nature structural & molecular biology

Functional interplay between MSL1 and CDK7 controls RNA polymerase II Ser5 phosphorylation

Sarantis Chlamydas[1,8], Herbert Holz[1,8], Maria Samata[1,2,8], Tomasz Chelmicki[1], Plamen Georgiev[1], Vicent Pelechano[3,4], Friederike Dündar[1,2], Pouria Dasmeh[1], Gerhard Mittler[1], Filipe Tavares Cadete[5], Fidel Ramírez[1], Thomas Conrad[1], Wu Wei[3,6], Sunil Raja[1], Thomas Manke[1], Nicholas M Luscombe[5,7], Lars M Steinmetz[3,6] & Asifa Akhtar[1]

Proper gene expression requires coordinated interplay among transcriptional coactivators, transcription factors and the general transcription machinery. We report here that MSL1, a central component of the dosage compensation complex in *Drosophila melanogaster* and *Drosophila virilis*, displays evolutionarily conserved sex-independent binding to promoters. Genetic and biochemical analyses reveal a functional interaction of MSL1 with CDK7, a subunit of the Cdk-activating kinase (CAK) complex of the general transcription factor TFIIH. Importantly, MSL1 depletion leads to decreased phosphorylation of Ser5 of RNA polymerase II. In addition, we demonstrate that MSL1 is a phosphoprotein, and transgenic flies expressing MSL1 phosphomutants show mislocalization of the histone acetyltransferase MOF and histone H4 K16 acetylation, thus ultimately causing male lethality due to a failure of dosage compensation. We propose that, by virtue of its interaction with components of the general transcription machinery, MSL1 exists in different phosphorylation states, thereby modulating transcription in flies.

Combinatorial action of *cis*-acting elements at gene promoters is the first step toward establishing efficient transcription. Accurate loading of RNA polymerase II (Pol II) at the transcription start sites (TSSs), as well as its controlled transition from initiation to active elongation, requires orchestrated cooperation among a large number of proteins including general transcription factors and histone-modifying enzymes. Importantly, these factors' recruitment to DNA, as well as their catalytic activity, can be often modulated via their post-translational modifications (PTMs). Extensive phosphorylation of the C-terminal domain (CTD) of Pol II's largest subunit, RBP1, is a prime example of a PTM determining enzymatic activity during the transcription cycle. Generation of fully functional mRNA requires Pol II to pass through a cycle of events starting with recruitment, promoter escape, active elongation, 5′-end capping, splicing and finally 3′-end polyadenylation, termination and recycling[1,2]. The ability of Pol II to sequentially execute transcription events relies on distinct phosphorylation patterns of its CTD[3], in which each of the serine residues within the CTD consensus sequence of heptad repeats (1-YSPTSPS-7) has the potential to be phosphorylated[3–5]. The network of metazoan CTD kinases involved in modulating Pol II activity is well characterized, and CDK7 and CDK9 play prominent roles. CDK7, cyclin H and Mat1 form a CAK subcomplex within TFIIH and mediate phosphorylation of Pol II Ser5 and Ser7 (forming Pol II Ser5p and Pol II Ser7p, respectively). This phosphorylation occurs concomitantly with early events of transcription[6–9], recruitment of the integrator complex and transcription of small nuclear RNAs[10,11]. The P-TEFb elongation complex, comprised of CDK9 and a cyclin T1 subunit, is responsible for CTD Ser2 phosphorylation (forming Pol II Ser2p), which is necessary for transcription elongation and 3′-end RNA processing[2,6,9,12].

The coordinated action between the chromatin-modifying machinery and the recruitment of general transcription factors can be appreciated chromosome wide during the process of dosage compensation in *D. melanogaster* (*dm*). The male-specific lethal (MSL) complex, composed of at least five proteins (MSL1, MSL2, MSL3, MLE and MOF) and two long noncoding RNAs (roX1 and roX2), increases the gene expression from the single X chromosome in males through histone H4 Lys16 acetylation (forming H4K16ac), thereby efficiently erasing gene-dose imbalance with respect to the two X chromosomes in females (reviewed in ref. 13). Interestingly, individual MSL-complex members exhibit distinct chromatin binding patterns across the X chromosome. For instance, MSL1 and MOF are enriched at gene promoters and throughout gene bodies, MSL3 peaks toward the 3′ ends of genes, and MSL2 and MLE mainly colocalize at high-affinity sites, which are recruitment hubs for the MSL complex[14,15]. Apart from the dosage-compensation role of the MSL complex, accumulating evidence suggests that MSL proteins may be involved in additional functions. In addition to X-linked binding, MSL1 and MOF have been demonstrated to recognize a number of autosomal gene promoters[14,16]. This recognition is also true for the mouse and human counterparts of MSLs: MSL1, MSL2 and MOF regulate transcription

[1]Max Planck Institute of Immunobiology and Epigenetics, Freiburg im Breisgau, Germany. [2]University of Freiburg, Faculty of Biology, Freiburg im Breisgau, Germany. [3]Genome Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany. [4]SciLifeLab, Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Solna, Sweden. [5]The Francis Crick Institute, London, UK. [6]Stanford Genome Technology Center, Stanford University, Palo Alto, California, USA. [7]UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, London, UK. [8]These authors contributed equally to this work. Correspondence should be addressed to A.A. (akhtar@ie-freiburg.mpg.de).

Received 25 November 2015; accepted 21 April 2016; published online 16 May 2016; doi:10.1038/nsmb.3233

through chromatin binding to gene promoters and *cis* regulatory elements[17–21]. However, beyond the role of the MSL complex in the deposition of H4K16ac, as well as the evolutionarily conserved MSL binding at TSSs, little is known about the direct links between MSL proteins and the general transcription machinery carrying out the promoter-proximal transcription events.

Here, we performed a series of *in vivo* and *in vitro* studies, which reveal a new functional link between CDK7 and MSL1. CDK7 and MSL1 interact *in vitro*, and the concomitant decrease of MSL1 and CDK7 protein levels by RNA interference (RNAi) leads to synthetic lethality in flies, thus suggesting that this interaction is biologically important. Furthermore, we demonstrate that MSL1 is a phosphoprotein and that inhibition of CDK7 through a small-molecule inhibitor leads to loss of MSL1 localization at chromatin. Moreover, mutation of the phosphorylation sites of MSL1 leads to protein mislocalization, loss of H4K16ac and male lethality, thus suggesting that this post-translational modification is important for MSL1 function in dosage compensation. Concurrently, we found that depletion of MSL1 results in a decrease in CDK7-catalyzed phosphorylation of Pol II Ser5 at MSL1-bound gene promoters. Thus, our data reveal that a complex functional interplay between MSL1 and CDK7 fine-tunes gene expression and has biological consequences in *Drosophila* dosage compensation. Nevertheless, the precise mechanism by which these two proteins cross-regulate each other requires further investigation.

RESULTS

MSL1 binding to promoters is evolutionarily conserved

The MSL1 protein is a core component of *Drosophila* and mammalian MSL complexes. Earlier structural studies of MSL1 have revealed its dual behavior in chromatin targeting. Although *Drosophila* X-linked genes undergoing dosage compensation display both promoter and gene-body MSL1 enrichments, autosomal MSL1 binding is restricted to promoters. Moreover, the ability of MSL1 to recognize promoters is sex independent[22]. To examine whether this phenomenon occurs globally, we performed MSL1 chromatin immunoprecipitation profiling (ChIP–seq) from third-instar larva salivary glands of male and female fruit flies. In addition to the X chromosomal gene-body binding, MSL1 also exhibited sex-independent promoter enrichments for X-linked and autosomal genes. In fact, MSL1 profiles largely resembled those of the MOF protein[23] (**Fig. 1a**). Interestingly, only MSL1 gene-body binding, but not TSS binding, was lost chromosome-wide in *mof*^{2}-mutant flies carrying a premature stop codon in the *mof* gene[24] (**Fig. 1a**). This finding is in agreement with previous observations obtained for selected genes[25]. These findings suggest that promoter binding of MSL1 is MOF independent and that it may play a general role in regulating transcription.

Next, we evaluated the MSL1 chromatin distribution across evolutionary time. Using a specific antibody raised against *D. virilis* (*dv*) MSL1, we generated ChIP–seq profiles from *D. virilis* male larvae. We chose *D. virilis* because it represents one of the earliest evolutionary divergences (~40 to 60 Myr) from *D. melanogaster* in the drosophilid phylogeny[26,27]. We observed predominantly promoter-restricted binding of *dv*MSL1 on all chromosomes (**Fig. 1b** and **Supplementary Fig. 1a–c**), thus suggesting that the TSS-focused binding of MSL1 is evolutionarily conserved. Further evidence supporting this notion came from the analysis of MSL1 in mouse embryonic stem cells (mESCs), in which MSL1 binds a subset of active gene promoters[17] (**Supplementary Figs. 1c,d** and **2e**).

Next, to assess the scope of MSL1-dependent gene expression control, we performed RNA sequencing in MSL1-depleted male S2 cells. Interestingly, MSL1 depletion led to global gene-expression changes (**Fig. 1c,d**). As expected, the X-linked genes were primarily downregulated. In contrast, autosomal genes were both up- and downregulated after MSL1 depletion. This observation held true regardless of the stringency threshold used (**Fig. 1d**). Moreover, the effect on X-linked genes was more pronounced than it was on the autosomal genes, most probably because of the contribution of the dosage-compensation machinery (**Fig. 1d**). These observations suggest that MSL1 binding to gene promoters is evolutionarily conserved and correlates with changes in gene expression of both X-linked and autosomal genes.

MSL1 depletion results in decreased RNA Pol II Ser5p

We next set out to evaluate the functional roles of the distinct *dm*MSL-complex members in transcription-associated events. Systematic depletion of MSL1, MSL2 and MOF in S2 cells did not substantially affect overall Pol II protein levels (**Fig. 2a** and **Supplementary Fig. 2a–c**). However, when we probed for Pol II phosphorylation status after MSL1 and MSL2 depletion, we observed a pronounced decrease in Pol II Ser5p (**Fig. 2a** and **Supplementary Fig. 2a–c**) but no obvious changes in Pol II Ser2p. Depletion of MSL1 in female Kc cells also resulted in a visible decrease in Pol II Ser5p (**Fig. 2b**). Notably, MOF depletion in male S2 cells did not affect Pol II phosphorylation levels (**Supplementary Fig. 2c**), thus indicating that MSL1 and MSL2 function independently of MOF in transcription regulation. Consistently with these data, a moderate, but visible reduction in Pol II Ser5p was also present in MSL1-mutant female flies (**Fig. 2c,d**). Because of lethality, an analogous experiment was not feasible in MSL1-mutant males.

Next, we tested whether the MSL1-dependent decrease in Pol II Ser5p is conserved in other *Drosophila* species and in mammals. Indeed, we observed a similar decrease in Pol II Ser5p after MSL1 depletion in *D. virilis* (**Fig. 2e**) but not in mESCs (**Supplementary Fig. 2f,g**). However, owing to the relatively small fraction of MSL1-bound genes in mESCs (**Supplementary Figs. 1d** and **2e**), we speculate that a global decrease in Pol II Ser5p may not be obvious.

To assess whether a decrease in Pol II Ser5p occurred at MSL1 target loci, we performed ChIP assays in S2 cells. Consistently with the results from western blot (WB) analysis, Pol II Ser5p ChIP and locus-specific qPCR after MSL1 depletion revealed a decrease in Pol II Ser5p at promoters of MSL1 targets (**Fig. 2f–h** and **Supplementary Fig. 2d**). In contrast, the overall levels of Pol II at the same gene loci were unaffected (**Fig. 2g**). These data establish that MSL1 is required for efficient phosphorylation of Pol II Ser5 at its target genes in flies.

msl-1 and *Cdk7* show genetic interaction in flies

Genetic studies have demonstrated that the presence of MSL1 is essential for *Drosophila* male viability, owing to its role in dosage compensation. In light of our observations that MSL1 binds to promoters of autosomal genes in male and female flies and that the loss of MSL1 causes a decrease in Pol II Ser5p, we set out to test whether MSL1-mutant females also have decreased viability. We characterized three different *msl-1* loss-of-function mutants, namely *msl-1*^{L60} (ref. 28), *msl-1*$^{\gamma 269}$ (ref. 29) and *msl-1*^{1} (ref. 30), as well as the effects of missing MSL1 maternal contribution (**Supplementary Fig. 3a,b**; details in Online Methods). Because none of these mutants lacked the full open reading frame, expression of stable truncated proteins with residual function might affect viability. Thus, we further complemented our analysis by using an upstream activating sequence (UAS)-*msl-1*^{RNAi} line with a strong ubiquitous *tub-Gal4* driver to silence the expression of *msl-1* (**Supplementary Fig. 3c**). In all the cases, males showed 100% lethality, whereas the viability of females was unaffected by the loss of MSL1. From these results, it appears that the loss of promoter-associated MSL1 in female flies is not sufficient to cause lethality.

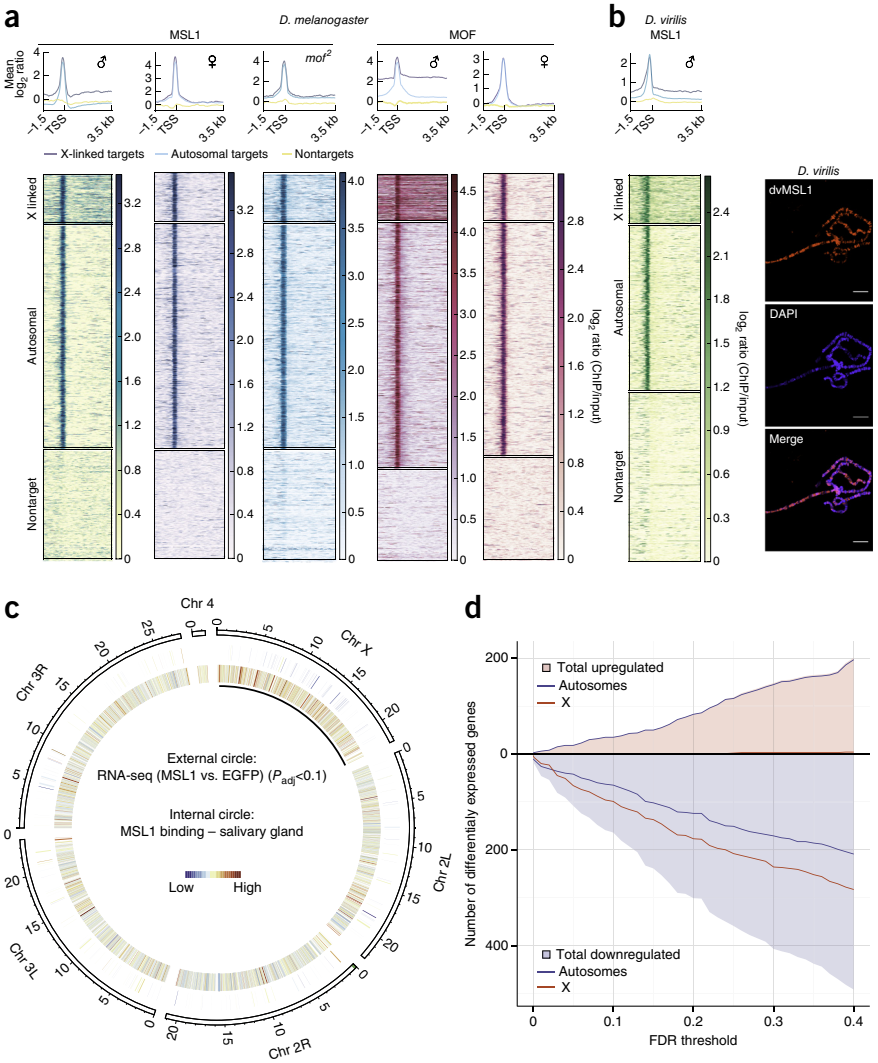

Figure 1 Evolutionarily conserved binding of MSL1 to target promoters. (**a**,**b**) Input-normalized ChIP–seq signals for MSL1 and MOF in different cell types for promoter-target genes and a randomly chosen subset (20%) of nontarget genes. Within the heat maps, the gene regions are sorted on the basis of their chromosome location; i.e., X-linked genes are on top and are followed by genes on chromosome 2, then 3 and so forth. Heat maps and summary plots were generated with the deepTools suite[59]. ChIP–seq of *D. melanogaster* salivary glands and *D. virilis* male larvae was performed on samples from two independent collections of larvae. Rabbit anti-MSL1 was used in the *D. melanogaster* ChIP–seq experiments. Additional data are in **Supplementary Figure 1b–d**. (**b**) Left, ChIP–seq profiles (as in **a**) from male *D. virilis* larvae. Right, immunostaining of *D. virilis* male larva polytene chromosomes with a *dv*MSL1–specific antibody (red). DNA is counterstained with DAPI (blue). Additional data are in **Supplementary Figure 1a**. Scale bars, 20 µm. (**c**) Genome-wide MSL1 binding (ChIP–seq; salivary glands, inner circle) and transcriptional effects (RNA-seq; S2 cells, outer circle) in *D. melanogaster*. Outer circle represents changes in expression measured by RNA-seq; only genes expressed at significantly different levels are shown (false discovery rate (FDR) <0.1 with Benjamini–Hochberg procedure (P_{adj}, adjusted *P* value); Online Methods). Chr, chromosome. (**d**) Gene expression (RNA-seq) analysis after MSL1 depletion in S2 cells. Numbers of differentially expressed genes at different threshold levels for adjusted *P* values (FDR with Benjamini–Hochberg procedure; Online Methods) are shown. The colored area depicts the total numbers of genes (pink, upregulated; blue, downregulated). Separate contributions from X-linked and autosomal genes are also depicted as solid lines (red and blue, respectively). RNA-seq from S2 cells was performed on samples from three independently dsRNA-transfected cell cultures.

Given the decrease in Pol II Ser5p after MSL1 loss in both male and female cells, we asked whether there might be a functional link between MSL1 and any of the kinases associated with Pol II Ser5p. Previous studies have demonstrated that Ser5 in the Pol II CTD is phosphorylated by at least three cyclin-dependent kinases: CDK7, CDK8 and CDK9 (ref. 31). Using *Drosophila* transgenic strains that allow double-stranded RNAi (dsRNAi)-induced silencing of *Cdk7*, *Cdk8* or *Cdk9* and two loss-of-function *msl-1* alleles (*msl-1*^{L60} *and msl-1*$^{\gamma 269}$), we addressed the possibility of genetic interaction. Strikingly, RNAi of *Cdk7* in *msl-1*^{L60} or *msl-1*$^{\gamma 269}$ heterozygotes resulted in enhanced lethality in both males and females. The lethality of RNAi-mediated *Cdk8* and *Cdk9* knockdown was not enhanced by the presence of an *msl-1* loss-of-function allele (**Fig. 3a**). When we compared the viability of compound-mutant heterozygous *msl-1*^{L60} females and *msl-1*$^{L60/\gamma 269}$ *trans*-heterozygous females, we also observed a strong enhancement of the lethality phenotype with *Cdk7* dsRNAi but not *Cdk9* dsRNAi (**Fig. 3b**). Unfortunately, this genetic approach could be carried out in only females, because *msl-1*$^{L60/\gamma 269}$ *trans*-heterozygous males are not viable. We also observed a genetic interaction between *Cdk7* and *msl-1* when using the temperature-sensitive *Cdk7*^{P140S} allele[32] at 25 °C (**Supplementary Fig. 4a**; details in Online Methods). Given that MSL1 depletion had no effect on the CDK7 protein levels in cells from both male and female *Drosophila* (**Fig. 2a**,**b**), these results together indicate a functionally relevant and sex-independent genetic interaction between *msl-1* and *Cdk7* and suggest a more general role of MSL1 in transcription, possibly extending beyond dosage compensation.

MSL1 chromosomal localization is sensitive to CDK7 inhibition

A positive genetic interaction between *msl-1* and *Cdk7* encouraged us to further probe for an interplay between these two factors. Because complete CDK7 depletion does not result in a decrease in Pol II Ser5p, owing to kinase redundancy[33,34], we chose an alternative approach of using small molecules to specifically inhibit CDK7 activity. This strategy also allowed us to delineate direct and indirect effects because of the short time frame required for CDK7 inhibition. To this end, we used a well-established mammalian CDK7 inhibitor, BS-181 (refs. 35,36). We first established that all key residues for binding CDK7 (refs. 35,36) are conserved between humans and flies (**Fig. 3c**). The incubation time was kept to a minimum to allow immediate and secondary effects to be distinguished. A short time course of treatment (5 to 20 min) in S2 cells revealed that administration of

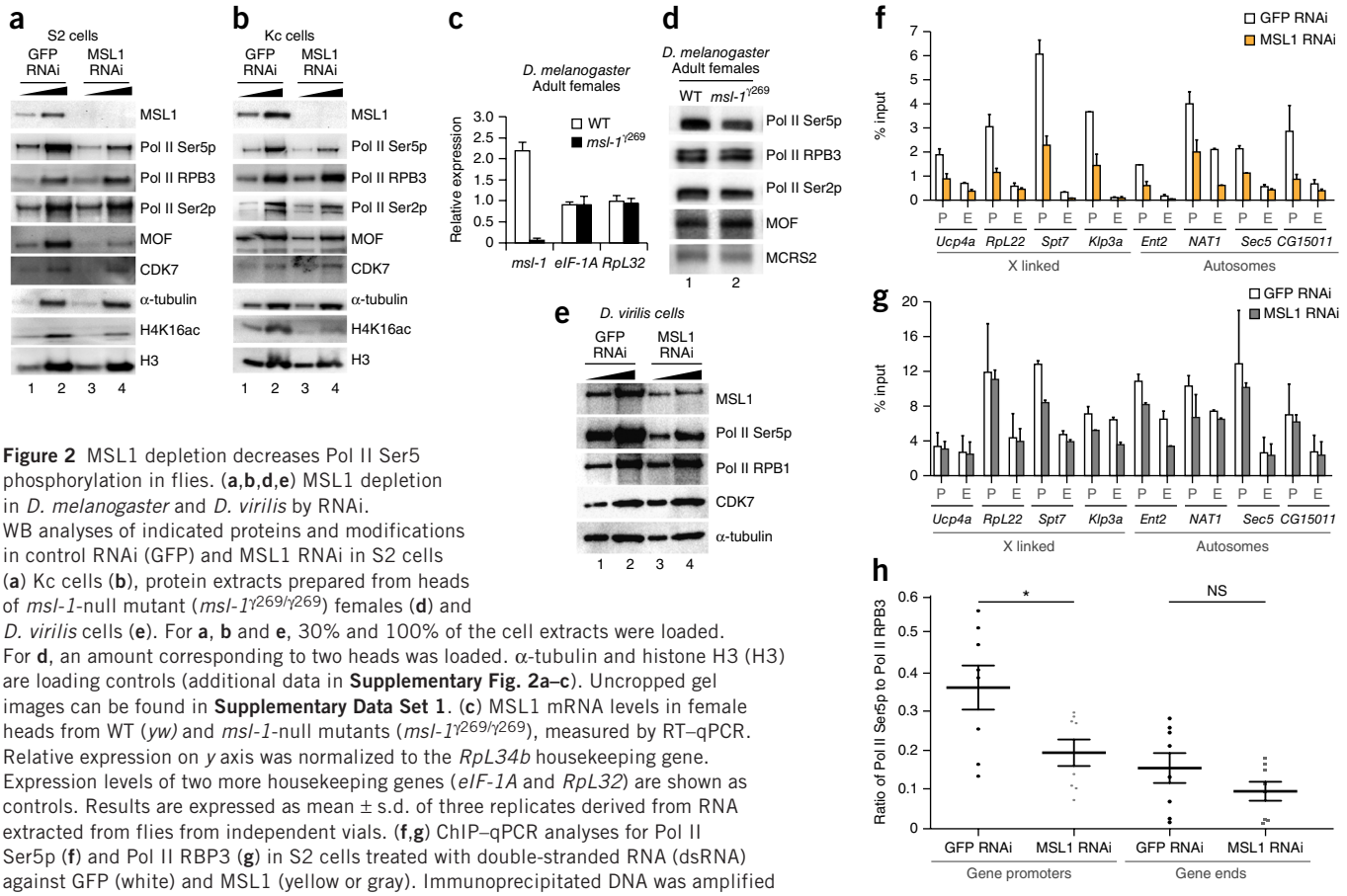

Figure 2 MSL1 depletion decreases Pol II Ser5 phosphorylation in flies. (**a**,**b**,**d**,**e**) MSL1 depletion in *D. melanogaster* and *D. virilis* by RNAi. WB analyses of indicated proteins and modifications in control RNAi (GFP) and MSL1 RNAi in S2 cells (**a**) Kc cells (**b**), protein extracts prepared from heads of *msl-1*-null mutant (*msl-1*$^{\gamma269/\gamma269}$) females (**d**) and *D. virilis* cells (**e**). For **a**, **b** and **e**, 30% and 100% of the cell extracts were loaded. For **d**, an amount corresponding to two heads was loaded. α-tubulin and histone H3 (H3) are loading controls (additional data in **Supplementary Fig. 2a–c**). Uncropped gel images can be found in **Supplementary Data Set 1**. (**c**) MSL1 mRNA levels in female heads from WT (*yw*) and *msl-1*-null mutants (*msl-1*$^{\gamma269/\gamma269}$), measured by RT–qPCR. Relative expression on *y* axis was normalized to the *RpL34b* housekeeping gene. Expression levels of two more housekeeping genes (*eIF-1A* and *RpL32*) are shown as controls. Results are expressed as mean ± s.d. of three replicates derived from RNA extracted from flies from independent vials. (**f**,**g**) ChIP–qPCR analyses for Pol II Ser5p (**f**) and Pol II RBP3 (**g**) in S2 cells treated with double-stranded RNA (dsRNA) against GFP (white) and MSL1 (yellow or gray). Immunoprecipitated DNA was amplified by qPCR with primer sets positioned at the promoter (P) and end (E) of indicated genes. Results are expressed as median ± s.d. of three experiments with independent cell cultures. (**h**) Scatter plot presenting all individual data points and the median ± s.e.m. of the individual ratios of Pol II Ser5p (**f**) and Pol II RBP3 (**g**) at gene promoters and gene ends in GFP (control) and MSL1-depleted cells. Asterisk represents statistical significance calculated by two-tailed Mann–Whitney test at gene promoters ($P = 0.049$; $n = 8$ gene promoters) or gene ends ($P = 0.2317$; $n = 8$ gene ends) for MSL1 compared with GFP depletion. NS, not significant. Additional data are in **Supplementary Figure 2d**. Source data for **c**, **f**, **g** and **h** are available online.

this compound led to a decrease in Pol II Ser5p in *Drosophila* cells (**Fig. 3d** and **Supplementary Fig. 4b**). Importantly, MSL1 was still able to coimmunoprecipitate MSL3 and MOF after BS-181 treatment (**Supplementary Fig. 4b**), thus suggesting that introduction of the inhibitor does not disrupt the MSL complex.

Next, to study MSL1 chromosomal localization, we incubated male *D. melanogaster* salivary glands with BS-181 and performed immunostaining. MSL1 normally displays a well characterized and pronounced staining of the X chromosome *in vivo*. Because salivary glands are mitotically inactive, this strategy eliminates the effects of CDK7 on the cell cycle. Interestingly, immunofluorescence experiments revealed that BS-181 diminished MSL1 staining on the X chromosome after only 10 min (**Supplementary Fig. 4c**). Further incubation with BS-181 for 30 min completely abolished MSL1 X-chromosomal localization. In contrast, treatment with DRB, a CDK9-specific inhibitor with little or no effect on human CDK7 *in vitro*[12], also decreased the MSL1 coating of the X chromosome, such that it was restricted to discrete loci (**Fig. 3e**). Together, these data provided the following important insights: first, MSL1 chromatin binding is dynamic and sensitive to CDK7 inhibitor, and second, inhibition of CDK7 activity did not alter the MSL complex formation but resulted in substantially reduced MSL1 chromatin binding, exemplified by compromised X-chromosomal enrichment.

MSL1 interacts with the CAK subcomplex of TFIIH

The above experiments provide strong evidence of a functional interaction between MSL1 and CDK7. Next, we asked whether we could detect a biochemical interaction between MSL1 and CDK7. Indeed, transiently expressed FLAG-CDK7 in S2 cells coimmunoprecipitated endogenous MSL1 but not the closely related NSL1 (**Fig. 4a**). Consistently with this result, CDK7 also coimmunoprecipitated endogenous MSL1 and Mat1, but not MOF and MSL3, in S2 cells (**Fig. 4b**). In contrast, MSL1 IP pulled down endogenous Mat1 but not CDK7. It is possible that the antibody to MSL1 interferes with the interaction surfaces *in vivo* or that this binding is dynamic and difficult to capture. Therefore, to probe the MSL1-CDK7 interaction in more detail, we also performed a series of *in vitro* binding assays, using baculovirus-expressed recombinant proteins. We found that that *Drosophila* and mouse MSL1, but not MOF, directly interacted with human CDK7 (**Figs. 4c** and **5d**). Importantly, the MSL1-CDK7 interaction was enhanced in the presence of the entire CAK complex (CycH–Mat1–CDK7). To determine the relevance of this interaction in the recruitment of CDK7 to chromatin, we further performed heterologous protein tethering assays in S2 cells cotransfected with a construct carrying LacI-V5-tagged MSL1 and a plasmid carrying 256 copies of the LacOp sequence. ChIP–qPCR analysis showed that MSL1-lacI-V5 was efficiently recruited to the LacOp repeats but not

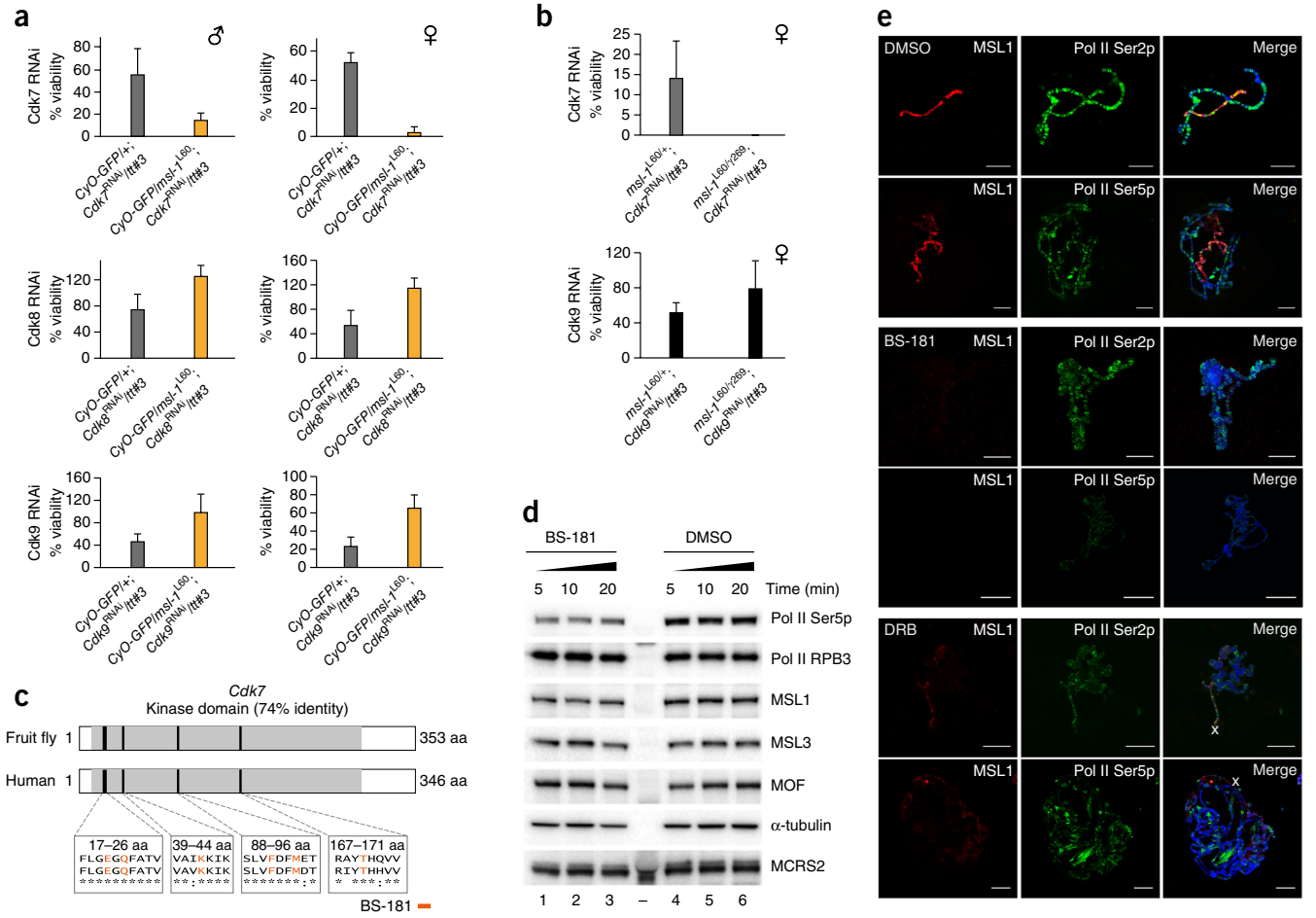

Figure 3 MSL1 and CDK7 show genetic interaction, and CDK7 inhibition causes decreased MSL1 localization on the X chromosome. (**a**) Relative percentages of viable adult male and female flies after *tub-Gal4*-driven dsRNAi knockdown of *Cdk7*, *Cdk8* and *Cdk9* in a wild-type (gray) or heterozygous *msl-1*^{L60} (yellow) background. Results are expressed as mean ± s.d. of at least three independent crosses (details in Source Data). The numbers of flies counted were 1,107 (*Cdk7*), 1,456 (*Cdk8*) and 2,208 (*Cdk9*). *tt#3*, recombinant chromosome *tub-Gal4*, *tub-Gal80*^{ts}. Details in Online Methods. (**b**) Relative percentages of viable adult female flies after *tub-Gal4*-driven dsRNA knockdown of *Cdk7* (top, gray bars) and *Cdk9* (bottom, black bars) in either heterozygous (*msl-1*$^{L60/+}$) or null-mutant (*msl-1*$^{L60/γ269}$) background. *msl-1*^{L60} heterozygous or homozygous flies not expressing *Cdk*^{dsRNAi} are internal controls with 100% viability. Error bars, s.d. of at least three independent crosses (details in Source Data). The numbers of flies counted were 796 (*Cdk7*) and 2,208 (*Cdk9*). *tt#3*, recombinant chromosome *tub-Gal4, tubGal80*^{ts}. Additional data are in **Supplementary Figure 4a**. Details in Online Methods. Source data for **a** and **b** are available online. (**c**) Evolutionary conservation of *Cdk7* in fruit flies and humans. The kinase domain (74% identity) is shaded in gray. Sequence alignment of regions marked in black is shown below. The binding sites and other key residues for the BS-181 inhibitor are indicated in orange. (**d**) Time course (5, 10 and 20 min) of DMSO (negative control) and BS-181 inhibitor (25 μM) treatment in S2 cells. Whole cell extracts were tested for bulk levels of Pol II Ser5p, Pol II RBP3, MSL1, MSL3 and MOF. α-tubulin and MCRS2 are loading controls. Uncropped gel images can be found in **Supplementary Data Set 1**. (**e**) Salivary glands dissected from third-instar larvae and treated with DMSO (negative control; top), BS-181 (50 μM; middle) or DRB (100 μM; bottom) for 30 min. Polytene chromosomal squashes are immunostained with antibodies against MSL1 (red), Pol II Ser5p (green) or Pol II Ser2p (green). Immunostaining with antibodies against Pol II Ser5p and Pol II Ser2p demonstrates the specificity of the two drugs. X, X chromosome; scale bars, 20 μm. Additional data in **Supplementary Figure 4b**,**c**.

to a sequence in the LacOp plasmid backbone. Ectopic tethering of MSL1 also led to the recruitment of endogenous MSL2, thus confirming that MSL1-lacI-V5 was functional. Interestingly, the presence of MSL1-lacI-V5 enhanced ectopic recruitment of endogenous CDK7 (**Fig. 4d**). In contrast, JIL-1, another kinase implicated in dosage compensation[37,38], scored negatively, thus indicating the specificity of the tethering assay. These results were consistent with those from a recent transcriptome analysis of JIL1-depleted cells, which has suggested that this kinase may be important for reinforcing rather than establishing active chromatin states[38].

Moreover, to test the interdependence of MSL1 and CDK7 at the chromatin level, we depleted MSL1 in S2 cells and performed ChIP–qPCR probing for CDK7 at MSL1 target loci. CDK7 chromatin binding, but not that of CDK9, was decreased after MSL1 depletion (**Fig. 4e**,**f**), thus suggesting that MSL1 plays a role in CDK7 chromatin recruitment to shared target genes in flies. Together, these experiments reinforce that MSL1 interacts with CDK7 and that MSL1 is important for optimal targeting of CDK7 to a subset of genes in *Drosophila*.

MSL1 is a phosphoprotein and a target of CDK7 *in vitro*

Phosphorylation of transcription factors is a well-established mechanism used by cells to regulate protein activity and DNA binding[39]. Because both *Drosophila* and mammalian MSL1 carry several SP/TP motifs, which have previously been identified as CDK targets[40,41], we hypothesized that MSL1 might undergo phosphorylation. Using multiple alignment tools, we compared MSL1 orthologs

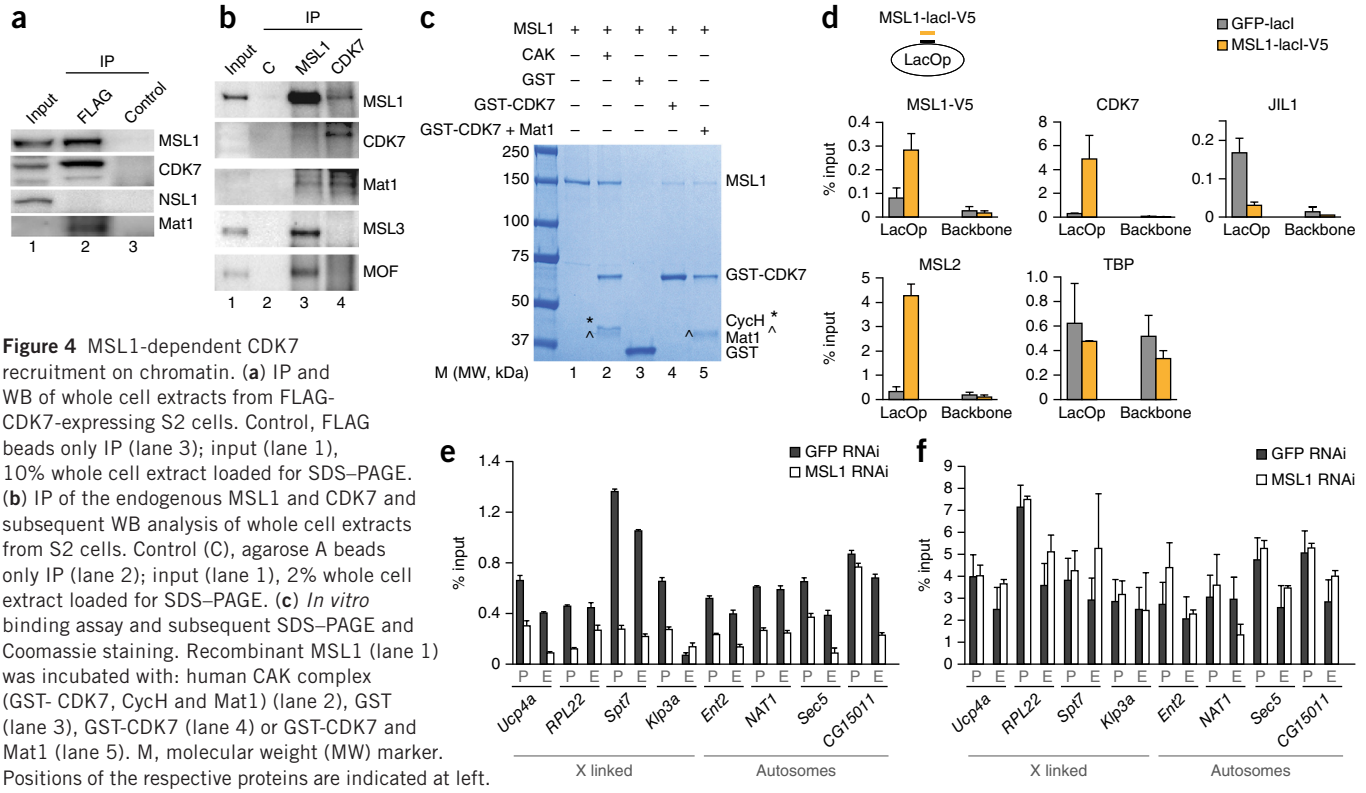

Figure 4 MSL1-dependent CDK7 recruitment on chromatin. (**a**) IP and WB of whole cell extracts from FLAG-CDK7-expressing S2 cells. Control, FLAG beads only IP (lane 3); input (lane 1), 10% whole cell extract loaded for SDS–PAGE. (**b**) IP of the endogenous MSL1 and CDK7 and subsequent WB analysis of whole cell extracts from S2 cells. Control (C), agarose A beads only IP (lane 2); input (lane 1), 2% whole cell extract loaded for SDS–PAGE. (**c**) *In vitro* binding assay and subsequent SDS–PAGE and Coomassie staining. Recombinant MSL1 (lane 1) was incubated with: human CAK complex (GST- CDK7, CycH and Mat1) (lane 2), GST (lane 3), GST-CDK7 (lane 4) or GST-CDK7 and Mat1 (lane 5). M, molecular weight (MW) marker. Positions of the respective proteins are indicated at left. (**d**) Heterologous protein tethering assay. GFP-lacI (gray) or MSL1-lacI-V5 (yellow) and a plasmid containing 256 copies of Lac operator arrays (LacOp) were transiently transfected into S2 cells. ChIP–qPCR was performed with primer pairs specific for Lac operator sequence (LacOp) or plasmid backbone (negative control). ChIP was performed with anti-V5 (for ectopic MSL1), anti-MSL2, anti-CDK7, anti-JIL1 and anti-TBP. Results are expressed as mean ± s.d. of three ChIP experiments prepared from independently dsRNA-transfected S2 cell cultures. (**e**,**f**) ChIP–qPCR analyses for CDK7 (top) and CDK9 (bottom) in S2 cells treated with dsRNA against GFP and MSL1. Immunoprecipitated DNA was amplified by qPCR with primer sets positioned at the promoter (P) and end (E) of indicated genes. Results are expressed as mean ± s.d. of three experiments with independent knockdown cell cultures performed on different days. Uncropped gel images of **a**–**c** can be found in **Supplementary Data Set 1**. Source data for **d**–**f** are available online.

across the *Drosophila* genus and found that five of these motifs cluster close (at a distance of ~100 amino acids (aa)) to the PEHE domain (**Fig. 5a**). We termed this region 'phosphorylation sites next to PEHE domain' (PNP). Our mass spectrometry analysis of MSL1 derived from S2 cells and *Drosophila* or mouse MSL1 *in vitro* phosphorylated by the CAK complex revealed multiple phosphorylation events occurring at SP/TP motifs (**Fig. 5b**, **Supplementary Fig. 5a** and **Supplementary Table 1**). Furthermore, when we cultured S2 cells with radiolabeled ATP and performed immunoprecipitation (IP) with antibodies recognizing two different regions of MSL1, we detected a radioactive signal with a molecular weight corresponding to the size of MSL1, thus indicating that MSL1 is phosphorylated *in vivo* (**Supplementary Fig. 5b**).

Our analysis suggested that MSL1 is phosphorylated at several residues. To address the functional importance of MSL1 phosphorylation, we generated and characterized *dm*MSL1-phosophomutant proteins. In the first derivative, MSL1-PH1, a single serine residue, Ser18, was mutated to alanine (S18A). Although, this residue is further away from the PNP region, interestingly, it is located in the highly conserved N-terminal part of MSL1 (ref. 42). Because phosphorylation sites are known to function both redundantly and cooperatively, in the second MSL1 derivative, MSL1-PH3, we mutated three PNP sites from threonines to alanines at positions 743, 747 and 751 (T743A, T747A and T751A) (**Fig. 5b**,**c**). Both MSL1-PH1 and MSL1-PH3 interacted with the CAK complex *in vitro*. We also included mouse MSL1 recombinant protein in our analysis, which also interacted with the CAK complex. We did not observe an interaction between the CAK complex and recombinant *Drosophila* MOF (**Fig. 5d**). Both MSL1-PH1 and MSL1-PH3 mutants, as well as $MSL1_{1-584}$ protein showed decreased phosphorylation by the CAK complex (**Fig. 5e**). The remaining phosphorylation signal detected in the mutants was probably a result of other residues in the MSL1 protein being phosphorylated in the *in vitro* assays.

To further characterize MSL1 phosphorylation, we generated a phosphoepitope-specific antibody that recognized only MSL1 phosphorylated at Ser18 (**Fig. 5f** and **Supplementary Fig. 6a**). This antibody immunoprecipitated and detected phosphorylated MSL1 in immunoprecipitated samples (**Fig. 5g**), thus indicating that MSL1 is indeed phosphorylated at Ser18 *in vivo*. We further verified the specificity of this antibody by monitoring the decrease in the detected IP signal in MSL1-depleted S2 cells (**Supplementary Fig. 6b**). Next, we sought to address whether a decrease in CDK7 would affect MSL1-PH1 levels *in vivo*. Because complete loss of CDK7 severely affects cell viability, we performed RNAi-mediated depletion of CDK7 for only 2 d, thus leading to an approximately 60% decrease in CDK7 (**Fig. 5h**), and subsequent IP with anti-MSL1-PH1. Interestingly, we observed a decrease in MSL1 bulk levels after CDK7 knockdown (**Fig. 5i**), which led to a concomitant decrease in the IP signal detected by antibodies to MSL1 and MSL1-PH1. Together, our data show that MSL1 is phosphorylated *in vivo* and that MSL1 protein levels are sensitive to CDK7 depletion.

Mutations of MSL1 phosphorylation sites cause male lethality

To address the relevance of the MSL1 phosphorylation *in vivo*, we generated flies carrying PH1- and PH3-mutant *msl-1* alleles

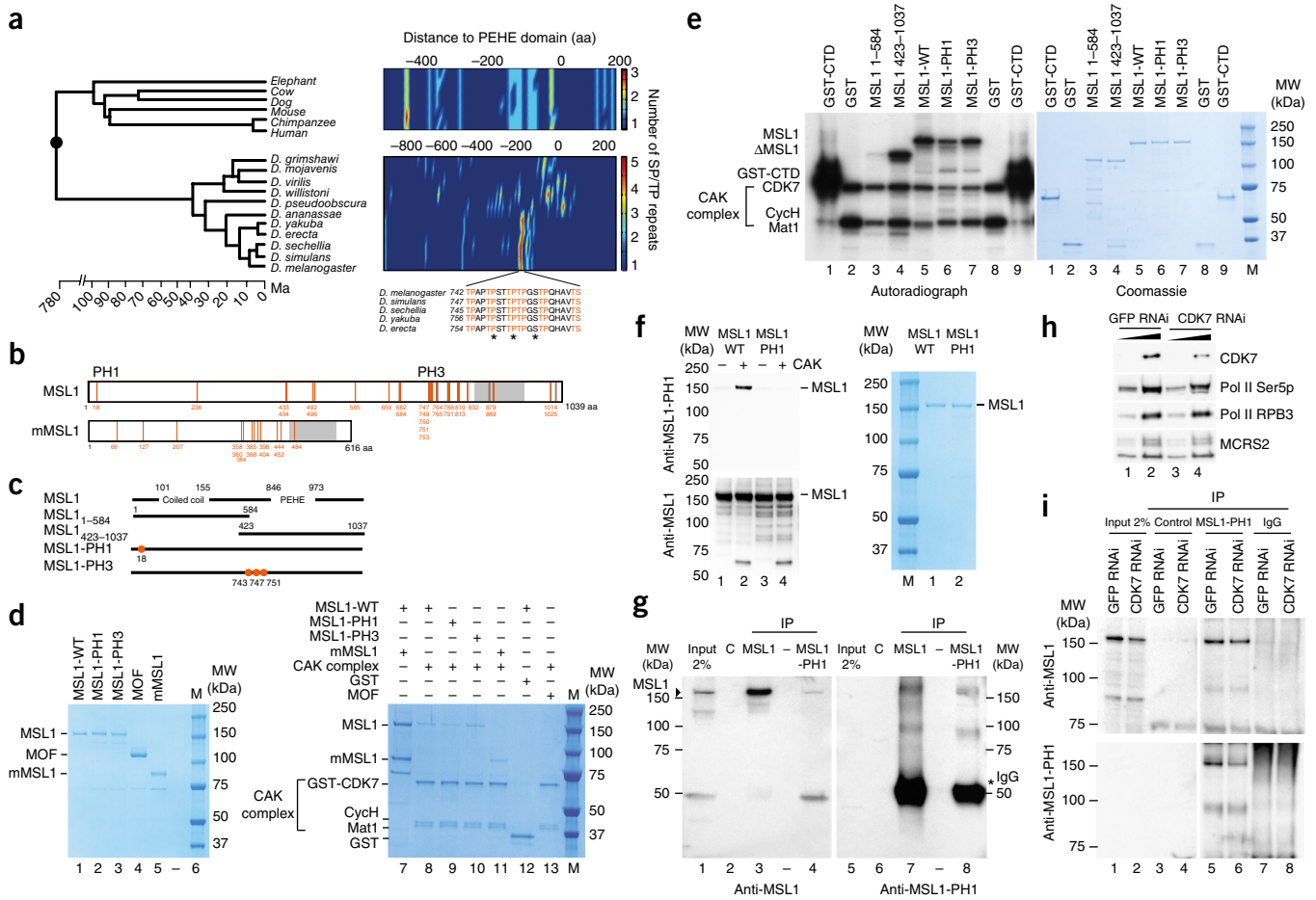

Figure 5 MSL1 is a phosphoprotein *in vivo* and can be phosphorylated by CDK7 *in vitro*. (**a**) Left, phylogenetic relationship among *Drosophila* (*D.*) species and mammals, dating to ~780 Ma. Black circle denotes a common ancestor. Right, protein sequences of MSL1 in 11 *Drosophila* species and 6 mammals, scanned for the presence of SP/TP sites in a window size of 10 aa. Asterisks indicate residues (MSL1-PH3) located in the highly conserved cluster of SP/TP repeats next to the PEHE domain. (**b**) Phosphorylated sites throughout FLAG-purified *Drosophila* and mouse MSL1 (mMSL1), mapped through MS/MS peptide sequencing. The positions of PH1 and PH3 are indicated in orange. Additional data are in **Supplementary Table 1**. (**c**) Illustration of *Drosophila* MSL1 derivatives used in this study. (**d**) *In vitro* binding assay and subsequent SDS–PAGE and Coomassie staining. Left, input showing quality control of proteins (lanes 1–5). Right, MSL1 WT, MSL1-PH1, MSL1-PH3, mMSL1 and MOF (lanes 8–11 and 13), copurified with the human CAK complex (GST-CDK7, CycH, Mat1) in GST pulldown assays. Recombinant GST (lane 12) is a negative control. Lane 7 shows input for MSL1 WT and mMSL1 coloaded in the same lane. (**e**) (Left) CAK-mediated *in vitro* phosphorylation of MSL1 derivatives illustrated in **c**. GST-CTD (lanes 1 and 9) and GST (lanes 2 and 8) are positive and negative controls, respectively. Right, Coomassie-stained gel, shown for protein quality control. (**f**) Left, *in vitro* phosphorylation assays with recombinant MSL1-WT and MSL1-PH1 phosphomutants incubated in the absence (lane 3 and 5) or presence (lane 4 and 6) of the CAK complex. WB analysis with anti-MSL1-PH1 (top) and anti-MSL1 (bottom). Right, Coomassie-stained gel for protein quality control. (**g**) WB analysis with anti-MSL1 (lanes 1–4) and anti-MSL1-PH1 (lanes 5–8). Samples are whole cell extracts from S2 cells, subjected to IP with anti-MSL1 (lanes 3 and 7) and anti-MSL1-PH1 (lanes 4 and 8), and subsequent SDS–PAGE. Control (C), IP with agarose A beads only (lanes 2 and 6); input, 2% whole cell extract; asterisks, IgG signal. (**h**) CDK7 depletion in S2 cells. Samples are whole cell extracts (30% and 100%) prepared from control RNAi (GFP) and CDK7 RNAi treated cells, subjected to WB analysis. MCRS2 is a loading control. (**i**) WB analysis with anti-MSL1 (top) and anti-MSL1-PH1 (bottom). Samples are whole cell extracts from S2 cells after RNAi depletion of GFP or CDK7, subjected to IP with anti-MSL1-PH1 (lanes 5 and 6). Agarose A beads (lanes 3 and 4) and preimmune serum (lanes 7 and 8) are control IPs. Input, 2% of whole cell extract. Uncropped gel images of **d** and **f**–**i** can be found in **Supplementary Data Set 1**.

under control of a UAS promoter. Using the *tub-Gal4* driver, we successfully overexpressed MSL1-PH1 and MSL1-PH3 in male and female flies (**Fig. 6a**,**b**). Interestingly, ectopic expression of either MSL1^{PH1} or MSL1^{PH3} did not rescue the decreased Pol II Ser5p observed in *msl-1*-null mutants (**Fig. 6a**,**b**). Moreover, in contrast to flies expressing wild-type *msl-1* (*msl-1*^{WT}), introduction of *msl-1*^{PH1} and *msl-1*^{PH3} to *msl-1*-null mutants (*msl-1*$^{L60/\gamma269}$) only partially rescued male lethality (**Fig. 6c**). These data confirm that phosphorylation of MSL1 at these sites is essential for its function in dosage compensation. Importantly, co-IP experiments of transiently expressed FLAG-tagged MSL1-PH1 and MSL1-PH3 proteins in male S2 cells showed that the MSL1 phosphomutants interacted with other MSL-complex members (**Fig. 6d**). Furthermore, both MSL1-PH1 and MSL1-PH3 derivatives were able to reconstitute the trimeric complex containing MSL1, MSL3 and MOF (**Fig. 6e**), thus indicating that mutation of these residues (S18 or T743 T747 T751) does not lead to misfolding of MSL1. To further interrogate and understand the male lethal phenotype, we performed immunostaining of polytene chromosomes from salivary glands of third-instar larvae expressing MSL1-PH1 and MSL1-PH3 in an *msl-1*-null background. We observed that in both cases, the MSL1 derivatives were unable to coat the X chromosome. Consistently with the loss of

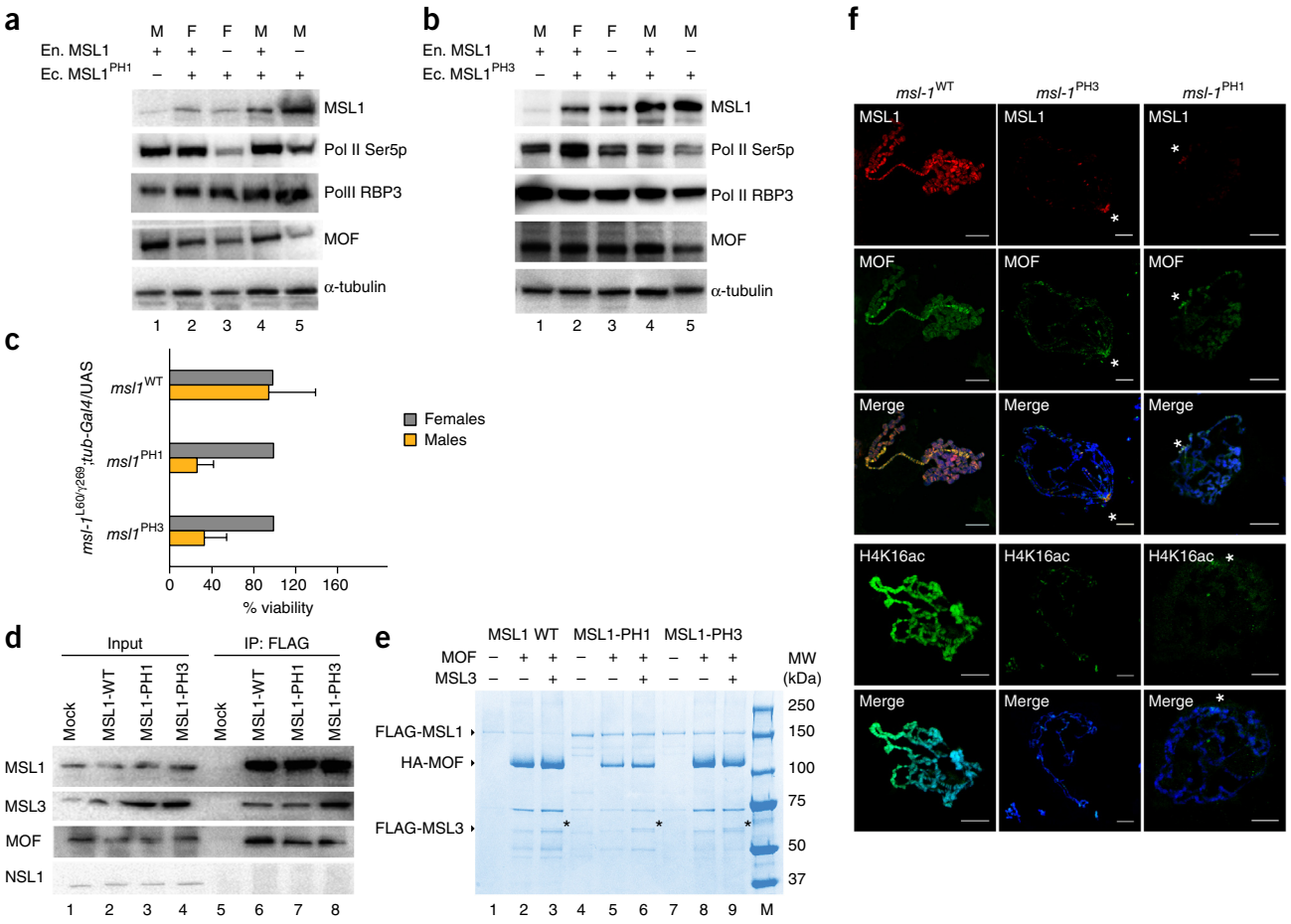

Figure 6 MSL1 phosphomutants show a decrease in Pol II Ser5p and do not rescue male viability. (**a**) WB analysis of whole cell extracts from *Drosophila* heads. Samples are from wild-type *Drosophila* males (M) (lane 1) and females (F) with *tub-Gal4*-driven UAS-*msl-1*^{PH1} expression in an *msl-1*-null heterozygous background (lane 2) and females with *tubGal4*-driven UAS-*msl-1*^{PH1} expression in an *msl-1*-null background (lane 3), *msl-1*-null heterozygous (lane 4) and *msl-1*-null homozygous 'escaper' mutant males expressing UAS-*msl-1*^{PH1} (lane 5). α-tubulin is a loading control. The presence or absence of endogenous (En.) or ectopically expressed (Ec.) MSL1 is indicated. (**b**) As in **a**, but with UAS-*msl-1*^{PH3}. (**c**) Viability (expressed as a relative percentage) of male (yellow) and female (gray) adult flies after *tub-Gal4*-induced ectopic overexpression of the MSL1^{PH3} and MSL1^{PH1} (*tub-Gal4*/UAS-*msl-1*$^{PH3, PH1}$) in an *msl-1*-null background (*msl-1*$^{L60/γ269}$). Flies not expressing *msl-1*$^{PH3, PH1}$ (*TM6-Tb*/UAS-*msl-1*^{PH3,PH1}) in heterozygous *msl-1*-null-mutant flies (*msl-1*^{L60} or *msl-1*^{γ269}*/CyO-GFP*) or wild-type background were used as internal controls with 100% viability. The numbers of flies counted were 1,390 (*msl-1*^{WT}), 1,798 (*msl-1*^{PH1}) and 1,449 (*msl-1*^{PH3}). Results are expressed as mean ± s.d. of at least three independent crosses (details in Online Methods and source data). (**d**) FLAG IP of whole cell extracts from S2 cells transiently transfected to express FLAG-MSL1-WT, MSL1-PH1 or MSL1-PH3. A mock transfection was used as a negative control IP. Input, lanes 1–4. (**e**) Reconstitution of MSL1-WT, MSL1-PH1 and MSL1-PH3 with MOF alone (lanes 2, 5 and 8) or MOF and MSL3 (lanes 3, 6 and 9) with baculovirus-expressed proteins. Asterisk, MSL3 band. FLAG-tagged MSL1, MSL1-PH1, MSL1-PH3 or MSL3 were used, and HA-tagged MOF was used as a bait. A Coomassie-stained SDS–PAGE gel is shown. (**f**) Polytene squashes from male larvae expressing *tub-Gal4*/UAS-*msl-1*^{WT}, *tub-Gal4*/UAS-*msl-1*^{PH3} or *tub-Gal4*/UAS-*msl-1*^{PH1} in an *msl-1*-null background. Immunofluorescence images with anti-MSL1 (red), anti-MOF (green) and anti-H4K16ac (green, bottom) are shown. Asterisk, chromocenter; scale bars, 20 μm. Uncropped gel images of **a**, **b** and **d** can be found in **Supplementary Data Set 1**.

MSL1, endogenous MOF and H4K16ac signals from the X chromosome were dramatically decreased (**Fig. 6f**). Together, our results suggest that loss of phosphorylation does not disrupt formation of the MSL complex but precludes its chromatin binding, thus suggesting new functional relevance of these residues.

DISCUSSION

Our study uncovers a new and conserved role of promoter-associated MSL1, whose regulatory scope extends beyond its canonical role in X-chromosome dosage compensation in *Drosophila*. Our earlier work has suggested that X-chromosomal enrichment of the *Drosophila* MSL complex contributes to enhanced Pol II recruitment, most probably because of MOF-mediated chromosome-wide H4K16ac deposition[23]. Using a combination of ChIP experiments, genetic and biochemical approaches, we established a functional link between MSL1 and the CDK7–CAK complex and demonstrated that the presence of MSL1 is important for efficient phosphorylation of Pol II Ser5 during early transcriptional events. We also determined that MSL1 is phosphorylated *in vivo* and showed that MSL1 phosphorylation is important for its X-chromosomal targeting and dosage compensation. Consequently, our work uncovered a new MSL1 PTM that plays an important role in modulating transcription: on the one hand, MSL1 is required for optimal CDK7 recruitment at target genes; on the other hand, CDK7 phosphorylates MSL1 *in vitro*, and MSL1-phosphorylation sites are required for the MSL-complex function *in vivo*. How these two mechanisms are linked is currently unclear and requires further

exploration. Importantly, because MSL1 can be phosphorylated at many sites, it is also possible that the regulatory landscape is quite complex, and several kinases may regulate the function of the MSL-complex members during different stages of transcription *in vivo*.

Although MSL1 is enriched at TSSs of both X-linked and autosomal genes in male and female *D. melanogaster* flies and in *D. virilis*, the functional relevance of this binding has been largely unexplored. Using fly genetics, we found that the presence of MSL1 is also biologically relevant in females, although their unaffected viability after MSL1 loss initially suggested otherwise. Interestingly, we observed enhanced lethality of both females and males when MSL1 depletion was accompanied by removal of CDK7 (**Fig. 3a**,**b**). These results indicate that MSL1 is functionally engaged in the regulatory pathway of CAK-associated CDK7, and only when levels of both MSL1 and CDK7 are decreased did the importance of MSL1 become apparent in females. We hypothesize that the more moderate effects of MSL1 loss in female flies may have been due to several reasons. First, the total amount of the MSL1 protein is substantially lower in female flies[28,43], thus suggesting a limited scope of its regulatory action. Second, MSL1 chromatin binding in female flies is restricted to TSSs, where loss of MSL1-mediated MOF recruitment and H4K16ac may potentially be compensated by the NSL complex, which is often recruited to the same genomic loci[44–46]. Second, the MSL holocomplex encompassing all MSL proteins and roX long noncoding RNAs assembles in only male flies[47,48], where it coats the X chromosome and increases the expression of most X-linked genes. Male lethality in the absence of the MSL complex is thus a consequence of severe transcriptional downregulation of the entire X chromosome[29], as well as of directly and indirectly enhanced misregulation of autosomal genes. This global cascade effect naturally does not apply to females, in which dosage compensation does not occur.

The observations that there is a genetic interaction between MSL1 and CDK7, and that MSL1 is a phosphoprotein containing several potential phosphorylation sites suggest a new mechanism by which MSL1 exerts regulatory functions. Here we showed that CDK7 phosphorylates MSL1 *in vitro* and that inhibition of CDK7 by a small-molecule inhibitor severely compromises MSL1 chromosomal localization. These data suggest both dynamic phosphorylation-dependent MSL1 localization and the interesting possibility that there may be distinct pools of MSL1 in the nucleus contributing to transcription regulation. It also remains unknown whether other kinases also modify MSL1, thus further modulating its regulatory properties during transcription and at different cell-cycle stages. Future studies are needed to distinguish the various possibilities and to unravel how phosphorylation of MSL1 affects transcription regulation *in vivo*.

Given the strong genetic interaction observed between MSL1 and CDK7, our observation of modest and substoichiometric biochemical interaction between endogenous MSL1 and CDK7 was at first puzzling. Nevertheless, from the combined results obtained by *in vitro* pulldown assays, *in vivo* tethering assays and ChIP analysis, we speculate that the MSL1-CDK7 interaction is likely to be dynamic. Treatment with a CDK7 inhibitor supported this hypothesis (**Fig. 3e** and **Supplementary Fig. 4c**). Because several proteins interact with CDK7–TFIIH[49], we suggest that by using several cofactors such as MSL1, the CDK7–TFIIH complex might achieve specificity for particular sets of target genes. This exciting possibility needs to be further explored in the future.

Interestingly, expression of MSL1 phosphomutants (MSL1^{PH1} S18A and MSL1^{PH3} T743A T747A T751A) was not sufficient to rescue the Pol II Ser5p defect observed in *msl-1*-null mutants (**Fig. 6c**), thus reinforcing that these residues may play an important role in cross-talk with CDK7 or other kinases *in vivo*. A functional interdependence between CDK7 and MSL1 may provide a plausible explanation for the decrease in Pol II Ser5p after MSL1 depletion. However, the mechanistic insights of the phosphorylation of MSL1 and its contribution to the catalytic activity of CDK7 remain unresolved and encourage further analysis.

Our study uncovers the importance of MSL1 phosphorylation, because MSL1 phosphomutants failed to properly establish H4K16ac on the male X chromosome, thereby resulting in male lethality (**Fig. 6c**,**f**). Although the phosphorylation sites studied with the MSL1^{PH1} and MSL1^{PH3} point mutants were in different regions of the protein, both mutants failed to restore Pol II Ser5p levels in the *msl-1*-null background and were unable to coat the X chromosome with MOF and H4K16ac. It is also possible that, through three-dimensional protein folding, these residues may cooperate in providing an interaction platform for optimal signal integration. Interestingly, a similar phenotype of H4K16ac mistargeting to the chromocenter has also been observed in roX mutants[50], thus suggesting that MSL1 phosphorylation may be important for roX RNA integration.

Acetylation of MSL3 and MOF[51–55] and ubiquitination of MSL1 and MSL2 (refs. 22,56–58) has recently been reported. Our study now adds to this actively growing list of PTMs of MSLs and raises the possibility that these modifications may contribute to diversifying MSL function at the X chromosome versus autosomes, as well as promoters versus gene bodies, in flies and mammals. Thus, the MSL complex may exist in several forms that are decorated with various PTMs and can thereby achieve functional specificity at target genes.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. The RNA-seq data have been deposited in the ArrayExpress database under accession code E-MTAB-3434. The ChIP–seq data sets have been deposited in the Gene Expression Omnibus database under accession code GSE61340.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank all members of the Akhtar laboratory for helpful discussions. We especially thank K. Lam and B. Sheikh for critical reading of the manuscript and helpful suggestions. We would also like to thank K. Adelman (NIEHS), K. Johansen (Iowa State University), J. Kadonaga (UCSD), J.T. Lis (Cornell University) and B. Suter (University of Bern) for kindly providing antibodies. This work was supported by an EU-funded EpiGeneSys awarded to A.A. and N.M.L., and DFG-BIOSS II, CRC992, CRC1140 and CRC746, awarded to A.A. L.M.S. acknowledges funding from the DFG and NIH.

AUTHOR CONTRIBUTIONS

S.C. performed ChIP–seq for MSL1 *D. virilis*, ChIP–qPCR, RNA-seq, co-IP assays, immunofluorescence microscopy and WB analyses; H.H. purified baculovirus-expressed proteins and performed binding assays, kinase assays, WB analyses and antibody characterization; M.S. performed genetic crosses, ChIP–qPCR, antibody characterization and WB analyses; T. Chelmicki performed and analyzed mammalian ChIP–seq and ChIP–qPCR, and analyzed WB data; P.G. performed genetic crosses, analysis and quantification of the phenotypes; V.P., F.D., F.T.C., F.R., W.W., N.M.L. and L.M.S. performed bioinformatics analyses and contributed to the corresponding manuscript sections; T.M. guided the development and implementation of deepTools for NGS analysis and quality controls; P.D. performed evolutionary analysis of MSL1; T. Conrad and S.R. performed *dm*MSL1 ChIP–seq experiments; G.M. performed and analyzed the mass spectrometry data; A.A., S.C., M.S., T. Chelmicki, P.G. and V.P. designed experiments and analyzed the data; A.A., T. Chelmicki and M.S. prepared the manuscript.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Orphanides, G. & Reinberg, D. A unified theory of gene expression. *Cell* **108**, 439–451 (2002).
2. Proudfoot, N.J., Furger, A. & Dye, M.J. Integrating mRNA processing with transcription. *Cell* **108**, 501–512 (2002).
3. Hsin, J.P., Xiang, K. & Manley, J.L. Function and control of RNA polymerase II C-terminal domain phosphorylation in vertebrate transcription and RNA processing. *Mol. Cell. Biol.* **34**, 2488–2498 (2014).
4. Hsin, J.P. & Manley, J.L. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* **26**, 2119–2137 (2012).
5. Hsin, J.P., Sheth, A. & Manley, J.L. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3′ end processing. *Science* **334**, 683–686 (2011).
6. Buratowski, S. Progression through the RNA polymerase II CTD cycle. *Mol. Cell* **36**, 541–546 (2009).
7. Fisher, R.P. Secrets of a double agent: CDK7 in cell-cycle control and transcription. *J. Cell Sci.* **118**, 5171–5180 (2005).
8. Glover-Cutter, K. *et al.* TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Mol. Cell. Biol.* **29**, 5455–5464 (2009).
9. Phatnani, H.P. & Greenleaf, A.L. Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* **20**, 2922–2936 (2006).
10. Egloff, S. *et al.* Serine-7 of the RNA polymerase II CTD is specifically required for snRNA gene expression. *Science* **318**, 1777–1779 (2007).
11. Egloff, S. *et al.* The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *J. Biol. Chem.* **285**, 20564–20569 (2010).
12. Larochelle, S. *et al.* Cyclin-dependent kinase control of the initiation-to-elongation switch of RNA polymerase II. *Nat. Struct. Mol. Biol.* **19**, 1108–1115 (2012).
13. Conrad, T. & Akhtar, A. Dosage compensation in *Drosophila melanogaster*: epigenetic fine-tuning of chromosome-wide transcription. *Nat. Rev. Genet.* **13**, 123–134 (2011).
14. Straub, T., Zabel, A., Gilfillan, G.D., Feller, C. & Becker, P.B. Different chromatin interfaces of the *Drosophila* dosage compensation complex revealed by high-shear ChIP-seq. *Genome Res.* **23**, 473–485 (2013).
15. Keller, C.I. & Akhtar, A. The MSL complex: juggling RNA-protein interactions for dosage compensation and beyond. *Curr. Opin. Genet. Dev.* **31**, 1–11 (2015).
16. Kind, J. *et al.* Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in *Drosophila*. *Cell* **133**, 813–828 (2008).
17. Chelmicki, T. *et al.* MOF-associated complexes ensure stem cell identity and Xist repression. *eLife* **3**, e02024 (2014).
18. Li, X. *et al.* The histone acetyltransferase MOF is a key regulator of the embryonic stem cell core transcriptional network. *Cell Stem Cell* **11**, 163–178 (2012).
19. Ravens, S. *et al.* Mof-associated complexes have overlapping and unique roles in regulating pluripotency in embryonic stem cells and during differentiation. *eLife* **3**, e02104 (2014).
20. Taylor, G.C., Eskeland, R., Hekimoglu-Balkan, B., Pradeepa, M.M. & Bickmore, W.A. H4K16 acetylation marks active genes and enhancers of embryonic stem cells, but does not alter chromatin compaction. *Genome Res.* **23**, 2053–2065 (2013).
21. Wang, Z. *et al.* Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019–1031 (2009).
22. Hallacli, E. *et al.* Msl1-mediated dimerization of the dosage compensation complex is essential for male X-chromosome regulation in *Drosophila*. *Mol. Cell* **48**, 587–600 (2012).
23. Conrad, T., Cavalli, F.M., Vaquerizas, J.M., Luscombe, N.M. & Akhtar, A. *Drosophila* dosage compensation involves enhanced Pol II recruitment to male X-linked promoters. *Science* **337**, 742–746 (2012).
24. Gu, W., Szauter, P. & Lucchesi, J.C. Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of *Drosophila melanogaster*. *Dev. Genet.* **22**, 56–64 (1998).
25. Conrad, T. *et al.* The MOF chromobarrel domain controls genome-wide H4K16 acetylation and spreading of the MSL complex. *Dev. Cell* **22**, 610–624 (2012).
26. Beverley, S.M. & Wilson, A.C. Molecular evolution in *Drosophila* and the higher Diptera II. A time scale for fly evolution. *J. Mol. Evol.* **21**, 1–13 (1984).
27. Russo, C.A., Takezaki, N. & Nei, M. Molecular phylogeny and divergence times of drosophilid species. *Mol. Biol. Evol.* **12**, 391–404 (1995).
28. Chang, K.A. & Kuroda, M.I. Modulation of MSL1 abundance in female *Drosophila* contributes to the sex specificity of dosage compensation. *Genetics* **150**, 699–709 (1998).
29. Palmer, M.J. *et al.* The male-specific lethal-one (*msl-1*) gene of *Drosophila melanogaster* encodes a novel protein that associates with the X chromosome in males. *Genetics* **134**, 545–557 (1993).
30. Belote, J.M. & Lucchesi, J.C. Male-specific lethal mutations of *Drosophila melanogaster*. *Genetics* **96**, 165–186 (1980).
31. Riedl, T. & Egly, J.M. Phosphorylation in transcription: the CTD and more. *Gene Expr.* **9**, 3–13 (2000).
32. Larochelle, S., Pandur, J., Fisher, R.P., Salz, H.K. & Suter, B. Cdk7 is essential for mitosis and for *in vivo* Cdk-activating kinase activity. *Genes Dev.* **12**, 370–381 (1998).
33. Schwartz, B.E., Larochelle, S., Suter, B. & Lis, J.T. Cdk7 is required for full activation of *Drosophila* heat shock genes and RNA polymerase II phosphorylation *in vivo*. *Mol. Cell. Biol.* **23**, 6876–6886 (2003).
34. Ganuza, M. *et al.* Genetic inactivation of Cdk7 leads to cell cycle arrest and induces premature aging due to adult stem cell exhaustion. *EMBO J.* **31**, 2498–2510 (2012).
35. Ali, S. *et al.* The development of a selective cyclin-dependent kinase inhibitor that shows antitumor activity. *Cancer Res.* **69**, 6208–6215 (2009).
36. Kwiatkowski, N. *et al.* Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* **511**, 616–620 (2014).
37. Jin, Y., Wang, Y., Johansen, J. & Johansen, K.M. JIL-1, a chromosomal kinase implicated in regulation of chromatin structure, associates with the male specific lethal (MSL) dosage compensation complex. *J. Cell Biol.* **149**, 1005–1010 (2000).
38. Regnard, C. *et al.* Global analysis of the relationship between JIL-1 kinase and transcription. *PLoS Genet.* **7**, e1001327 (2011).
39. Whitmarsh, A.J. & Davis, R.J. Regulation of transcription factor function by phosphorylation. *Cell. Mol. Life Sci.* **57**, 1172–1183 (2000).
40. Larochelle, S. *et al.* T-loop phosphorylation stabilizes the CDK7-cyclin H-MAT1 complex *in vivo* and regulates its CTD kinase activity. *EMBO J.* **20**, 3749–3759 (2001).
41. Bisteau, X. *et al.* CDK4 T172 phosphorylation is central in a CDK7-dependent bidirectional CDK4/CDK2 interplay mediated by p21 phosphorylation at the restriction point. *PLoS Genet.* **9**, e1003546 (2013).
42. Li, F., Parry, D.A. & Scott, M.J. The amino-terminal region of *Drosophila* MSL1 contains basic, glycine-rich, and leucine zipper-like motifs that promote X chromosome binding, self-association, and MSL2 binding, respectively. *Mol. Cell. Biol.* **25**, 8913–8924 (2005).
43. Palmer, M.J., Richman, R., Richter, L. & Kuroda, M.I. Sex-specific regulation of the male-specific lethal-1 dosage compensation gene in *Drosophila*. *Genes Dev.* **8**, 698–706 (1994).
44. Mendjan, S. *et al.* Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Mol. Cell* **21**, 811–823 (2006).
45. Raja, S.J. *et al.* The nonspecific lethal complex is a transcriptional regulator in *Drosophila*. *Mol. Cell* **38**, 827–841 (2010).
46. Lam, K.C. *et al.* The NSL complex regulates housekeeping genes in *Drosophila*. *PLoS Genet.* **8**, e1002736 (2012).
47. Beckmann, K., Grskovic, M., Gebauer, F. & Hentze, M.W. A dual inhibitory mechanism restricts msl-2 mRNA translation for dosage compensation in *Drosophila*. *Cell* **122**, 529–540 (2005).
48. Kelley, R.L., Wang, J., Bell, L. & Kuroda, M.I. *Sex lethal* controls dosage compensation in *Drosophila* by a non-splicing mechanism. *Nature* **387**, 195–199 (1997).
49. Thomas, M.C. & Chiang, C.M. The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41**, 105–178 (2006).
50. Menon, D.U. & Meller, V.H. Imprinting of the Y chromosome influences dosage compensation in roX1 roX2 *Drosophila melanogaster*. *Genetics* **183**, 811–820 (2009).
51. Kadlec, J. *et al.* Structural basis for MOF and MSL3 recruitment into the dosage compensation complex by MSL1. *Nat. Struct. Mol. Biol.* **18**, 142–149 (2011).
52. Buscaino, A., Legube, G. & Akhtar, A. X-chromosome targeting and dosage compensation are mediated by distinct domains in MSL-3. *EMBO Rep.* **7**, 531–538 (2006).
53. Yang, C., Wu, J., Sinha, S.H., Neveu, J.M. & Zheng, Y.G. Autoacetylation of the MYST lysine acetyltransferase MOF protein. *J. Biol. Chem.* **287**, 34917–34926 (2012).
54. McCullough, C.E. & Marmorstein, R. Molecular basis for histone acetyltransferase regulation by binding partners, associated domains, and autoacetylation. *ACS Chem. Biol.* **11**, 632–642 (2016).
55. Lu, L. *et al.* Modulations of hMOF autoacetylation by SIRT1 regulate hMOF recruitment and activities on the chromatin. *Cell Res.* **21**, 1182–1195 (2011).
56. Wu, L., Zee, B.M., Wang, Y., Garcia, B.A. & Dou, Y. The RING finger protein MSL2 in the MOF complex is an E3 ubiquitin ligase for H2B K34 and is involved in crosstalk with H3 K4 and K79 methylation. *Mol. Cell* **43**, 132–144 (2011).
57. Lai, Z. *et al.* Msl2 is a novel component of the vertebrate DNA damage response. *PLoS One* **8**, e68549 (2013).
58. Villa, R. *et al.* MSL2 combines sensor and effector functions in homeostatic control of the *Drosophila* dosage compensation machinery. *Mol. Cell* **48**, 647–654 2012).
59. Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).

ONLINE METHODS

ChIP–seq read mapping. For ChIP–seq samples of MSL1 in *D. melanogaster* (male, female, *mof*^{2} mutants) and male *D. virilis*, reads were aligned to the dm3 assembly (*D. melanogaster*) and the *D. virilis* scaffolds (D. vir1.2) with Bowtie version 2.0.5 with default parameters. Reads that mapped multiple times to the reference genome were kept, as recommended. For the processing of the aligned reads and their normalization, we used the deepTools package (https://deeptools.readthedocs.io/). Briefly, reads with a mapping quality <5 were discarded, and remaining reads were extended to a fragment length of 200 bp. The genome was divided into 10-bp bins, and overlapping reads were counted. To compute $\log_2$ fold changes between ChIP and input samples, the larger sample was scaled to the smaller one. The scaling factor for each comparison was estimated with the signal-extraction method proposed by Diaz and colleagues[60].

Peak calling on ChIP–seq samples. For *D. melanogaster*, MACS2 (version 2.0.10) was used with the following command: macs2 callpeak –f BAM –g 158736537–qvalue 0.01–keep-dup all–broad –t ChIP –c Input. For *D. virilis*, MACS2 (version 2.0.10) was used on reads assigned to those scaffolds associated with known Muller elements (scaffolds 10322, 10324, 12472, 12723, 12726, 12758, 12822, 12823, 12855, 12875, 12928, 12930, 12932, 12954, 12963, 12970, 13042, 13047, 13049, 13052, 13246 and 13324, according to http://insects.eugenes.org/species/data/dros-synteny-data/muller-elements-output/dvir-dmel-segments-r5.html/). The effective genome size was adjusted to 146,202,119 bp, which equals the number of base pairs of the aforementioned scaffold with nonzero coverage in the input-sequencing file. Peaks with a Q value >15 were used for downstream analyses.

Definition of target genes in ChIP–seq. Target genes were defined by an overlap of significant binding sites (peaks) for the respective ChIP sample with promoter regions (defined as 150 bp upstream of TSSs in *D. melanogaster* and *D. virilis*) annotated genes. Nontarget genes had no overlap of any ChIP–seq peak in either female or male samples with their promoter or gene-body regions.

Western blot analysis and antibodies. WB analyses were performed with whole cell extracts. The Invitrogen NuPAGE precast gel system was used for SDS–PAGE. The 4–12% gradient or 12% Bis-Tris gels (Invitrogen) were loaded with samples supplemented with Roti-Load reducing sample buffer (Roth). After the protein transfer, membranes were blocked in blocking solution composed of PBS containing 5% nonfat dry milk or 5% BSA and 0.3% Tween-20 and incubated with primary antibodies. Next, membranes were washed and incubated with suitable HRP-coupled secondary antibodies. Secondary antibodies used were anti-mouse IgG HRP (NXA931), anti-rat IgG HRP (NA935V) and anti-rabbit IgG HRP (NA934) from Sigma. Protein signals were visualized with Lumi-Light Plus Western Blotting Substrate with a Gel Doc XR+ System. The following primary antibodies were used: anti-Pol II RBP3 (kind gift from K. Adelman), anti-CDK7 (kind gift from J.T. Lis), anti-CDK7 (kind gift from B. Suter), anti-TBP (kind gift from J. Kadonaga), anti-*dm*JIL kinase (kind gift from K. Johansen), rabbit and rat anti-*dm*NSL1 and anti-*dm*NSL3 (previously characterized in ref. 45), anti-Pol II (N-20 X, sx-899 X; Santa Cruz) anti-Pol II Ser5p (Ab5131, Abcam), anti-Pol II Ser2p (Ab5095, Abcam), anti-CDK7 (C7089, Sigma), anti-α-tubulin (Ab125267, Abcam), anti-MOF (A3000992A; Bethyl), anti-MSL2 (HPA003413; Sigma), anti-GAPDH (A300-639A; Bethyl), anti-H2B (07-371, Millipore) anti-H2Bub (05-1312, Millipore), anti-H3 (Ab1791, Abcam), anti-H3K36me3 (Ab9050, Abcam), anti-H3K4me3 (Ab8580, Abcam), anti-H4 (Ab31830, Abcam) and anti-H4K16ac (07-329, Millipore). For commercial antibodies, validation is provided on the manufacturers' websites. For anti-mMSL1 (previously validated in ref. 17), the final bleed was used in the experiments. The antibody to *dv*MSL1 was raised against the GST-fused N-terminal fragment (265 aa) of the *dv*MSL1 expressed with the PGEX-6P1 vector. Cloning details are available upon request. The antibody against the C-terminal fragment of MSL1 was a kind gift from P. Heun. The phosphospecific anti-MSL1-PH1 was raised against the peptide KRANYLES(p)PYPHI by Eurogentec. The original blots can be found in **Supplementary Data Set 1**.

***Drosophila* RNAi knockdown experiments.** dsRNA was prepared according to the DRSC protocol (http://www.flyrnai.org/DRSC-PRS.html/). Primers are listed in the **Supplementary Note**. Reactions were performed per the DRSC protocol (http://www.flyrnai.org/DRSC-PRR.html/).

Preparation of protein extracts from adult *Drosophila* heads for WB analysis. Twenty adult male or female flies aged 12–24 h after eclosion were collected and placed on ice. Flies were decapitated, and heads were homogenized in 100 µl 2× Roti-Load reducing sample buffer (Roth) with antifoam A (Sigma) and Roche protease-inhibitor cocktail. Homogenates were incubated for 5 min at 95 °C, sonicated for 10 min and centrifuged for 5 min (16,000*g*). Supernatants were used for SDS–PAGE or stored at −20 °C.

RNA and genomic-DNA isolation and real-time PCR. RNA and genomic DNA extraction was performed with an Allprep DNA/RNA kit (Qiagen). Briefly, 300 ng of total RNA was used for reverse-transcription reactions. Validation of real-time PCR was performed with SYBR-Green PCR master mix (Roche) and an ABI7500 real-time PCR thermocycler (Applied Biosystems).

Tethering assays and ChIP–qPCR experiments for *Drosophila*. Tethering experiments were performed as previously described[61]. Briefly, a pLacOp plasmid carrying 256 bp of LacOp sequences was cotransfected with a pMT-MSL1-LacI-V5 plasmid. Schneider S2 cells were grown at 25 °C in Schneider's *Drosophila* medium (Serva). Cells were transfected with FuGENE6 Transfection Reagent (Roche), and after 48 h, chromatin was isolated as previously described[45]. Chromatin sonication was performed with a Branson 250 sonifier (40 pulses; intensity, 5). Equal amounts of chromatin were used for all ChIP experiments. Purified DNA was subjected to qPCR amplification (Applied Biosystems). Input was used for normalization.

Preparation and immunoprecipitation of S2 whole cell extracts. Total cell extracts were prepared from S2 cells (DRSC) as previously described[22]. Briefly, 1×10^7 cells were harvested and washed with ice-cold PBS supplemented with PhosStop (Roche) and resuspended in 1 ml of HMG K75 buffer (25 mM HEPES-NaOH, pH 7.6, 12.5 mM $MgCl_2$, 10% glycerol, 0.5% Tween-20, 75 mM KCl and protease-inhibitor cocktail (Roche) with PhosStop). Samples then were subjected to three freeze-thaw cycles and then to benzonase treatment (Sigma). Subsequently, samples were pelleted, and the supernatant was collected. Preincubation of the extract with Protein A/G–agarose bead mix (Roche) was used for preclearing. For MSL1 and CDK7 IP experiments $2{,}5 \times 10^6$ cells were used for extract preparation. For the anti-MSL1-PH1 IP experiments, 1×10^7 cells were lysed in HMG K150 buffer. As a negative control, Protein A/G–agarose beads (Roche) were used.

Preparation and immunoprecipitation of S2 whole cell extracts after transient transfection. Whole cell extracts were prepared from S2 cells as previously described[22]. Briefly, harvested cells were washed with ice-cold PBS and resuspended in 1 ml of HMG150 buffer (25mM HEPES-NaOH, pH 7.6, 12.5 mM $MgCl_2$, 10% glycerol, 0.2% Tween-20, 150 mM KCl and Roche protease-inhibitor-cocktail tablet). Samples were subjected to three freeze-thaw cycles and pelleted, and then the supernatant was collected. Anti-FLAG IP was performed with 30 µl of FLAG-agarose bead resin (Sigma).

Immunofluorescence and confocal microscopy. Preparation of polytene chromosomes from third-instar larvae in different *Drosophila* strains was performed as previously described[62]. Briefly, *Drosophila* polytene chromosome spreads were incubated with primary antibodies (1:100 in PBS, 0.1% Triton-X-100 and 1% milk). Images were obtained with a Leica Sp5 (Leica Microsystems) with an apochromat 1.32-NA oil-immersion objective.

For inhibitor experiments, salivary glands from wild-type male larvae were treated with 50 µM BS-181 or 100 M DRB for 10 or 30 min. Control experiments were performed with DMSO. Primary and secondary antibodies were used in 1:100 and 1:200 dilutions, respectively. Details on the antibodies used are provided in the 'Western blot and antibodies' section.

Cloning and DNA constructs. pMT-GFP-LacI was obtained by cloning of a GFP-LacI PCR product into a modified pMT/V5 hygro vector, as previously described[61]. The MSL1 PCR product was cloned via the XhoI and SacII sites of the pMT-LacI_V5 vector. The pMT-Cdk7-FLAG construct was created with a pMT-Cid-FLAG construct (kind gift from P. Heun) and a PCR CDK7 product. The cloning was performed through the Not1 and Spe1 sites. For different derivatives of *Drosophila* MSL1, we used: (i) $MSL1_{1–584}$, which is a truncated form

of MSL expressing aa 1–584; (ii) $MSL1_{423–1037}$, which expresses aa 423–1037; (iii) MSL1-PH1 mutant generated by a point mutation of Ser18 to alanine (S18A); (iv) MSL1-PH3 mutant generated by point mutations of three threonines to alanines at positions 743, 747 and 751 (T743A T747A T751A).

Expression and purification of *Drosophila* recombinant proteins with a baculovirus expression system. Full-length MSL1, MSL1-PH1 and MSL1-PH3 were cloned into the FLAG-pFastBac vectors and expressed with the baculovirus system. Recombinant baculoviruses were generated per the manufacturer's instructions (Bac-to-Bac Baculovirus Expression System, Life Technologies). Viruses were used to infect the SF21 insect cells. Cells were harvested 48 h after infection by dissolving the cell pellets in HEMG K200 buffer (25 mM HEPES, pH 7.6, 0.1 mM EDTA, 12.5 mM $MgCl_2$, 10% glycerol, 200 mM KCl, 0.5% IGEPAL and protease-inhibitor-cocktail tablet (Roche)). For the purification of the recombinant proteins, FLAG-agarose beads (Sigma) were incubated with the cell extracts for 2 h and washed with HEMG K500 and HEMG K200 buffers. FLAG peptides (400 ng/µl, Sigma) were used to elute recombinant proteins. All recombinant proteins were stored in HEMG K200 buffer.

***In vitro* phosphorylation assays with nonradioactive ATP.** Depending on the experiment, 100 ng of the respective recombinant protein was incubated with 100 ng of the human CAK-complex trimer (ProQinase). All reactions were carried out in 30-µl volumes in 70 mM HEPES buffer, pH 7.6, 3 mM $MgCl_2$, 3 mM $MnCl_2$, 1.5 µM PEG, 1.2 mM DTT and 100 µM ATP. Okadaic acid was used as an inhibitor of PP1 and PP2A phosphatases. The reactions were carried out in an Eppendorf ThermoMixer at 300 r.p.m. for 30 min at 30 °C. The reaction products were further analyzed with mass spectrometry and western blotting or dot blotting.

***In vitro* phosphorylation assays with ^{32}P-labeled γ-ATP.** Assays were as described above, except as phosphate, we used radioactive ^{32}P (1 µl of [γ-^{32}P]ATP (PerkinElmer, 3,000 Ci/mmol)).

***In vivo* labeling of *Drosophila* S2 cells.** *Drosophila* S2 cells (2×10^6 cells/ml) were used to perform the *in vivo* labeling with [γ-^{32}P]ATP (PerkinElmer, 3,000 Ci/mmol). 40×10^6 cells were incubated with 20 µl of [γ-^{32}P]ATP at room temperature for 3 h. Phosphatases were blocked with PhosStop phosphatase-inhibitor cocktail for 10 min, washed with PBS and lysed in 1 ml of HEMG K150 buffer. Cell extract from ~10×10^6 cells was used for each IP assay. IPs were performed with anti-MSL1 (rat and rabbit; previously characterized in ref. 16), rat preimmune serum and Sepharose G beads (P3296, Sigma). After SDS–PAGE, the gels were dried, exposed on a phosphor-screen and scanned with a Typhoon scanner (GE Healthcare).

Publicly available ChIP–seq data. Raw reads from published ChIP–seq experiments were downloaded from ArrayExpress, accession number E-MTAB-911 for MOF from male and female *D. melanogaster* samples and the corresponding input. BigWig files of input-normalized read coverage from the ChIP–seq samples for MOF and MSL1 as well as peak regions in male mouse ESCs were taken from ref. 17.

Annotation files. Annotation files were downloaded from FlyBase (Dvir1.2 for *D. virilis*) and UCSC (dm3 for *D. melanogaster*, mm9 for mice).

Strand specific RNA-seq. Strand specific RNA-seq libraries were prepared as previously described[63]. Briefly, polyadenylated RNA was isolated from 10 µg of total RNA with an Oligotex midi kit (Qiagen) according to the manufacturer's instructions. Poly(A)-enriched RNA was fragmented by incubation of the samples at 80 °C for 4 min in the presence of RNA-fragmentation buffer (40 mM Tris-acetate, pH 8.1, 100 mM KOAc and 30 mM MgOAc). The fragmented RNA was purified with 1.8× (v/v) Ampure XP Beads (Beckman Coulter Genomics) and eluted in 25 µl elution buffer (EB; 10 mM Tris-HCl, pH 8) according to the manufacturer's protocol. 24 µl of eluted RNA was reverse transcribed with 1 µl of random hexamers (30 ng/µl). The samples were denatured at 70 °C for 5 min and transferred to ice. 2 µl dNTPs (10 mM), 8 µl 5× first-strand buffer, 4 µl DTT (0.1 M), 0.5 µl actinomycin D (1.25 mg/ml) and 0.5 µl RNasin (Promega) were added to each sample, and the samples were then incubated at 25 °C for 2 min. Subsequently, 0.5 µl Superscript III reverse transcriptase (200 U/µl, Invitrogen) was added. The retrotranscription was carried out at 25 °C for 10 min, then at 55 °C for 60 min, and was inactivated at 75 °C for 15 min. The samples were purified with 1.8× of Ampure XP beads and eluted in 20 µl EB. For producing the second cDNA strand, 19 µl of sample was mixed with 2.5 µl of 10× NEB Next Second Strand Synthesis (dNTP-free) Reaction buffer (NEB), 1.5 µl of dNTPs (containing dUTPs instead of dTTPs, 10 mM), 0.5 µl of RNaseH (10 U/µl) and 0.5 µl of *Escherichia coli* DNA polymerase I (10 U/µl, Fermentas). The samples were incubated at 16 °C for 2.5 h, 80 °C for 20 min purified with 1.8× Ampure XP beads and eluted in 17 µl EB. 2 µl end-repair buffer and 1 µl end-repair enzyme mix (NEBNext DNA Sample Prep Master Mix Set 1, NEB) were added, and the samples were incubated at 20 °C for 30 min. The samples were purified with 1.8× Ampure XP and resuspended in 17 µl EB. 2 µl dA tailing buffer (10× NEBuffer 2 from NEB and 0.2 mM dATP) and 1 µl Klenow fragment 3′-5′ exo– (5 U/µl, NEB) were added, and the samples were incubated at 37 °C for 30 min. The samples were purified with 1.8× Ampure XP and resuspended in 19.5 µl EB. 2.5 µl 10× T4 DNA ligase buffer (NEB), 1 µl multiplexed PE Illumina adaptors (1 µM) and 2 µl T4 DNA ligase were added (2,000 U/µl, NEB) and incubated at 16 °C for 1 h. The dUTPs of the second strand were hydrolyzed by incubating the samples at 37 °C for 15 min with 1 µl USER enzyme (1 U/µl, NEB) and 5 min at 95 °C. The samples were purified with 1.8× Ampure XP beads and eluted in 20 µl EB. Enrichment PCR was performed with 10 µl of sample, 25 µl Phusion Master Mix 2× (NEB), 0.5 µl each of oligos PE 1.0 and PE 2.0 (10 µM, Illumina) and water up to 50 µl final volume. The PCR program was 30 s at 98 °C, 15 cycles of 10 s at 98 °C, 30 s at 65 °C and 30 s at 72 °C, and 5 min at 72 °C. The PCR products were size-selected with double Ampure XP size selection (0.6×–0.8×) and eluted in 15 µl EB. Libraries with an average size of 420 bp were submitted for Illumina 50-bp paired-end sequencing. Sequences were trimmed and then aligned to the reference genome (r5.49) with novoalign V2.07.10 (http://www.novocraft.com/) with default parameters. Differential expression was analyzed with DESeq (http://bioconductor.org/packages/release/bioc/html/DESeq.html), and *P* values adjusted for multiple testing with the Benjamini–Hochberg procedure were computed to control the false discovery rate (FDR) with default parameters.

Lentiviral-based RNAi in ESCs. shRNA constructs were obtained from Sigma in pLKO.1. For production of lentiviral particles, 70–80% confluent HEK293FT cells grown in 10-cm culture dishes were cotransfected with 3.33 µg lentiviral construct, 2.5 µg psPAX2 plasmid and 1 µg pMD2.G plasmid with Lipofectamine 2000 (Invitrogen). ESC medium without LIF was added to the HEK293FT cells, and the supernatant with lentiviral particles was collected after 48 h, filter-sterilized and added to the ESCs. 12 h after infection, cells were treated with 1.0 µg/ml puromycin to remove uninfected cells. ESCs were harvested on day 5 after knockdown. The following shRNA sequences were used for the knockdowns: *msl-1* shRNA, Sigma, TRCN0000241378; *msl-2* shRNA, Sigma, TRCN0000243429; nontargeting control, Sigma, SHC002.

Dot blot assays. Protran BA79 nitrocellulose membranes (0.1 µm; Whatman, Fisher Scientific) were spotted with PBS, Roche protease-inhibitor cocktail and Roche PhosStop phosphatase-inhibitor cocktail solutions of MSL1 unmodified peptide, MSL1 Ser18p peptide and recombinant full-length MSL1 (with or without CAK complex and ATP). After being dried, membranes were blocked with 5% milk in PBS with 0.3% Tween (PBS-T) for 1 h at room temperature (RT), washed three times in PBS-T and then incubated with anti-MSL1-PH1.

***Drosophila* stocks and cell lines.** S2 and Kc167 were the modENCODE cell lines of *D. melanogaster* and were grown in Serva medium supplemented with 10% FCS at 25 °C. The *D. virilis* cell line WR_DV-1 was obtained from the *Drosophila* Genomics Resource Center. Cells were grown in M3 + BPYE medium supplemented with 10% FCS at 25 °C. The *D. virilis* strain was obtained from the *Drosophila* Species Stock Center (University of California, San Diego, wild-type, stock number 15010-1051.00) The *w;;*UAS-*msl-1*^{PH1} and *w;;*UAS-*msl-1*^{PH3} transgenic lines carrying mutant versions of *msl-1* were generated through phiC31 integrase-mediated germline transformation as previously described[64] with y^1 *M[vas–int. Dm]ZH–2A w;PBac[y$^+$–attP–3B]VK00033* (Bloomington stock no. 24871), which carries an attP site at position 65B2 on chromosome 3L[65] (and a *Drosophila* codon-optimized ΦC31 integrase driven in the germline by the *vasa* promoter[66]). The generation of the *w;;*UAS-*msl-1*^{WT} line carrying a wild-type *msl-1* transgene

doi:10.1038/nsmb.3233

has previously been reported[22]. The following stocks were obtained from the Bloomington stock center or kindly donated: *y*^{1} *w*; P[tubP-Gal4]LL7/TM3, Sb*^{1} (Bloomington stock no. 5138), *msl-1*^{L60}*/CyO* (M. Kuroda), *msl-1*$^{\gamma 269}$ *cn*^{1}*bw*^{1}*/CyO and msl-1*^{1}*/CyO* (J. Lucchesi), *w*; In(2LR)noc*^{4L}*Sco*^{rv9R}*, b*^{1}*/CyO, P[ActGFP]JMR1* (Bloomington stock no. 4533),*w*;sna*^{Sco}*/CyO;P[tubP–Gal80*^{ts}*]7* (Bloomington stock no. 7018), *y*^{1}*, sc*, v*^{1}*; P[TRiP.GL00073]attP2 flies* (UAS-*Cdk7*^{dsRNAi}) (Bloomington stock no. 35199), *y*^{1}*, sc*, v*^{1}*; P[TRiP.GL00231]attP2 flies* (UAS-*Cdk8*^{dsRNAi}) (Bloomington stock no. 35324), *y*^{1}*, sc*, v*^{1}*; P[TRiP.GL00230]attP2/TM3, Sb*^{1} *flies* (UAS-*Cdk9*^{dsRNAi}) (Bloomington stock no. 35323), *Df(1)JB254, P[snf*$^{+}$*, dhd*$^{+}$*]SL2, w*/FM7c;; P[Cdk7*^{P140S}*]SL1, Sb*^{1}*/TM3, Ser*^{1} *flies* (Beat Suter), *y*^{1}*, sc*, v*^{1}*; P[TRiP.HMS01930]attP40 flies* (UAS-*msl-1*^{RNAi}) (Bloomington stock no. 39012). The latter stock carries a transgene encoding a temperature-sensitive *Cdk7* (*Cdk7*^{P140S}) in the background of a *Cdk7* loss-of-function allele (*Df(1)JB254*). All UAS-*Cdk*^{dsRNAi} mutants have been generated as part of the Transgenic RNAi Project by the DRSC at Harvard Medical School. All lines used in this study were generated via standard genetic crosses from the above stocks.

***Drosophila* rearing conditions and genetics.** Unless otherwise specified, flies (*D. melanogaster*) were reared on a standard cornmeal fly medium at 25 °C, 70% relative humidity and 12 h dark/12 h light cycle.

To access the viability of males and females homozygous and heterozygous for one of three *msl-1* loss-of-function alleles, namely *msl-1*^{L60} (ref. 28), *msl-1*$^{\gamma 269}$ (ref. 29) and *msl-1*^{1} (ref. 30), 3-d-old *msl-1*^{L60}*/CyO-GFP* virgin females were crossed with 3-d-old *msl-1*$^{\gamma 269}$*/CyO-GFP* males in a population cage and were allowed to lay eggs on yeast-supplemented apple-juice plates. Collections (0–6 h) were used 24 h later to separate fluorescent (*msl-1*^{L60} or *msl-1*$^{\gamma 269}$*/CyO-GFP*) from nonfluorescent (*msl-1*$^{L60/\gamma 269}$) larvae under a fluorescence stereomicroscope. The larvae were placed in standard fly food vials, 75 per vial. A total of 150 larvae were collected for each genotype. Males and females were counted after eclosion, and the percentage viability was scored, taking 75 as 100%.

To test for an effect of a lack of maternal MSL1 contribution to the viability of female flies, *msl-1*$^{L60/\gamma 269}$ *trans*-heterozygous virgin females were collected from the progeny of *msl-1*^{L60}*/CyO-GFP* and *msl-1*$^{\gamma 269}$*/CyO-GFP* parents and crossed to either *msl-1*^{L60}*/CyO-GFP* or *msl-1*$^{\gamma 269}$*/CyO-GFP* males. The offspring from at least three independent crosses were counted every other day for a period of 10 d from the start of the eclosion. The total number of non-*Cy* females was divided by the total number of *Cy* females, which was used as an internal control with 100% viability.

To elucidate the effect of dsRNAi-mediated *msl-1* silencing on fly viability, five virgin females carrying a strong ubiquitous *tub-Gal4* driver (*tub-Gal4/TM6*, *Tb*, GFP) were crossed to five males carrying a UAS-*msl-1*^{RNAi} transgene. The total number of non-*Tb* female flies was divided by the total number of *Tb* females and was used as an internal control with 100% viability.

To obtain flies ectopically expressing *msl-1*^{WT}, *msl-1*^{PH1} or *msl-1*^{PH3} in an *msl-1*-null background, *y*^{1} w*; *msl-1*^{L60}*/CyO, P[ActGFP]JMR1; [tubP-Gal4]LL7/TM6B, Tb*^{1} virgin females were crossed with *y*^{1} w*; *msl-1*$^{\gamma 269}$ *cn*^{1} *bw*^{1}/CyO, *P[ActGFP]JMR1*; *P*[UAS-*msl-1*]*65B2 males, where * stands for WT, PH1 or PH3. To determine the relative viability after ectopic expression of *msl-1**, male and female adult flies from three independent crosses were counted every other day for a period of 10 d from the start of eclosion. The total number of non-*Cy*, non-*Tb* males and females (*msl-1*$^{L60/\gamma 269}$; *tub-Gal4*/UAS-*msl-1** ectopically expressing *msl-1** in a null-mutant background) was divided by the total number of *Cy*, *Tb* (*msl-1*^{L60} or *msl-1*$^{\gamma 269}$*/CyO-GFP TM6-Tb*/UAS-*msl-1**, with endogenous *msl-1* and no ectopic expression) males and females respectively, which were used as internal controls with 100% viability.

To assess the genetic interaction between CDK7 and MSL1, *y*^{1} w*; *msl-1*^{L60}*/CyO, P[ActGFP]JMR1; P[tubP-Gal4]LL7, P[tubP-Gal80*^{ts}*]7/TM6B, Tb*^{1} and *y*^{1} w*;; *P[tubP-Gal4]LL7, P[tubP-Gal80*^{ts}*]7/TM6B, Tb*^{1} virgin females were crossed with *y*^{1} w*; *msl-1*$^{\gamma 269}$ *cn*^{1}*bw*^{1}*/CyO, P[ActGFP]JMR1*; *P*[UAS-*Cdk**]attP2 males, where * stands for 7, 8 or 9, and flies were allowed to develop at 24 °C. The progeny were counted as described for the above cross, and the viability of the *Cy*, non-*Tb* males and females from the two crosses was evaluated relative to the *Cy*, *Tb* males and females used as an internal control with 100% viability. All non-*Cy* males die because of the lack of MSL1 and dosage compensation, so the effect of Cdk* RNAi could be scored in only non-*Cy*, non-*Tb* females relative to the non-*Cy*, *Tb* females, which were used as a 100%-viable internal control.

Notably, although heterozygous *msl-1*-null mutant alleles do not affect viability of flies, they have previously been shown to exhibit decreased levels of MSL1, thus making them amenable to genetic interaction analysis[43]. When a strong ubiquitous driver, *tub-Gal4*, was used for dsRNAi induction at 25 °C, this resulted in 100% lethality (data not shown), thus recapitulating the loss-of-function phenotype previously reported for *Cdk7*, *Cdk8* and *Cdk9*. When a combination of *tub-Gal4* and *tub-Gal80*^{ts} was used at 24 °C, partial lethality was observed (50% for *Cdk7*, 60% for *Cdk8* and 40% for *Cdk9*; **Fig. 3a**), thus allowing for enhancers or suppressors of this phenotype to be identified after introduction in this genetic background.

In an alternative approach to genetically assess the interaction between CDK7 and MSL1, *Df(1)JB254, P[snf*$^{+}$*, dhd*$^{+}$*]SL2, w*/FM7c;; P[Cdk7*^{P140S}*]SL1, Sb*^{1}*/TM3, Ser*^{1} virgin females were crossed with either *w y; msl-1*^{L60}*/CyO, Act5C-GFP* or *w y; If/CyO, Act5C-GFP* males, and the progeny were counted as described for the above cross. Ectopic expression of the *Cdk7*^{P140S} allele in a *Cdk7*-null background [*Df(1)JB254*] at a permissive temperature (18 °C) resulted in viable flies phenotypically indistinguishable from wild type. However, inactivation of *Cdk7*^{P140S} at the restrictive temperature (29 °C) (when the only source of CDK7 comes from the mutant allele *Cdk7*^{P140S}) causes embryonic and larval lethality[32]. Partial lethality (~60%) was observed at 25 °C, and the surviving flies exhibited a range of wing, cuticular and bristle phenotypes, thus allowing for enhancement or suppression of viability to be scored in the presence of a mutant *msl-1* allele or a dosage-compensation-independent mutation '*If*' as a control (**Supplementary Fig. 4a**). However, because the *Cdk7* gene is on the X chromosome (position 4F4), synthetic lethality could be scored only in males that lacked endogenous *Cdk7* and whose only source of Cdk7 was the temperature-sensitive autosomal transgene. Females were heterozygous for the *Cdk7*-null mutation with Cdk7 originating from both the endogenous *Cdk7* and the transgene, thus resulting in levels promoting 100% viability; hence, these flies were used as an internal control.

***Drosophila* primers.** *Drosophila* primer sequences are listed in the **Supplementary Note**.

60. Diaz, A., Nellore, A. & Song, J.S. CHANCE: comprehensive software for quality control and validation of ChIP-seq data. *Genome Biol.* **13**, R98 (2012).
61. Padeken, J. *et al.* The nucleoplasmin homolog NLP mediates centromere clustering and anchoring to the nucleolus. *Mol. Cell* **50**, 236–249 (2013).
62. Zink, D. & Paro, R. *Drosophila* Polycomb-group regulated chromatin inhibits the accessibility of a *trans*-activator to its target DNA. *EMBO J.* **14**, 5660–5671 (1995).
63. Parkhomchuk, D. *et al.* Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res.* **37**, e123 (2009).
64. Groth, A.C. & Calos, M.P. Phage integrases: biology and applications. *J. Mol. Biol.* **335**, 667–678 (2004).
65. Venken, K.J., He, Y., Hoskins, R.A. & Bellen, H.J. P[acman]: a BAC transgenic platform for targeted insertion of large DNA fragments in D. melanogaster. *Science* **314**, 1747–1751 (2006).
66. Bischof, J., Maeda, R.K., Hediger, M., Karch, F. & Basler, K. An optimized transgenesis system for *Drosophila* using germ-line-specific phiC31 integrases. *Proc. Natl. Acad. Sci. USA* **104**, 3312–3317 (2007).