Tissue-specific expression of heterozygous Z-inversions in the zebra finch

Heidi M. Viitaniemi 1,2\*, Erica H. Leder 2,3,4\*, Ondřej Kauzál1, Romana Stopková 5, Pavel Stopka 5, Jan T. Lifjeld 4, Tomáš Albrecht 1,5

1 Institute of Vertebrate Biology, Czech Academy of Sciences, Brno, Czech Republic

2 Section of Ecology and Evolution, Department of Biology, University of Turku, Turku, Finland

3 Tjärnö Marine Laboratory, Linnaeus Centre for Marine Evolutionary Biology, University of Gothenburg, Strömstad, Sweden

4 Natural History Museum, University of Oslo, Oslo, Norway

5 Department of Zoology,Faculty of Science, Charles University, Prague, Czech Republic

\*Joint first authors

Corresponding author: Heidi M. Viitaniemi, [hmviit@gmail.com](mailto:hmviit@gmail.com)

**Abstract**

Chromosomal inversions have been identified in many natural populations and can be responsible for novel traits and rapid adaptation. In zebra finch, an inversion in the Z chromosome has a pleiotropic effect on multiple traits but especially on sperm. Males heterokaryotypic for the inversion have sperm with a longer midpiece and higher velocity. They have higher fertility and siring success than homokaryotypic males. To understand the effect the Z inversion on these traits, we examined young males at different stages of testis development using two homokaryotypic lines and heterokaryotypes. We sampled both testis and liver tissue to identify gene expression differences. In testis, 520 differentially expressed genes were found and most of them were located on chromosome Z. They were functionally enriched for sperm-related traits. In liver, there were some enriched functions and some overrepresentation on chromosome Z, but only when examining specific contrasts. In both tissues, the overrepresentation was located near the distal end of Z but also in the middle of the chromosome. For the heterokaryotype, we observed several genes with expression patterns more similar to one or the other karyotype, based on similarity to homokaryotype invidividuals. This was confirmed with SNPs for two genes, and interestingly one gene, *DMGDH*, had allele-specific expression originating from only one inversion-allele in the testis, yet both inversion alleles were expressed in the liver. This allele-specific difference in tissue-specific expression demonstrates a pleiotropic effect of the inversion and thus suggests a mechanism for divergent phenotypic effects resulting from an inversion.

**Introduction**

Phenotypic polymorphism, where individuals of the same population fall into two or more distinct categories in terms of trait expression (Leimar, 2005), has been identified in many taxa. Intraspecific phenotypic polymorphism observed in nature has been explained by genetic polymorphism ranging in size from single to short nucleotide polymorphisms (*e.g.* Chan et al., 2010) to large chromosomal rearrangements and inversions (*e.g.* Koch et al., 2021). Chromosomal inversions protect inverted sequences from recombination and typically link multiple genetic elements that control a suite of traits (Faria et al., 2019; Wellenreuther & Bernatchez, 2018).

Chromosomal inversions have been identified in many natural populations, and their effects on phenotypes and adaptation have been studied in a variety of systems (Berdan et al., 2021; Crow et al., 2020; Koch et al., 2021; Lavington & Kern, 2017; Puig et al., 2015). Depending on the location and size of the inversion, it can either disrupt a gene sequence in the vicinity of breakpoints or distort gene regulation by affecting the regulatory element landscape up- or downstream from the gene (Puig et al., 2015). Regulation of transcription is affected by the genome architecture, chromatin landscape and cis-regulatory regions all of which result in tissue specific gene expression patterns (Buchberger et al., 2019; Kleinjan & van Heyningen, 2005) and can thus be impacted by an inversion in a tissue-specific manner. For example, an inversion may disrupt the 3D conformation involving one tissue-specific enhancer but not that of another tissue. Discerning how inversions contribute to evolutionary novelty thus requires an understanding of the regulatory machinery across varied tissues to investigate the pleiotropic effects of inversions.

In zebra finch (*Taeniopygia guttata*), an Australian grassfinch in the family Estrildidae, inversions on the Z chromosome seem to have a pleiotropic effect on multiple traits. Various inversions explain as much as 90% of variance in sperm midpiece length in zebra finches as well as aspects of individual physiology, such as visible fat deposition rates, body mass, and beak length (Knief et al., 2016). Three common segregating inversion haplotypes have been characterized from wild zebra finch populations with A representing the ancestral form, and B and C representing the derived inversion haplotypes (Knief *et al.* 2016). A fourth inversion, type D, also exists, but it is not as common as the other three (Knief et al., 2017). Sperm morphology is affected by the inversion such that males with AA karyotype have long flagellum and short midpiece while BB karyotype has short flagellum and long midpiece and heterokaryotypes AB and AC have intermediate flagellum and long midpiece (Kim et al., 2017; Knief et al., 2017). Because of the phenotypic similarity of the sperm from B and C heterokaryotypes, they have also been referred to as AB\* which represents either AB or AC heterokaryotypes (Knief et al., 2016).

Due to the striking variation of sperm phenotypes due to the inversion karyotype, we wanted to use this system for investigating genes involved in sperm morphology. Spermatozoa have evolved to be one of the most diverse cell types in the animal kingdom (Pitnick et al., 2009), and high variation in sperm size is often detected across species (Kahrl et al., 2022), but also within species (Laskemoen et al., 2013) or within populations of the same species (Lifjeld et al., 2010; Míčková et al., 2023). The inversion polymorphism on the Z chromosome in zebra finch is likely maintained in the population by overdominance, with males heterozygous for the inversion producing “super sperm” (Fisher, 2017) with a proportionally large sperm mitochondrial part (midpiece) and increased mitochondrial loading (Knief et al., 2017, 2021). “Super sperm” exhibits the highest velocity compared to homokaryotypes of the inversion (Kim et al., 2017; Knief et al., 2017), and heterokaryotypic males achieve the highest fertility and the highest siring success, both within- and extra-pair matings (Knief et al., 2017). Additionally, several SNPs with an important role in spermatogenesis from the inversion region were previously identified by GWAS and by examining gene expression (Kim et al., 2017).

Although previous studies have focused on the description of the Z chromosome inversion and its effects on sperm performance, the exact molecular mechanisms responsible for phenotypic polymorphism in zebra finch sperm remain unclear. Once mature, individuals constantly synthesize new sperm cells, resulting in a mixture of differentially mature spermatozoa and an extremely complex transcriptomic profile in testicular tissue. This complexity can be reduced by examining sperm cells and testes transcriptomes at a uniform stage, i.e., by examining the testes of males as they mature and synthesize sperm for the first time. Indeed, by sampling testes at different time points through the first wave of spermatogenesis, specific genes involved in the various phases of sperm development were discovered in mice (Laiho et al., 2013).

In this study, we set out to identify genes that are differentially expressed as sperm cells are being generated in the zebra finch. Specifically, we characterized expression transcripts at four different, histologically-characterized timepoints of testis development, as a proxy for the first wave of spermatogenesis, for three karyotypes (ancestral homozygous, heterozygous and derived homozygous; (Knief et al., 2017)). Additionally, we analyzed gene expression in a somatic tissue, liver, to potentially identify differentially expressed genes involved in other processes as well as to examine allele-specific bias between the tissues. Our expectation was that the inversion karyotypes would result in regulatory differences and would thus be observed as differential expression between the inversion karyotypes across spermatogenesis. We found that the inversion affects gene expression in both testis and in liver, and the inversion karyotype also affects which allele is expressed in the heterokaryotype in some genes within the testes, but not within the liver. Our results demonstrate a pleiotropic regulatory effect due to the inversion.

**Material and Methods**

*Bird husbandry and sampling*

Zebra finch with different inversion karyotypes were reared at Institute of Vertebrate Biology (Studenec, Czechia). These birds originate from the domesticated ‘Krakow’ population that was generated by hybridizing between Krakow and Seewiesen populations (population 11 and 8 in (Forstmeier et al., 2007)). The lines contain the homokaryotype line AA, the ancestral karyotype, and the derived inversion types B and C, referred to as the homokaryotype line B\*B\* as described in (Knief et al., 2017).

Crosses were made between individuals in order to produce AA, AB\* and B\*B\* males and inversion genotyping for all samples was done using 6 tag SNPs (Knief et al., 2017). Male offspring from the crosses were sampled between 49-69 days of age to capture different timepoints of testis development from immature to fully mature. Birds were euthanized by using cervical dislocation, and both right testis and liver tissue were frozen in liquid nitrogen and stored at -80°C until further analysis. Left testis was fixed in Davidson’s Fixative, and later prepared for histological determination of testis development. Based on histological sections of the testes, the presence of developing sperm was checked, and developmental stage was determined as follows: timepoint 1 = Inactive with no elongation, no sperm; timepoint2 = Accelerating I with no elongation, no sperm; timepoint 3 = Accelerating II with elongating spermatids visible and maturation index -0.7 – 0.1; timepoint 4 = Active with elongating spermatids visible and maturation index 0.1 – 0.7. The final dataset consisted of 32 birds which spanned across the testis development stage in each of the inversion karyotypes as described in table 1.

*Table 1. Final dataset, across genotypes and testis development stage, consisted of 32 birds. B\* indicates the number of derived chromosome Z inversions an individual carries. B\* can be B or C inversion type.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Karyotype | Timepoint 1: Inactive | Timepoint 2: Accelerating I | Timepoint 3: Accelerating II | Timepoint 4: Active |
| AA | 3 | 3 | 3 | 3 |
| AB\* | 2 | 3 | 2 | 3 |
| B\*B\* | 2 | 2 | 3 | 3 |

*RNA extraction and sequencing*

RNA from testis (right) and liver tissue was extracted with a phenol- based phase separation (Tri-reagent, Ambion) following the Trizol-protocol from SIGMA with minor modifications as follows: approximately 100 mg piece of liver/whole testis was taken and 1ml of Tri was added on the frozen tissue along with a metal bead. Tissue was homogenized using a TissueLyzer (Qiagen) for 30 s with frequency 30/s three times for the liver and one times for 30 s with frequency 30/s for testis. The bead was removed, and 100 µl of 1-bromo-3 chloropropane (SIGMA) was added to the sample. The sample was then shaken vigorously for 15 seconds and incubated at room temperature (RT) for 10 minutes and centrifuged at 4° C for 15 minutes at 12 000 g. The aqueous layer was carefully pipetted out to a clean tube and 500 µl of isopropanol (SIGMA) was added. Tubes were inverted a few times to allow sufficient mixing. After 10 minutes incubation at RT, tubes were centrifuged at 4° C for 10 minutes at 12 000 g to pellet the RNA. To clean the pellet, 1.5 ml of 75% EtOH was added, and the sample was left overnight at 4° C.

To ensure no traces of DNA were left in the extracted RNA, all samples were treated with DNAse (RQ1, Promega) as follows: EtOH was removed from samples after centrifuging at 4° C for 5 minutes at 7 500 g. Samples were left to dry for 5-10 minutes in RT after which 25 µl of nuclease free H2O (Ambion) was added along with 3 µl RQ1 buffer and 2.5 µl of RQ1 DNAse. Samples were mixed gently and incubated 37° C for 45 minutes after which the samples were cooled on ice. To remove traces of the enzyme and DNA, all samples were extracted a second time following the previous protocol but scaling all the reagent volumes down to 400 µl of Tri-Reagent. RNA samples were stored in 75% EtoH at -20° C until prepared for analyses by dissolving into 17 µl of nuclease free H2O.

The RNA samples were quantified with Nanodrop and qualified with Bioanalyzer 2100 (Agilent). In total, 200 ng of RNA was used for TruSeq® Stranded mRNA (Illumina) library preparation and sequenced 100PE with two lanes of NovaSeq 6000 S1 (Illumina). Library preparation and sequencing were performed at the Finnish Functional Genomics Centre (Turku, Finland).

Reads were trimmed with cutadapt v2.7 (Martin, 2011) prior to alignment against the zebra finch genome (Ensembl bTaeGut1\_v1.p, assembly GCA\_003957565.2). Alignment was performed with STAR v2.7.2 (Dobin et al., 2013) using the primary assembly and the corresponding gtf annotation file (Ensembl release 100 bTaeGut1\_v1.p). Read counts per gene were calculated with *quantMode* option provided in STAR.

*Differential expression analysis*

Prior to normalization and differential expression analysis, the count data was prefiltered to account for individual variation in read depths within a gene and to focus on genes where the coverage is high. In each sample all values below 10 in the count data were transformed into 0 after which edgeR v3.40.2 (Robinson et al., 2009) *filterByExpr* was run with minimum count of 10 and total count of 320 to obtain a gene list for downstream analysis. This prefiltered gene list was then used to select genes to retain for differential expression analysis but using the original counts data (*i.e*., not prefiltered). Genes retained in the analysis were normalized using TMM method, and then voom (Law et al., 2014) was run to obtain mean-variance trend to add that to the differential analysis to account for additional variance in genes. Normalization was visually inspected by plotting the raw and normalized count data for testis and liver (Figure S1 and Figure S2). All analysis were performed in R 4.2.2 (R Core Team, 2022).

Liver and testis tissue were analyzed separately for differential expression. Analysis was done with limma v3.54.2 (Smyth, 2005) using *lmFit* and *eBayes.* Contrasts in limma were done using *makeContrasts.* For testis, comparison among the three inversion genotypes was performed within each developmental timepoint. For example, within timepoint 1 (inactive testis) the following contrasts were examined: AA-AB\*, AA-B\*B\*, AB\*-B\*B\*. In the liver samples, we first examined the same contrasts as for the testis tissue, but since there was no effect of testis development time in the liver tissue (Figure 1), all timepoints within an inversion genotype were merged. Therefore, the contrast matrix was determined between inversion genotypes AA-AB\*, AA-B\*B\* and AB\*-B\*B\* by averaging the expression for each inversion genotype across the development groups. Significance was determined separately in each tissue and across the contrast matrix (12 contrasts for testis; 3 contrasts for liver) with *decideTests* for the fits using method “global” and p-value 0.05.

*SNP calling along the Z chromosome*

To be able to estimate allele-specific expression from the RNAseq reads, variant calling was done in both tissues separately. Using bcftools v1.16 (Danecek et al., 2021), *mpileup* and *call* were used to call the variants from the bam files of the mapped reads, and *filter* was used to remove low quality (QUAL<999) loci, indels and low coverage (INFO/DP<640) variants .

To examine the overall SNP pattern across the Z chromosome for estimating the karyotype in our data, bam files for liver and testis were merged to increase the coverage, and variant calling was performed the same way as above for the separate tissues. Variants were filtered for individual depth of 8 (DP<8) and for genotype quality of 10 (GQ<10). This resulted in 8388 SNPs. These were used in the R package GenotypePlot v0.2.1 (Whiting, 2022) to identify the inversion pattern across the Z chromosome and to verify the karyotypes. For the genotype plotting, missingness was set to 0.9 and invariant sites were not allowed, and genotypes were polarized towards the AA individuals to better visualize the pattern of homozygous and heterozygous regions among the AA, AB\* and B\*B\* individuals. We first examined the full chromosome Z (Figure S7) and then focused on the region between 7Mb until the end of the chromosome (Figure 6).

*Functional annotation*

ShinyGO v0.77 (Ge et al., 2018) was used to identify Gene Ontology (GO) term enrichment using the differentially expressed genes from the testis (16360) and liver (14404). Species was set to zebra finch and ensembl geneID was used as input identifiers. The list of the expressed genes in testis and liver were used as the background for enrichment for testis and liver, respectively. For the functional enrichment analysis, all available Gene Ontology sets (*i.e.* Biological Process, Cellular Component, and Molecular Function) were selected for pathway database and pathway minimum size was set to 2 and maximum size to 2000. Redundancy removal was allowed. For statistical enrichment in genomic locations a window size of 4Mb with 4 steps in a window were selected for testis and window size of 8Mb with 4 steps in a window was chosen for liver with FDR cutoff for windows set to 0.01 for both tissues.

Functional overlap of GO terms between the contrasts were investigated with R package ViSEAGO v1.12.0 (Brionne et al., 2019) using the semantic clustering of enriched GO terms identified from each contrasts. This was done as ShinyGo does not allow multiple input lists to be compared to each other. To retain concordance with ShinyGO database, Ensembl2GO was chosen as database and zebra finch was chosen as the species. All genes expressed in testis and liver were used as background, similar to ShinyGO. First all GO terms were identified with ontology as Biological Process and node size of 5 and then statistical significance was determined with the *classic* algorithm and *fishers* test. Results from contrasts were merged within liver and within testis. For identifying clustering in GO terms, we calculated semantic similarity distance between the terms using the Wang method and then used the Ward2 method for aggregation of the similar terms.

*Transcription factors*

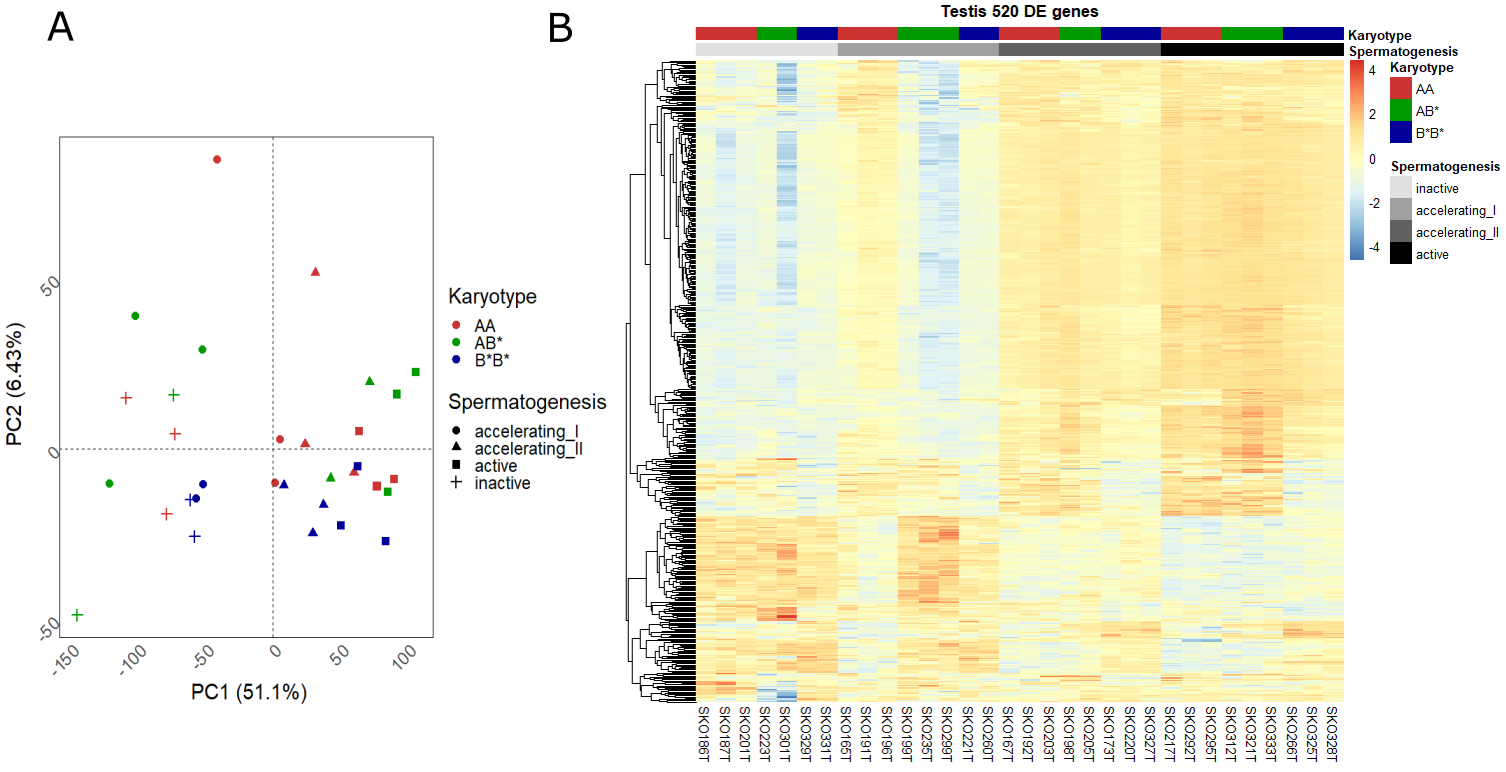
For genes which were found significant in both liver and testis (14 genes) potential upstream regulators were investigated by extracting a 5 Kb region upstream from the gene start using bedtools *getFasta* with the reference genome. For genes on the reverse strand the sequence was reverse complemented. Sequences were input to SEA (Bailey & Grant, 2021) using a shuffled background, setting parameter ‘order to’ 2 for determining the significance and choosing right alignment for identifying the motifs from upstream sequences. For the motif database, JASPAR 2022 CORE (Castro-Mondragon et al., 2022) for vertebrates was chosen.

**Results**

On average, testis samples had 35 184 128 reads and liver samples 29 790 781 reads with average mapping rates being 85.21% and 84.58%, respectively (Table S1). After filtering and normalization in edgeR, reads were compared in a principal component analysis. Reads clustered based on tissue type along PC1 (Figure S3), whereas the testis reads were spread out along PC2 following the progression of the testis development. The liver reads were clustered together as one group along both axes (Figure S3). The timepoints of testis development did not seem to have an impact on liver tissue, even though there could potentially have been a metabolic change as the males develop. After this exploratory analysis, differential expression analysis was done separately for both tissues. Developmental timepoint and inversion karyotype were used for testis, and only differences due to karyotype were examined in the liver.

*Differential expression in testis*

A total of 16 360 genes were expressed in the testis after removing lowly expressed genes. (Table S2). In the PCA analysis PC1 explained 51.1% of the variation and the largest difference among adjacent timepoints was between accelerating I and accelerating II, whereas accelerating II and active were more similar, and inactive and accelerating I were more similar (Figure 1A). Of the expressed genes 520 were differentially expressed (DE) among karyotypes in one or more timepoints. Most of the testis DE genes were found in the accelerating I developmental stage (timepoint 2) between AA-AB\* (Figure 1B).



*Figure 1. Expression in testis. A: PCA plot of all genes expressed in testis. PC1 and PC2 explained 58% of the variation which matches progression of the developmental timepoints we sampled. Stages of spermatogenesis are indicated by shapes: plus for inactive, circle for accelerating I, triangle for accelerating II and square for active. Color indicates the different genotypes AA=red, AB\*=green and B\*B\*=blue. B: Heatmap of the 520 DE genes in testis identified across the 12 contrasts ran in limma where significance was determined across all the contrasts with adjusted p-value <0.05. Color scale is Z-score which is calculated across all samples within each gene. Top rows indicate samples karyotype with AA=red, AB\*=green and B\*B\*=blue and testis developmental stage with light = inactive, light grey =accelerating I, dark grey = accelerating II and black = active.*

There were 48 significant GO categories identified from the functional enrichment analysis for the testis DE genes (Table S4). This enrichment is driven largely by the differential expression between genotypes in accelerating I stage between AA and AB\* (Figure 2) as similar terms were found enriched when comparing the contrasts to each other (Table S5 and Figure S4). In the active stage (timepoint 4), most of the DE genes are between the AB\* and B\*B\* and the AA and B\*B\* comparisons (Figure S4). From the semantic clustering of GO terms, we chose 3 cluster groups which were related to spermatid development and sperm motility for closer investigation. Gene expression of DE genes assigned to the chosen GO term clusters are represented in Figure 2.

Statistically significant enrichment based on gene location was found from chromosome Z when using the 520 DE genes (Figure S5). Two regions were overlapping, and both are located close to the end of the chromosome Z between 64.0 Mb and 68.5 Mb. This area contains 10 genes: DMGDH, ENSTGUG00000024790, GFM2, GCNT4, MAP1B, ENSTGUG00000021944, SMARCA2, DMRT3, DOCK8, GLIS3. A third region with enrichment of differentially expressed genes is identified before the centromere region in chromosome Z between 24.0 and 27.9 Mb (estimated from (Itoh et al., 2011), and it contains 7 genes all of which were lncRNAs: ENSTGUG00000021152, ENSTGUG00000020830, ENSTGUG00000020572, ENSTGUG00000020596, ENSTGUG00000000970, ENSTGUG00000023043.

Kuva, joka sisältää kohteen teksti, kuvakaappaus, viiva, diagrammi

Kuvaus luotu automaattisesti

*Figure 2. Heatmaps of testis gene expression in three functional clusters. Clusters were based on semantic GO term clustering of differentially expressed genes with ViSEAGO using 4 testis contrast comparisons where a minimum of 20 GO terms were identified (Figure S4). A: The gene cluster for spermatid development contains 15 genes. B: The gene cluster for cilium dependent motility contains 25 genes. C: The gene cluster for axoneme, dynein and cilium assembly contains 34 genes. Color scale is Z-score which is calculated across samples within each gene. Top rows indicate the sample’s karyotype with AA=red, AB\*=green and B\*B\*=blue and testis developmental stage with light = inactive, light grey =accelerating I, dark grey = accelerating II and black = active.*

When looking across testis development, eight of the testis DE genes had a consistent pattern of expression in the heterozygote AB\* being towards either A or B\* inversion karyotype (Figure 3). Six of these genes seemed to be B\* inversion specific in the heterozygote AB\* as expression was higher in AB\* than AA, but highest in B\*B\* across all testis development stages: ENSTGUG00000021152 (lncRNA), ENSTGUG00000021693 (novel gene) and *FREM1* (ENSTGUG00000004583) which are located on chromosome Z and ENSTGUG00000020011 (lncRNA) and ENSTGUG00000021745 (lncRNA) located on chromosome 2 and BIN1 (ENSTGUG00000004882) on chromosome 7. Two genes, *DMGDH* (ENSTGUG00000003802) and ENSTGUG00000020830 (lncRNA) were located on chromosome Z and showed an A inversion-specific pattern. Expression was higher in AB\* than in B\*B\* and AB\* is closer to expression of AA across all timepoints.

Kuva, joka sisältää kohteen teksti, diagrammi, viiva, Fontti

Kuvaus luotu automaattisesti

*Figure 3. Inversion specific expression for ENSTGUG00000021152 (lncRNA), ENSTGUG00000021693 (novel gene), DMGDH (ENSTGUG00000003802), ENSTGUG00000020830 (lncRNA) and FREM1 (ENSTGUG00000004583) which are located on chromosome Z and ENSTGUG00000020011 (lncRNA) and ENSTGUG00000021745 (lncRNA) located on chromosome 2 and BIN1 (ENSTGUG00000004882) on chromosome 7. For each developmental stage within each karyotype the mean expression in logCPM is represented and error bars indicate the minimum and maximum value in the group. Karyotype is indicated with AA=red, AB\*=green and B\*B\*=blue.*

*Differential expression in liver*

A total of 14 404 genes were identified from liver (Table S3) with 420 genes differentially expressed among the karyotypes. In general, liver was not clearly separated by testis development stage but rather by karyotype, and the first PC explained 11% of the variation among the expressed genes (Figure 4). No significant enrichment for gene ontology nor statistical enrichment for genomic locations was found among all the 420 DE genes, but when contrasts were examined separately, there are some enriched functions and well as gene regions. From the semantic clustering of GO terms identified in total 21 clusters which were characterized by small overlap between the contrasts. Functions were related to *e.g.* regulation of metabolic processes, fatty acid metabolism and regulation and development of tissues and cells (Table S6). For the comparison between AA-B\*B\* we identified three genomic locations where there was an enrichment of DE genes compared to the background (Figure S6). These were located on chromosome Z and matched to the locations identified in testis (Figure S5). In liver there was also overlap between the DE genes for AA-B\*B\* and AB\*-B\*B\* highlighting differences in B\*B\* from the others (Table S3).

Kuva, joka sisältää kohteen teksti, kuvakaappaus, diagrammi

Kuvaus luotu automaattisesti

*Figure 4. PCA plot of liver genes. A: PC analysis was done for all expressed genes in liver. PC1 and PC2 explained 22% of the variation. In liver, the developmental timepoints did not affect sample distribution. Stages of spermatogenesis are indicated by shapes: plus for inactive, circle for accelerating I, triangle for accelerating II and square for active. Color indicates the different genotypes AA=red, AB\*=green and B\*B\*=blue. B Heatmap of the 420 DE genes in liver identified across the 3 contrasts ran in limma where significance was determined across all the contrasts with adjusted p-value <0.05. Color scale is Z-score which is calculated across all samples by gene. Top rows indicate samples karyotype with AA=red, AB\*=green and B\*B\*=blue and testis developmental stage with light = inactive, light grey =accelerating I, dark grey = accelerating II and black = active.*

*Effect of karyotype between tissues*

To further examine the allele-specific expression pattern due to karyotype, variant calling was performed on the significant genes on the Z chromosome. For *dmgdh,* allele-specific expression in the heterozygote AB\* was found in testis, but not in liver tissue (Figure 5). For gene ENSTGUG00000021693, allele-specific expression was observed in both tissue types in the heterozygote AB\* (Figure 5). For *FREM1,* expression of both karyotype-specific alleles was observed in the heterozygote AB\* (Figure 5). For the remaining five out of the eight genes, no SNPs were identified and allele-specific expression from a specific karyotype could not be verified. For other expressed genes, the patterns for allele-specific expression were not determined as clearly as not all variants could be reliably assigned to either A or B\* inversion with our data.

Kuva, joka sisältää kohteen teksti, kuvakaappaus, diagrammi, Tontti

Kuvaus luotu automaattisesti

*Figure 5. Expression compared with allele frequency in testis and liver for genes DMGDH (ENSTGUG00000003802), FREM1 (ENSTGUG00000004583) and ENSTGUG00000021693. Allele frequency on the x-axis indicates the proportion of the derived Z inversion genotype specific variation obtained from the sequenced reads. Expression level is represented by logCPM on y-axis. Liver=circle and testis=triangle, and* *color indicates the different genotypes AA=red, AB=green and BB=blue.*

*Chromosome Z inversion landscape in taegut1.1*

Plotting of the chromosome Z genotypes from the merged transcriptome data indicated that our AA and B\*B\* individuals were what we expected, but in the AB\* karyotype, the pattern between 20 Mb and 25 Mb is different in 4 individuals compared to the rest of the AB\* samples (Figure 6). The region from 45 Mb to 70 Mb at the distal end of Z discriminates the AA and B\*B\* from each other whereas the region at 2Mb in the beginning of chromosome Z in B\*B\* results from the family relationships in the selection lines (Figure S7, Table S1). As we cannot completely neglect the effect of allele specific expression to our variant calls, particularly in the case of AB\* individuals, we cannot identify the exact the inversion breakpoints with our transcriptome data. Therefore, we attempted to match our data with that of (Kim et al., 2018).

Kuva, joka sisältää kohteen teksti, kuvakaappaus, viiva, Värikkyys

Kuvaus luotu automaattisesti

*Figure 6. Genotype plot from 4011 SNPs from 7Mb until the end of the chromosome. SNPs were acquired after merging the liver and testis transcriptomes.* *Samples are represented vertically, and the physical location along chromosome Z is at the top. The downward lines indicate thechromosome position among the SNPs as the SNP density along the Z is not constant. Each SNP genotype homozygous for the tagut1.1 reference is represented by red, homozygous for the alternative allele is in blue and the heterozygote is in green. Black boxes indicate the regions where overrepresentation of DE genes were identified in both tissues. Full length chromosome Z picture is provided in Figure S7.*

Localization of differentially expressed genes in chromosome Z as well as the patterns of SNP genotypes from our transcriptomes (Figure 6) did not conform to the inversion landscape which was laid out based on the previous zebra finch assembly taegut3.4.2 with respect to breakpoint positions. Rematching the SNP data published by Kim *et al.*, (2017) to the current zebra finch assembly produced a genotype pattern in chromosome Z which was also seen in the transcriptomes (Supplementary File 1 and Figure S8, S9). Based on the pattern of pairwise LD patterns (Figure S9) and genotypes (Figure S8) in Kim *et al.*, (2017) SNP data after reanalyzing to the taegut1.v1 assembly, potential inversion regions are plausible for where we see overrepresentation of differentially expressed gene in testis and liver. In comparison to the genotype plot from transcriptomes only, the end of chromosome Z is similar to the reanalyzed data from Kim *et al.*, (2017) and separates our samples into three groups based on the inversion karyotypes.

*Transcription factor motifs*

In total 14 genes were identified as differentially expressed in both tissues, and the direction of expression between the karyotypes was similar across tissues (Table S7). Upstream motif analysis identified 19 putative binding motifs for these genes (Table S8). All transcription factor genes associated to these motifs had an ortholog in the zebra finch. Of the identified transcription factors, 12 were found expressed in both tissues, three were only expressed in testis and four were not expressed in our data (Table S8).

**Discussion**

As the various inversion karyotypes on the Z chromosome involve many genes, it is not surprising that the effects of the inversions on gene expression are not just restricted to the testis but are also observed in liver. Our most striking finding was that some genes show an expression pattern more similar to one of the karyotypes when the individual is heterokaryotypic rather than an intermediate pattern of expression. Interestingly, this is not only toward one inversion karyotype but changes in a gene-specific manner suggesting that regulation favors the best allele, a ‘best of both worlds’ scenario. This karyotypic expression pattern also occurs for some genes that are not on the Z chromosome which suggests a trans-regulatory role, possibly from one of the many lncRNAs on the Z chromosome, some of which also have allele-specific expression by karyotype. Additionally, it seems that one gene (*dmgdh*) has transcripts that are only expressed from one karyotype in the testis, but transcripts from both karyotypes are expressed in the liver in the heterokaryotypic individuals.

When comparing the locations of the differentially expressed genes and genotypes called from the transcriptomes to the pattern of LD and alleles along the Z chromosome (using SNP data from (Kim *et al.*, 2017), it is plausible that the inversions contribute to the expression patterns observed between the AA, B\*B\* and AB\* karyotypes through differences in physical location or SNP variation in regulatory motifs. Further, our transcriptome data together with SNP data (Kim et al., 2018) shows that the chromosome Z landscape has a different pattern than what the old taegut3.4.2 assembly indicated. In general, inversions can affect gene expression in multiple ways: *e.g.* disrupting the gene sequence, distorting gene regulation through the 3D regulatory conformation or altering chromatin landscape (Puig et al., 2015). Distortion of gene regulation is the most plausible explanation for the expression patterns observed (Figure 3 and 5) since if a gene is disrupted, it is likely non-functional. Pleiotropic effects across tissues could be explained by regulators, such as transcription factors or enhancers. Although we did not identify causative transcription factors from Z chromosome as upstream regulators for the 14 genes differentially expressed in both tissues, it is likely that enhancer sites are more important than transcription binding sites for *cis*-regulatory effects due to inversions since enhancers can be very distant from the gene and are known to be important for temporal and tissue-specific gene regulation (Kosuthova & Solc, 2022; Sabarís et al., 2019). Unfortunately, the techniques required to examine enhancers are not possible within the scope of this study, and we cannot address the 3D structure and the regulatory landscape at this time (Kleinjan & van Heyningen, 2005).

Our results concerning allele-specific expression due to inversion are not unique. In seaweed flies (*Coelopa frigida),* inversion *Cf-Inv(1)* resulted in differential gene expression between the karyotypes across life stage and sex: both *cis*- and *trans*-regulatory effects were identified, and they differed between the sexes and life-stages (Berdan et al., 2021). The *Cf-Inv(1)* inversion is similar in size to the zebra finch Z inversion that it covers most of the chromosome, but the inversion in seaweed fly is autosomal. Another feature in common with our data and others is that the effect of inversion on gene expression is genome-wide and not just restricted to the chromosome containing the inversion (Berdan et al., 2021; Crow et al., 2020; Fuller et al., 2016; Lavington & Kern, 2017). Studying homokaryotypes together with established hetereokaryotypes enables one to disentangle the evolutionarily novel gene expression patterns and how they result from an inversion (Berdan et al., 2021; Fuller et al., 2016), and our results add further support for this approach.

Differences in gene regulation could also result from inversions relocating the gene near heterochromatin which would suppress the expression (Puig et al., 2015; Shanta et al., 2020). Unfortunately testing this would require knowledge of, or at least predicting, the zebra finch Topological Associating Domains (TADs) for chromosome Z and for this reason it was out of scope for the current study. Given that many long-non-coding RNA genes were also differentially expressed due to the inversion suggests that they affect gene regulation in a cell specific manner, as seen in humans (Zhang et al., 2019), or affect the epigenetic pattern of gene regulation (Fatica & Bozzoni, 2014; Wei et al., 2017). Epigenetic regulation is also a potentially important mechanisms for differential gene expression. In humans gene expression in inversion regions is affected by inversion-allele specific methylation which also showed tissue specific effects (Carreras-Gallo et al., 2022). Similarly, in white-throated sparrows inversion ZAL2m showed allele-specific differential methylation in cis-regulatory element of the vasoactive intestinal peptide gene which was previously shown to be differentially expressed between individuals carrying different inversion alleles (Prichard et al., 2022).

We found more DE genes in testis than in liver, and given the distinct sperm phenotypes among the inversion karyotypes as well as the high degree of transcriptional activity in testis, this may be expected. Although in the liver, since the timepoints were merged, there should potentially be more power to detect differences in expression among the karyotypes, which we did not observe. In testis, the DE genes were involved with functions related to cilium (GO:0097729, GO:0060294, GO:0060285) and microtubule organization (GO:0005874, GO:0120031, GO:0007288) and assembly (GO:0035082, GO:0036159, GO:0060271), thus relating the DE genes to cell motility (GO:0048870, GO:0001539 ) and more specifically to sperm motility (GO:0097722, GO:0030317) (Table S4, S5 and Figure 3), and genes involved in these GO terms have also been identified to be involved in the human sperm microtubulome (Jumeau et al., 2017). As we set out to identify genes responsible for the sperm traits, the observed functional enrichment corresponds to the described differences of sperm characteristics of the different inversions (Kim et al., 2017; Knief et al., 2017). Furthermore, the functions identified from our DE genes in testis match those identified in a previous transcriptome analysis comparing long and short sperm phenotypes. (Kim et al., 2017). Regardless of the functional overlap, hardly any significant genes were in common between our data and that of (Kim et al., 2017).

Many of the genes associated to the above-mentioned GO terms are related to defects in sperm mobility, such as CCDC65 (Bower et al., 2018), and flagellum MEIG1 (Salzberg et al., 2010) and TEKT2 (Shimasaki et al., 2010) in human and mouse. These genes cause ‘primary ciliary dyskinesia’ (OMIM entry #244400) which is characterized by defects in motile cilium and decreased sperm motility in males. From the point of sperm movement, genes RSPH9 and ODAD2 are related to ciliary beat pattern. The *rose-comb* mutation in chickens results from an inversion on chromosome 7 which distorts the expression pattern of CFAP65 further causing structural changes in cilia and flagellum leading to the typical comb pattern but also in reduced sperm motility (Imsland et al., 2012).

Most of the DE genes in testis were identified in the AA and AB\* comparison in the accelerating I stage. These genes were also enriched for cilium and microtubule organization and assembly as well as to sperm motility and axonemal assembly (Table S5, Figure 3). When looking at expression across the sampled developmental stages in testis, genes for cilium and flagellum are mostly expressed in the two later timepoints, namely the accelerating II and active timepoints, but in the accelerating I timepoint, AA stands out by having expression pattern more similar to the later timepoints in the other karyotypes (Figure 3B and 3C). In the accelerating I stage, which is characterized by the primary spermatocytes in meiotic division and some round spermatids, but no elongating spermatids or sperm cells, the enriched functions match with previous studies on the stages of spermatogenesis (Laiho et al., 2013). However, a few genes have the opposite pattern with higher expression in AB\* and B\*B\* in the accelerating I stage: DCX and MYO10 in the axoneme, dynein and cilium assembly cluster (Figure 3C) and DND1 and WNT4 in spermatid development (Figure 3A). DCX likely has a role in microtubule dynamics in sperm flagellum (Jumeau et al., 2017) whereas MYO10 is expressed in Lydig cells, Sertoli cells, and spermatogonia in humans (Uhlén et al., 2015). Despite the relevant expression in human gamete formation, no specific function is attributed to spermatogenesis. However, MYO10 is involved in promoting germline-somatic contact in developing oocytes (Granados-Aparici et al., 2022) and is involved in cytoskeleton remodeling in cervical cancer (He et al., 2020) and both of these functions would also be relevant for sperm development in the testis. DND1 is a RNA binding enzyme which inhibits microRNA-mediated repression especially in germline for many genes and causes defects in sperm development if its expression is disrupted (Kedde & Agami, 2008; Westerich et al., 2023). WNT4 is a well-known signaling gene involved in sex-determination (Chue & Smith, 2011), but its correct expression is also important for proper sperm development (Jameson et al., 2012; Wang et al., 2013). From the expression patterns among the karyotypes for these 4 genes, it seems that AA is not expressing these genes as much at the accelerating I stage, and thus variation in the duration of expression temporally is another way which could account for phenotypic differences in sperm, in this case midpiece length which is longer in AB\* and B\*B\* compared to AA.

Although the pattern of differential expression was driven by the accelerating I stage in testis with the majority of the DE genes observed between the AA-AB\* contrast, in the accelerating II and active stages the DE genes are mainly found in the AA-B\*B\* and AB\*-B\*B\* contrasts (Figure 3) which is also the case in liver (Figure 6). From the viewpoint of progression of onset of spermatogenesis in our sampling design, the first two timepoints mainly contain early sperm cell stages whereas the active stage contains cell types from all of the previous stages as well (Aire, 2014), although genes seen for the first time in stage 4 should be involved in the final stages of sperm cell maturation. Although our analysis does not have the resolution of a single-cell analysis, we can discriminate differences in expression which likely reflect the different developmental status of cells along spermatogenesis in the inversion karyotypes. This is also indicated in the heatmaps (Figure 1 and Figure 2) where testis samples form a continuum of changing gene expression across the sampled timepoints from inactive to active in contrast to the pattern in the liver (Figure 4).

Kim *et al.*, (2017) found eight genes in GWAS from the Z inversion which were found associated to sperm morphology: GADD45G, LRRC2, C9orf3, FBXL17 (associated with midpiece length), DMRT2, LINGO2, ZNF462 and RAD23B. None of these eight genes were differentially expressed in our testis data or in liver. In the same study Kim *et al.*, (2017) identified over 100 DE genes between long (AA, AB or AC karyotypes) and short (BB karyotype) sperm phenotypes using a microarray, but of these, only six were differentially expressed in our testis transcripts. Our final timepoint (active) is comparable to other studies which are looking at mature testis where spermatogenesis is ongoing and thus contains a mixture of cells at different developmental stages (Aire, 2014). Although we did not have large overlap between the DE genes identified by Kim *et al.*, (2017), they found *DMGDH* as the most differentially expressed gene between the long (AA, AB or AC karyotypes) and short (BB karyotype) sperm phenotypes sampled from adult males. Our results corroborate this finding although *DMGDH* was not the most strongly differentially expressed gene in our analysis in testis, but in the active stage, which would match Kim et al (2017), the expression difference between AA and AB\* was not very large (log2 FC 0.82, Table S2) whereas to B\*B\* the difference was much larger (log2 FC 3.9, Table S2). GOterms associated to *DMGDH* contain GO:0042426 choline catabolic process and GO:0047865 dimethylglycine dehydrogenase activity indicating that if *dmgdh* is important for sperm function it is likely through the choline pathway. In a human study, a variant in the choline dehydrogenase, *CHDH,* lead to decreased expression which affected sperm function and mobility through reduced ATP concentration (Johnson et al., 2012). So, there may be a functional link between increased *DMGDH* expression and increased sperm motility in AB\* karyotype individuals through the choline metabolism pathway. Given that *DMGDH* gene expression in liver is stably higher than in testis (Figure 4), the true involvement of *DMGDH* for “super sperm” in zebra finch remains unresolved as its function in sperm has not been clearly discerned.

Our data provide verification to the previous study (Kim et al., 2017) but also provide novel aspects to tissue-specific effects of the zebra finch Z chromosome inversions. Due to the focus on testis development, we also provide insight into the potential mechanism behind the overdominance of the heterokaryotypic sperm morphology and function. However, our data contain some limitations due to the difficulty identifying the spermatogenesis stage without histology, since calendar age is not an accurate indicator of spermatogenesis stage. We tried to limit the number of birds sacrificed for the experiment, and this resulted in small sample sizes in the inactive and accelerating I stages for some karyotypes. Another confounding factor is identifying the Z chromosome inversion breakpoints. Our SNP data from the transcriptomes provide some resolution for identifying distinct patterns of homozygosity/heterozygosity along the Z chromosome, but the density of SNPs and the read depth within the transcriptomes is limited. Since we were unable to unequivocally distinguish between the AB or AC heterokaryotype in our data, we retain caution in our interpretation and retain a similar strategy to that of (Knief et al., 2017) such that the main comparison with the heterokaryotypes is the effect of carrying one ancestral and one derived inversion in comparison to homozygotes. Future exploration of the effects of inversions shaping the regulatory landscape and gene expression patterns in hetereokaryotypes would require characterizing the sequence and gene order of the different inversion types A, B, C and D, preferably in females where there is no need to phase the haplotypes.

In conclusion, our unique sampling of zebra finch Z chromosome inversion karyotypes in two different tissues across testis development revealed a pleiotropic regulatory effect due to the inversion. Many genes involved in spermatogenesis were differentially expressed among the karyotypes demonstrating that the inversion alters the regulatory machinery among the karyotypes. The gene enrichment coincides with gene functions that would be strong candidates for the phenotypic differences observed in the sperm, i.e., flagellum and midpiece length. Additionally, we identify several hundred differentially expressed genes in the liver resulting from the various inversion karyotypes. These genes may be influencing some of the physiological phenotypes previously observed such as fat deposition rates and body mass (Knief et al. 2016) since some of the enriched functions from the semantic similarity clustering are related to regulation of metabolic processes and fatty acid metabolism. Further investigation into specific karyotypes should allow for a better understanding of how inversions, by disrupting the existing regulatory landscape, can promote evolutionary novelty.

**Author Contributions**

**TA, EHL, JL designed the original research plan. OK and TA maintained and sampled the birds used for the experiment at the Institute of Vertebrate Biology, CZ. HMV designed and conducted the laboratory work, analyzed the data, and led the writing of the manuscript together with EHL. All authors contributed critically to the drafts and gave final approval for publication.**

**Funding**

This study was funded by the Czech Science Foundation (GAČR) project 19-22538S.

**Data availability**

Sequences, including metadata, produced in this study have been deposited to SRA and are on embargo until acceptance of the manuscript. All code used in the analysis is available upon request.

**Supporting/Supplemental Information**

This manuscript contains supplementary material to support the main text.

Supplementary\_Figures.docx

Supplementary\_Tables.xlsx

Supplementary\_MaterialMethods.docx

**References**

Aire, T. A. (2014). Spermiogenesis in Birds. *Spermatogenesis*, *4*(1), e34346. https://doi.org/10.4161/spmg.34346

Bailey, T. L., & Grant, C. E. (2021). SEA: Simple Enrichment Analysis of motifs. *BioRxiv*, 2021.08.23.457422. https://www.biorxiv.org/content/10.1101/2021.08.23.457422v1%0Ahttps://www.biorxiv.org/content/10.1101/2021.08.23.457422v1.abstract

Berdan, E. L., Mérot, C., Pavia, H., Johannesson, K., Wellenreuther, M., & Butlin, R. K. (2021). A large chromosomal inversion shapes gene expression in seaweed flies ( Coelopa frigida ). *Evolution Letters*, *5*(6), 607–624. https://doi.org/10.1002/evl3.260

Bower, R., Tritschler, D., VanderWaal Mills, K., Heuser, T., Nicastro, D., & Portera, M. E. (2018). DRC2/CCDC65 is a central hub for assembly of the nexin–dynein regulatory complex and other regulators of ciliary and flagellar motility. *Molecular Biology of the Cell*, *29*(2), 137–153. https://doi.org/10.1091/mbc.E17-08-0510

Brionne, A., Juanchich, A., & Hennequet-Antier, C. (2019). ViSEAGO: A Bioconductor package for clustering biological functions using Gene Ontology and semantic similarity. *BioData Mining*, *12*(1), 1–13. https://doi.org/10.1186/s13040-019-0204-1

Buchberger, E., Reis, M., Lu, T. H., & Posnien, N. (2019). Cloudy with a chance of insights: Context dependent gene regulation and implications for evolutionary studies. In *Genes* (Vol. 10, Issue 7). https://doi.org/10.3390/genes10070492

Carreras-Gallo, N., Cáceres, A., Balagué-Dobón, L., Ruiz-Arenas, C., Andrusaityte, S., Carracedo, Á., Casas, M., Chatzi, L., Grazuleviciene, R., Gutzkow, K. B., Lepeule, J., Maitre, L., Nieuwenhuijsen, M., Slama, R., Stratakis, N., Thomsen, C., Urquiza, J., Wright, J., Yang, T., … González, J. R. (2022). The early-life exposome modulates the effect of polymorphic inversions on DNA methylation. *Communications Biology*, *5*(1). https://doi.org/10.1038/s42003-022-03380-2

Castro-Mondragon, J. A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L., Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Perez, N. M., Fornes, O., Leung, T. Y., Aguirre, A., Hammal, F., Schmelter, D., Baranasic, D., Ballester, B., Sandelin, A., Lenhard, B., … Mathelier, A. (2022). JASPAR 2022: The 9th release of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*, *50*(D1), D165–D173. https://doi.org/10.1093/nar/gkab1113

Chan, Y. F., Marks, M. E., Jones, F. C., Villarreal, G., Shapiro, M. D., Brady, S. D., Southwick, A. M., Absher, D. M., Grimwood, J., Schmutz, J., Myers, R. M., Petrov, D., Jonsson, B., Schluter, D., Bell, M. A., & Kingsley, D. M. (2010). Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a Pitx1 Enhancer. *Science*, *327*(5963), 302–305. https://doi.org/10.1126/science.1182213

Chue, J., & Smith, C. A. (2011). Sex determination and sexual differentiation in the avian model. *FEBS Journal*, *278*(7), 1027–1034. https://doi.org/10.1111/j.1742-4658.2011.08032.x

Crow, T., Ta, J., Nojoomi, S., Aguilar-Rangel, M. R., Rodríguez, J. V. T., Gates, D., Rellán-Álvarez, R., Sawers, R., & Runcie, D. (2020). Gene regulatory effects of a large chromosomal inversion in highland maize. *PLoS Genetics*, *16*(12 December), 1–28. https://doi.org/10.1371/journal.pgen.1009213

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2), 1–4. https://doi.org/10.1093/gigascience/giab008

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Faria, R., Johannesson, K., Butlin, R. K., & Westram, A. M. (2019). Evolving Inversions. *Trends in Ecology and Evolution*, *34*(3), 239–248. https://doi.org/10.1016/j.tree.2018.12.005

Fatica, A., & Bozzoni, I. (2014). Long non-coding RNAs: New players in cell differentiation and development. *Nature Reviews Genetics*, *15*(1), 7–21. https://doi.org/10.1038/nrg3606

Fisher, H. S. (2017). Supergene yields super sperm. *Nature Ecology & Evolution*, *1*(8), 1064–1065. https://doi.org/10.1038/s41559-017-0261-0

Forstmeier, W., Segelbacher, G., Mueller, J. C., & Kempenaers, B. (2007). Genetic variation and differentiation in captive and wild zebra finches (Taeniopygia guttata). *Molecular Ecology*, *16*(19), 4039–4050. https://doi.org/10.1111/j.1365-294X.2007.03444.x

Fuller, Z. L., Haynes, G. D., Richards, S., & Schaeffer, S. W. (2016). Genomics of Natural Populations: How Differentially Expressed Genes Shape the Evolution of Chromosomal Inversions in Drosophila pseudoobscura. *Genetics*, *204*(1), 287–301. https://doi.org/10.1534/genetics.116.191429

Ge, S. X., Jung, D., Jung, D., & Yao, R. (2018). ShinyGO: a graphical enrichment tool for ani-mals and plants. *BioRxiv*, *36*(December 2019), 315150. https://doi.org/10.1101/082511.Lai

Granados-Aparici, S., Volodarsky-Perel, A., Yang, Q., Anam, S., Tulandi, T., Buckett, W., Son, W. Y., Younes, G., Chung, J. T., Jin, S., Terret, M. E., & Clarke, H. J. (2022). MYO10 promotes transzonal projection-dependent germ line-somatic contact during mammalian folliculogenesis. *Biology of Reproduction*, *107*(2), 474–487. https://doi.org/10.1093/biolre/ioac078

He, J. H., Chen, J. G., Zhang, B., Chen, J., You, K. L., Hu, J. M., Xu, J. W., & Chen, L. (2020). Elevated MYO10 Predicts Poor Prognosis and its Deletion Hampers Proliferation and Migration Potentials of Cells Through Rewiring PI3K/Akt Signaling in Cervical Cancer. *Technology in Cancer Research and Treatment*, *19*(106), 1–9. https://doi.org/10.1177/1533033820936773

Imsland, F., Feng, C., Boije, H., Bed’hom, B., Fillon, V., Dorshorst, B., Rubin, C. J., Liu, R., Gao, Y., Gu, X., Wang, Y., Gourichon, D., Zody, M. C., Zecchin, W., Vieaud, A., Tixier-Boichard, M., Hu, X., Hallböök, F., Li, N., & Andersson, L. (2012). The Rose-comb mutation in chickens constitutes a structural rearrangement causing both altered comb morphology and defective a strcutucal change in which affects the comb sperm motility. *PLoS Genetics*, *8*(6). https://doi.org/10.1371/journal.pgen.1002775

Itoh, Y., Kampf, K., Balakrishnan, C. N., & Arnold, A. P. (2011). Karyotypic polymorphism of the zebra finch Z chromosome. *Chromosoma*, *120*(3), 255–264. https://doi.org/10.1007/s00412-010-0308-3

Jameson, S. A., Lin, Y. T., & Capel, B. (2012). Testis development requires the repression of Wnt4 by Fgf signaling. *Developmental Biology*, *370*(1), 24–32. https://doi.org/10.1016/j.ydbio.2012.06.009

Johnson, A. R., Lao, S., Wang, T., Galanko, J. A., & Zeisel, S. H. (2012). Choline dehydrogenase polymorphism rs12676 is a functional variation and is associated with changes in human sperm cell function. *PLoS ONE*, *7*(4), 1–12. https://doi.org/10.1371/journal.pone.0036047

Jumeau, F., Chalmel, F., Fernandez-Gomez, F. J., Carpentier, C., Obriot, H., Tardivel, M., Caillet-Boudin, M. L., Rigot, J. M., Rives, N., Buée, L., Sergeant, N., & Mitchell, V. (2017). Defining the human sperm microtubulome: An integrated genomics approach. *Biology of Reproduction*, *96*(1), 93–106. https://doi.org/10.1095/biolreprod.116.143479

Kahrl, A. F., Snook, R. R., & Fitzpatrick, J. L. (2022). Fertilization mode differentially impacts the evolution of vertebrate sperm components. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-34609-7

Kedde, M., & Agami, R. (2008). Interplay between microRNAs and RNA-binding proteins determines developmental processes. *Cell Cycle*, *7*(7), 899–903. https://doi.org/10.4161/cc.7.7.5644

Kim, K.-W., Bennison, C., Hemmings, N., Brookes, L., Hurley, L. L., Griffith, S. C., Burke, T., Birkhead, T. R., & Slate, J. (2017). A sex-linked supergene controls sperm morphology and swimming speed in a songbird. *Nature Ecology and Evolution*, *1*(8), 1168–1176. https://doi.org/10.1038/s41559-017-0235-2

Kim, K.-W., Bennison, C., Hemmings, N., Brookes, L., Hurley, L. L., Griffith, S. C., Burke, T., Birkhead, T. R., & Slate, J. (2018). *Data from: A sex-linked supergene controls sperm morphology and swimming speed in a songbird*. Dryad. https://doi.org/https://doi.org/10.5061/dryad.p4238

Kleinjan, D. A., & van Heyningen, V. (2005). Long-Range Control of Gene Expression: Emerging Mechanisms and Disruption in Disease. *The American Journal of Human Genetics*, *76*(1), 8–32. https://doi.org/10.1086/426833

Knief, U., Forstmeier, W., Kempenaers, B., & Wolf, J. B. W. (2021). A sex chromosome inversion is associated with copy number variation of mitochondrial DNA in zebra finch sperm. *Royal Society Open Science*, *8*(9), 211025. https://doi.org/10.1098/rsos.211025

Knief, U., Forstmeier, W., Pei, Y., Ihle, M., Wang, D., Martin, K., Opatová, P., Albrechtová, J., Wittig, M., Franke, A., Albrecht, T., & Kempenaers, B. (2017). A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. *Nature Ecology & Evolution*, *1*(8), 1177–1184. https://doi.org/10.1038/s41559-017-0236-1

Knief, U., Hemmrich-Stanisak, G., Wittig, M., Franke, A., Griffith, S. C., Kempenaers, B., & Forstmeier, W. (2016). Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biology*, *17*(1), 199. https://doi.org/10.1186/s13059-016-1056-3

Koch, E. L., Morales, H. E., Larsson, J., Westram, A. M., Faria, R., Lemmon, A. R., Lemmon, E. M., Johannesson, K., & Butlin, R. K. (2021). Genetic variation for adaptive traits is associated with polymorphic inversions in Littorina saxatilis. In *Evolution Letters* (Vol. 5, Issue 3, pp. 196–213). https://doi.org/10.1002/evl3.227

Kosuthova, K., & Solc, R. (2022). Inversions on human chromosomes. In *American Journal of Medical Genetics, Part A*. https://doi.org/10.1002/ajmg.a.63063

Laiho, A., Kotaja, N., Gyenesei, A., & Sironen, A. (2013). Transcriptome Profiling of the Murine Testis during the First Wave of Spermatogenesis. *PLoS ONE*, *8*(4). https://doi.org/10.1371/journal.pone.0061558

Laskemoen, T., Albrecht, T., Bonisoli-Alquati, A., Cepak, J., de Lope, F., Hermosell, I. G., Johannessen, L. E., Kleven, O., Marzal, A., Mousseau, T. A., Møller, A. P., Robertson, R. J., Rudolfsen, G., Saino, N., Vortman, Y., & Lifjeld, J. T. (2013). Variation in sperm morphometry and sperm competition among barn swallow (Hirundo rustica) populations. *Behavioral Ecology and Sociobiology*, *67*(2), 301–309. https://doi.org/10.1007/s00265-012-1450-0

Lavington, E., & Kern, A. D. (2017). The effect of common inversion polymorphisms In(2L)t and In(3R)Mo on patterns of transcriptional variation in Drosophila melanogaster. *G3: Genes, Genomes, Genetics*, *7*(11), 3659–3668. https://doi.org/10.1534/g3.117.1133

Law, C. W., Chen, Y., Shi, W., & Smyth, G. K. (2014). Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology*, *15*(2), R29. https://doi.org/10.1186/gb-2014-15-2-r29

Leimar, O. (2005). The Evolution of Phenotypic Polymorphism: Randomized Strategies versus Evolutionary Branching. *The American Naturalist*, *165*(6), 669–681. https://doi.org/10.1086/429566

Lifjeld, J. T., Laskemoen, T., Kleven, O., Albrecht, T., & Robertson, R. J. (2010). Sperm length variation as a predictor of extrapair paternity in passerine birds. *PLoS ONE*, *5*(10). https://doi.org/10.1371/journal.pone.0013456

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Míčková, K., Tomášek, O., Jelínek, V., Šulc, M., Pazdera, L., Albrechtová, J., & Albrecht, T. (2023). Age-related changes in sperm traits and evidence for aging costs of sperm production in a sexually promiscuous passerine. *Frontiers in Ecology and Evolution*, *11*(February), 1–9. https://doi.org/10.3389/fevo.2023.1105596

Pitnick, S., Hosken, D. J., & Birkhead, T. R. (2009). Sperm morphological diversity. In T. R. Birkhead, D. J. Hosken, & S. Pitnick (Eds.), *Sperm Biology: An Evolutionary Perspective* (1st ed., pp. 69–149). Elsevier.

Prichard, M. R., Grogan, K. E., Merritt, J. R., Root, J., & Maney, D. L. (2022). Allele-specific cis-regulatory methylation of the gene for vasoactive intestinal peptide in white-throated sparrows. *Genes, Brain and Behavior*, *21*(8), 1–12. https://doi.org/10.1111/gbb.12831

Puig, M., Casillas, S., Villatoro, S., & Cáceres, M. (2015). Human inversions and their functional consequences. *Briefings in Functional Genomics*, *14*(5), 369–379. https://doi.org/10.1093/bfgp/elv020

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Sabarís, G., Laiker, I., Preger-Ben Noon, E., & Frankel, N. (2019). Actors with Multiple Roles: Pleiotropic Enhancers and the Paradigm of Enhancer Modularity. *Trends in Genetics*, *35*(6), 423–433. https://doi.org/10.1016/j.tig.2019.03.006

Salzberg, Y., Eldar, T., Karminsky, O. D., Itach, S. B. S., Pietrokovski, S., & Don, J. (2010). Meig1 deficiency causes a severe defect in mouse spermatogenesis. *Developmental Biology*, *338*(2), 158–167. https://doi.org/10.1016/j.ydbio.2009.11.028

Shanta, O., Noor, A., & Sebat, J. (2020). The effects of common structural variants on 3D chromatin structure. *BMC Genomics*, *21*(1), 95. https://doi.org/10.1186/s12864-020-6516-1

Shimasaki, S., Yamamoto, E., Murayama, E., Kurio, H., Kaneko, T., Shibata, Y., Inai, T., & Iida, H. (2010). Subcellular localization of Tektin2 in rat sperm flagellum. *Zoological Science*, *27*(9), 755–761. https://doi.org/10.2108/zsj.27.755

Smyth, G. K. (2005). limma: Linear Models for Microarray Data. In *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* (Issue 2005, pp. 397–420). Springer-Verlag. https://doi.org/10.1007/0-387-29362-0\_23

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I. M., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A. K., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., … Pontén, F. (2015). Tissue-based map of the human proteome. *Science*, *347*(6220). https://doi.org/10.1126/science.1260419

Wang, X. N., Li, Z. S., Ren, Y., Jiang, T., Wang, Y. Q., Chen, M., Zhang, J., Hao, J. X., Wang, Y. B., Sha, R. N., Huang, Y., Liu, X., Hu, J. C., Sun, G. Q., Li, H. G., Xiong, C. L., Xie, J., Jiang, Z. M., Cai, Z. M., … Gao, F. (2013). The Wilms Tumor Gene, Wt1, Is Critical for Mouse Spermatogenesis via Regulation of Sertoli Cell Polarity and Is Associated with Non-Obstructive Azoospermia in Humans. *PLoS Genetics*, *9*(8). https://doi.org/10.1371/journal.pgen.1003645

Wei, J. W., Huang, K., Yang, C., & Kang, C. S. (2017). Non-coding RNAs as regulators in epigenetics (Review). *Oncology Reports*, *37*(1), 3–9. https://doi.org/10.3892/or.2016.5236

Wellenreuther, M., & Bernatchez, L. (2018). Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends in Ecology and Evolution*, *33*(6), 427–440. https://doi.org/10.1016/j.tree.2018.04.002

Westerich, K. J., Reinecke, S., Emich, J., Wyrwoll, M. J., Stallmeyer, B., Meyer, M., Oud, M. S., Fietz, D., Pilatz, A., Kliesch, S., Reichman-Fried, M., Tarbashevich, K., Limon, T., Stehling, M., Friedrich, C., Tüttelmann, F., & Raz, E. (2023). Linking human Dead end 1 (DND1) variants to male infertility employing zebrafish embryos. *Human Reproduction*, *38*(4), 655–670. https://doi.org/10.1093/humrep/dead031

Whiting, J. R. (2022). *JimWhiting91/genotype\_plot: Genotype Plot*. https://doi.org/10.5281/zenodo.5913504

Zhang, X., Zhang, P., Song, D., Xiong, S., Zhang, H., Fu, J., Gao, F., Chen, H., & Zeng, X. (2019). Expression profiles and characteristics of human lncRNA in normal and asthenozoospermia sperm. *Biology of Reproduction*, *100*(4), 982–993. https://doi.org/10.1093/biolre/ioy253