegren & Fridolfsson 1997).

DNA extraction and genotyping

DNA was extracted from blood or tissue samples of all triploid individuals and their parents using the NucleoSpin Blood QuickPure Kit (Macherey-Nagel). Both the Type-it Microsatellite PCR Kit (Qiagen) and the Multiplex PCR Kit (Qiagen) were used for genotyping following manufacturer's instructions (with the exception of an extension step of 60 °C for 30 minutes instead of 72 °C for 10 minutes with the Multiplex PCR Kit). Details on the PCR protocol for each multiplex are given in Supplementary Table S1.

Determination of parental origin

We first determined whether the supernumerary haploid chromosome set was inherited from the mother or the father. For that we considered those markers as being informative which showed the genotype AB in one parent, CD or CC in the other parent and ABC or ABD in the offspring. Parental origin of the additional chromosome set could be determined in 32 out of the 37 individuals with at least two markers per individual being informative (Supplementary Table S2). Of the remaining five individuals, two were found to be

tetraploid with one additional chromosome set inherited from the mother and one from the father, and were hence still useful for the current study (2011_180, 2011_251). The other three individuals (K2012/13_125, 2011_289, 2006_584) had to be excluded because they were uninformative at all marker loci or appeared to be a mixture of digynic and diandric origin of the third chromosome set.

Determination of mechanism of origin

Triploidy may arise from non-disjunction of homologous chromosomes at meiosis I or by non-disjunction of sister-chromatids at meiosis II. Since diandric triploidies may also result from dispermy, in which case a half-tetrad cannot be recovered, they are not useful for centromere mapping and were excluded from further analyses and will be described elsewhere (n=12).

In the remaining 20 digynic triploids and the two tetraploids (2011_180, 2011_251), those markers located close to the known centromeres on the ten largest chromosomes (*Tgu1*–*Tgu8*, *Tgu1A* and *TguZ*) were used to distinguish between non-disjunction of homologous chromosomes at meiosis I or non-disjunction of sister-chromatids at meiosis II (Figure 1). For that purpose we assumed that the centromeric markers were in complete linkage with the centromere. Hence, whenever the mother was heterozygous at a centromeric marker and passed on both her alleles to the triploid offspring, we took it as evidence for an error in the first meiotic division. Each time she passed on only one of her two alleles, it was pointing to an error in meiosis II (see Introduction for the underlying logic).

Female birds carry one Z and one W chromosome. In zebra finches, the Z and the W chromosome pair during meiosis I (Pigozzi & Solari 1998) and a mandatory recombination event happens in the pseudoautosomal region (PAR) (Pigozzi 2008). Since the PAR is located at one end of chromosome *TguZ* (minimum range 1,213,256 – 1,464,488 bp; Stapley *et al.* 2008) and the centromere is located around 28 Mb, a centromeric marker will always be located on chromosome *TguZ* and not recombine with chromosome *TguW*. If non-disjunction happens in meiosis I females will always inherit a single Z and a single W chromosome. Meiosis II errors should lead to the inheritance of either two Z or two W chromatids with equal probabilities.

Mapping of centromeres

We used the maximum-likelihood method in Chakravarti *et al.* (1989) to estimate the distance of our markers to the centromere under complete interference, i.e. that only a single cross-over between the marker and the centromere is allowed. Complete interference is a reasonable assumption since usually a single cross-over happens per chromosome arm in the zebra finch (Calderón & Pigozzi 2006). However, one should keep in mind that the estimated genetic distances are restricted to 50 cM (if there is one cross-over

in any meiosis between two markers then they are 50 cM apart) and may be underestimated because of occasional double or triple cross-overs.

In order to estimate the genetic distance of our markers from the centromere, we define m_1 as being the number of non-reduced triploid individuals and m_2 being the number of reduced triploid individuals at a specific marker resulting from an error in meiosis I and $m = m_1 + m_2$. Similarly, we define n_1 as being the number of non-reduced triploid individuals and n_2 being the number of reduced triploid individuals at a specific marker resulting from an error in meiosis II and $n = n_1 + n_2$. Then we calculated the maximum likelihood estimate of y, the probability of a recombinant meiotic tetrad, by solving the equation $(m + n) * y^2 - (3 * (m + n) - (2m_1 + n_2)) * y + 2 * (m_2 + n_1) = 0$. The variance in y is given by $Var(y) = y * (1 - y) * (2 - y) / (n + (m + n) * (1 - y))$ (Chakravarti *et al.* 1989). By assuming complete cross-over interference y can be translated into the marker-centromere distance (w; in cM) with $w = y / 2 * 100$ (Chakravarti & Slaugenhaupt 1987). The variance in w is given by $Var(w) = Var(y) / 4 * 100$ (Deka *et al.* 1990).

The locations of several microsatellite markers were not covered by the linkage map. Thus, we inferred the genetic location of those microsatellites by extrapolating linearly from the closest two markers in the linkage map.

Results

Parent and mechanism of origin

20 out of the 37 triploid individuals inherited the supernumerary haploid chromosome set from their mother with at least three markers per individual indicating an error in the maternal meiosis (and no marker indicating an error in the paternal meiosis; Supplementary Table S2). Two additional individuals were tetraploid, with one additional chromosome set passed on from their mother and one from their father (dispermy). For the purpose of this study, we will refer to these two individuals subsequently as digynic triploid since the additional paternal chromosome set is not of relevance for centromere mapping.

In 12 out of the 22 digynic triploid cases the error occurred in the first meiotic division and in 10 cases it occurred in the second meiotic division (Table 1). One of those ten triploid individuals (individual G8-3-4) could not be assigned with maximal confidence to be the result of an error in meiosis II. Since seven out of eight informative chromosomes indicated an error in the second meiotic division, a cross-over on chromosome *Tgu8* between the centromeric marker (located 1.38Mb from the chromosome end) and the centromere seems the most parsimonious explanation (Table 1). As expected, all triploid individuals resulting from an error in female meiosis I inherited both a Z and a W chromosome and those triploids originating from an error in the second meiotic division either got two Z or presumably two W chromosomes (we have no markers on chromosome *TguW* to prove the

presence of two W chromosomes). For two individuals the status of the sex chromosomes could not be inferred unambiguously but was consistent with the expectation (individuals B2013_088 and G8-3-4; Table 1).

Centromere positions and comparison to the linkage map

Female non-disjunction in both the first and second meiotic division is informative for centromere mapping. For each chromosome we had 2–16 informative meioses at the centromere and in total 8–39 informative inheritance events at the centromere and the distal end taken together. We hence were able to determine the location of centromeres on all but three chromosomes (these three were chromosomes *Tgu1B* and *Tgu16* and *Tgu27*; Figure 2 and Table 2). The centromeric markers on chromosomes *Tgu8*, *Tgu13*, *Tgu21* and *Tgu25* were not completely linked to the centromere (especially on chromosome *Tgu13*), yet the distal end markers contained enough information to localize the centromere unambiguously (the two markers on these chromosomes were at least 18.34 cM apart).

We compared the genetic distance between the two markers on each chromosome estimated from the published linkage map with the here estimated genetic distance from the centromere-marker-mapping. The correlation was highly significant (Pearson's r = 0.75, 95% confidence interval 0.53–0.88, df = 26, $P=4*10^{-6}$; Figure 3), even though the estimated genetic distances between markers from the centromere-marker-mapping are restricted to be maximally 50 cM (see Methods).

Discussion

We here make use of naturally occurring triploid zebra finches to map the location of centromeres in reference to the physical genome assembly. By using centromere-marker-mapping techniques we were able to map the centromere position on almost all of the 32 assembled chromosomes in the current zebra finch genome assembly (WUSTL v3.2.4).

Triploidy is one of the most common chromosome abnormalities in spontaneous human abortions, estimated to occur in 1–2% of all conceptions (Jacobs *et al.* 1982). In their study on zebra finches Forstmeier & Ellegren (2010) found 4 triploids among 331 embryos that died during development (1.2%) and in an independent sample from the same population of 115 embryos that died during development we found 3 triploids (2.6%), which is not significantly different from the first estimate (Fisher's exact test P = 0.38) and similar to rates found in chicken (1.6% vs 2.7% in zebra finch and chicken (Thorne *et al.* 1991), respectively, Fisher's exact test P = 0.25).

In humans, the relative importance of diandric to digynic triploidies is still a matter of debate, probably resulting from ascertainment bias and differing sampling schemes (Zaragoza *et al.* 2000). Estimates for diandric origin range from around 20% to 89% of all

triploidies with a mean of 64.4% (Joergensen *et al.* 2014; McFadden *et al.* 1993; McFadden & Langlois 2000; McFadden & Robinson 2006 and references therein). Digynic triploidies result from errors both in the first and second meiotic division with a slight bias towards errors in the second meiotic division (51 vs 63 cases, respectively; calculated from Joergensen *et al.* 2014; McFadden *et al.* 1993; McFadden & Langlois 2000; McFadden & Robinson 2006 and references therein). In our sample of 34 triploid zebra finches, 41.2% had a diandric origin and the 22 digynic triploidies resulted from errors in meiosis I and II with about equal frequencies. Thus, digynic meiotic errors may be a more common cause of triploidy in zebra finches than in humans (Fisher's exact test P = 0.004), which is also the case in chicken (Fechheimer 1981). Rates of meiosis I and II errors were similar between humans and zebra finches, contrasting results in chicken where errors in meiosis II predominate (Bloom 1972; Fechheimer 1981; Thorne *et al.* 1991).

For chromosome *Tgu9* we had only two informative meioses at the centromere, yet these two meioses indicated perfect linkage of our marker with the centromere and the second marker on chromosome *Tgu9* is located 39 cM away from the centromere. Chromosome *Tgu9* is acrocentric (Pigozzi 2008) and its linkage map spans almost the whole assembled chromosome (Backström *et al.* 2010). Thus, the genetic positions from the centromere-marker-mapping should correspond to the linkage map positions in Backström *et al.* (2010), and they agree reasonably well (0 cM vs 0.3 cM and 39 cM vs 55 cM for the first and second microsatellite marker, respectively). Thus, also for chromosome *Tgu9* we are confident that we localized the centromere at the correct end of the chromosome. All other chromosomes had at least four informative meioses at the centromere, indicating high reliability of our mapping results.

From estimates of the repeat content of the zebra finch genome it seems possible that the main satellite sequences are still missing from the genome assembly (Warren *et al.* 2010). In line with this, the smallest microchromosomes in chicken are at least 3.4 Mb in size as measured by pulse-field electrophoresis (Pichugin *et al.* 2001) and this is probably also true for the zebra finch given that the karyotype and genome size between the two species are highly conserved (Peterson *et al.* 1994; Pigozzi & Solari 1998). There are three chromosomes which are shorter than 3.4 Mb in the zebra finch reference genome, namely *Tgu1B*, *Tgu16* and *Tgu25* (1.08 Mb, 9 kb and 1.28 Mb, respectively). For chromosome *Tgu16* we did not even attempt to map the centromere since more than 99% of this difficult-to-assemble chromosome is missing from the assembly. Chromosome *Tgu1B*, in our linkage map, appears to be linked to chromosome *Tgu1* (Backström *et al.* 2010) and it thus might not even have a centromere. Our failure to map its centromere might be regarded as further support that chromosome *Tgu1B* is not an independent chromosome. On chromosome *Tgu25* our centromeric marker was estimated to be 5.3 cM away from the actual centromere. Given that some parts of the chromosome are absent from the reference assembly (the terminal marker in our linkage map at the centromeric side of the

chromosome maps to *Tgu25_random*; Backström *et al.* 2010), even markers at the end of each chromosome could be separated from the centromere and cross-overs may occasionally happen between marker and centromere. This is probably also the case for chromosomes *Tgu8*, *Tgu13* and *Tgu21*. For chromosomes *Tgu8* and particularly *Tgu13* we have direct evidence from our linkage map that at the centromeric side of the chromosome parts are missing in the reference assembly because the terminal markers are located on *Tgu8_random* and *Tgu13_random*, respectively (Backström *et al.* 2010). Chromosome *Tgu21* is among those chromosomes with the most amount of sequence unordered on its random chromosome, indicating that also sequence between the centromere and the marker may be missing.

Similarly, even though the microsatellite markers on the acrocentric chromosomes were located at most only 1.91 Mb from the chromosome ends, several chromosomes are estimated to be genetically shorter than 50 cM, which is the minimum genetic size of a chromosome, since at least one cross-over is required for proper chromosome segregation (Petronczki *et al.* 2003). First, this could be due to the fact that there is subtelomeric sequence missing in the current genome assembly (Warren *et al.* 2010). Second, with only two markers per chromosome we were unable to identify double cross-overs and could not distinguish single cross-overs from triple cross-overs which leads to an underestimation of the genetic length of a chromosome (Danzmann & Gharbi 2001).

In summary, we here report the approximate location of an additional 19 centromeres in the zebra finch reference genome, meaning that in total 29 of the 32 assembled chromosomes can now be oriented according to their centromere position. The newly developed centromeric and distal telomeric microsatellite markers can now be used for studies in which the centromere position is of crucial importance, e.g. in studies of meiotic drive or species divergence (Axelsson *et al.* 2010; Ellegren *et al.* 2012).

Acknowledgements

We are grateful to K. Martin, J. Rutkowska and M. Ihle for help with breeding and collecting the embryo samples and S. Bauer, E. Bodendorfer, A. Grötsch, A. Kortner, K. Martin, P. Neubauer, F. Weigel, and B. Wörle for animal care and help with breeding. M. Pigozzi generously provided her time and expertise to discuss cytogenetic methods and concepts. We thank B. Kempenaers for providing facilities and various other support. This study was funded by financial support from the Max Planck Society. U.K. is part of the International Max Planck Research School for Organismal Biology.

References

Axelsson E, Albrechtsen A, Van AP*, et al.* (2010) Segregation distortion in chicken and the evolutionary consequences of female meiotic drive in birds. *Heredity* **105**, 290–298.

Backström N, Forstmeier W, Schielzeth H*, et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research* **20**, 485–495.

Bloom SE (1972) Chromosome abnormalities in chicken (*Gallus domesticus*) embryos: types, frequencies and phenotypic effects. *Chromosoma* **37**, 309–326.

Calderón PL, Pigozzi MI (2006) MLH1-focus mapping in birds shows equal recombination between sexes and diversity of crossover patterns. *Chromosome Research* **14**, 605–612.

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.

Chakravarti A, Majumder PP, Slaugenhaupt SA*, et al.* (1989) Gene-centromere mapping and the study of non-disjunction in autosomal trisomies and ovarian teratomas. *Prog Clin Biol Res* **311**, 45–79.

Chakravarti A, Slaugenhaupt SA (1987) Methods for studying recombination on chromosomes that undergo nondisjunction. *Genomics* **1**, 35–42.

Christidis L (1986) Chromosomal evolution within the family Estrildidae (Aves) 1. The Poephilae. *Genetica* **71**, 81–97.

Côté GB, Edwards JH (1975) Centrometric linkage in autosomal trisomies. *Annals of Human Genetics* **39**, 51–59.

Danzmann RG, Gharbi K (2001) Gene mapping in fishes: a means to an end. *Genetica* **111**, 3–23.

Deka R, Chakravarti A, Surti U*, et al.* (1990) Genetics and biology of human ovarian teratomas. II. Molecular analysis of origin of nondisjunction and gene-centromere mapping of chromosome I markers. *American Journal of Human Genetics* **47**, 644–655.

Ekblom R, Stapley J, Ball AD*, et al.* (2011) Genetic mapping of the major histocompatibility complex in the zebra finch (*Taeniopygia guttata*). *Immunogenetics* **63**, 523–530.

Ellegren H, Fridolfsson AK (1997) Male-driven evolution of DNA sequences in birds. *Nature Genetics* **17**, 182–184.

Ellegren H, Smeds L, Burri R*, et al.* (2012) The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* **491**, 756–760.

Fechheimer NS (1981) Origins of heteroploidy in chicken embryos. *Poultry Science* **60**, 1365–1371.

Forstmeier W, Ellegren H (2010) Trisomy and triploidy are sources of embryo mortality in the zebra finch. *Proceedings of the Royal Society B-Biological Sciences* **277**, 2655–2660.

Forstmeier W, Segelbacher G, Mueller JC, Kempenaers B (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular Ecology* **16**, 4039–4050.

Girndt A, Knief U, Forstmeier W, Kempenaers B (2014) Triploid ZZZ zebra finches *Taeniopygia guttata* exhibit abnormal sperm heads and poor reproductive performance. *Ibis* **156**, 472–477.

Itoh Y, Arnold AP (2005) Chromosomal polymorphism and comparative painting analysis in the zebra finch. *Chromosome Research* **13**, 47–56.

Jacobs PA, Morton NE (1977) Origin of human trisomics and polyploids. *Hum Hered* **27**, 59–72.

Jacobs PA, Szulman AE, Funkhouser J, Matsuura JS, Wilson CC (1982) Human triploidy: relationship between parental origin of the additional haploid complement and development of partial hydatidiform mole. *Annals of Human Genetics* **46**, 223–231.

Joergensen MW, Niemann I, Rasmussen AA*, et al.* (2014) Triploid pregnancies, genetic and clinical features of 158 cases. *Am J Obstet Gynecol*.

Johnson SL, Gates MA, Johnson M*, et al.* (1996) Centromere-linkage analysis and consolidation of the zebrafish genetic map. *Genetics* **142**, 1277–1288.

Krasikova A, Deryusheva S, Galkina S*, et al.* (2006) On the positions of centromeres in chicken lampbrush chromosomes. *Chromosome Research* **14**, 777–789.

Mather K (1938) Crossing-over. *Biological Reviews of the Cambridge Philosophical Society* **13**, 252–292.

McFadden DE, Kwong LC, Yam IYL, Langlois S (1993) Parental origin of triploidy in human fetuses: evidence for genomic imprinting. *Human Genetics* **92**, 465–469.

McFadden DE, Langlois S (2000) Parental and meiotic origin of triploidy in the embryonic and fetal periods. *Clinical Genetics* **58**, 192–200.

McFadden DE, Robinson WP (2006) Phenotype of triploid embryos. *Journal of Medical Genetics* **43**, 609–612.

Peterson DG, Stack SM, Healy JL, Donohoe BS, Anderson LK (1994) The relationship between synaptonemal complex length and genome size in four vertebrate classes (Osteicthyes, Reptilia, Aves, Mammalia). *Chromosome Research* **2**, 153–162.

Petronczki M, Siomos MF, Nasmyth K (2003) Un ménage à quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112**, 423–440.

Pichugin AM, Galkina SA, Potekhin AA*, et al.* (2001) Estimation of the minimal size of chicken *Gallus gallus domesticus* microchromosomes via pulsed-field electrophoresis. *Russian Journal of Genetics* **37**, 535–538.

Pigozzi MI (2008) Relationship between physical and genetic distances along the zebra finch Z chromosome. *Chromosome Research* **16**, 839–849.

Pigozzi MI, Solari AJ (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.

Shang WH, Hori T, Toyoda A*, et al.* (2010) Chickens possess centromeres with both extended tandem repeats and short non-tandem-repetitive sequences. *Genome Research* **20**, 1219–1228.

Stapley J, Birkhead TR, Burke T, Slate J (2008) A linkage map of the zebra finch *Taeniopygia guttata* provides new insights into avian genome evolution. *Genetics* **179**, 651–667.

Thorne MH, Collins RK, Sheldon BL (1991) Chromosome analysis of early embryonic mortality in layer and broiler chickens. *British Poultry Science* **32**, 711–722.

Warren WC, Clayton DF, Ellegren H*, et al.* (2010) The genome of a songbird. *Nature* **464**, 757–762.

Zaragoza MV, Surti T, Redline RW*, et al.* (2000) Parental origin and phenotype of triploidy in spontaneous abortions: predominance of diandry and association with the partial hydatidiform mole. *American Journal of Human Genetics* **66**, 1807–1820.

Zhao HY, Speed TP (1998) Statistical analysis of half-tetrads. *Genetics* **150**, 473–485.

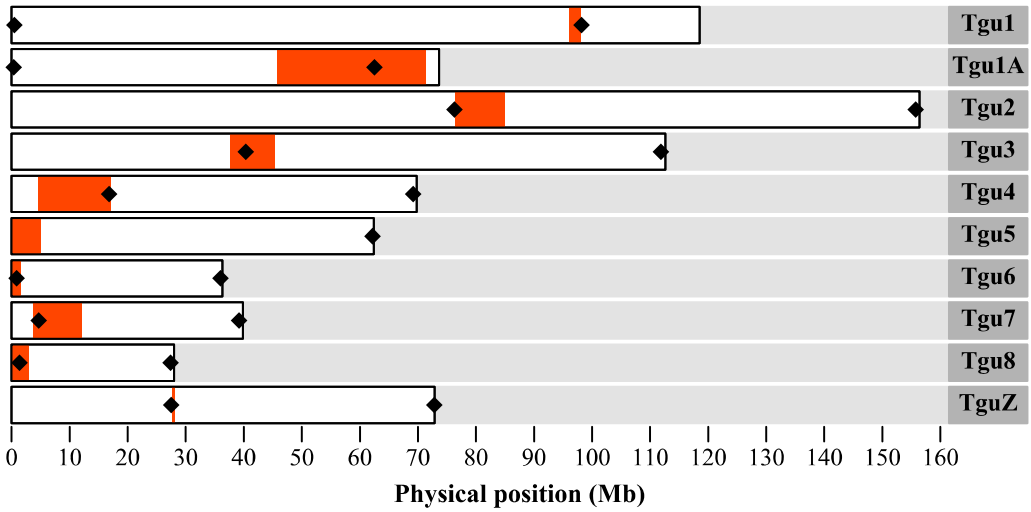


Figure 1: *The ten chromosomes in the current zebra finch genome assembly (WUSTL v3.2.4) with a known centromere position in reference to the physical map. The centromere positions have been inferred by FISH (Warren et al. 2010). The intervals between the FISH-probes closest flanking the centromeres are indicated in red. Black diamonds indicate the positions of microsatellite markers used in this study. On chromosome Tgu5 the microsatellite marker proximal to the centromere is located on Tgu5_random and thus not indicated in the figure (see main text for an explanation). Chromosome nomenclature follows the one introduced by Warren et al. (2010) and Itoh & Arnold (2005).*

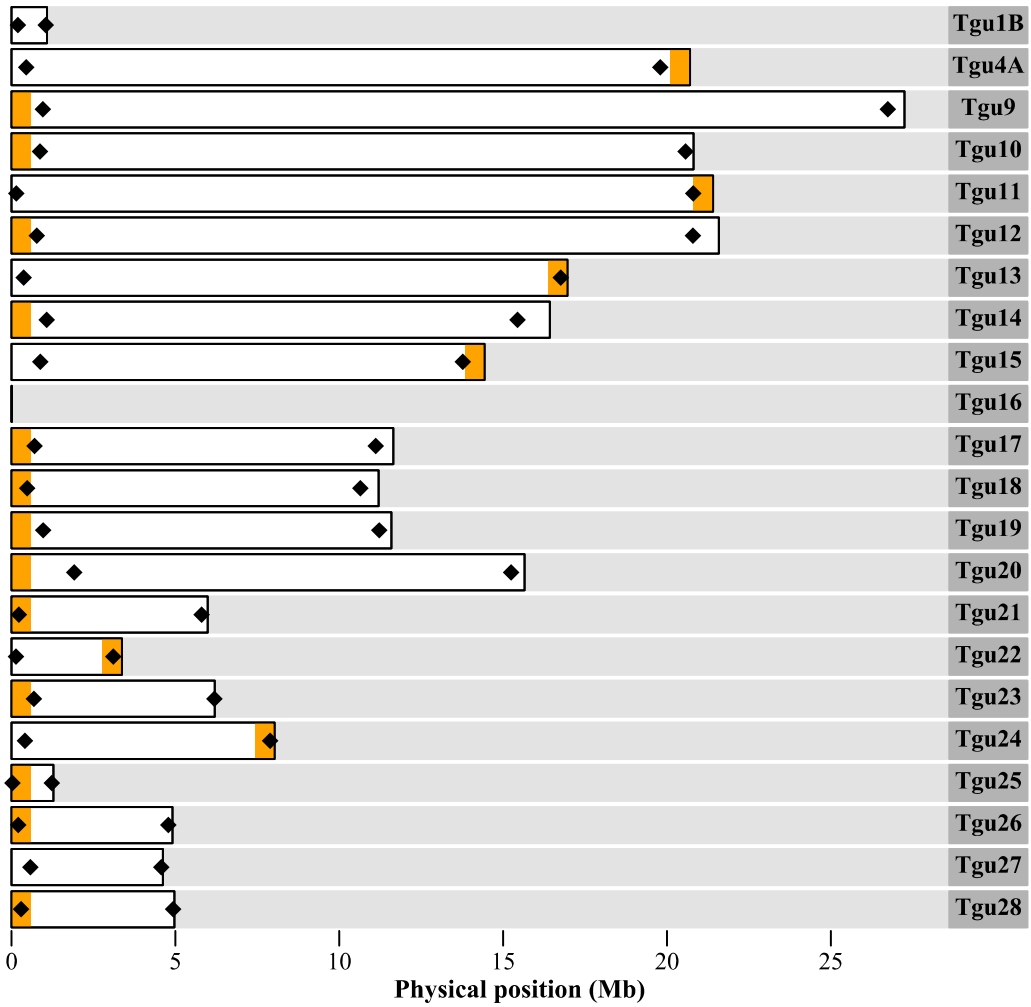


Figure 2: *The 22 acrocentric chromosomes in the current zebra finch genome assembly (WUSTL v3.2.4) with an unknown centromere position in reference to the physical map. For 19 of these chromosomes the positions of the centromeres were mapped and are indicated in orange. For clarity, each centromere position is indicated by a 600kb wide interval, which does not reflect the true extent of the centromere though. Black diamonds indicate the positions of microsatellite markers used in this study. Chromosome nomenclature follows the one introduced by Warren et al. (2010) and Itoh & Arnold (2005).*

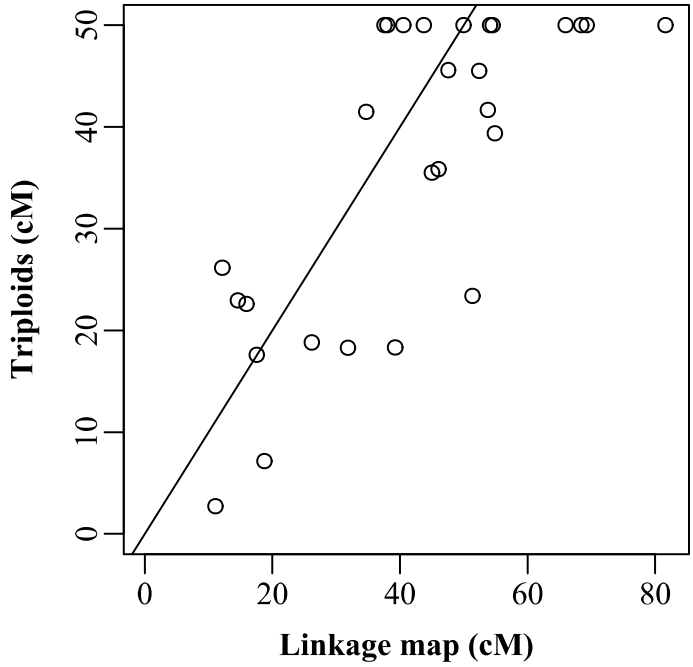


Figure 3: *Comparison of genetic map distances between the two microsatellite markers for each chromosome taking estimates from the linkage map (Backström et al. 2010) and estimates from the here presented marker-centromere-mapping using triploids. The line represents the diagonal line of equality.*

Table 1: *Digynic triploid individuals resulting from non-disjunction in the first or second meiotic division with their sex chromosome karyotype and the numbers of reduced and non-reduced markers at known centromeres and distal ends (from nine chromosomes as in Figure 1 except chromosome TguZ). Bold print highlights the key observation for inferring errors in the first meiotic division (MI) versus the second meiotic division (MII). These individuals and the information about MI or MII errors were subsequently used for mapping the location of centromeres on additional chromosomes.*

Individual ID	Sex chromosome†	Known centromere			Known distal end			Meiotic error
		not-reduced	reduced	not informative	not-reduced	reduced	not informative	
B2012_130	*ZZW*	6	**0**	3	1	2	6	MI
B2011_258a	*ZZW*	5	**0**	4	4	4	1	MI
B2013_088	*ZZW/ZWW*	5	**0**	4	4	0	5	MI
B2013_207	*ZZW*	5	**0**	4	4	0	5	MI
B2012_129	*ZZW*	5	**0**	4	1	2	6	MI
B2013_086	*ZZW*	4	**0**	5	2	4	3	MI
B2013_198	*ZZW*	4	**0**	5	3	2	4	MI
2006_486	*ZZW*	4	**0**	5	3	1	5	MI
2011_328	*ZZW*	4	**0**	5	3	1	5	MI
2011_183	*ZZW*	4	**0**	5	1	2	6	MI
B2011_017	*ZZW*	1	**0**	8	4	1	4	MI
2011_180	*ZZZW*	3	**0**	6	5	2	2	MI + Polyspermy
2006_550	*ZZZ*	**0**	7	2	5	1	3	MII
B2013_236	*ZWW*	**0**	7	2	3	2	4	MII
B2011_187	*ZWW*	**0**	6	3	5	3	1	MII
B2013_227	*ZWW*	**0**	6	3	2	4	3	MII
G12-1-1	*ZWW*	**0**	6	3	4	2	3	MII
2011_205	*ZZZ*	**0**	5	4	4	2	3	MII
2005_118	*ZWW*	**0**	5	4	4	1	4	MII
2011_308	*ZZZ*	**0**	4	5	3	2	4	MII

G8-3-4	*ZWW/ZZW*	**1***	7	1	6	3	0	MII
2011_251	*ZZZZ*	**0**	5	4	2	1	6	MII + Polyspermy

* Since seven chromosomes indicate an error in meiosis II, a cross-over between the marker and the centromere on chromosome *Tgu8* is the most parsimonious explanation.

† In contrast to the Z chromosome, we only have markers that determine the presence of a W chromosome but not any polymorphic markers to distinguish the presence of one versus two W chromosomes, so the latter was inferred by logic whenever possible.

Table 2: *Physical and genetic position of all microsatellite markers used in this study. The genetic position and its standard deviation are in reference to the centromere and were calculated from the numbers of reduced and not-reduced chromosomes in meiosis I and II (Chakravarti et al. 1989). A genetic position of 0 cM indicates complete linkage to the centromere. Bold print highlights the key observation for inferring linkage to the centromere.*

Chromosome	Marker	Position (Mb)	MI error			MII error			Position (cM ± SD)	Linked cytogenetic feature
			not-reduced	reduced	not informative	not-reduced	reduced	not informative		
Tgu1	1_cen_98.17	98.17	4	**0**	8	**0**	4	6	0.00	Centromere
Tgu1A	1A_cen_62.53	62.53	7	**0**	5	**0**	4	6	0.00	Centromere
Tgu2	2_cen_76.29	76.29	1	**0**	11	**0**	7	3	0.00	Centromere
Tgu3	3_cen_40.34	40.34	7	**0**	5	**0**	9	1	0.00	Centromere
Tgu4	4_cen_16.82	16.82	6	**0**	6	**0**	8	2	0.00	Centromere
Tgu4A	4A_en_19.79	19.79	1	**0**	11	**0**	3	7	0.00	Centromere
Tgu5_random	5rand_cen_0.26	0.26	7	**0**	5	**0**	8	2	0.00	Centromere
Tgu6	6_cen_0.89	0.89	6	**0**	6	**0**	5	5	0.00	Centromere
Tgu7	7_cen_4.65	4.65	2	**0**	10	**0**	8	2	0.00	Centromere
Tgu8	8_cen_1.38	1.38	10	**0**	2	**1**	5	4	4.65 ± 4.43	Centromere
Tgu9	9_st_0.96	0.96	2	**0**	10	**0**	0	10	0.00	Centromere
Tgu10	10_st_0.86	0.86	10	**0**	2	**0**	6	4	0.00	Centromere
Tgu11	11_en_20.8	20.8	4	**0**	8	**0**	7	3	0.00	Centromere
Tgu12	12_st_0.77	0.77	4	**0**	8	**0**	5	5	0.00	Centromere
Tgu13	13_en_16.75	16.75	3	**5**	4	**2**	5	3	27.04 ± 8.08	Centromere
Tgu14	14_st_1.07	1.07	6	**0**	6	**0**	8	2	0.00	Centromere
Tgu15	15_en_13.76	13.76	1	**0**	11	**0**	3	7	0.00	Centromere
Tgu17	17_st_0.7	0.7	8	**0**	4	**0**	7	3	0.00	Centromere
Tgu18	18_st_0.48	0.48	1	**0**	11	**0**	6	4	0.00	Centromere
Tgu19	19_st_0.97	0.97	5	**0**	7	**0**	3	7	0.00	Centromere

Tgu20	20_st_1.91	1.91	2	**0**	10	**0**	5	5	0.00	Centromere
Tgu21	21_st_0.22	0.22	6	**1**	5	**1**	7	2	8.53 ± 5.63	Centromere
Tgu22	22_en_3.1	3.1	5	**0**	7	**0**	1	9	0.00	Centromere
Tgu23	23_st_0.68	0.68	9	**0**	3	**0**	6	4	0.00	Centromere
Tgu24	24_en_7.89	7.89	5	**0**	7	**0**	4	6	0.00	Centromere
Tgu25	25_st_0.03	0.03	3	**1**	8	**0**	7	3	5.31 ± 5.17	Centromere
Tgu26	26_st_0.2	0.2	3	**0**	9	**0**	5	5	0.00	Centromere
Tgu28	28_st_0.29	0.29	9	**0**	3	**0**	6	4	0.00	Centromere
TguZ	Z_cen_27.51	27.51	3	**0**	9	**0**	4	6	0.00	Centromere
Tgu1	1_st_0.48	0.48	5	2	5	2	5	3	18.28 ± 7.72	Distal end
Tgu1A	1A_st_0.38	0.38	8	2	2	5	1	4	35.84 ± 7.86	Distal end
Tgu2	2_en_155.77	155.77	6	0	6	4	5	1	17.61 ± 7.09	Distal end
Tgu3	3_en_111.84	111.84	2	1	9	3	3	4	26.16 ± 9.46	Distal end
Tgu4	4_en_69.2	69.2	3	2	7	3	5	2	22.60 ± 7.96	Distal end
Tgu4A	4A_st_0.45	0.45	4	5	3	3	0	7	50.00	Distal end
Tgu5	5_en_62.17	62.17	5	5	2	7	0	3	50.00	Distal end
Tgu6	6_en_35.99	35.99	3	2	7	7	0	3	50.00	Distal end
Tgu7	7_en_39.18	39.18	2	6	4	6	1	3	45.58 ± 5.13	Distal end
Tgu8	8_en_27.41	27.41	1	1	10	1	2	7	22.98 ± 12.95	Distal end
Tgu9	9_en_26.74	26.74	2	0	10	5	1	4	39.37 ± 8.12	Distal end
Tgu10	10_en_20.56	20.56	2	5	5	6	0	4	50.00	Distal end
Tgu11	11_st_0.14	0.14	3	4	5	5	3	2	35.51 ± 7.33	Distal end
Tgu12	12_en_20.79*	20.79	0	1	11	0	0	10	50.00	Distal end
Tgu13	13_st_0.37	0.37	2	6	4	7	0	3	50.00	Distal end
Tgu14	14_en_15.44	15.44	5	6	1	6	0	4	50.00	Distal end
Tgu15	15_st_0.88	0.88	5	4	3	6	0	4	50.00	Distal end
Tgu17	17_en_11.11	11.11	0	3	9	2	1	7	41.67 ± 10.06	Distal end
Tgu18	18_en_10.64	10.64	1	4	7	4	0	6	50.00	Distal end
Tgu19	19_en_11.22	11.22	3	5	4	3	0	7	50.00	Distal end
Tgu20	20_en_15.24	15.24	1	3	8	5	0	5	50.00	Distal end

Tgu21	21_en_5.8	5.8	0	2	10	2	0	8	50.00	Distal end
Tgu22	22_st_0.13	0.13	4	2	6	8	0	2	50.00	Distal end
Tgu23	23_en_6.19	6.19	2	4	6	1	5	4	23.38 ± 8.77	Distal end
Tgu24	24_st_0.41	0.41	4	5	3	7	0	3	50.00	Distal end
Tgu25	25_en_1.22	1.22	4	0	8	5	4	1	24.12 ± 7.76	Distal end
Tgu26	26_en_4.78	4.78	1	6	5	5	1	4	45.52 ± 5.57	Distal end
Tgu28	28_en_4.93									Distal end
TguZ	Z_en_72.81	72.81	10	0	2	0	9	1	0.00	Distal end
Tgu1B	1B_en_1.05	1.05	5	4	3	2	5	3	23.57 ± 7.85	
Tgu1B	1B_st_0.19	0.19	1	7	4	0	6	4	26.27 ± 8.53	
Tgu27	27_st_0.58	0.58	1	1	10	0	0	10	50.00	
Tgu27	27_en_4.57	4.57	0	5	7	0	0	10	50.00	

* The microsatellite 12_en_20.79 is duplicated in the genome. Since we do not know whether the second copy is also located on chromosome *Tgu12*, it is not possible to infer triploidy by the occurrence of all alleles from a parent. Yet the marker is informative if only a single allele gets inherited (because then it was reduced).

Chapter 6: Association mapping of morphological traits in wild and captive zebra finches: reliable within but not between populations

Abstract

Identifying causal genetic variants underlying heritable phenotypic variation is a longstanding goal in evolutionary genetics. In previous linkage analyses we identified several quantitative trait loci (QTLs) in the genome of a captive population of zebra finches (Taeniopygia guttata) that were associated with variation in wing and tarsus length, beak morphology, digit ratio and body mass. We here follow up on these studies with the aim to identify quantitative trait nucleotides (QTNs) in one wild and four captive populations. First, we performed an association study using 672 SNPs within candidate genes located in the previously identified QTL regions in a sample of 939 wild-caught zebra finches. Then, we validated the most promising SNP-phenotype-associations (n = 25 SNPs) in 5,229 birds from four captive populations. Genotype-phenotype associations were generally weak in the wild population, where linkage disequilibrium (LD) spans only short genomic distances. In contrast, in captive populations, where LD blocks are large, SNP-effects on morphological traits were highly repeatable within the same population (tested within the captive population used for QTL mapping). Most of those SNPs also showed significant associations with the same trait in other captive populations, but the direction and magnitude of these effects differed. This suggests that the tested SNPs are not the causal QTNs but rather are physically linked to them, and that LD between SNP and causal variant differs between populations due to founder effects (genetic drift). We conclude that the identification of causal variants remains challenging in non-model organisms, even with large genotyping efforts, because effect sizes of single variants on quantitative traits are typically small.

Prepared as: Knief U, G Hemmrich-Stanisak, M Wittig, A Franke, SC Griffith, B Kempenaers, W Forstmeier: Association mapping of morphological traits in wild and captive zebra finches: reliable within but not between populations.

Introduction

Identifying the functional genetic variants underlying heritable phenotypic variation (quantitative trait nucleotides, QTN) in natural populations is a primary goal in evolutionary genetics (Feder & Mitchell-Olds 2003; Slate 2005; Slate *et al.* 2009). Knowledge of the loci underlying natural variation in phenotypic traits can offer insights reaching beyond those gained by classical quantitative genetics. For example, Slate *et al.* (2010) formulated several research questions that can be only be addressed after the causal variants have been identified, such as: Do the same or linked QTN cause pleiotropy? Do the same QTN contribute to adaptive evolution in different populations? How frequent and strong are gene-by-environment interactions at individual QTN? Ultimately, addressing these questions will lead to a better understanding of the maintenance of variation within populations (e.g. Johnston *et al.* 2013) and of how populations respond to selection (Robinson *et al.* 2014).

In previous studies on the zebra finch (*Taeniopygia guttata*), we have successfully used whole-genome linkage analyses to identify genomic regions that are associated with highly heritable phenotypic traits (QTL regions for beak color, Schielzeth *et al.* 2011; wing length, Schielzeth *et al.* 2012; beak morphology, Knief *et al.* 2012; unpublished data). Linkage analyses generally suffer from poor resolution (Slate *et al.* 2009), which is especially pronounced in the zebra finch, because crossover rates are heavily skewed towards the telomeres on the macrochromosomes (Backström *et al.* 2010). Accordingly, although these skewed recombination rates lent high power to identify genomic regions associated with phenotypic variation, our linkage analyses identified large linkage blocks covering hundreds or even thousands of genes. Identifying the causal variants in our captive zebra finch population is difficult; it would require an unrealistically larger pedigree to get sufficient amounts of recombination to break up the linkage blocks in the center of the macrochromosomes.

Alternatively, association studies can be performed that do not rely on familiar inheritance patterns but rather on linkage disequilibrium (LD) between the causal and the genotyped variants (Lander & Schork 1994). By exploiting historical recombination events in a population (Lander & Schork 1994), association mapping in a sample of seemingly unrelated individuals offers great precision but is crucially dependent on sufficient LD between the typed markers and the causal variants (Pritchard & Przeworski 2001; Vasemägi & Primmer 2005). In wild zebra finches, LD is extremely low even within a range of 300 bp (Balakrishnan & Edwards 2009). Consequently, if a genotyped variant is statistically associated with a phenotypic trait in a sample of unrelated individuals, it is highly likely either the QTN itself or very close to it. However, to cover the whole genome (1.2 Gb, Warren *et al.* 2010) with a sufficiently dense marker panel requires genotyping millions of variants (Schielzeth & Husby 2014), imposing an immense multiple testing problem and an unrealistically large sample size, especially because effect sizes of single QTN are expected to be small (Park *et al.* 2010; Yang *et al.* 2010). Thus, a whole-genome association study in wild zebra finches is not

feasible because it would require hundreds of thousands of individuals to be phenotyped. As an alternative, we here focus on candidate genes that reside within the QTL regions previously identified in our captive population. This is a commonly advocated research strategy also in human genetics (Cardon & Bell 2001).

The aim of our study was to find the causal variants underlying previously identified QTL regions. To this end, we first identified around 23 million single nucleotide polymorphisms (SNPs) in a sample of 100 wild-caught zebra finches, and selected a subset of those SNPs that resided in candidate genes within the previously identified QTL regions for several morphological traits (which were identified in a captive population). We then genotyped this subset of SNPs in around 1,000 wild-caught zebra finches and performed association mapping (the most powerful method for fine-mapping when causative variants are included; Uleberg & Meuwissen 2010). Replication within and across populations is regarded as the gold standard of genetic association studies, because it leads to convincing statistical evidence, rules out bias and improves effect size estimates (Kraft *et al.* 2009; NCI-NHGRI working group on replication in association studies *et al.* 2007). Thus, in a last step we genotyped the most promising candidate SNPs in four captive populations of zebra finches, and performed an association analysis. Our study addresses the following questions. (1) How much variance is explained by a single candidate SNP within a previously identified QTL region? (2) Are QTL effects replicable within a population? (3) Are the most promising QTN from the wild population replicable across different populations of captive zebra finches? If so, it is highly likely that we identified the causal variants.

Material and Methods

Initial association study in wild zebra finches

Study population, phenotyping and sample collection of wild birds

We collected blood samples from 1,059 wild adult zebra finches (530 females, 529 males) at Fowlers Gap, NSW, Australia, in two locations (S 30°57’ E 141°46’ and S 31°04’ E 141°50’) during October/December 2010 and April/May 2011. In the following we refer to this population as “Fowlers Gap”. More details on the study sites and catching procedure using walk-in traps at feeders are given in Griffith *et al.* (2008) and Mariette & Griffith (2012). For each individual we measured the following traits (once, unless stated otherwise; Table 1). (1) Tarsus length (the length of the right tarsus from the bent foot to the proximal edge of the tarsometatarsus, including the intertarsal joint), measured to the nearest 0.1 mm using a ruler as described in Forstmeier *et al.* (2007; method 3). Birds with swollen joints or otherwise injured tarsi were removed from further analyses. (2) Wing length (maximum wing chord: flattened right wing from the carpal joint to the longest primary), measured to the nearest 0.5 mm using a wing ruler as described in Schielzeth *et al.* (2012). Birds whose primaries were obviously damaged were excluded from further analyses. (3) Beak length (from the proximal end of the rhamphotheca, i.e. the horny part of the beak, to the tip of

the culmen), beak depth (maximal depth at the base) and beak width (at upper mandible), measured to the nearest 0.01 mm using digital calipers as described in Knief *et al.* (2012).
(4) Digit ratio (the ratio of the length of the second to fourth digit of the right foot) as described in Forstmeier (2005). First, we measured both the second and fourth digit by aligning all three front toes on top of a ruler and pressing the basal pad of the hind toe gently against the edge of the ruler, while holding the hind toe perpendicular to the other three toes. Second, we measured the length difference between the second and fourth digit directly by holding the bird with its ventral side up, stretching its right leg and aligning the second and fourth digit, gently pushing the third toe further dorsally and measuring the distance between the distal tip of the two digits (excluding the claw). Using digital calipers, we measured each bird twice around 5 minutes apart to reduce memory effects. We then averaged each of the two measurements and calculated the digit ratio as the mean of ((length of fourth toe - length difference) / length of fourth toe) and (length of second tow / length of fourth toe). (5) Body mass, measured to the nearest 0.1 g using a digital scale (Table 1).

Juvenile zebra finches develop their full adult plumage until an age of around 100 to 110 days (Zann 1996). Thus, we scored each bird's age according to its appearance (based on its beak and eye color and the extent of plumage ornaments). Birds that appeared to be fully grown were assigned an age of 120 days. Measurements were taken by U.K. and W.F, but only birds measured by U.K. were included in the initial study (see below).

Selection of candidate genes
We selected candidate genes in QTL regions that were identified in previous studies on a population of zebra finches held at the Max Planck Institute for Ornithology in Seewiesen, Germany (in Supplement; Knief *et al.* 2012; Schielzeth *et al.* 2012). Candidate genes were selected based on published associations with particular phenotypes in chicken (*Gallus gallus*) and other bird species (Table S1). Because we could not genotype all SNPs within a candidate gene, we restricted candidate SNPs to promoter sequences, 3' and 5' UTR regions, exons and regions that are conserved compared to other vertebrates. To determine conserved regions we used the chicken genome assembly WUGSC2.1/galGal3 (Hillier *et al.* 2004) and extracted sequences reaching from 10 kb upstream to 5 kb downstream of the target gene that were conserved between seven vertebrate species according to a phylogenetic hidden Markov model ('phastCons7way' table in UCSC's genome browser). We used BLAT (v34x13; Kent 2002) to map these sequences to the zebra finch genome assembly (WUSTL v3.2.4; Warren *et al.* 2010) and filtered hits according to their position (within 10 kb upstream to 5 kb downstream of the target gene) and E-value (a measure of sequence identity; $E < 10^{-5}$). Because each BLAT-query can produce multiple hits in close proximity to each other, we kept the longest range, thus spanning also small regions that were not conserved between chicken and zebra finch (at maximum 4.4 kb, median = 26 bp). By focusing only on SNPs in promoters, exons, UTRs and conserved regions we tried to enrich

for SNPs with strong effects on either the transcription level or protein function of the candidate genes (see McCauley *et al.* (2007), who use multi-species conserved sequences for the same reason).

SNP discovery

We sequenced a pooled, non-barcoded sample of equal amounts of whole genomic DNA from 100 wild-caught individuals on the Illumina HiSeq 2000 platform (8 lanes = 1 flow cell) to obtain an initial dense SNP panel of around 23 million SNPs (Knief *et al.* 2015). Details of the SNP discovery pipeline are given in Knief *et al.* (2015). Briefly, reads were mapped using BWA (v0.5.9; Li & Durbin 2009; settings: bwa aln -n 4 -q 20 -l 5000) to the WUSTL v3.2.4 zebra finch genome assembly, which yielded an average genome coverage of 247.5 x calculated using the BEDTools' (v2.17.0; Quinlan & Hall 2010) coverageBed function after removing alignment gaps. Coverage was evenly distributed across regions. SNPs were called and filtered using GATK (v2.1-11-g13c0244; McKenna *et al.* 2010; UnifiedGenotyper settings: -stand_call_conf 50.0 -stand_emit_conf 10.0 -dcov 1000 -mbq 10 -mmq 10 -glm BOTH; for the VariantFiltration settings see Supplement).

SNP chip design

From the 23 million SNPs discovered in the pooled sequencing we selected 5,289 SNPs to be genotyped in all 1,059 individuals on an Illumina Infinium iSelect HD Custom BeadChip (n = 6,000 bead types). We included 884 SNPs that covered our candidate genes (Table S1) and another 4,405 SNPs that covered all assembled chromosomes (except chromosome *Tgu16* which only spans 9.9 kb in the current genome assembly). For the latter, we attempted to get at least 40 physically evenly spaced SNPs on each chromosome, yet this was not possible for chromosomes *Tgu1B* (n = 33 SNPs) and *Tgu25* (n = 24 SNPs) because too few SNPs passed our filtering procedure (see below). In order to reduce ascertainment bias, we did not exclude C/G and A/T SNPs which require two so-called 'bead types' for genotyping, which explains the discrepancy between the number of SNPs genotyped and the number of beads assayed on the chip.

We selected SNPs based on the following criteria.

(1) For the whole-genome SNP set, we removed all SNPs that either had another SNP or an Insertion/Deletion (InDel) in the flanking 60 bp (up- or downstream of the target SNP).
(2) For both the candidate-gene and whole-genome SNP set, we removed all SNPs that had more than ten ambiguous bases (IUPAC code 'N') in the flanking 60 bp (up- or downstream of the target SNP).
(3) After using BLAT (Kent 2002) on the up- and downstream sequences (60 bp) of each target SNP against the zebra finch genome (WUSTL v3.2.4 assembly), we removed from both the candidate-gene and whole-genome SNP set all SNPs with duplicated

flanking sequences, defined as an alignment length of more than 40 bp with either less than ten gap openings or less than ten mismatches.

(4) After running Illumina's Assay Design Tool (ADT) on the remaining SNPs, we removed all SNPs that could not be genotyped (error code ≤ 199) and those that had a success score of <0.8 (candidate gene SNP set) or <0.95 (whole-genome SNP set).

(5) For the whole-genome SNP set, we removed all SNPs that had more than ten bases marked as repeats in the flanking 60 bp (up- or downstream).

The final chip delivered by Illumina contained 4,553 SNPs (86% of the attempted SNPs, Figure S1). Drop-outs were randomly distributed along chromosomes (Supplement).

Median marker spacing of SNPs on the chip was 241.29 kb (interquartile range IQR = 11.71–342.28 kb) on macrochromosomes (*Tgu1*–*Tgu5*, *Tgu1A*), 267.65 kb (IQR = 50.31–356.06 kb) on microchromosomes (all other autosomes), and 174.63 kb (IQR = 161.11–179.40 kb) on chromosome *TguZ* (Figure S2).

Genotyping Illumina and quality control

Genotyping was done at the Institute of Clinical Molecular Biology (IKMB) at Kiel University, Germany. Cluster plots, which separate homozygous from heterozygous individuals, were automatically analyzed for each of the 4,553 SNPs using Illumina's GenomeStudio software (v2011.1, genotyping module 1.9.4) and manually inspected using R (v3.0.2; R Core Team 2013). We removed 111 individuals with a missing call rate larger than 0.05. These were mainly (92%) individuals measured by W.F., in which problems with DNA extraction occurred. We also removed the last nine individuals phenotyped by W.F. in order to have a data set in which all birds were measured by the same author (U.K.), leaving 939 individuals. The R package GWASTools (v1.6.2; Gogarten *et al.* 2012) was used for further quality checks (i.e., the occurrence of heterozygous deletions, batch effects) and to test for Hardy-Weinberg-disequilibrium (Ziegler *et al.* 2010). We also checked whether females were consistently called as homozygous for Z-linked SNPs, which was the case for all individuals. One SNP (*TguZ*: 1,768,144 bp), potentially located within the pseudo-autosomal region because several females were called as heterozygote (see also Stapley *et al.* 2008), was removed from further analyses because it could also be an assembly error. We removed an additional 151 SNPs because they did not form defined clusters (n = 42), had high missing call rates (missing rate > 0.1, n = 26), were monomorphic (n = 19), deviated strongly from Hardy-Weinberg equilibrium (Fisher's exact test $P < 0.05/4{,}553$, n = 63), or because they were Z-linked although the zebra finch genome assembly mapped them to chromosome *Tgu28* (n = 1, at *Tgu28*: 2,105,389 bp), resulting in 672 SNPs to be tested for phenotype-genotype-associations and 3,729 SNPs covering all assembled chromosomes.

The minor allele frequencies of the 4,401 SNPs that passed all quality checks were highly correlated with their minor allele counts from the pooled sequencing (Pearson's r = 0.96)

and deviations were mainly within the range of the expected sampling noise (Figure S3), suggesting that both the pooled sequencing and individual genotyping worked properly.

Kinship matrix construction

To control for unobserved relatedness between individuals from the wild we constructed a genetic similarity matrix (GSM, also called genetic relatedness matrix; Speed & Balding 2015), which was fitted as a random effect in our association models (see below). First, we filtered the autosomal markers to include only SNPs in approximate linkage equilibrium using the R package SNPRelate (v0.9.14; Zheng *et al.* 2012) with a composite linkage disequilibrium cutoff of 0.2 within 500 kb, resulting in 3138 autosomal SNPs. An inclusion of candidate markers in the GSM may lead to a loss of power ("proximal contamination"; Listgarten *et al.* 2012; Yang *et al.* 2014). Thus, for each trait we also removed all SNPs within candidate genes and applied Method 1 described in Robinson *et al.* (2013) to construct the GSM (which is equivalent to Method 2 in VanRaden 2008). Note that the resulting GSM is highly correlated (for all traits Pearson's $r \geq 0.93$) with the identical-by-state (IBS) allele-sharing matrix used by the efficient mixed-model association (EMMA) R package (v1.1.2; Kang *et al.* 2008), which has been shown to efficiently control for unobserved relatedness (Eu-ahsunthornwattana *et al.* 2014). The GSM has the advantage of being analogous to the numerator relationship matrix used in animal models for additive genetic variance component estimation (VanRaden 2008).

Statistical models and software

Genotype quality control and kinship matrix construction were done in R (v3.0.2; R Core Team 2013). For the association analyses we used ASReml-R (v3; Gilmour *et al.* 2009) to fit univariate mixed-effects linear models of the form: phenotypic trait ~ mean + fixed effects + SNP + GSM + error. Each SNP was either fitted as a covariate using one degree of freedom (additive effect) or as a factor using two degrees of freedom (dominance effects). As fixed effects we included age (covariate), sex (factor with two levels) and occasion (factor with two levels referring to the catching seasons November/December 2010 and April/May 2011) and for body mass we also included time of day (covariate). To reduce the impact of overall body size (Forstmeier 2011) we fitted body mass as a covariate for tarsus length, wing length and the three beak dimensions and in case of body mass being the independent variable we used tarsus length as the covariate. The GSM was fitted as a random effect. We estimated the heritability of each trait using the pin-module in ASReml-R after accounting for fixed effects. The variance explained by a SNP was calculated as

$$V_{SNP} = 2 \times p \times (1-p) \times \beta^2$$

where p is the minor allele frequency and β the estimated slope (i.e. the average effect of allelic substitution; Lynch & Walsh 1998) from the above mixed-effects model for the SNP-

effect, which takes additive and dominance effects into account (Falconer & Mackay 1996). We calculated the standard error of V_{SNP} as

$SE(V_{SNP}) = SE(\beta) \times \beta$.

To obtain the proportion of phenotypic variance explained by a SNP we divided V_{SNP} by the sum of the polygenic and residual variance of a null-model without fitting the SNP as a covariate, i.e. the phenotypic variance after controlling for all other fixed effects. Since SNPs within candidate genes showed varying degrees of linkage disequilibrium, we used the simpleM-algorithm with default settings to estimate the effective number of tests performed and used this estimate for controlling the type I error rate (Gao *et al.* 2008). For visualization of the association results we used log quantile-quantile P-value plots (PP-plots; e.g. Balding 2006). Assuming a uniform distribution of P-values in the interval [0, 1] as the null distribution, we plot the *i*th largest $-\log_{10}$(P-value) against its expectation, which is $-\log_{10}(i / (n + 1))$, where n is the number of SNPs tested. Confidence intervals can be obtained from a beta distribution (using the qbeta() function in R) assuming independent and identically distributed random variables (Casella & Berger 2002). In the presence of linkage disequilibrium these confidence intervals are conservative, i.e. they are too narrow. SNPs above the identity line in the top right corner of the PP-plot indicate significant associations.

Validation study in multiple captive populations

Validation study: populations and phenotyping

We followed-up on those candidate SNPs which showed the strongest associations with the respective phenotypes in the "Fowlers Gap" population (see below for the inclusion criteria), and genotyped them in additional 5,229 (mostly captive) zebra finches from four populations. (1) A captive population held at the Max Planck Institute for Ornithology in Seewiesen (n = 3,233 individuals; study population 18 in Forstmeier *et al.* 2007) with a complete pedigree covering eight generations, of which the last seven were genotyped. In the following, we refer to this population as "Seewiesen". The first four of the genotyped generations (n = 1,207) have been used for linkage map construction (Backström *et al.* 2010) and QTL-mapping (Knief *et al.* 2012; Schielzeth *et al.* 2012). (2) A recently wild-derived population held at the Max Planck Institute for Ornithology in Seewiesen (n = 1,096 individuals; originating from study population 4 in Forstmeier *et al.* 2007) with a complete pedigree covering six generations, of which the last four generations were genotyped. We refer to this population as "Bielefeld". (3) A population that was produced by crossing individuals from a captive population held in Cracow, Poland (study population 11 in Forstmeier *et al.* 2007) with the "Seewiesen" population (n = 634 individuals) with a complete pedigree covering three generations (all genotyped). We refer to this population as "Cracow". (4) Wild-caught birds (about half from Fowlers Gap and half from Sturt National Park, NSW) held at Macquarie University, Sydney, Australia, and additional wild

birds from Fowlers Gap (measured by W.F. in 2010, n = 265 individuals without pedigree information). In the following, we refer to these birds jointly as “Sydney”.

All birds were measured using the exact same methodology as for the wild birds. Each measurement was taken once (twice for digit ratio) per individual such that phenotypic values and their measurement errors are comparable between the initial and the validation study (Falconer & Mackay 1996). Descriptive statistics for each trait are summarized in Table 1.

Validation study: SNP selection, genotyping and quality control

We selected SNPs from the initial candidate gene scan that showed the strongest associations with the respective phenotypes in the “Fowlers Gap” population. We used a false-discovery (FDR) cut-off of 0.5, resulting in 26 SNPs to be genotyped in the validation sample of birds (for 29 associations, because four SNPs showed an association with two phenotypes). An FDR of 0.5 is rather high, but we expected *a priori* many associations to be true because the SNPs reside in previously identified QTL regions (Sham & Purcell 2014). We also included 41 SNPs for genotyping in the validation sample for a different purpose (will be described elsewhere).

We designed three Sequenom assays consisting of 62 SNPs according to the manufacturer’s users guide for the Sequenom MassARRAY iPLEX platform (Gabriel *et al.* 2009). Five additional SNPs were genotyped using the TaqMan assay (Holland *et al.* 1991) according to the manufacturer’s recommendations. All genotyping was done at the Institute of Clinical Molecular Biology (IKMB) at Kiel University. Genotypes were called using the MassARRAY Typer (v4.0) and TaqMan SDS (v2.4) software with standard settings. Genotyping of one SNP in the gene *RALDH3* failed, resulting in 25 SNPs for testing phenotype-genotype associations.

For all birds (except the ”Sydney” population) pedigree information was available and we checked the inheritance of every SNP using PedCheck (v1.00; O'Connell & Weeks 1998). Mendelian errors were mainly due to the occurrence of null-alleles and we inferred null-alleles such that no Mendelian inconsistencies remained. If null-alleles were present, we also marked those individuals for which the genotypes could not be inferred unambiguously (excluded, see below). After accounting for null-alleles the missing rate in the remainder of the SNPs was 1.62% (Figure S4, Figure S5).

Pedigree founders of the three captive populations were also included in the initial Illumina genotyping. This allowed us to compare the genotypes of 239 individuals that were genotyped on both the Illumina and the Sequenom platform. We found six inconsistencies out of 16,013 genotypes (0.037%) due to errors in the Illumina genotyping (in all six cases the birds were called homozygous in the Illumina genotyping and heterozygous in the

Sequenom genotyping; in five cases they passed on either the alternative allele or both alleles to their offspring). For ninety-nine individuals that were genotyped twice on the Sequenom platform, we found one inconsistency out of 6,317 genotypes (0.016%).

Statistical models and software

Genotype quality control was done in R (v3.0.2; R Core Team 2013). For each individual we calculated its inbreeding coefficient from pedigree data using the R package pedigreemm (v0.3-1; Vazquez *et al.* 2010). For the association analyses we used ASReml-R (v3; Gilmour *et al.* 2009) to fit univariate mixed-effects linear models. For the three captive populations ("Seewiesen", "Bielefeld" and "Cracow") we included the observer of the measurement (factor with maximally five levels) and the individual's sex (factor), age (covariate) and , inbreeding coefficient (covariate) as fixed effects. For body mass we included time of day (covariate). In the "Seewiesen" population we included whether the bird was measured dead or live (factor) for the three beak morphology traits (see Knief *et al.* (2012) for an explanation). We controlled for size as explained above. Instead of a GSM we fitted an additive genetic relatedness matrix to control for relatedness between individuals in the three captive populations. The proportion of variance explained by a SNP was calculated as explained above. For the "Sydney" birds we used the same fixed effects as in the initial sample of wild birds and added the observer of the measurement as a factor (two levels). We fitted a linear model without random effects (i.e. GSM or pedigree) in these birds.

We ran every model including and excluding the individuals for which genotypes were ambiguous due to null-alleles (see above). Results did not change substantially, so we here report only results excluding these individuals. Meta-analyses of average effects of allelic substitution of each SNP were carried out in the rmeta R-package (v2.16; Lumley 2012) using a fixed effect model, because random effect models are deflated when only a small number of studies are combined (Kraft *et al.* 2009). For each meta-analyzed effect estimate we obtained a P-value from the corresponding Z-score by P = 2 x pnorm(-abs(Z)) in R (v3.0.2; R Core Team 2013).

Study design

The design of our validation study followed the guidelines provided by the NCI-NHGRI working group on replication in association studies *et al.* (2007). We began by testing whether the selected candidate SNPs actually explained some of the phenotypic variation in those birds used for the QTL mapping (from the "Seewiesen" population). The initial QTL results were most likely biased upwards due to the Beavis effect (Beavis 1998; Slate 2013), which is due to the winner's curse but may also be due to gene-by-gene or gene-by-environment interactions (Göring *et al.* 2001). Individual SNP effects should be lower than the corresponding QTL effects, but generally pick up some of the QTL effects.

In the next step we replicated association results within the same population as the one used for the QTL mapping ("Seewiesen" population), but birds were taken from succeeding generations. Since the environment was kept constant, gene-by-environment interactions should have been negligible in this comparison.

Replication in a single captive population is not a valid procedure to determine whether the selected SNPs are indeed the causal variants for the phenotypic traits, because relatedness and linkage structure preclude fine-mapping in captivity. Thus, we compared average allelic effects between the "Seewiesen" and a second captive population ("Bielefeld") with a different genetic ancestry (Forstmeier *et al.* 2007) but that should exhibit a similar LD structure and was kept in the same environment. If the selected SNPs had similar effects in both captive populations, they would likely be QTN (NCI-NHGRI working group on replication in association studies *et al.* 2007).

Although sample sizes in the "Seewiesen" and "Bielefeld" populations were reasonably large, power might have been still low given the small effect sizes of QTN found in humans to date (Park *et al.* 2010; Yang *et al.* 2010). Thus, as a final analysis, we meta-analytically combined all average allelic substitution effects for every SNP across the four populations from the validation sample ("Seewiesen", "Bielefeld", "Cracow", "Sydney") and compared the SNP-effects to the estimates from the initial wild "Fowlers Gap" population.

Calculation of linkage disequilibrium (LD)

Association studies critically depend on LD between causal variants and genotyped SNPs. To contrast LD between wild and captive birds, we calculated LD between all candidate SNPs (n=672 SNPs) on a chromosome for the wild "Fowlers Gap" birds and between 1,269 SNPs that had been previously genotyped in 1,067 birds from the "Seewiesen" population (Backström *et al.* 2010; Schielzeth *et al.* 2011). We calculated LD including all related individuals to get an estimate that reflects the actual LD structure in the mapping populations.

SNP genotypes were not resolved into haplotypes and all LD calculations were thus based on composite LD (Weir 1996), such that double-heterozygous individuals were ambiguous in terms of their haplotypes. We calculated standardized LD as the squared Pearson's correlation coefficient (r^2) between every two SNPs within candidate genes on a chromosome (Weir 1979; Wellek & Ziegler 2009). LD measured as r^2 takes differences in allele frequencies into account, and is thus more comparable between markers and more informative in genotype-phenotype associations than the standardized disequilibrium coefficient D' (Remington *et al.* 2001; Zaykin 2004). Under low levels of mutation and drift-recombination equilibrium the expected value of r^2 ($E(r^2)$) can be approximated by:

$$E(r^2) = \left[\frac{10 + C}{(2 + C) * (11 + C)}\right] * \left[1 + \frac{(3 + C) * (12 + 12 * C + C^2)}{n * (2 + C) * (11 + C)}\right]$$

where C is the product of the population recombination rate ρ and the distance between SNPs in bp and n is the number of haplotypes sampled. The population recombination rate can be calculated as

ρ = 4 x N x c

where N is the effective population size and c the recombination fraction between sites (Hill & Weir 1988; Remington *et al.* 2001). We obtained least-squares estimates of ρ by fitting nonlinear models using the nls() function in R (R code provided by Marroni *et al.* 2011). The model does not provide an accurate estimate of ρ, especially in the captive population, because pairs of SNPs are not independent and populations are not at equilibrium (Weir & Hill 1986). Nonetheless, the model is informative for characterizing the decay of LD with physical distance (Remington *et al.* 2001).

Results

Linkage disequilibrium (LD) in wild and captive populations

The LD structure differed markedly between the wild birds from the "Fowlers Gap" population and those from the "Seewiesen" population. In the wild birds, LD decayed rapidly (ρ = 0.023, half-decay at 90.5 bp) and $r^2 > 0.1$ extended for maximally 185 kb (Figure 1A), whereas in the captive "Seewiesen" population LD extended over large physical distances (ρ = 5.6 x 10^{-6}, half-decay at around 368 kb) with $r^2 > 0.1$ even between SNPs that were separated by more than 133 Mb (Figure 1B).

Associations in wild birds

As expected for morphological traits, we found considerable additive genetic variance for each of the phenotypic traits studied in the "Fowlers Gap" population (Table S2). After accounting for multiple testing using the effective number of tests performed, only a single SNP on chromosome *Tgu12* in the 3'UTR of the *WNT5A* gene showed a significant association with wing length (P = 0.02; Figure 2, wing length in Figure 2E). However, this association also became non-significant after correcting for the number of phenotypic traits tested.

PP-plots for tarsus length and body mass showed unexpected deviations (Figure 2F, G), such that P-values were larger than expected under the assumption of uniformly distributed P-values. This is most likely a consequence of SNPs being in linkage disequilibrium with each other, such that individual tests were not completely independent.

Because effect sizes of single SNPs on quantitative traits are expected to be small, our initial scan may be underpowered. Hence, we selected the 25 SNPs with the strongest association signals (for 29 associations since four SNPs were associated with two phenotypes) for the validation study in the same wild population and in three captive zebra finch populations.

Comparison of QTL with SNP-effects in a captive population

We compared the proportion of phenotypic variance explained by each of the SNPs in the first four generations of the "Seewiesen" population with the proportion of phenotypic variance explained by their respective QTL regions. Single SNPs explained much less of the phenotypic variance than the QTL as a whole (maximally 4.2%, equivalent to 45% of the respective QTL effect; Figure 3A), which was expected given that SNPs are rarely in perfect LD with the entire QTL effect. Effect sizes of SNPs and QTL were positively correlated, but the correlation was not significant (Pearson's r = 0.34, n = 25, P = 0.097).

Replication of SNP-effects within a captive population

We checked whether the effects of allelic substitution for all 25 SNPs were replicable between those individuals in the "Seewiesen" population initially used for the linkage mapping and another set of 1,903 individuals from the next three generations of the same population. Effect sizes of all SNPs were highly correlated (Pearson's r = 0.86, n = 14, P = 7.3 x 10^{-5}; Figure 3B, Table 2). Therefore, we analyzed the entire "Seewiesen" population as one sample to improve the accuracy of allelic substitution effect estimates for each SNP (Table 2, Seewiesen combined).

Replication of SNP-effects across populations

We compared the average effects of allelic substitution between the "Seewiesen" and the "Bielefeld" population, because the two populations are ancestrally different and sample sizes for both were large. Although there are eight SNPs that showed a significant association in the "Bielefeld" population, only two of them replicated the allelic substitution effects found in the "Seewiesen" population (Pearson's r = 0.45, n = 17, P = 0.073; Figure 4A, Table 2).

We meta-analytically combined all average allelic substitution effects for every SNP across the four populations from the validation sample ("Seewiesen", "Bielefeld", "Cracow", "Sydney") and compared the SNP-effects to the estimates from the initial wild "Fowlers Gap" population. We found that the SNP-effects generally did not replicate between the initial discovery sample and the captive populations (Pearson's r = 0.24, n = 29, P = 0.21; Figure 4B, Table 2).

Validation of specific genotype-phenotype associations

The significant association between the SNP in the *WNT5A* gene and wing length in the "Fowlers Gap" population was not confirmed in any of the captive populations or in the meta-analysis (Association ID 25 in Table 2; meta-analysis: P=0.69).

One SNP in the conserved region close to gene *SHH* was significantly associated with beak length and beak depth in the meta-analysis and effect size estimates were similar both in sign and magnitude in the "Fowlers Gap" population (Association IDs 10 and 13 in Table 2; meta-analysis: P = 0.021 and P = 0.012). Another SNP in exon five of gene *CTNNB1* showed a similar pattern with beak depth and beak width (Association IDs 16 and 19 in Table 2; meta-analysis: P = 0.0016 and P = 0.057). However, none of these associations would be significant under a strict Bonferroni-correction.

Three SNPs in the conserved region close to gene *LIN28B* showed a significant association with digit ratio after Bonferroni-correction (Association IDs 6, 7 and 8 in Table 2; meta-analysis: $P = 9.6 \times 10^{-16}$, $P = 7.1 \times 10^{-5}$ and $P = 2.9 \times 10^{-5}$), but also significant heterogeneity between populations (Cochran's Q test of homogeneity for Association ID 6, P = 0.038). In any case, effects in the initial discovery analyses ("Fowlers Gap") were similar in sign and magnitude (Table 2). Although two SNPs in the conserved region near gene *ESR1* were significantly associated with digit ratio even after Bonferroni-correction (Association ID 3 and 4 in Table 2; meta-analysis: $P = 1.5 \times 10^{-10}$ and $P = 7.7 \times 10^{-8}$), there was considerable heterogeneity between populations (Cochran's Q test of homogeneity for Association ID 4, P = 0.0054). Moreover, in the initial discovery analysis in the "Fowlers Gap" population the SNP alleles had an opposite effect (Table 2).

Discussion

The aim of this study was to identify the actual causal variants underlying several quantitative morphological traits in zebra finches. Starting from 672 SNPs tested in a discovery sample of 939 unrelated wild-caught birds, we narrowed down on the most promising 25 SNPs in a validation sample of 5,229 birds from four captive populations with different genetic ancestry. Only a single SNP survived Bonferroni correction in the discovery stage, but its effect could not be replicated in any of the four populations. Within the captive "Seewiesen" population, SNP effects were highly repeatable but this was generally not the case in a cross-population context, suggesting that we did not identify the causal variants underlying the phenotypic traits of interest, but rather SNPs that were in linkage disequilibrium (LD) with them. Not surprisingly, the LD structure was markedly different between wild and captive populations, with rapid decay of LD in wild Australian zebra finches (see also Balakrishnan & Edwards 2009 for a smaller study) and high levels of LD in the captive populations due to recent founder events (Forstmeier *et al.* 2007; Zann 1996). Founder effects and associated high levels of LD may also explain the large differences in individual SNP effects between populations, especially since most of the birds in the

validation sample were kept in the same environment, reducing the potential impact of gene-by-environment interactions. Indeed, the prominent role of LD structure in association studies has been highlighted previously (Kraft *et al.* 2009; Kruglyak 1999).

Additive genetic variance in wild and captive populations

Heritability estimates for all the morphological traits were considerably lower in the wild ("Fowlers Gap" population) than in the captive populations. This lower heritability was not due to an increased environmental variance in the wild (coefficients of variation of phenotypic measures were not larger in the wild birds as compared to captive populations; Table 1) but to a lower additive genetic variance component (Table S2). This is counterintuitive because, if anything, the captive populations should have lost some of the additive genetic variance due to founder effects (Forstmeier *et al.* 2007; Lynch & Walsh 1998). It has been shown that heritability estimates are biased downwards when only a limited number of loci in imperfect LD with the causal variants is used for constructing the genetic similarity matrix (de los Campos *et al.* 2015), and the same effect was also found in other wild bird populations using even higher marker densities (Husby *et al.* 2015; Robinson *et al.* 2013). Thus, the low heritability estimates in the wild "Fowlers Gap" population is rather a methodological artefact than a real difference between wild and captive populations.

Association mapping within candidate genes in a wild population

Overall we found only a single SNP that was significantly associated with one of the phenotypic traits, namely wing length. Wing length was also the trait for which the least number of SNPs was tested for an association, suggesting that the Bonferroni correction after multiple testing might have led to too conservative P-values for the other phenotypic traits. We therefore decided to use an anti-conservative significance threshold for the selection of SNPs to be included in the confirmatory study. Individual genetic association studies (even on humans) suffer from being underpowered (Kraft *et al.* 2009) and replication is needed to establish a valid genotype-phenotype association (NCI-NHGRI working group on replication in association studies *et al.* 2007). This means that also marginally significant SNPs may be truly associated with the phenotype of interest and the absence of a genome-wide significant association at the first stage of an association study should not be overvalued (Kraft *et al.* 2009). However, the effect of the one significant association in the initial scan could not be replicated in any of the four validation populations. Consequently, the significant association in the initial scan is either spurious (i.e. a type I error) or real but absent from the captive populations because the effect is dependent on a specific environment (Falconer & Mackay 1996).

How much variance is explained by a single SNP within a previously identified QTL region?

Before assessing effect sizes of single SNPs within and across populations, we compared these effect sizes with those obtained from previous QTL studies conducted on the same set

of individuals ("Seewiesen" population), which removes the impact of gene-by-environment and gene-by-gene interactions. All SNPs that were genotyped within QTL regions explained much less of the phenotypic variance than the linkage analyses, which was expected because (1) effect size estimates from linkage studies are known to be inflated when they are estimated from the same data used for localizing the QTL regions (Beavis effect; Beavis 1998; Slate 2013) and (2) because a single QTL is probably composed of multiple linked causal variants. Even if only a single causal variant contributes to the QTL region and even if we had genotyped this variant or one in perfect LD with it in the same population, effect size estimates from the linkage study would still be biased upwards (Bogdan & Doerge 2005), because they had been estimated at the location where the test statistic was maximal (Knief *et al.* 2012; Schielzeth *et al.* 2012).

Are QTL effects replicable within a population?

Individual SNPs explained some of the QTL effect, which is necessary but not sufficient for being a causal variant. In the next step we replicated the associations in an additional three generations from the same captive population ("Seewiesen"). Allelic substitution effects of individual SNPs were consistent across generations and effect size estimates in the first four generations were not biased upwards. Thus, we conclude that the identified associations in the "Seewiesen" population and their associated QTL regions are real, as already suggested based on simulations (Slate 2013).

Are the most promising QTN replicable across different populations?

The main purpose of the study was to identify causal variants underlying phenotypic traits. In order to establish a causal relationship between genotype and phenotype, the association must at least be replicated in a comparable population (NCI-NHGRI working group on replication in association studies *et al.* 2007). When comparing SNP effects between populations (Figure 4), we found little repeatability, in strong contrast to the high repeatability of effects seen within one population (Figure 3B). The low repeatability cannot be caused exclusively by gene-by-environment interactions because even within the same captive environment two populations showed no concordance in average allelic substitution effects ("Seewiesen" versus "Bielefeld"). It rather indicates that most of the selected SNPs are not the causal variants underlying the traits, but may be linked to actual QTN. This interpretation is consistent with multiple observations. (1) LD in the captive "Seewiesen" population was high (due to founder effects), and although we were not able to test this in the other captive populations, there is no reason to suspect that LD is lower in those populations. Since LD (measured as r^2) is inversely proportional to the power of an association study (Pritchard & Przeworski 2001), it should be possible to detect significant associations due to LD between causal and genotyped markers frequently. (2) The captive populations have been founded from a large panmictic population (Zann 1996), with high genetic diversity and low LD (this study; Balakrishnan & Edwards 2009). Since the captive populations show some degree of genetic differentiation (Forstmeier *et al.* 2007), it is

reasonable to assume that they also differ in their haplotype and linkage structure. Given that we did not genotype a QTN this may lead to associations in some but not other populations.

Because it was not feasible to collect samples from another wild population that was similar in relatedness and LD structure, we meta-analytically summarized results from several genetically distinct captive populations. By considering several captive populations, spurious associations due to linkage in the population founders should have been broken up. The meta-analytically summarized allelic substitution effects from the four captive populations did not replicate the estimates from the wild "Fowlers Gap" population. Because our validation sample was four times larger, the detection power should have been high, even if effect sizes in the initial association mapping in the "Fowlers Gap" population were probably overestimated (Göring *et al.* 2001).

Conclusion

Although we identified several SNPs that were related to morphological measures in zebra finches, the selected candidate QTN are probably not the causal variants underlying these traits. Nevertheless, it may be worth following up on some of the potential QTN that showed consistent effects across populations. Two of these had pleiotropic effects on two beak dimensions, which are strongly genetically correlated (Knief *et al.* 2012). Furthermore, the candidate genes could still harbor causal variants in intronic non-conserved regions, which were not covered by the present study.

Even in times of large-scale genotyping efforts, an unequivocal identification of causal variants remains challenging in non-model organisms because effect sizes of single variants on quantitative traits are small and because of the difficulty of separating epistatic and gene-by-environment effects from initial false positive associations. Detailed knowledge about the LD structure in the mapping population is crucial because it determines statistical power, genomic resolution and coverage. Whereas wild populations offer high precision resulting from low levels of LD, they suffer from low power and genome coverage. Single captive populations are suitable for low-resolution mapping because LD is generally high and combining estimates from several captive populations with different genetic backgrounds may offer great potential for mapping causal variants.

Acknowledgements

We thank C. Beckmann, A. Sager, and M. Mariette for assistance with sampling the wild birds. Wild birds were sampled and banded under approval of the Macquarie University Animal Ethics Committee, the Australian Bird and Bat Banding Scheme, and a Scientific

License from NSW National Parks and Wildlife Service. We are grateful to K. Martin, J. Rutkowska, M. Ihle and J. Schreiber for help with breeding and phenotyping and S. Bauer, E. Bodendorfer, A. Grötsch, A. Kortner, K. Martin, P. Neubauer, F. Weigel, and B. Wörle for animal care and help with breeding. We further thank M. Schneider for laboratory work in Seewiesen and M. Schilhabel, the Next-Generation Sequencing team and the Genotyping team at the IKMB in Kiel for laboratory work. This study was funded by the Max Planck Society (B.K.), with the zebra finch study at Fowlers Gap funded by support to S.C.G. from the Australian Research Council. U.K. is part of the International Max Planck Research School for Organismal Biology.

References

Backström N, Forstmeier W, Schielzeth H*, et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research* **20**, 485–495.

Balakrishnan CN, Edwards SV (2009) Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* **181**, 645–660.

Balding DJ (2006) A tutorial on statistical methods for population association studies. *Nature Reviews Genetics* **7**, 781–791.

Beavis WD (1998) QTL analysis: power, precision, and accuracy. In: *Molecular dissection of complex traits* (ed. Paterson AH), pp. 145–162. CRC Press, Boca Raton, FL.

Bogdan M, Doerge RW (2005) Biased estimators of quantitative trait locus heritability and location in interval mapping. *Heredity* **95**, 476–484.

Cardon LR, Bell JI (2001) Association study designs for complex diseases. *Nature Reviews Genetics* **2**, 91–99.

Casella G, Berger RL (2002) *Statistical inference*, 2nd edn. Duxbury Press, Pacific Grove, California.

de los Campos G, Sorensen D, Gianola D (2015) Genomic heritability: what is it? *Plos Genetics* **11**, e1005048.

Eu-ahsunthornwattana J, Miller EN, Fakiola M*, et al.* (2014) Comparison of methods to account for relatedness in genome-wide association studies with family-based data. *Plos Genetics* **10**, e1004445.

Falconer D, Mackay T (1996) *Introduction to quantitative genetics*, 4th edn. Longmann, Harlow, UK.

Feder ME, Mitchell-Olds T (2003) Evolutionary and ecological functional genomics. *Nature Reviews Genetics* **4**, 651–657.

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.

Forstmeier W (2011) Women have relatively larger brains than men: a comment on the misuse of general linear models in the study of sexual dimorphism. *Anatomical Record* **294**, 1856–1863.

Forstmeier W, Segelbacher G, Mueller JC, Kempenaers B (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular Ecology* **16**, 4039–4050.

Gabriel S, Ziaugra L, Tabbaa D (2009) SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Curr Protoc Hum Genet* **Chapter 2**, Unit 2 12.

Gao XY, Stamier J, Martin ER (2008) A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genetic Epidemiology* **32**, 361–369.

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009) *Asreml user guide release 3.0*. VSN International Ltd, Hemel Hempstead, UK.

Gogarten SM, Bhangale T, Conomos MP*, et al.* (2012) GWASTools: an R/Bioconductor package for quality control and analysis of genome-wide association studies. *Bioinformatics* **28**, 3329–3331.

Göring HHH, Terwilliger JD, Blangero J (2001) Large upward bias in estimation of locus-specific effects from genomewide scans. *American Journal of Human Genetics* **69**, 1357–1369.

Griffith SC, Pryke SR, Mariette M (2008) Use of nest-boxes by the Zebra Finch (*Taeniopygia guttata*): implications for reproductive success and research. *Emu* **108**, 311–319.

Hill WG, Weir BS (1988) Variances and covariances of squared linkage disequilibria in finite populations. *Theoretical Population Biology* **33**, 54–78.

Hillier LW, Miller W, Birney E*, et al.* (2004) Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**, 695–716.

Holland PM, Abramson RD, Watson R, Gelfand DH (1991) Detection of specific polymerase chain-reaction product by utilizing the 5' --> 3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 7276–7280.

Husby A, Kawakami T, Ronnegard L*, et al.* (2015) Genome-wide association mapping in a wild avian population identifies a link between genetic and phenotypic variation in a life-history trait. *Proc Biol Sci* **282**.

Johnston SE, Gratten J, Berenos C*, et al.* (2013) Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* **502**, 93–95.

Kang HM, Zaitlen NA, Wade CM*, et al.* (2008) Efficient control of population structure in model organism association mapping. *Genetics* **178**, 1709–1723.

Kent WJ (2002) BLAT—The BLAST-like alignment tool. *Genome Research* **12**, 656–664.

Knief U, Hemmrich-Stanisak G, Wittig M*, et al.* (2015) Quantifying realized inbreeding in wild and captive animal populations. *Heredity* **114**, 397–403.

Knief U, Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2012) QTL and quantitative genetic analysis of beak morphology reveals patterns of standing genetic variation in an Estrildid finch. *Molecular Ecology* **21**, 3704–3717.

Kraft P, Zeggini E, Ioannidis JPA (2009) Replication in genome-wide association studies. *Statistical Science* **24**, 561–573.

Kruglyak L (1999) Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nature Genetics* **22**, 139–144.

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. *Science* **265**, 2037–2048.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760.

Listgarten J, Lippert C, Kadie CM*, et al.* (2012) Improved linear mixed models for genome-wide association studies. *Nature Methods* **9**, 525–526.

Lumley T (2012) rmeta: Meta-analysis. *R package version 2.16*.

Lynch M, Walsh B (1998) *Genetics and analysis of quantitative traits*. Sinauer, Sunderland, MA.

Mariette MM, Griffith SC (2012) Conspecific attraction and nest site selection in a nomadic species, the zebra finch. *Oikos* **121**, 823–834.

Marroni F, Pinosio S, Zaina G*, et al.* (2011) Nucleotide diversity and linkage disequilibrium in *Populus nigra* cinnamyl alcohol dehydrogenase (CAD4) gene. *Tree Genetics & Genomes* **7**, 1011–1023.

McCauley JL, Kenealy SJ, Margulies EH*, et al.* (2007) SNPs in multi-species conserved sequences (MCS) as useful markers in association studies: a practical approach. *Bmc Genomics* **8**, Artn 266.

McKenna A, Hanna M, Banks E*, et al.* (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297–1303.

NCI-NHGRI working group on replication in association studies, Chanock SJ, Manolio T*, et al.* (2007) Replicating genotype-phenotype associations. *Nature* **447**, 655–660.

O'Connell JR, Weeks DE (1998) PedCheck: A program for identification of genotype incompatibilities in linkage analysis. *American Journal of Human Genetics* **63**, 259–266.

Park JH, Wacholder S, Gail MH*, et al.* (2010) Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nature Genetics* **42**, 570–U139.

Pritchard JK, Przeworski M (2001) Linkage disequilibrium in humans: models and data. *American Journal of Human Genetics* **69**, 1–14.

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842.

R Core Team (2013) R: a language and environment for statistical computing. *3.0.2*.

Remington DL, Thornsberry JM, Matsuoka Y*, et al.* (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 11479–11484.

Robinson MR, Santure AW, DeCauwer I, Sheldon BC, Slate J (2013) Partitioning of genetic variation across the genome using multimarker methods in a wild bird population. *Molecular Ecology* **22**, 3963–3980.

Robinson MR, Wray NR, Visscher PM (2014) Explaining additional genetic variation in complex traits. *Trends in Genetics* **30**, 124–132.

Schielzeth H, Forstmeier W, Kempenaers B, Ellegren H (2012) QTL linkage mapping of wing length in zebra finch using genome-wide single nucleotide polymorphisms markers. *Molecular Ecology* **21**, 329–339.

Schielzeth H, Husby A (2014) Challenges and prospects in genome-wide quantitative trait loci mapping of standing genetic variation in natural populations. *Year in Evolutionary Biology* **1320**, 35–57.

Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2011) Data from: QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. *Dryad Digital Repository*.

Sham PC, Purcell SM (2014) Statistical power and significance testing in large-scale genetic studies. *Nature Reviews Genetics* **15**, 335–346.

Slate J (2005) Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Molecular Ecology* **14**, 363–379.

Slate J (2013) From Beavis to beak color: a simulation study to examine how much QTL mapping can reveal about the genetic architecture of quantitative traits. *Evolution* **67**, 1251–1262.

Slate J, Gratten J, Beraldi D*, et al.* (2009) Gene mapping in the wild with SNPs: guidelines and future directions. *Genetica* **136**, 97–107.

Slate J, Santure AW, Feulner PGD*, et al.* (2010) Genome mapping in intensively studied wild vertebrate populations. *Trends in Genetics* **26**, 275–284.

Speed D, Balding DJ (2015) Relatedness in the post-genomic era: is it still useful? *Nature Reviews Genetics* **16**, 33–44.

Stapley J, Birkhead TR, Burke T, Slate J (2008) A linkage map of the zebra finch *Taeniopygia guttata* provides new insights into avian genome evolution. *Genetics* **179**, 651–667.

Uleberg E, Meuwissen THE (2010) Fine mapping and detection of the causative mutation underlying quantitative trait loci. *Journal of Animal Breeding and Genetics* **127**, 404–410.

VanRaden PM (2008) Efficient methods to compute genomic predictions. *Journal of Dairy Science* **91**, 4414–4423.

Vasemägi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology* **14**, 3623–3642.

Vazquez AI, Bates DM, Rosa GJM, Gianola D, Weigel KA (2010) Technical note: An R package for fitting generalized linear mixed models in animal breeding. *Journal of Animal Science* **88**, 497–504.

Warren WC, Clayton DF, Ellegren H*, et al.* (2010) The genome of a songbird. *Nature* **464**, 757–762.

Weir BS (1979) Inferences about linkage disequilibrium. *Biometrics* **35**, 235–254.

Weir BS (1996) *Genetic data analysis II: methods for discrete population genetic data*. Sinauer Associates, Sunderland, USA.

Weir BS, Hill WG (1986) Nonuniform recombination within the human beta-globin gene cluster. *American Journal of Human Genetics* **38**, 776–781.

Wellek S, Ziegler A (2009) A genotype-based approach to assessing the association between single nucleotide polymorphisms. *Human Heredity* **67**, 128–139.

Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL (2014) Advantages and pitfalls in the application of mixed-model association methods. *Nature Genetics* **46**, 100–106.

Yang JA, Benyamin B, McEvoy BP*, et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* **42**, 565–U131.

Zann RA (1996) *The zebra finch: a synthesis of field and laboratory studies*. Oxford University Press, USA.

Zaykin DV (2004) Bounds and normalization of the composite linkage disequilibrium coefficient. *Genetic Epidemiology* **27**, 252–257.

Zheng XW, Levine D, Shen J*, et al.* (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328.

Ziegler A, König IR, Pahlke F (2010) *A Statistical Approach to Genetic Epidemiology: Concepts and Applications*, 2nd edn. Wiley-VCH, Weinheim, Germany.

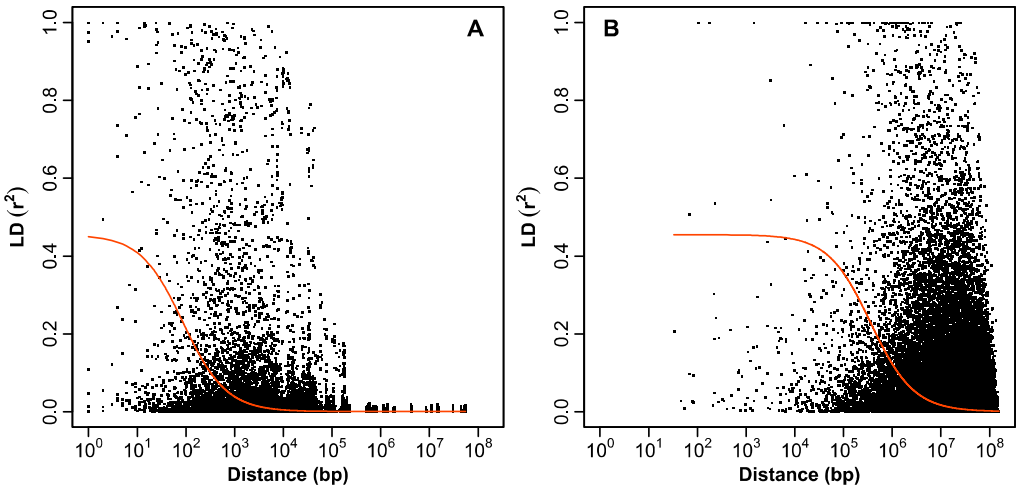


Figure 1: *Decay of linkage disequilibrium (LD) with (log_{10}) distance in (A) a sample of 939 wild-caught zebra finches ("Fowlers Gap") estimated using 38,464 pairwise comparisons between 672 SNPs in candidate genes on chromosomes Tgu1, Tgu2, Tgu3, Tgu5, Tgu5_random, Tgu6, Tgu7, Tgu10, Tgu12, Tgu20 and (B) a sample of 1,067 captive zebra finches ("Seewiesen") estimated using 54,495 pairwise comparisons between 1,269 SNPs on all annotated autosomes excluding Tgu16 and Tgu22. The red lines indicate nonlinear regressions of r^2 on distance based on a mutation-recombination drift model (Hill & Weir 1988; Remington et al. 2001).*

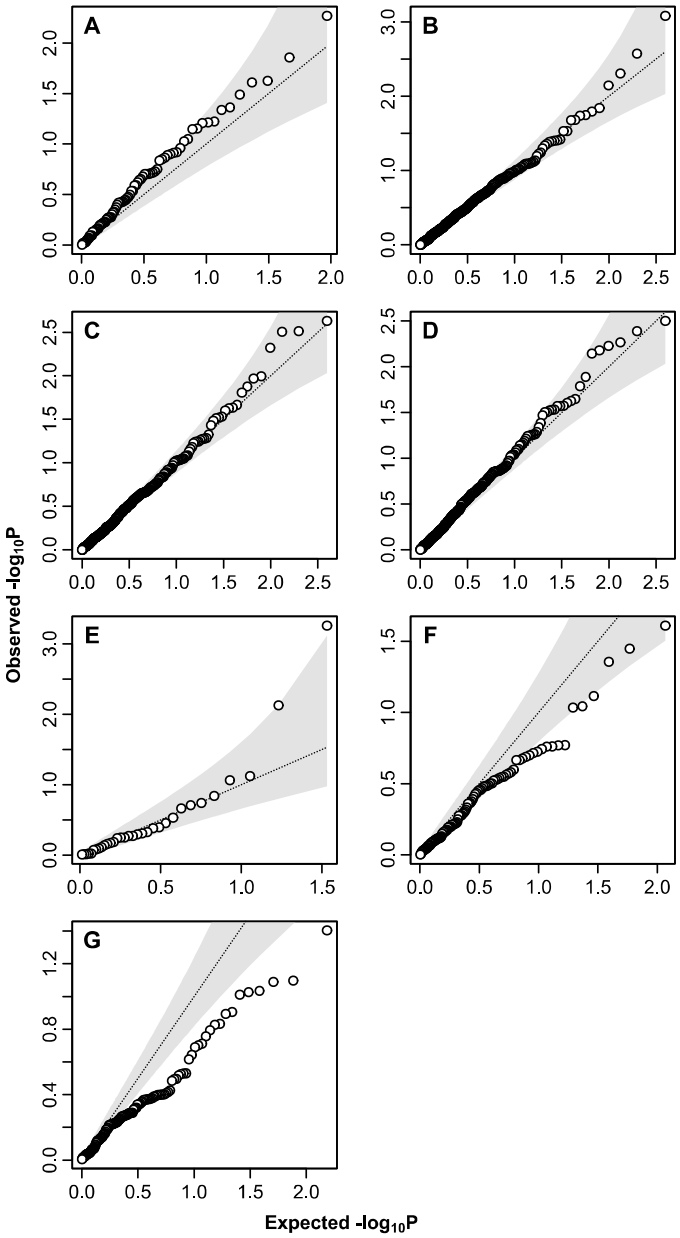


Figure 2: *PP-Plots of observed versus expected -log$_{10}$(P-values) in the initial association study in the "Fowlers Gap" population of wild birds for (A) digit ratio, (B) beak length, (C) beak depth, (D) beak width, (E) wing length, (F) body mass and (G) tarsus length. The dashed line indicates identity and the grey shading 95% confidence intervals, assuming P-values are independent and identically distributed (Casella & Berger 2002). In the presence of linkage disequilibrium (LD) these confidence intervals are conservative (i.e. too narrow). Judged by LD estimates from the simpleM algorithm (Gao et al. 2008), the confidence intervals for body mass (F), tarsus length (G) and digit ratio (A) are biased downwards most strongly.*

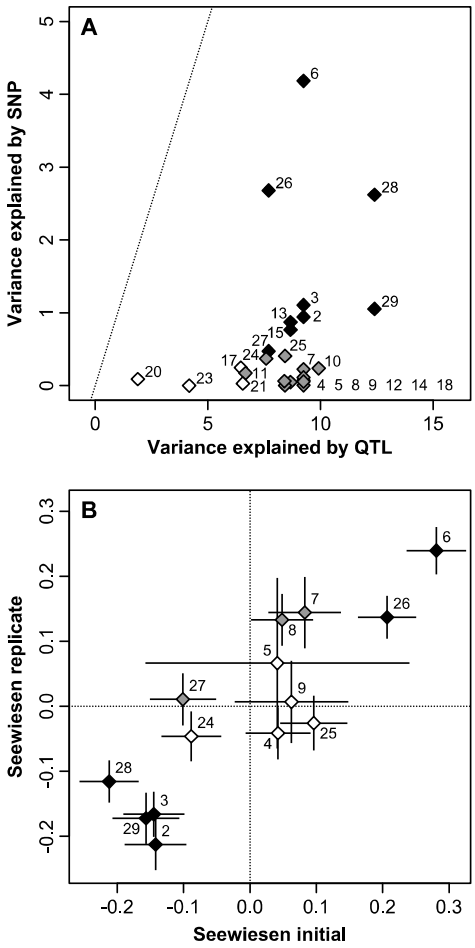


Figure 3: *Replication of SNP-effects within the "Seewiesen" population. (A) Relationship between the percentage phenotypic variance explained by the QTL-regions and by their corresponding individual SNPs. The dashed line indicates identity. (B) Relationship between the average effects of gene substitution (± SE) in the birds used for QTL-mapping (n = 1,156 phenotyped individuals) and in birds of the following four generations (n = 1,903 phenotyped individuals). A major axis regression slope was not significantly different from unity (β = 0.93, 95% CI = 0.65–1.33). Gene content (SNP genotypes) in the mixed-models are coded as -1 for individuals being homozygous for the minor allele in the wild birds ("Fowlers Gap" population), 0 for heterozygotes and 1 for individuals being homozygous for the major allele in the wild birds. Effects that are significant in both the QTL-mapping and replicate population are highlighted in black, effects that are significant in one of the two populations are coded grey and effects that are significant in neither of the two populations are white. Numbers at each data point refer to Association IDs in Table 2.*

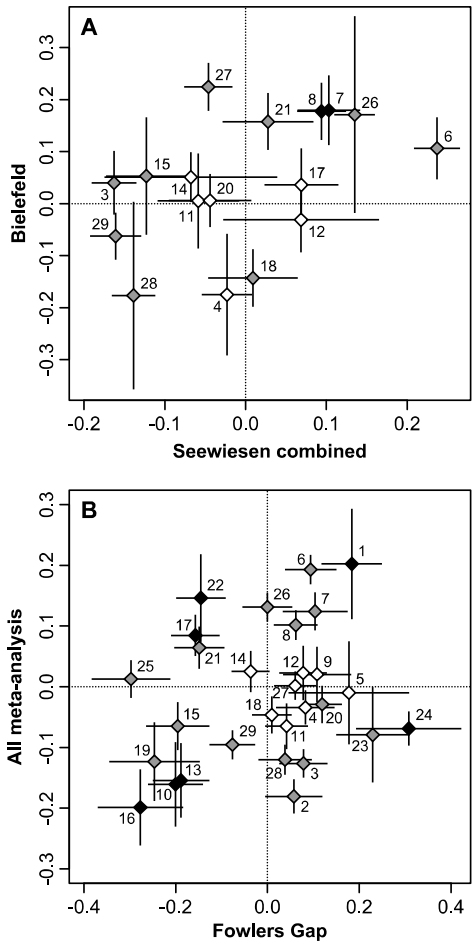


Figure 4: *Replication of SNP-effects across populations. (A) Relationship between the effects of gene substitution (± SE) in the combined "Seewiesen" population (n = 3,059 phenotyped individuals) and the "Bielefeld" population (n = 913 phenotyped individuals). (B) Relationship between the effects of gene substitution (± SE) in the initial "Fowlers Gap" population (n = 939 phenotyped individuals) and the meta-analytical combined "Seewiesen", "Bielefeld", "Cracow" and "Sydney" populations (n = 4,753 phenotyped individuals). Gene content (SNP genotypes) in the mixed-models are coded as -1 for individuals being homozygous for the minor allele in the wild birds ("Fowlers Gap" population), 0 for heterozygotes and 1 for individuals being homozygous for the major allele in the wild birds. Effects that are significant in both displayed populations are highlighted in black, effects that are significant in one of the two populations are coded grey and those that are significant in neither of the two populations are white. Numbers at each data point refer to Association IDs in Table 2.*

Table 1: *Descriptive statistics of phenotypic traits of five zebra finch populations.*

Population	Trait	Mean ± SD	N individuals	Mean age ± SD (d)	N authors *
Fowlers Gap	Digit ratio (mm)	0.90 ± 0.03	916	NA	1
	Beak length (mm)	10.14 ± 0.42	939	NA	1
	Beak depth (mm)	7.90 ± 0.27	938	NA	1
	Beak width (mm)	6.20 ± 0.19	939	NA	1
	Wing length (mm)	55.05 ± 1.49	937	NA	1
	Tarsus length (mm)	16.66 ± 0.49	931	NA	1
	Body mass (g)	12.07 ± 0.94	910	NA	1
Seewiesen	Digit ratio (mm)	0.93 ± 0.04	3,030	165 ± 202	2
	Beak length (mm) †	10.42 ± 0.50	1,086	1450 ± 583	1
	Beak depth (mm) †	8.34 ± 0.29	1,087	1450 ± 584	1
	Beak width (mm) †	6.61 ± 0.22	1,089	1450 ± 583	1
	Wing length (mm)	58.15 ± 1.49	3,085	102 ± 55	5
	Tarsus length (mm)	17.14 ± 0.58	3,095	113 ± 46	4
	Body mass (g)	15.87 ± 1.85	3,091	121 ± 38	2
Bielefeld	Digit ratio (mm)	0.92 ± 0.04	888	255 ± 73	2
	Beak length (mm)	10.12 ± 0.36	806	252 ± 65	1
	Beak depth (mm)	7.84 ± 0.30	806	252 ± 65	1
	Beak width (mm)	6.18 ± 0.20	806	252 ± 65	1
	Wing length (mm)	55.84 ± 1.40	909	255 ± 72	2
	Tarsus length (mm)	16.26 ± 0.63	908	255 ± 72	2
	Body mass (g)	12.11 ± 0.96	913	255 ± 72	2
Cracow	Digit ratio (mm)	0.90 ± 0.03	499	129 ± 8	1
	Beak length (mm)	10.51 ± 0.37	516	129 ± 9	1
	Beak depth (mm)	8.31 ± 0.29	516	129 ± 9	1
	Beak width (mm)	6.74 ± 0.23	516	129 ± 9	1
	Wing length (mm)	58.73 ± 1.42	510	129 ± 9	1
	Tarsus length (mm)	17.34 ± 0.73	513	129 ± 8	1
	Body mass (g)	15.95 ± 1.72	516	129 ± 9	1
Sydney	Digit ratio (mm)	0.91 ± 0.03	263	NA	2
	Beak length (mm)	10.12 ± 0.37	260	NA	2
	Beak depth (mm)	7.78 ± 0.34	259	NA	2
	Beak width (mm)	6.24 ± 0.21	260	NA	2
	Wing length (mm)	55.23 ± 1.45	265	NA	2
	Tarsus length (mm)	16.42 ± 0.53	263	NA	2
	Body mass (g)	12.27 ± 1.11	258	NA	2

* Whenever there is only a single author, birds were exclusively measured by U.K.

† Birds were measured either dead or alive. Although there was a significant difference between dead and live measurements we here report overall means and standard deviations. In the association models we controlled for status.

Table 2: *Minor allele substitution effects of all SNPs followed-up in the four populations and in the initial wild "Fowlers Gap" population. N refers to the combined number of individuals from the initial and validation study. Effects printed in bold are significant at α = 0.05.*

Trait	Gene	Association	Chromosome	Position	Fowlers Gap	Seewiesen initial	Seewiesen replicate	Seewiesen combined	Bielefeld	Cracow	Sydney	N
Digit ratio	*ESR1*	1	*Tgu3*	56,300,200	**0.184**					0.102	**0.292**	5,449
Digit ratio	*ESR1*	2*	*Tgu3*	56,410,679	0.0573	**-0.142**	**-0.213**	**-0.184**		**-0.275**	-0.0363	5,444
Digit ratio	*ESR1*	3	*Tgu3*	56,480,773	0.0783	**-0.145**	**-0.166**	**-0.163**	0.0399	**-0.23**	0.0118	5,418
Digit ratio	*ESR1*	4	*Tgu3*	56,481,129	0.0825	0.0425	-0.0414	-0.0231	-0.175	-0.0945	0.0259	5,443
Digit ratio	*LIN28B*	5*	*Tgu3*	71,185,384	0.177	0.0414	0.0663	0.0393		0.0133	-0.183	5,443
Digit ratio	*LIN28B*	6	*Tgu3*	71,190,570	0.094	**0.281**	**0.239**	**0.237**	0.106	0.0522	0.0737	5,367
Digit ratio	*LIN28B*	7	*Tgu3*	71,223,764	0.104	0.0826	**0.144**	**0.103**	**0.18**	0.0838	0.168	5,350
Digit ratio	*LIN28B*	8	*Tgu3*	71,225,705	0.0615	0.0485	**0.133**	**0.094**	**0.177**	0.176	-0.0427	5,364
Digit ratio	*LIN28B*	9	*Tgu3*	71,271,648	0.108	0.0628	0.00667	0.0132		-0.251	**0.318**	5,451
Beak length	*SHH*	10	*Tgu2*	8,964,100	**-0.201**	-0.135		-0.135		0.0487	**-0.2410**	3,413
Beak length	*DKK3*	11*	*Tgu5_random*§	1,639,897	0.0416	-0.0585		-0.0585	0.00503	-0.0499	-0.1950	3,406
Beak length	*BMP7*	12*	*Tgu20*	13,034,130	0.0777	0.0688		0.0688	-0.0312	-0.034	0.147	3,394
Beak depth	*SHH*	13	*Tgu2*	8,964,100	**-0.189**	**-0.258**		**-0.258**		0.0544	-0.101	3,412
Beak depth	*TGFBR2*	14*	*Tgu2*	60,184,328	-0.0369	-0.0677		-0.0677	0.0506	0.0237	0.011	3,412
Beak depth	*TGFBR2*	15	*Tgu2*	60,185,947	**-0.196**	**-0.123**		**-0.123**	0.0527	0.135	-0.0712	3,295
Beak depth	*CTNNB1*	16	*Tgu2*	64,387,922	**-0.277**				**-0.202**		-0.171	3,406
Beak depth	*BMP2*	17	*Tgu3*	25,959,864	**-0.157**	0.069		0.069	0.0358	0.143	**0.201**	3,427
Beak depth	*DKK3*	18*	*Tgu5_random*§	1,638,471	0.00935	0.0091		0.0091	**-0.143**	-0.0129	-0.015	3,119
Beak width	*CTNNB1*	19	*Tgu2*	64,387,922	**-0.246**				**-0.133**		-0.00894	3,410
Beak width	*BMP2*	20	*Tgu3*	25,954,108	**0.119**	-0.0441		-0.0441	0.00595	-0.0407	-0.070	3,428
Beak width	*DKK3*	21	*Tgu5_random*§	1,644,170	**-0.149**	0.0277		0.0277	**0.158**	-0.0549	-0.0532	3,431
Beak width	*DKK3*	22	*Tgu5_random*§	1,645,324	**-0.145**				**0.254**	0.164	0.0024	3,376
Beak width	*BMP7*	23	*Tgu20*	13,002,363	**0.23**	-0.00242		-0.00242		-0.137	-0.15000	3,374
Wing length	*WNT6*	24	*Tgu7*	10,279,234	**0.308**	-0.0882	-0.0463	**-0.0674**		-0.194	0.1440	5,304
Wing length	*WNT5A*	25	*Tgu12*	7,944,291	**-0.298**	0.0964	-0.0259	0.0185		-0.0599	-0.093	5,397

Body mass	*FOXO1A*	26*	*Tgu1*	54,663,134	0.000465	**-0.207**	**-0.137**	**-0.135**	-0.171	-0.157	-0.0245	5,404
Body mass	*INTS6*	27*	*Tgu1*	55,214,922	0.0605	**-0.101**	0.0106	-0.0461	**0.224**	**-0.202**	**-0.227**	5,389
Tarsus length	*FOXO1A*	28†	*Tgu1*	54,663,134	0.0383	**-0.212**	**-0.116**	**-0.139**	-0.177	-0.14	**0.222**	5,513
Tarsus length	*INTS6*	29†	*Tgu1*	55,214,922	-0.0766	**-0.157**	**-0.172**	**-0.161**	-0.0626	-0.012	**0.208**	5,498

* Associations were found in the Fowlers Gap population when genotype was fitted as a factor using two degrees of freedom (dominance effect).

† *FOXO1A* and *INTS6* are candidate genes for tarsus length and reside within QTL peaks in the Seewiesen population. However, there was no association in birds from Fowlers Gap.

§ SNPs in *DKK3* on chromosome *Tgu5_random* are linked to a polymorphic inversion on chromosome *Tgu5*, which extends from around 0.9–15 Mb. QTL peaks on chromosome *Tgu5* span almost the entire chromosome, starting at the first SNP in our linkage map. Consequently, *DKK3* resides within the QTL peak region.

Supplement

Material and Methods

GATK VariantFiltration settings

-filterExpression "(AF > 1.00)" -filterName "filter_AF" -filterExpression "(DP < 10.0 || DP > 600.0)" -filterName "filter_DP" -filterExpression "(HRun > 10.0)" -filterName "filter_HRun" -filterExpression "(MQ0 > 50.0 || ((MQ0/(1.0*DP)) > 0.50))" -filterName "filter_MQ0" -filterExpression "(QD < 0.04)" -filterName "filter_QDlow" -filterExpression "(QD > 39.0)" -filterName "filter_QDhigh" -filterExpression "(MQ < 10.0)" -filterName "filter_MQ" -filterExpression "(Dels > 40.0)" -filterName "filter_Dels" -clusterWindowSize 0 -mask InDels -maskName "InDel" -maskExtension 0

Testing whether drop-outs of SNPs on the iSelect chip are randomly distributed

We tested whether drop-outs were randomly distributed along chromosomes by permuting the positions of drop-outs 10,000 times for each chromosome and recording the distance between all drop-outs. We then compared the variance in distance between empirical drop-outs with the simulated ones. If drop-outs were clustered we would expect the empirical distribution of distances to be enriched for small (within clusters) and large (between clusters) distances, resulting in a larger variance in distance. The variance in distance between drop-outs on each chromosome was within the 95% quartile range of the expected variance based on the simulation (except for chromosomes *Tgu17* and *TguZ* where the empirical variances were in the top 3.7% and 3.1% of simulated variances, respectively). We concluded that drop-outs were randomly distributed along chromosomes.

References

Abzhanov A, Kuo WP, Hartmann C*, et al.* (2006) The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* **442**, 563–567.

Abzhanov A, Protas M, Grant BR, Grant PR, Tabin CJ (2004) *Bmp4* and morphological variation of beaks in Darwin's finches. *Science* **305**, 1462–1465.

Abzhanov A, Tabin CJ (2004) *Shh* and *Fgf8* act synergistically to drive cartilage outgrowth during cranial development. *Developmental Biology* **273**, 134–148.

Forstmeier W, Mueller JC, Kempenaers B (2010) A polymorphism in the oestrogen receptor gene explains covariance between digit ratio and mating behaviour. *Proceedings of the Royal Society B-Biological Sciences* **277**, 3353–3361.

Knief U, Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2012) QTL and quantitative genetic analysis of beak morphology reveals patterns of standing genetic variation in an Estrildid finch. *Molecular Ecology* **21**, 3704–3717.

Lamichhaney S, Berglund J, Almén MS*, et al.* (2015) Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature* **518**, 371–375.

Liu X, Zhang H, Li H*, et al.* (2008) Fine-mapping quantitative trait loci for body weight and abdominal fat traits: effects of marker density and sample size. *Poultry Science* **87**, 1314–1319.

Mallarino R, Campàs O, Fritz JA*, et al.* (2012) Closely related bird species demonstrate flexibility between beak morphology and underlying developmental programs. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 16222–16227.

Mallarino R, Grant PR, Grant BR*, et al.* (2011) Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 4057–4062.

Medland SE, Zayats T, Glaser B*, et al.* (2010) A variant in *LIN28B* is associated with 2D:4D finger-length ratio, a putative retrospective biomarker of prenatal testosterone exposure. *American Journal of Human Genetics* **86**, 519–525.

Schielzeth H, Forstmeier W, Kempenaers B, Ellegren H (2012) QTL linkage mapping of wing length in zebra finch using genome-wide single nucleotide polymorphisms markers. *Molecular Ecology* **21**, 329–339.

Song Y, Hui JN, Fu KK, Richman JM (2004) Control of retinoic acid synthesis and FGF expression in the nasal pit is required to pattern the craniofacial skeleton. *Developmental Biology* **276**, 313–329.

Wu P, Jiang TX, Shen JY, Widelitz RB, Chuong CM (2006) Morphoregulation of avian beaks: comparative mapping of growth zone activities and morphological evolution. *Developmental Dynamics* **235**, 1400–1412.

Wu P, Jiang TX, Suksaweang S, Widelitz RB, Chuong CM (2004) Molecular shaping of the beak. *Science* **305**, 1465–1466.

Xie L, Luo CL, Zhang CG*, et al.* (2012) Genome-wide association study identified a narrow chromosome 1 region associated with chicken growth traits. *Plos One* **7**, e30910.

Zhang H, Liu SH, Zhang Q*, et al.* (2011) Fine-mapping of quantitative trait loci for body weight and bone traits and positional cloning of the *RB1* gene in chicken. *Journal of Animal Breeding and Genetics* **128**, 366–375.

Zhang H, Wang SZ, Wang ZP*, et al.* (2012) A genome-wide scan of selective sweeps in two broiler chicken lines divergently selected for abdominal fat content. *Bmc Genomics* **13**, Artn 704.

Zhang H, Zhang YD, Wang SZ*, et al.* (2010) Detection and fine mapping of quantitative trait loci for bone traits on chicken chromosome one. *Journal of Animal Breeding and Genetics* **127**, 462–468.

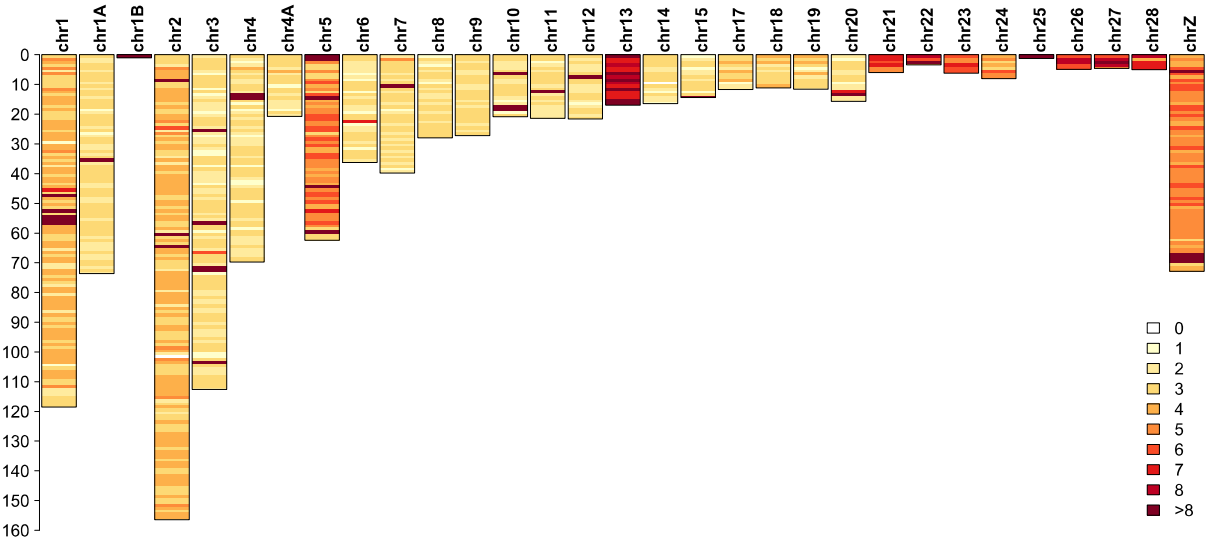


Figure S1: *Distribution of 4,553 SNPs genotyped with an Illumina Infinium iSelect HD Custom BeadChip in 1Mb windows along the annotated chromosomes of the zebra finch genome (genome build WUSTL v3.2.4). No SNP was placed on chromosome Tgu16 because this chromosome is assembled for only 9.9 kb.*

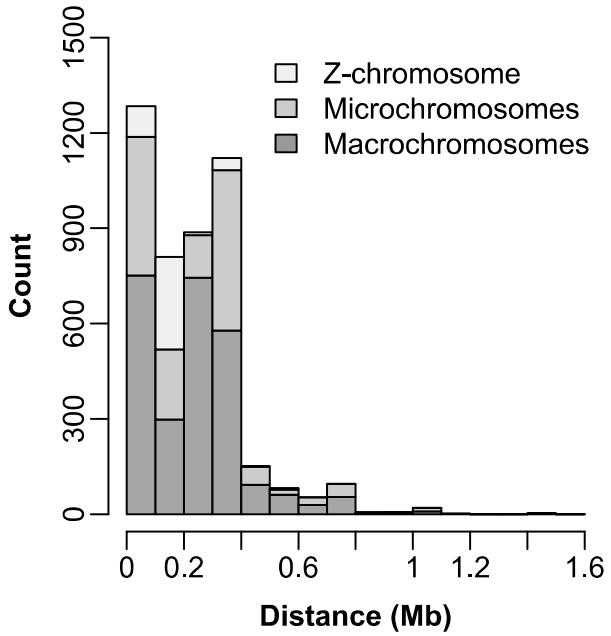


Figure S2: *Minimum pairwise distances between 4,553 SNPs on the Illumina Infinium iSelect HD Custom BeadChip separated for autosomal macrochromosomes (chromosomes Tgu1–Tgu5, Tgu1A), microchromosomes and the Z-chromosome. The median distance between adjacent SNPs is 227.09 kb (50% of the distances are between 29.01 kb and 345.61 kb, interquartile range).*

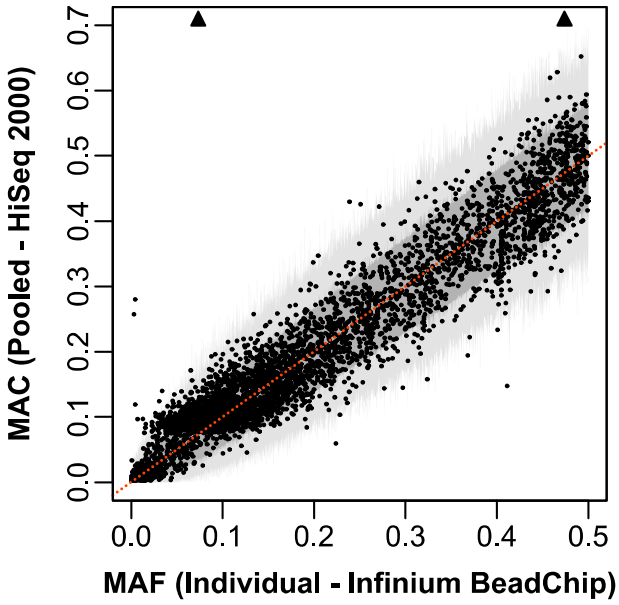


Figure S3: *Correlation of the minor allele frequency (MAF) of 4,401 SNPs estimated by genotyping 948 wild-caught zebra finches using an Illumina Infinium iSelect HD Custom BeadChip with the minor allele count (MAC) obtained by pooled sequencing of a subset of 100 individuals to an average coverage of 247.5 x by using the Illumina HiSeq 2000 platform (Pearson's r = 0.96). To get an estimate of the nominal 95% (dark grey) and Bonferroni-corrected 95% (light grey) quantile range we drew randomly 100 from the 948 individuals and sampled for each of their SNPs 248 alleles with replacement (to mimic pooled sequencing of 100 individuals with an average coverage of 247.5 x). We repeated this 10,000 times and recorded in each round the MAC for each SNP. The red dashed line represents the line of identity. Although we call the ordinate "minor allele count", its values may be larger than 0.5 because we refer to the minor allele obtained via genotyping the 948 individuals. Two SNPs at [0.073, 0.81] and at [0.47, 0.90] are outside the plotting area.*

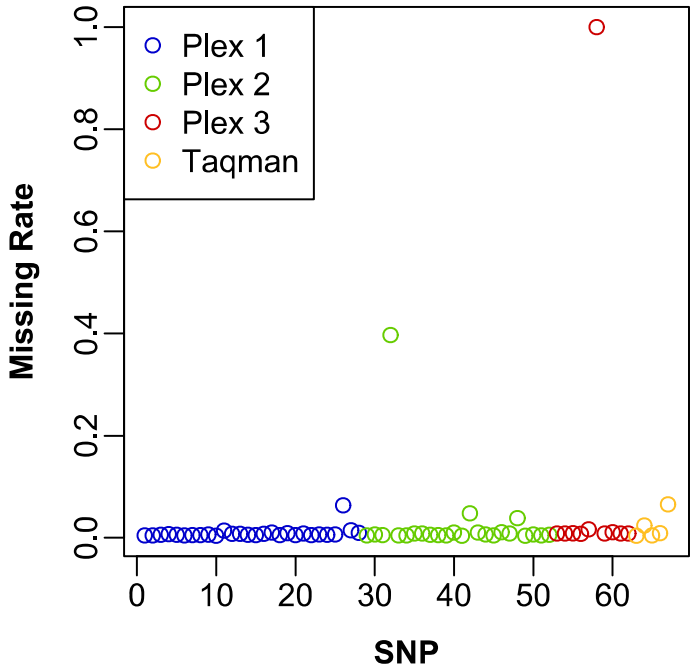


Figure S4: *Missing call rate per SNP after accounting for null-alleles in the validation sample of 5,229 birds. Genotyping of one SNP in gene RALDH3 failed.*

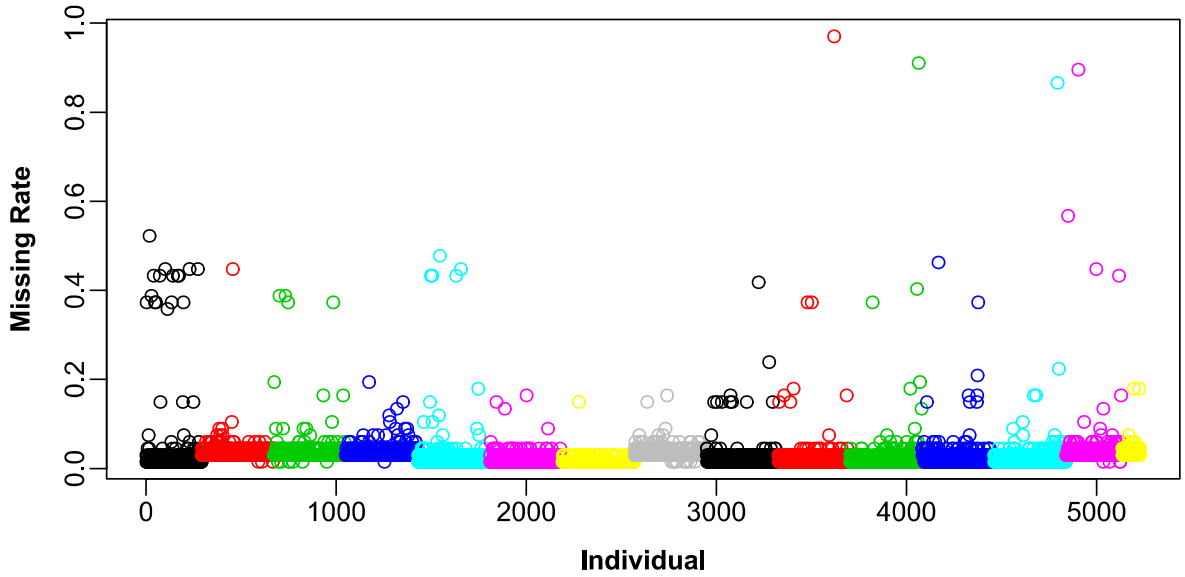


Figure S5: *Missing call rate per individual (n = 5,229, validation sample). For some individuals complete assays did not work, indicating technical genotyping errors.*

Table S1: *Description of the candidate genes studied in wild and captive zebra finch populations. For each gene, its position in the WUSTL 3.2.4 zebra finch genome assembly is given together with the number of SNPs that were intended to be present and that were actually present on the Illumina Infinium iSelect HD Custom BeadChip.*

Gene	Trait	Position	N SNPs intended	N SNPs on Illumina chip					References
				Total	Promoter	Exonic	3'UTR	Conserved	
ESR1	Digit ratio	*chr3*: 56296005-56481396	50	38	9	9	8	12	1
LIN28B	Digit ratio	*chr3*: 71181842-71277129	79	69	0	2	0	67	2
TGFBR2	Beak morphology	*chr2*: 60123432-60186353	104	89	0	15	35	39	3, 4, 5
CTNNB1	Beak morphology	*chr2*: 64381075-64413468	88	76	0	15	1	60	3, 4, 5
SHH	Beak morphology	*chr2*: 8957599-8988447	14	13	0	1	0	12	3, 6, 7
BMP2	Beak morphology	*chr3*: 25954020-25979222	20	15	0	1	1	13	3, 8
CALM1	Beak morphology	*chr5*: 44692021-44705897	14	13	0	2	5	6	3, 4, 9, 10
BMP4	Beak morphology	*chr5*: 59381704-59395667	7	6	0	1	0	5	3, 4, 8, 11
DKK3	Beak morphology	*chr5_random*: 1621546-1645796	55	47	0	8	6	33	3, 4, 5
FGF8	Beak morphology	*chr6*: 22049349-22064936	4	4	0	0	0	4	3, 6
IHH	Beak morphology	chr7: 10426278-10431101	3	2	0	1	0	1	4
RALDH2	Beak morphology	*chr10*: 6777382-6844027	85	74	0	6	7	61	3, 12
RALDH3	Beak morphology	*chr10*: 17967161-18010202	76	66	0	6	34	26	3, 12
BMP7	Beak morphology	*chr20*: 12988263-13038222	50	40	0	1	0	39	3, 8
Twist2	Wing length	*chr7*: 1553207-1586581	4	3	0	0	2	1	13
WNT6	Wing length	*chr7*: 10265481-10279398	18	13	0	6	0	7	13
WNT5A	Wing length	*chr12*: 7942566-7974364	28	19	0	2	6	11	13
RB1	Tarsus length / Body mass	*chr1*: 56557927-56634069	44	36	0	11	6	19	14, 15, 16, 17
FOXO1A	Tarsus length / Body mass	*chr1*: 54662825-54747562	29	27	0	8	11	8	14, 18
INTS6	Tarsus length / Body mass	*chr1*: 55213859-55254685	40	33	0	2	3	28	18
RNASEH2B	Tarsus length / Body mass	*chr1*: 55409553-55448059	50	43	0	4	31	8	18
KPNA3	Tarsus length / Body mass	*chr1*: 56015350-56043313	22	20	0	2	3	15	18

1 Forstmeier *et al.* (2010), 2 Medland *et al.* (2010), 3 Knief *et al.* (2012) and references therein, 4 Mallarino *et al.* (2012), 5 Mallarino *et al.* (2011), 6 Abzhanov & Tabin (2004), 7 Wu *et al.* (2006), 8 Abzhanov *et al.* (2004), 9 Lamichhaney *et al.* (2015), 10 Abzhanov *et al.* (2006), 11 Wu *et al.* (2004), 12 Song *et al.* (2004), 13 Schielzeth *et al.* (2012) and references therein, 14 Zhang *et al.* (2010), 15 Zhang *et al.* (2012), 16 Liu *et al.* (2008), 17 Zhang *et al.* (2011), 18 Xie *et al.* 2012.

Table S2: *Additive genetic variance and narrow-sense heritability for each phenotype studied.*

Population	Trait	V_A ± SE	h^2 ± SE	N individuals
Fowler's Gap	Digit ratio (mm)	0.00043 ± 0.000085	0.41 ± 0.073	916
	Beak length (mm)	0.018 ± 0.009	0.15 ± 0.074	910
	Beak depth (mm)	0.016 ± 0.0042	0.29 ± 0.074	909
	Beak width (mm)	0.0088 ± 0.0023	0.29 ± 0.075	910
	Wing length (mm)	0.5 ± 0.13	0.31 ± 0.10	908
	Tarsus length (mm)	0.027 ± 0.016	0.13 ± 0.077	904
	Body mass (g)	0.16 ± 0.05	0.26 ± 0.086	887
Seewiesen	Digit ratio (mm)	0.00085 ± 0.00006	0.72 ± 0.030	3,030
	Beak length (mm)	0.087 ± 0.017	0.4 ± 0.066	1,085
	Beak depth (mm)	0.045 ± 0.0057	0.65 ± 0.055	1,086
	Beak width (mm)	0.028 ± 0.0032	0.7 ± 0.05	1,088
	Wing length (mm)	1.4 ± 0.089	0.75 ± 0.035	3,073
	Tarsus length (mm)	0.18 ± 0.013	0.63 ± 0.032	3,083
	Body mass (g)	1.3 ± 0.1	0.54 ± 0.045	3,041
Bielefeld	Digit ratio (mm)	0.00082 ± 0.00013	0.59 ± 0.067	888
	Beak length (mm)	0.081 ± 0.013	0.68 ± 0.07	806
	Beak depth (mm)	0.035 ± 0.0057	0.58 ± 0.066	806
	Beak width (mm)	0.017 ± 0.0026	0.61 ± 0.065	806
	Wing length (mm)	0.98 ± 0.15	0.66 ± 0.08	909
	Tarsus length (mm)	0.18 ± 0.024	0.67 ± 0.057	908
	Body mass (g)	0.36 ± 0.061	0.54 ± 0.074	908
Cracow	Digit ratio (mm)	0.00056 ± 0.00014	0.56 ± 0.10	450
	Beak length (mm)	0.11 ± 0.022	0.75 ± 0.097	466
	Beak depth (mm)	0.051 ± 0.010	0.8 ± 0.092	466
	Beak width (mm)	0.04 ± 0.0065	0.85 ± 0.074	466
	Wing length (mm)	1.5 ± 0.24	0.82 ± 0.09	461
	Tarsus length (mm)	0.23 ± 0.072	0.46 ± 0.13	463
	Body mass (g)	1.1 ± 0.29	0.43 ± 0.14	463

Chapter 7: Large inversion polymorphisms in the zebra finch genome: effects on morphology and fitness

Abstract

Inversion polymorphisms constitute an evolutionary puzzle: Albeit they potentially increase embryo mortality in heterokaryotypic individuals they are ubiquitously found in nature. Here we describe four large (12–63 Mb) intraspecific inversion polymorphisms in wild Australian zebra finches (Taeniopygia guttata), which in total span at least 8.7% of the genome and 8.1% of all annotated genes. Two of them had been identified previously in cytogenetic screens, because they shift the position of the centromere. Although all polymorphisms were found to be in Hardy-Weinberg equilibrium in a sample of 939 wild birds, their remarkably high (and also similar) allele frequencies (range of the major allele 0.53–0.60) may indicate some sort of balancing selection acting on them. Patterns of diversity and heterozygosity along the inverted regions are consistent with the idea of balancing selection. Using inversion genotypes of 6168 wild and captive individuals we find tentative evidence that the two largest inversions increase embryo mortality in heterokaryotypic males (but not so in females), which makes neutral evolution even less likely. We find strong additive effects of inversion genotypes on several morphological traits, but no dominant gene action on morphology or several aspects of fitness in three captive populations, which makes simple heterosis unlikely to stabilize the inversion polymorphism. The inversion tagging SNPs that we describe here can be used in other wild and captive populations of this model species to test for phenotypic associations and to identify the selective forces balancing this polymorphism.

Prepared as: Knief U, G Hemmrich-Stanisak, M Wittig, A Franke, SC Griffith, B Kempenaers, W Forstmeier: Large inversion polymorphisms in the zebra finch genome: effects on morphology and fitness.

Introduction

Between-individual genetic variation is the substrate for selection and ranges in size from single nucleotides (SNPs) to large scale insertions, deletions or rearrangements that span several millions of basepairs (Conrad & Hurles 2007; White 1977). Among these structural variants, inversions play a prominent role, since they have long been recognized as drivers of local adaptation and speciation (reviewed in Hoffmann & Rieseberg 2008). Inversions are an intrachromosomal rearrangement, in which a segment of the chromosome has been turned around by 180°, such that the genic content of the chromosome does not change but is reversely ordered.

In heterokaryotypic individuals (those that are heterozygous for an inversion) recombination within the inverted region is suppressed, either because homologous pairing is partially inhibited or because crossovers give rise to unbalanced gametes (carrying deletions or duplications) which will lead to the death of the zygote (White 1977). These two processes are not mutually exclusive and their prevalence depends, amongst others, on the size and location of the inverted region (Anton *et al.* 2005; Navarro *et al.* 1997; Navarro & Ruiz 1997; Roberts 1967). In particular, a distinction between those inversions which cover both chromosome arms and thus include the centromere (pericentric inversions) and those which are restricted to a single chromosome arm (paracentric inversions) has often been made (Swanson *et al.* 1981): A single crossover within a pericentric inversion leads to the formation of two chromatids with duplications and deficiencies, and two normal chromatids, whereas in paracentric inversions an acentric fragment and a dicentric chromatid along with two normal chromatids are formed (Swanson *et al.* 1981). In species with an ordered (linear) tetrad in the female meiosis [e.g. *Drosophila sp.* or maize (*Zea mays*)] paracentric inversions often do not cause reduced fertility in females because the dicentric chromatid is preferentially passed into the second polar body (Roberts 1967; Swanson *et al.* 1981). On the other hand, pericentric inversions often lead to decreased fertility in females (Navarro & Ruiz 1997; Roberts 1967), which may also explain the preponderance of polymorphic paracentric over pericentric inversions in species like *Drosophila sp.* that lack male recombination (Krimbas & Powell 1992). In contrast, humans and maize recombine in the male meiosis, and heterokaryotypic males for both pericentric and paracentric inversions produce a higher percentage of unbalanced gametes and hence inviable embryos (Anton *et al.* 2005; Bhatt *et al.* 2009; Morel *et al.* 2007; Morgan 1950; Yapan *et al.* 2014), with recombination frequency being the highest in the largest inversions (absolute or proportional to the total chromosome size).

Limited exchange of genetic material between the two chromosomal arrangements (= gene flux) is possible in single large-scale inversions either through gene conversion or multiple crossovers within the inverted region (Andolfatto *et al.* 2001; Navarro *et al.* 1997): Whenever two crossovers involve the same two chromatids, one would undo the deleterious effect of the other, leading exclusively to balanced gametes, two of which have

exchanged genetic material between arrangements (Navarro *et al.* 1997; Swanson *et al.* 1981). On the other hand, an arrangement consisting of multiple inversions on a single chromosome is able to suppress gene flux between the two chromosomal arrangements over several megabases almost completely (e.g. the *t*-complex in the house mouse *Mus musculus*; Lyon 2003).

Despite their presumed heterozygous fitness costs, which would ultimately lead to the loss of the minor allele from the population, inversion polymorphisms are ubiquitously found within species (reviewed in Hoffmann & Rieseberg 2008). Consequently, in order to get established or to be maintained as a polymorphism, the inverted region should either confer a fitness advantage to the organism or exhibit segregation distortion (drift as the sole force is unlikely but may contribute in small populations; Kirkpatrick 2010). The most prominent feature of an inversion is its ability to suppress recombination within the inverted region, and thereby linkage disequilibrium between locally adapted combinations of alleles may be preserved, which may eventually lead to the spread of an inversion (with epistatic fitness interactions: Dobzhansky 1970, without epistasis: Feder *et al.* 2011; Kirkpatrick & Barton 2006).

Most of the know inversion polymorphisms within a species vary latitudinally (Balanya *et al.* 2006; Cheng *et al.* 2012; Fang *et al.* 2012; McAllister *et al.* 2008; Umina *et al.* 2005), locally (Lowry & Willis 2010) or seasonally (Dubinin & Tiniakov 1946), apparently in response to a changing environment. However, there are also examples of polymorphisms within single populations which could be stabilized via frequency dependent (disruptive) selection (Joron *et al.* 2011; Kopp & Hermisson 2006), antagonistic pleiotropy (Avelar *et al.* 2013), mate choice (Lowther 1961; Thorneycroft 1975), recessive deleterious mutations captured by or accumulating on the inverted haplotype ("associative overdominance", Kirkpatrick 2010; Sturtevant & Mather 1938), overdominance (i.e. the heterokaryotypic individuals have higher fitness than both homozygotes; Falconer & Mackay 1996; Sturtevant & Mather 1938) or under several scenarios involving segregation distortion (Lyon 2003; Traulsen & Reed 2012).

In birds, intraspecific inversion polymorphisms have been regarded as common (Price 2008; Shields 1982), yet the only well-studied polymorphism is found in white-throated sparrows (*Zonotrichia albicollis*), which are in fact polymorphic for at least two large pericentric inversions: Whereas one of the rearrangements (on chromosome 3, which should be equivalent to zebra finch chromosome *Tgu1* assuming homology within passerines; Griffin *et al.* 2007) may exhibit segregation distortion (Thorneycroft 1975), the other inversion on chromosome 2 (homologous to zebra finch chromosome *Tgu3*) covers around 1,000 genes and almost completely suppresses recombination between arrangements (which differ in fact by at least two included pericentric inversions; Huynh *et al.* 2011). Intriguingly, the inversion on chromosome 2 exhibits plumage and behavioral phenotypes and is kept

polymorphic by disassortative mating between birds with the two arrangements (e.g. Huynh *et al.* 2011).

The zebra finch (*Taeniopygia guttata*) belongs to the family of grass finches (*Estrildidae*), which are rich in fixed and polymorphic inversions (Christidis 1986a, b, 1987). Until now two polymorphic pericentric inversions have been described cytogenetically in zebra finches on chromosomes *Tgu5* (which is the sixth largest chromosome in the karyotype; Christidis 1986a; Itoh & Arnold 2005) and the sex chromosome *TguZ* (Itoh *et al.* 2011). The inversion on chromosome *TguZ* was found to be polymorphic over the whole Australian continent and, with a different allele frequency, in the subspecies from Timor (*T. guttata guttata*).

Here we perform a genome-wide scan for inversion polymorphisms by searching for unusual patterns of long-range linkage disequilibrium using 4,553 SNPs in a wild population of 939 zebra finches. After identifying four large linkage blocks (two of which are the known inversion polymorphisms on chromosomes *Tgu5* and *TguZ*) we infer the inversion genotypes for every individual by principle component analysis, select unique tagging SNPs and genotype an additional set of 5,229 birds stemming from four different captive populations. We then study the phenotypic and fitness consequences of the four inversion polymorphisms. (1) Heterokaryotypic individuals should exhibit increased embryo mortality rates if they are not able to suppress recombination within the inverted region completely or if they are unable to remove the unbalanced meiotic products. We test this prediction by analyzing the occurrence of natural embryonic deaths in 9,764 eggs. (2) Effect sizes in association studies between inversion genotypes and polygenic traits are expected to be higher than those of single SNPs in collinear parts of the genome, because multiple causal variants will be linked together by the inversion. Thus, inversions offer a nice opportunity to study the relative importance of additive versus dominance effects in a defined genomic region. We perform association studies between inversion genotypes and eight morphological traits and compare the contribution of additive and dominance effects on phenotypic variation. (3) Heterosis could balance inversion polymorphisms, and we test for dominance effects of the inversion genotypes on several fitness parameters (fecundity, siring success, or numbers of offspring).

Results

Detection and description of inversion polymorphisms

We screened the entire annotated zebra finch genome for the occurrence of large-scale inversion polymorphisms by first searching for high linkage disequilibrium (LD) over large distances and second by performing principle component analyses (PCA) for every chromosome to identify distinct groups.

Linkage disequilibrium patterns

In collinear parts of the zebra finch genome, LD (measured as r^2) > 0.1 extends maximally for 185 kb (Knief et al. unpublished). Chromosome *Tgu2* is depicted as an example thereof (Figure S1A). In contrast, four chromosomes (*Tgu5*, *Tgu11*, *Tgu13* and *TguZ*) showed extraordinarily large linkage blocks, spanning several megabases (12–63 Mb) and exhibiting the typical LD structure of an inversion (Andolfatto *et al.* 2001; Bansal *et al.* 2007; Stefansson *et al.* 2005), with LD being highest near the presumed inversion breakpoints and lower in the central parts of the inverted region (Figure 1A–D). Specifically, for chromosome *Tgu5* the LD block reached from 0.96–16.50 Mb, which is equivalent to 25% of the assembled chromosome and covered 325 genes. The inversion most likely includes the centromere, which is located maximally 5.12 Mb from the proximal chromosome end (Table 1; Warren *et al.* 2010). On chromosome *Tgu11*, the region of high LD extended from 0.086–12.29 Mb (equivalent to 57% of the assembled chromosome, spanning 250 genes), covering the most proximal SNP that had been genotyped. The centromere on chromosome *Tgu11* is located at the distal end of the chromosome at around 20 Mb (Chapter 5) and is thus located outside the LD block (Table 1). On chromosome *Tgu13*, almost the complete assembled chromosome was part of one large LD region (99%, covering 312 genes), starting from the second proximal SNP and covering the most distal SNP being genotyped (0.15–16.91 Mb). The centromere on chromosome *Tgu13* is located at the distal end of the chromosome, but since parts of the genome assembly are missing at this position (Backström *et al.* 2010, Chapter 5), we could not clarify whether the centromere was part of the LD region (Table 1). Finally, the physically largest LD block was found on the sex chromosome *TguZ*, extending from 5.91–68.83 Mb, which is equivalent to 86% of the total chromosome length and covering 619 genes. The centromere on chromosome *TguZ* is located at 27.62–28.12 Mb and is thus included in the LD region (Table 1).

Weaker signals of long range LD were also found on chromosomes *Tgu26* and *Tgu27* (Figure S1C, E), covering 2.05 Mb (42% of the total chromosome length, covering 57 genes) and 2.74 Mb (59% of the total chromosome length, covering 166 genes), respectively. The centromere on chromosome *Tgu26* is located at the proximal end of the chromosome (Chapter 5) and it is thus not clear whether it is covered by the LD region, and the centromere position on chromosome *Tgu27* is unknown. Since marker density was relatively low on both chromosomes, we could not clarify whether these LD regions were indeed inversions or rather assembly errors and we did not analyze them further (Figure S1D, F).

Principle component analyses

The four chromosomes found in the LD scan also showed inversion-typical patterns in the PCA (Figure 1E–H). The three autosomal inversions had two main homozygote haplotype clusters (with the heterozygous individuals in between) and the sex chromosome split into three main homozygote haplotype clusters (with the heterozygous individuals in between). The clusters were well defined on the autosomes but on chromosome *TguZ* the least common haplotype (haplotype C in Figure 1H) seemed to allow some recombination with

each of the two other haplotypes, making the clusters more diffuse. However, both the low average heterozygosity within each cluster of homozygotes compared to heterozygotes (Table 2) and median-joining networks (using Network v4.6.1.1 with standard settings; Bandelt *et al.* 1999) on phased SNP data at the inversion breakpoint (using Beagle v3.3.2; Browning & Browning 2007, Figure S2) further support the interpretation that the LD regions represent inversion polymorphisms. It should also be noted that chromosomes *Tgu5* and *TguZ* had been previously found cytogenetically to carry pericentric inversions and the breakpoints match to the LD region boundaries (Christidis 1986a; Itoh & Arnold 2005; Itoh *et al.* 2011).

At the moment we do not know which arrangement is ancestral and we thus name them according to their allele frequency (A = major haplotype, B = minor haplotype, C = least common haplotype on chromosome *TguZ*, Figure 1E–H, Table 2). The major alleles of all four inversion polymorphisms had similar frequencies in the range between 0.53–0.60 (Table 2). On chromosome *TguZ*, the least common allele (haplotype C) was rare (frequency 0.074, Table 2). All inversion polymorphisms were in Hardy-Weinberg equilibrium (Table 2).

Pooled heterozygosity and minor allele counts at inversion breakpoints

We calculated pooled heterozygosity (ZH_p) in 50 kb non-overlapping sliding windows along each chromosome (Figure 2A). Low values of ZH_p are indicative for regions with a high degree of fixation, for example due to positive selection (Rubin *et al.* 2010). To the contrary, high values of ZH_p would be expected, for example, in regions under balancing selection. We found pronounced peaks in ZH_p at the presumed breakpoints of the inversions on chromosomes *Tgu5*, *Tgu11* and *Tgu13*, which dropped almost to genome-wide average ZH_p in the interior of the inversions. Chromosome *Tgu11* had only one such peak, suggesting that the proximal breakpoint is missing in the current genome assembly. Diversity (the number of SNPs per 50 kb window) was slightly reduced at the presumed breakpoints of every inversion compared to the inversion's interior (mean diversity ± SD at breakpoints vs interior on *Tgu5*: 866 ± 225 vs 998 ± 237, *Tgu11*: 287 ± 182 vs 901 ± 207, *Tgu13*: 814 ± 289 vs 1092 ± 198, collinear autosomal genome-wide average diversity 975 ± 423). On chromosome *TguZ*, the entire inversion interior had high ZH_p values, which only dropped to the genome-wide average outside the inverted region. Diversity on *TguZ* was markedly reduced all along the inverted region including the presumed breakpoints and increased to genome-wide average outside the inversion (107 ± 74 vs 1077 ± 428, respectively).

The minor allele count frequency (MAC) spectra at the breakpoint regions for all autosomal inversion polymorphisms showed an admixture of the background MAC spectrum (Figure 2F) with a second MAC distribution whose local maximum matches with remarkable accuracy the allele frequency of the inversion types in those 100 individuals that had been used for the pooled-population sequencing (*Tgu5*: MAC local maximum at 0.34–0.36, frequency of the minor (B) haplotype in the sample 0.35, Figure 2B; *Tgu11*: MAC local

maximum at 0.48–0.50, frequency of minor (B) haplotype in the sample 0.47, Figure 2C; *Tgu13*: MAC local maximum at 0.48–0.50, frequency of minor (B) haplotype in sample 0.50, Figure 2D). The MAC spectrum along the entire inverted region on chromosome *TguZ* showed two local maxima at 0.28–0.30 and 0.42–0.44 (frequency of the B haplotype in the sample 0.30 and frequency of the major (A) haplotype in the sample 0.63, Figure 2E).

Association analyses and fitness consequences

In a first step, we tested whether heterokaryotypic individuals had higher embryo mortality rates than homozygotes in three captive populations of zebra finches. Then we performed association analyses between inversion genotypes and morphological phenotypes in the three captive and two wild populations of zebra finches. Finally, we tested for associations between inversion genotypes and multiple aspects of fitness in the three captive populations. We only briefly describe the statistical models in this section and provide complete model specifications in the Methods section.

Associations with embryo mortality

We fitted generalized linear mixed-effects models using embryo mortality as a binomial response variable (0 = embryo survived until hatch, 1 = embryo died naturally, $\bar{x}$ = 30%, n = 9,764 eggs analyzed) and the inversion genotypes of both parents as two predictors, coded as 1 = heterozygous and 0 = homozygous in three captive populations ("Seewiesen": $\bar{x}$ = 31.5%, n = 6,334 eggs analyzed, "Bielefeld": $\bar{x}$ = 22.9%, n = 1,170 eggs analyzed, "Cracow": $\bar{x}$ = 18.4%, n = 2,260 eggs analyzed).

Neither the mother's nor the father's inversion genotype had an effect on embryo mortality that survived strict Bonferroni correction (Figure 3). Notably, the two inversions which cover almost entire chromosomes (*Tgu13* and *TguZ*) had a weak effect in the expected direction in heterokaryotypic males (*Tgu13*: meta-analytic odds ratio, OR [95% confidence interval] = 1.17 [1.01–1.36], P = 0.040; *TguZ*: OR = 1.16 [0.99–1.36], P = 0.065).

Associations with phenotypes

To test whether inversion genotypes had an effect on morphological traits we fitted generalized linear mixed-effects models using eight different Z-transformed phenotypes as response variables (body mass, tarsus length, wing length, beak length, beak depth, beak width, digit ratio and visible fat deposition) and the inversion genotype simultaneously as an additive (-1 = homozygous for the minor allele, 0 = heterozygous, 1 = homozygous for the major allele, using one degree of freedom) and a dominance (0 = homozygous, 1 = heterozygous) predictor in three captive populations ("Seewiesen": n = 3,233 individuals, "Bielefeld": n = 1,096 individuals, "Cracow": n = 634 individuals) and two wild populations ("Fowlers Gap": n = 939 individuals, "Sydney": n = 265 individuals).

The inversions on chromosomes *Tgu5*, *Tgu11* and *TguZ* had strong additive effects on six out of the eight phenotypes. In total, nine out of 40 associations survived a strict Bonferroni correction (Figure 4). The major allele A of the inversion on chromosome *TguZ* had the strongest effects overall and increased visible fat deposition (nominal $P = 1 \times 10^{-16}$) and body mass (nominal $P = 2 \times 10^{-14}$) and had an opposing effect on tarsus length (nominal $P = 4 \times 10^{-6}$).

None of the inversions exhibited significant dominance effects on any of the phenotypes (Figure S3).

Associations with fitness parameters

We fitted generalized linear mixed-effects models using four measures of fitness components as response variables and the inversion genotype simultaneously as an additive and a dominance predictor (see above) in the three captive populations ("Seewiesen, "Bielefeld", "Cracow"). For females we included fecundity (number of eggs laid) and fitness (number of chicks that survived until an age of 35 days). For males we used siring success (number of eggs sired) and fitness (number of chicks that survived until an age of 35 days). Sample sizes are given in Table 3.

Neither in females nor in males did any of the inversion polymorphisms exhibit fitness effects (Figure 5, Figure S4).

Summary across morphological and fitness phenotypes

The effects of inversion genotypes on morphology and fitness could be so small that we missed them in our association studies due to low power (thereby committing a type II error). This is especially true for the fitness components because sample sizes were more limited and also we expected smaller effect sizes, at least for the additive genetic component (since natural selection should reduce the amount of additive genetic variance in fitness; Fisher 1930). In order to get the null distribution of effect sizes, we permuted inversion genotypes 100 times within each sex, ran the same mixed-models including the same fixed and random effects as for the empirical data and extracted additive and dominance effect size estimates.

The 40 empirical additive effect size estimates on morphology (5 comparisons between inversions x 8 phenotypes) clearly exceeded the random expectation, highlighting that probably also several of the associations which did not survive a strict Bonferroni correction were in fact real (Figure 6). In contrast, the 40 empirical dominance effects on morphology did not deviate from the random expectation (Figure 6). In fact, the mean of the 40 dominance effects was slightly smaller than the mean of the random expectation. Similarly, neither the 20 empirical additive effect size estimates on fitness (5 comparisons between inversions x 4 fitness components) nor the 16 dominance effects (the same 20 associations

excluding dominance effects on female fitness components on chromosome *TguZ*) deviated from the random expectation (Figure 6). Further, since we specifically expected to find positive dominance effects (heterosis) for fitness, we calculated the overall mean of the 16 dominance effects and found an effect size that was very close to zero (weighted d = 0.0019, P = 0.91).

Discussion

Here we describe four large inversion polymorphisms (12–63 Mb) in wild Australian zebra finches using molecular and population genetic tools. Two of them had been identified previously in cytogenetic screens since they shift the position of the centromere (Christidis 1986a; Itoh & Arnold 2005; Itoh *et al.* 2011). In total, the inverted regions span at least 8.7% of the zebra finch genome and 8.1% of all annotated genes (based on the Ensembl80 gene predictions). Although all polymorphisms are in Hardy-Weinberg equilibrium, their remarkably low (and also similar) major allele frequencies (range of the major allele 0.53–0.60) may indicate some sort of balancing selection acting on them. We find tentative evidence that the largest two inversions increase embryo mortality in heterokaryotypic males (but not so in females), which makes their high minor allele frequencies even less likely to be due to drift alone. However, although the inversions have an additive effect on several morphological traits, we do not find any dominant gene action and no balancing effects on several aspects of fitness in three captive populations.

Inversion polymorphisms on chromosomes Tgu5 and Tgu11 – small and simple

The inverted segments on chromosomes *Tgu5* and *Tgu11* span rather small proportions of the corresponding total chromosome lengths (covering 25% and 57% of the total chromosome, respectively). In the principle component analyses, individuals were only separated along PC1. PC2 was a single normal distribution, which indicates that there is no additional population substructure (Zhu & Zhang 2010; Zhu *et al.* 2002) due to a second rearrangement included, overlapping or independent of the first. Accordingly, the median-joining networks formed only two separated haplotype clusters and also the linkage disequilibrium patterns suggested that these are simple single inversions. Linkage disequilibrium and pooled heterozygosity were highest at the presumed breakpoints and dropped to the central regions, which are typical signs of gene flux due to double crossovers between two simple arrangements (Navarro *et al.* 1997). Parsimoniously, we would also expect that single detrimental crossovers should occur occasionally between the arrangements, leading to unbalanced gametes and embryo mortality, but we did not observe any increased embryo mortality rate in heterokaryotypic individuals.

At least a single crossover per chromosome is needed to ensure the proper segregation of homologous chromosomes in meiosis (Petronczki *et al.* 2003). However, on chromosomes *Tgu5* and *Tgu11* the collinear parts of the chromosome are large and a crossover is likely to

be initiated there, with no adverse effect on the meiotic products (White 1977). Alternatively, the inversions may be too rigid to synapse (a loop structure needs to be formed within the inverted region), thereby suppressing recombination (Anton *et al.* 2005). In either case, we suspect that single crossovers within the inverted region in heterokaryotypic individuals happen so rarely that they fall below the detection limit in our pedigrees.

On chromosome *Tgu5*, heterozygosity (Table 2) and diversity (the spread in the median-joining network) within cluster B is low, suggesting that it increased in frequency in the population and that type A is the ancestral state. Using the same line of argument, on chromosome *Tgu11* type A spread in the population and type B is the ancestral state.

Inversion polymorphisms on chromosomes Tgu13 and TguZ – complex and costly

The inverted regions on chromosomes *Tgu13* and *TguZ* are large in relation to the corresponding total chromosome lengths (covering 99% and 86% of the total chromosome, respectively). Our principle component analysis for chromosome *TguZ* showed at least three large haplotype clusters, but also the higher principle components (e.g. PC3) deviated from a normal distribution, suggesting an even more complex situation and an independent linkage block at the distal inversion breakpoint (data not shown). Itoh *et al.* (2011) described a single large pericentric inversion on chromosome *TguZ* and we suppose that they had identified the two most common types A and B, because our breakpoint locations match with the data in Itoh *et al.* (2011). However, in that case the allele frequency estimates by Itoh *et al.* (2011) in wild Australian zebra finches deviate from ours. Assuming no allele frequency change over time, their allele frequency estimates suggest that they had lumped haplotypes A and B together versus haplotype C.

The median-joining network and the number of shared SNPs suggest that haplotypes B and C on chromosome *TguZ* are more closely related with each other than with haplotype A. Judging from the fuzzy clusters formed in the PCA, gene flux between arrangements seems to happen, either between haplotypes A and C or between haplotypes B and C or between both of the pairs. Thus, inversion types B and C could be more related because of their shared ancestry or because of gene flux, and in the end we cannot separate these two possibilities.

The principle component analysis on chromosome *Tgu13* separated individuals largely along PC1. However, PC2 distinguished between at least two groups within inversion type A; yet these groups were not completely separated, indicating some gene flux between them. The higher principle components were normally distributed, suggesting that there is no additional population substructure (Zhu & Zhang 2010; Zhu *et al.* 2002). The LD patterns on chromosome *Tgu13* suggest that there is gene flux between the two main arrangements (types A and B) due to double crossovers.

There is tentative evidence that in heterokaryotypic males for both chromosomes *Tgu13* and *TguZ* embryo mortality rates are increased (by a weighted average of 4.5% for each of the chromosomes *Tgu13* and *TguZ* across populations). We suspect that these effects are not due to type I errors, because also in human males noticeable rates of unbalanced gametes are produced only when an inversion (both para- and pericentric) spans more than half of the chromosome (Anton *et al.* 2005; Morel *et al.* 2007; Yapan *et al.* 2014). However, the effect in humans is an order of magnitude (12-fold) larger than in zebra finches (Anton *et al.* 2005), indicating that zebra finches evolved a way to decrease recombination within inversion heterokaryotypes. Interestingly, the median-joining networks and PCA results suggest that there is a succession of inversions on chromosomes *Tgu13* (within haplotype A) and *TguZ* (haplotypes B and C appear to be more closely related), and accumulating inversions on a chromosome may be a way to more and more suppress recombination between inversion haplotypes (as, for example, in white-throated sparrows; Thomas *et al.* 2008).

Heterokaryotypic female zebra finches for the inversion on chromosome *Tgu13* did not show increased rates of embryo mortality, which indicates that they are either able to pass on abnormal meiotic products to the second polar body (as it has been found in *Drosophila* and maize in case of paracentric inversions; Morgan 1950; Roberts 1967) or that they are able to shut down recombination within the inverted region almost completely (as suggested for pericentric inversions in grasshoppers; White 1977).

For both chromosome *Tgu13* and *TguZ* it is difficult to conclude which haplotypes represent the ancestral states from patterns of diversity or the median-joining networks. Within the inversion on chromosome *TguZ*, diversity was reduced tenfold compared to the autosomes and also to the collinear parts of *TguZ* (which we also confirm in the data of Balakrishnan & Edwards 2009). Thus, the low diversity cannot simply be explained by the reduced effective population size for the Z chromosome (0.75 x N_e of the autosomes assuming strict monogamy and constant population size; Ellegren 2009) or by the smaller number of *TguZ* chromosomes in the pooled sequencing experiment (52 females x 1 *TguZ* + 48 males x 2 *TguZ* = 148 instead of 200 of each autosome; a reduction of 148/200 = 0.74, assuming a proportionally reduced coverage on *TguZ*). Even if in birds diversity on the Z chromosome is generally reduced three- to fourfold as compared to the autosomes (Corl & Ellegren 2012; Ellegren 2009; Wright *et al.* 2015), the reduction in zebra finches seems extraordinary, and together with the high heterozygosity indicates that rearrangements spread rather recently due to positive selection. However, the minimal sojourn time of at least one of the rearrangements is 1.2–2.8 million years, which is the estimated split time, with little subsequent gene flow, between Timor and Australian zebra finches (Balakrishnan & Edwards 2009), which are both polymorphic for one of the rearrangements (Itoh *et al.* 2011).

Fitness effects – no heterosis for viability in the wild or for fitness in captivity

The inversions on chromosomes *Tgu13* and *TguZ*, and probably also to a lesser extent the ones on chromosomes *Tgu5* and *Tgu11*, are costly in terms of increased embryo mortality whenever they are in the heterozygous state in males. Given an effective population size (N_e) of wild Australian zebra finches of $1.3 \times 10^6 – 7 \times 10^6$ (Balakrishnan & Edwards 2009) it is unlikely that the polymorphisms would have escaped purifying selection and be at frequencies (0.53–0.60) close to their fitness minimum (at an allele frequency of 0.5 the maximal number of individuals are heterozygous), if they do not confer a fitness advantage to their carriers. The minor allele count (MAC) frequency spectra at the presumed breakpoints of the autosomal inversions and along the whole inversion on chromosome *TguZ* showed an excess of alleles segregating at high frequency, which is characteristic for a balanced polymorphism (Fijarczyk & Babik 2015).

The simplest condition for a balanced polymorphism with two alleles is given when both homozygotes have lower fitness than heterokaryotypic individuals (heterosis; White 1977). In our sample of wild zebra finches, all four inversion polymorphisms were in Hardy-Weinberg equilibrium (HWE), indicating that there was no heterosis for viability at the time of sampling the individuals. However, it is possible that heterosis can only be measured after stressful environmental conditions in the wild, such as severe drought or heat. Such selective events could be so rare that they did not happen during the few years in which the sampled individuals lived. Furthermore, deviations from HWE are not necessarily expected if heterotic superiority depends on fecundity or siring success rather than viability (White 1977). Thus, we tested whether the inversions showed heterotic superiority in respect to several other aspects of fitness (female fecundity, male siring success and the number of offspring produced) in three captive populations of zebra finches, but the average of all effect sizes was close to zero (weighted d = 0.0019, P = 0.91). Thus, although we cannot rule out heterosis completely, other forms of balancing selection are more likely to keep the inversion polymorphic, which do not require heterokaryotypic superiority and do not lead to deviations from HWE, such as (negative) frequency dependent selection (White 1977).

Morphological effects – all additive, no dominant gene action

We found remarkable additive genetic effects of the inversion genotypes on several morphological traits, which were highly consistent across populations. Only one test for heterogeneity between populations was significant after Bonferroni correction, which was the association between beak length and haplotype A vs B on chromosome *TguZ* (Cochran's Q test P = 0.026; Cochran 1954). However, in contrast to the strong additive effects on morphology, all four inversion polymorphisms did not exhibit any dominant gene action.

Additive genotype-phenotype association studies typically find small effects of individual SNPs on a phenotype (Park *et al.* 2010; Yang *et al.* 2010), and associations are often difficult

to replicate between populations due to differences in LD structure (Kraft *et al.* 2009). Recently, we had tested several promising causal SNPs in collinear parts of the zebra finch genome for an additive association with the same morphological phenotypes and in the same populations as the ones studied here, and individual SNP effects were small and not consistent across populations (Chapter 6). The contrast between the diminishingly little additive effects of individual SNPs in collinear parts of the genome and the large and consistent additive effects of inversions is most likely due to differences in LD and highlights the polygenic nature of the quantitative traits that we studied: Whereas in collinear parts of the zebra finch genome LD decays rapidly (Balakrishnan & Edwards 2009, Chapter 6) making associations hard to detect, inversions capture hundreds of genes in extended defined haplotypes, which do not or hardly ever recombine. Thereby they combine the additive effects of many causal alleles and allelic interactions (epistasis).

Conclusion and Outlook

Large inversion polymorphisms are abundant in the *Estrildid* finch family (Christidis 1986a, b, 1987). Here we describe the inversion polymorphisms in one species belonging to this family, the zebra finch. We find polymorphic inversions on four out of its 32 annotated chromosomes. In each case, a novel haplotype has spread to about 50% allele frequency and has persisted for an extended period of time, but the selective forces maintaining the polymorphisms remain unclear. We neither find signs of heterosis for viability in a wild population nor for other fitness-related traits in captivity. Assuming some benefit to the individual (unmeasured heterosis or frequency dependent selection) or to the genotype itself (segregation distortion; Knief *et al.* 2015b), there remains a cost: heterokaryotypic males produce a higher proportion of inviable embryos, potentially due to single crossovers within the inverted region. It appears like past selection has effectively minimized this cost: (1) “Small” inversions (chromosomes *Tgu5* and *Tgu11*) do not observably increase the proportion of inviable embryos produced by heterokaryotypic individuals. These inversions maybe do not synapse regularly in meiosis, thereby reducing the risk of detrimental crossovers. (2) Heterokaryotypic females do not exhibit increased rates of embryo mortality even for the largest inversion on chromosome *Tgu13* and thus may have found a way to deposit the abnormal meiotic products (the dicentric single-crossover chromatids in case of a paracentric inversion) to the polar bodies. (3) The effects on embryo mortality in heterokaryotypic males for the two largest inversions on chromosomes *Tgu13* and *TguZ* are an order of magnitude smaller than in humans. We suspect that this could be due to selection favoring repeated inversions on the same chromosome, thereby effectively suppressing pairing of the inversion types during meiosis and inhibiting detrimental crossovers. Additionally, the highly skewed distribution of recombination events towards the chromosome ends in zebra finches (Backström *et al.* 2010) may be an adaptation to minimize crossovers in the inverted regions.

Material and Methods

Inversion discovery in wild zebra finches

Study population and phenotypes

We took blood samples from 1,059 wild adult zebra finches (530 females, 529 males) at Fowlers Gap, NSW, Australia, in two contiguous places (S 30°57' E 141°46' and S 31°04' E 141°50') between October to December 2010 and in April/May 2011. A detailed description of the study sites and catching procedure using a walk-in trap at feeders is provided in Griffith *et al.* (2008) and Mariette & Griffith (2012). In the following we refer to this population as "Fowlers Gap".

The following phenotypes were measured on all birds: right tarsus length, right wing length, beak length, beak depth, beak width, ratio of the length of the second to fourth digit of the right foot (measured twice and averaged) and body mass. Further details on the measurement procedures and summary statistics are given in Chapter 6. We included a score-based measure of visible fat on the ventral side at the furcular depression and at the abdomen (Bairlein 1995).

Population-level SNP data and sequencing

We sequenced pooled non-barcoded DNA samples from 100 of the 1,059 individuals on the Illumina HiSeq 2000 platform at the Institute of Clinical Molecular Biology (IKMB) at Kiel University, Germany. Software input parameters are provided in Knief *et al.* (2015a). Briefly, after mapping reads to the zebra finch genome assembly (WUSTL 3.2.4; Warren *et al.* 2010) using bwa (v0.5.9; Li & Durbin 2009), we calculated an average genome coverage of 247.5 x (using BEDTools v2.17.0; Quinlan & Hall 2010) and called around 23 million SNPs using GATK (v2.1-11-g13c0244; McKenna *et al.* 2010). SNPs with a minor allele count frequency (MAC) below 0.1 were rarer than expected due to an ascertainment bias in the SNP discovery pipeline (Knief *et al.* 2015a).

Pooled-population sequencing allows estimating diversity and allele frequencies across the genome (Schlötterer *et al.* 2014). Although individual-based data was missing we calculated a measure of heterozygosity (pooled heterozygosity, H_p) in 50 kb non-overlapping sliding windows along the autosomes (Rubin *et al.* 2010) as $H_p = 2 \times \sum n_{MAJ} \times \sum n_{MIN} / (\sum n_{MAJ} + \sum n_{MIN})^2$, where n_{MAJ} and n_{MIN} are counts of reads covering the major and minor allele, respectively, and $\sum n_{MAJ}$ and $\sum n_{MIN}$ are the sum of all these counts in a 50 kb window. We transformed the H_p-values into Z-scores (ZH_p) as $ZH_p = (H_p - \mu H_p) / \sigma H_p$.

SNP chip design

From the 23 million SNPs we designed an Illumina Infinium iSelect HD Custom BeadChip with 6,000 attempted bead types (Chapter 6). In short, 884 SNPs resided within candidate genes for an association study (Chapter 6) and were not used for the present study and

4,405 SNPs covered all assembled chromosomes except chromosome *Tgu16*. We attempted to position at least 40 physically evenly spaced SNPs on each chromosome, yet this was not possible for chromosomes *Tgu1B* (n = 33 SNPs) and *Tgu25* (n = 24 SNPs) because too few SNPs passed our filtering procedure (Chapter 6). In regions of the genome where the pooled heterozygosity was exceptionally high we increased the SNP density. Overall we intended to genotype 5,289 SNPs (which summed up to 6,000 bead types because we did not exclude C/G and A/T SNPs that require two bead types for genotyping) and the final chip delivered by Illumina contained 4,553 of these SNPs, with drop-outs being randomly distributed along chromosomes (Chapter 6).

Median marker spacing of SNPs on the chip was 243.17 kb (interquartile range [IQR] = 16.68–343.70 kb) on macrochromosomes (chromosomes *Tgu1*–*Tgu5*, *Tgu1A*), 239.03 kb (IQR = 20.57–355.14 kb) on microchromosomes (all other autosomes) and 174.63 kb (IQR = 161.11–179.40 kb) on chromosome *TguZ* (Chapter 6).

Individual genotyping and quality control

All 1,059 "Fowlers Gap" individuals were genotyped for the 4,553 SNPs at the IKMB at Kiel University. Quality control was done using the R package GWASTools (v1.6.2; Gogarten *et al.* 2012) and details are provided in Chapter 6. In summary, we removed 111 individuals with a missing call rate larger than 0.05 (which was due to DNA extraction problems but these birds were genotyped in the follow-up study and are included in the "Sydney" sample, see below), leaving 948 individuals, and 152 SNPs that did not form defined genotype clusters, had high missing call rates (missing rate > 0.1), were monomorphic, deviated strongly from Hardy-Weinberg equilibrium (Fisher's exact test $P < 0.05/4,553$), or because their position in the zebra finch genome assembly was likely not correct, leaving 4,401 SNPs (Chapter 6).

Linkage disequilibrium (LD) calculations

Inversion polymorphisms lead to extensive LD across the inverted region, with the highest LD near the inversion breakpoints because recombination in these regions is almost completely suppressed in inversion heterozygotes (Andolfatto *et al.* 2001; Bansal *et al.* 2007; Stefansson *et al.* 2005).

In order to screen for inversion polymorphisms we did not resolve genotypic data into haplotypes and thus based all LD calculation on composite LD (Weir 1996). We calculated the squared Pearson's correlation coefficient (r^2) as a standardized measure of LD between every two SNPs on a chromosome genotyped in the 948 individuals (Weir 1979; Wellek & Ziegler 2009).

Principle component analyses (PCA)

Inversion polymorphisms appear as a kind of localized population substructure within a genome because the two inversion haplotypes do not or only rarely recombine (Deng *et al.*

2008; Ma & Amos 2012), and this substructure is amenable to principle component analysis (PCA; Price *et al.* 2006). In case of an inversion polymorphism, we expected three clusters when plotting principle component 1 (PC1) against principle component 2 (PC2) – the two inversion homozygotes at both sides with their mixture, the heterozygotes, in between. Subsequently, the principal component scores allowed us to classify every individual as being either homozygous for one or the other inversion genotype or as being heterozygous (Ma & Amos 2012).

We performed PCA on the quality-checked SNP set of the 948 individuals using the R package SNPRelate (v0.9.14; Zheng *et al.* 2012). On the macrochromosomes, we first used a sliding window approach analyzing 50 SNPs at a time, moving 5 SNPs to the next window. Since the sliding window approach did not provide us with more details than including all SNPs on a chromosome at once in the PCA, we only present the results from the full SNP set per chromosome. On the microchromosomes, the number of SNPs was restricted and thus from the start we performed PCA only including all SNPs residing on a chromosome.

In collinear parts of the genome composite LD > 0.1 does not extend beyond 185 kb (Figure S1A; Chapter 6). Thus, we also filtered the SNP set to include only SNPs in the PCA that were spaced by more than 185 kb (filtering was done using the "earliest finish time" greedy algorithm). Both the full and the filtered SNP set gave qualitatively the same results and in the manuscript we only present results based of the full SNP set for clarity and because tag SNPs (see below) were defined on these data (we present PCA plots in the Supplement, Figure S5).

Tag SNP selection

For each of the identified inversion polymorphisms we selected combinations of SNPs that uniquely identified the inversion types (composite LD of individual SNPs $r^2 > 0.9$). For each inversion polymorphism we calculated standardized composite LD between the eigenvector of PC1 (and PC2 in case of three inversion types) and the SNPs on the respective chromosome as the squared Pearson's correlation coefficient. Then, for each chromosome, we selected SNPs that tagged the inversion haplotypes uniquely. We tried to pick tag SNPs in both break point regions of an inversion, spanning the largest physical distance possible (Table S1). Using only information from the tag SNPs and a majority vote decision rule all individuals from Fowlers Gap were assigned to the correct inversion genotypes for chromosomes *Tgu5*, *Tgu11* and *Tgu13* (Figure S6A, B, C). Since clusters are not as well defined for chromosome *TguZ* as for the other three autosomes, there is some ambiguity in cluster borders. Using a unanimity decision rule, the inferred inversion genotypes from the tag SNPs correspond very well with the PCA results, leaving some individuals uncalled (Figure S6D).

Follow-up genotyping and phenotyping in captive populations

Study populations

In order to study phenotypic and fitness effects of the inversion polymorphisms we genotyped all 15 tag SNPs (Table S1) in an additional 5,229 birds stemming from four different populations: (1) A captive population held at the Max Planck Institute for Ornithology in Seewiesen, Germany, (n=3,233 individuals; study population 18 in Forstmeier *et al.* 2007) with a complete pedigree covering eight generations, of which the last seven were genotyped completely. In the following, we refer to this population as "Seewiesen". (2) A recently wild-derived population held at the Max Planck Institute for Ornithology in Seewiesen (n=1,096 individuals; originating from study population 4 in Forstmeier *et al.* 2007) with a complete pedigree covering six generations, of which the last four generations were genotyped completely. We refer to this population as "Bielefeld". (3) A population that was produced by crossing individuals from a captive population held in Cracow (study population 11 in Forstmeier *et al.* 2007) with the "Seewiesen" population (n=634 individuals) with a complete pedigree covering three generations, of which all generations were genotyped completely. Hereafter, we refer to this population as "Cracow". (4) Wild birds held at Macquarie University in Sydney, Australia, and additional wild birds from Fowlers Gap, which were excluded in the initial genotyping due to DNA extraction problems (see above; n=265 individuals without pedigree information). In the following, we refer to these birds as "Sydney".

Genotyping, quality control and inversion haplotype inference

Tag SNPs (n = 15, chromosomes *Tgu5* + *Tgu11* + *Tgu13* + *TguZ* = 3 + 3 + 3 + 6) were included in three Sequenom genotyping assays (plexes) which in total consisted of 62 SNPs. All 5,229 individuals were genotyped according to the manufacturer's users guide on the Sequenom MassARRAY iPLEX platform (Gabriel *et al.* 2009) at the IKMB at Kiel University. Genotypes were called using the MassARRAY Typer (v4.0) software with standard settings.

The quality control procedure of genotype calls has been described previously and involved inheritance checks using PedCheck (v1.00; O'Connell & Weeks 1998), the inference of null alleles and a comparison of 16,013 genotype calls of individuals that were genotyped using both the Illumina and Sequenom genotyping platform. All tests indicated high genotyping accuracy (Chapter 6).

We inferred inversion genotypes for each individual as in the "Fowlers Gap" population using a majority vote decision rule. Founders of all four populations that produced offspring (n = 239 individuals) were run on both the Illumina and Sequenom genotyping platform. Thus, we used the SNP-loadings on PC1 and PC2 from the PCA of the "Fowlers Gap" birds on the population founders to calculate a PCA score for each individual (Figure S7A–D) and compared the inversion genotypes inferred by PCA and tag SNPs. There was complete agreement between the two methods for the autosomal inversion genotypes (Table S2). In the "Bielefeld" population a recombinant haplotype for chromosome *TguZ* was common (26

out of 74 founder individuals, Figure S7D) and we changed the majority vote decision rule to a unanimity decision rule, which reduced the number of individuals assigned to a specific inversion genotype and removed all wrongly assigned individuals (Figure S6D, Table S2). 1,062 "Seewiesen" individuals had been previously genotyped with a different set of 37 SNPs on chromosome *TguZ* (Backström *et al.* 2010) and we compared the inversion haplotype inference between Sequenom and the PCA results using these 37 SNPs and they agreed completely (Table S2). Using these inference rules there was not a single inheritance error of an inversion genotype out of 35,584 inheritance events.

Morphological phenotyping

The same morphological phenotypes as for the wild birds from the "Fowlers Gap" population were taken on every bird using the same methodology. Each phenotypic measurement was taken once (twice for digit ratio) per individual such that phenotypic values and their measurement errors between the "Fowlers Gap" and the other populations are comparable (Falconer & Mackay 1996). Descriptive statistics for each trait (except visible fat deposition) are summarized in Chapter 6.

Embryo mortality and fitness parameters

Embryo mortality and fitness measures were taken from the three captive populations held at the Max Planck Institute for Ornithology ("Seewiesen", "Bielefeld" and "Cracow"). Data was collected in four different experimental set-ups (Table 3): (1) Cage laying: Pairs that were allowed to lay eggs in isolated cages, and whose eggs were removed after four to five days or whose eggs were cross-fostered. Of those pairs we analyzed fecundity (i.e. the number of eggs laid including infertile eggs) as a female fitness trait and embryo mortality rate. (2) Cage breeding: Pairs that were allowed to lay eggs and raise offspring in isolated cages. We analyzed fecundity (including infertile eggs) as a female fitness trait, the number of fledglings (≥ 35 days of age) as female and male fitness and embryo mortality rate. (3) Aviary laying: Pairs that laid eggs in communal aviaries, and whose eggs were removed after four to five days or cross-fostered and parentage was assigned genetically. We analyzed fecundity (i.e. the number of eggs laid excluding infertile eggs) as a female fitness trait, siring success as a male fitness trait and embryo mortality rate. (4) Aviary breeding: Pairs that bred in communal aviaries, and who raised offspring. We analyzed fecundity (excluding infertile eggs) as a female fitness trait, siring success as a male fitness trait and the number of fledglings (≥ 35 days of age) as female and male fitness and embryo mortality rate. Sample sizes are given in Table 3.

Association analyses and software

Software

All analyses were performed in R (v3.0.2; R Core Team 2013): Mixed-effects models were fitted using ASReml-R (v3; Gilmour *et al.* 2009). Inbreeding coefficients for each individual were calculated using the pedigreemm package (v0.3-1; Vazquez *et al.* 2010). Meta-analyses

using a fixed-effect model were done with the meta.summaries() function in the rmeta package (v2.16; Lumley 2012).

Embryo mortality

We fitted mixed-effects generalized linear models with embryo mortality as the binomial response variable (0 = embryo survived until hatch, 1 = embryo died naturally) and the mother's and the father's inversion genotypes as two predictors (underdominance effect coded as 0 = homozygous for either inversion genotype, 1 = heterozygous). We included the pedigree-based inbreeding coefficient of the embryo (as a covariate) and the pair identity and the mother's and father's additive genetic relatedness matrices as random effects. We also controlled for experimental setup (cage versus aviary breeding and cage versus aviary laying) by fitting it as a factor with four levels and the specific experiment as an additional random effect.

Morphological phenotypes

We used the Z-transformed morphological phenotypes as the response variable in univariate mixed-effects linear models and the individual's inversion genotype simultaneously as an additive effect (coded as -1 = homozygous for the minor allele, 0 = heterozygous, 1 = homozygous for the major allele using one degree of freedom) and a dominance effect (coded as 0 = homozygous for either inversion genotype, 1 = heterozygous) as two predictors.

In the wild "Fowlers Gap" population we included sex (factor with two levels), the individual's age (covariate) and season (factor with two levels referring to the catching seasons November/December 2010 and April/May 2011) as fixed effects and for body mass we also included time of day (covariate). We corrected for overall body size by fitting body mass as a covariate (when body mass was the independent variable we used tarsus length instead). We controlled for relatedness by fitting a genetic relatedness matrix as a random effect (see Chapter 6 for details).

In the "Sydney" population we used the same fixed effects as in the "Fowlers Gap" analyses and additionally included the observer of the measurement (factor with two levels). Since we neither had a pedigree nor genome-wide SNP data for these birds, we did not control for relatedness.

In the three captive populations ("Seewiesen", "Bielefeld", and "Cracow") we fitted the individual's sex (factor), the observer of the measurement (factor with maximally five levels), the individual's pedigree-based inbreeding coefficient (covariate) and its age (covariate) as fixed effects. For body mass we included time of day (covariate) and in the "Seewiesen" population we included whether the bird was measured dead or live (factor) for the three beak morphology traits as fixed effects (see Knief *et al.* 2012 for details). We

controlled for body size as in the "Fowlers gap" population. We controlled for relatedness by fitting a additive genetic relatedness matrix as a random effect.

Fitness parameters

We fitted univariate mixed-effects linear models using each of the four fitness parameters (female fecundity, male siring success, female fitness, male fitness) as the independent variable in the three captive populations and the inversion genotype of the individual coded as an additive (-1 = homozygous for the minor allele, 0 = heterozygous, 1 = homozygous for the major allele using one degree of freedom) and dominance (0 = homozygous for either inversion genotype, 1 = heterozygous) effect as two predictors. We first square-root-transformed the independent variables to improve model fit and then Z-transformed them prior to analysis but also fitted Poisson models, which qualitatively gave the same results.

As fixed effects we included the pedigree-based inbreeding coefficient of the individual (covariate) and for female fecundity, male siring success, and female and male fitness we added the number of days an individual spent in the respective experiment. We fitted an additive genetic relatedness matrix as a random effect and since we had multiple measures per individual we also fitted a permanent environment random effect. We controlled for experimental setup (cage versus aviary breeding and cage versus aviary laying) by fitting it as a fixed effect (factor with four levels) and the specific experiment as an additional random effect.

Acknowledgements

We thank C. Beckmann, A. Sager, and M. Mariette for assistance with sampling the wild birds. Wild birds were sampled and banded under approval of the Macquarie University Animal Ethics Committee, the Australian Bird and Bat Banding Scheme, and a Scientific License from NSW National Parks and Wildlife Service. We are grateful to K. Martin, J. Rutkowska, M. Ihle and J. Schreiber for help with breeding and phenotyping and S. Bauer, E. Bodendorfer, A. Grötsch, A. Kortner, K. Martin, P. Neubauer, F. Weigel, and B. Wörle for animal care and help with breeding. We further thank M. Schneider for laboratory work in Seewiesen and M. Schilhabel, the Next-Generation Sequencing team and the Genotyping team at the IKMB in Kiel for laboratory work. This study was funded by the Max Planck Society (B.K.), with the zebra finch study at Fowler's Gap funded by support to S.C.G. from the Australian Research Council. U.K. is part of the International Max Planck Research School for Organismal Biology.

References

Andolfatto P, Depaulis F, Navarro A (2001) Inversion polymorphisms and nucleotide variability in *Drosophila*. *Genetical Research* **77**, 1–8.

Anton E, Blanco J, Egozcue J, Vidal F (2005) Sperm studies in heterozygote inversion carriers: a review. *Cytogenetic and Genome Research* **111**, 297–304.

Avelar AT, Perfeito L, Gordo I, Ferreira MG (2013) Genome architecture is a selectable trait that can be maintained by antagonistic pleiotropy. *Nature Communications* **4**, Artn 2235.

Backström N, Forstmeier W, Schielzeth H*, et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research* **20**, 485–495.

Bairlein F (1995) Manual of field methods. *European-African songbird migration network. Institut für Vogelforschung, Wilhelmshaven*.

Balakrishnan CN, Edwards SV (2009) Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* **181**, 645–660.

Balanya J, Oller JM, Huey RB, Gilchrist GW, Serra L (2006) Global genetic change tracks global climate warming in *Drosophila subobscura*. *Science* **313**, 1773–1775.

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37–48.

Bansal V, Bashir A, Bafna V (2007) Evidence for large inversion polymorphisms in the human genome from HapMap data. *Genome Research* **17**, 219–230.

Bhatt S, Moradkhani K, Mrasek K*, et al.* (2009) Breakpoint mapping and complete analysis of meiotic segregation patterns in three men heterozygous for paracentric inversions. *European Journal of Human Genetics* **17**, 44–50.

Browning SR, Browning BL (2007) Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *American Journal of Human Genetics* **81**, 1084–1097.

Cheng CD, White BJ, Kamdem C*, et al.* (2012) Ecological genomics of *Anopheles gambiae* along a latitudinal cline: a population-resequencing approach. *Genetics* **190**, 1417–1432.

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.

Christidis L (1987) Chromosomal evolution within the family Estrildidae (Aves) III. The Estrildae (waxbill finches). *Genetica* **72**, 93–100.

Cochran WG (1954) The combination of estimates from different experiments. *Biometrics* **10**, 101–129.

Conrad DF, Hurles ME (2007) The population genetics of structural variation. *Nature Genetics* **39**, S30–S36.

Corl A, Ellegren H (2012) The genomic signature of sexual selection in the genetic diversity of the sex chromosomes and autosomes. *Evolution* **66**, 2138–2149.

Deng LB, Zhang YZ, Kang J*, et al.* (2008) An unusual haplotype structure on human chromosome 8p23 derived from the inversion polymorphism. *Human Mutation* **29**, 1209–1216.

Dobzhansky T (1970) *Genetics of the evolutionary process*. Columbia University Press, New York, NY, USA.

Dubinin NP, Tiniakov GG (1946) Seasonal cycle and inversion frequency in populations. *Nature* **157**, 23–24.

Ellegren H (2009) The different levels of genetic diversity in sex chromosomes and autosomes. *Trends in Genetics* **25**, 278–284.

Falconer D, Mackay T (1996) *Introduction to quantitative genetics*, 4th edn. Longmann, Harlow, UK.

Fang Z, Pyhäjärvi T, Weber AL*, et al.* (2012) Megabase-scale inversion polymorphism in the wild ancestor of maize. *Genetics* **191**, 883–894.

Feder JL, Gejji R, Powell THQ, Nosil P (2011) Adaptive chromosomal divergence driven by mixed geographic mode of evolution. *Evolution* **65**, 2157–2170.

Fijarczyk A, Babik W (2015) Detecting balancing selection in genomes: limits and prospects. *Molecular Ecology* **24**, 3529–3545.

Fisher RA (1930) *The genetical theory of natural selection*. Oxford University Press, Oxford, UK.

Forstmeier W, Segelbacher G, Mueller JC, Kempenaers B (2007) Genetic variation and differentiation in captive and wild zebra finches (*Taeniopygia guttata*). *Molecular Ecology* **16**, 4039–4050.

Gabriel S, Ziaugra L, Tabbaa D (2009) SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Curr Protoc Hum Genet* **Chapter 2**, Unit 2 12.

Gelman A (2008) Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* **27**, 2865–2873.

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009) *Asreml user guide release 3.0*. VSN International Ltd, Hemel Hempstead, UK.

Gogarten SM, Bhangale T, Conomos MP*, et al.* (2012) GWASTools: an R/Bioconductor package for quality control and analysis of genome-wide association studies. *Bioinformatics* **28**, 3329–3331.

Griffin DK, Robertson LBW, Tempest HG, Skinner BM (2007) The evolution of the avian genome as revealed by comparative molecular cytogenetics. *Cytogenetic and Genome Research* **117**, 64–77.

Griffith SC, Pryke SR, Mariette M (2008) Use of nest-boxes by the Zebra Finch (*Taeniopygia guttata*): implications for reproductive success and research. *Emu* **108**, 311–319.

Hoffmann AA, Rieseberg LH (2008) Revisiting the impact of inversions in evolution: from population genetic markers to drivers of adaptive shifts and speciation? *Annual Review of Ecology Evolution and Systematics* **39**, 21–42.

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

Itoh Y, Arnold AP (2005) Chromosomal polymorphism and comparative painting analysis in the zebra finch. *Chromosome Research* **13**, 47–56.

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

Joron M, Frezal L, Jones RT*, et al.* (2011) Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **477**, 203–207.

Kirkpatrick M (2010) How and why chromosome inversions evolve. *PLoS Biol* **8**, e1000501.

Kirkpatrick M, Barton N (2006) Chromosome inversions, local adaptation and speciation. *Genetics* **173**, 419–434.

Knief U, Hemmrich-Stanisak G, Wittig M*, et al.* (2015a) Quantifying realized inbreeding in wild and captive animal populations. *Heredity (Edinb)* **114**, 397–403.

Knief U, Schielzeth H, Ellegren H, Kempenaers B, Forstmeier W (2015b) A prezygotic transmission distorter acting equally in female and male zebra finches *Taeniopygia guttata*. *Molecular Ecology* **24**, 3846–3859.

Knief U, Schielzeth H, Kempenaers B, Ellegren H, Forstmeier W (2012) QTL and quantitative genetic analysis of beak morphology reveals patterns of standing genetic variation in an Estrildid finch. *Molecular Ecology* **21**, 3704–3717.

Kopp M, Hermisson J (2006) The evolution of genetic architecture under frequency-dependent disruptive selection. *Evolution* **60**, 1537–1550.

Kraft P, Zeggini E, Ioannidis JPA (2009) Replication in genome-wide association studies. *Statistical Science* **24**, 561–573.

Krimbas CB, Powell JR (1992) *Drosophila inversion polymorphism*. CRC Press, Boca Raton, Florida, USA.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760.

Lowry DB, Willis JH (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol* **8**, e1000500.

Lowther JK (1961) Polymorphism in white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Canadian Journal of Zoology* **39**, 281–292.

Lumley T (2012) rmeta: Meta-analysis. *R package version 2.16*.

Lyon MF (2003) Transmission ratio distortion in mice. *Annual Review of Genetics* **37**, 393–408.

Ma JZ, Amos CI (2012) Investigation of inversion polymorphisms in the human genome using principal components analysis. *Plos One* **7**, e40224.

Mariette MM, Griffith SC (2012) Conspecific attraction and nest site selection in a nomadic species, the zebra finch. *Oikos* **121**, 823–834.

McAllister BF, Sheeley SL, Mena PA, Evans AL, Schlotterer C (2008) Clinal distribution of a chromosomal rearrangement: a precursor to chromosomal speciation? *Evolution* **62**, 1852–1865.

McKenna A, Hanna M, Banks E*, et al.* (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**, 1297–1303.

Morel F, Laudier B, Guerif F*, et al.* (2007) Meiotic segregation analysis in spermatozoa of pericentric inversion carriers using fluorescence in-situ hybridization. *Human Reproduction* **22**, 136–141.

Morgan DT (1950) A cytogenetic study of inversions in *Zea mays*. *Genetics* **35**, 153–174.

Navarro A, Betran E, Barbadilla A, Ruiz A (1997) Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics* **146**, 695–709.

Navarro A, Ruiz A (1997) On the fertility effects of pericentric inversions. *Genetics* **147**, 931–933.

O'Connell JR, Weeks DE (1998) PedCheck: A program for identification of genotype incompatibilities in linkage analysis. *American Journal of Human Genetics* **63**, 259–266.

Park JH, Wacholder S, Gail MH*, et al.* (2010) Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nature Genetics* **42**, 570–U139.

Petronczki M, Siomos MF, Nasmyth K (2003) Un ménage à quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112**, 423–440.

Price AL, Patterson NJ, Plenge RM*, et al.* (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* **38**, 904–909.

Price T (2008) *Speciation in birds*. Roberts and Company, Greenwood Village, Colorado, USA.

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842.

R Core Team (2013) R: a language and environment for statistical computing. *3.0.2*.

Roberts PA (1967) A positive correlation between crossing over within heterozygous pericentric inversions and reduced egg hatch of *Drosophila* females. *Genetics* **56**, 179–187.

Rubin CJ, Zody MC, Eriksson J*, et al.* (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* **464**, 587–593.

Schielzeth H (2010) Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* **1**, 103–113.

Schlötterer C, Tobler R, Kofler R, Nolte V (2014) Sequencing pools of individuals — mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics* **15**, 749–763.

Shields GF (1982) Comparative avian cytogenetics: a review. *Condor* **84**, 45–58.

Stefansson H, Helgason A, Thorleifsson G*, et al.* (2005) A common inversion under selection in Europeans. *Nature Genetics* **37**, 129–137.

Sturtevant AH, Mather E (1938) The interrelations of inversions, heterosis and recombination. *American Naturalist* **72**, 447–452.

Swanson CP, Merz T, Young WJ (1981) *Cytogenetics: the chromosome in division, inheritance, and evolution*. Prentice-Hall, Englewood Cliffs, NJ, USA.

Thomas JW, Caceres M, Lowman JJ*, et al.* (2008) The chromosomal polymorphism linked to variation in social behavior in the white-throated sparrow (*Zonotrichia albicollis*) is a complex rearragement and suppressor of recombination. *Genetics* **179**, 1455–1468.

Thorneycroft HB (1975) Cytogenetic study of white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Evolution* **29**, 611–621.

Traulsen A, Reed FA (2012) From genes to games: cooperation and cyclic dominance in meiotic drive. *Journal of Theoretical Biology* **299**, 120–125.

Umina PA, Weeks AR, Kearney MR, McKechnie SW, Hoffmann AA (2005) A rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. *Science* **308**, 691–693.

Vazquez AI, Bates DM, Rosa GJM, Gianola D, Weigel KA (2010) Technical note: An R package for fitting generalized linear mixed models in animal breeding. *Journal of Animal Science* **88**, 497–504.

Warren WC, Clayton DF, Ellegren H*, et al.* (2010) The genome of a songbird. *Nature* **464**, 757–762.

Weir BS (1979) Inferences about linkage disequilibrium. *Biometrics* **35**, 235–254.

Weir BS (1996) *Genetic data analysis II: methods for discrete population genetic data*. Sinauer Associates, Sunderland, USA.

Wellek S, Ziegler A (2009) A genotype-based approach to assessing the association between single nucleotide polymorphisms. *Human Heredity* **67**, 128–139.

White MJD (1977) *Animal cytology and evolution*. Cambridge University Press, UK, Cambridge.

Wright AE, Harrison PW, Zimmer F*, et al.* (2015) Variation in promiscuity and sexual selection drives avian rate of faster-Z evolution. *Molecular Ecology* **24**, 1218–1235.

Yang JA, Benyamin B, McEvoy BP*, et al.* (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics* **42**, 565–U131.

Yapan CC, Beyazyurek C, Ekmekci CG, Kahraman S (2014) The largest paracentric inversion, the highest rate of recombinant spermatozoa. Case report: 46,Xy, Inv(2)(Q21.2q37.3) and literature review. *Balkan Journal of Medical Genetics* **17**, 55–61.

Zheng XW, Levine D, Shen J*, et al.* (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* **28**, 3326–3328.

Zhu X, Zhang S (2010) Population-based association studies. In: *Handbook on analyzing human genetic data: computational approaches and software.* (eds. Lin S, Zhoa H). Springer, Berlin, Heidelberg, Germany.

Zhu XF, Zhang SL, Zhao HY, Cooper RS (2002) Association mapping, using a mixture model for complex traits. *Genetic Epidemiology* **23**, 181–196.

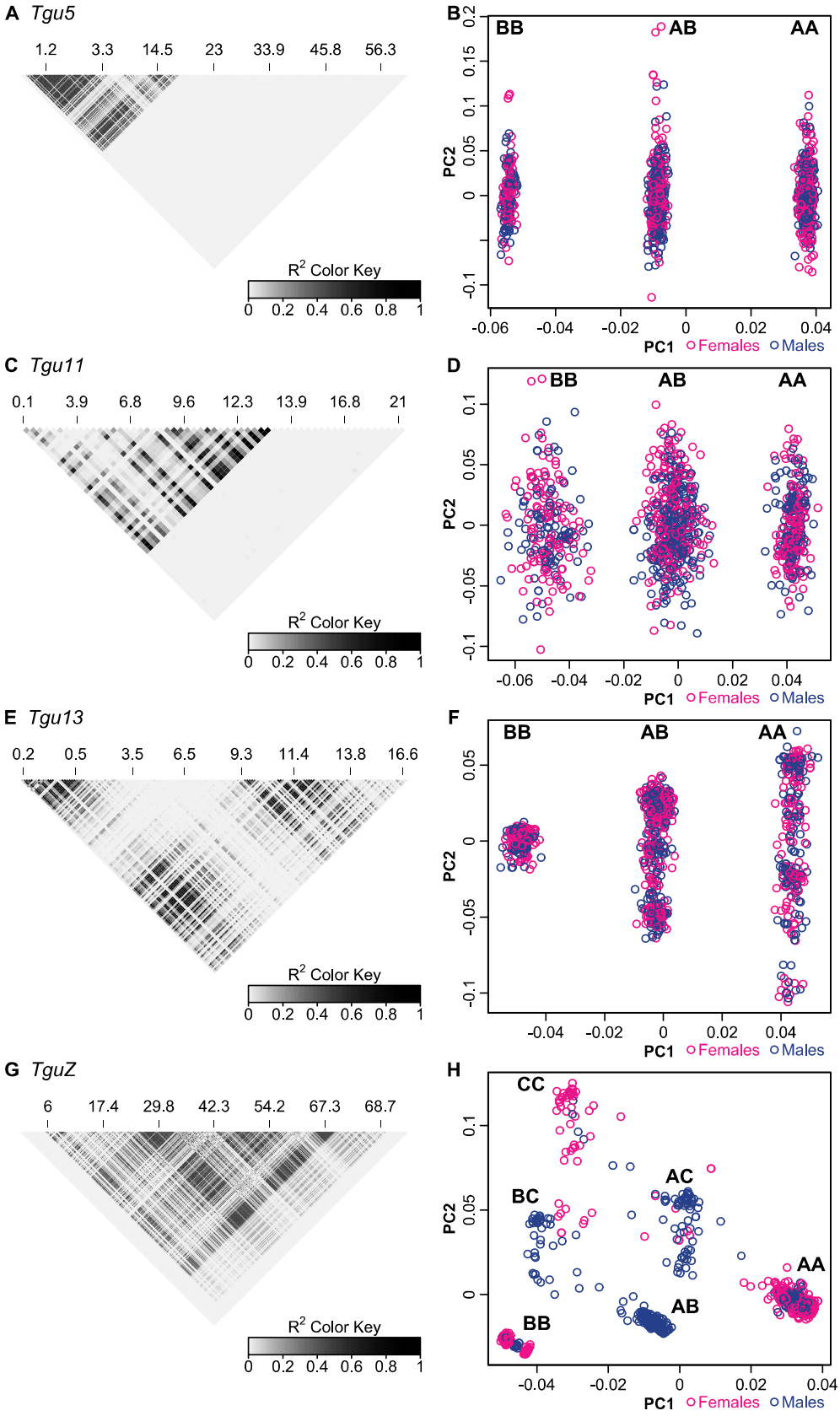


Figure 1: *The left panel depicts linkage disequilibrium (LD) and the right panel principal component analysis (PCA) results along chromosomes (A, B) Tgu5, (C, D) Tgu11, (E, F) Tgu13 and (G, H) TguZ. Above the LD plots marker positions in Mb are given. PCA included all SNPs*

on the respective chromosome. Note that in (H) there are a few (n=18) females that were called as heterozygous for the inversion. These are carriers of occasional double-crossovers.

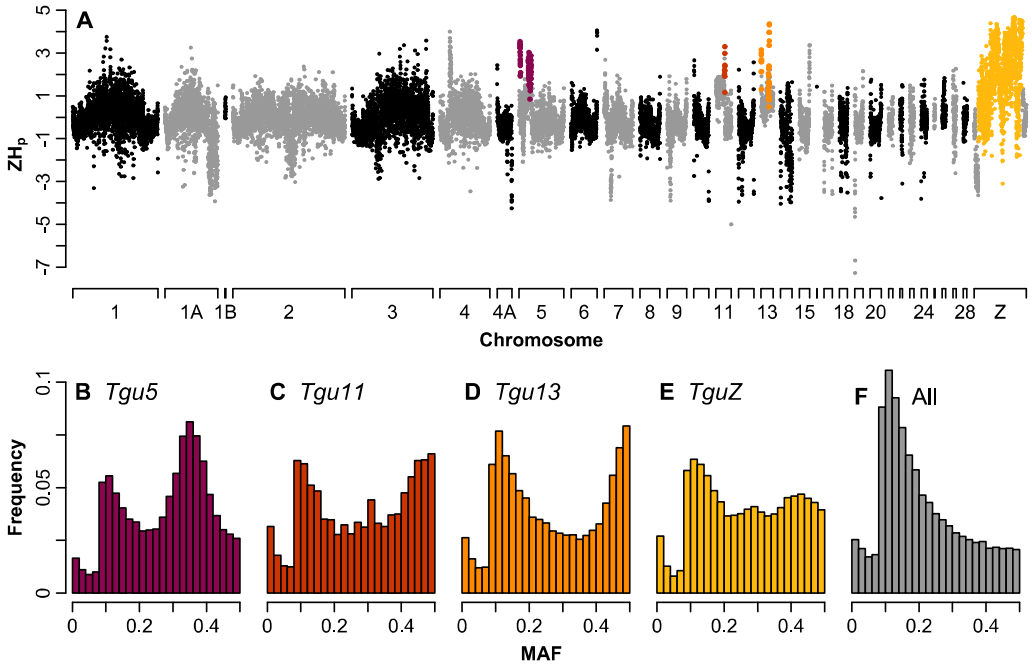


Figure 2: *(A) Pooled heterozygosity (ZH_p) in 50 kb sliding windows along each chromosome in the zebra finch genome. For the highlighted areas, which are the presumed inversion breakpoints on the autosomes and the entire inversion interior on the sex chromosome, the minor allele count frequency spectra (MAC) are shown for (B) chromosome Tgu5, (C) Tgu11, (D) Tgu13 and (E) TguZ. (F) For comparison, the MAC of all remaining SNPs, which peaks at an allele frequency of around 0.1 because SNPs with a lower frequency were not unambiguously called.*

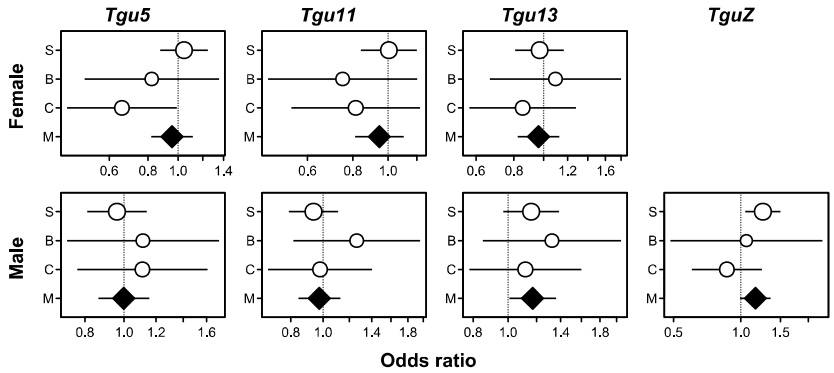


Figure 3: *Dominance effects (Odds ratio ± 95% confidence intervals) of mother's and father's inversion karyotype on embryo mortality in three captive populations (S = "Seewiesen", B = "Bielefeld", C = "Cracow" and M = meta-analytic summary). An odds ratio > 1 indicates an increased rate of embryo mortality in the offspring of females (top row) or males (bottom row) that are heterozygous for one of the four inversions on chromosomes Tgu5, Tgu11, Tgu13 and TguZ. The point sizes reflect log-transformed sample sizes.*

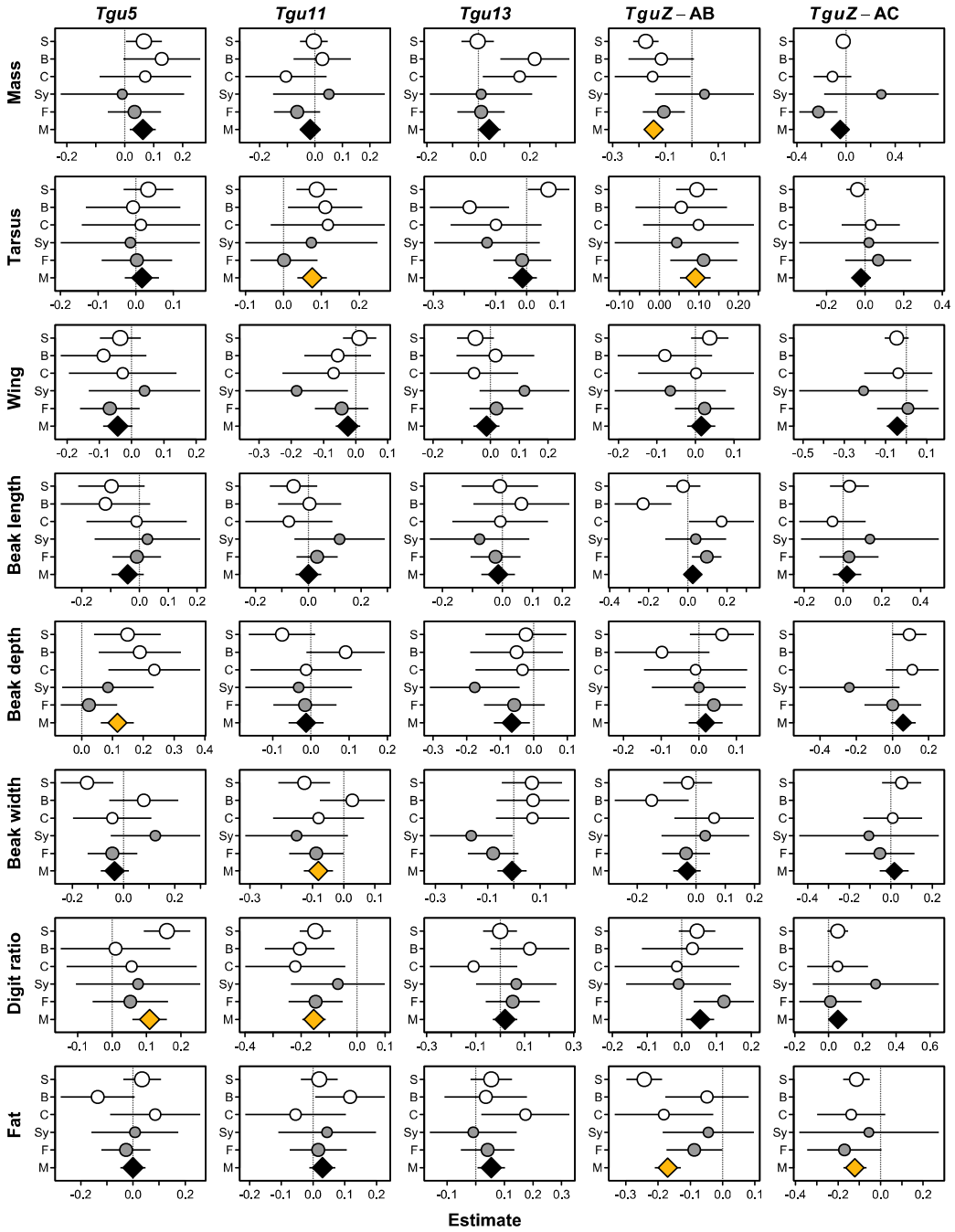


Figure 4: *Additive effects of the minor inversion allele ± 95% confidence intervals on morphological phenotypes in four captive and a wild populations (S = "Seewiesen", B = "Bielefeld", C = "Cracow", Sy = "Sydney", F = "Fowlers Gap" and M = meta-analytic summary). Effect size estimates are regression slopes of Z-transformed phenotypes over inversion genotypes (while simultaneously fitting dominance effects). The point sizes reflect log-transformed sample sizes. Meta-analytic summaries that survive strict Bonferroni correction are highlighted in yellow.*

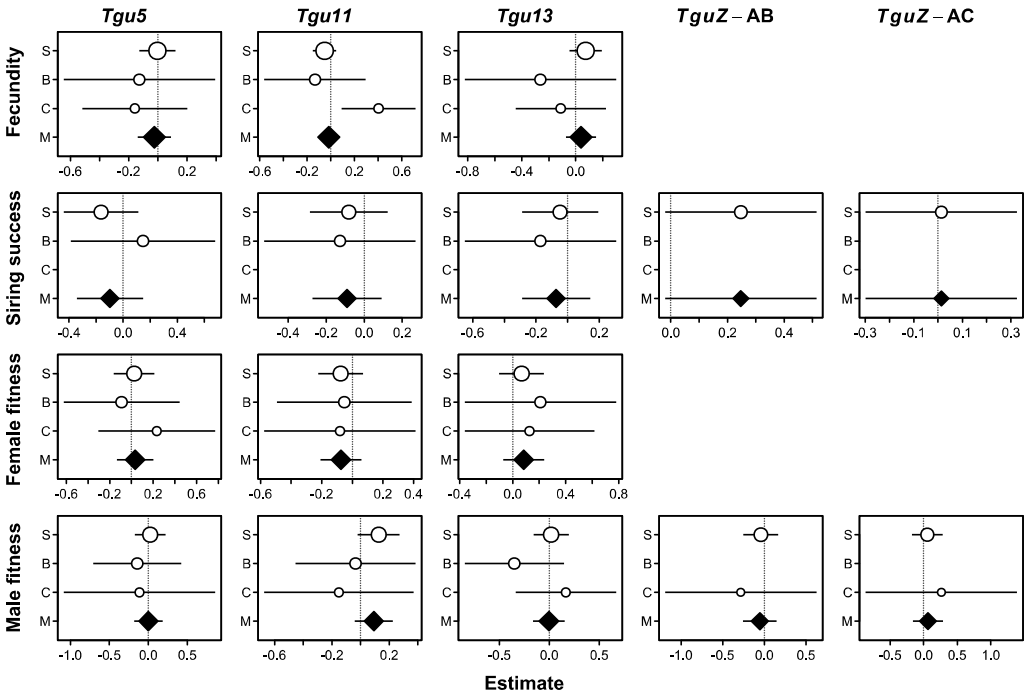


Figure 5: *Dominance effects ± 95% confidence intervals on different fitness parameters in three captive populations (S = "Seewiesen", B = "Bielefeld", C = "Cracow" and M = meta-analytic summary). Effect sizes are the factor level estimates of square-rooted and Z-transformed fitness components over inversion heterozygosity (while simultaneously fitting additive effects).The point sizes reflect log-transformed sample sizes.*

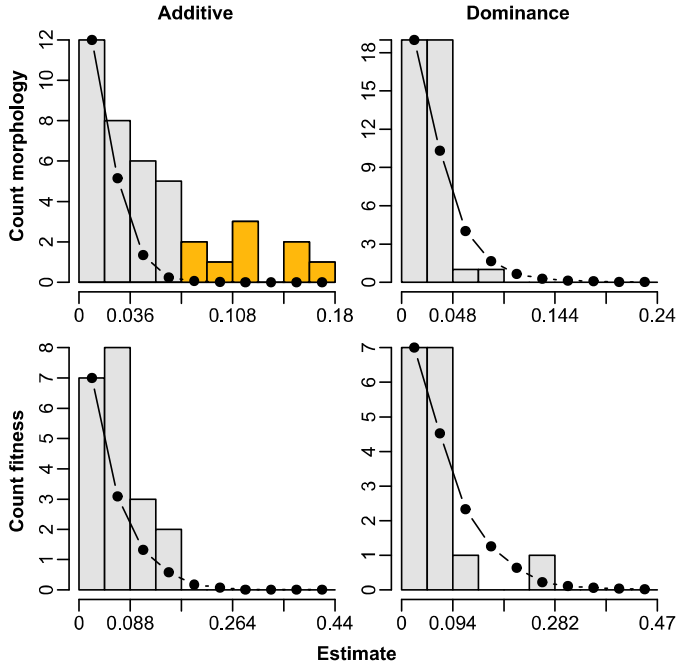


Figure 6: *Summary of additive (left column) and dominance (right column) effect sizes from association studies between inversion genotypes and morphological traits (40 estimates = 8 phenotypes x 5 inversions; top row) and of the additive and dominance effect sizes from associations between inversion genotypes and the fitness traits (20 [16] estimates = 4 fitness parameters x 5 inversions [minus 4 TguZ dominance effects in females]; bottom row). Empirical effect sizes are the light grey bars overlaid with the null distribution as a black line. Effects that survived strict Bonferroni correction are highlighted in yellow. For the null distribution we permuted the inversion genotypes within sexes 100 times and ran the same mixed models as for the empirical data set. The null distribution has been scaled to overlap the first bar in the histogram of the empirical estimates completely. Partial regression coefficients of additive and dominance effects are not directly comparable the way we standardized and fitted them and thus their null distributions deviate (dominance effects reach higher values than additive effects because their variance is smaller; compare left vs right column and see also Gelman 2008; Schielzeth 2010).*

Table 1: *Description of the four large linkage blocks (resulting from inversion polymorphisms) on chromosomes Tgu5, Tgu11, Tgu13 and TguZ and the two smaller ones on chromosomes Tgu26 and Tgu27. Inversions on chromosomes Tgu5 and TguZ had been previously found cytogenetically (Christidis 1986a; Itoh & Arnold 2005; Itoh et al. 2011). Centromere positions were taken from Chapter 5 and Warren et al. (2010). For each chromosome we list the first and the last SNP that is in linkage disequilibrium with the LD region (defined as composite LD $r^2 > 0.1$, and we indicate its maximal value). n SNPs is the number of SNPs genotyped and contributing to the LD region.*

Chromosome	Inversion type	SNP ID	Position (bp)	Maximal r^2	n SNPs
Tgu5	pericentric	WZF00178137	962,370	0.996	152
		WZF00169812	16,503,169	0.184	
Tgu11	paracentric	WZF00035574	86,193	0.187	38
		WZF00031807	12,290,125	0.985	
Tgu13	unknown	WZF00041237	150,262	0.904	163
		WZF00041448	16,906,706	0.130	
TguZ	pericentric	WZF00231767	5,913,912	0.285	383
		WZF00239958	68,830,532	0.261	
Tgu26	unknown	WZF00114713	657,240	0.101	16
		WZF00114507	2,710,851	0.244	
Tgu27	unknown	WZF00115125	358,632	0.133	23
		WZF00114985	3,097,302	0.553	

Table 2: *Population genetic descriptive statistics of the four inversion polymorphisms. Heterozygosity is the average heterozygosity of all individuals within the respective PCA score cluster. Hardy-Weinberg-equilibrium (HWE) was tested using a chi-square test with the indicated degrees of freedom.*

Chromosome	Genotype	Heterozygosity (95% CI)	Genotype counts		Allele frequency	HWE Test
			males	females		
Tgu5	AA	0.131 (0.0855, 0.178)	154	192	0.595	X^2_1=1.99, P=0.16
	AB	0.689 (0.644, 0.737)	232	204		
	BB	0.0523 (0.0263, 0.0855)	82	84	0.405	
Tgu11	AA	0.0790 (0.000, 0.158)	124	143	0.526	X^2_1=0.40, P=0.53
	AB	0.493 (0.368, 0.605)	245	218		
	BB	0.214 (0.105, 0.342)	99	119	0.474	
Tgu13	AA	0.180 (0.119, 0.240)	129	128	0.525	X^2_1=0.28, P=0.59
	AB	0.469 (0.411, 0.527)	243	238		
	BB	0.170 (0.117, 0.216)	96	114	0.475	
*TguZ**	AA / AW	0.162 (0.103, 0.230)	140	266	0.596	X^2_3=4.42, P=0.22
	AB	0.592 (0.521, 0.639)	174	0		
	BB / BW	0.0657 (0.0395, 0.0953)	36	155	0.33	
	AC	0.555 (0.496, 0.596)	38	0		
	BC	0.294 (0.265, 0.332)	19	0		
	CC / CW	0.108 (0.0868, 0.143)	4	29	0.074	
	untyped		57	30		

* heterozygosity data taken from males only

Table 3: *Sample sizes for the association analyses with embryo mortality and fitness parameters in the three captive populations.*

Population	**Parameter**	**Cage**		**Aviary**	
		Laying	**Breeding**	**Laying**	**Breeding**
Seewiesen	Embryo mortality (# eggs)	930	4121	765	518
	Egg volume (# eggs)	9838	7527	3894	653
	Female fecundity (# eggs)	10108	7539	3875	655
	Male siring success (# eggs)			3869	655
	Female fitness (# chicks)		1843		276
	Male fitness (# chicks)		1843		276
Bielefeld	Embryo mortality (# eggs)				1170
	Egg volume (# eggs)				1295
	Female fecundity (# eggs)				1295
	Male siring success (# eggs)				1295
	Female fitness (# chicks)				556
	Male fitness (# chicks)				556
Cracow	Embryo mortality (# eggs)	1674	586		
	Egg volume (# eggs)	133	773		
	Female fecundity (# eggs)	133	776		
	Male siring success (# eggs)				
	Female fitness (# chicks)		343		
	Male fitness (# chicks)		343		

Supplement

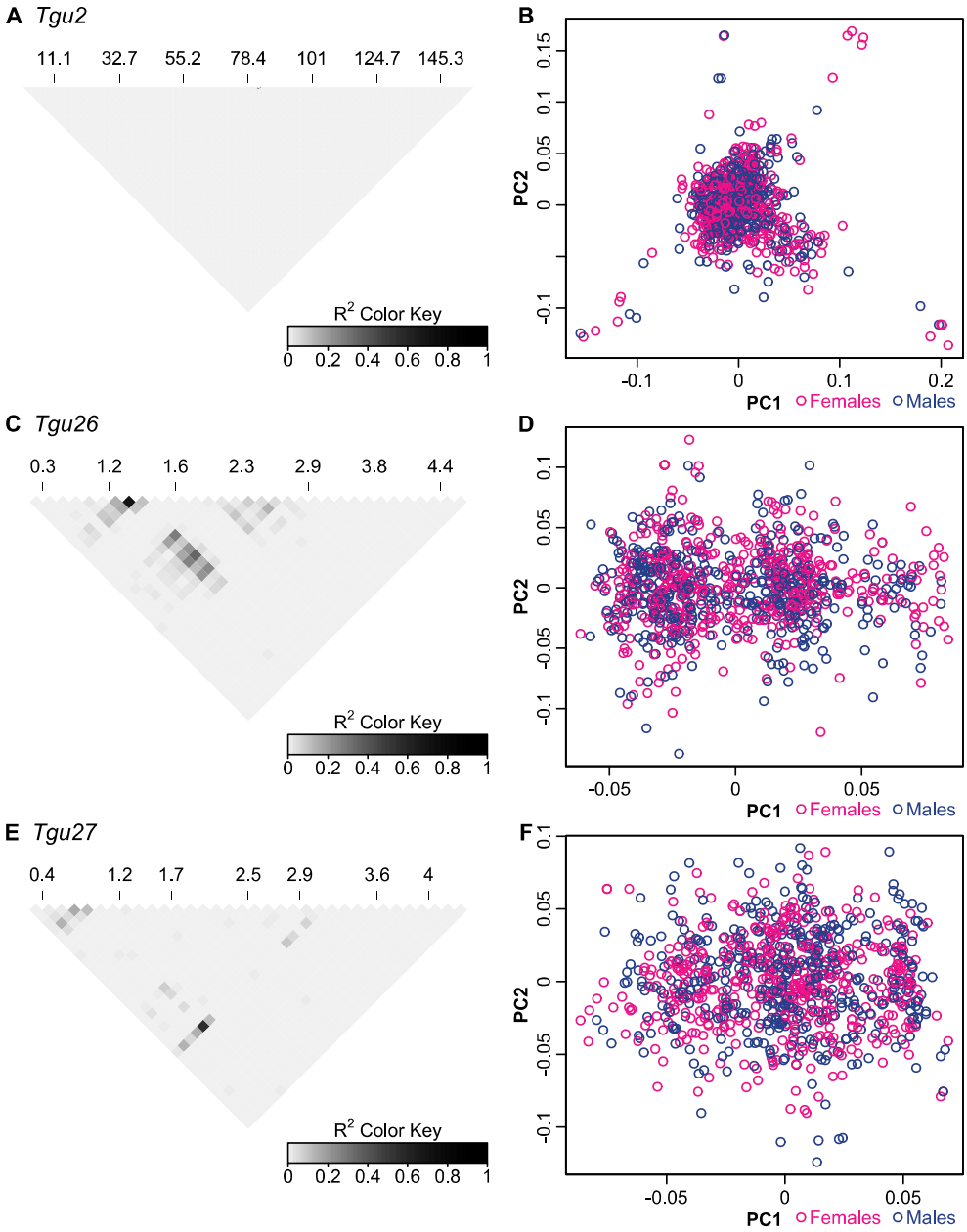


Figure S1: *The left panel depicts linkage disequilibrium (LD) and the right panel principal component analysis (PCA) results along chromosomes (A, B) Tgu2, (C, D) Tgu26 and (E, F) Tgu27. Above the LD plots marker positions in Mb are given. PCA included all SNPs on the respective chromosome. Chromosome Tgu2 is shown as an example lacking any linkage blocks (except a small region around the centromere around 82 Mb).*

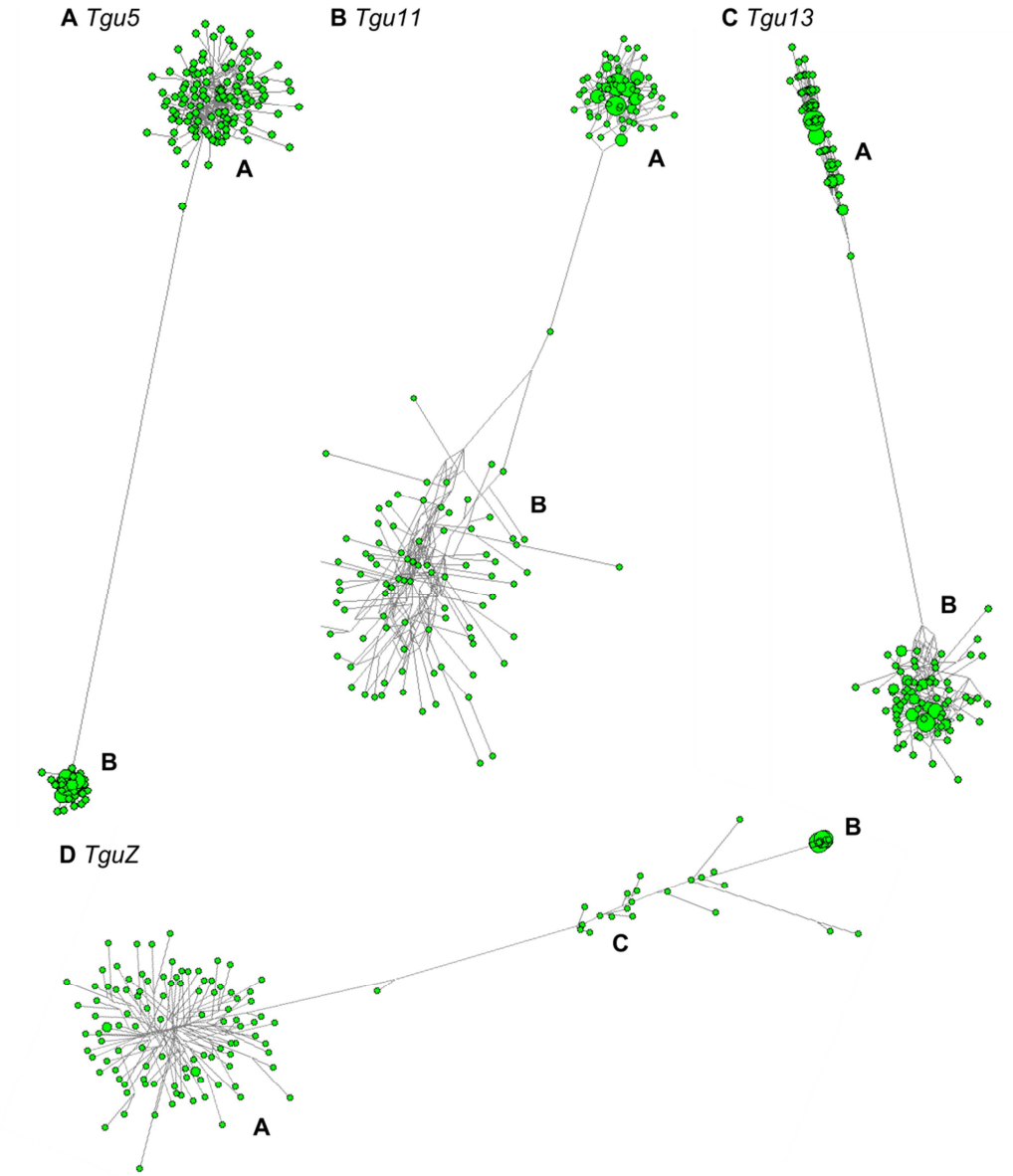


Figure S2: *Median-joining networks of (A) the phased SNPs in the presumed breakpoint regions of chromosome Tgu5, (B) all phased SNPs within the inverted region on chromosome Tgu11 (there were too few SNPs at the breakpoint), (C) the phased SNPs in the presumed breakpoint regions of chromosome Tgu13, and (D) all SNPs within the inverted region on chromosome TguZ in females (breakpoint regions are unclear on chromosome TguZ). For each chromosome 10% (n=188) of all haplotypes are shown and the haplotype is named as in Table 2.*

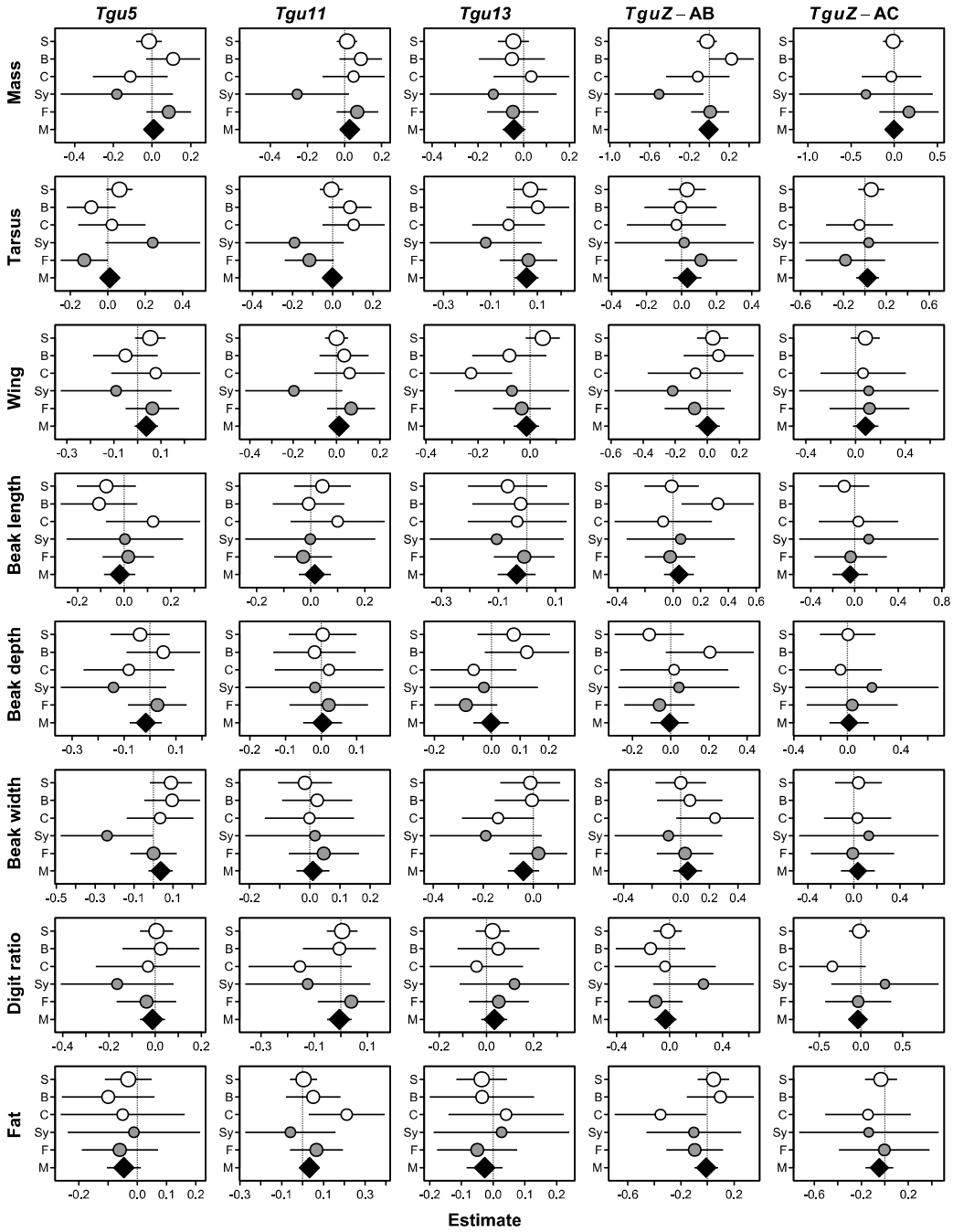


Figure S3: *Heterotic effects of the minor inversion allele ± 95% confidence intervals on morphological phenotypes in three captive and two wild populations (S = "Seewiesen", B = "Bielefeld", C = "Cracow", Sy = "Sydney", W = "Fowlers Gap" and M = meta-analytic summary). The point size reflects log-transformed sample sizes. None of the meta-analytic summary estimates survived strict Bonferroni correction.*

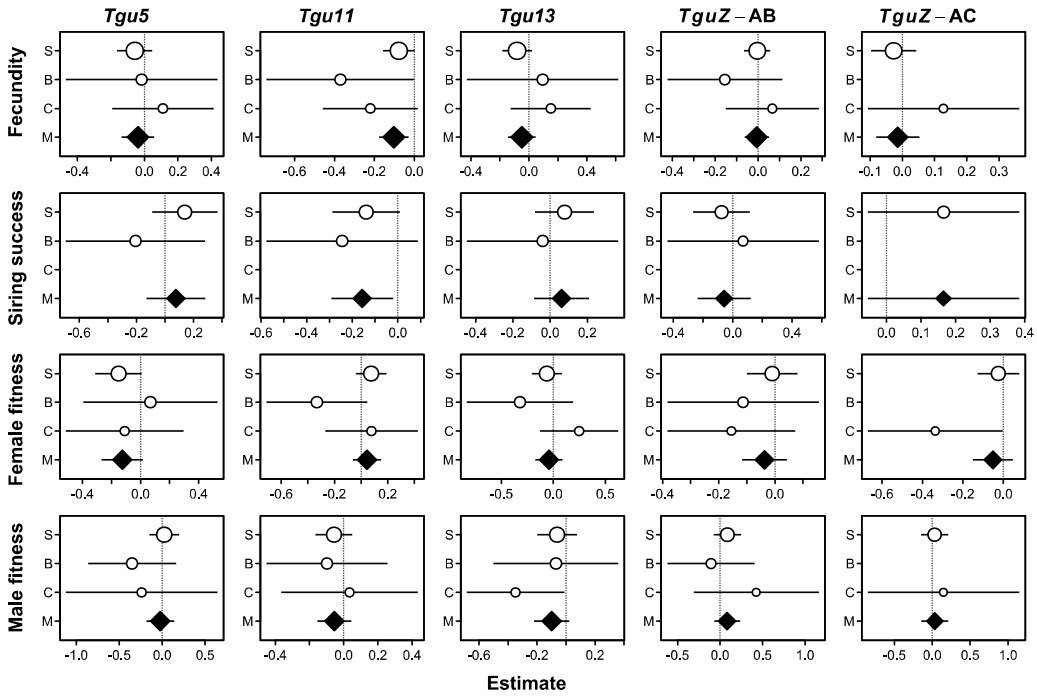


Figure S4: *Additive effects of the minor inversion allele ± 95% confidence intervals on different fitness parameters in the three captive populations (S = "Seewiesen", B = "Bielefeld", C = "Cracow" and M = meta-analytic summary). The point size reflects log-transformed sample sizes. None of the meta-analytic summary estimates survived strict Bonferroni correction*.

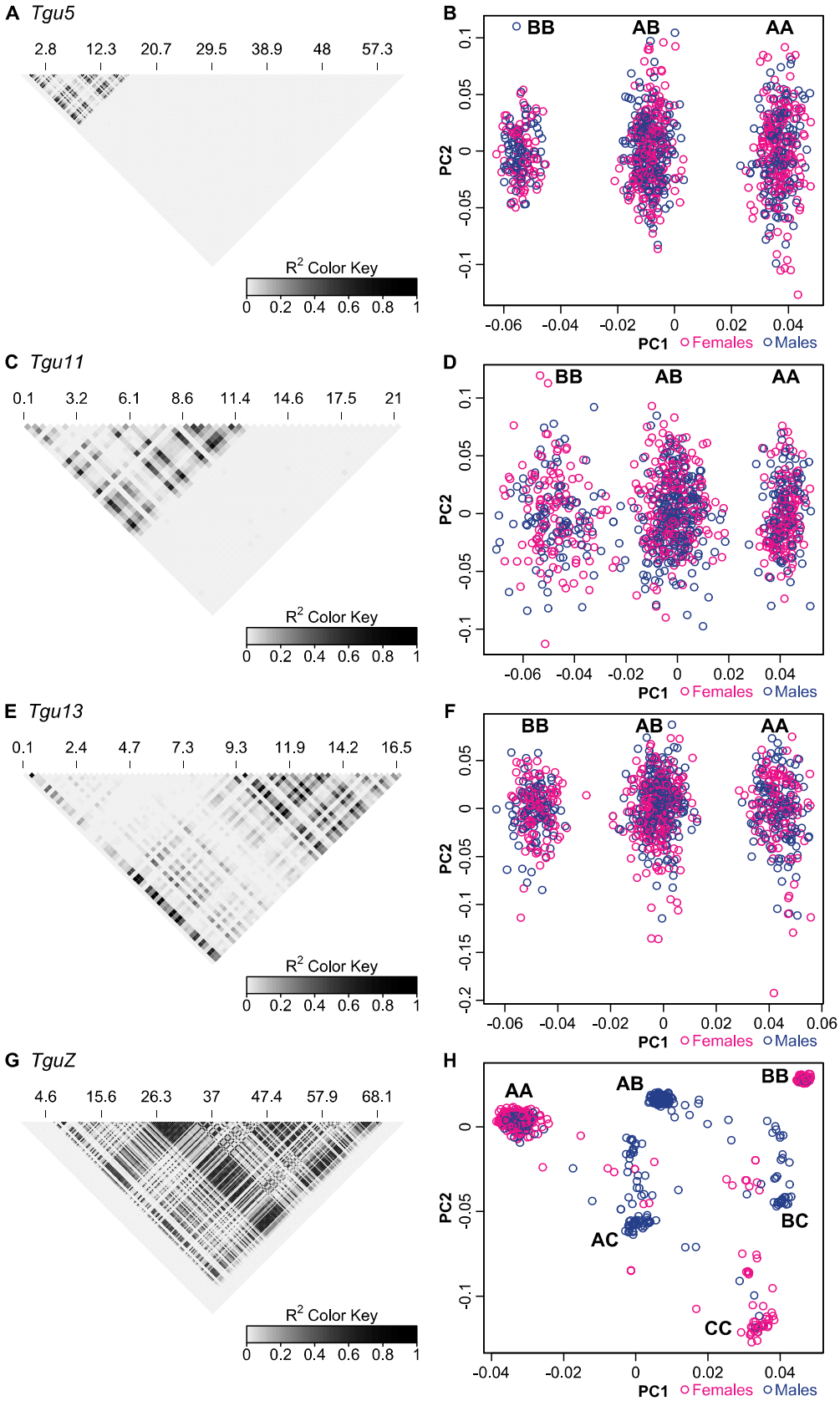


Figure S5: *The left panel depicts linkage disequilibrium (LD) and the right panel principal component analysis (PCA) results along chromosomes (A, B) Tgu5, (C, D) Tgu11, (E, F) Tgu13 and (G, H) TguZ. Above the LD plots marker positions in Mb are given. SNPs had been filtered*

prior to analyses using the "earliest finish time" greedy algorithm to include only those SNPs that were separated by minimally 185 kb.

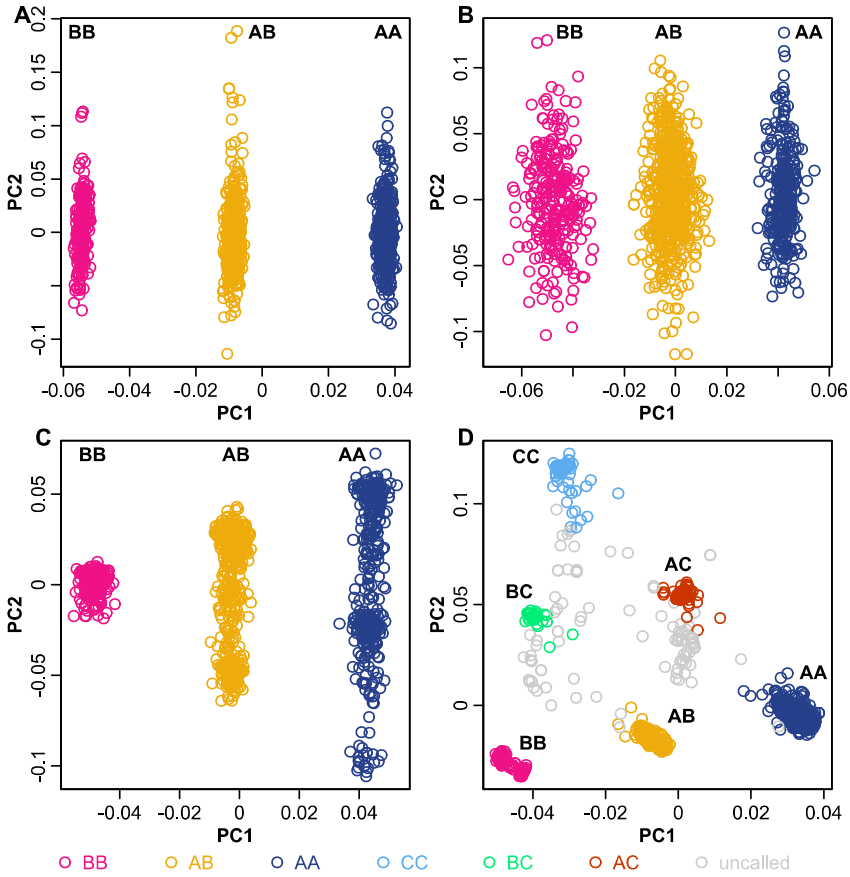


Figure S6: *Principle component analysis (PCA) results from the wild "Fowlers Gap" birds along chromosome (A) Tgu5, (B) Tgu11, (C) Tgu13, (D) TguZ. Founders of all four populations that produced offspring (n = 239 individuals) were run on both the Illumina and Sequenom genotyping platform and we used the SNP-loadings on PC1 and PC2 from the PCA of the "Fowlers Gap" birds on the population founders to calculate a PCA score for each individual and added them to the figure (see also Figure S6). Colors represent the inversion haplotype calling using only information from the tag SNPs.*

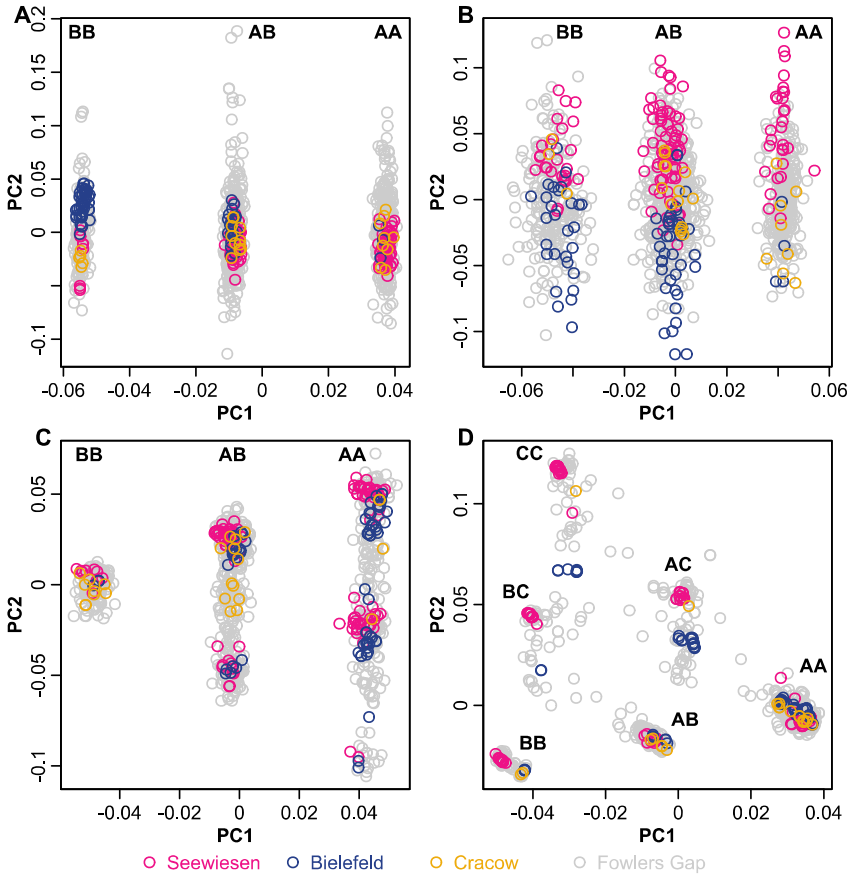


Figure S7: *Principle component analysis (PCA) results from the wild "Fowlers Gap" birds along chromosome (A) Tgu5, (B) Tgu11, (C) Tgu13, (D) TguZ. Founders of all four populations that produced offspring (n = 239 individuals) were run on both the Illumina and Sequenom genotyping platform and we used the SNP-loadings on PC1 and PC2 from the PCA of the "Fowlers Gap" birds on the population founders to calculate a PCA score for each individual and added them in different colors to the figure.*

Table S1: *Description of the 15 tag SNPs. r^2 is the composite LD with the inversion genotype. Type A, Type B and Type C indicate the number of A alleles in a homozygous individual for inversion genotypes A, B and C, respectively.*

Chromosome	SNP ID	Position	r^2	A allele	B allele	Type A	Type B	Type C
Tgu5	WZF00167975	1,000,220	1	A	C	2	0	
	WZF00170082	1,778,553	1	A	C	0	2	
	WZF00169329	14,526,432	1	T	G	2	0	
Tgu11	WZF00031778	12,252,712	0.972	A	G	2	0	
	WZF00031788	12,268,156	0.901	A	G	2	0	
	WZF00031805	12,289,339	0.994	T	C	2	0	
Tgu13	WZF00041460	171,215	1	A	G	0	2	
	WZF00040683	11,363,650	0.996	T	C	2	0	
	WZF00040731	11,536,148	0.998	T	C	2	0	
TguZ	WZF00231991	5,954,002	1	A	G	2	0	2
	WZF00218433	37,400,431	0.986	T	C	2	2	0
	WZF00218859	38,265,467	1	A	C	0	2	0
	WZF00222123	44,773,107	0.986	T	C	2	2	0
	WZF00223156	46,527,787	0.986	A	G	0	2	2
	WZF00237790	65,844,335	0.986	A	G	0	2	2

Table S2: *Comparison between Sequenom calling of inversion genotypes using the tag SNPs and Illumina calling of inversion types using PCA of all SNPs genotyped on the chromosome in n = 127 Seewiesen, n = 74 Bielefeld and n = 25 Cracow founder individuals. There were additional 10 Fowler's Gap individuals. 1,062 Seewiesen individuals had been previously genotyped with 37 SNPs on chromosome TguZ and we compared the calling between Sequenom and the PCA results using these 37 SNPs.*

Chromosome	n Errors	n Uncalled	n Comparisons	Assay	Calling algorithm
Tgu5	0	0	234	Illumina iSelect (same SNPs)	Majority vote
Tgu11	0	0	236	Illumina iSelect (same SNPs)	Majority vote
Tgu13	0	0	236	Illumina iSelect (same SNPs)	Majority vote
TguZ	0	2	200	Illumina iSelect (same SNPs)	Majority vote; in Bielefeld restrict calling to individuals where all SNPs are genotyped and fit, which removes 26 individuals ("recombinants")
TguZ	0	0	1062	Illumina GoldenGate (different SNPs)	Majority vote

General discussion

In my PhD thesis, I studied the effects of quantitative and molecular genetic variation on complex phenotypes in zebra finches. I used a combination of forward and reverse genetic approaches for whole-genome screening, combined with detailed confirmatory studies to lower the false-discovery rates. Genome-wide scans per se are powerful methods to generate hypotheses in an unbiased way, but testing these hypotheses is a necessary separate step towards a better understanding of the genetic architecture of complex traits and the evolution of genomic landscapes. **Chapter 4** exemplary illustrates the need for confirmatory sampling after performing genome scans: The initial finding of one locus showing a significant deviation from Mendelian transmission turned out to be a false-positive finding since it was not replicated in a more than twice as large confirmatory sample. In the following, I will highlight the most important results from the forward and reverse genetic approaches for the study of genetic architecture, build connections between the chapters and suggest further avenues of research.

Forward genetics approaches to study genetic architecture

Linkage and association mapping

In **Chapter 1** I used quantitative genetics and linkage mapping to gain insights into the genetic architecture of phenotypic and genotypic variation in beak morphology (beak length, depth and width). In the captive population I studied, genetic correlations between the beak dimensions were high (0.46–0.69), but there was also substantial conditional heritability for all three traits, meaning that they could potentially evolve independently of each other (Hansen & Houle 2008). QTL maps confirmed the quantitative genetic analyses: The genomic regions linked to variation in the three beak dimensions overlapped in large parts, which constituted the genetic correlations. However, each beak dimension had also at least one genomic region that was exclusively linked to it, and these regions contributed to the conditional heritability.

Although the linkage analyses in **Chapter 1** narrowed down the variants underlying phenotypic variation, QTL regions covered thousands of genes (see also Schielzeth *et al.* 2012), which necessitated high-resolution association mapping to potentially identify the actual causal variants. Thus, I tested SNPs within candidate genes located in the QTL regions for an association with their respective phenotypic traits, which is a commonly advocated research strategy for gene mapping (Cardon & Bell 2001). In order to achieve high resolution in the association mapping, I sampled birds from a wild Australian population in which linkage disequilibrium (LD) decays rapidly (Balakrishnan & Edwards 2009; **Chapter 6**). To confirm those associations found in the wild, I set out to replicate effects in several captive

populations in which LD is higher but expected to differ in phase between populations (**Chapter 6**). Most of the tested variants showed an association in one or more of the populations. Effects were replicable within populations but generally not consistent across populations, indicating that the tested SNPs were not by themselves causal but linked to causal variants.

In the forward genetics approach outlined above I focused on the standing additive genetic variation within several populations of wild and captive zebra finches. Previous studies in other (model) organisms generally found small effects of individual SNPs on quantitative traits (Flint & Mackay 2009), largely consistent with the infinitesimal model of quantitative genetic variation (Falconer & Mackay 1996; Robinson *et al.* 2014) and consistent with my findings in zebra finches. However, heritability and the percentage of phenotypic variance explained by individual variants are population-specific parameters and are shaped by population genetic and demographic processes (Lynch & Walsh 1998; Pardo-Diaz *et al.* 2015). Thus, loci with strong effects on phenotypes may segregate in populations and contribute to the standing genetic variation, as I have shown for captive populations with high genome-wide LD (**Chapters 1 and 6**) and for a wild outbred population with very low genome-wide LD (**Chapter 7**). After identifying the causal variants underlying phenotypic variation (**Chapters 6 and 7**) we can study their relative contribution to the standing genetic variation, their pleiotropic effects and gain insights into the forces maintaining them in a population (Lee *et al.* 2014; Slate *et al.* 2010). In this respect, the two SNPs identified in **Chapter 6** which had consistent pleiotropic effects on two beak dimensions could be causal variants contributing to the genetic correlations observed in **Chapter 1** between the three beak dimensions. They should be followed up by functional assays (such as expression analyses or *in situ* hybridization; Pardo-Diaz *et al.* 2015) in order to see whether they are differentially expressed in the beak forming tissues (which are the prenasal cartilage and the premaxillary bone; Mallarino *et al.* 2012) of long- and short-beaked birds. The inversion polymorphisms described in **Chapter 7** were extraordinary in terms of their effects on phenotypes and seemed to be balanced polymorphisms. Understanding the selective forces maintaining them at intermediate allele frequencies is an exciting avenue for further research (see below).

The reverse genetics approach to study genetic architecture

Genome scans for inbreeding

In **Chapter 2** I focused on identifying regions in the genome which are identical-by-decent (IBD). I did this by considering several closely linked SNPs jointly as a single, highly polymorphic marker (Kong *et al.* 2008). The probability of a large number of densely spaced SNPs (in my case I used 56–75 SNPs) being homozygous due to chance alone (which means identical-by-state, IBS) is diminishingly small and thus indicates homozygosity due to common ancestry (i.e. IBD; Broman & Weber 1999). Within those regions all recessive

deleterious mutations which occurred before the last common ancestor are homozygous, and thus fully express their deleterious effect, leading to inbreeding depression. These genomic scans for regions being IBD do not rely on pedigree information. I showed analytically that they may capture inbreeding loops that reach back in time much further than typical pedigree information (see also Hayes *et al.* 2003; **Chapter 2**).

In **Chapter 3** I used simulations to study how these marker-based estimates of IBD compare to estimates gained from pedigree information (the pedigree-based inbreeding coefficient F) in heterozygosity-fitness correlations. Specifically, I focused on three sources of measurement error affecting the inbreeding estimate of an individual: the Mendelian sampling noise, the marker sampling noise and the IBD-IBS discrepancy (see the Introduction to **Chapter 3** for details). Ordinary least squares regressions of a fitness-related trait over the pedigree-based estimate of inbreeding yielded unbiased slopes, whereas marker-based estimates of inbreeding gave slopes which were biased downwards, meaning that they underestimated the inbreeding load (Szulkin *et al.* 2010). However, a rather small number of markers was better in predicting individual inbreeding levels than the pedigree-based estimate.

Quantifying inbreeding and inbreeding depression using microsatellites or other molecular markers gains much attention in the current literature (discussed in Forstmeier *et al.* 2012; Kardos *et al.* 2015). However, a clear distinction between IBS and IBD is usually not drawn but may offer more precise estimates of individual inbreeding levels and an unbiased estimate of the inbreeding load in a population (**Chapter 3**). The method I described in **Chapter 2** creates highly informative markers for IBD detection and is a way to reduce the IBD-IBS discrepancy which is the fraction of markers being homozygous just by chance (**Chapter 3**). Precision and accuracy in heterozygosity-fitness correlations could either be increased by reducing the marker sampling noise by genotyping more markers at independent loci across the genome. Alternatively, the IBD-IBS discrepancy could be minimized by genotyping high-density genetic marker panels in small genomic regions. It should be tested empirically which of the two approaches leads to less biased results in heterozygosity-fitness correlation. The answer to this question depends partly on the magnitude of the IBD-IBS discrepancy but to my knowledge it has never been assessed how well IBS at a specific locus (such as a microsatellite) does reflect common ancestry in a wild population. Here, high density SNP panels could be used to address this question empirically in wild populations.

Genome scans for outlier regions

Deviations from Mendelian segregation

In **Chapter 4** I described an allele located on chromosome *Tgu2* which heterozygous individuals pass on to their offspring more often than expected by fair Mendelian transmission. The transmission bias was found in all developmental stages (embryos, chicks

and adult birds) and individuals heterozygous for the overtransmitted allele did not produce a higher fraction of (apparently) infertile eggs than homozygotes. Thus, I concluded that the transmission distorter was acting prezygotically.

These types of selfish genetic elements have been described in several organisms (Burt & Trivers 2006), but usually preferential transmission is restricted to one sex since female and male meiosis are fundamentally different (Lyttle 1993). From a single progenitor cell, females produce only a single egg cell and three polar bodies (an evolutionary dead-end), but males produce four sperm cells. Thus, an effective segregation distorter in females should have a higher probability to end up in the egg cell, whereas in males the sperm cells carrying the overtransmitted allele should disrupt or otherwise outperform those sperm not carrying the driving allele (Pardo-Manuel de Villena & Sapienza 2001). Surprisingly, however, both heterozygous females and males passed on the allele I described in **Chapter 4** more frequently than expected under Mendelian segregation, making it more likely that the bias occurs at an early stage of meiosis in which both sexes are still similar, rather than after gamete formation. Studying female meiosis is difficult but in males heterozygous for the driving allele one could examine testes for increased (selective) apoptosis using the TUNEL assay or sequence ejaculates with high coverage (around 350 x for 80% power at a transmission ratio of 0.567) to check whether the transmission bias is already established at such an early stage.

Transmission distorters typically consist of multiple linked genes that do not recombine with the undertransmitted haplotype (Burt & Trivers 2006). This may explain the large number of segregation distorters on the non-recombining sex chromosomes (Frank 1991; Hurst & Pomiankowski 1991), but this finding could also be due to an ascertainment bias, since deviations from an equal sex ratio are readily observable (Burt & Trivers 2006). On the autosomes, transmission distorters are usually protected from recombination by a single or multiple chromosomal inversions (for example the *t*-haplotype in mice [*Mus musculus*; Lyon 2003] or the *SD* locus in *Drosophila melanogaster*; Presgraves *et al.* 2009). The zebra finch has been shown to harbor several inversion polymorphisms (Christidis 1986a; Itoh & Arnold 2005; Itoh *et al.* 2011; **Chapter 7**) and also the transmission distortion locus on chromosome *Tgu2* seems to be located in a structural variant which appears to be absent from the wild population I studied in **Chapter 7**. A principle component analysis using 59 SNPs in 1,057 captive birds (for a rational of this approach see Ma & Amos 2012 and **Chapter 7**) suggests that the undertransmitted allele is part of a defined rare haplotype (Figure 1). However, to gain a better understanding of the molecular organization of this structural variant, cytogenetic (for example fluorescence *in situ* hybridization on mitotic or meiotic chromosomes) and long-range sequencing methods covering several kilobases are needed.

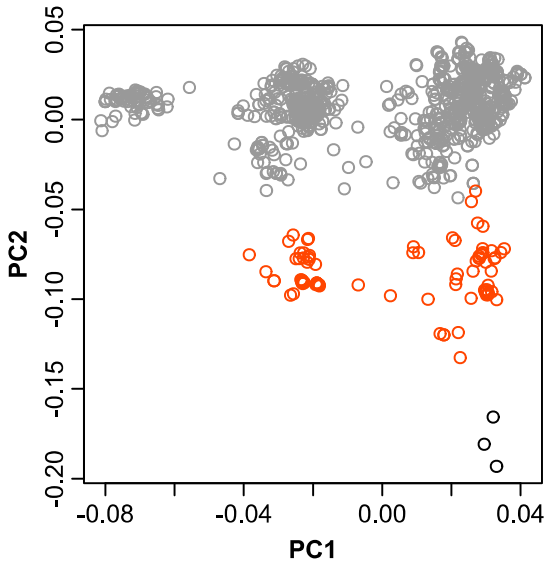


Figure 1: *Principal component analysis results using 1,057 individuals and 59 SNPs on chromosome Tgu2 (0.1–44.9 Mb). Grey points are individuals homozygous for the overtransmitted allele, red points are individuals who are heterozygous for the driving allele, black points are individuals homozygous for the undertransmitted allele.*

Deviations in diversity

It has been recognized that differentiation between species using relative measures of divergence (such as F_{ST}) often showed pronounced peaks at centromeres and telomeres (Cruickshank & Hahn 2014), which has been interpreted either as signs of adaptive divergence ("islands of speciation"; Turner *et al.* 2005), transmission distortion due to meiotic drive (Henikoff & Malik 2002) or background selection in the absence of recombination (Charlesworth *et al.* 1993). Since all three selective processes lead to a reduction in diversity near centromeres (Cruickshank & Hahn 2014), their positions should be mapped in order to distinguish between the three selective forces.

I conducted a genome scan for diversity in the zebra finch genome which showed markedly reduced diversity levels at both ends of almost all microchromosomes and at the known centromeres of the macrochromosomes (**General Introduction**). From cytogenetic studies it is known that all zebra finch microchromosomes are acrocentric (Pigozzi 2008), meaning that the centromeres are located close to the end of a chromosome. In **Chapter 5** I used triploidy mapping to locate the centromeres on almost all chromosomes in the annotated zebra finch genome. Bird genomes are highly syntenic (Romanov *et al.* 2014) and the zebra finch genome is often used as a reference assembly for other songbirds (for example Delmore *et al.* 2015). Since genome scans for diversity and differentiation are getting increasingly popular in birds (Delmore *et al.* 2015; Ellegren *et al.* 2012; Poelstra *et al.* 2014), and since centromeres strongly influence diversity levels (Cruickshank & Hahn 2014), this is a valuable resource for further studies not limited to zebra finches.

Deviations in heterozygosity

In **Chapter 7** I combined the results of a genome scan for heterozygosity in a wild population of zebra finches with the power of individual SNP genotyping in multiple captive populations (**Chapter 6**) to screen for and understand genomic outlier regions in terms of heterozygosity and linkage disequilibrium. Genome scans for heterozygosity are usually conducted to identify selective sweeps which are regions of low diversity and heterozygosity due to positive selection acting in the past (Qanbari *et al.* 2012; Rubin *et al.* 2012; Rubin *et al.* 2010). Instead, I focused on regions of high heterozygosity which may be under balancing selection (Fijarczyk & Babik 2015), allowing to measure selection in contemporary populations.

On four different chromosomes I identified large genomic regions of high LD which showed peaks in heterozygosity at their boundaries. Two of these LD regions had been identified as polymorphic pericentric inversions on macrochromosomes *Tgu5* and *TguZ* (Christidis 1986a; Itoh & Arnold 2005; Itoh *et al.* 2011) and from patterns of LD I concluded that the other two regions (on microchromosomes *Tgu11* and *Tgu13*) were also inversion polymorphisms. The heterozygosity at the inversion breakpoints coincided with their allele frequencies in the wild population. All of them had low and remarkably similar major allele frequencies (range of the major allele 0.53–0.60), indicating some form of balancing selection rather than drift or a transitional stage (replacement of one allele by another due to positive selection; Falconer & Mackay 1996). However, the polymorphisms have not been maintained at their frequency for a very long time (i.e. more than ten times the effective population size in generations), because diversity levels at the breakpoints were slightly decreased rather than increased, which would be expected under long-term balancing selection (Andolfatto *et al.* 2001).

The two largest inversions (relative to their total physical chromosome size) increased embryo mortality rates each by 4.5% in heterokaryotypic males but not in females. All inversions were at an allele frequency close to 0.5 at which the maximal number of individuals is heterozygous and thus the population-level costs in terms of embryo mortality are highest. This strongly argues against drift having taken the inversions to their current allele frequencies. Thus, I tested for heterotic fitness effects in the wild and in captivity: There was no heterosis for viability in the wild since all inversion polymorphisms were in Hardy-Weinberg equilibrium, and no heterosis on several other fitness traits (fecundity, siring success, chicks fledged) in captivity.

Hence, the selective forces maintaining the inversions in the polymorphic state are unclear: In general, there could still be heterotic superiority in the wild if the selective forces acted only sporadically (like extreme drought or heat) and not within the lifetime of those birds studied here, leaving no signs of departure from Hardy-Weinberg equilibrium (reviewed in Waples 2015). Alternatively, frequency dependent selection or selection in heterogeneous

environments which favor different alleles (annidation; Ludwig 1950) do not require heterokaryotypic superiority and do not necessarily lead to deviations from Hardy-Weinberg equilibrium (White 1977). An alternative to selection at the organismal level (i.e. heterosis, frequency dependent selection, and annidation) could be selection at the gene level (leading to segregation distortion; **Chapter 4**) which could be balanced by reduced individual fitness (Traulsen & Reed 2012). Single and also epistatically interacting inversions have been shown to bias transmission ratios in a diverse range of organisms including birds (reviewed in Burt & Trivers 2006; Nankivell 1967; Thorneycroft 1975) and most likely also zebra finches, as I have described in **Chapter 4**. However, I did not observe significant transmission distortion on chromosomes *Tgu11*, *Tgu13* or *TguZ* (**Chapter 4**). Further, although the only locus deviating significantly from Mendelian segregation in the combined-sexes genome-wide scan was located on chromosome *Tgu5* (**Chapter 4**), this SNP was unlinked to the inversion polymorphism (composite LD $r^2 = 0.0027$). Thus, segregation distortion is unlikely the current driving force keeping the inversions in the population. However, it may have played a role in establishing them in the population and is now suppressed by the rest of the genome (cryptic drive; Hartl 1975), which has been frequently observed in several *Drosophila* species (for example Tao *et al.* 2001). In order to find out whether the inversion polymorphisms in the zebra finch are indeed such cryptic drive systems, they need to be introgressed in a naïve genetic background of a separated population or a different species, in which the suppressors of the segregation distorter did not evolve.

Besides the selective forces keeping the inversions polymorphic in the wild, two additional aspects should be emphasized here:

(1) As theory predicts (Swanson *et al.* 1981), and as it has been observed in humans (Anton *et al.* 2005), the inversions increase embryo mortality rates in heterokaryotypic individuals but only if they span more than half of the chromosome and only in males. The effect, however, is twelve times weaker than in humans, raising the question how this is achieved mechanistically. Average recombination rates in the zebra finches are slightly higher than in humans (1.98 cM/Mb vs 1.30 cM/Mb, respectively; **Chapter 3**) but these estimates are misleading because the zebra finch has almost double the number of chromosomes. Assuming one obligate cross-over per chromosome per meiosis (Petronczki *et al.* 2003) and using the autosomal linkage maps from Calderón & Pigozzi (2006) for zebra finches (2,272.5 cM) and Matise *et al.* (2007) for humans (3,595 cM), zebra finches show on average only 1.17 (= 2,272.5 cM / 39 chromosomes / 50 cM) but humans 3.27 (= 3,595 cM / 22 chromosomes / 50 cM) recombination events per chromosome per meiosis. Furthermore, recombination events in zebra finches are highly skewed towards the chromosome ends (Backström *et al.* 2010). It also appears as if the large and costly inversions on chromosomes *Tgu13* and *TguZ* accumulated more and more inversions, which could potentially inhibit homologous pairing and thus detrimental cross-overs in heterokaryotypic individuals. In combination, these factors may explain the reduced embryo mortality rate in zebra finches

but causation remains unclear: did the low and skewed recombination landscape allow the spread of inversions because costs were rather low (Cáceres *et al.* 1997) or did the emergence of multiple inversion polymorphisms select for a low rate and skewed distribution of recombination events? An answer to this question may be gained by studying species closely related to the zebra finch and there is some evidence that also in related species recombination is skewed towards the telomeres (Singhal *et al.* 2015), which could explain the preponderance of polymorphic inversions in the whole family of grassfinches (Christidis 1986a, b, 1987; Hooper & Price 2015).

(2) Each inversion linked hundreds of genes (and their allelic states in coupling phase) in one defined haplotype, creating kind of a "supergene", which had large additive effects on morphological phenotypes but exhibited no dominant gene action. In order to detect overdominance (heterotic effects) of inversion polymorphisms on phenotypes (**Chapter 7**), I estimated additive and dominance effects simultaneously within the same statistical model, which allows to distinguish between an additive and a dominant mode of gene action (Vitezica *et al.* 2013). In standard quantitative genetic analyses, an allelic effect is estimated as the slope of a least squares regression of the genotypic value on gene content, which is called the average effect of allelic substitution, and which is directly related to the additive genetic variance (Lynch & Walsh 1998). However, the slope estimate does not distinguish between an additive and a dominant mode of gene action which I was interested in. Also, although heritability estimates of the phenotypes studied in **Chapter 7** were high, it does not necessarily mean that there is little dominant gene action at individual loci (Hill *et al.* 2008; Nietlisbach & Hadfield 2015), which is especially true for loci with allele frequencies close to 0.5 (Lynch & Walsh 1998). In any case, the absence of a dominant gene action of the inversion polymorphisms on morphology further points to no heterotic effects.

Conclusion

The combination of unbiased genome-wide approaches and detailed confirmatory sampling is a fruitful avenue for research in evolutionary genetics. The first step should be thought of as the hypothesis-generating stage of a study, making the second step essential for hypothesis testing (Province 2000; Thomas *et al.* 1985), but it is all too often overlooked in the current literature. Its importance cannot be overemphasized in genomic studies, since they suffer an immense multiple-testing burden and particularly since both forward and reverse genetic approaches focus on genome-wide outlier regions: Forward genetics on the more subtle outliers in terms of effect sizes on phenotypic traits and reverse genetics on outliers from the genome-wide average in terms of some population genetic measure (for example diversity or heterozygosity), making both approaches prone to overestimate effect sizes (Open Science Collaboration 2015) and the regression towards the mean phenomenon (Galton 1886).

I used forward and reverse genetics alone or in combination to localize regions in the genome which are associated with quantitative traits and to actually visualize genetic correlations (**Chapter 1**), to quantify the amount of inbreeding over many generations without the need for pedigrees (**Chapters 2 and 3**), to identify regions in the genome that are subject to selection (also below the organismal level, **Chapters 4 and 7**), to map centromere positions (**Chapter 5**), to quantify the mode of gene action and effects of individual SNPs on quantitative traits (**Chapter 6**), and to study large-scale structural variants which may follow different evolutionary trajectories than individual SNPs (**Chapters 7**).

These are exciting times for evolutionary genetics, in which technological advances open up new lines of research and allow making discoveries of "things we didn't know we didn't know" (Wray 2013). In this respect, any genome scan may be informative, and more so for non-trivial genetic phenomena like segregation distortion (**Chapter 4**) or large-scale structural variants (**Chapter 7**), which are species- or population-specific and cannot be studied in established model organisms. In all this excitement, however, we should not forget the necessary statistical rigor and that replication is the gold standard leading to scientific progress.

"There are more things in heaven and earth, Horatio, than are dreamt of in your philosophy."
— Shakespeare, Hamlet, Act 1, Scene V

References

Andolfatto P, Depaulis F, Navarro A (2001) Inversion polymorphisms and nucleotide variability in *Drosophila*. *Genetical Research* **77**, 1–8.

Anton E, Blanco J, Egozcue J, Vidal F (2005) Sperm studies in heterozygote inversion carriers: a review. *Cytogenetic and Genome Research* **111**, 297–304.

Backström N, Forstmeier W, Schielzeth H*, et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research* **20**, 485–495.

Balakrishnan CN, Edwards SV (2009) Nucleotide variation, linkage disequilibrium and founder-facilitated speciation in wild populations of the zebra finch (*Taeniopygia guttata*). *Genetics* **181**, 645–660.

Broman KW, Weber JL (1999) Long homozygous chromosomal segments in reference families from the centre d'étude du polymorphisme humain. *American Journal of Human Genetics* **65**, 1493–1500.

Burt A, Trivers R (2006) *Genes in conflict: the biology of selfish genetic elements*. Belknap Press, Harvard University Press, Cambridge, Massachusetts, London, England.

Cáceres M, Barbadilla A, Ruiz A (1997) Inversion length and breakpoint distribution in the *Drosophila buzzatii* species complex: is inversion length a selected trait? *Evolution* **51**, 1149–1155.

Calderón PL, Pigozzi MI (2006) MLH1-focus mapping in birds shows equal recombination between sexes and diversity of crossover patterns. *Chromosome Research* **14**, 605–612.

Cardon LR, Bell JI (2001) Association study designs for complex diseases. *Nature Reviews Genetics* **2**, 91–99.

Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303.

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.

Christidis L (1987) Chromosomal evolution within the family Estrildidae (Aves) III. The Estrildae (waxbill finches). *Genetica* **72**, 93–100.

Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology* **23**, 3133–3157.

Delmore KE, Hübner S, Kane NC*, et al.* (2015) Genomic analysis of a migratory divide reveals candidate genes for migration and implicates selective sweeps in generating islands of differentiation. *Molecular Ecology* **24**, 1873–1888.

Ellegren H, Smeds L, Burri R*, et al.* (2012) The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* **491**, 756–760.

Falconer D, Mackay T (1996) *Introduction to quantitative genetics*, 4th edn. Longmann, Harlow, UK.

Fijarczyk A, Babik W (2015) Detecting balancing selection in genomes: limits and prospects. *Molecular Ecology* **24**, 3529–3545.

Flint J, Mackay TFC (2009) Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Research* **19**, 723–733.

Forstmeier W, Schielzeth H, Mueller JC, Ellegren H, Kempenaers B (2012) Heterozygosity-fitness correlations in zebra finches: microsatellite markers can be better than their reputation. *Molecular Ecology* **21**, 3237–3249.

Frank SA (1991) Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. *Evolution* **45**, 262–267.

Galton F (1886) Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute of Great Britain and Ireland*, 246–263.

Hansen TF, Houle D (2008) Measuring and comparing evolvability and constraint in multivariate characters. *Journal of Evolutionary Biology* **21**, 1201–1219.

Hartl DL (1975) Modifier theory and meiotic drive. *Theoretical Population Biology* **7**, 168–174.

Hayes BJ, Visscher PM, McPartlan HC, Goddard ME (2003) Novel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Research* **13**, 635–643.

Henikoff S, Malik HS (2002) Centromeres: selfish drivers. *Nature* **417**, 227.

Hill WG, Goddard ME, Visscher PM (2008) Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics* **4**, e1000008.

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

Hurst LD, Pomiankowski A (1991) Causes of sex ratio bias may account for unisexual sterility in hybrids: a new explanation of Haldane's rule and related phenomena. *Genetics* **128**, 841–858.

Itoh Y, Arnold AP (2005) Chromosomal polymorphism and comparative painting analysis in the zebra finch. *Chromosome Research* **13**, 47–56.

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

Kardos M, Luikart G, Allendorf FW (2015) Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity* **115**, 63–72.

Kong A, Masson G, Frigge ML*, et al.* (2008) Detection of sharing by descent, long-range phasing and haplotype imputation. *Nature Genetics* **40**, 1068–1075.

Lee YW, Gould BA, Stinchcombe JR (2014) Identifying the genes underlying quantitative traits: a rationale for the QTN programme. *Aob Plants* **6**, plu004.

Ludwig W (1950) Zur Theorie der Konkurrenz: die Annidation (Einnischung) als fünfter Evolutionsfaktor. *Neue Ergebisse Probleme Zool* **145**, 516–537.

Lynch M, Walsh B (1998) *Genetics and analysis of quantitative traits*. Sinauer, Sunderland, MA.

Lyon MF (2003) Transmission ratio distortion in mice. *Annual Review of Genetics* **37**, 393–408.

Lyttle TW (1993) Cheaters sometimes prosper: distortion of mendelian segregation by meiotic drive. *Trends in Genetics* **9**, 205–210.

Ma JZ, Amos CI (2012) Investigation of inversion polymorphisms in the human genome using principal components analysis. *PLoS One* **7**, e40224.

Mallarino R, Campàs O, Fritz JA*, et al.* (2012) Closely related bird species demonstrate flexibility between beak morphology and underlying developmental programs. *Proceedings of the National Academy of Sciences of the USA* **109**, 16222–16227.

Matise TC, Chen F, Chen WW*, et al.* (2007) A second-generation combined linkage–physical map of the human genome. *Genome Research* **17**, 1783–1786.

Nankivell RN (1967) A terminal association of two pericentric inversions in first metaphase cells of the Australian grasshopper *Austroicetes interioris* (*Acrididae*). *Chromosoma* **22**, 42–68.

Nietlisbach P, Hadfield JD (2015) Heritability of heterozygosity offers a new way of understanding why dominant gene action contributes to additive genetic variance. *Evolution* **69**, 1948–1952.

Open Science Collaboration (2015) Estimating the reproducibility of psychological science. *Science* **349**, aac4716.

Pardo-Diaz C, Salazar C, Jiggins CD (2015) Towards the identification of the loci of adaptive evolution. *Methods in Ecology and Evolution* **6**, 445–464.

Pardo-Manuel de Villena F, Sapienza C (2001) Nonrandom segregation during meiosis: the unfairness of females. *Mammalian Genome* **12**, 331–339.

Petronczki M, Siomos MF, Nasmyth K (2003) Un ménage à quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112**, 423–440.

Pigozzi MI (2008) Relationship between physical and genetic distances along the zebra finch Z chromosome. *Chromosome Research* **16**, 839–849.

Poelstra JW, Vijay N, Bossu CM*, et al.* (2014) The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **344**, 1410–1414.

Presgraves DC, Gerard PR, Cherukuri A, Lyttle TW (2009) Large-scale selective sweep among *Segregation Distorter* chromosomes in African populations of *Drosophila melanogaster*. *PLoS Genetics* **5**, e1000463.

Province MA (2000) A single, sequential, genome-wide test to identify simultaneously all promising areas in a linkage scan. *Genetic Epidemiology* **19**, 301–322.

Qanbari S, Strom TM, Haberer G*, et al.* (2012) A high resolution genome-wide scan for significant selective sweeps: an application to pooled sequence data in laying chickens. *PLoS One* **7**, e49525.

Robinson MR, Wray NR, Visscher PM (2014) Explaining additional genetic variation in complex traits. *Trends in Genetics* **30**, 124–132.

Romanov MN, Farré M, Lithgow PE*, et al.* (2014) Reconstruction of gross avian genome structure, organization and evolution suggests that the chicken lineage most closely resembles the dinosaur avian ancestor. *BMC Genomics* **15**, Artn 1060.

Rubin CJ, Megens HJ, Barrio AM*, et al.* (2012) Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences of the USA* **109**, 19529–19536.

Rubin CJ, Zody MC, Eriksson J*, et al.* (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* **464**, 587–593.

Schielzeth H, Forstmeier W, Kempenaers B, Ellegren H (2012) QTL linkage mapping of wing length in zebra finch using genome-wide single nucleotide polymorphisms markers. *Molecular Ecology* **21**, 329–339.

Singhal S, Leffler E, Sannareddy K*, et al.* (2015) Stable recombination hotspots in birds. bioRxiv doi: http://dx.doi.org/10.1101/023101.

Slate J, Santure AW, Feulner PGD*, et al.* (2010) Genome mapping in intensively studied wild vertebrate populations. *Trends in Genetics* **26**, 275–284.

Swanson CP, Merz T, Young WJ (1981) *Cytogenetics: the chromosome in division, inheritance, and evolution*. Prentice-Hall, Englewood Cliffs, NJ, USA.

Szulkin M, Bierne N, David P (2010) Heterozygosity-fitness correlations: a time for reappraisal. *Evolution* **64**, 1202–1217.

Tao Y, Hartl DL, Laurie CC (2001) Sex-ratio segregation distortion associated with reproductive isolation in *Drosophila*. *Proceedings of the National Academy of Sciences of the USA* **98**, 13183–13188.

Thomas DC, Siemiatycki J, Dewar R*, et al.* (1985) The problem of multiple inference in studies designed to generate hypotheses. *American Journal of Epidemiology* **122**, 1080–1095.

Thorneycroft HB (1975) Cytogenetic study of white-throated sparrow, *Zonotrichia albicollis* (Gmelin). *Evolution* **29**, 611–621.

Traulsen A, Reed FA (2012) From genes to games: cooperation and cyclic dominance in meiotic drive. *Journal of Theoretical Biology* **299**, 120–125.

Turner TL, Hahn MW, Nuzhdin SV (2005) Genomic islands of speciation in *Anopheles gambiae*. *PLoS Biol* **3**, e285, 1572–1578.

Vitezica ZG, Varona L, Legarra A (2013) On the additive and dominant variance and covariance of individuals within the genomic selection scope. *Genetics* **195**, 1223–1230.

Waples RS (2015) Testing for Hardy-Weinberg proportions: have we lost the plot? *Journal of Heredity* **106**, 1–19.

White MJD (1977) *Animal cytology and evolution*. Cambridge University Press, UK, Cambridge.

Wray GA (2013) Genomics and the evolution of phenotypic traits. *Annual Review of Ecology, Evolution, and Systematics* **44**, 51–72.

Summary

Evolutionary genetics has been transformed from a statistical description of genetic variation in terms of variances and covariances to a data-rich discipline which makes use of extensive genotypic data. However, the overarching goal did not change, namely to understand how genotypic variation contributes to complex phenotypic traits. In the absence of any genotypic data, quantitative genetics provides a statistical framework to separate phenotypic variation into its genetic and environmental component and to further partition the genetic component into an additive and a non-additive effect. For this to work, some implicit and simplifying assumptions need to be accepted, such as the infinitesimal model of additive genetic effects. With the advent of large-scale genotyping possibilities for non-model organisms we can now combine the large body of quantitative genetic theory with extensive genotypic data to identify the genetic variants contributing to complex phenotypes (forward genetics). However, we can also screen genomes for outliers from the assumptions made in quantitative genetics, try to identify the phenotypes influenced by these outliers and finally learn something about the selective forces shaping genotypic and phenotypic variation (reverse genetics). Both the forward and reverse genetics approach benefit from the current increase in genotyping possibilities, in such a way that they can now be used in an unbiased genome-wide manner with no prior knowledge on candidate genes.

My thesis combines quantitative genetics with molecular genetic approaches to study the genetic architecture of phenotypic traits in zebra finches (*Taeniopygia guttata*). I made use of both forward and reverse genetics to narrow down on causal genetic variants influencing morphology (**Chapters 1, 6 and 7**). Specifically, in **Chapters 1 and 6** I first estimated heritabilities and genetic correlations between beak length, depth and width in a captive population, then performed linkage mapping in this population, and finally mapped causal variants contributing to the identified linkage peaks in a wild and several captive populations of zebra finches. This part of my thesis highlights the power of linkage mapping to identify the regions associated with complex phenotypes and to visualize genetic correlations. Yet it also demonstrates the difficulties faced when narrowing further down to the causal variants contributing to phenotypic variation. In **Chapter 7** I made use of the reverse genetics approach and first identified four outlier regions in the zebra finch genome in terms of heterozygosity and linkage disequilibrium which turned out to be inversion polymorphisms segregating at intermediate allele frequencies in the wild. Then I used individual genotyping to study the effects of these polymorphisms on morphology, embryo mortality and fitness. All inversions had strong additive effects on morphology, probably resulting from multiple alleles in coupling phase within the inverted haplotype, but surprisingly no dominant genetic effects on morphology or fitness, as would be expected

under heterotic selection. However, the largest inversions increased embryo mortality rates in heterozygous males, meaning that they are underdominant, as expected from theory and as it has also been found in humans. Yet the effect was an order of magnitude smaller than in humans, which may be due to the low and highly skewed recombination rate in zebra finches and an accumulation of multiple clustered inversions, which may effectively suppress pairing and detrimental cross-overs between homologous chromosomes carrying different inversion types in meiosis.

As an alternative to selection at the organismal level keeping the inversions polymorphic, there could be selection at the gene level if one of the two alleles of a diploid organism gets transmitted to the next generation more often than expected under fair Mendelian segregation. This advantage may be balanced by reduced organismal fitness. I performed a genome scan for deviations from Mendelian transmission in a captive population in **Chapter 4** and identified an allele which was passed on more frequently than expected by chance and which was potentially linked to an inversion. Yet this inversion was none of those identified in **Chapter 7** and it appeared to be absent in the wild.

In **Chapters 2 and 3** I first introduced a new molecular method to quantify levels of inbreeding within populations without the need for pedigree information, tested it in a wild and a captive population of zebra finches and explored its performance in comparison to traditional pedigree-based estimates of inbreeding in heterozygosity-fitness correlations. The method treats multiple closely spaced genetic variants as one highly polymorphic marker which, when homozygous, implies common ancestry and thus inbreeding. For the purpose of heterozygosity-fitness correlations, all inbreeding loops are relevant which are younger than the average recessive deleterious mutation in a population. I showed analytically that the method was able to capture these inbreeding loops reaching back hundreds of generations. Interestingly, depending on the aim of a study on inbreeding, pedigree- or molecular marker-based estimates of inbreeding performed better: Regressing fitness over the pedigree-based estimates of inbreeding yielded unbiased slopes, but not so when using marker-based estimates of inbreeding. However, molecular markers performed better in predicting an individual's inbreeding coefficient than the pedigree-based estimate.

As a follow-up on a genome scan for diversity, I mapped the positions of almost all centromeres on the annotated chromosomes in the zebra finch genome in **Chapter 5**. Diversity levels dropped markedly at known zebra finch centromeres and at both ends of the microchromosomes, which are all acrocentric. However, the chromosomal end at which the centromere is located was unknown. Using naturally occurring triploid individuals and genetic markers at both ends of all chromosomes, I was able to link the centromeres to one or the other chromosomal end. Centromeres often stick out in genome scans for diversity or differentiation between populations and species, but it has not been conclusively clarified which selective forces are responsible for this pattern. Mapping the positions of

centromeres has to be the first step towards a better understanding of centromere evolution.

In general, my thesis shows the power and promise of unbiased genome-wide approaches in combination with detailed confirmatory sampling for evolutionary genetics. In this respect, every genome scan is informative and offers new insights, especially for non-trivial genetic phenomena like segregation distortion or large-scale structural variants, which are species- or population-specific and cannot be studied in established model organisms. Genome-wide scans are the hypothesis-generating step of a study and must be confirmed by additional independent sampling to test specific hypotheses; an all too often overlooked statistical principle, which is of immense merit especially in large-scale genomic studies which suffer an immense multiple testing burden. May this approach be tedious, it is the only way towards scientific progress.

Addresses of co-authors

Hans Ellegren
Department of Evolutionary Biology
Uppsala University
Uppsala, Sweden
Email: hans.ellegren@ebc.uu.se

Wolfgang Forstmeier
Department of Behavioural Ecology & Evolutionary Genetics
Max Planck Institute for Ornithology
Seewiesen, Germany
Email: forstmeier@orn.mpg.de

Andre Franke
Institute of Clinical Molecular Biology
Christian-Albrechts-University
Kiel, Germany
Email: a.franke@mucosa.de

Simon C. Griffith
Department of Biological Sciences
Macquarie University
Sydney, NSW, Australia

School of Biological, Earth & Environmental Sciences
University of New South Wales
Sydney, NSW, Australia
Email: simon.griffith@mq.edu.au

Georg Hemmrich-Stanisak
Institute of Clinical Molecular Biology
Christian-Albrechts-University
Kiel, Germany
Email: g.hemmrich-stanisak@ikmb.uni-kiel.de

Bart Kempenaers
Department of Behavioural Ecology & Evolutionary Genetics
Max Planck Institute for Ornithology

Seewiesen, Germany
Email: b.kempenaers@orn.mpg.de

Holger Schielzeth
Department of Evolutionary Biology
Bielefeld University
Bielefeld, Germany
Email: holger.schielzeth@uni-bielefeld.de

Michael Wittig
Institute of Clinical Molecular Biology
Christian-Albrechts-University
Kiel, Germany
Email: m.wittig@mucosa.de

Author contributions

Chapter 1: W.F. conceived the study. U.K. collected data. U.K. and H.S. analyzed the data with input from B.K., H.E. and W.F. U.K. and W.F. wrote the first draft of the manuscript. All authors read and commented on the first draft of the manuscript. All authors approved the final manuscript.

Chapter 2: S.C.G. managed the wild zebra finch population and organized the collection of blood samples. U.K. and W.F. collected blood samples. G.H.S., U.K. and A.F. performed next-generation sequencing and mapping. U.K. designed the SNP chip. M.W. and A.F. performed genotyping. U.K. analyzed the genotype data. U.K. and W.F. wrote the first draft of the manuscript. W.F. and B.K. conceived of the study. All authors read and commented on the first draft of the manuscript. All authors approved the final manuscript.

Chapter 3: U.K., B.K. and W.F. jointly conceived of the study. U.K. designed and implemented the simulation script. U.K. and W.F. analyzed the data. U.K. and W.F. wrote the first draft of the manuscript. All authors read and commented on the first draft of the manuscript. All authors approved the final manuscript.

Chapter 4: U.K., H.S., H.E., B.K. and W.F. jointly conceived of the study. U.K. analyzed the data with input from H.S. and W.F. U.K. wrote the manuscript with input from all authors.

Chapter 5: U.K. and W.F. conceived the study. W.F. selected individuals. U.K. analyzed genotype data. U.K. wrote the manuscript with input from W.F.

Chapter 6: S.C.G. managed the wild zebra finch population and organized the collection of blood samples. U.K. and W.F. collected blood samples and phenotyped the birds. G.H.S., U.K. and A.F. performed next-generation sequencing and mapping. U.K. designed the SNP chip. M.W. and A.F. performed genotyping. U.K. analyzed the genotype data. U.K. and W.F. wrote the first draft of the manuscript. W.F., U.K. and B.K. conceived of the study. All authors read and commented on the first draft of the manuscript. All authors approved the final manuscript.

Chapter 7: S.C.G. managed the wild zebra finch population and organized the collection of blood samples. U.K. and W.F. collected blood samples and phenotyped the birds. G.H.S., U.K. and A.F. performed next-generation sequencing and mapping. U.K. designed the SNP chip. M.W. and A.F. performed genotyping. U.K. analyzed the genotype data. U.K. and W.F. wrote the first draft of the manuscript. W.F., U.K. and B.K. conceived of the study. All

authors read and commented on the first draft of the manuscript. All authors approved the final manuscript.

…….	…….
Bart Kempenaers	Johann Ulrich Knief
Doktorvater	Doktorand

Acknowledgments

My deepest thanks go to Wolfgang Forstmeier whose intellect, critical mind and scientific breadth are outstanding, sometimes surprising and a constant source of inspiration and motivation. In his enthusiastic way he taught me how rigorous science should be done in times of an ever increasing pressure to publish significant results. I am grateful for his patience and trust in me, also in situations where progress was slower than anticipated. I will miss the daily scientific discussions. You are the best supervisor I can imagine.

Second, I want to thank Bart Kempenaers for his positive and also critical comments on my research. He was always open for scientific discussions, supported me in all my scientific projects and enabled me to build the genetic resources I needed for my work. Thank you very much for your faith in me and my skills.

Third, my thanks go to my PhD advisory committee members Jakob Müller and Holger Schielzeth who invested a lot of time in discussing about my research projects and whom I could always ask for advice. Our meetings were always in a nice and friendly atmosphere: Fortiter in re, suaviter in modo.

I also want to thank my co-authors at the Institute of Clinical Molecular Biology in Kiel (Georg Hemmrich-Stanisak, Michael Wittig, and Andre Franke) who always welcomed me and trained me in next-generation sequencing, genotyping and data handling. Your technical expertise made my scientific projects run smoothly.

Most of the tedious work with the birds and in the lab was done by two tireless technical assistants: Katrin Martin and Melanie Schneider to whom I am deeply thankful. Katrin, thank you for breeding the birds, the daily nest checks, the exact data collection and the maintenance of the databases. Melanie, thank you for your efficient and precise lab work, the design and preparation of hundreds of DNA plates and the genotyping of almost endless microsatellites.

My work would not have been possible without a large group of care takers looking after the birds every day: Sonja Bauer, Edith Bodendörfer, Jane Didsbury, Annemarie Grötsch, Andrea Kortner, Frank Lehmann, Petra Neubauer, Katharina Piehler, Frances Weigel and Barbara Wörle. I am very grateful for your work.

Nothing in biology makes sense except you also have a life outside science and I want to thank several people for inspiring me, their friendship and making me feel comfortable throughout (in alphabetic order): ADS for being an exceptional French not only in Japan; AG

for Madrugada and sailing through rough waters; FS for Turbo dance nights with skin and flowers on the carpet, #MTFBWY; HFP for being the good soul of 22B; IS for proofreading, long phone calls, wonderful times and moral guidance; JJN for DTK and everything else; KM for excessive use of vowels and exclamation marks and for bringing Hanseatic flavor to Bavaria; MF for final proofreading, picturesque support from Oz, delayed May pole beers, and strong opinions not only about plants and art; MH for unforgettable Norderoog experiences; MI for being my passionate and supportive binôme; MM for brunch and discussions at the green table, for being Ms. Knief for the mailman, and bearing with my daily German insanity, seriously; RQV for BBQ, jam sessions and drumming; the SFT for fitness and frisbee; SZ for spinning top competitions and the best evening ever; TH for long walks, Tatort nights, and a passion for German dialects; WF for Döner and football evenings; YAA for teaching me philosophy, life and how to use a towel.

Last but definitely not least I want to thank my parents who supported me in all my decisions and put me back on track whenever it was needed. I would not be where I am today without you.

Ulrich Knief
Seewiesen
August 2015

Statutory declaration and statement

Ehrenwörtliche Versicherung

Ich versichere hiermit ehrenwörtlich, dass die von mir vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt worden ist.

München, den ....

Johann Ulrich Knief

Erklärung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist. Im 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 ....

Johann Ulrich Knief