

NEWS AND COMMENTARY

Flap about chicken MHM

Evaluating dosage compensation on the chicken Z chromosome: should effective dosage compensation eliminate sexual bias?

E Melamed, D Elashoff and AP Arnold

Heredity (2009) 103, 357–359; doi:10.1038/hdy.2009.91; published online 12 August 2009

A variety of species with XX female/XY male sex chromosomes have independently evolved effective mechanisms of sex chromosome dosage compensation (DC). This suggests that a large-scale sexual imbalance in gene expression, resulting from the sex difference in copy number of X genes, is typically not tolerated during evolution (Meyer *et al.*, 2004). The imbalance has frequently and perhaps inevitably led to sex-specific selection pressures to reduce the sex difference in gene expression. It was surprising, therefore, when the lack of a chromosome-wide mechanism of DC was recently reported in the ZZ male/ZW female system of birds (Itoh *et al.*, 2007), where Z gene expression is constitutively higher in males. This finding has now been confirmed in a second study in chickens (Ellegren *et al.*, 2007) and detected in another ZZ/ZW system, the silkworm *Bombyx mori* (Zha *et al.*, 2009).

Although birds lack a global mechanism of DC, we investigated if they possess any regional differences in DC on the Z chromosome (Melamed and Arnold, 2007). Earlier studies in chickens suggested that the MHM (male hypermethylated) locus on the Z chromosome might be involved in DC (Bisoni *et al.*, 2005) because the MHM noncoding RNA is expressed only in females, is associated with the MHM locus and is tightly spatially correlated with acetylation of histone 4 lysine 16 (H4K16), a histone mark correlated with increased gene expression (Teranishi *et al.*, 2001; Bisoni *et al.*, 2005). Thus, acetylation of H4K16 could lead to upregulation of gene expression in hemizygous females in the region of MHM, in a manner similar to H4K16 acetylation in male hemizygous *Drosophila* where it is part of the male-specific upregulation to achieve DC (Smith *et al.*, 2000). We reported that the MHM region on chromosome arm Zp has an unusually low incidence of genes with

high M/F ratios so that this region shows better than average compensation of gene expression in three embryonic tissues, and that Zq has an unusual number of genes with high (uncompensated) M/F ratios (Melamed and Arnold, 2007). These results supported the idea that the MHM ncRNA is involved in regional DC of Zp near MHM.

A recent study by Mank and Ellegren (2009) did not confirm that conclusion. The authors considered three tissues in chickens at several ages, but used different data analysis and statistical methods compared to our study. They confirmed the presence of an uncompensated neighborhood of genes on Zq, but did not detect a statistically significant region of DC on Zp. The authors concluded that no specialized region of DC occurs on Zp. We have attempted to reconcile the differences in the two studies by reanalyzing the data produced by both studies using methods similar to those of both groups.

A major issue is that the two studies differ in their view of the process of DC, and therefore, in the choice of parameters used to assess the degree of DC. Because chromosome-wide DC results from selection pressures that reduce M/F ratios (either by increasing expression in females or reducing expression in males, or both), we judged that genes with low (female-biased) M/F ratios were compensated. In contrast, Mank and Ellegren (2009) state that if a neighborhood of DC exists on the Z chromosome, it would have little or no sex bias (sexual inequality) of gene expression. They therefore assessed sex bias using an 'amplitude' metric (the absolute value of the \log_2 M/F ratio of expression), which measures the deviation of M/F ratios from sexual equality irrespective of the direction of the deviation. In their view, DC exists only when the amplitude metric approaches zero (that is, when the M/F ratio equals 1).

Thus, genes with M/F ratios well below 1 are considered uncompensated. In our view the amplitude metric is inappropriate, because even in XX/XY systems with exceptionally good DC sexual bias is not eliminated and the amplitude metric is not zero (Figures 1a and b). Rather, effective DC leads to matching of the distribution of M/F ratios (the amount of sexual bias) on the X chromosome and autosomes in a tissue-specific manner (Figures 1a and b), possibly because of network interactions between X and autosomal (A) genes (Itoh *et al.*, 2007). Moreover, the selection pressure toward DC of the Z chromosomes is directional because it stems from a directional inequality of Z genomic dose ($M > F$). That inequality will lead only to selection pressure favoring mutations that reduce M/F ratios, rather than increase M/F ratios for genes with ratios below 1. The amplitude metric assumes that genes with low M/F ratios below 1 are not compensated. Our view is that genes with low M/F ratios are compensated with respect to the inequality of Z chromosome number, but that other gene-specific forces act to reduce the ratio below 1.

The two studies also differed in the hypotheses tested. Previous studies suggested that the MHM region will show greater DC than other regions of the Z chromosome ('MHM hypothesis'). On the basis of that hypothesis, we evaluated whether a region of low M/F ratios occurs near MHM. We found that the region of low M/F ratios was unlikely to occur by chance anywhere on the Z chromosome, and was especially unlikely to occur near MHM as predicted. In contrast, Mank and Ellegren did not evaluate the MHM hypothesis.

In addition, the two studies differed in methods of analysis, including algorithms for data filtering and normalization, grouping probes that measure the same gene and statistical tests to find regionally significant DC. To assess if these factors might have led to different conclusions, we reanalyzed data sets from both studies using algorithms that mimicked, as closely as possible, those originally used in each study. We were unable to replicate the probe aggregation methods of Mank and Ellegren, because they were not fully described. Therefore we used here the probe aggregation from Melamed and Arnold (2007). For similar reasons, we were unable to replicate the statistical tests used by Mank and Ellegren, and present new statistical methods that augment those used previously (Melamed and Arnold, 2007).

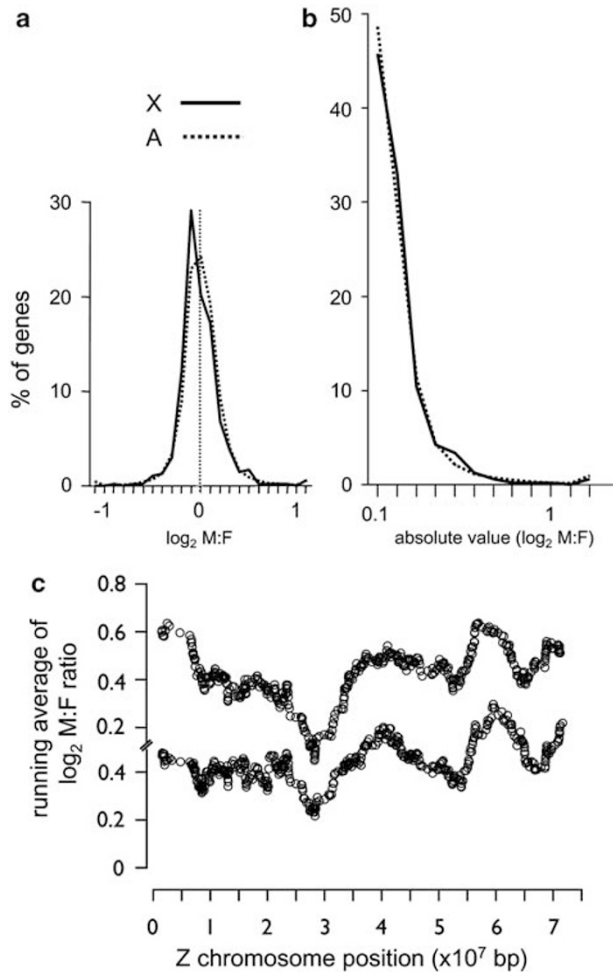


Figure 1 (a) The distribution of \log_2 M/F ratios of gene expression for autosomal (A) and X genes in mouse liver (reprinted from Figure 2a of Itoh *et al.*, 2007). Effective dosage compensation of the X chromosome leads to close matching of the distribution of the degree of sex bias in X and A expression, but does not eliminate sex bias. The distribution (b) of the ‘amplitude’ of sex bias (= absolute values of \log_2 M/F ratios from (a), the metric used by Mank and Ellegren (2009)) illustrates that effective dosage compensation does not reduce the amplitude to zero. The mean absolute value of M/F ratio was 0.15–0.16 for X and A genes. (c) The running average of \log_2 M/F ratios of gene expression in chick embryonic brain, plotted as a function of position on the Z chromosome. The mean is calculated in a window of 30 genes, shifted 1 gene at a time along the chromosome. The upper curve shows data from Mank and Ellegren (2009), and lower curve data from Melamed and Arnold (2007), DChip normalized with 60% present call. The upper part of the split ordinate refers to the upper curve, and the lower part to the lower curve. The two data sets give very similar curves, and the dip in gene expression near 2.8×10^7 bp is statistically significant in each case.

We analyzed comparable chick embryonic brain data from each study (day 14 from our study, day 18 from Mank and Ellegren (2009)). We normalized both data sets using DChip software as in our previous study, as well as the RMA method in the R package Bioconductor (R Development Core Team, 2009; Gentleman *et al.*, 2004), approximating the methods of Mank and Ellegren. We filtered both data sets to eliminate probe sets that had fewer than 50 or 60% ‘present’ call. To decrease gene redundancy, we averaged expression values of Affymetrix probes mapping to the same GeneEntrez ID. Finally, because Mank and Ellegren

attributed the dip in M/F ratios near MHM to two genes with especially low M/F ratios, we conservatively excluded genes in the region between 25 and 35 Mb that had a \log_2 M/F < 0.4. DChip normalization generally resulted in a higher number of aggregated probes compared to RMA normalization (range 392–463 Z genes). Mank and Ellegren reported 635 Z genes, probably because of different probe aggregation methods.

We then performed statistical tests on the two data sets to evaluate whether there is a statistically significant dip in M/F ratios near MHM (Figure 1c). On the basis of our visual assessment of both brain data sets, the region between 2.7

and 3.1 Mb consistently appeared to show a lack of genes with high M/F ratios regardless of methods of normalization and filtration or logging status. To compare the region at 2.7–3.1 Mb to the rest of the Z chromosome, we calculated the *t*-test *P*-value (called ‘empirical *P*-value’) of M/F ratios in this region compared to the rest of Z chromosome. Next, to obtain an appropriate reference distribution for this test statistic that takes into account the fact that a specific region was chosen for statistical testing, we randomly permuted the chromosome locations of M/F ratios on the Z chromosome 1000 times. For each permutation cycle, we recorded the lowest *P*-value and compared it to the empirical *P*-value. In both data sets, the lowest *P*-value in the 1000 permutations was rarely lower than the empirical ($P < 0.003$). Therefore, we judge that the dip in M/F ratios is unlikely to occur by chance anywhere along the Z chromosome. We also performed a second statistical test, evaluating runs of 30 genes at a time with a step of 1 gene, permuting the data in the same way as described above. This test resulted in the same conclusion, that the difference actually observed between M/F ratios near MHM (at 2.7–3.1 Mb) compared to the rest of the Z chromosome is unlikely to occur by chance ($P < 0.001$) anywhere along the Z chromosome. In general, the data obtained in the two studies were highly comparable no matter which analysis method was applied (Figure 1c).

We reaffirm our earlier conclusion that the Zp contains a region near MHM with low M/F ratios that are not random. The dip in M/F ratios is not caused by the occurrence of a few genes with low M/F ratios, but is due to the relative lack of genes with high M/F ratios (Melamed and Arnold, 2007). Because M/F ratios are potentially influenced by both regional mechanisms of DC and gene-specific mechanisms, it is not possible to conclude yet whether the MHM ncRNA is involved in regional compensation, or whether the MHM region has low M/F ratios for some other reason. Future molecular studies in birds will be necessary to resolve this interesting issue.

Drs E Melamed and AP Arnold are at the Department of Physiological Science and Laboratory of Neuroendocrinology of the Brain Research Institute, UCLA, Los Angeles, CA, USA and Dr D Elashoff is at the Department of Biostatistics, School of Public Health, UCLA, Los Angeles, CA, USA.

e-mail: emelamed@ucla.edu

- Bisoni L, Batlle-Morera L, Bird AP, Suzuki M, McQueen HA (2005). Female-specific hyperacetylation of histone H4 in the chicken Z chromosome. *Chromosome Res* 13: 205–214.
- Ellegren H, Hultin-Rosenberg L, Brunström B, Dencker L, Kultima K, Scholz B (2007). Faced with inequality: chicken do not have a general DC of sex-linked genes. *BMC Biol* 5: 40.
- Gentleman R, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S *et al.* (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80.
- Itoh Y, Melamed E, Yang X, Kampf K, Wang S, Yehya N *et al.* (2007). DC is less effective in birds than in mammals. *J Biol* 6: 2.
- Mank JE, Ellegren H (2009). All DC is local: gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. *Heredity* 102: 312–320.
- Melamed E, Arnold AP (2007). Regional differences in DC on the chicken Z chromosome. *Genome Biol* 8: R202.
- Meyer BJ, McDonel P, Csankovszki G, Ralston E (2004). Sex and X-chromosome-wide repression in *Caenorhabditis elegans*. *Cold Spring Harb Symp Quant Biol* 69: 71–79.
- R Development Core Team (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Smith ER, Pannuti A, Gu W, Steurnagel A, Cook RG, Allis CD *et al.* (2000). The drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. *Mol Cell Biol* 20: 312–318.
- Teranishi M, Shimada Y, Hori T, Nakabayashi O, Kikuchi T, Macleod T *et al.* (2001). Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. *Chromosome Res* 9: 147–165.
- Zha X, Xia Q, Duan J, Wang C, He N, Xiang Z (2009). Dosage analysis of Z chromosome genes using microarray in silkworm, *Bombyx mori*. *Insect Biochem Mol Biol* 39: 315–321.