Compensation of Dosage-Sensitive Genes on the Chicken Z Chromosome

Fabian Zimmer*,[1], Peter W. Harrison[1], Christophe Dessimoz[1,2,3], and Judith E. Mank[1]

[1]Department of Genetics Evolution and Environment, University College London, London, United Kingdom

[2]Department of Ecology and Evolution & Center for Integrative Genomics, University of Lausanne, Biophore 1015, Lausanne, Switzerland

[3]Swiss Institute of Bioinformatics, Biophore, 1015 Lausanne, Switzerland

*Corresponding author: E-mail: fabian.zimmer.12@ucl.ac.uk.

Accepted: March 26, 2016

Data deposition: This project has been deposited at the NCBI Short Read Archive under the accession SRP065394.

Abstract

In many diploid species, sex determination is linked to a pair of sex chromosomes that evolved from a pair of autosomes. In these organisms, the degeneration of the sex-limited Y or W chromosome causes a reduction in gene dose in the heterogametic sex for X- or Z-linked genes. Variations in gene dose are detrimental for large chromosomal regions when they span dosage-sensitive genes, and many organisms were thought to evolve complete mechanisms of dosage compensation to mitigate this. However, the recent realization that a wide variety of organisms lack complete mechanisms of sex chromosome dosage compensation has presented a perplexing question: How do organisms with incomplete dosage compensation avoid deleterious effects of gene dose differences between the sexes? Here we use expression data from the chicken (*Gallus gallus*) to show that ohnologs, duplicated genes known to be dosage-sensitive, are preferentially dosage-compensated on the chicken Z chromosome. Our results indicate that even in the absence of a complete and chromosome wide dosage compensation mechanism, dosage-sensitive genes are effectively dosage compensated on the Z chromosome.

Key words: dosage sensitivity, whole genome duplication, sex chromosomes, ohnologs.

Introduction

Heteromorphic sex chromosomes have evolved independently in many species (Bachtrog et al. 2014; Beukeboom and Perrin 2014). In some cases, recombination has been suppressed along the majority of the length of the sex chromosomes, leading to a large-scale loss of active genes from the sex-limited Y and W chromosomes (Charlesworth et al. 2005; Bachtrog et al. 2011). This results in large differences in size, with one large, gene-rich chromosome (X or Z chromosome), and one smaller chromosome, lacking many genes (Y or W chromosome).

The decay of Y and W chromosome gene content leads to differences in gene dose between the sexes, where the heterogametic sex has one half of the dose of all genes lost from the sex-limited chromosome compared with the homogametic sex. For many loci, gene dose correlates with gene expression (Pollack et al. 2002; Birchler et al. 2005; Torres et al. 2007; Malone et al. 2012), therefore the reduced gene dose on the X or Z chromosome should result in reduced gene expression in the heterogametic sex. When dosage-sensitive genes are affected, this could lead to a reduction in fitness in the heterogametic sex, and result in selective pressures favoring the evolution of dosage compensation mechanisms (Ohno 1967; Charlesworth 1978, 1996, 1998). These mechanisms should equalize the expression between the sex chromosomes and the autosomes, thereby restoring them to the ancestral level before the decay of gene content on the W or Y chromosome. Second, they should equalize the expression of individual dosage-sensitive genes between males and females.

Although it was once assumed that complete and global dosage compensation would always be associated with sex chromosome evolution (Ohno 1967), there is considerable variation in the mechanism and completeness of dosage compensation across species. For example, in *Drosophila melanogaster* (Conrad and Akhtar 2012) and *Caenorhabditis elegans* (Meyer 2010), dosage balance is achieved through regulatory mechanisms affecting the entire X chromosome (Straub and Becker 2007). In these cases, differences in gene dose of the

sex chromosome are compensated for and expression is on average balanced between the sexes for the X chromosome, and between the single X and the diploid autosomes in males, the heterogametic sex. However, it is now clear that complete mechanisms of dosage compensation are rare, and many organisms, including birds (Ellegren et al. 2007; Itoh et al. 2007; Naurin et al. 2011; Wolf and Bryk 2011; Uebbing et al. 2013; Wright et al. 2015), snakes (Vicoso et al. 2013), many insects (Vicoso and Bachtrog 2015), and fish (Leder et al. 2010; Chen et al. 2014), have incomplete dosage compensation (reviewed in Mank 2013).

Incomplete dosage compensation was first documented in chicken (Ellegren et al. 2007; Itoh et al. 2007) and subsequently confirmed in several other avian species (Naurin et al. 2011; Wolf and Bryk 2011; Uebbing et al. 2013; Wright et al. 2015). In birds, which are a model for studies of incomplete dosage compensation, there is a significant reduction in average expression of the Z chromosomes in females, the heterogametic sex, relative to the autosomes as well as to the male Z chromosome average (Ellegren et al. 2007; Itoh et al. 2007; Wolf and Bryk 2011; Uebbing et al. 2013, 2015). The realization that many organisms with heteromorphic sex chromosomes have not in fact evolved complete and global dosage compensation mechanisms is perplexing as it is unclear how these organisms cope with negative dose effects. A reduction in gene dose often does not produce an observable difference in expression for many genes (Malone et al. 2012), and it was unclear whether certain loci are actively dosage-compensated or simply lack dose effects.

One possible explanation proposed by Mank and Ellegren (2008) is that instead of requiring a global mechanism of dosage compensation, the regulation of gene dose might occur on a gene-by-gene basis. A more targeted, local mechanism of dosage compensation should primarily affect the expression of dosage-sensitive genes (Mank et al. 2011). The role of dosage-sensitivity for the evolution of dosage compensation mechanisms has been discussed by a number of reviews (Mank 2013; Pessia et al. 2013; Ercan 2015; Veitia et al. 2015) and was investigated in a range of species. For example, in mammals X chromosomal expression is reduced compared with the autosomes in both males and females (Xiong et al. 2010; Julien et al. 2012), possibly as a consequence of X chromosome inactivation. However, dosage-sensitive genes, such as protein–complexes, show evidence of a higher degree of dosage-compensation (Lin et al. 2012; Pessia et al. 2012), compared with other gene categories. Recent studies in nematodes (Albritton et al. 2014) and fish (White et al. 2015) also showed similar patterns of compensated dosage-sensitive genes.

Dosage-sensitivity can result from interactions with other genes or gene products (Veitia 2004), such as in the case of transcription factors and large protein complexes (Papp et al. 2003). Individual duplications of these dosage-sensitive genes are likely to be rare, as they disrupt the stoichiometric balance and may disturb gene networks (Birchler et al. 2001; Papp et al. 2003; Birchler and Veitia 2012). However, dosage-sensitive genes should be preferentially retained after whole genome duplications (WGDs) (Edger and Pires 2009; Birchler and Veitia 2012). In contrast, dosage-insensitive genes that do not exhibit neo- or sub-functionalization are often lost after WGD (Dehal and Boore 2005). WGDs have occurred in a wide range of lineages (Wolfe and Shields 1997; Kellis et al. 2004; Dehal and Boore 2005; Cui et al. 2006; Van de Peer et al. 2009), including two rounds of WGD events roughly 500 MYA ago (Dehal and Boore 2005), which gave rise to roughly 16–34% of the chicken genome (Singh et al. 2015).

Preferentially retained gene duplicates originating from WGDs, also known as ohnologs (Wolfe 2000, 2001), are skewed toward gene families associated with dosage-sensitive functions such as signaling and development (Blomme et al. 2006) and protein–complexes (Makino et al. 2009). The dosage sensitivity of ohnologs (Blomme et al. 2006; Makino et al. 2009) is well established and makes them particularly useful in assessing the effectiveness of incomplete dosage compensation. We therefore use ohnologs to investigate the effectiveness of compensation on the chicken Z chromosome and to understand the evolution of incomplete sex chromosome dosage compensation mechanisms in general.

Results

We generated RNA-Seq gene expression profiles from multiple male and female biological replicates for four different tissues (spleen, heart, liver, and gonad) in chicken (*Gallus gallus*), recovering on average 17 million paired-end mappable reads per sample. We removed genes that were not expressed on average in all male and female above at least two counts per million (CPM). The number of genes expressed on the autosomes and Z chromosome for each tissue are shown in supplementary table S1, Supplementary Material online.

Incomplete Dosage Compensation in Females and Reduced Z Expression in Males

Dosage compensation has been assessed in a variety of ways, often depending on the system being studied. We used two approaches to assess dosage compensation status. First, complete dosage compensation should equalize female Z-linked and autosomal expression. Second, dosage compensation can also act on a local gene-by-gene basis, balancing the individual gene expression in males and females, which may be the dominant mechanism for dosage-sensitive genes.

Consistent with previous studies showing the incomplete dosage compensation in chicken, we detected lower average expression of Z-linked genes in comparison to autosomal genes in all female tissues (spleen $P < 0.0001$, Z-score = 11.19; heart $P < 0.0001$, Z-score = 11.22; liver $P < 0.0001$, Z-score = 8.88; ovaries $P < 0.0001$, Z-score = 9.20; Wilcoxon

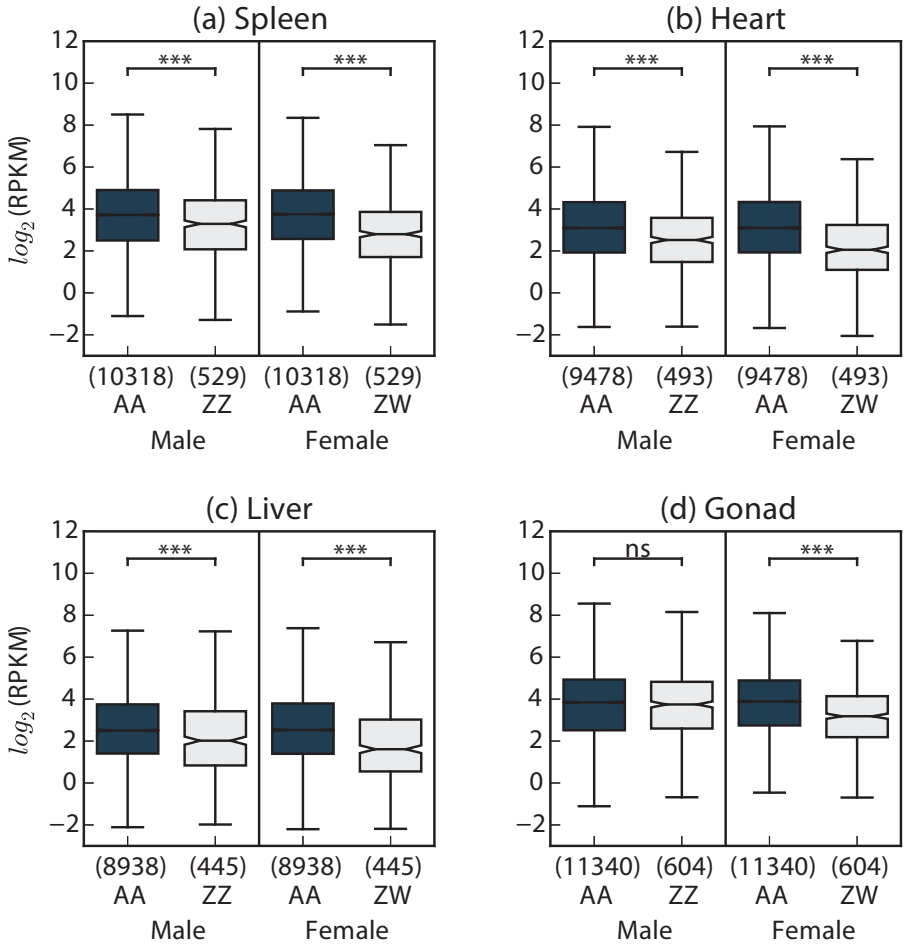

Fig. 1.—Comparison of gene expression measured for autosomal genes (dark grey) and Z-linked genes (light grey) in (*a*) spleen, (*b*) heart, (*c*) liver, and (*d*) gonad tissue in males and females. In all tissues, gene expression for Z-linked genes is significantly lower in comparison to autosomal genes in females. In males, gene expression of Z-linked genes is significantly lower in comparison to autosomal genes in all somatic tissues but not in gonad. Significance levels are indicated as stars ($*P < 0.05$, $**P < 0.001$, $***P < 0.0001$), differences between distributions were tested using Wilcoxon Rank Sum tests. The number of genes expressed on the autosomes and Z chromosome(s) are given in brackets for each distribution. Boxes show the interquartile range, notches represent the median of the distribution and whiskers extend to 1.5 times the interquartile range (Q3 + 1.5 × IQR, Q1–1.5 × IQR). Outliers are not shown for clarity, but included in all statistical comparisons.

Rank Sum Test, fig. 1, supplementary fig. S1 and table S2, Supplementary Material online). We also expect that the average expression of the Z chromosomes in males is similar to the autosomal average, as two Z chromosomes are present. In line with this prediction, we find that the distribution of male expression is not significantly different to the autosomes in testes ($P = 0.79$, Z-score = 0.27, Wilcoxon Rank Sum Test). However, a previous study has indicated that in some tissues, expression of the Z in males is also less than the autosomal average (Julien et al. 2012), and we also recovered a significant reduction in average expression of Z-linked loci compared with average autosomal expression in all somatic tissues in males (spleen $P < 0.0001$, Z-score = 5.50; heart $P < 0.0001$, Z-score = 6.69; liver $P < 0.0001$, Z-score = 5.02; Wilcoxon Rank Sum Test). When we compared the average expression level of all autosomes and the Z chromosomes, it is clear that the Z chromosome expression in both males and females is outside the autosomal spectrum for all somatic tissues (supplementary fig. S1, Supplementary Material online).

One possible explanation for the low *Z* expression could be the inclusion of lowly expressed genes, but the median *Z*:*A* ratios for males (*ZZ*:*AA*) and females (*Z*:*AA*) across a range of higher CPM expression thresholds (supplementary fig. S2, Supplementary Material online) is similar, suggesting that a minimum CPM threshold >2 is effective in filtering out lowly expressed genes. The difference in male and female Z-linked gene expression is also robust across expression quartiles, except for gonad expression quartile one (supplementary fig. S3, Supplementary Material online). The reduction in Z expression in males is also consistent with the possible inactivation of

one Z chromosome in males, analogous to the X inactivation observed in therian females (Cooper et al. 1993; Deakin et al. 2009). Male Z chromosome inactivation has been suggested by previous work on a limited number of Z-linked loci (Livernois et al. 2013) and we investigated the potential for Z inactivation using our RNA-Seq data. If one copy of the Z chromosome were partially inactivated in males, we would expect to find SNPs with a significantly greater contribution to the total expression from one allele at heterozygous sites. Our analyses of allele-specific expression (ASE) indicate that only a limited number of Z-linked genes exhibit ASE, and there is no robust evidence that the proportion is greater than that observed for the autosomes (Supplementary Material online). This suggests that the reduction in male expression on the Z chromosome is not due to chromosomal inactivation.

Ohnologs Are Preferentially Dosage-Compensated

If incomplete dosage compensation is sufficient for compensating dosage-sensitive genes, we might expect the proportion of dosage-compensated ohnologs on the Z chromosome to be higher in comparison to nonohnologs. We tested whether ohnologs are more often dosage-compensated using our expression data and ohnologs obtained from the OhnologsDB (Singh et al. 2015). The chicken genome contains 5,228 (33.71%) annotated ohnologs, of which 223 are annotated on the Z chromosome. Z chromosome ohnologs show over-enrichment for Gene Ontology terms compared with all genes, such as cell motility and locomotion, which may be important in dosage sensitivity (supplementary table S3, Supplementary Material online).

In order to determine whether ohnologs are preferentially dosage-compensated, we first compared the $\log_2$ fold change between female and male expressions for Z-linked ohnologs and nonohnologs (fig. 2). The difference in expression between females and males ($\log_2$FC) was significantly lower for ohnologs than nonohnologs (spleen $P < 0.0001$, Z-score = 5.95; heart $P < 0.0001$, Z-score = 4.57; liver $P < 0.0001$, Z-score = 5.22; gonad $P < 0.0001$, Z-score = 4.89; Wilcoxon Rank Sum Test), suggesting a higher degree of dosage compensation. In addition, the proportion of dosage-compensated ohnologs ($\log_2$FC range from −0.5 to 0.5) was significantly higher when compared with nonohnologs in all tissues (P-value < 0.0001 in all comparisons; Fisher's Exact test, table 1). This is also the case when we used a wider range of $\log_2$FC (−0.6 to 0.6), similar to the mean expression change for female one-dose genes reported by Malone et al. (2012) (supplementary table S4, Supplementary Material online). In addition, we used the strict set of ohnologs from the OhnologsDB, with 2,489 ohnologs annotated in the chicken genome and 106 on the Z chromosome, recovering similar results (supplementary fig. S4 and Table S5, Supplementary Material online).

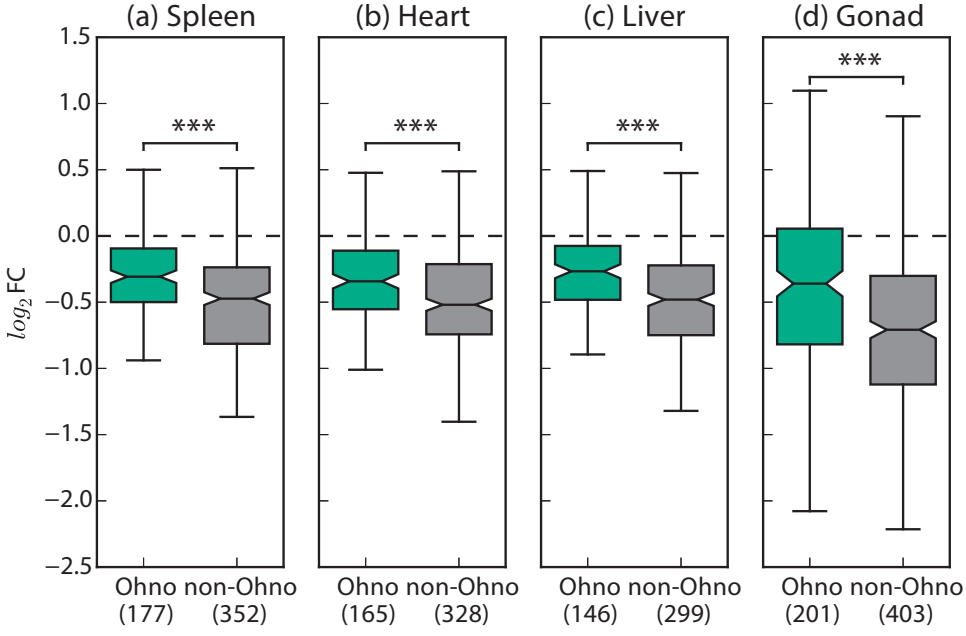

Fig. 2.—Comparison of $\log_2$-transformed fold change between female and male expressions for ohnologs (green) and nonohnologs (grey) on the Z chromosome in (*a*) spleen, (*b*) heart, (*c*) liver, and (*d*) gonad. The number of genes in the distributions is given in brackets. Negative fold changes indicate higher male expression; positive fold changes indicate stronger female expression. Significance levels are indicated as stars (*$P < 0.05$, **$P < 0.001$, ***$P < 0.0001$), differences between distributions were tested using Wilcoxon Rank Sum tests. Outliers are not shown for clarity, but included in all statistical comparisons.

Table 1

Contingency Tables for All Four Tissues, Comparing the Proportion of Dosage-Compensated (DC) and Uncompensated (U) Ohnologs to Non-ohnologs Using a Fisher's Exact Test

	Ohnolog		Non-ohnolog		*P* value	Odds ratio
	DC	U	DC	U		
Spleen	126 (71.19%)	51 (28.81%)	180 (51.14%)	172 (48.86%)	**1.08×10^{-5}**	2.36
Heart	111 (67.27%)	54 (32.73%)	152 (46.34%)	176 (53.66%)	**1.06×10^{-5}**	2.38
Liver	105 (71.92%)	41 (28.08%)	147 (49.16%)	152 (50.84%)	**6.52×10^{-6}**	2.65
Gonad	86 (42.79%)	115 (57.21%)	103 (25.56%)	300 (74.44%)	**2.57×10^{-5}**	2.18

NOTE—Significant *P* values are reported in bold.

An alternative explanation for the high degree of dosage compensation among ohnologs is that all paralogs, even those that originate in single-gene duplications, are dosage-compensated. We tested this hypothesis by extracting Z-linked paralogs from the Ensembl database (Cunningham et al. 2015) that originated in single-gene duplication events. These paralogs do not show a higher proportion of dosage compensation ($P > 0.05$ in all comparisons; Fisher's Exact test; supplementary table S6, Supplementary Material online) compared with all other genes on the Z chromosome. This indicates that the higher degree of dosage compensation among ohnologs is not a property of paralogs in general, and that the mode of duplication has an important impact on the evolution of gene-by-gene dosage compensation.

Older Z Chromosome Parts Contain Fewer Ohnologs

Sex chromosome divergence can drive the movement of some gene classes off the sex chromosomes (Emerson et al. 2004; Potrzebowski et al. 2008; Vibranovski et al. 2009) and we might expect an out of Z migration for dosage-sensitive genes. Overall, the proportion of ohnologs is not significantly different between the Z (764 coding genes) and the genomic background (14,744 coding genes) ($P = 0.19$, odds ratio = 0.89; Fisher's Exact test), suggesting that the Z chromosome is not depleted of ohnologs and that dosage-sensitive gene have not moved off the Z. However, the Z chromosome contains at least four strata, where recombination was suppressed between the Z and W at different times, spanning roughly 130 million years (Wright et al. 2012). We divided the chromosome into an old and young parts along the border of stratum 3, resulting in two almost equally sized regions of the Z chromosome. Given 223 ohnologs located on the Z chromosome, we expect that half of these would be located in the old and half in the young part of the chromosome. However, the number of ohnologs in the older half of the chromosome is significantly less than expected ($\chi^2 = 22.605$, $P < 0.0001$; Chi-square test), and also significantly less when accounting for the difference in gene content ($P < 0.05$, odds ratio = 0.62; Fisher's Exact test). This could indicate that some ohnologs may have relocated during the early evolution of the Z chromosome. When we compared the proportion of dosage-compensated ohnologs between old and young parts of the Z chromosome, we do not detect a significantly higher proportion of dosage-compensated ohnologs in older parts ($P > 0.05$ in all comparisons; Fisher's Exact test), suggesting that dosage compensation of ohnologs occurs relatively quickly following W chromosome gene loss. Alternatively, this bias could be an artifact of the ancestral ohnolog distribution, as the WGD events precede the formation of the sex chromosome system.

Dosage Compensation of Ohnologs across Tissues

The degree of dosage compensation is similar in all somatic tissues ($P > 0.05$ in all comparisons; Fisher's Exact test; supplementary table S7, Supplementary Material online), and greater in the soma compared with the gonad ($P < 0.0001$ in all comparisons; Fisher's Exact test; supplementary table S7, Supplementary Material online). Tissues can be seen as a form of functional compartmentalization, and the same gene can show a diverse range of expression patterns in different tissues. For this reason, similar overall dosage compensation could hide an underlying pattern of pleiotropic expression. Dosage sensitivity may in fact be tissue dependent and can result in gene-by-gene dosage compensation (Mank and Ellegren 2008).

We also investigated the overlap of dosage-compensated ohnologs across tissues. A set of 68 of 223 ohnologs was dosage-compensated in all somatic tissues; however, we detected substantial variation (fig. 3). Of the 68 ohnologs that are dosage-compensated in all somatic tissues, only 36 are also dosage-compensated in gonad, showing that only a small core set of ohnologs are dosage-sensitive across all tissues. In gonad, a unique set of 50 ohnologs was dosage-compensated. In combination with the overall lower degree of dosage compensation in gonad, this suggests different dosage compensation patterns when compared with the somatic tissues.

Discussion

Our analyses of dosage compensation and ohnologs on the chicken Z chromosome provide novel insights into the nature of incomplete dosage compensation. We confirm previous

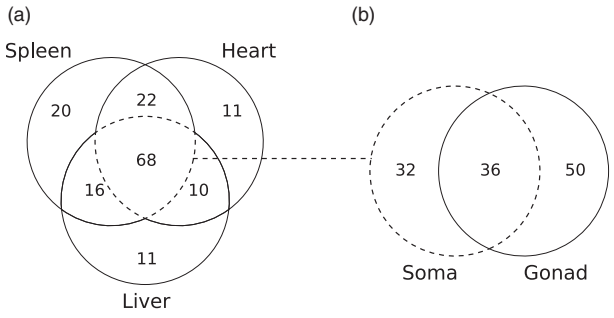

Fig. 3.—(*a*) Overlap between dosage-compensated ohnologs in the three somatic tissues. (*b*) Overlap between dosage-compensated genes in the soma (spleen, heart, and liver) and gonad tissue. Circles represent the total of dosage-compensated ohnologs in a tissue and numbers indicate the overlap between sets.

reports of incomplete dosage compensation in chicken (Ellegren et al. 2007; Itoh et al. 2007; Uebbing et al. 2015) and show that ohnologs are preferentially dosage-compensated on the chicken Z chromosome, indicating that incomplete dosage compensation can effectively balance dosage-sensitive genes. Even though the average expression of the Z chromosome is consistently lower in females as a function of incomplete dosage compensation, a considerable number of Z-linked genes show equal expression between males and females. Moreover, selection for compensation of dosage-sensitive genes appears to act relatively quickly, as there is no significant difference in the proportion of dosage-compensated ohnologs in younger regions of the avian Z chromosome compared with older regions.

The X chromosomal expression in mammals is reduced compared with the autosomes, potentially as a consequence of X inactivation (Xiong et al. 2010; Julien et al. 2012). It has been suggested that selection for the compensation of dosage-sensitive genes could have driven the evolution of X inactivation in therian mammals. Similarly, we also observe a reduction in Z expression in somatic tissues in males (Itoh et al. 2007). The reduced expression of the Z chromosome compared with the autosomes in males is not as pronounced as in females (fig. 1, supplementary table S2, Supplementary Material online) and there are several possible explanations for this pattern. The reduction has been suggested to result by partial *Z* inactivation that affects parts of the chromosome (Livernois et al. 2013; Graves 2014). However, our assessment of ASE suggests that inactivation is not a major mechanism affecting Z chromosome expression in males. An alternative explanation for the lower Z expression may be that the ancestral expression level of the Z chromosome, before the differentiation of the sex chromosomes, was already on average on the lower end of the expression spectrum (Brawand et al. 2011; Julien et al. 2012). Finally, it is possible that dosage sensitive genes have moved off the Z, as the mammalian X chromosome is depleted of genes requiring high transcription rates as a result of haploid expression in females (Hurst et al. 2015). Our analysis suggests that although there is some potential for movement of dosage-sensitive genes off the Z chromosome, the effect is confined to the oldest regions of the Z chromosome and is not substantial enough to explain the reduced expression in males.

It is important to keep in mind that the detection of ohnologs in vertebrate genomes remains challenging due to the age of the two rounds of WGD. All tools for the detection of ohnologs depend on the analysis of preserved gene order (synteny) among paralogs to distinguish single-gene duplicates from WGD. Large intra-genomic rearrangements may complicate these analyses, and may result in the underestimation of the number of ohnologs. Avian genomes, however, are relatively stable and compact, with fewer repeats and more coding DNA compared with other amniotes (Hillier et al. 2004; Ellegren 2005; Organ et al. 2007), suggesting that these issues are less prevalent. In addition, the detection of ohnologs depends on the selection of one or more outgroups that did not undergo a WGD to distinguish between genes that were duplicated before the WGD events. The outgroup selection can influence the number of ohnologs (Makino and McLysaght 2010) and the OhnologsDB mitigates that issue by using multiple outgroups.

Conclusion

Our results are consistent with gene-by-gene dosage compensation (Mank and Ellegren 2008; Mank 2013; Uebbing et al. 2013) and demonstrate that selection for dosage compensation of ohnologs does not necessitate the evolution of a global dosage compensation mechanism. This in turn leads to the interesting question why some organisms exhibit complex mechanisms of complete dosage compensation that require regulation of the entire X chromosome when such mechanisms are not necessarily evolutionarily required.

Methods

RNA-Seq Analysis and Gene Expression Estimates

We collected heart, liver, and spleen samples from White Leghorn chicken (*G. gallus*) embryonic day 19 eggs incubated under standard conditions. Embryos were sexed visually and based on expression of W-linked genes. For each tissue, four biological samples were collected for both males and females. One female liver sample was excluded from the analyses because it showed only spurious W expression and when investigating the *Z:A* ratio it was clearly masculinized. All samples were first stored in RNAlater (Qiagen) and then total RNA was extracted (Qiagen Animal Tissue RNA kit).

Library construction and Illumina sequencing was done at the Wellcome Trust Centre of Human Genetics (WTCHG), Oxford. Each sample was normalized to 2.5 μg total RNA

prior to a PolyA isolation using an NEB Magnetic mRNA Isolation Kit. PCR was carried out over 15 cycles using custom-indexed primers (WTCHG). Libraries were quality controlled with picogreen and tapestation, and were subsequently normalized equimolarly into 12-plex pools for Illumina HiSeq sequencing. Heart, liver, and spleen samples were sequenced using an Illumina HiSeq 2000 as paired-end 100-bp reads. 51-bp paired end reads of gonadal samples from the same development stage were obtained from Moghadam et al. (2012).

We trimmed each library using Trimmomatic v0.22 (Lohse et al. 2012) removing leading and trailing bases with a Phred score <3 and trimming using a sliding window approach when the average Phred score over four bases was <15. Reads were kept if they were at least 36 bases after trimming. Libraries were quality-inspected manually using FASTQC v0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The trimmed libraries were aligned against the chicken reference genome Ensembl version 75 Galgal4 (Cunningham et al. 2015) using TopHat v2.0.11 (Kim et al. 2013) and bowtie2 v2.2.2 (Langmead and Salzberg 2012) allowing five mismatches to the reference genome, with on average 17 million paired-end mappable reads per sample. Multi-mapping reads were removed and we then sorted and indexed the resulting alignment files for each library separately using Samtools v0.1.18/9 (Li et al. 2009).

We extracted reads mapping to annotated genes using HTseq-Count v0.6.1p1 (Anders et al. 2014) and normalized all tissues separately using the trimmed mean of M-values method available in edgeR v3.2.4 (Robinson et al. 2010). We estimated differential expression between males and females in all tissues using edgeR's exactTest method and exported the $\log_2$ fold change ($\log_2$FC; female–male expression), average $\log_2$ count per million (logCPM), FDR corrected *P*-values from the exactTest function and individual CPM values for all samples and genes. Genes were only included when the average CPM was >2 across all males and females, filtering out loci with low expression. When comparing groups of genes to each other, we normalized the CPM values by gene length, resulting in reads per kilobase of transcript per million mapped reads values (RPKM). Only genes annotated to the autosomes and the Z chromosome were assessed. Individual genes were defined as dosage-compensated on the Z chromosome if the female:male $\log_2$ fold change ranged from −0.5 to 0.5 (Wright et al. 2015). We defined genes as sex-biased if the edgeR exactTest was significant after FDR correction ($q < 0.05$) and the $\log_2$ fold change was >1 for female-biased genes or <−1 for male-biased genes.

Identification of Ohnologs and Other Paralogs

We used the Ohnologs database (http://ohnologs.curie.fr/) (Singh et al. 2015) to obtain ohnologs present in the chicken genome. We used the relaxed set of ohnologs as the primary dataset, in order to maximize the number of ohnologs. In addition, we used the Ensembl REST API (accessed February 2015) (Yates et al. 2015) to identify all paralogs in the chicken genome, which also includes those homologs originated in single-gene duplications.

Functional Annotation of Ohnologs

We used the G:profiler toolkit (Reimand et al. 2011) to perform GO Term (Ashburner et al. 2000) overrepresentation analyses. All ohnologs on the Z chromosome were provided as an input list and compared with the entire genomic background, using only genes with annotated GO terms in the comparison. Standard settings were used and GO Terms were only considered if they had a significant *P* value after multiple testing correction via G:Profiler's G:SCS method (*P* value <0.05). We additionally used the CORUM database (Ruepp et al. 2010), version from February 2012, to annotate protein complexes in the chicken genome. The CORUM database contains only mammalian data and we used the Ensembl REST API (Yates et al. 2015) to detect the corresponding chicken homologs, where possible.

SNP Calling and Estimation of ASE

In order to detect ASE from RNA-Seq data we modified a pipeline from Quinn et al. (2014). As we were interested in detecting ASE on the Z chromosome, we only called SNPs in the homogametic sex (males) for each tissue. SNPs were called using Samtools mpileup v0.1.18 (Li et al. 2009) and VarScan2 v2.3.6 (Koboldt et al. 2012). SNPs were called separately for each tissue using all four available male samples. We required minimum coverage of 2 and minimum Phred score of 20 (–min-avg-qual 20) to call an SNP and also required a minimum frequency of 0.9 to call a homozygote (–min-freq-for-hom 0.9). The resulting variant call formatted files were then filtered further to remove noise and increase SNP call confidence. In a first step, we filtered out SNPs using a combination of a fixed minimum threshold of 17 reads per site (the combination of major and minor allele) in all samples, as our power analysis indicates that a 17 read coverage for an SNP results in 73% power to detect allele specific-expression and also excluded all SNPs with more than two alleles. We additionally used a variable threshold that accounts for the likelihood of observing a second allele because of sequencing errors an error probability of 1 in 100 (Quinn et al. 2014) and a maximum coverage of 100,000. RNA-Seq data have an intrinsic bias for the estimation of ASE, because those reads that resemble the reference genome have a higher probability of aligning successfully. In order to remove this bias, we eliminated clusters of SNPs if there were >5 SNPs in a window of 100 bp (Stevenson et al. 2013). We used BEDtools intersect v2.20.1 (Quinlan and Hall 2010) to filter out all SNPs that were not located in a known transcript.

If both chromosomes are active to the same degree, we expect that the probability of observing reads from one or the other chromosome is 0.5. We therefore used a two-tailed binomial test to show significant deviations from this expected distribution ($P < 0.05$). Binomial tests were corrected for multiple testing on the autosomes, because of the larger number of testable sites. In order to account for the fact that binomial tests will be significant even for very small deviations in the observed distribution when the sample size, in our case the alignment depth, is big enough, we also employed a minimum threshold of 70% reads stemming from one allele to call significant ASE. In addition, we used a power analysis to ensure that our ability to detect ASE is sufficient. At a minimum coverage of 17 reads per site our power for detecting ASE is >73%, which suggests that we are able to detect patterns of ASE successfully in most cases. We only included genes in the analysis if at least one SNP showed consistent ASE across all samples.

All analyses and statistical comparisons were performed using Python, Matplotlib (Hunter 2007) and R (R Core Team 2015), code and iPython notebooks (Pérez and Granger 2007) are available on GitHub at https://github.com/qfma/ohnologdc. All sequencing data used in the analyses are available in the NCBI Short Read Archive under accession number SRP065394.

Supplementary Material

Supplementary tables S1–S7 and figures S1–S4 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

We thank Stephen Montgomery, Alison Wright, Natasha Bloch, Vicencio Oostra, Iulia Darolti, and Kevin Fowler for helpful comments on the manuscript. This study was supported by a European Research Council Framework 7 grant (agreement 260233) to J.E.M, University College London and a Swiss National Science Foundation grant to C.D. (PP00P3_150654). Sequencing was performed by the Wellcome Trust Centre for Human Genetics Sequencing Hub (funded by Wellcome Trust Grant reference 090532/Z09/Z and MRC Hub Grant G0900747 91070). The authors acknowledge the use of the UCL *Legion* High Performance Computing Facility (Legion@UCL) and associated support services, in the completion of this study.

Literature Cited

Albritton SE, et al. 2014. Sex-biased gene expression and evolution of the X chromosome in nematodes. Genetics doi: 10.1534/genetics.114.163311.

Anders S, Pyl PT, Huber W. 2014. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 31(2):166–169.

Ashburner M, et al. 2000. Gene Ontology: tool for the unification of biology. Nat Genet. 25:25–29.

Bachtrog D, et al. 2011. Are all sex chromosomes created equal? Trends Genet. 27:350–357.

Bachtrog D, et al. 2014. Sex determination: why so many ways of doing it? PLoS Biol. 12:e1001899.

Beukeboom L, Perrin N. 2014. The evolution of sex determination. Cambridge (MA): Oxford University Press.

Birchler JA, Bhadra U, Bhadra MP, Auger DL. 2001. Dosage-dependent gene regulation in multicellular eukaryotes: implications for dosage compensation, aneuploid syndromes, and quantitative traits. Dev Biol. 234:275–288.

Birchler JA, Riddle NC, Auger DL, Veitia RA. 2005. Dosage balance in gene regulation: biological implications. Trends Genet. 21:219–226.

Birchler JA, Veitia RA. 2012. Gene balance hypothesis: connecting issues of dosage sensitivity across biological disciplines. Proc Natl Acad Sci U S A. 109:14746–14753.

Blomme T, et al. 2006. The gain and loss of genes during 600 million years of vertebrate evolution. Genome Biol. 7:R43.

Brawand D, et al. 2011. The evolution of gene expression levels in mammalian organs. Nature 478:343–348.

Charlesworth B. 1978. Model for evolution of Y chromosomes and dosage compensation. Proc Natl Acad Sci U S A. 75:5618–5622.

Charlesworth B. 1996. The evolution of chromosomal sex determination and dosage compensation. Curr Biol. 6:149–162.

Charlesworth B. 1998. Sex chromosomes: evolving dosage compensation. Curr Biol. 8:R931–R933.

Charlesworth D, Charlesworth B, Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. Heredity 95:118–128.

Chen S, et al. 2014. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. Nat Genet. 46:253–260.

Conrad T, Akhtar A. 2012. Dosage compensation in *Drosophila melanogaster*: epigenetic fine-tuning of chromosome-wide transcription. Nat Rev Genet. 13:123–134.

Cooper DW, Johnston PG, Watson JM, Graves JAM. 1993. X-inactivation in marsupials and monotremes. Semin Dev Biol. 4:117–128.

Cui L, et al. 2006. Widespread genome duplications throughout the history of flowering plants. Genome Res. 16:738–749.

Cunningham F, et al. 2015. Ensembl 2015. Nucleic Acids Res. 43:D662–D669.

Deakin J, Chaumeil J, Hore T, Marshall Graves J. 2009. Unravelling the evolutionary origins of X chromosome inactivation in mammals: insights from marsupials and monotremes. Chromosome Res. 17:671–685.

Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. 3:e314.

Edger P, Pires JC. 2009. Gene and genome duplications: the impact of dosage-sensitivity on the fate of nuclear genes. Chromosome Res. 17:699–717.

Ellegren H. 2005. The avian genome uncovered. Trends Ecol Evol. 20:180–186.

Ellegren H, et al. 2007. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. BMC Biol. 5:40.

Emerson JJ, Kaessmann H, Betrán E, Long M. 2004. Extensive gene traffic on the mammalian X chromosome. Science 303:537–540.

Ercan S. 2015. Mechanisms of X chromosome dosage compensation. J Genomics 3:1–19.

Graves JM. 2014. Avian sex, sex chromosomes, and dosage compensation in the age of genomics. Chromosome Res. 22:45–57.

Hillier LW, et al. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695–716.

Hunter JD. 2007. Matplotlib: a 2D graphics environment. Comput Sci Eng. 9:90–95.
Hurst L, Tashakkori Ghanbarian A, Forrest ARR, Huminiecki L. 2015. The constrained maximal expression level owing to haploidy shapes gene content on the mammalian X chromosome. PLoS Biol. 13(12):e1002315.
Itoh Y, et al. 2007. Dosage compensation is less effective in birds than in mammals. J Biol. 6:2.
Julien P, et al. 2012. Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. PLoS Biol. 10:e1001328.
Kellis M, Birren BW, Lander ES. 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. Nature 428:617–624.
Kim D, et al. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14:R36.
Koboldt DC, et al. 2012. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. 22:568–576.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.
Leder EH, et al. 2010. Female-biased expression on the X chromosome as a key step in sex chromosome evolution in threespine sticklebacks. Mol Biol Evol. 27:1495–1503.
Li H, et al. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.
Lin F, Xing K, Zhang J, He X. 2012. Expression reduction in mammalian X chromosome evolution refutes Ohno's hypothesis of dosage compensation. Proc Natl Acad Sci U S A. 109:11752–11757.
Livernois AM, et al. 2013. Independent evolution of transcriptional inactivation on sex chromosomes in birds and mammals. PLoS Genet. 9:e1003635.
Lohse M, et al. 2012. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Res. 40:W622–W627.
Makino T, Hokamp K, McLysaght A. 2009. The complex relationship of gene duplication and essentiality. Trends Genet. 25:152–155.
Makino T, McLysaght A. 2010. Ohnologs in the human genome are dosage balanced and frequently associated with disease. Proc Natl Acad Sci U S A. 107:9270–9274.
Malone J, et al. 2012. Mediation of *Drosophila* autosomal dosage effects and compensation by network interactions. Genome Biol. 13:R28.
Mank JE. 2013. Sex chromosome dosage compensation: definitely not for everyone. Trends Genet. 29:677–683.
Mank JE, Ellegren H. 2008. All dosage compensation is local: gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. Heredity 102:312–320.
Mank JE, Hosken DJ, Wedell N. 2011. Some inconvenient truths about sex chromosome dosage compensation and the potential role of sexual conflict. Evolution 65:2133–2144.
Meyer BJ. 2010. Targeting X chromosomes for repression. Curr Opin Genet Dev. 20:179–189.
Moghadam HK, et al. 2012. W chromosome expression responds to female-specific selection. Proc Natl Acad Sci U S A. 109:8207–8211.
Naurin S, et al. 2011. The sex-biased brain: sexual dimorphism in gene expression in two species of songbirds. BMC Genomics 12:37.
Ohno S. 1967. Sex chromosomes and sex-linked genes. Berlin: Springer.
Organ CL, et al. 2007. Origin of avian genome size and structure in non-avian dinosaurs. Nature 446:180–184.
Papp B, Pál C, Hurst LD. 2003. Dosage sensitivity and the evolution of gene families in yeast. Nature 424:194–197.
Pérez F, Granger BE. 2007. IPython: a system for interactive scientific computing. Comput Sci Eng. 9:21–29.
Pessia E, Engelstädter J, Marais GAB. 2013. The evolution of X chromosome inactivation in mammals: the demise of Ohno's hypothesis? Cell Mol Life Sci. 71(8):1383–1394.
Pessia E, et al. 2012. Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. Proc Natl Acad Sci U S A. 109:5346–5351.
Pollack JR, et al. 2002. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proc Natl Acad Sci U S A. 99:12963–12968.
Potrzebowski L, et al. 2008. Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. PLoS Biol. 6:e80.
Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842.
Quinn A, Juneja P, Jiggins FM. 2014. Estimates of allele-specific expression in *Drosophila* with a single genome sequence and RNA-seq data. Bioinformatics 30:2603–2610.
R Core Team 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing.
Reimand J, Arak T, Vilo J. 2011. g:Profiler—a web server for functional interpretation of gene lists (2011 update). Nucleic Acids Res. 39:W307–W315.
Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.
Ruepp A, et al. 2010. CORUM: the comprehensive resource of mammalian protein complexes—2009. Nucleic Acids Res. 38:D497–D501.
Singh PP, Arora J, Isambert H. 2015. Identification of ohnolog genes originating from whole genome duplication in early vertebrates, based on synteny comparison across multiple genomes. PLoS Comput. Biol. 11:e1004394.
Stevenson K, Coolon J, Wittkopp P. 2013. Sources of bias in measures of allele-specific expression derived from RNA-seq data aligned to a single reference genome. BMC Genomics 14:536.
Straub T, Becker PB. 2007. Dosage compensation: the beginning and end of generalization. Nat Rev Genet. 8:47–57.
Torres EM, et al. 2007. Effects of aneuploidy on cellular physiology and cell division in haploid yeast. Science 317:916–924.
Uebbing S, et al. 2015. Quantitative mass spectrometry reveals partial translational regulation for dosage compensation in chicken. Mol Biol Evol. 32:2716–2725.
Uebbing S, Künstner A, Mäkinen H, Ellegren H. 2013. Transcriptome sequencing reveals the character of incomplete dosage compensation across multiple tissues in flycatchers. Genome Biol Evol. 5:1555–1566.
Van de Peer Y, Maere S, Meyer A. 2009. The evolutionary significance of ancient genome duplications. Nat Rev Genet. 10:725–732.
Veitia RA. 2004. Gene dosage balance in cellular pathways: implications for dominance and gene duplicability. Genetics 168:569–574.
Veitia RA, Veyrunes F, Bottani S, Birchler JA. 2015. X chromosome inactivation and active X upregulation in therian mammals: facts, questions, and hypotheses. J Mol Cell Biol. 7:2–11.
Vibranovski MD, Zhang Y, Long M. 2009. General gene movement off the X chromosome in the *Drosophila* genus. Genome Res. 19:897–903.
Vicoso B, et al. 2013. Comparative sex chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global dosage compensation. PLoS Biol. 11:e1001643.
Vicoso B, Bachtrog D. 2015. Numerous transitions of sex chromosomes in diptera. PLoS Biol. 13:e1002078.
White MA, Kitano J, Peichel CL. 2015. Purifying selection maintains dosage-sensitive genes during degeneration of the threespine stickleback Y chromosome. Mol Biol Evol. 32:1981–1995.
Wolf J, Bryk J. 2011. General lack of global dosage compensation in ZZ/ZW systems? Broadening the perspective with RNA-seq. BMC Genomics 12:91.

Wolfe K. 2000. Robustness—it's not where you think it is. Nat Genet. 25:3–4.
Wolfe KH. 2001. Yesterday's polyploids and the mystery of diploidization. Nat Rev Genet. 2:333–341.
Wolfe KH, Shields DC. 1997. Molecular evidence for an ancient duplication of the entire yeast genome. Nature 387:708–713.
Wright AE, Moghadam HK, Mank JE. 2012. Trade-off between selection for dosage compensation and masculinization on the avian Z chromosome. Genetics 192:1433–1445.
Wright AE, Zimmer F, Harrison PW, Mank JE. 2015. Conservation of regional variation in sex-specific sex chromosome regulation. Genetics 201:587–598.
Xiong Y, et al. 2010. RNA sequencing shows no dosage compensation of the active X-chromosome. Nat Genet. 42:1043–1047.
Yates A, et al. 2015. The ensembl REST API: ensembl data for any language. Bioinformatics 31:143–145.

Associate editor: Soojin Yi