Comment on "*Drosophila* Dosage Compensation Involves Enhanced Pol II Recruitment to Male X-Linked Promoters"

F. Ferrari,[1]* Y. L. Jung,[1,2]* P. V. Kharchenko,[1,3] A. Plachetka,[2,4] A. A. Alekseyenko,[2,4] M. I. Kuroda,[2,4]† P. J. Park[1,2]†

Conrad *et al*. (Reports, 10 August 2012, p. 742) reported a doubling of RNA polymerase II (Pol II) occupancy at X-linked promoters to support 5′ recruitment as the key mechanism for dosage compensation in *Drosophila*. However, they employed an erroneous data-processing step, overestimating Pol II differences. Reanalysis of the data fails to support the authors' model for dosage compensation.

In *Drosophila melanogaster,* dosage compensation is achieved by a ~2-fold up-regulation of genes on the male X chromosome (*1*, *2*). The male-specific lethal (MSL) complex is critical for dosage compensation (*3*), but the specific mechanisms and stage of the transcriptional process at which it acts are less clear. Global run-on sequencing (GRO-seq) analyses suggested that facilitated transcriptional elongation plays an important role (*4*). The Report by Conrad *et al*. (*5*) appeared to refute this, based on RNA polymerase II (Pol II) occupancy [chromatin immunoprecipitation sequencing (ChIP-seq)] in male and female salivary glands. A 2-fold increase was found in Pol II binding starting at promoters and continuing throughout X-linked genes in males compared with females, or male X versus autosomes. Thus, an entirely initiation-based model was suggested, in which differential elongation plays no role.

We noticed that the analysis by Conrad *et al*. reported unusually high Pol II ChIP enrichment levels. The average enrichment at the promoters of bound genes was reported to be ~30,000-fold over input (~15 on a log2 scale), orders of magnitude higher than what is typical of robust ChIP-seq experiments. We reexamined the data of Conrad *et al*. using an independent analysis pipeline (*6*), finding an average promoter enrichment ratio of ~8-fold rather than ~30,000. More importantly, we found that the difference in the Pol II occupancy between male X-linked and autosomal genes was ~1.2, which is substantially lower than the ~2-fold increase reported by Conrad *et al*. The reduced ratio does not support an initiation-based model but rather is inconclusive regarding the steps in the transcription cycle necessary to achieve a 2-fold increase in male X-linked transcription.

This major discrepancy can be traced to an erroneous normalization step in the analysis of Conrad *et al*. (*5*) that resulted in overestimation of the Pol II ratios. We repeated the analysis using the same gene lists, annotations, parameters, and code that were provided to us by the authors as well as their previously published pipeline (*7*). We found that both discrepancies described above can be traced to a single normalization step in their ChIP-seq data processing. Their pipeline calculated log2 ratios of the normalized read counts (IP over input) in 25-bp (base pair) bins and used a moving average with a 400-bp window to smooth the resulting profiles. The empty bins and nonfinite log2 ratios were removed by the earlier preprocessing steps to avoid introducing bias in the moving average estimates. The moving averages were then multiplied by the square root of the number of remaining bins within each moving-average window (*7*). However, this multiplication was mistakenly performed directly on the log2-transformed ratios, resulting in exponential amplification of the IP over input (Pol II enrichment) ratios and distorting the comparisons subsequently performed (*5*).

Omitting this multiplication step reduces the average Pol II enrichment levels by three orders of magnitude, bringing them into the 2- to 20-fold enrichment range typical of most ChIP-seq studies (Fig. 1, A versus B, upper panels). The resulting Pol II ratios for male X-linked genes compared to autosomal genes are around ~1.2-fold (Fig. 1, A versus B, lower panels), as are Pol II ratios for X-linked genes in males versus females (Fig. 1, C versus D). The RNA interference (RNAi) experiments were affected in a similar way (Fig. 2, A and B). The ~1.2-fold ratio is consistent with what we obtained from the independent analysis pipeline as described above (Fig. 2C). The difference between males and females based on a rank-based test is statistically significant in both cases, but the modest ratio of ~1.2 fails to support a doubling of RNA pol II occupancy at male X-linked promoters, which is the central evidence for the initiation model.

In a corroborating line of evidence, Conrad *et al*. examined short RNAs for a panel of genes by quantitative polymerase chain reaction, showing a 2-fold increase at X-linked genes in males compared with females. However, genome-wide data from the Adelman laboratory (*8*) show that X and autosomal genes exhibit very similar levels of paused RNA, with no evidence of a doubling, on average, at male X-linked genes (Fig. 2D). Finally, the reanalysis of our GRO-seq data (*5*) is also problematic, because the index for elongation effect they employed ignores the last 500 bp of each gene, in addition to the usual first 500 bp. This is highly relevant, as genes as short as 1.1 Kb were considered and the effect is greatest toward the 3′ ends. Thus, the overall evidence in Conrad *et al*. is insufficient to support the authors' model for dosage compensation.

[1]Center for Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA. [2]Division of Genetics, Brigham and Women's Hospital, Boston, MA 02115, USA. [3]Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA 02115, USA. [4]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA.

*These authors contributed equally to this work.
†Corresponding author. E-mail: mkuroda@genetics.med.harvard.edu (M.I.K.); peter_park@harvard.edu (P.J.P.)

References and Notes

1. M. E. Gelbart, M. I. Kuroda, *Development* **136**, 1399 (2009).
2. T. Conrad, A. Akhtar, *Nat. Rev. Genet.* **13**, 123 (2012).
3. A. Hilfiker, D. Hilfiker-Kleiner, A. Pannuti, J. C. Lucchesi, *EMBO J.* **16**, 2054 (1997).
4. E. Larschan *et al*., *Nature* **471**, 115 (2011).
5. T. Conrad, F. M. G. Cavalli, J. M. Vaquerizas, N. M. Luscombe, A. Akhtar, *Science* **337**, 742 (2012).
6. P. V. Kharchenko, M. Y. Tolstorukov, P. J. Park, *Nat. Biotechnol.* **26**, 1351 (2008).
7. R. Ali, F. M. G. Cavalli, J. M. Vaquerizas, N. M. Luscombe, A pipeline for ChIP-seq data analysis (Prot 56) (2012); www.epigenesys.eu/index.php/en/protcols/bio-informatics/411-a-pipeline-for-chip-seq-data-analysis.
8. S. Nechaev *et al*., *Science* **327**, 335 (2010).

Acknowledgments: This work was supported by a Sloan Research Fellowship to P.J.P. and NIH grant GM45744 to M.I.K.

22 October 2012; accepted 27 February 2013
10.1126/science.1231815

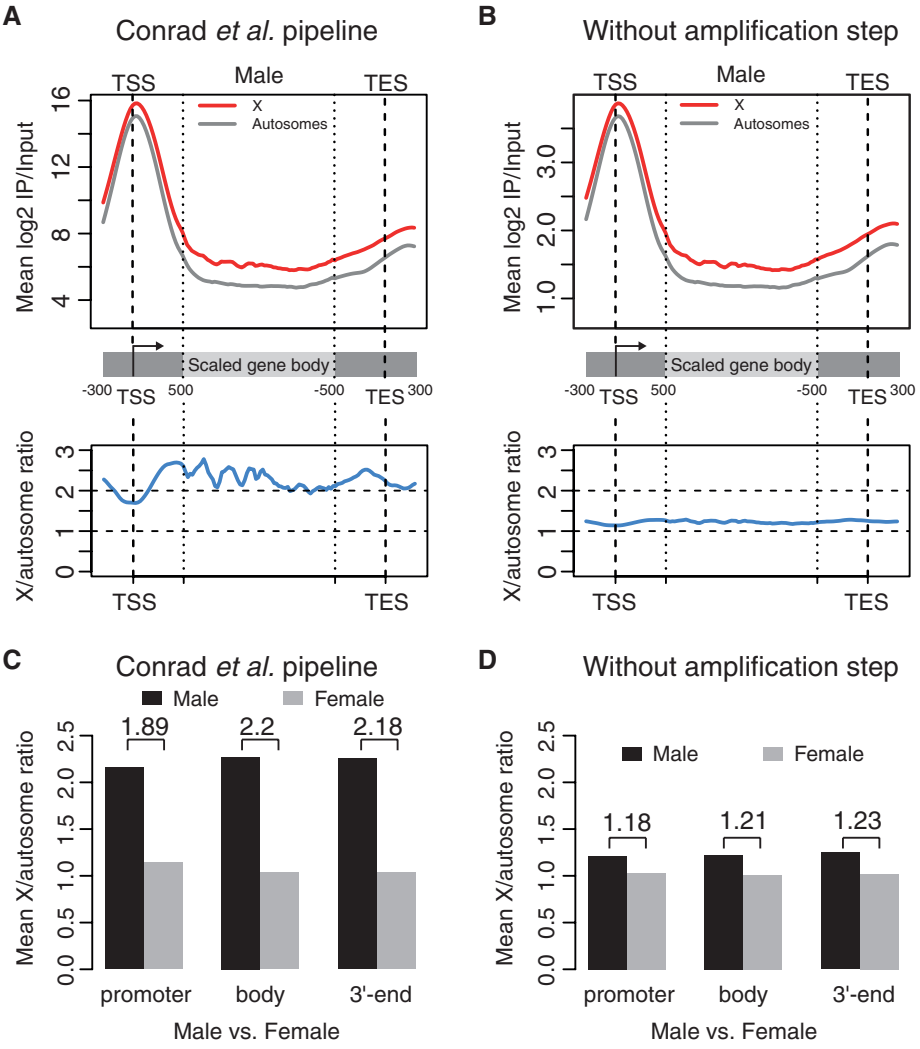

Fig. 1. Effect of the erroneous multiplication step on the ratio of Pol II binding on X-linked genes compared with autosomal genes. (**A**) We have reproduced the analysis pipeline in (*5*). The upper plot reports the average Pol II ChIP over input log2 enrichment ratio along X-linked (red) and autosomal (gray) genes in male salivary glands. The lower plot reports the ratio between X and autosomes. The same annotations and gene lists from figure 2 by Conrad *et al*. (*5*) are used here. (**B**) Same data as in (A), but processed without the erroneous multiplication step. (**C**) Average Pol II binding ratios between X and autosomal genes in male (black) and female (gray) salivary glands, as measured at the promoter, gene body, and 3′-end regions. Male over female ratios are also indicated. The data were processed using the ChIP-seq processing pipeline in (*5*). (**D**) Same as (C), but processed without the erroneous amplification step.

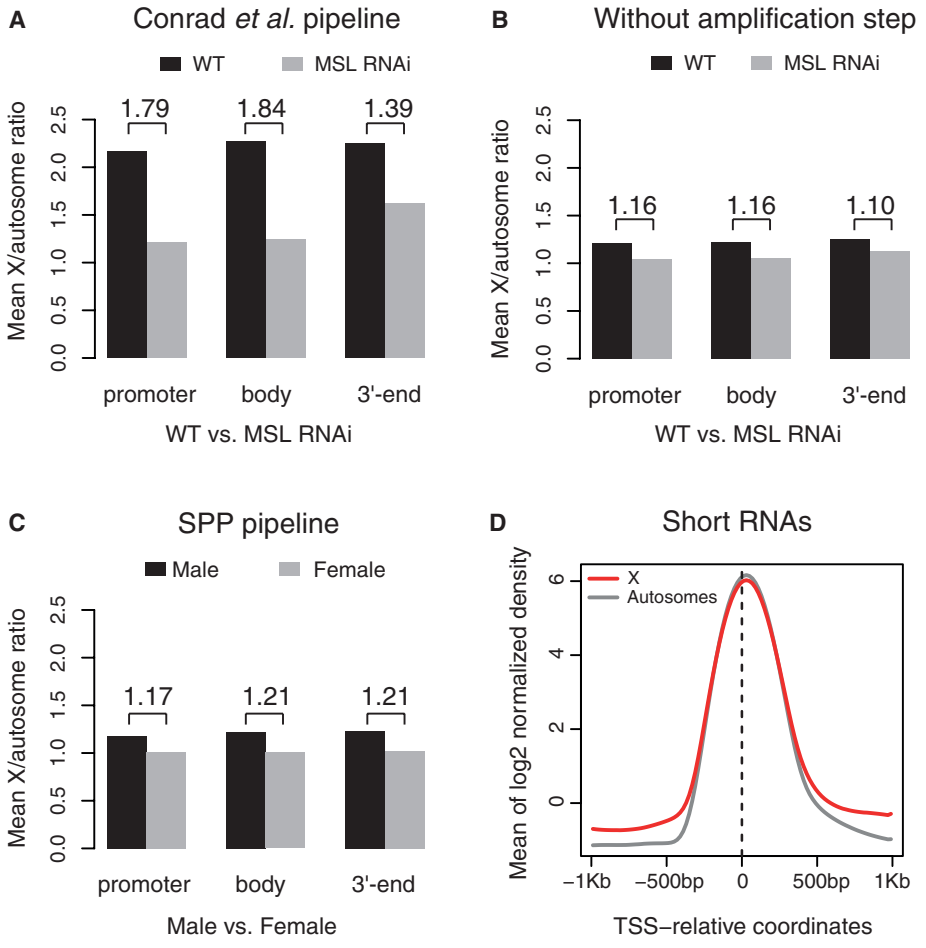

Fig. 2. Additional analyses of ChIP-seq and paused RNAs fail to support the initiation model for dosage compensation. (**A**) Average Pol II binding ratios between X and autosomal genes in untreated wild-type (black) and MSL2 RNAi–treated (gray) male cells, using the same annotations and gene lists as in figure 4 by Conrad *et al*. (*5*). (**B**) The same as in (A), but processed without the erroneous amplification step. (**C**) Average Pol II binding ratios between chromosome X and autosomal genes in male (black) and female (gray) salivary glands, analyzed with an alternative pipeline. Normalized ChIP over input enrichment was computed using the maximum likelihood estimate (MLE) analysis implemented in the SPP software package (*6*). The promoter, gene body and 3′-end regions are analyzed. The numbers on top specify the ratio of male over female values. (**D**) Short RNAs associated with paused Pol II from (*8*) show similar density per gene copy around the transcription start sites (TSS) on active X-linked and autosomal genes in males. The same annotations and gene lists from figure 2 by Conrad *et al*. (*5*) were used; data were processed with SPP (*6*); normalization over genomic control was used to account for differences in copy number.

Comment on "*Drosophila* Dosage Compensation Involves Enhanced Pol II Recruitment to Male X-Linked Promoters"

F. Ferrari, Y. L. Jung, P. V. Kharchenko, A. Plachetka, A. A. Alekseyenko, M. I. Kuroda and P. J. Park

Science **340** (6130), 273.
DOI: 10.1126/science.1231815