Cell Reports
Report

"Jump Start and Gain" Model for Dosage Compensation in *Drosophila* Based on Direct Sequencing of Nascent Transcripts

Francesco Ferrari,[1,6] Annette Plachetka,[2,3,6] Artyom A. Alekseyenko,[2,3,6] Youngsook L. Jung,[1,2] Fatih Ozsolak,[4] Peter V. Kharchenko,[1,5] Peter J. Park,[1,2,*] and Mitzi I. Kuroda[2,3,*]
[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]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA
[4]Helicos BioSciences Corporation, One Kendall Square, Cambridge, MA 02139, USA
[5]Division of Hematology/Oncology, Boston Children's Hospital, Boston, MA 02115, USA
[6]These authors contributed equally to this work
*Correspondence: peter_park@hms.harvard.edu (P.J.P.), mkuroda@genetics.med.harvard.edu (M.I.K.)
http://dx.doi.org/10.1016/j.celrep.2013.09.037

SUMMARY

Dosage compensation in *Drosophila* is mediated by the MSL complex, which increases male X-linked gene expression approximately 2-fold. The MSL complex preferentially binds the bodies of active genes on the male X, depositing H4K16ac with a 3′ bias. Two models have been proposed for the influence of the MSL complex on transcription: one based on promoter recruitment of RNA polymerase II (Pol II), and a second featuring enhanced transcriptional elongation. Here, we utilize nascent RNA sequencing to document dosage compensation during transcriptional elongation. We also compare X and autosomes from published data on paused and elongating polymerase in order to assess the role of Pol II recruitment. Our results support a model for differentially regulated elongation, starting with release from 5′ pausing and increasing through X-linked gene bodies. Our results highlight facilitated transcriptional elongation as a key mechanism for the coordinated regulation of a diverse set of genes.

INTRODUCTION

In *Drosophila melanogaster,* expression from the single male X chromosome (XY) is upregulated about 2-fold to match transcription of the two Xs in females (XX). This process, termed dosage compensation (DC), makes X-linked transcription equivalent between the sexes and also balances X and autosomal transcription (Gupta et al., 2006; reviewed in Lucchesi et al., 2005; Gelbart and Kuroda, 2009). Upregulation of the male X is mediated by the MSL complex, which consists of at least five protein subunits (MSL1, MSL2, MSL3, MLE, and MOF) and two noncoding RNAs (roX1 and 2) (reviewed in Gelbart and Kuroda, 2009). MOF has histone acetyltransferase activity and modifies histone 4 at lysine 16 (H4K16), enriching this mark along the male X (Gu et al., 1998; Smith et al., 2000). H4K16ac is thought to play a key role in the upregulation of genes on the male X because it enhances transcription in vitro and in vivo (Akhtar and Becker, 2000; Smith et al., 2001; Dou et al., 2005).

The steps in transcription that are targeted by the DC mechanism have been controversial. The major mechanistic steps in transcription are listed in Figure 1A. In brief, recruitment of RNA polymerase II (Pol II) to promoters by general transcription factors is followed by conversion to an initiating complex via TFIIB stimulation of RNA synthesis (Sainsbury et al., 2013). Once the RNA grows to 12–13 nucleotides, it triggers TFIIB displacement and elongation complex formation. However, in *Drosophila* and mammals elongating Pol II subsequently pauses on a majority of transcribed genes, with a prominent Pol II peak seen around +50 relative to the TSS (reviewed in Gilmour, 2009). Surprisingly, many highly expressed genes in *Drosophila* display pausing, suggesting that it could be an obligate step during active transcription, perhaps to allow time for association of key elongation and splicing factors. Release of paused Pol II is mediated by phosphorylation of NELF, DSIF, and serine 2 in the C-terminal domain (CTD) of Pol II by the kinase P-TEFb (Peterlin and Price, 2006). DSIF is composed of subunits SPT4 and SPT5 that act positively in subsequent elongation by closing the Pol II active cleft to render the elongation complex stable and processive (Martinez-Rucobo et al., 2011). Transcription termination involves coordination between the exit of elongation factors and recruitment of cleavage and termination factors (Mayer et al., 2010, 2012). There is no single method to simultaneously analyze regulation of each of these steps of transcription in vivo. However, recent advances allow a composite picture of regulation genome-wide. In Figure 1A, we list these methods (described below) and their abilities to distinguish between steps in the transcription cycle.

Two competing models have been proposed for how MSL complex can coordinately control hundreds of functionally

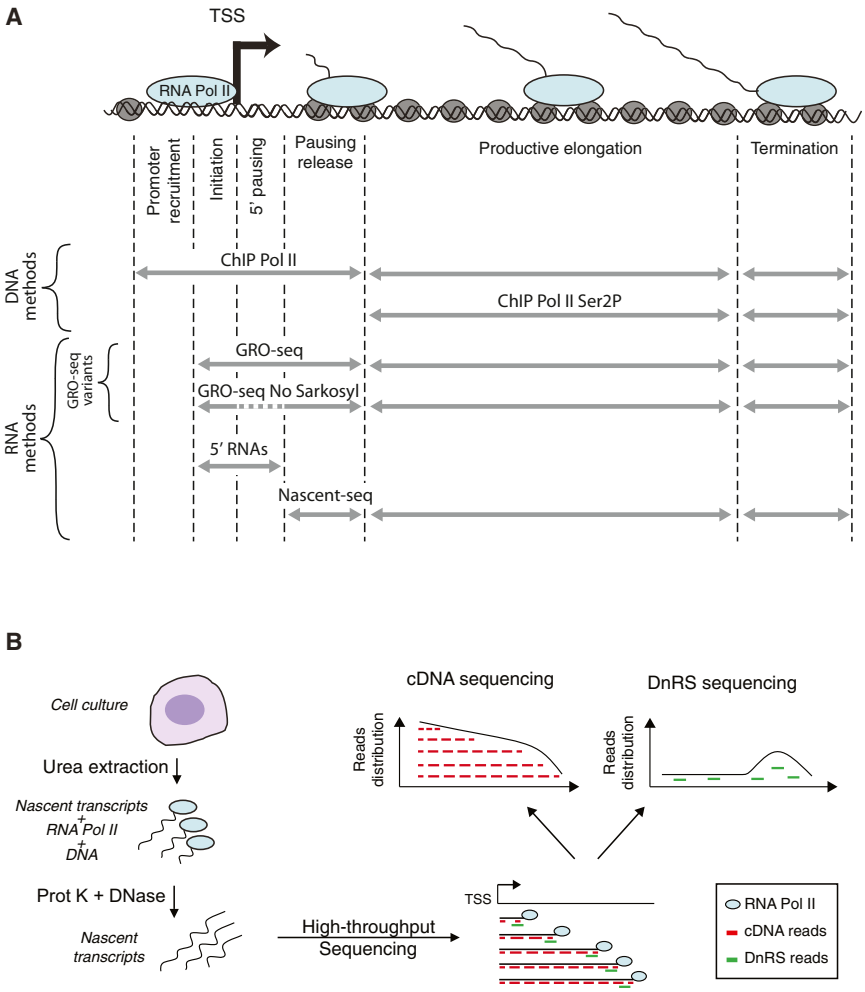

Figure 1. Transcription Analysis

(A) Summary of genome-wide methods for studying transcription is shown. Each method can provide information about specific stages of transcription (horizontal arrows). In some stages, and especially in the initial phases (on the left), the available methods do not allow discrimination between different steps of transcription. These resolution limitations are depicted as arrows spanning boundaries between multiple stages. The dashed arrow in "GRO-seq No Sarkosyl" indicates the lack of data for paused Pol II. ChIP Pol II, chromatin immunoprecipitation targeting Pol II followed by either sequencing (ChIP-seq) or microarray hybridizations (ChIP-chip); Ser2P, serine 2 phosphorylation marking elongating Pol II; GRO-seq, global run-on sequencing with sarkosyl treatment to detect paused and active Pol II or without sarkosyl to detect active Pol II only; 5′ RNAs, sequencing of short 5′ capped nuclear transcripts to measure 5′ pausing; Nascent-seq, sequencing of nascent transcripts.

(B) A schematic representation of the cDNA-based and direct nascent RNA sequencing (DnRS) protocols is shown. Nascent transcripts were isolated using the same protocol, followed by two different sequencing approaches. In cDNA sequencing, strand-specific cDNA is synthesized by random priming of the entire nascent transcript, with a resulting decrease in read density from the 5′ end to 3′ end (red). In contrast, the 3′ end of the nascent transcript is sequenced directly in DnRS, thus providing a precise map of Pol II position (green).

See also Figure S1.

unrelated genes: one focused on differential promoter recruitment of Pol II (Conrad et al., 2012), and a second featured enhanced transcriptional elongation (Larschan et al., 2011; Prabhakaran and Kelley, 2012). To distinguish between them, we combine three high-resolution methods: 5′ paused RNA sequencing, GRO-seq, and Nascent-seq. For the recruitment model, each should demonstrate a similar fold increase in X versus autosomal signal along genes, indicating that DC is fully implemented at the recruitment step. In contrast, the elongation model predicts that a differential effect may be seen from 5′ to 3′. Early steps such as 5′ pausing may not be rate limiting and may not exhibit compensation (e.g., X ≈A), whereas later steps may show increasing compensation (e.g., X ≈2A). Thus, Pol II behavior on X and autosomal genes in male cells should allow us to discriminate between the two models.

RESULTS AND DISCUSSION

Male S2 and female Kc cells are robust models for the study of DC in *Drosophila*. S2 cells express the MSL complex, which binds to the bodies of X-linked genes (Alekseyenko et al., 2006; Gilfillan et al., 2006) to increase their transcription (Hamada et al., 2005; Straub et al., 2005). If MSL2 is induced in female Kc cells, it also increases X transcription relative to autosomes (Alekseyenko et al., 2012). However, female cells normally do not express MSL2, and therefore levels of X transcription are on average similar to autosomes (Conrad et al., 2012, and see below). Thus, the key measurement for DC is the differential behavior of the X and autosomes within male cells.

The choice of method(s) is crucial (Figure 1A). Genome-wide ChIP analysis of Pol II is commonly used to study its density along genes; however, this approach is strongly dependent on antibody quality and has limited resolution owing to fragmented chromatin size. In addition, ChIP-seq does not allow a precise distinction between different activity states of Pol II, although antibodies for phosphorylated forms of the CTD can help distinguish between elongating or paused isoforms (Weeks et al., 1993; Buratowski, 2009). In contrast, GRO-seq (Core et al., 2008) is independent of antibody specificity and maps the position and density of paused and elongating Pol II along genes in a quantitative manner. A strength of this method is its ability to measure transcriptionally engaged but paused Pol II at 5′ ends of genes. An independent assay for 5′ pausing involves the isolation and sequencing of short 5′ capped nuclear transcripts (Nechaev et al., 2010). Nascent-seq (Khodor et al., 2011) isolates Pol

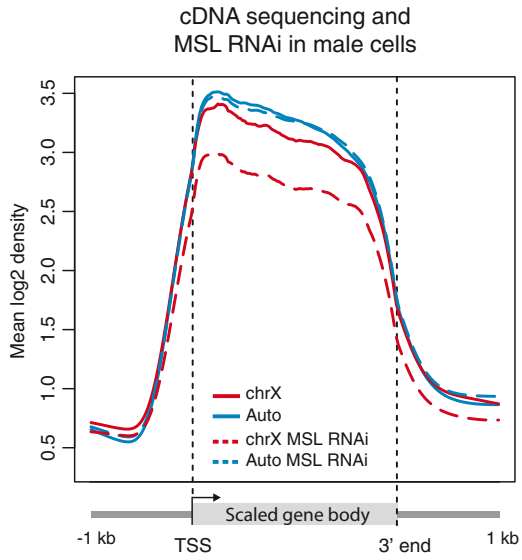

Figure 2. Nascent-Seq cDNA Analysis of Transcription in Male S2 Cells after Control or MSL RNAi

Average cDNA sequencing read-density profiles of actively transcribed X (red; n = 379) and autosomal (blue; n = 1,538) genes in control RNAi (solid lines) and MSL RNAi (dashed) samples in male (S2) cells are shown. The metagene profile is generated by rescaling the gene body to a fixed width between transcription start site (TSS) and transcript end (3′ end), excluding introns. Because intron removal reduces the size of the represented region, the whole gene body is scaled; if unscaled, the 5′ and 3′ ends would leave little to no central gene body for analysis.

See also Figure S2.

II and nascent transcripts by virtue of their exceptional stability on the chromatin template during transcription (Wuarin and Schibler, 1994; Carrillo Oesterreich et al., 2010) and can be used as a sensitive method to measure transcriptional elongation (see below).

We prepared nascent transcripts (Wuarin and Schibler, 1994) from control and MSL-RNAi-treated S2 cells, converted the RNA samples to strand-specific cDNA, and sequenced the cDNA on a HeliScope Single Molecule Sequencer. In the resulting profile, the density of the sequenced reads reflects the nascent RNA abundance. Assuming no RNA degradation, all nascent RNAs for a given transcript share the same 5′ end, but differ in their 3′ end depending on the position of elongating Pol II. Therefore, the abundance of RNA sequences should decrease from the 5′ to the 3′ ends of genes (Figure 1B), as shown for the *roX2* gene (Figure S1A). To summarize the genome-wide pattern, we plotted the metagene profile of read densities, in which actively transcribed genes > 2.5 kb long (> 0.5 kb after intron removal) were rescaled to have the same length (Figure 2). The intronic regions were removed because cotranscriptional splicing (Carrillo Oesterreich et al., 2010; Khodor et al., 2011; Bhatt et al., 2012) results in underrepresentation of intron sequences, thus distorting the profile related to Pol II processivity alone (Figure S2A). The profiles for cDNA sequencing show the expected 5′ to 3′ decline (Figure 2 and Figure S2B).

Male S2 cells have a relative X:A copy number ratio of 1:2. However, we observed a relative increase in X chromosome sequencing reads approaching the level of autosomal reads (X ≈2A), rather than being half as abundant (Figure 2 and Figures S2C–S2D). In contrast, after MSL1 and MSL2 RNAi, we found a decrease of X sequencing reads (red dashed line), which now fail to approach the 2A autosomal level (Figure 2). Thus, Nascent-seq successfully documents MSL-dependent DC at the level of nascent X versus autosomal transcription. However, sequencing cDNA derived from whole nascent transcripts did not allow us to identify the stages in the transcription cycle where critical differences occur.

To obtain a mechanistic picture of the affected steps, we needed a method to map the position of elongating Pol II with nucleotide resolution. We accomplished this by replicating the Nascent-seq results using direct RNA sequencing (DRS) (Ozsolak et al., 2009), designating this approach as direct nascent RNA sequencing, or DnRS (see the Experimental Procedures). DnRS starts from and is restricted to the 3′ end of the isolated transcript, and thus the expected profile (Figures 1B and S1A, right) reflects the actual position of Pol II at nucleotide resolution. We plotted the S2 cell DnRS metagene profile of active genes longer than 2.5 kb (Figure 3A). Because only the 3′ end of the nascent transcript is sequenced, intron removal is not necessary because cotranscriptional splicing does not interfere with Pol II localization. The resulting average profile shows progression of RNA Pol II along active genes (Figures 3A and 3B).

To determine which steps in transcription might be differentially regulated on X and autosomes, we compared their metagene profiles in S2 cells. When plotting the nascent RNA abundance in cells with a relative X:A copy number of 1:2, we found that there are fewer X-chromosomal (red) compared to autosomal (blue) nascent transcript reads mapped to the 5′ ends of the metagenes (Figure 3A), but the difference is less than 2-fold, suggesting partial DC (X > A). As transcription progresses, this difference narrows. In the last portion of the metagene, nascent transcript reads increase on both X and autosomes as Pol II approaches the 3′ polyadenylation site, possibly slowing in preparation for processing and termination (Core et al., 2008; Carrillo Oesterreich et al., 2010; Mayer et al., 2012). Interestingly, we found that X chromosomal nascent transcription, reflecting Pol II density, approaches autosomal levels (X ≈2A) (Figure 3C). This representation without gene body rescaling allows a better visualization of true absolute distances from the gene end (Figure 3C). In contrast to male S2 cells, DnRS of female Kc cells shows similar transcription of X and autosomal genes (2X ≈2A) during all phases of transcription (Figures 3B and 3D), validating our approach. We found these results to be significant and robust over variations of several analysis parameters, such as transcriptional threshold, gene length, and distance to neighboring gene (Figure S3).

The simplest explanation for our results is that Nascent-seq reveals increasing DC along the bodies of X-linked genes in male cells, with increased density of Pol II at steady state correlating with the increase in mRNA output. One possibility is that H4K16ac increases retention of X-linked Pol II that might otherwise prematurely terminate during elongation. Alternatively, increased efficiency of successfully elongating Pol II may result in positive feedback to Pol II molecules waiting to

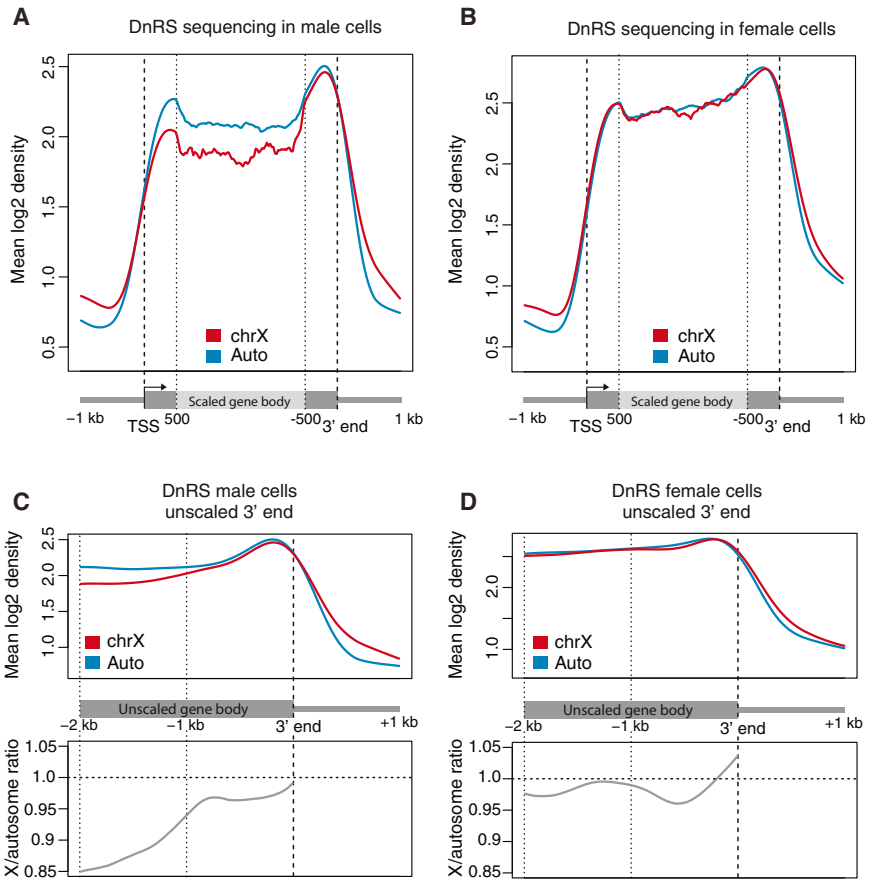

Figure 3. Nascent-Seq DnRS Analysis of Transcription in Male and Female Cells

(A and B) Average read-density profiles of actively transcribed genes for DnRS data in male (S2) cells (A) and female (Kc) cells (B) for the X (red; male n = 471, female n = 428) and autosomes (blue; male n = 1,612, female n = 1,552) are shown. The common genes between male and female are 405 and 1420 for X chromosome and autosomes, respectively. The metagene profile is generated by rescaling the gene coordinates to a fixed width between TSS and 3′ end. The initial and final 500 bp of genes is presented without scaling of the coordinates.

(C and D) Average DnRS read density near the 3′ end of actively transcribed genes (top panels) in male (S2) cells (C) and female (Kc) cells (D) is shown. Unscaled coordinates (real distance) for the same genes are shown. Ratio between average read density of X and autosomes is reported in the bottom panels.

See also Figure S3.

engage in earlier steps in the transcription cycle. Distinguishing between these two possibilities will require future development of methods to measure the kinetics of Pol II processivity genome-wide, for example, by visualizing a single round of transcription rather than the steady-state density of Pol II.

S2 and Kc cells are known to be polyploid (on average 2X:4A and 4X:4A, respectively) and to harbor numerous variations in copy number for individual genes. When assessing X to A transcription above, we simply compared total X and autosomal sequence reads per cell type because it is not possible to normalize for direct RNA sequencing efficiency or biases using genomic DNA sequencing. We also tested normalization by gene copy number and observed similar results, possibly dampened by a contribution from genome-wide compensation for aneuploidy, previously observed in S2 cells (Zhang et al., 2010) (Figures S4A and S4C). When assessing nascent RNA levels from each chromosome arm separately, 3R appears to be an outlier using correction per gene copy (Figure S4E), whereas 3L is an outlier using total read density (Figure S4G). The fourth chromosome is also quite variable, but this has negligible effect on the overall autosomal average because there are few active genes on the fourth, accounting for only 2.7% of the autosomal genes used in these analyses (Figure S4I). Because we observed these normalization differences, we believe that the specific values associated with X and A differences in our genomic analyses should be interpreted with caution. However, the general trends, such as a difference from 5′ to 3′ along X-linked genes (Figure S3 and discussed below), are consistent across all methods.

Nascent-seq demonstrated a clear effect across gene bodies and also suggested that regulation was starting early in the transcription cycle. Precisely where the regulation occurs, however, could not be evaluated using this technique (Figure 1A). Therefore, we analyzed potential X and autosome differences in initiation, 5′ pausing, and pausing release using several independent data sets from male S2 cells. We previously found that the average X-chromosomal level of 5′ paused Pol II is equivalent to the average autosomal level in male S2 cells by GRO-seq analysis when normalized to gene dose (X = A), consistent with lack of compensation at that step (Larschan et al., 2011). GRO-seq can detect engaged but paused Pol II, which is released from pausing by sarkosyl detergent. However, the large peaks of promoter-proximal paused Pol II might mask differences in nonpaused Pol II at early steps in transcription. Therefore, we used recently published data (Core et al., 2012) in which S2 cells were analyzed using two GRO-seq protocols: either the original GRO-seq protocol (Figure 4A) or actively elongating RNA Pol II only (GRO-seq protocol without sarkosyl treatment) (Figure 4B). The independent GRO-seq data are concordant with our previous observations (with the original protocol), showing no compensation of X-linked genes at the pausing step (Larschan et al., 2011). The ratio over genomic sequencing control shows that X = A at the 5′ end (Figure 4A). However, in the absence of sarkosyl (Figure 4B), X and autosomes differ at the 5′ end of the metagene, indicating that the population of Pol II engaged in active transcription at this early point already shows partial compensation (X > A). These results are consistent with a model in which release from 5′ pausing is a key rate-limiting step in

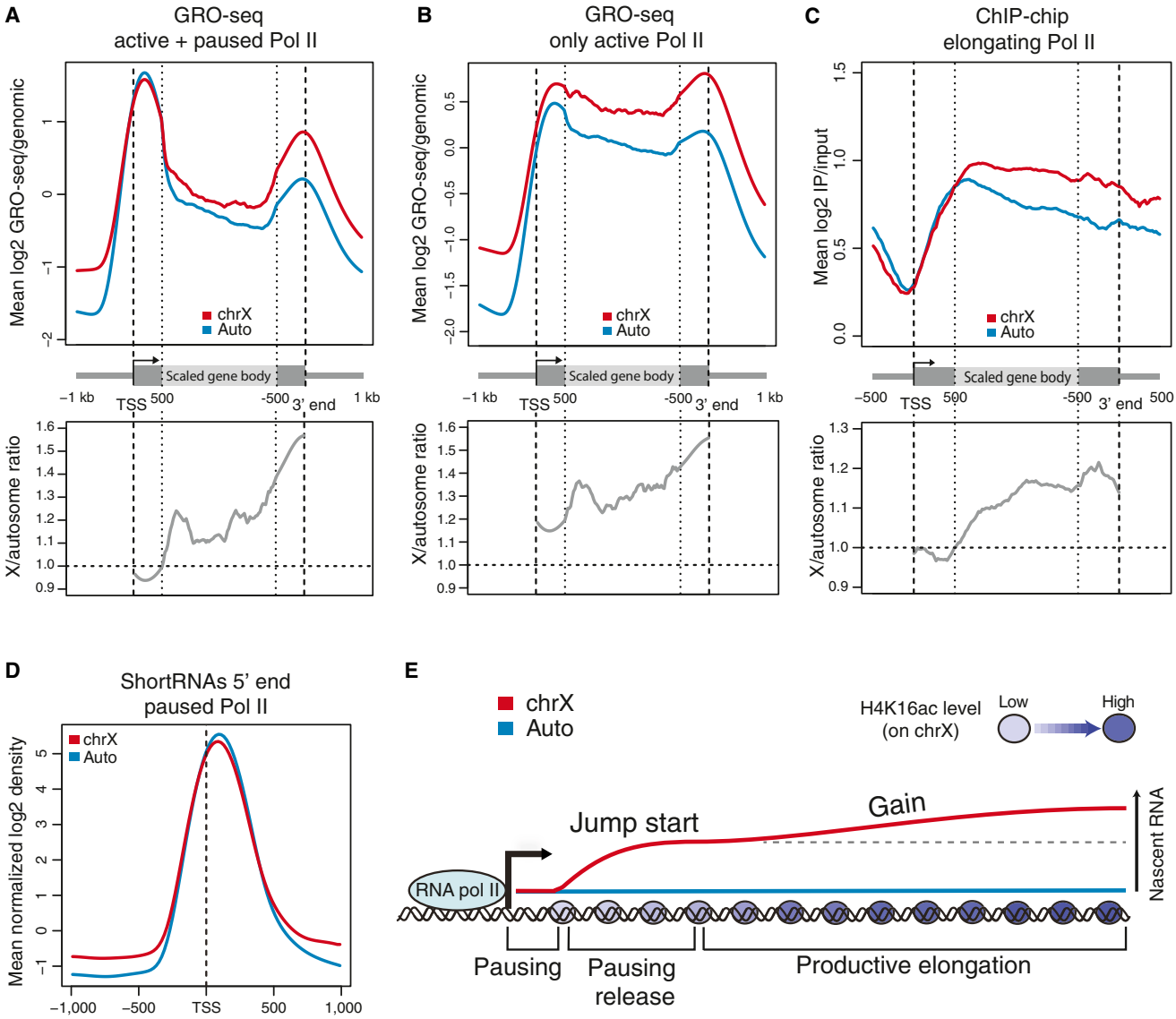

Figure 4. Public Data sets Confirm Partial Compensation at Early Steps of Transcription and Augmented Pol II Density at Later Stages

(A and B) Comparison of X (red) and autosomal (blue) Pol II density using recently published data (Core et al., 2012) in which GRO-seq in male (S2) cells was performed in both the presence and absence of sarkosyl to track paused plus active Pol II (A) or only active Pol II (B), respectively. The average $\log_2$ ratio over genomic sequencing control is shown (top panels) to account for the differences in the copy number for genes on chromosome X or autosomes. The metagene profile is generated as in Figure 3. The bottom panels show the ratio between the average profiles of the X (n = 398) and autosomes (n = 1,657) reported in the upper panels.

(C) The top panel shows the metagene profile of the RNA Pol II Ser2P occupancy on the X chromosome and autosomes in S2 cells using published Pol II Ser2P ChIP-chip data (Regnard et al., 2011). The bottom panel shows the ratio between the average occupancy profiles of the X (n = 507) and autosomes (n = 2,132).

(D) Data for short RNAs associated with paused Pol II from Nechaev et al. (2010) are shown. The average $\log_2$ ratio over genomic sequencing control is shown around transcription start sites (TSS), considering reads from the 5′ ends of short RNAs in X-linked (red line; n = 507) and autosomal (blue; n = 2,132) actively transcribed genes.

(E) Jump start and gain model: Paused Pol II is not augmented on the male X chromosome, but early elongation is increased on the male X relative to autosomes. Pol II release from pausing and entry into the elongation phase ("jump start") is facilitated by X-specific enrichment of H4K16 acetylation in the gene bodies. H4K16ac levels increase over the bodies of genes, therefore continuing to reduce steric hindrance and leading to a "gain" of progression or processivity of Pol II. See also Figure S4.

transcription, which can be facilitated on male X-linked genes. Both GRO-seq protocols also confirm a relative increase of Pol II density along the bodies of genes on X versus autosomes (Figures 4A and 4B, bottom panels).

We also compared X and autosomal levels of pausing from a published study on male S2 cells using an alternative approach in which 5′ capped RNAs < 100 bp in length were quantified (Nechaev et al., 2010). We found that normalized levels of

promoter-proximal short RNAs from the X chromosome are equivalent to autosomal genes, thus further supporting the GRO-seq results (Figure 4D). Taken together, GRO-seq, short RNA, and Nascent-seq results from independent data sets all point to transcriptional elongation as the process that is differentially regulated by DC, starting with pausing release and continuing along the gene body.

A recently published paper proposed an alternative model in which RNA Pol II recruitment is the key regulated step, with initiation, 5′ pausing, and all subsequent steps reflecting the initial increase (Conrad et al., 2012). This was supported by a 2-fold increase of RNA Pol II ChIP at X-linked promoters in male salivary glands, which was later revised to 1.2-fold after correction of an erroneous processing step (Ferrari et al., 2013; Straub and Becker 2013; Vaquerizas et al., 2013). Notably, the data from Conrad et al. (2012) provide no support for the elongation model (Ferrari et al., 2013; Straub and Becker 2013; Vaquerizas et al., 2013). Detection of elongation differences by ChIP might require antibodies specific for elongating Pol II phosphorylated at serine-2 (Ser2P) (Marshall et al., 1996; Lee and Greenleaf, 1997). We analyzed published ChIP-chip data for elongating Pol II Ser2P (possibly in conjunction with Ser5P) in male S2 cells from the Becker lab (Regnard et al., 2011), which show a clear increase of Pol II over X-linked gene bodies (Figure 4C). Taken together, Pol II Ser2P ChIP results are concordant with our Nascent-seq and GRO-seq analyses.

In summary, we have systematically dissected the mechanism of DC during distinct steps of transcription. Multiple high-resolution, genome-wide approaches converge on the following model (Figure 4E): paused Pol II is not augmented in general on the male X, but Pol II release from pausing ("jump-start" in our model) appears to be a key rate-limiting step that is facilitated by X-specific enrichment of H4K16ac in gene bodies. The increasing MSL and H4K16ac levels over the bodies of genes further reduce steric hindrance, leading to a "gain" of Pol II density. Currently, we cannot determine whether this gain is the result of (1) increased processivity (reduced termination) or (2) positive feedback to 5′ Pol II, to further increase pausing release. In either case, we believe that facilitated elongation through an acetylated chromatin template enables coordinate control of X-linked genes with widely differing mechanisms of individual, gene-specific regulation.

EXPERIMENTAL PROCEDURES

Cell Culture

S2 and Kc-167 cells were cultured at 25°C in Schneider's medium (Gibco) supplemented with 10% (v/v) heat-inactivated FBS (JRH).

Nascent RNA Isolation

The isolation of nascent RNA was adapted from Wuarin and Schibler (Wuarin and Schibler, 1994). See the Supplemental Experimental Procedures for details.

RNAi

RNAi treatment was performed as described previously (Gelbart et al., 2009; Larschan et al., 2011). See the Supplemental Experimental Procedures for details including primer sequences.

Helicos cDNA-Based Sequencing and Direct Nascent RNA Sequencing

Nascent transcripts were converted to strand-specific cDNA by random priming and modified with a poly-A tail according to the company's protocol prior to loading on the HeliScope Single Molecule Sequencer. For DnRS, the recovered nascent RNAs were submitted to Helicos. DnRS sequencing is similar to the published DRS protocol (Ozsolak et al., 2009), except that DRS was tailored to sequence polyadenylated mRNAs, whereas in DnRS an additional step was included before sequencing to add poly(A) tails to the RNA molecules.

Processing of Nascent-Seq Data

The sequenced reads were filtered and aligned using the Helisphere tools suite by Helicos (http://sourceforge.net/projects/openhelisphere/). See the Supplemental Experimental Procedures.

Gene Annotations and Estimation of Transcriptional Magnitude for Nascent-Seq and GRO-Seq Data

Refined annotations for mRNA transcripts were obtained from Graveley et al. (2011). The read count per gene was computed taking into account strand specificity. Comparison of the read counts between replicates confirmed good reproducibility (Figure S1B) and replicates were merged. To estimate transcription magnitude from Nascent-seq data, reads per kilobase per million mapped reads (RPKM) values were computed adjusting for mappability. The distribution of RPKM values was examined and RPKM ≥ 2, with at least three reads, were chosen as thresholds to select actively transcribed genes from Nascent-seq data. The same procedure was applied to GRO-seq data to define actively transcribed genes. See also Supplemental Experimental Procedures.

Metagene Profiles

To compute the average Nascent-seq read-density profile we used normalized Gaussian smoothing of read-mapping positions. Actively transcribed genes were further filtered by length and distance from neighboring loci. For cDNA sequencing data, introns were excluded from specific metagene profiles as reported in the figure legends. Scaling of genes was achieved by splitting smoothed profiles into 200 bins for each gene and flanking regions, and then computing the average for each bin in individual genes. Read density was then $\log_2$ transformed, after adding a pseudocount of 1 to each bin to avoid log transformation of zero values. The value of the Nth bin in the metagene profile is then the average of the $\log_2$ read density in the Nth bins across the gene set. The same procedure was used for unscaled metagene profiles centered at gene ends, but equally sized bins were used. For GRO-seq data, the average $\log_2$ ratio between GRO-seq and genomic control read densities was computed for each gene and for each bin to facilitate comparison with other public data sets (ChIP-chip and short RNAs) (Figure 4). See the Supplemental Experimental Procedures for more details.

ChIP-chip Data

The data set for Pol II ser-2P ChIP-chip profiles used in this study were obtained from Regnard et al. (2011). To determine gene expression level, we used public RNA-seq data (GSE15596 in the Gene Expression Omnibus). See also Supplemental Experimental Procedures.

ACCESSION NUMBERS

The sequencing data were deposited in the Sequence Read Archive under accession number for SRA062950.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.09.037.

ACKNOWLEDGMENTS

We are very grateful to Drs. S. Churchman, J. Gray, J. Lis, E. Larschan, and A. Mayer for critically reading the manuscript. This work was supported by grants from the National Institutes of Health (GM45744 to M.I.K. and HG005230 to F.O.), and a fellowship to A.P. from the German Academic Exchange Service. F.O. was an employee of Helicos BioSciences.

Received: April 24, 2013
Revised: August 13, 2013
Accepted: September 25, 2013
Published: October 31, 2013

REFERENCES

Akhtar, A., and Becker, P.B. (2000). Activation of transcription through histone H4 acetylation by MOF, an acetyltransferase essential for dosage compensation in Drosophila. Mol. Cell *5*, 367–375.

Alekseyenko, A.A., Larschan, E., Lai, W.R., Park, P.J., and Kuroda, M.I. (2006). High-resolution ChIP-chip analysis reveals that the Drosophila MSL complex selectively identifies active genes on the male X chromosome. Genes Dev. *20*, 848–857.

Alekseyenko, A.A., Ho, J.W., Peng, S., Gelbart, M., Tolstorukov, M.Y., Plachetka, A., Kharchenko, P.V., Jung, Y.L., Gorchakov, A.A., Larschan, E., et al. (2012). Sequence-specific targeting of dosage compensation in Drosophila favors an active chromatin context. PLoS Genet. *8*, e1002646.

Bhatt, D.M., Pandya-Jones, A., Tong, A.-J., Barozzi, I., Lissner, M.M., Natoli, G., Black, D.L., and Smale, S.T. (2012). Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. Cell *150*, 279–290.

Buratowski, S. (2009). Progression through the RNA polymerase II CTD cycle. Mol. Cell *36*, 541–546.

Carrillo Oesterreich, F., Preibisch, S., and Neugebauer, K.M. (2010). Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. Mol. Cell *40*, 571–581.

Conrad, T., Cavalli, F.M.G., Vaquerizas, J.M., Luscombe, N.M., and Akhtar, A. (2012). Drosophila dosage compensation involves enhanced Pol II recruitment to male X-linked promoters. Science *337*, 742–746.

Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848.

Core, L.J., Waterfall, J.J., Gilchrist, D.A., Fargo, D.C., Kwak, H., Adelman, K., and Lis, J.T. (2012). Defining the status of RNA polymerase at promoters. Cell Rep. *2*, 1025–1035.

Dou, Y., Milne, T.A., Tackett, A.J., Smith, E.R., Fukuda, A., Wysocka, J., Allis, C.D., Chait, B.T., Hess, J.L., and Roeder, R.G. (2005). Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell *121*, 873–885.

Ferrari, F., Jung, Y.L., Kharchenko, P.V., Plachetka, A., Alekseyenko, A.A., Kuroda, M.I., and Park, P.J. (2013). Comment on "Drosophila dosage compensation involves enhanced Pol II recruitment to male X-linked promoters". Science *340*, 273.

Gelbart, M.E., and Kuroda, M.I. (2009). Drosophila dosage compensation: a complex voyage to the X chromosome. Development *136*, 1399–1410.

Gelbart, M.E., Larschan, E., Peng, S., Park, P.J., and Kuroda, M.I. (2009). Drosophila MSL complex globally acetylates H4K16 on the male X chromosome for dosage compensation. Nat. Struct. Mol. Biol. *16*, 825–832.

Gilfillan, G.D., Straub, T., de Wit, E., Greil, F., Lamm, R., van Steensel, B., and Becker, P.B. (2006). Chromosome-wide gene-specific targeting of the Drosophila dosage compensation complex. Genes Dev. *20*, 858–870.

Gilmour, D.S. (2009). Promoter proximal pausing on genes in metazoans. Chromosoma *118*, 1–10.

Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). The developmental transcriptome of Drosophila melanogaster. Nature *471*, 473–479.

Gu, W., Szauter, P., and Lucchesi, J.C. (1998). Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of Drosophila melanogaster. Dev. Genet. *22*, 56–64.

Gupta, V., Parisi, M., Sturgill, D., Nuttall, R., Doctolero, M., Dudko, O.K., Malley, J.D., Eastman, P.S., and Oliver, B. (2006). Global analysis of X-chromosome dosage compensation. J. Biol. *5*, 3.

Hamada, F.N., Park, P.J., Gordadze, P.R., and Kuroda, M.I. (2005). Global regulation of X chromosomal genes by the MSL complex in Drosophila melanogaster. Genes Dev. *19*, 2289–2294.

Khodor, Y.L., Rodriguez, J., Abruzzi, K.C., Tang, C.-H.A., Marr, M.T., 2nd, and Rosbash, M. (2011). Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in Drosophila. Genes Dev. *25*, 2502–2512.

Larschan, E., Bishop, E.P., Kharchenko, P.V., Core, L.J., Lis, J.T., Park, P.J., and Kuroda, M.I. (2011). X chromosome dosage compensation via enhanced transcriptional elongation in Drosophila. Nature *471*, 115–118.

Lee, J.M., and Greenleaf, A.L. (1997). Modulation of RNA polymerase II elongation efficiency by C-terminal heptapeptide repeat domain kinase I. J. Biol. Chem. *272*, 10990–10993.

Lucchesi, J.C., Kelly, W.G., and Panning, B. (2005). Chromatin remodeling in dosage compensation. Annu. Rev. Genet. *39*, 615–651.

Marshall, N.F., Peng, J., Xie, Z., and Price, D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J. Biol. Chem. *271*, 27176–27183.

Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C.M., and Cramer, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. EMBO J. *30*, 1302–1310.

Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J., and Cramer, P. (2010). Uniform transitions of the general RNA polymerase II transcription complex. Nat. Struct. Mol. Biol. *17*, 1272–1278.

Mayer, A., Heidemann, M., Lidschreiber, M., Schreieck, A., Sun, M., Hintermair, C., Kremmer, E., Eick, D., and Cramer, P. (2012). CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. Science *336*, 1723–1725.

Nechaev, S., Fargo, D.C., dos Santos, G., Liu, L., Gao, Y., and Adelman, K. (2010). Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. Science *327*, 335–338.

Ozsolak, F., Platt, A.R., Jones, D.R., Reifenberger, J.G., Sass, L.E., McInerney, P., Thompson, J.F., Bowers, J., Jarosz, M., and Milos, P.M. (2009). Direct RNA sequencing. Nature *461*, 814–818.

Peterlin, B.M., and Price, D.H. (2006). Controlling the elongation phase of transcription with P-TEFb. Mol. Cell *23*, 297–305.

Prabhakaran, M., and Kelley, R.L. (2012). Mutations in the transcription elongation factor SPT5 disrupt a reporter for dosage compensation in Drosophila. PLoS Genet. *8*, e1003073.

Regnard, C., Straub, T., Mitterweger, A., Dahlsveen, I.K., Fabian, V., and Becker, P.B. (2011). Global analysis of the relationship between JIL-1 kinase and transcription. PLoS Genet. *7*, e1001327.

Sainsbury, S., Niesser, J., and Cramer, P. (2013). Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. Nature *493*, 437–440.

Smith, E.R., Pannuti, A., Gu, W., Steurnagel, A., Cook, R.G., Allis, C.D., and Lucchesi, J.C. (2000). The Drosophila MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage compensation. Mol. Cell. Biol. *20*, 312–318.

Smith, E.R., Allis, C.D., and Lucchesi, J.C. (2001). Linking global histone acetylation to the transcription enhancement of X-chromosomal genes in Drosophila males. J. Biol. Chem. *276*, 31483–31486.

Straub, T., Gilfillan, G.D., Maier, V.K., and Becker, P.B. (2005). The Drosophila MSL complex activates the transcription of target genes. Genes Dev. *19*, 2284–2288.

Straub, T., and Becker, P.B. (2013). Comment on "Drosophila dosage compensation involves enhanced Pol II recruitment to male X-linked promoters". Science *340*, 273.

Vaquerizas, J.M., Cavalli, F.M.G., Conrad, T., Akhtar, A., and Luscombe, N.M. (2013). Response to Comments on "Drosophila Dosage Compensation Involves Enhanced Pol II Recruitment to Male X-Linked Promoters". Science *340*, 273.

Weeks, J.R., Hardin, S.E., Shen, J., Lee, J.M., and Greenleaf, A.L. (1993). Locus-specific variation in phosphorylation state of RNA polymerase II in vivo: correlations with gene activity and transcript processing. Genes Dev. *7*(12A), 2329–2344.

Wuarin, J., and Schibler, U. (1994). Physical isolation of nascent RNA chains transcribed by RNA polymerase II: evidence for cotranscriptional splicing. Mol. Cell. Biol. *14*, 7219–7225.

Zhang, Y., Malone, J.H., Powell, S.K., Periwal, V., Spana, E., Macalpine, D.M., and Oliver, B. (2010). Expression in aneuploid Drosophila S2 cells. PLoS Biol. *8*, e1000320.

Update

Cell Reports

Volume 5, Issue 4, 27 November 2013, Page 1157

DOI: https://doi.org/10.1016/j.celrep.2013.11.024

Cell Reports
Correction

"Jump Start and Gain" Model for Dosage Compensation in *Drosophila* Based on Direct Sequencing of Nascent Transcripts

Francesco Ferrari, Annette Plachetka, Artyom A. Alekseyenko, Youngsook L. Jung, Fatih Ozsolak, Peter V. Kharchenko, Peter J. Park,* and Mitzi I. Kuroda*
*Correspondence: peter_park@hms.harvard.edu (P.J.P.), mkuroda@genetics.med.harvard.edu (M.I.K.)
http://dx.doi.org/10.1016/j.celrep.2013.11.024

(Cell Reports *5*, 629–636; October 31, 2013)
In the originally published version of this article, the text "Jump start" in Figure 4E was cut off and illegible. Figure 4 has now been corrected online.

The journal apologizes for any confusion this error may have caused.