

Harvey, R, Schuster, E and Jennings, B

Pleiohomeotic Interacts with the Core Transcription Elongation Factor Spt5 to Regulate Gene Expression in *Drosophila*

Harvey, R, Schuster, E and Jennings, B (2013) Pleiohomeotic Interacts with the Core Transcription Elongation Factor Spt5 to Regulate Gene Expression in *Drosophila*. *PLoS ONE*, 8 (7). pp. 1-10.

This version is available: <https://radar.brookes.ac.uk/radar/items/9509ce0b-fdd3-41df-b21a-99e539ca04ab/1/>

Available on RADAR: July 2016

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the published version of the journal article.

# Pleiohomeotic Interacts with the Core Transcription Elongation Factor Spt5 to Regulate Gene Expression in *Drosophila*

Robert Harvey<sup>1</sup>✉, Eugene Schuster<sup>2</sup>✉, Barbara H. Jennings<sup>1\*</sup>

**1** Transcriptional Regulation Group, UCL Cancer Institute, University College London, London, United Kingdom, **2** Department of Genetics, Evolution and Environment, University College London, London, United Kingdom

## Abstract

The early elongation checkpoint regulated by Positive Transcription Elongation Factor b (P-TEFb) is a critical control point for the expression of many genes. Spt5 interacts directly with RNA polymerase II and has an essential role in establishing this checkpoint, and also for further transcript elongation. Here we demonstrate that *Drosophila* Spt5 interacts both physically and genetically with the Polycomb Group (PcG) protein Pleiohomeotic (Pho), and the majority of Pho binding sites overlap with Spt5 binding sites across the genome in S2 cells. Our results indicate that Pho can interact with Spt5 to regulate transcription elongation in a gene specific manner.

**Citation:** Harvey R, Schuster E, Jennings BH (2013) Pleiohomeotic Interacts with the Core Transcription Elongation Factor Spt5 to Regulate Gene Expression in *Drosophila*. PLoS ONE 8(7): e70184. doi:10.1371/journal.pone.0070184

**Editor:** Helen White-Cooper, Cardiff University, United Kingdom

**Received:** April 11, 2013; **Accepted:** June 17, 2013; **Published:** July 22, 2013

**Copyright:** © 2013 Harvey et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was funded by a MRC Career Development Award to Barbara Jennings (G0800310) and a MRC Career Development Award to Eugene Schuster (G0800339). <http://www.mrc.ac.uk/index.htm>. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** Barbara Jennings is currently an Academic Editor for PLOS ONE. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

\* E-mail: [barbara.jennings@ucl.ac.uk](mailto:barbara.jennings@ucl.ac.uk)

✉ These authors contributed equally to this work.

✉ Current address: MRC Toxicology Unit, University of Leicester, Leicester, United Kingdom

## Introduction

The regulation of the early phase of transcriptional elongation is used to control the expression of many genes. When this process fails it leads to death or severe defects during development and contributes to cancer pathogenesis in adult animals [1].

Once transcription has been initiated by recruitment of the pre-initiation complex (PIC), RNA polymerase II (RNAP II) transcribes 20–40 base pairs but then must pass through a checkpoint regulated by Positive Transcription Elongation Factor b (P-TEFb) to produce full-length transcripts (recently reviewed in [2,3,4]). Two protein complexes act together to inhibit transcript elongation beyond ~25–40 nucleotides after initiation. One of these is made up of the Spt5 and Spt4 proteins and is sometimes referred to as “DSIF” [5,6], and the other, Negative Elongation Factor (NELF), contains four subunits (NELF-A, NELF-B, NELF-C/D, NELF-E; [7]). For further elongation to occur, P-TEFb must phosphorylate specific residues in NELF, Spt5, and RNAP II. This induces the dissociation of NELF from the polymerase complex, the switch in Spt5 from being a negative to positive regulator of transcription, and production of the full-length transcript by RNAP II. Spt5 tracks along with the RNAP II elongation complex until transcription termination.

Spt5 is required to establish promoter proximal polymerase pausing at the P-TEFb checkpoint, however, it is essential for productive transcription from all genes. Spt5 is conserved across the three domains of life [Eukaryotes, Archaea and Bacteria

(NusG)] and is recruited by RNA polymerases I, II and III [5]. Recent structural studies have shown that the NGN domain of Spt5 sits over the DNA and RNA bound in the active site of RNA polymerases, where it can directly control the rate of transcript elongation [8,9].

It is well established that the P-TEFb checkpoint is a key point of regulation for many genes. However, the factors that determine which genes are subject to rate-limiting regulation at the P-TEFb checkpoint are largely unknown, as is how they interact with the RNAP II elongation complex to establish promoter proximal pausing.

Missense mutations in *Spt5* that give rise to specific developmental defects have been isolated in zebrafish and *Drosophila* [10,11] providing evidence that Spt5 activity is responsive to contextual factors controlling gene expression. Zebrafish homozygous for the *Spt5*<sup>oggy(m306)</sup> allele develop quite normally, however they do exhibit a distinctive neural phenotype (excess dopaminergic neurons and fewer serotonergic neurons) and eventually die of vascular defects thought to be a secondary consequence of abnormal neuronal function [10]. Meanwhile, *Drosophila* embryos derived from maternal germline clones homozygous for the *Spt5*<sup>W049</sup> mutation (thus, all protein in the embryo prior to the onset of zygotic transcription is mutant), exhibit segmentation defects stemming from aberrant expression of *even-skipped* (*eve*) and *run* (*run*). The effects of *Spt5*<sup>W049</sup> are gene-specific, (gap gene and *hairy* expression are normal in *Spt5*<sup>W049</sup> germline clones) and appear to be enhancer-specific for *eve* expression [11]. The single

amino acid substitutions found in the Foggy and W049 mutant proteins map close together in the C-terminal region of Spt5, which is conserved in higher metazoans including *Drosophila*, but not found in yeast or *C. elegans*. This region is distinct from the domain in Spt5 that is subject to phosphorylation by P-TEFb, which is sometimes referred to as the Spt5 CTR or CTD domain. Thus to avoid confusion, we will refer to the extreme C-terminal domain of Spt5 found in higher metazoans as the Developmental Domain (DD).

In this study we characterise an interaction between Spt5 and the transcription factor Pleiohomeotic (Pho) that we uncovered using the yeast 2-hybrid assay. We demonstrate that Spt5 and Pho act together in vivo during adult maturation and PcG repression, and that the majority of sites bound by Pho in the genome colocalize to Spt5 and NELF binding sites.

## Results

### Spt5 Interacts with Pho

We performed a yeast 2-hybrid screen using the C-terminal 153 amino acids of *Drosophila* Spt5 as bait to identify factors that interact with the DD. In frame fragments of Pho were recovered from the screen multiple times and did not retest as false positives. Pho is an ortholog of mammalian Ying Yang 1 (YY1) and like Spt5, is a ubiquitously expressed nuclear protein that has been implicated in both transcriptional activation and repression [12,13].

Full length Spt5 interacts with full length Pho in the yeast 2-hybrid assay (Figure 1A) and the DD also interacted weakly but specifically with Pho in GST pull down assays (Figure 1B). The interaction between full length Spt5 and Pho was further validated by expressing tagged proteins in *Drosophila* S2 cells and performing co-immunoprecipitation assays (Figure 1C). We mapped the interaction with Spt5 to the N-terminal 351-amino acids of Pho using co-immunoprecipitation of tagged proteins (Figure 1C). Sequences within this region have previously been shown to interact with Polycomb (Pc), Polyhomeotic (Ph) and the Brahma (BRM) complex [14]. The C-terminal region of Pho (remaining 169 amino acids), which does not interact with Spt5, contains the 4 zinc finger motif (C2H2-like) that is highly conserved with human YY1 and has been shown to bind DNA [12].

The DD carrying the *W049* (G994D) mutation is able to interact with Pho in the yeast 2-hybrid and GST pull down assays, and full length W049 protein interacts with Pho in the yeast 2-hybrid and co-immunoprecipitation assays (Figure 1D and data not shown). Thus a failure of the interaction between Spt5 and Pho is not likely to explain the phenotypes observed in *Spt5<sup>W049</sup>* mutants. The *W049* mutation may be affecting the ability of Spt5 to interact with other as yet unidentified factors.

### Spt5 Contributes to Pho Mediated Repression of PcG Targets in vivo

We looked for genetic interactions between mutant alleles of *Spt5* and *pho* to assess if they function together in vivo. *pho<sup>cv</sup>* is a hypomorphic allele that is homozygous viable but male sterile [12]. Homozygous *pho<sup>cv</sup>* males exhibit the classic polycomb phenotype of ectopic sex combs on the middle (mesothoracic) and rear (metathoracic) legs due to de-repression of the *Sex combs reduced* (*Scr*) gene [15]. Multiple crosses were done in parallel in uncrowded vials (~3–6 females and ~2–4 males) at 25°C. Siblings were scored to reduce effects caused by genetic background and environment.

We counted the number of flies carrying ectopic sex combs in homozygous *pho<sup>cv</sup>* males and homozygous *pho<sup>cv</sup>* males heterozygous

for *Spt5* alleles to determine if *Spt5* interacts with *pho* during PcG repression in vivo. A two-proportion hypothesis test was applied to determine the significance of any differences observed in these frequencies. There was no significant increase in the frequency of ectopic sex combs observed in *pho<sup>cv</sup>/pho<sup>cv</sup>* males that are heterozygous for a null allele of Spt5 (*Spt5<sup>MGE-3</sup>*) [16], indicating that halving the dose of Spt5 does not compromise residual Pho activity (Figure 2 and Table 1). This is perhaps not unexpected as Spt5 is expressed at moderate or moderately high levels during larval development [17] and *Spt5<sup>MGE-3</sup>/+* flies appear wild-type. However, *Spt5<sup>W049</sup>/+*; *pho<sup>cv</sup>/pho<sup>cv</sup>* males did show a significant increase in the number of ectopic sex combs, revealing that the *W049* variant protein can disrupt the repressive activity of Pho in vivo (Figure 2 and Table 1). We also observed a small but significant increase in the number of ectopic sex combs in *pho<sup>cv</sup>* males heterozygous for *NELF-A<sup>KG</sup>* [18] (Figure 2, Table 1).

The W049 protein has a significantly reduced repressive activity on transcription in vitro, and in some contexts in vivo [11]. W049 allows RNAP II to continue transcribing through the P-TEFb checkpoint in the presence of a P-TEFb inhibitor [5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB)] in nuclear extracts [11]. Thus, the presence of W049 protein has the potential to interfere with repression dependent on the P-TEFb checkpoint in heterozygous flies. *Spt5<sup>W049</sup>/+* flies resemble wild-type, so this effect is only apparent when the function of other factors involved is compromised.

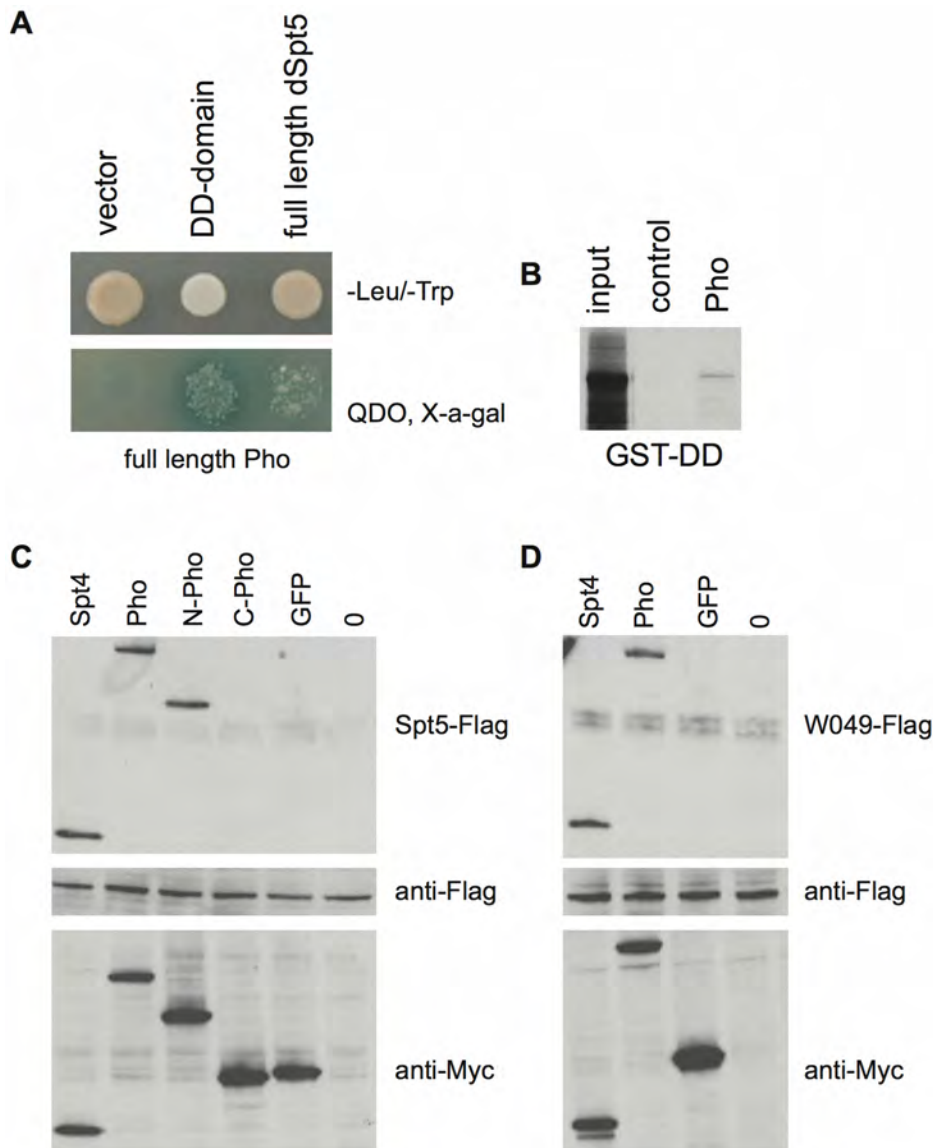
Kwong et al., 2008 observed an enrichment of Pho binding just downstream of the start of transcription of *Scr* in T3 imaginal discs around the predicted site of the P-TEFb checkpoint [19]. Furthermore, the additional ectopic sex combs observed in *NELF-A<sup>KG</sup>/+*; *pho<sup>cv</sup>/pho<sup>cv</sup>* flies is consistent with the model that inhibition of this checkpoint is critical for Pho-mediated PcG repression of *Scr*. Thus, we propose a model in which Pho acts together with Spt5 and NELF to prevent RNAP II transcribing through the P-TEFb checkpoint to maintain PcG repression.

### Spt5 Genetically Interacts with *pho* during Wing Maturation

While assessing the various genotypes for Polycomb phenotypes we noticed that approximately 10% of *pho<sup>cv</sup>* homozygotes exhibit a phenotype resulting from aberrant wing inflation and deflation during hatching from the pupal case (eclosion). Introduction of a single copy of *Spt5<sup>W049</sup>* or *Spt5<sup>MGE-3</sup>* into this background increased the frequency to 30% and 28% respectively, demonstrating a significant genetic interaction between the *pho* and *Spt5* loci (Figure 3 and Table 1). All elements of the wing (veins, bristles, and hairs) are present and normal in *pho<sup>cv</sup>* homozygotes, but affected wings were noticeably ruffled along the posterior edge and had regions where the dorsal and ventral surfaces were coming apart. The extent of this phenotype was variable, with some wings also being folded and/or containing small blisters.

A wing inflation phenotype has not previously been described for *pho<sup>cv</sup>*, however the phenotypes of escaper flies homozygous for stronger *pho* alleles support a role for *pho* in wing development. Flies homozygous for *pho<sup>b</sup>* allele [12] die as pharate adults; they make it all the way through development on the maternally supplied Pho, but fail to eclose.

Similarly, expression of *UAS-RNAi-pho* driven ubiquitously throughout development by *da-GAL4* is generally lethal at 18°C, with flies dying as pharate adults. The vast majority of escapers that hatch are unable to fully inflate their wings and remain pale and juvenile looking (91%; n=67) in addition to having the phenotypes previously described for *pho* mutants including ectopic sex combs and partial homeotic transformations of abdominal



**Figure 1. PhO physically interacts with Spt5.** A) Yeast 2-hybrid assay showing binding of full length PhO with the Spt5 DD domain and full length Spt5. Vector (pGBKT7) containing no insert was used as a control to demonstrate that PhO does not activate reporter gene expression in the absence of Spt5. B) PhO binds to immobilized GST-DD. Ten percent of the input PhO is run in left lane, immobilized GST in middle lane incubated with PhO as negative control. C) Western blots of co-immunoprecipitation (co-IP) assays from S2 cell extracts of Flag-tagged Spt5 with Myc-Spt4 (positive control), Myc-Pho, Myc-N-Pho (amino acids 1–351), Myc-C-Pho (351–520), Myc-GFP (negative control) and no protein. D) Western blots of co-IP assays from S2 cell extracts of Flag-tagged W049 variant of Spt5 with Myc-Spt4 (positive control), Myc-Pho, Myc-GFP (negative control) and no protein.  
doi:10.1371/journal.pone.0070184.g001

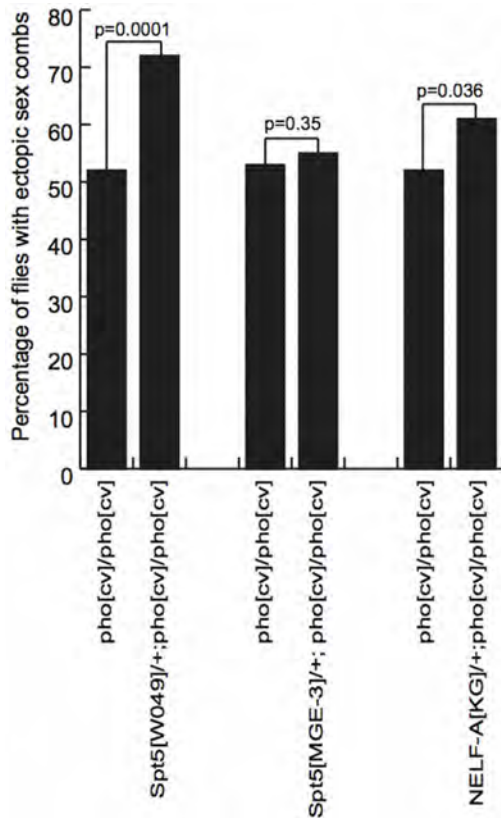
segments (Figure 3F and 3G, [20]). Driving ubiquitous expression of *UAS-RNAi-pho* recapitulates the phenotype of strong *pho* alleles in vivo.

There are no obvious wing defects when *765-Gal4* drives *UAS-RNAi-pho* expression broadly in wing imaginal discs (Figure 3C). However, expression of *UAS-Pho-RNAi* under the control of *386Y-Gal4*, which drives expression in peptidergic neurons that control wing inflation [21] leads to an inflation phenotype in 51% of flies ( $n = 136$ ) (Figure 3D). Knock-down of *Spt5* expression by *UAS-RNAi-Spt5* is cell lethal, similarly clones of cells homozygous for null alleles of *Spt5* do not survive (Figure 4), so we were unable to determine if the genetic interaction between *Spt5* and *pho* occurs specifically in peptidergic neurons.

Due to the technical difficulties of studying pupal development, the gene networks that drive eclosion and wing inflation are poorly understood. However, a number of other transcriptional regulators have been implicated in these processes including CREB binding protein (CBP) and the trithorax group protein Ash1 [21]. Our observations demonstrate for the first time that *pho* plays a key role in eclosion, including the process of wing inflation and deflation.

#### Pho and Spt5 Bind Overlapping Sites across the Genome

We performed meta-analysis of Pho and Spt5 data from published chromatin immunoprecipitation (ChIP) experiments in *Drosophila* S2 culture cells to determine if Pho and Spt5 ever co-localize to the same sites in the genome in a given cell type. Peaks



**Figure 2. Modification of the extra sex combs phenotype of *pho<sup>cy</sup>/pho<sup>cy</sup>* mutants by *Spt5* and *NELF* mutant alleles.** A chart representing the frequency of ectopic sex combs in *pho<sup>cy</sup>/pho<sup>cy</sup>* mutants and siblings heterozygous for *Spt5<sup>W049</sup>*, *Spt5<sup>MGE-3</sup>* or *NELF-A<sup>KG</sup>* over wild-type chromosomes. p values from two proportion z-tests are shown.

doi:10.1371/journal.pone.0070184.g002

of Spt5 binding were identified from data in Gilchrist et al., 2010 and the binding data set for Pho was available as part of the modENCODE project [22]. We identified 5590 binding sites for Spt5 and 1862 for Pho in S2 cells. The vast majority of Pho binding sites (1424/1862; 76%) overlap with Spt5 peaks (Figure 5A), while conversely 25% of Spt5 sites overlap with Pho peaks.

Previous studies have demonstrated that Spt5 binds around the transcription start site (TSS) of genes that recruit RNAP II, and also within the gene bodies of actively transcribed genes [23,24,25]. Pho binds target sequences associated with the establishment of PcG complexes, but peaks of binding are also found around the TSS and within the gene body of many genes [19,26,27,28]. Heat maps of Spt5 and Pho binding illustrate that Spt5 and Pho frequently bind overlapping sites at or within 200 bp of the TSS (Figure 5B).

The NELF complex has a well documented role in establishing promoter proximal paused RNAP II in higher eukaryotes including *Drosophila* [7,29,30,31]. Spt5 and NELF co-localize around the TSS of many paused genes in *Drosophila* [25]. We compared the peaks of Pho binding to the peaks of NELF (NELF-B) in S2 cells identified in [25]. The majority of Pho peaks overlapped with peaks of NELF (72%), and 67% of Pho peaks overlap with both NELF and Spt5 (Figure 5C). We also compared peaks of Pho binding to data for NELF-B and NELF-E binding reported in [31]. In this data set, Pho peaks overlap with 74% of NELF-B, 75% of NELF-E and 72% with both NELF-B and NELF-E. Thus the vast majority of Pho binding sites co-localise with binding sites for factors known to regulate pausing. There are many more binding sites for Spt5 and NELF than for Pho in S2 cells, indicating that Pho is not a core component of the machinery regulating transcription elongation, but rather a factor that may influence its activity at a subset of genes.

The ability of Pho to bind Polycomb Response Elements (PREs) in chromatinized DNA is augmented by the GAGA factor (GAF; encoded by *Trl*) [32]. Mutations in *pho* and *Trl* interact in genetic assays, but a direct physical interaction between the proteins has not been detected [32,33,34]. GAF associates with 39% of the genes that have NELF, and GAF binding is often associated with promoter proximal paused polymerase [31,35,36]. We observe that 38% of Pho peaks overlap GAF peaks in S2 cells (Figure 5D). Although GAF may facilitate Pho binding at some sites, its presence is not always necessary for Pho recruitment.

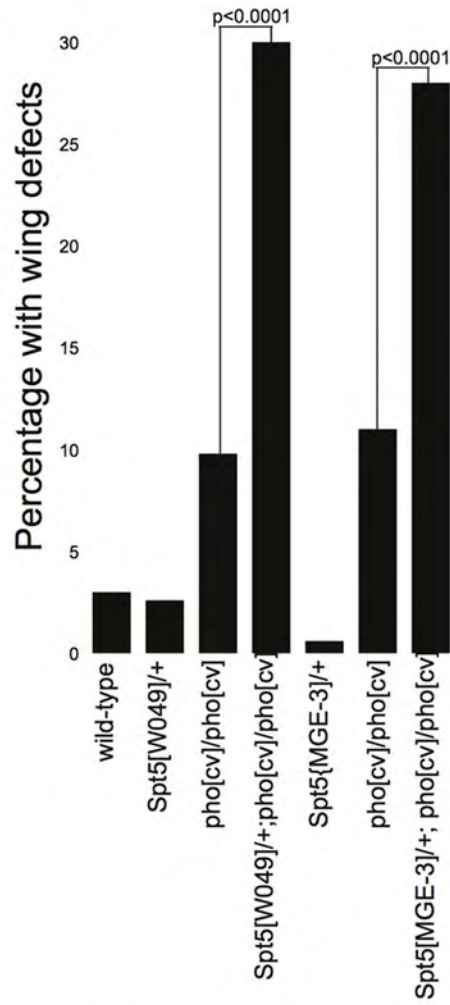
## Discussion

We have detected a physical association of Pho and Spt5 in three different assays; yeast 2-hybrid, GST-pull down and co-immunoprecipitation of tagged proteins. Unfortunately we were unable to co-immunoprecipitate the endogenous proteins as the antibodies generously made available to us against the endogenous proteins were all rabbit polyclonals making co-immunoprecipitation

**Table 1. Genetic Interactions between *Spt5*, *NELF-A* alleles and *pho<sup>cy</sup>*.**

Genotype	Total (n)	Wing Phenotype	Ectopic sex combs
wild-type	233	7 (3%)	0
<i>Spt5</i> [W049]/+	191	5 (2.6%)	0
<i>pho</i> [cy]/ <i>pho</i> [cy]	194	19 (9.8%)	102 (53%)
<i>Spt5</i> [W049]/+; <i>pho</i> [cy]/ <i>pho</i> [cy]	149	45 (30%)	108 (72%)
<i>Spt5</i> [MGE-3]/+	170	1 (0.6%)	0
<i>pho</i> [cy]/ <i>pho</i> [cy]	184	21 (11%)	102 (55%)
<i>Spt5</i> [MGE-3]/+; <i>pho</i> [cy]/ <i>pho</i> [cy]	198	56 (28%)	106 (54%)
<i>NELF-A</i> [KG]/+	166	17 (10%)	0
<i>pho</i> [cy]/ <i>pho</i> [cy]	175	22 (12%)	92 (53%)
<i>NELF-A</i> [KG]/+; <i>pho</i> [cy]/ <i>pho</i> [cy]	288	61 (21%)	176 (61%)

doi:10.1371/journal.pone.0070184.t001



**E**

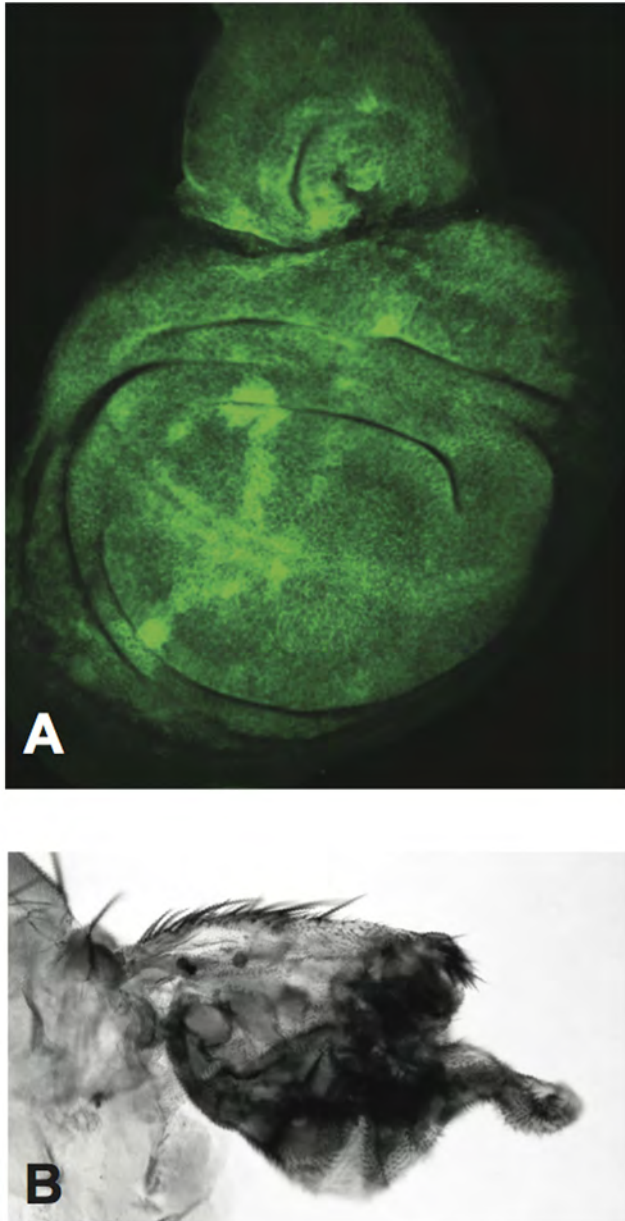


**F**



**G**

**Figure 3. Pho and Spt5 function together in wing maturation.** A wing inflation phenotype is observed in approximately 10% of *pho<sup>cv</sup>/pho<sup>cv</sup>* B) and 51% of *386Y-Gal4>UAS-RNAi pho* males (n = 136), but not in *765-Gal4>UAS-RNAi-pho*. E) Percentage of flies of indicated genotypes displaying wing inflation phenotypes. F) Ventral view of *da-Gal4>UAS-RNAi-pho* male, red arrow points to ectopic sex comb on middle (mesothoracic) leg. G) Dorsal view of *da-Gal4>UAS-RNAi-pho* male displaying homeotic transformations in the abdominal segments. doi:10.1371/journal.pone.0070184.g003



**Figure 4. Depletion of Spt5 leads to cell death in vivo.** A) Homozygous clones of the *Spt5<sup>MGE</sup>* null allele are not viable. Attempts were made to make clones of homozygous *Spt5<sup>MGE</sup>* cells using the FLP/FRT technique [61]. Third instar imaginal wing disk (anterior to the left and dorsal to the top) stained for GFP. All cells stain green and are thus either heterozygous or homozygous (bright green) for the *FRT42B*, *GFP* chromosome; loss of GFP would mark clones of homozygous *FRT42B*, *Spt5<sup>MGE</sup>*. Similarly, when we induced homozygous germ-line clones of *Spt5<sup>MGE</sup>* in females using the FLP/FRT/ovoD technique [62], they did not lay any eggs indicating that homozygous *Spt5<sup>MGE</sup>* clones are cell lethal (data not shown). B) Residual wing stub from fly expressing *765-Gal4>UAS-RNAi-Spt5* at 18°C the portion of the wing expressing *765-Gal4* does not develop as there is a deficit of cells consistent with expression of *UAS-RNAi-Spt5* being lethal to cells. doi:10.1371/journal.pone.0070184.g004

impractical. However, we do find that Spt5 and Pho co-localize to over 1000 peaks of binding in *Drosophila* S2 cells, supporting the model that they can interact directly. We have also detected a genetic interaction between alleles of *pho* and *Spt5* during PcG repression and wing maturation, indicating that they function together in vivo.

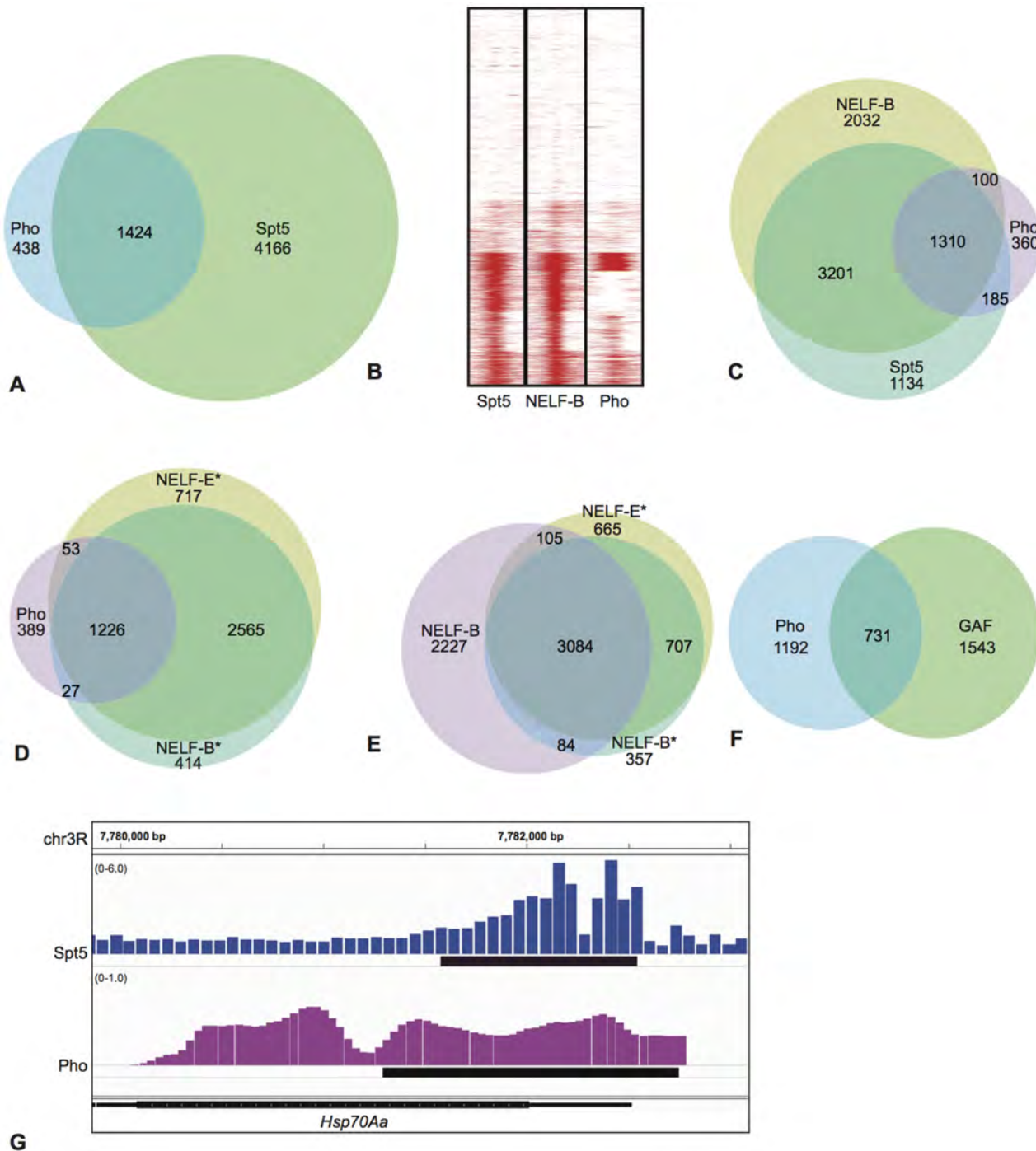
Previous studies have generated speculation about a direct interaction between PcG proteins and the transcription elongation complex. In mouse embryonic stem cells (ESCs), there is a well-established link between PcG repression and polymerase pausing at bivalent genes [37,38,39]. However, the composition of PcG complexes differs between flies and mice, and YY1 (the mouse orthologue of Pho) is not as commonly associated with PcG complexes as Pho [40]. Thus the observations made in mouse may have limited relevance with respect to our observations in *Drosophila*.

In *Drosophila*, the observation that stalled RNAP II persists in tissues where *Ubx* and *Abd-B* are silenced by the PcG complex lead to the supposition that RNAP II elongation factors “somehow communicate with the PcG-silencing complex” [41]. Others noted that PRC1 preferentially binds to promoters associated with stalled RNAP II in *Drosophila* S2 cells [42].

We have confirmed that there is indeed a direct physical interaction between at least one of the RNAP II elongation factors (Spt5) and one member of the PcG complex (Pho) in *Drosophila*. We have also detected a genetic interaction between the *Spt5<sup>W049</sup>* and *pho<sup>cv</sup>* alleles in vivo. The W049 variant of Spt5 causes ectopic transcription through the P-TEFb checkpoint [11]. Thus, we propose a model in which Spt5 acts together Pho to prevent RNAP II transcribing through the P-TEFb checkpoint to maintain PcG repression. In *Spt5<sup>W049</sup>/+; pho<sup>cv</sup>/pho<sup>cv</sup>* flies, the effects of the greatly reduced levels of Pho on PcG repression are exacerbated by a proportion of the remaining Pho interacting with the W049 variant of Spt5 that allows aberrant transcription through the P-TEFb checkpoint.

Pho also functions independently of PcG complexes. One example of this is Pho’s function during the recovery from heat shock to repress heat shock gene expression to basal levels [26]. The mechanism to establish recovery from heat shock involves inducing RNAP II to pause at the P-TEFb checkpoint [43]. Observations made by Beisel et al., lead to a speculative model that Pho interacts directly with the RNAP II elongation complex or a remodeling complex [26]. Our observation that Pho interacts with Spt5 supports this model. Mutations in *Spt5* lead to a greatly diminished heat shock response, making it difficult to evaluate the role of Spt5 in heat shock recovery ([11] and BHJ unpublished data). However, Spt5 and Pho co-localize around the TSS of the *Hsp70Aa* gene in S2 cells that have not been heat shocked, consistent with a model in which they interact to establish pausing (Figure 5G).

Spt5 is recruited to RNAP II during the transition from initiation to early elongation [44] and is involved in all transcription irrespective of promoter proximal pausing, thus it is unlikely that Spt5 recruitment is directly dependent on Pho. Pho is a sequence specific DNA binding protein [12]. However, Pho is also found spread across actively transcribed genes, including *hsp70*, where it is involved with re-establishing polymerase pausing after heat shock [26]. It is possible that Spt5 recruits Pho to the



**Figure 5. Meta-analysis of ChIP data for Pho [22], Spt5 [25], NELF [25,31] and GAF [31] binding across the genome of *Drosophila* S2 cells.** Asterisks denote the NELF data from [31]. A) Venn diagram showing peaks the overlap of Pho binding with Spt5. B) Heat maps shows peaks of Spt5, NELF-B and Pho binding relative to the TSS (centre of each column) for all coding genes annotated in the genome from the Ensembl database (Release 5.48). Plots show 200 bp up and downstream of the TSS. SEQMINER was used to cluster and visualise the data using the default settings (the 'Kmeans raw' clustering normalization method with 10 expected clusters) [60]. C) Venn diagram showing overlap of Pho and Spt5 peaks with NELF-B binding. D) Overlap of Pho peaks with peaks derived from NELF-B\* and NELF-E\* datasets from [31]. Differences in the profiles of NELF binding between these two studies may be due to the use of different antibodies and/or experimental conditions to perform ChIP-chip. E) Venn diagram showing the overlap between the NELF-B peaks from [25] and NELF-B\* and NELF-E\* datasets from [31]. Differences in the profiles of NELF binding between these two studies may be due to the use of different antibodies and/or experimental conditions to perform ChIP-chip. F) Overlap between Pho and GAF peaks in S2 cells. Note: the total number of peaks for an individual factor can vary due to the merging of several overlapping peaks into a single peak if multiple peaks of one factor overlap with a single peak of another. G) Overlap of ChIP peaks for Spt5 and Pho binding at the *Hsp70Aa* gene. Note: There are differences in the publicly available track format of Spt5 (bedgraph) and Pho (smoothed wig) data, and differences in the tiling array used for the ChIP-chip experiments (Spt5: Nimblegen Henikoff\_Dmel\_r52\_ChIP tiling design, mean probe length = 53 bp, mean distance between probes = 12 bp; Pho: Affymetrix *Drosophila* v2.0R tiling array, probe length = 25 bp, mean distance between probes = 38 bp). doi:10.1371/journal.pone.0070184.g005

gene body of *hsp70*, but depletion of Spt5 is lethal to cells (Figure 4) making it difficult to evaluate the role of Spt5 in Pho recruitment. Alternatively, Pho and Spt5 may be recruited to target genes independently, but interact when recruited in close proximity.

Further studies are required to determine the precise details of how Pho influences polymerase pausing, however our current knowledge of which factors Pho interacts with suggests that it could act by helping to tether the polymerase complex close to TSSs, and/or act by nucleosome remodelling.

It has been proposed that paused polymerase is physically held by factors bound to DNA at promoters, since conditions that disrupt protein-protein and protein-DNA interactions allow transcription to run-on [45]. Furthermore, insertion of spacer sequences into the promoter of the *Hsp70* gene in *Drosophila* does not change the site of transcription initiation but does shift the site where the polymerase pauses [36]. Pho could form part of the tethering complex when it binds close to the TSS and interacts with Spt5. Polymerase pausing is not always associated with repression of gene expression; indeed the majority of NELF target genes show decreased expression after NELF RNAi [30]. Thus Pho and Spt5 may interact to promote pausing at genes where Pho maintains transcriptional activity [13], although we have no formal evidence of this.

Pho/YY1 has also been shown to associate with the INO80 nucleosome remodelling complex in *Drosophila* and mammalian cells [46,47]. The INO80 complex has been implicated in PcG repression of HOX genes in *Drosophila* [48]. Promoter proximal pausing of RNAP II is linked to a distinctive pattern of nucleosome arrangement around the TSS [25,49,50]. GAF has been shown to cooperate with NURF to remodel nucleosomes and increase DNA accessibility at the paused *Hsp70* promoter [51]. However, GAF is not associated with all genes with paused RNAP II. Very little is known about which factors help to establish the nucleosome architecture at genes with paused RNAP II in general, so a role for the Pho/INO80 complex can not yet be excluded.

## Materials and Methods

### Cloning

cDNA clones were obtained for *Spt5* (GH15287), *pho* (RE17954) *Spt4* (LD44495), and *NELF-B* (GH10333) from the *Drosophila* Genomics Resource Center (DGRC).

The region coding for the last 153 amino acids of Spt5 was amplified using KAPA HiFi mastermix (Anachem) from the GH15287 cDNA clone as the template and cloned in frame into pGBKT7 (Clontech) for use as the bait in the yeast-2 hybrid screen. The entire coding regions of *Spt5* and *pho* were also amplified from the cDNA templates and cloned in frame into pGBKT7 and pGADT7 respectively.

To construct plasmids for expression of tagged proteins in *Drosophila* cell culture, the coding regions of these cDNAs were amplified and appropriate primers and cloned into pENTR-D using a pENTR(TM)/D-TOPO(R) Cloning Kit (Life Technologies - Invitrogen Division). The coding regions were subsequently cloned into pAMW (N-6xMyc tag) and pAFW (N-FLAG tag) (*Drosophila* Gateway Vector Collection distributed by the DGRC) using the Gateway LR Clonase II kit (Life Technologies - Invitrogen Division). The Myc-tagged GFP clone was a gift from Nic Tapon. All primer sequences used for cloning are available on request.

### Yeast 2-hybrid Screen

The yeast 2-hybrid screen was performed using the Matchmaker Gold Yeast Two-Hybrid System from Clontech following the

manufacturer's protocols. Approximately 6.4 million colonies of the Clontech Universal *Drosophila* (Normalized) Mate & Plate Library were screened with the C-terminal 153 amino acids of *Drosophila* Spt5 cloned in frame into pGBKT7. Three independent clones expressing in frame sequences of *pho* were recovered from the screen. These clones did not activate expression of reporter genes in the yeast in the absence of Spt5.

### In Vitro Protein-Protein Interactions

GST pull-down experiments were performed as described previously [52]. pGADT7-pho was translated using the TNT T7 Quick Coupled Transcription/Translation System (Promega). Co-immunoprecipitations were performed essentially as described by [53]. S2R+ cells were grown in Schneider's insect medium (Sigma) supplemented with 10% FCS and 1% Penicillin/Streptomycin, were transfected using Effectene (Qiagen) according to the manufacturer's instructions. Myc-tagged proteins were detected using anti-Myc Antibody (A-14): sc-789 rabbit polyclonal IgG from Santa Cruz and FLAG-tagged proteins using Monoclonal ANTI-FLAG(R) M2 antibody produced in mouse (Sigma). Western blots were visualized using the ECL plus kit (GE Healthcare) and Kodak(R) BioMax(TM) MR film.

### Drosophila Strains

The *pho<sup>ex/ci<sup>D</sup></sup>* stock was a gift from Ana Busturia and *P{neoFRT}82B cu<sup>1</sup> sr<sup>1</sup> NELF-A<sup>KG</sup>/TM3* flies [18] a gift from Peter Gergen. The *Spt5<sup>W049</sup>* stock has been described previously [11]. *y<sup>1</sup> w<sup>1118</sup>; P{lacW}M64 P{lacW}G38 P{lacW}J29 P{EP}wch<sup>EP813</sup> P{lacW}K61 Spt5<sup>MGE-3</sup>/SM1* were obtained from the Bloomington *Drosophila* Stock Center. *Spt5<sup>MGE-3</sup>* was recombined on to *FRT42B* to clean up the chromosome and to facilitate clonal analysis. The *765-Gal4* driver line has been previously described in [54] and the *386Y-Gal4* and *da-GAL4* driver lines were obtained from the Bloomington *Drosophila* Stock Center. The *UAS-RNAi-pho* flies used in this study were *w<sup>1118</sup>; P{GD1509}v39529* from the Vienna *Drosophila* RNAi Center (VDRC) [55]. Somatic clone induction, immunohistochemistry and germ-line clone analysis were carried out as described previously [56]. The *UAS-RNAi-Spt5* flies used in this study were *P{KK101304}VIE-260B* from the Vienna *Drosophila* RNAi Center (VDRC) [55].

### Analysis of Published ChIP-chip Data

Peaks were identified with Ringo [57] using the default parameters and a threshold cutoff of 1.5 for the ratio of Spt5 to input binding on the Nimblegen *Drosophila* tiling array [25] (GEO accession number GSE20472). Similarly peaks were called for NELF-B [25] (GEO accession number GSE20472) with Ringo using the default parameters. The average probe signals were smoothed, and the 95<sup>th</sup> quantile of the log2 immunoprecipitated/input ratio was used as the cutoff value for the detection of peaks. For all other ChIP data, peaks were taken from the published literature. The data for Pho was from modENCODE [22] (<http://www.modencode.org>), NELF-B\*, NELF-E\* and GAF binding was from [31] ArrayExpress ([www.ebi.ac.uk/arrayexpress/](http://www.ebi.ac.uk/arrayexpress/)) under accession number E-MEXP-1547. Identification of overlapping peaks was determined by the ChIPpeakAnno package [58] and if multiple peaks from one factor overlapped with a single peak from another factor, then the peaks would be merged in the calculation of the number of overlapping peaks. Venn diagrams were generated with vennuler [59]. Heat maps of binding sites relative to TSS were generated with the SeqMiner program [60].

## Acknowledgments

We would like to thank Ana Busturia and Peter Gergen for the gift of fly stocks, Birgit Aerne and Nic Tapon for help setting up the co-immunoprecipitation experiments, and Lili Fodor and Helen Stephens for preparing the material for Figure 4. Many thanks to Aamna Kaul, David Ish-Horowitz and Ivana Bjedov for helpful comments on the manuscript.

## References

- Jennings BH (2013) Pausing for thought: Disrupting the early transcription elongation checkpoint leads to developmental defects and tumorigenesis. *BioEssays* 35: 553–560.
- Levine M (2011) Paused RNA polymerase II as a developmental checkpoint. *Cell* 145: 502–511.
- Adelman K, Lis JT (2012) Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* 13: 720–731.
- Zhou Q, Li T, Price DH (2012) RNA polymerase II elongation control. *Annu Rev Biochem* 81: 119–143.
- Werner F (2012) A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J Mol Biol* 417: 13–27.
- Hartzog GA, Fu J (2013) The Spt4-Spt5 complex: a multi-faceted regulator of transcription elongation. *Biochim Biophys Acta* 1829: 105–115.
- Yamaguchi Y, Shibata H, Handa H (2013) Transcription elongation factors DSIF and NELF: promoter-proximal pausing and beyond. *Biochim Biophys Acta* 1829: 98–104.
- Hirtreiter A, Damsma GE, Cheung AC, Klose D, Grohmann D, et al. (2010) Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. *Nucleic Acids Res* 38: 4040–4051.
- Martinez-Rucobo FW, Sainsbury S, Cheung AC, Cramer P (2011) Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J* 30: 1302–1310.
- Guo S, Yamaguchi Y, Schilbach S, Wada T, Lee J, et al. (2000) A regulator of transcriptional elongation controls vertebrate neuronal development. *Nature* 408: 366–369.
- Jennings BH, Shah S, Yamaguchi Y, Seki M, Phillips RG, et al. (2004) Locus-specific requirements for Spt5 in transcriptional activation and repression in *Drosophila*. *Curr Biol* 14: 1680–1684.
- Brown JL, Mucci D, Whiteley M, Dirksen ML, Kassis JA (1998) The *Drosophila* Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. *Mol Cell* 1: 1057–1064.
- Fujioka M, Yusibova GL, Zhou J, Jaynes JB (2008) The DNA-binding Polycomb-group protein Pleiohomeotic maintains both active and repressed transcriptional states through a single site. *Development* 135: 4131–4139.
- Mohd-Sarip A, Venturini F, Chalkey GE, Verrijzer CP (2002) Pleiohomeotic can link polycomb to DNA and mediate transcriptional repression. *Mol Cell Biol* 22: 7473–7483.
- Pattatucci AM, Kaufman TC (1991) The homeotic gene *Sex combs reduced of Drosophila melanogaster* is differentially regulated in the embryonic and imaginal stages of development. *Genetics* 129: 443–461.
- Mahoney MB, Parks AL, Ruddy DA, Tiong SY, Esengil H, et al. (2006) Presenilin-based genetic screens in *Drosophila melanogaster* identify novel notch pathway modifiers. *Genetics* 172: 2309–2324.
- Graveley BR, Brooks AN, Carlson JW, Duff MO, Landolin JM, et al. (2011) The developmental transcriptome of *Drosophila melanogaster*. *Nature* 471: 473–479.
- Wang X, Hang S, Prazak L, Gergen JP (2010) NELF potentiates gene transcription in the *Drosophila* embryo. *PLoS ONE* 5: e11498.
- Kwong C, Adryan B, Bell I, Meadows L, Russell S, et al. (2008) Stability and dynamics of polycomb target sites in *Drosophila* development. *PLoS Genet* 4: e1000178.
- Girton JR, Jeon SH (1994) Novel embryonic and adult homeotic phenotypes are produced by pleiohomeotic mutations in *Drosophila*. *Dev Biol* 161: 393–407.
- Bantignies F, Goodman RH, Smolik SM (2000) Functional interaction between the coactivator *Drosophila* CREB-binding protein and ASH1, a member of the trithorax group of chromatin modifiers. *Mol Cell Biol* 20: 9317–9330.
- Celniker SE, Dillon LA, Gerstein MB, Gunsalus KC, Henikoff S, et al. (2009) Unlocking the secrets of the genome. *Nature* 459: 927–930.
- Mayer A, Lidschreiber M, Siebert M, Leike K, Soding J, et al. (2010) Uniform transitions of the general RNA polymerase II transcription complex. *Nat Struct Mol Biol* 17: 1272–1278.
- Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, et al. (2010) c-Myc regulates transcriptional pause release. *Cell* 141: 432–445.
- Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, et al. (2010) Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* 143: 540–551.
- Beisel C, Bunes A, Roustan-Espinosa IM, Koch B, Schmitt S, et al. (2007) Comparing active and repressed expression states of genes controlled by the Polycomb/Trithorax group proteins. *Proc Natl Acad Sci U S A* 104: 16615–16620.

## Author Contributions

Conceived and designed the experiments: RH BJ. Performed the experiments: RH ES BJ. Analyzed the data: RH ES BJ. Contributed reagents/materials/analysis tools: ES. Wrote the paper: BJ.

- Oktaba K, Gutierrez L, Gagneur J, Girardot C, Sengupta AK, et al. (2008) Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in *Drosophila*. *Dev Cell* 15: 877–889.
- Schuettengruber B, Ganapathi M, Leblanc B, Portoso M, Jaschek R, et al. (2009) Functional anatomy of polycomb and trithorax chromatin landscapes in *Drosophila* embryos. *PLoS Biol* 7: e13.
- Yamaguchi Y, Takagi T, Wada T, Yano K, Furuya A, et al. (1999) NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* 97: 41–51.
- Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, et al. (2008) NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev* 22: 1921–1933.
- Lee C, Li X, Hechmer A, Eisen M, Biggin MD, et al. (2008) NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*. *Mol Cell Biol* 28: 3290–3300.
- Mahmoudi T, Zuijderduijn LM, Mohd-Sarip A, Verrijzer CP (2003) GAGA facilitates binding of Pleiohomeotic to a chromatinized Polycomb response element. *Nucleic Acids Res* 31: 4147–4156.
- Mishra RK, Mihaly J, Barges S, Spierer A, Karch F, et al. (2001) The iab-7 polycomb response element maps to a nucleosome-free region of chromatin and requires both GAGA and pleiohomeotic for silencing activity. *Mol Cell Biol* 21: 1311–1318.
- Busturia A, Lloyd A, Bejarano F, Zavortink M, Xin H, et al. (2001) The MCP silencer of the *Drosophila* Abd-B gene requires both Pleiohomeotic and GAGA factor for the maintenance of repression. *Development* 128: 2163–2173.
- Hendrix DA, Hong JW, Zeitlinger J, Rokhsar DS, Levine MS (2008) Promoter elements associated with RNA Pol II stalling in the *Drosophila* embryo. *Proc Natl Acad Sci U S A* 105: 7762–7767.
- Kwak H, Fuda NJ, Core IJ, Lis JT (2013) Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* 339: 950–953.
- Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, et al. (2007) Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nature Cell Biol* 9: 1428–1435.
- Brookes E, de Santiago I, Hebenstreit D, Morris KJ, Carroll T, et al. (2012) Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates metabolic genes in ESCs. *Cell Stem Cell* 10: 157–170.
- Min IM, Waterfall JJ, Core IJ, Munroe RJ, Schimenti J, et al. (2011) Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev* 25: 742–754.
- Kassis JA, Brown JL (2013) Polycomb group response elements in *Drosophila* and vertebrates. *Adv Genet* 81: 83–118.
- Chopra VS, Hong JW, Levine M (2009) Regulation of Hox gene activity by transcriptional elongation in *Drosophila*. *Curr Biol* 19: 688–693.
- Enderle D, Beisel C, Stadler MB, Gerstung M, Athri P, et al. (2011) Polycomb preferentially targets stalled promoters of coding and noncoding transcripts. *Genome Res* 21: 216–226.
- Ghosh SK, Misra A, Gilmour DS (2011) Negative elongation factor accelerates the rate at which heat shock genes are shut off by facilitating dissociation of heat shock factor. *Mol Cell Biol* 31: 4232–4243.
- Grohmann D, Nagy J, Chakraborty A, Klose D, Fielden D, et al. (2011) The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation. *Mol Cell* 43: 263–274.
- Rougiv AE, Lis JT (1988) The RNA polymerase II molecule at the 5' end of the uninduced *hsp70* gene of *D. melanogaster* is transcriptionally engaged. *Cell* 54: 795–804.
- Klymenko T, Papp B, Fischle W, Kocher T, Schelder M, et al. (2006) A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. *Genes Dev* 20: 1110–1122.
- Cai Y, Jin J, Yao T, Gottschalk AJ, Swanson SK, et al. (2007) YY1 functions with INO80 to activate transcription. *Nat Struct Mol Biol* 14: 872–874.
- Bhatia S, Pawar H, Dasari V, Mishra RK, Chandrasekaran S, et al. (2010) Chromatin remodeling protein INO80 has a role in regulation of homeotic gene expression in *Drosophila*. *Genes Cells* 15: 725–735.
- Gilchrist DA, Adelman K (2012) Coupling polymerase pausing and chromatin landscapes for precise regulation of transcription. *Biochimica et biophysica acta* 1819: 700–706.
- Core IJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, et al. (2012) Defining the status of RNA polymerase at promoters. *Cell Rep* 2: 1025–1035.
- Tsukiyama T, Becker PB, Wu C (1994) ATP-dependent nucleosome disruption at a heat-shock promoter mediated by binding of GAGA transcription factor. *Nature* 367: 525–532.

52. Jennings BH, Pickles LM, Wainwright SM, Roe SM, Pearl LH, et al. (2006) Molecular recognition of transcriptional repressor motifs by the WD domain of the Groucho/TLE corepressor. *Mol Cell* 22: 645–655.
53. Genevet A, Wehr MC, Brain R, Thompson BJ, Tapon N (2010) Kibra is a regulator of the Salvador/Warts/Hippo signaling network. *Dev Cell* 18: 300–308.
54. Jennings BH, Tyler DM, Bray SJ (1999) Target specificities of *Drosophila* enhancer of split basic helix-loop-helix proteins. *Mol Cell Biol* 19: 4600–4610.
55. Dietzl G, Chen D, Schnorrer F, Su KC, Barinova Y, et al. (2007) A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448: 151–156.
56. Jennings BH, Wainwright SM, Ish-Horowicz D (2008) Differential *in vivo* requirements for oligomerization during Groucho-mediated repression. *EMBO Rep* 9: 76–83.
57. Toedling J, Skylar O, Krueger T, Fischer JJ, Sperling S, et al. (2007) Ringo—an R/Bioconductor package for analyzing ChIP-chip readouts. *BMC Bioinformatics* 8: 221.
58. Zhu LJ, Gazin C, Lawson ND, Pages H, Lin SM, et al. (2010) ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* 11: 237.
59. Wilkinson L (2012) Exact and approximate area-proportional circular Venn and Euler diagrams. *IEEE transactions on visualization and computer graphics* 18: 321–331.
60. Ye T, Krebs AR, Choukrallah MA, Keime C, Plewniak F, et al. (2011) seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res* 39: e35.
61. Xu T, Rubin GM (1993) Analysis of genetic mosaics in developing and adult *Drosophila* tissues. *Development* 117: 1223–1237.
62. Chou TB, Perrimon N (1996) The autosomal FLP-DFS technique for generating germline mosaics in *Drosophila melanogaster*. *Genetics* 144: 1673–1679.