A Conserved Nuclear Cyclophilin Is Required for Both RNA Polymerase II Elongation and Co-transcriptional Splicing in *Caenorhabditis elegans*

Jeong H. Ahn[1,2], Andreas Rechsteiner[3], Susan Strome[3], William G. Kelly[1] *

1 Biology Department, Emory University, Atlanta, Georgia, United States of America, **2** Program in Genetics and Molecular Biology, Emory University, Atlanta, Georgia, United States of America, **3** Department of Molecular, Cell and Developmental Biology, University of California Santa Cruz, Santa Cruz, California

* bkelly@emory.edu

Abstract

The elongation phase of transcription by RNA Polymerase II (Pol II) involves numerous events that are tightly coordinated, including RNA processing, histone modification, and chromatin remodeling. RNA splicing factors are associated with elongating Pol II, and the interdependent coupling of splicing and elongation has been documented in several systems. Here we identify a conserved, multi-domain cyclophilin family member, SIG-7, as an essential factor for both normal transcription elongation and co-transcriptional splicing. In embryos depleted for SIG-7, RNA levels for over a thousand zygotically expressed genes are substantially reduced, Pol II becomes significantly reduced at the 3' end of genes, marks of transcription elongation are reduced, and unspliced mRNAs accumulate. Our findings suggest that SIG-7 plays a central role in both Pol II elongation and co-transcriptional splicing and may provide an important link for their coordination and regulation.

Author Summary

mRNA splicing can occur co-transcriptionally; i.e., splicing occurs as the RNA emerges from the RNA Polymerase II holoenzyme during transcription elongation. Recent studies suggest that defective splicing can cause defective transcription elongation, suggesting an interdependency of the two mechanisms. The *C. elegans* gene *sig-7* encodes a nuclear cyclophilin, a highly conserved protein family characterized by the presence of both a peptidyl isomerase domain and an RNA-recognition motif. Studies on *sig-7* homologs in plants and fission yeast have shown these proteins to interact with RNA Polymerase II, and indicate they regulate the phosphorylation status of its C-terminal domain. We show that SIG-7 activity is essential for both efficient co-transcriptional splicing and normal RNA Polymerase II elongation and may provide an important link between the two processes.

Introduction

Transcription by RNA Polymerase II (Pol II) is a highly regulated process involving coordination of multiple processes that together modulate the level of gene expression and its temporal and spatial control [1–4]. Epigenetic mechanisms play important roles in both transcription initiation and elongation, with various histone modifications both guiding and resulting from these processes [5–8]. Kinases also regulate both stages by modifying the C-terminal domain (CTD) of Pol II's catalytic subunit and phosphorylating other factors that regulate the transitions accompanying Pol II transcription, including promoter-proximal pausing [9–16]. The CTD is composed of a conserved heptapeptide repeat, and phosphorylation of specific serines and threonines within the repeats correlates with these transitions [2, 17–25].

A connection between the modifications of the CTD and mRNA splicing has long been observed. It has been shown that the association of SR (Serine/Arginine-rich) splicing factors with the CTD requires phosphorylation of Ser2 of the heptapeptide repeat [26–28]. It has also been observed that some splicing factors are required for normal RNA Pol II elongation, suggesting a reciprocal mechanistic relationship between RNA processing and transcription elongation [29–32]. RNA processing in the nucleus is largely co-transcriptional, so an interdependency of splicing and Pol II elongation represents a potentially important mode of transcription regulation.

In addition to kinases and histone modifying enzymes, peptidyl proline isomerases (PPIs) can regulate Pol II during transcription progression. The nuclear parvulin family of PPIs direct *cis-trans* isomerization of prolines in the context of Ser/Thr, such as those found in the Pol II CTD heptapeptide repeats, and the activity of these PPIs is affected by the phosphorylation of Ser/Thr [33–36]. These PPIs are thought to contribute to structural regulation of the CTD, participating in a "CTD code" that controls the recruitment of various factors to Pol II during elongation and transcript processing [33, 37–39].

The nuclear cyclophilin PPI family, characterized by having an RNA-recognition motif (RRM) in addition to a PPI domain has also been implicated in regulation of Pol II through interactions with the CTD. Members of this highly conserved family include KIN241 in *Paramecium tetraurelia*, AtCyp59 in *Arabidopsis thaliana* and Rct1 in *Schizosaccharomyces pombe* [40–43]. AtCyp59 interacts with Pol II, and its overexpression causes defective regulation of Pol II CTD phosphorylation [41]. AtCyp59 also interacts with RNA through its RRM domain and has PPI activity, but whether the PPI domain is required for AtCyp59 function is unclear [42]. The *S. pombe* Rct1 also interacts with and affects Pol II CTD phosphorylation, and the effect on phosphorylation is dependent on Rct1's PPI domain [43, 44].

Here we present the first *in vivo*, genome-wide analysis of a *C. elegans* nuclear cyclophilin, SIG-7, and show that it is essential for normal transcription and RNA processing during embryogenesis. Loss of SIG-7 results in a genome-wide decrease in mRNA production that is correlated with both defective elongation and defective co-transcriptional splicing. Our results identify SIG-7 as a conserved and important factor for both efficient Pol II transcription elongation and co-transcriptional splicing.

Results

sig-7 encodes a conserved nuclear cyclophilin that is essential for development

A recessive mutation in *sig-7 (*silencer in germline 7) was recovered years ago from an unpublished screen for defective silencing of repetitive transgenes in germ cells. The *sig-7(cc629)* allele exhibits multiple developmental defects in both soma and germline (S1 Fig). The *cc629*

mutation was mapped to a small region of LG I, and a mutation in a predicted splice acceptor site in gene F39H2.2 was identified by sequencing (S2B Fig). Sequencing of F39H2.2 cDNAs from *cc629* animals showed the predicted splicing error in most, but a few correctly spliced cDNAs were also recovered, consistent with the lack of a strict requirement for a canonical AG dinucleotide at the 3' end of pre-mRNA introns [45]. RNAi targeting F39H2.2 resulted in embryonic lethality with rare "escapers" growing up to exhibit the same spectrum of pleiotropic phenotypes observed in the *cc629* animals. The *cc629* allele is thus hypomorphic, with partial maternal rescue of homozygous offspring produced by heterozygous mothers (see below). An additional deletion allele, *n5037* (a deletion allele from L. Ma and R. Horvitz, S2B Fig), causes early larval arrest. The early embryonic arrest caused by *F39H2.2 RNAi* and the larval arrest of *n5037* homozygous offspring from heterozygous mothers is indicative of both maternal and zygotic requirements for F39H2.2 function. Because F39H2.2 (hereafter called *sig*-7) is the third gene in an operon (CEOP1492; S2A Fig), a *sig-7::gfp::3XFLAG* translational fusion transgene was generated from a fosmid clone (TransgeneOme Project) encompassing the entire operon [46]. This transgene, when integrated as a single copy using MosSCI, rescues *sig-7(n5037)* animals to fertile adults. The rescued *n5037* deletion strain expressing SIG-7:: GFP::3xFLAG was used in the experiments described below.

SIG-7 possesses an N-terminal peptidyl prolyl isomerase (PPI) domain and an adjacent RNA-recognition motif (RRM) (S2C Fig). The C-terminal region is of low complexity, characterized by the presence of many charged residues including RS and RD dipeptides (S3 Fig). SIG-7 is the sole *C. elegans* ortholog of a highly conserved family of PPI and RRM domain-containing nuclear cyclophilins found in most eukaryotes from fission yeast to humans, but notably absent from *S. cerevisiae* [47]. SIG-7 homologs share greater than 39% overall amino acid identity, most of which is concentrated within the PPI and RRM domains, with the highest degree of sequence identity in the RRM domain (S2D and S3 Figs).

SIG-7 localizes to transcriptionally active chromatin

A. thaliana AtCyp59 and *S. pombe* Rct1 interact with the CTD of Pol II and serve roles in Pol II regulation (41, 43). SIG-7::GFP::3XFLAG transgene localizes to the nucleus in all tissues including germ cells (S4 Fig). In adult germ cells, the localization overlaps with DNA in mitotic and meiotic nuclei, but with a somewhat broader distribution than DAPI-intense chromatin (Figs 1 and S4D). In diakinetic oocytes, SIG-7::GFP loses its chromatin association and becomes diffuse within the nucleoplasm (S4D Fig). This transition correlates with the loss of Pol II from chromatin and the presumed global cessation of transcription in late oogenesis [48]. We further examined the correlation between SIG-7 staining and transcriptionally active chromatin in meiotic germ cells. Transcriptional activity is repressed on the X chromosomes relative to the autosomes during *C. elegans* meiosis: the X chromosomes are easily identified by their significantly lower levels of AMA-1 (the catalytic subunit of *C. elegans* Pol II), and H3K36me3 and H3K4me2, chromatin marks of transcription [48]. SIG-7::GFP localization in meiotic chromatin exhibits the same pattern as AMA-1 and H3K36 and H3K4 methylation i.e., abundant on all autosomes and depleted from the X chromosomes (Fig 1). Thus, SIG-7 associates with chromatin and co-localizes with active transcription.

We next examined *sig-7(RNAi)* embryos for defects in gastrulation, a sensitive indicator of zygotic transcription defects in embryos. Gastrulation in *C. elegans* consists of the inward migration of a few peripheral cells, including the P4 cell, which is the progenitor of the two primordial germ cells, Z2 and Z3 [49, 50]. Gastrulation is largely completed by the ~80 cell-stage, with P4 having migrated to the interior and subsequently divided to yield internally localized Z2/Z3, which are readily identified using antibodies that recognize germline-specific P-granules
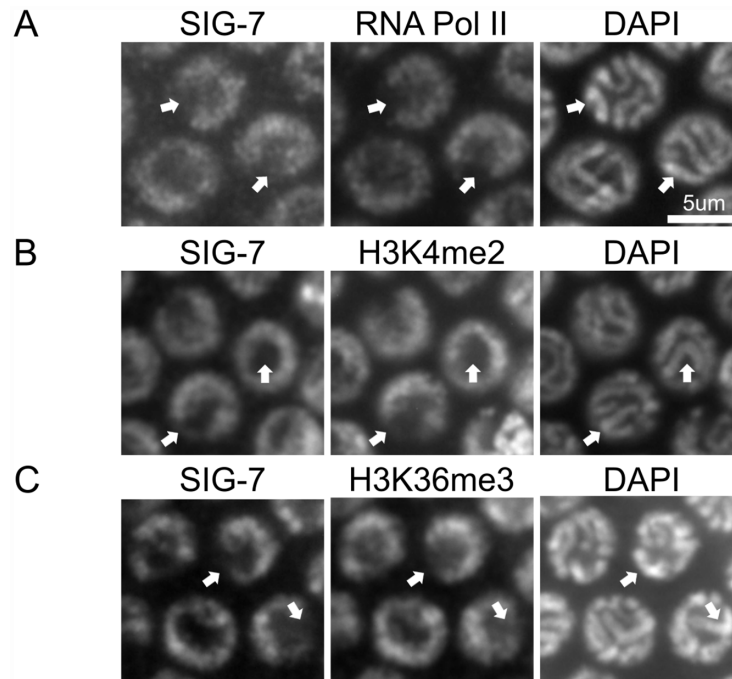
Fig 1. SIG-7 is associated with transcriptionally active chromatin. A-C) The distribution of SIG-7::GFP in pachytene-stage germ cells was compared to DAPI staining of DNA and the distribution of RNA Polymerase II (A), H3K4me2 (B), and H3K36me3 (C). Arrows point to the pair of X chromosomes, which in germ cells have lower transcriptional activity than the autosomes.

doi:10.1371/journal.pgen.1006227.g001

(Fig 2A, L4440 control). Zygotic gene activation is required for gastrulation, and disruption of embryonic Pol II or other essential transcription activities in embryos results in a failure of P4 to migrate internally, causing Z2 and Z3 to be born at the periphery (Fig 2A; *ama-1(RNAi)*). RNAi targeting of either *ama-1* or *sig-7* caused a highly penetrant gastrulation phenotype, yielding 92.5% and 86.15% gastrulation-defective embryos, respectively (Fig 2A and 2B). Thus, SIG-7 is required for normal zygotic transcription during early embryonic development.

The *sig-7(RNAi)* gastrulation phenotype could be due to inactivation of one or a few genes specifically involved in gastrulation. As a first test for a more widespread defect, we quantified transcript levels of a panel of genes with strictly zygotic expression in *sig-7(RNAi)* embryos [51–53]. Significant decreases for all tested zygotic transcripts were observed (Fig 2C). The decreases were substantial, albeit not as dramatic as those observed in *ama-1(RNAi)* embryos. Thus, both molecular and phenotypic data indicate that loss of SIG-7 activity leads to reduced levels of zygotic gene expression.

SIG-7 physically interacts with Pol II *in vivo*

We next tested for physical interactions between SIG-7 with Pol II in *C. elegans* by immunopre-cipitating AMA-1 from transgene-rescued *sig-7(n5037)* animals, followed by probing western blots of the precipitated material with anti-FLAG antibodies to detect SIG-7::GFP::3xFLAG. SIG-7 co-precipitated with AMA-1 (Fig 3), and in reciprocal experiments AMA-1 co-precipi-tated with SIG-7 (S5 Fig). These results indicate that SIG-7 interacts with Pol II *in vivo*.
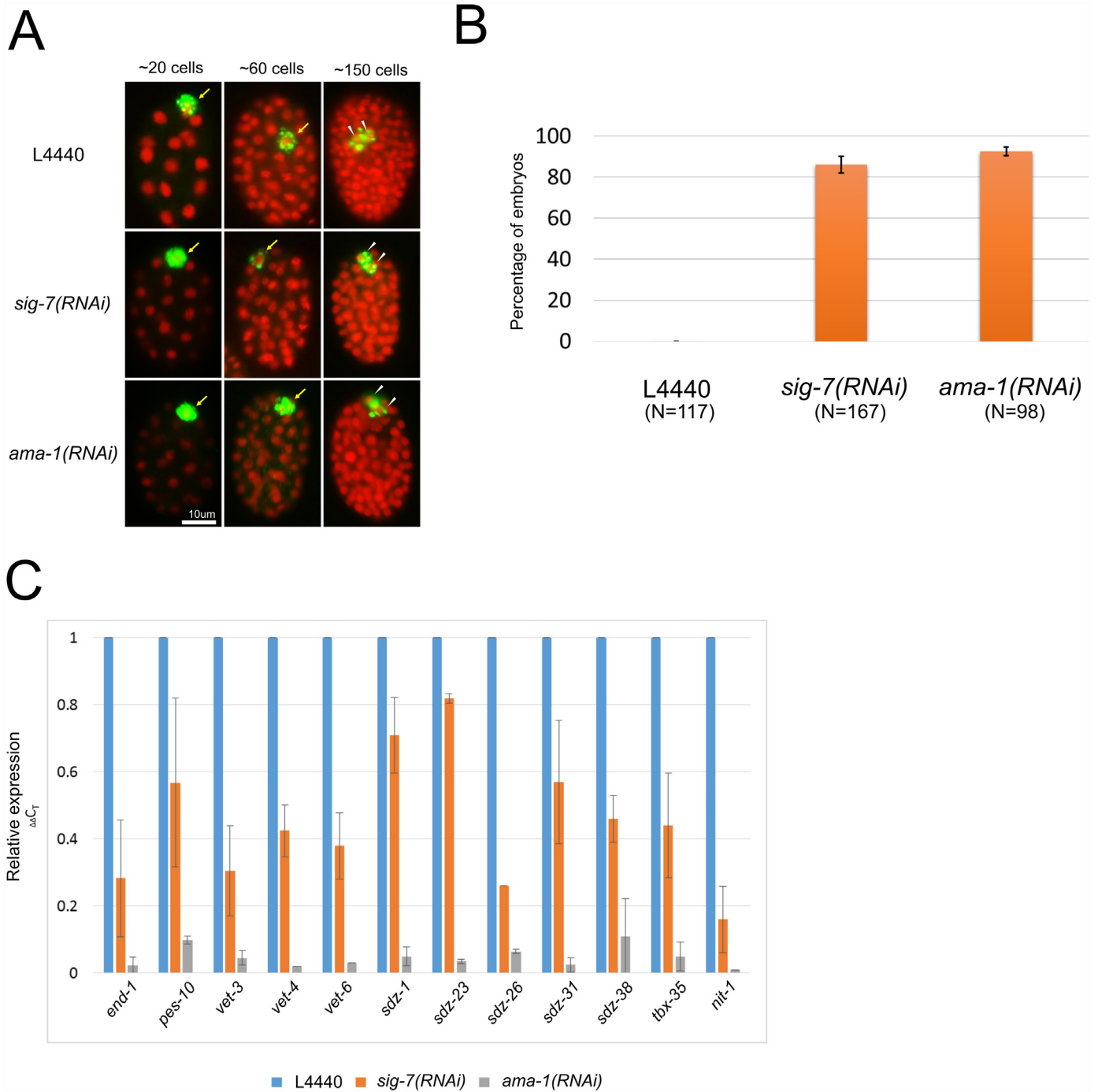
Fig 2. *sig-7* is required for hallmarks of zygotic transcription. A) RNAi control (L4440), *sig-7(RNAi)* and *ama-1(RNAi)* embryos were fixed and stained with DAPI (red) and antibody OIC1D4, a marker of germ cells (green). In L4440 controls embryos, the germline blastomere P4 is initially at the periphery of the embryo (~20 cells, arrow), migrates to the interior of the embryo (~60 cells, arrow) and then divides to produce the primordial germ cells Z2 and Z3 (~150 cells, arrowheads). In *sig-7(RNAi)* and *ama-1(RNAi)* embryos, both P4 and its daughters Z2 and Z3 remain at the periphery of the embryo. B) Quantification of the % of embryos in which P4 failed to migrate into the interior of the embryo. Error bars = S.D. from two biological replicates. C) The RNA levels of a panel of zygotically expressed genes were measured by qRT-PCR in control (L4440), *sig-7(RNAi)*, and *ama-1(RNAi)* embryos. Each sample was normalized to18S RNA levels, and *sig-7(RNAi)* and *ama-1(RNAi)* were plotted relative to L4440 control. Error bars = S.D. from two technical replicates each of two biological replicates.

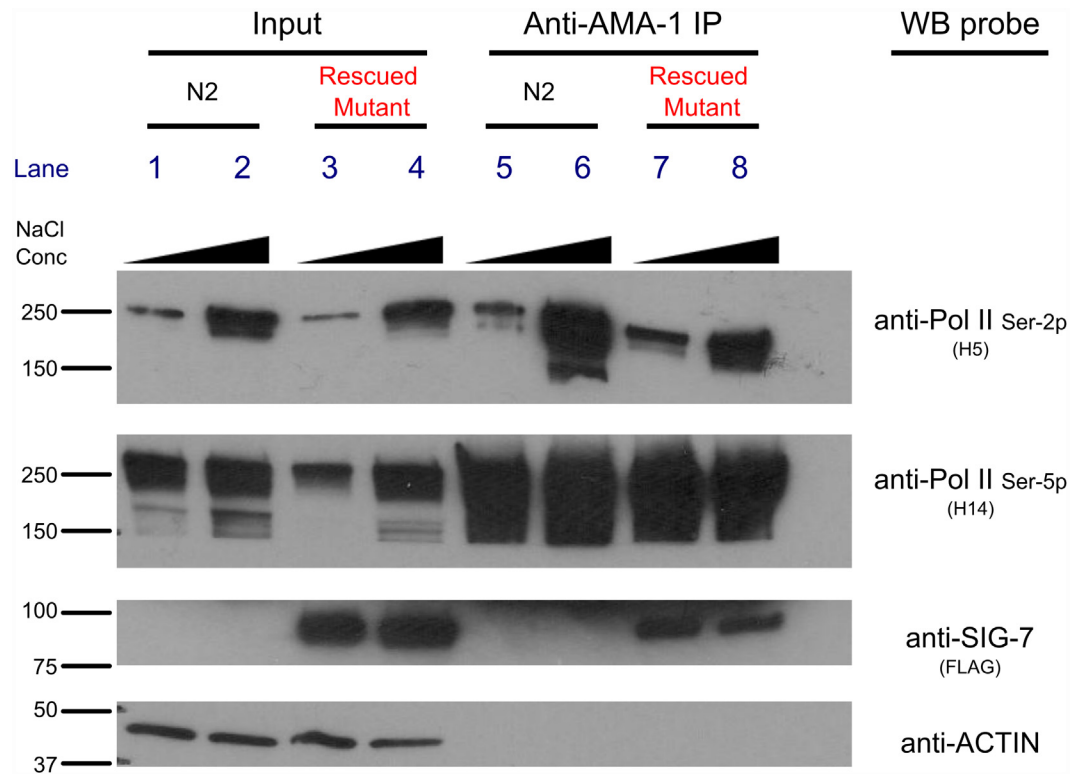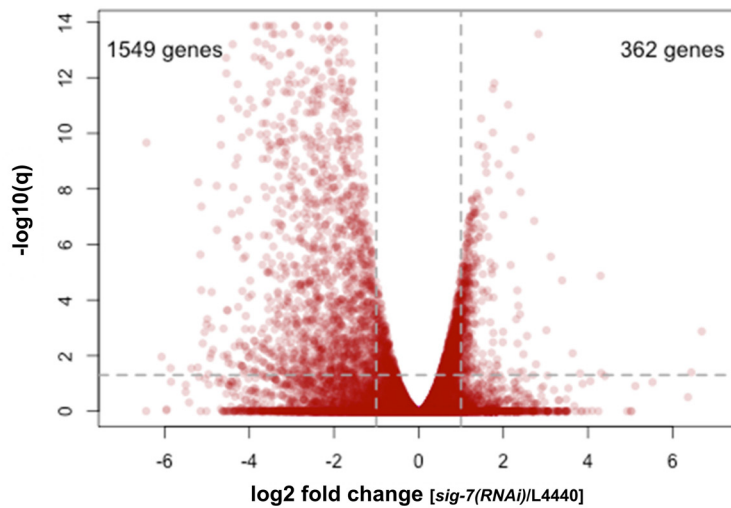doi:10.1371/journal.pgen.1006227.g002

Fig 3. SIG-7 associates with RNA Pol II *in vivo*. Embryo lysates were prepared from wild type (N2) or *sig-7 (n5037)* mutants rescued by a *sig-7*::GFP::3XFLAG transgene, using either normal salt (150 mM NaCl; odd lanes) or high salt (420 mM NaCl; even lanes) extraction conditions. RNA Pol II was immunoprecipitated (IP) using an antibody that recognizes all isoforms (anti-AMA-1), and the IP material was probed by western blot using the antibodies indicated. 5% input was used for SDS-PAGE. The blots were also probed with anti-actin both to normalize for the amount of total protein used for the IP (bottom panel, lanes 1–4) and to determine the specificity of the co-IP (bottom panel; lanes 5–8). The expected sizes of the proteins assayed are as follows: SIG-7::3XFLAG:: GFP, 84 KD; AMA-1 (phosphorylated), 210–250 KD; Actin, 42 KD.

doi:10.1371/journal.pgen.1006227.g003

sig-7 RNAi causes a global decrease in embryonic transcript levels

To further explore the extent of SIG-7's role in gene expression, we next performed RNA-seq on *sig-7(RNAi)* embryos and on L4440 RNAi control embryos ([Fig 4](#)). The results revealed that *sig-7* RNAi causes a global defect in embryonic gene expression. Of the 45,627 annotated genes (including non-coding RNAs, etc.), 10,703 had sufficient read representation for further analysis. Of these, 3,045 genes displayed significantly different RNA accumulation ($q \leq 0.05$) between *sig-7(RNAi)* and L4440 control RNAi samples ([S2 Table](#)). Many more genes were down-regulated at least 2-fold (1549) than were up-regulated at least 2-fold (362) in *sig-7(RNAi)* ([Fig 4A](#)). We sorted these genes into gene categories based on published evidence for either zygotic expression during embryonic development ("soma-specific", "embryo–expressed", and "X-linked"), exhibiting "ubiquitous" expression, or displaying expression enriched in or restricted to the germline ("germline-enriched" and "germline-specific", respectively) [54]. X-linked genes show a distinct bias for either having weak expression in germ cells or only being expressed in somatic lineages [55]. Genes categorized as soma-specific, embryo-expressed, or X-linked were significantly over-represented among the down-regulated genes, and genes categorized as germline-expressed were significantly under-represented ([Fig 4B](#), left panel). This
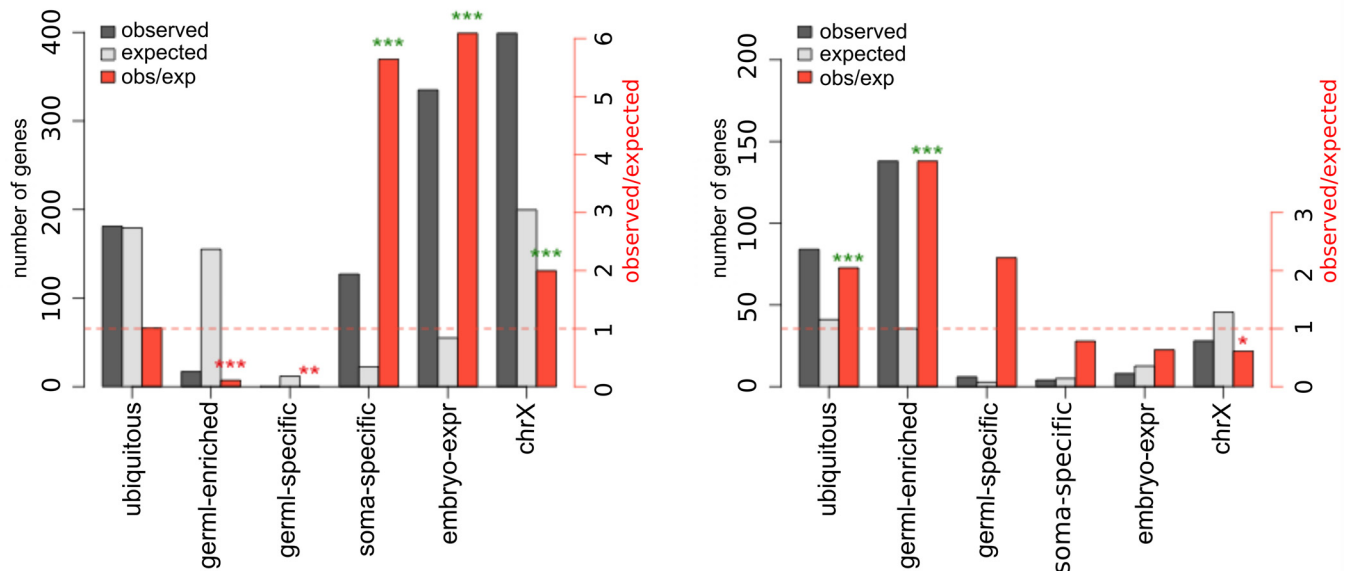
Fig 4. RNA-seq analysis of *sig-7(RNAi)* embryos reveals a global transcription defect. RNA-seq was performed on RNA purified from *sig-7 (RNAi)* and L4440 control RNAi embryos and the profiles compared. A) Volcano plot of normalized RNA-seq results showing the number of genes exhibiting significantly decreased (1549) and increased (362) RNA abundance in *sig-7(RNAi)* compared to L4440 (>2-fold differences with q value < 0.05). (B) Genes exhibiting significantly decreased (left) or increased (right panel) transcript accumulation in *sig-7(RNAi)* compared to L4440 were classified as ubiquitous, germline (germl)-enriched, germline(germl)-specific, soma-specific (broadly defined; soma-all), embryo-expressed (embryo-expr), or X-linked (chrX). Gene class definitions are described in Materials and Methods. The number of genes observed and expected for each class (left y-axis), and the ratio of observed/expected (right y-axis) are shown. The gene classes with zygotic expression are highly overrepresented in those showing reduced expression in *sig-7(RNAi)* embryos, while genes with germline expression are highly overrepresented in those showing higher expression in *sig-7(RNAi)* embryos. The significance of the enrichment (green asterisks) or depletion (red asterisks) as determined by the hypergeometric distribution for each gene class is indicated by asterisks (* = p ≤ 0.01, ** = p ≤ 10^{-5}, *** = p ≤ 10^{-10}).

doi:10.1371/journal.pgen.1006227.g004

pattern was reversed for the up-regulated genes: germline-expressed genes, including ubiquitous and germline-enriched genes, were significantly over-represented (Fig 4B, right panel).

We investigated whether the different effects of loss of SIG-7 on germline versus somatic transcripts in embryos reflected a different requirement of those tissues for SIG-7, or weaker germline RNAi effects using the standard feeding technique from L3 stage. Favoring the latter possibility, we found that extended RNAi for longer periods as adults resulted in significant reduction of germline-expressed genes in both *sig-7* and *ama-1* RNAi adult animals (S6B Fig), and RNAi starting from earlier stages caused sterility. Thus, SIG-7 is required for efficient RNA production in larval and adult germ cells as well as in embryos. The differential effect of loss of SIG-7 from mothers on germline versus soma transcripts in embryos is probably due to lower efficiency of RNAi in adult germ cells using standard feeding protocols from the L3 stage.

We also considered whether the embryonic arrest phenotype was skewing the effect on genes expressed in later-stage embryos, since we compared *sig-7(RNAi)* embryos that mostly arrest at ~200–300 cells with control embryos that can continue to develop. The impact of stage differences on our results is probably low. The embryos used in our experiments were isolated from young adults with developing embryos in their uterus; these embryos are highly enriched for stages prior to the *sig-7(RNAi)* arrest point. Analysis of the embryo stage distributions from independent RNAi experiments showed the expected bias for early stages (e.g. S7 Fig). We further compared the affected genes from our experiments with those analyzed in a landmark study examining transcript dynamics at early *C. elegans* embryonic developmental time points, all of which are earlier than the *sig-7(RNAi)* arrest point [51]. We focused on three narrowly defined gene sets: "strictly maternal" (expressed in the ovary and degraded in the early embryo), "maternal/embryonic" (expressed in both ovary and embryos), and "strictly embryonic" (expressed only by the embryo with no maternal contribution). Of the genes classified as "strictly embryonic" that in our analyses showed >2-fold changes in *sig-7(RNAi)* embryos, 328/339 (96.7%) were down-regulated and only 11/339 (3.2%) were significantly up-regulated (S8 Fig). "Maternal/embryonic" and "strictly maternal" genes showed less bias, with 202/310 (65%) and 77/163 (47.3%) showing down-regulation, respectively, in *sig-7(RNAi)* embryos. The increase of several strictly maternal genes in *sig-7(RNAi)* embryos was verified by qRT-PCR (e.g. S6A Fig). The increased abundance of strictly maternal RNAs in *sig-7(RNAi)* embryos may be an indirect effect of defective zygotic transcription-driven development, causing impaired degradation of maternal RNAs [56].

sig-7 RNAi causes changes in RNA processing

Our RNA-seq analyses also revealed a role for SIG-7 in RNA splicing. 1431 of the 1549 genes down-regulated in *sig-7(RNAi)* also registered "isoform differences" in our Cuffdiff analyses (S2 Table). Upon closer examination, many of these *sig-7(RNAi)*-dependent "isoform differences" appeared to be caused by decreases in exon reads without corresponding decreases in intron reads (S9A Fig). Genome-wide analyses also revealed this trend: the average exon read coverage of genes showing decreased expression in *sig-7(RNAi)* embryos showed the expected decrease, but the intron read coverage showed little change relative to controls (S9B Fig). Thus, although the amount of RNA for these genes was decreased, the ratio of intron to exon reads for these RNAs increased. Importantly, many of the intron read sequences were linked to exon sequences, indicating they were from unprocessed transcripts, rather than from abnormal persistence of spliced-out intron segments.

We confirmed an increase in intron abundance compared to exons for several affected genes by qRT-PCR (Fig 5A). We used primer sets that span intron-exon junctions and exon-exon junctions to distinguish unspliced primary transcripts (pre-mRNAs) from spliced mature
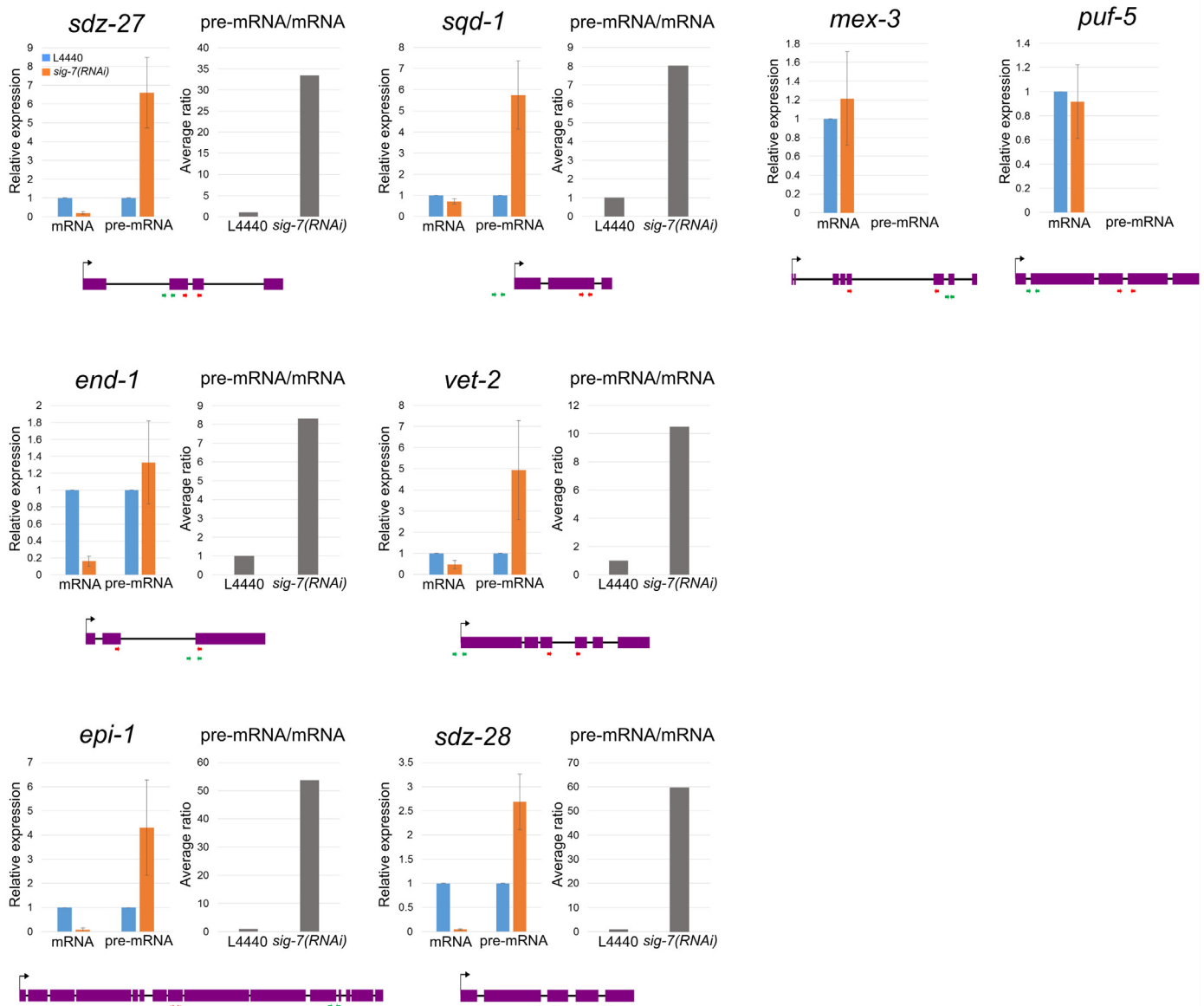
Fig 5. qRT-PCR analysis of the effect of *sig-7(RNAi)* on splicing. qRT-PCR was performed on several strictly embryonic (A), and strictly maternal genes (B) affected by *sig-7* RNAi. In order to distinguish unprocessed pre-mRNA from mRNA, primers were designed to amplify either outron, introns, or intron-exon junctions as indicated with green arrows. Other primers were designed to amplify only spliced exon-exons of mRNAs as indicated with red arrow. The relative abundances measured by qRT-PCR of respective RNA are shown for each gene as are the ratios of pre-mRNA/mRNA. For maternally expressed genes, no pre-mRNAs were detected. The expression for each was normalized to 18S RNA and plotted relative to L4440 controls in each experiment. Error bars = S.D. from three technical replicates each of four biological replicates.

doi:10.1371/journal.pgen.1006227.g005

mRNAs (mRNAs), respectively. The six embryonic genes tested (*sdz-27*, *sdz-28*, *epi-1*, *sqd-1*, *vet-2*, *end-1*) displayed the markedly reduced mRNA levels in *sig-7(RNAi)* embryos observed by RNA-seq. All six also showed significantly increased levels of unspliced RNAs, confirming that the reduced transcripts present in *sig-7(RNAi)* embryos are enriched for defectively processed RNAs (Figs 5A and S9A).

In contrast, RNAs from "upregulated" genes, such as strictly maternal genes, showed the opposite trend. RNA-seq results for these genes showed an increase in exon reads, yet their intron reads stayed relatively constant in *sig-7(RNAi)* embryos, suggesting a relative enrichment for spliced RNAs relative to controls (S9B Fig). Indeed, intron sequences could not be detected by qRT-PCR for two maternal genes tested (Fig 5A). This result is also consistent with abnormal persistence of fully spliced maternal products, resulting in an apparent enrichment for exon reads relative to intron reads compared to controls.

Reads from sequences 5' to the first exon of many of the down-regulated genes also increased in *sig-7(RNAi)* embryos relative to controls (Figs 5A and S9A). These reads represent 5' outrons, which like introns are removed from the primary transcripts. *C. elegans* exhibits co-transcriptional *trans*-splicing, in which a common spliced leader transcript serves as a 5' splice donor, leading to a common 5' exon that is present on the majority of mRNAs [57–59]. In *C. elegans*, approximately 70% of mRNAs are reported to be trans-spliced [60]. The outron reads thus represent 5' nascent transcript sequences that are normally removed by trans-splicing and replaced by spliced leader sequences during transcription. Indeed the 5' reads enriched in *sig-7 (RNAi)* embryos precisely mark the transcription start sites (TSSs) recently identified by GRO-seq and related methods [61–63]. The relative increase of RNA-seq reads corresponding to introns and outrons indicates that depletion of SIG-7 causes defects in both *cis-* and *trans*-splicing, the latter of which is only known to occur co-transcriptionally [64, 65]. Since SIG-7 interacts with Pol II, this strongly suggests that SIG-7 plays an important role in transcription-coupled RNA processing events.

A transcription defect could indirectly cause splicing defects by reducing the production of essential splicing factors. This seemed unlikely, since splicing factors in early embryos are available from maternal stores and thus would fall into the class of genes either unaffected or slightly enriched in *sig-7(RNAi)* embryos. Indeed, our RNA-seq data confirmed this: of 18 conserved *C. elegans* splicing factors [66] for which significant RNA levels could be detected in control embryos, 8 factors showed a slight increase in *sig-7(RNAi)* embryos, and 10 factors showed no significant difference between *sig-7(RNAi)* and control (S1 Table). Thus, the splicing defects in *sig-7(RNAi)* embryos are unlikely to be due to reduced expression of splicing factors and are instead likely to be directly due to defects in transcription coupled processing.

sig-7 RNAi causes a global change in Pol II occupancy and distribution within gene bodies

Numerous reports indicate that co-transcriptional splicing is mechanistically coupled to Pol II elongation, and it has recently been proposed that defects in co-transcriptional splicing can affect Pol II elongation [67–74]. We therefore analyzed the genome-wide distribution of Pol II by anti-AMA-1 ChIP-seq in *sig-7(RNAi)* and control embryos. The ChIP-seq data showed a strong correlation with the RNA-seq data; i.e., genes showing down-regulation by RNA-seq also showed decreased Pol II occupancy by ChIP-seq (S10 and S11 Figs). We next performed metagene analyses of the Pol II distribution within the body of genes in five of the expression categories described above (Fig 6). Genes classified as either "soma-specific" or "ubiquitous" showed substantial changes. In these genes, the 3' enrichment of Pol II observed in control embryos was significantly reduced in *sig-7(RNAi)* embryos, with 3' depletion the most obvious in the "soma-specific" class (Fig 6). 5' localization was also reduced for the "soma-specific" class, but the effect was less marked than the 3' reduction.

In contrast, genes classified as "germline-enriched" showed little change in Pol II distribution. This result indicates that, as with the RNA-seq data, there is a disproportionate effect of
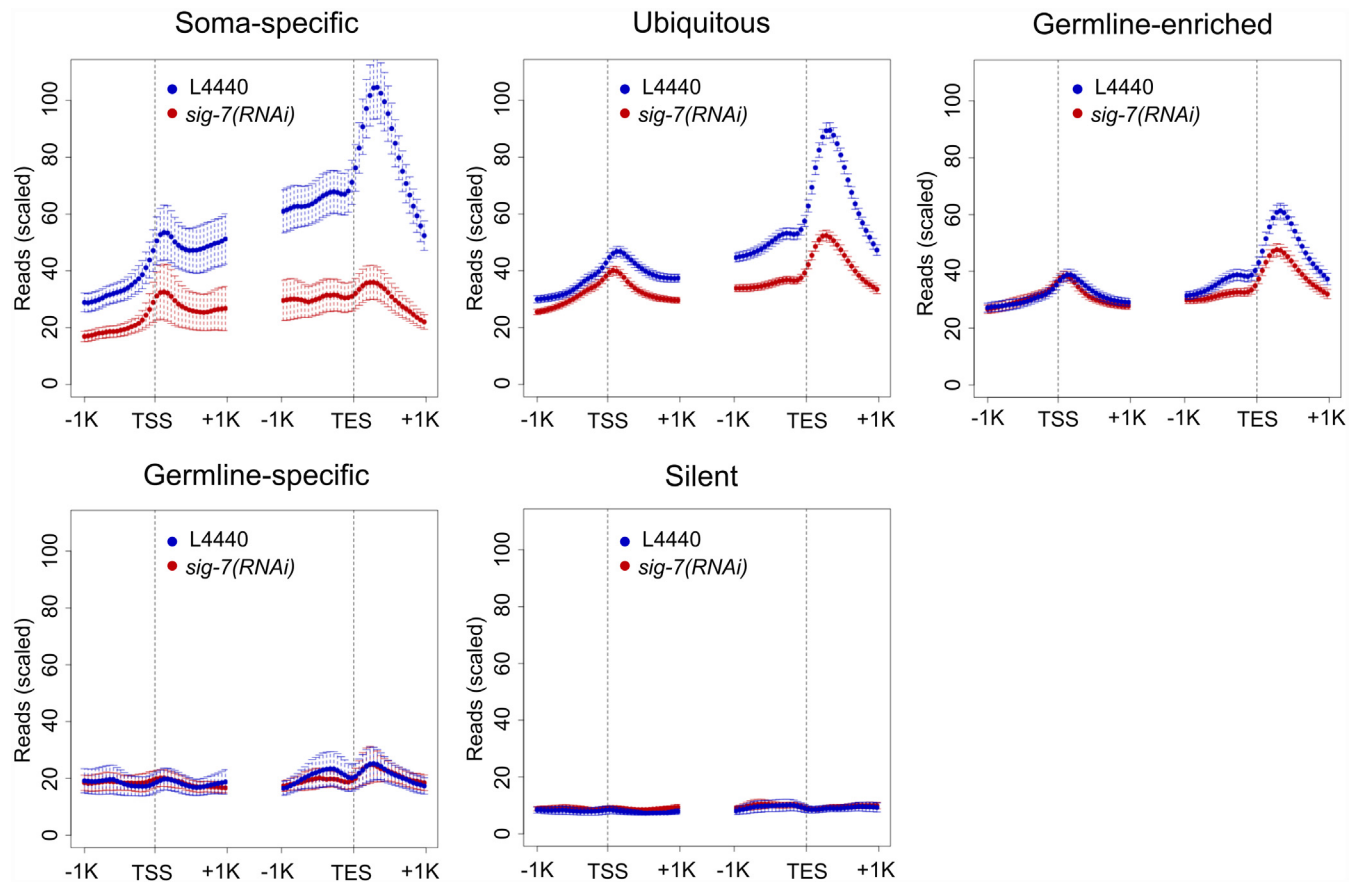
Fig 6. RNA Pol II distribution within genes is altered by *sig-7(RNAi)*. Metagene displays of RNA Pol II ChIP-seq enrichment 1kb upstream and downstream of the Transcription Start Site (TSS) and Transcription End Site (TES) for different classes of genes in L4440 RNAi control (blue) and *sig-7(RNAi)* (red). Genes were categorized as in Fig 4. Graphs illustrate combined results from 2 biological replicates. Error bars indicate the 95% confidence interval of the mean signal, indicated by the circles. Reads were normalized as indicated in Materials and Methods. Note that "germline-enriched" genes include genes expressed in all tissues that exhibit enhanced expression in germ cells.

doi:10.1371/journal.pgen.1006227.g006

sig-7(RNAi) on genes expressed in embryos, including an effect on steady-state localization of Pol II within gene bodies. The lack of effect on "germline-enriched" loci is not as easy to ascribe to reduced *sig-7* RNAi efficiency in parental germlines compared to embryos, since many of these genes include ubiquitously expressed genes transcribed in embryos. The reduced effect of SIG-7 depletion for these genes may be related to the different modes of Pol II regulation observed for germline- and ubiquitously-expressed genes compared to soma-specific genes, the latter of which involve tissue-specific modes of gene regulation [75].

Depletion of SIG-7 causes a reduction in Pol II isoforms and histone modifications associated with transcription elongation

The decrease in Pol II at the 3' end of gene bodies observed by ChIP-seq suggested that *sig-7* RNAi affects the elongation phase of transcription. The phosphorylation of specific residues in the Pol II CTD correlates with different stages of the transcription cycle; e.g., Ser-5P correlates with initiation and Ser-2P increases with elongation [2, 17, 76–78]. We assessed the relative abundances of these CTD phospho-epitopes in *sig-7(RNAi)* and L4440 control embryos using monoclonal antibodies specific for the different phosphorylated isoforms of AMA-1 (Fig 7A). We observed similar levels of AMA-1 protein in experimental and control lanes, indicating
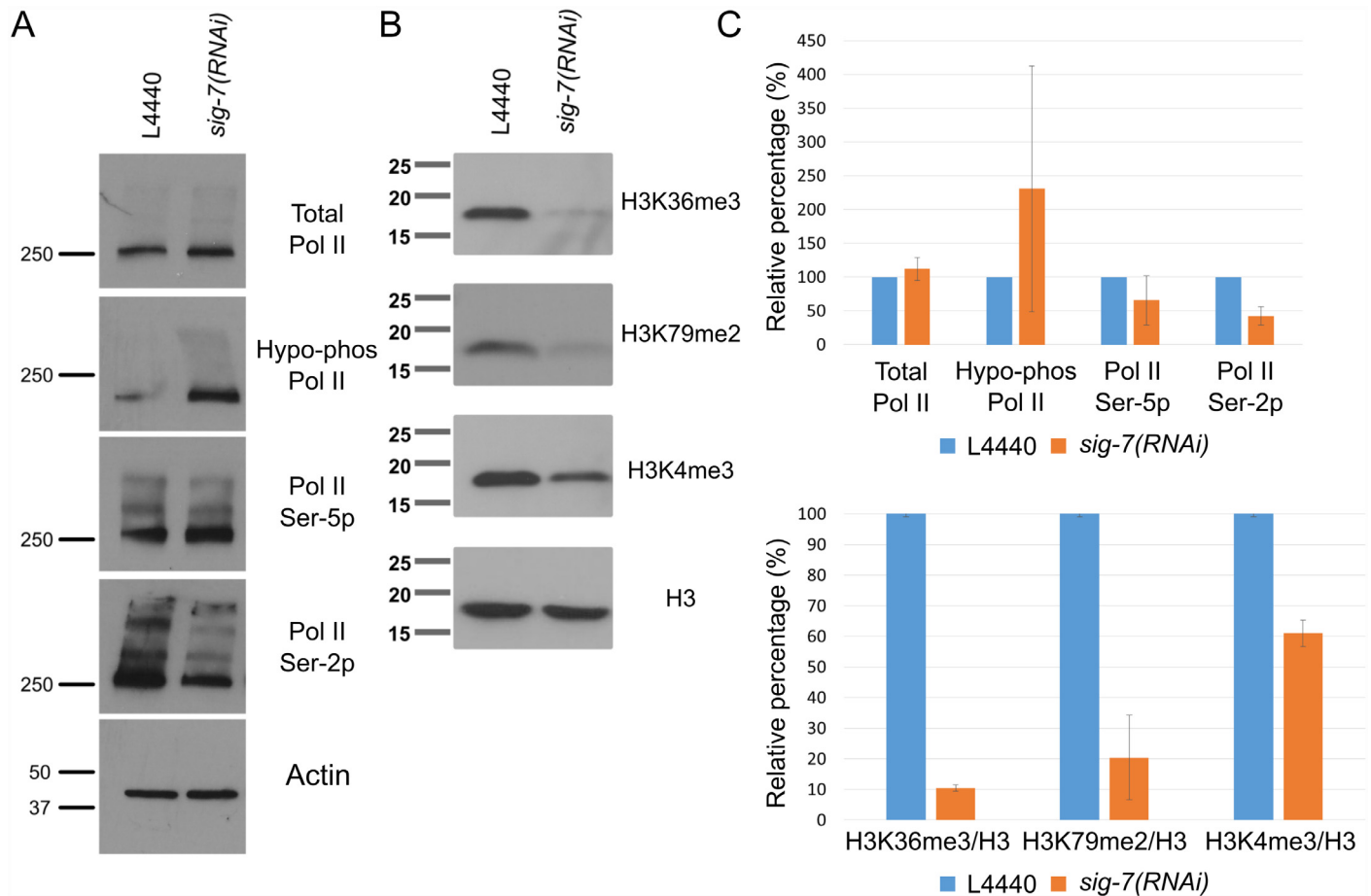
Fig 7. RNA Pol II phosphoepitopes and histone H3 modifications associated with transcription elongation are altered by *sig-7(RNAi)*. A) Equal amounts of lysates from L4440 control RNAi and *sig-7(RNAi)* embryos were analyzed by western blots probed with antibodies against all RNA Pol II isoforms (anti-AMA-1), hypo-phosphorylated Pol II (8WG16), Pol II Ser5-P (H14), and Pol II Ser2-P (H5). Anti-actin antibodies show equivalent protein loading. B) Lysates were analyzed by western blots as in (A) using antibodies against histone H3 modifications associated with transcription elongation, H3K36me3 and H3K79me2, and the 5' end of active genes, H3K4me3. Antibodies against total histone H3 (H3) show similar protein loading. C) Quantification of the probe signals from each RNA Pol II antibody relative to control L4440 in A (Top panel). Quantification of probe signals from each histone modification relative to control L4440 and normalized to total H3 for each sample in B (Bottom panel). Error bars = S.D. from two separate experiments.

doi:10.1371/journal.pgen.1006227.g007

that SIG-7 depletion has little effect on embryonic AMA-1 protein levels. The amount of hypo-phosphorylated Pol II (hypo-phos; 8WG16) was variable between experiments but often higher in *sig-7(RNAi)* embryos relative to controls. Pol II Ser-5P levels were also variable, but with a slight decrease often observed. In contrast, a significant decrease in levels of Pol II Ser-2P was consistently observed in *sig-7(RNAi)* embryos (Fig 7A and 7C). The decrease in Pol II Ser-2P is consistent with the decreased 3' Pol II profile observed by ChIP-seq, and indicates that the elongation phase of transcription is altered in *sig-7(RNAi)* embryos.

In yeast, the addition of Ser-5P to the CTD by TFIIH correlates with recruitment of the histone H3K4-specific methyltransferase Set1, which in turn leads to an enrichment of H3K4me3 in nucleosomes near the promoter [79–81]. Elongation and increased phosphorylation of Ser2 in turn correlates with recruitment of the H3K36 methyltransferase Set2 and a resulting enrichment of H3K36me3 within the body of the gene as elongation proceeds [13, 81, 82]. H3K79me2 is also added to gene body nucleosomes during Pol II elongation [83–85]. We examined *sig-7(RNAi)* and control embryos by western blot analysis, using antibodies specific
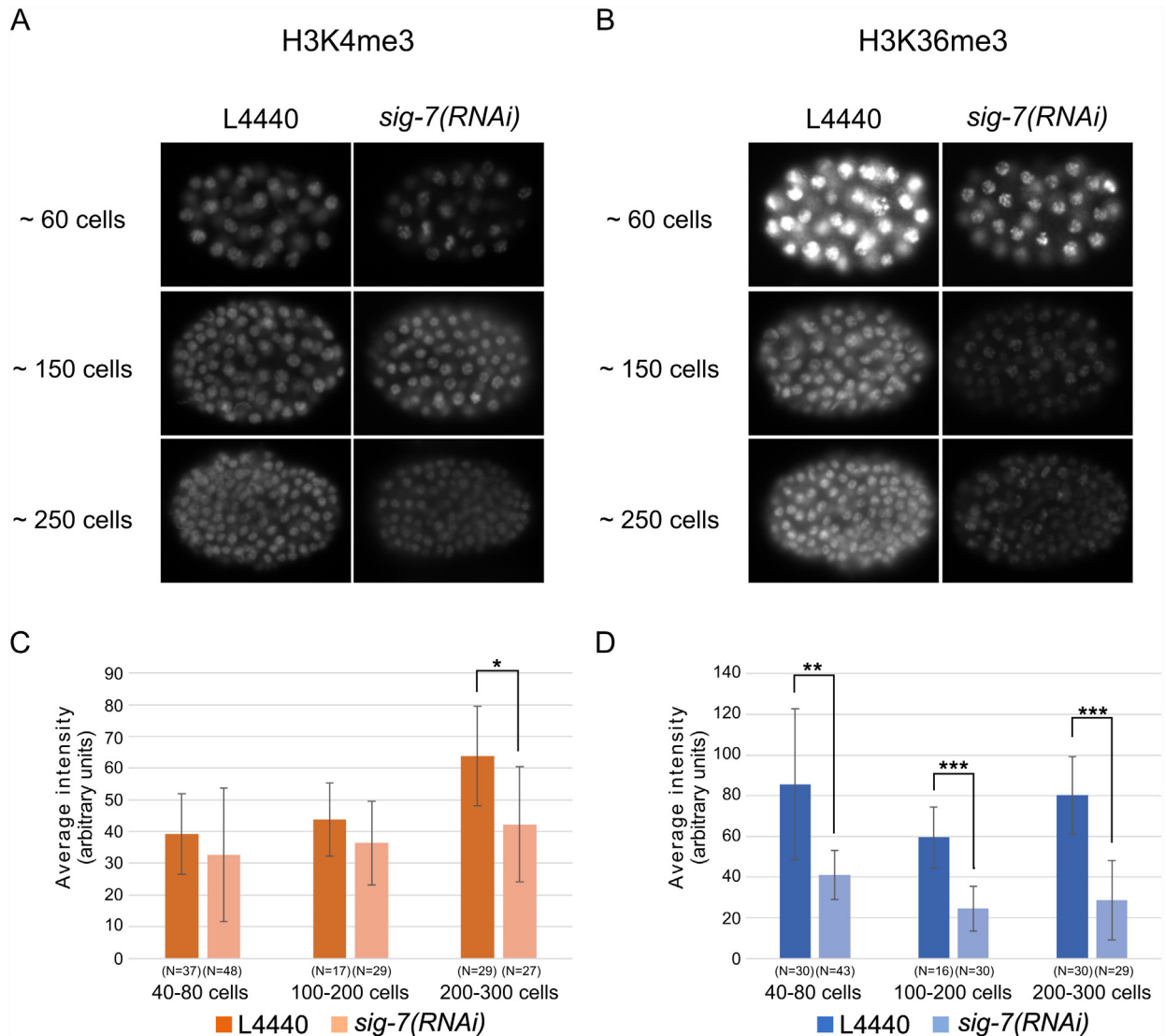
Fig 8. H3K36me3 is more significantly affected than H3K4me3 during embryonic development in *sig-7(RNAi)*. Anti-H3K4me3 (A) and anti-H3K36me3 (B) immunofluorescence patterns are shown for various embryonic stages. The significant difference is only observed in older embryonic stage between L4440 control and *sig-7(RNAi)*. Quantification of the average intensity of H3K4me3(C) and H3K36me3(D) staining from various stages are shown. N = total number of embryos quantified. Error bars = S.D. The statistical significance of the difference in average intensity is indicated by asterisks (* = p ≤ 0.01, ** = p ≤ 10^{-5}, *** = p ≤ 10^{-10}).

doi:10.1371/journal.pgen.1006227.g008

for H3K4me3, H3K36me3, or H3K79me2 and compared these to total histone H3 (Fig 7B and 7C). We observed a slight decrease in H3K4me3 and a substantial decrease in H3K36me3 and H3K79me2 in *sig-7(RNAi)* embryos. Thus, like elongation-dependent phosphorylation of Ser2 in the Pol II CTD, elongation-dependent histone modifications are also disproportionately affected by *sig-7(RNAi)*. We also looked at H3K4me3 and H3K36me3 levels in embryos by immunofluorescence. Consistent with our western blot result, we observed slight decreases in the level of H3K4me3, most notably in the ~250 cell stage embryos, while H3K36me3 levels were observed to be decreased in all stages after the ~60 cell stage (Fig 8). The H3K36me3 in the early embryo is predominantly provided by MES-4, a transcription-independent H3K36 methyltransferase, whereas transcription-dependent H3K36me3 predominates in later stages.

The minimal effect on H3K4me3, a promoter-proximal mark, and the dramatic reduction in H3K36me3, a mark enriched toward the 3'end of transcribed genes, is consistent with a role for SIG-7 in normal Pol II elongation.

In summary, depletion of SIG-7 from *C. elegans* embryos causes a developmental arrest, likely due to widespread defects in splicing accompanied by a global decrease in transcription of genes required for normal embryogenesis. This transcription defect correlates with a marked decrease in Pol II at the 3' end of genes and decreases in Pol II CTD phospho-epitopes and chromatin modifications that are hallmarks of elongating Pol II. SIG-7 physically associates with Pol II *in vivo* and is enriched in chromatin in patterns consistent with association with active transcription, and loss of SIG-7 causes defects in co-transcriptional splicing. SIG-7 is thus required for both transcription and splicing, and while it could directly impact just one process and indirectly the other, it's possible it coordinates both processes to promote accurate and efficient mRNA production.

Discussion

We report the first genome-wide analysis of a highly conserved, multi-domain nuclear cyclophilin, SIG-7, that is required for efficient transcription elongation and transcript splicing in *C. elegans*. Similar to the *S. pombe* and *A. thaliana* SIG-7 orthologs (Rct1 and AtCyp59, respectively), SIG-7 is an essential protein implicated in regulation of the phosphorylation status of important serines in the CTD of Pol II [41, 43]. Both a maternal supply and zygotic production of SIG-7 are required for normal development at all stages.

Depletion of SIG-7 by RNAi treatment results in a substantial decrease in embryonically produced transcripts in embryos. This decrease is accompanied by defective patterns of mRNA splicing, including co-transcriptional *trans*-splicing. A change in the distribution of Pol II within gene bodies is also observed that, along with reduced Pol II CTD Ser-2 phosphorylation and H3K36 and H3K79 methylation, are consistent with defects in Pol II elongation. The correlation between elongation defects and splicing defects could suggest an interdependency of these two processes in *C. elegans*, with SIG-7 providing an essential link. The extent to which mRNA processing and Pol II elongation are co-dependent in any organism is controversial, and indeed our results cannot rule out separable functions for SIG-7 in both processes [32, 73, 86–91]. This will likely remain controversial, because it is challenging to experimentally discriminate between a splicing defect directly causing an elongation defect versus an elongation defect causing a splicing defect. However, several reports have indicated that a primary defect in splicing can cause defective Pol II elongation. In cultured fibroblasts, depletion of a known splicing factor, SC35, results in attenuation of Pol II elongation through gene bodies in mammalian cells [29]. In HeLa cells, inhibition of splicing using spliceostatin (SSA) or antisense oligos targeting snRNAs resulted in defects remarkably similar to those caused by SIG-7 depletion: early dissociation of Pol II leading to its 3' depletion and decreases in Ser2P [92]. It thus seems likely that the primary defect in *sig-7* mutants is defective splicing, which leads to defective elongation and Pol II dissociation from genes. Metagene analysis of our RNA-seq data showed no increased reads past the annotated TES in RNAi versus control embryos (S12 Fig), indicating that in the reduced instances when Pol II completed elongation, termination was largely unaffected.

While the mechanistic roles of nuclear cyclophilins in any organism remain to be determined, yet our results and those from other studies provide important clues. The SIG-7-type nuclear cyclophilins all have a conserved RNA-recognition motif (RRM domain) in addition to the peptidyl-prolyl *cis-trans* isomerase (PPI) domain. Studies in *S. pombe* and *A. thaliana* demonstrate a role for Rct1 and AtCyp59, respectively, in the regulation of Pol II CTD

phosphorylation [41, 43], and *in vitro* binding experiments show that the CTD of Pol II interacts with the PPI domain of Rct1 [44]. Furthermore, it was shown that Rct1-dependent effects on Pol II CTD phosphorylation is dependent on the PPI domain, indicating that this motif is important for the association of Rct1 with Pol II and regulation of Pol II phosphorylation [44]. The other motifs in the SIG-7 orthologs are involved with RNA interactions. The RRMs of both Rct1 and AtCyp59 were shown to bind a motif present in ~70% of all mRNAs, and AtCyp59 has been shown to interact with pre-mRNAs, supporting a general role in co-transcriptional RNA processing *in vivo* [42]. Thus, a potential model for SIG-7 is that it binds to the CTD of Pol II through its PPI domain and employs its RRM domain to capture emerging RNAs, perhaps to efficiently recruit them to the spliceosome machinery attached to the CTD. In the absence of SIG-7, the coordinated interactions between emerging transcripts and the splicing machinery may be compromised, leading to decreased splicing efficiency and as a consequence, elongation may be disrupted through an as yet unknown mechanism.

An alternative model is that the PPI domain's catalytic function in isomerization of prolines may target the Pol II CTD repeat, which in turn may affect Pol II elongation via CTD structural alterations. PPI activity may be regulated by RNA binding and/or RNA processing, and this could provide a mechanistic link between elongation and splicing. Indeed, binding of RNA to AtCyp59 affects the isomerase activity of the PPI domain *in vitro* [42]. Future studies should investigate the importance of the PPI domain's catalytic activity and a requirement for it to be structurally linked to the RRM domain.

While SIG-7 homologs are found in most eukaryotes, there is no obvious homolog in budding yeast. Introns are relatively rare in budding yeast genes (present in just ~4% of protein-coding genes) and the few introns present are small in size [93]. Indeed, SIG-7/Rct1 is among a number of conserved spliceosome components and related proteins that are present in fission yeast, but have been lost from budding yeast [47]. SIG-7 function is thus dispensable in *S. cerevisiae*, as might be predicted for a protein responsible for coordinating transcription elongation with efficient splicing, as budding yeast have relatively few, small introns to process. The predominance of intron-less genes in budding yeast would presumably make maintaining a protein that was central to linking proper splicing with efficient elongation no longer essential.

Materials and Methods

Worm strains and maintenance

C. elegans strains were maintained at 20°C. Worms were grown on NGM (Nematode Growth Medium) plates unless stated otherwise.

Strains: wild type N2 (Bristol), KW1317: *sig-7*(*cc629*) I/*hT2*[*bli-4(e937) let-?(q782) qIs48*](I; III), KW2230: *sig-7*(*n5037*) I/*hT2* [*bli-4*(*e937*) *let-?*(*q782*) *qIs48*](I, III), PD7271: *pha-1(e2123ts)* III; ccEx(pBK48.1::pC1), KW2309: *sig-7(n5037)I; ckIs33(unc-119, sig-7:gfp:3XFLAG)II (below)*.

sig-7::*GFP*::*3XFLAG* transgenic strain

Repetitive sequences within intron 1 of the *sig-7* gene prevented PCR-based cloning of the whole gene. We used a fosmid clone (construct ID: 15087717651452437 A06) containing *sig-7* engineered with a 3' GFP::3XFLAG tag obtained from the TransgeneOme project [46]. The fosmid was cut with SphI to generate an 18KB fragment containing the entire operon with *sig-7* and two neighboring genes (CEOP1492). This fragment was blunt ended and inserted into the pCFJ151 MosSCI targeting vector cut with PvuII. This construct, pJA8, was integrated into an LGII MosSCi targeting site (*ttTi5605*) by standard Mos-SCI integration techniques [94].

RNAi-mediated depletion

RNAi was performed by feeding HT115 bacteria transformed with plasmids expressing dsRNA targeting the corresponding gene, or carrying the empty L4440 RNAi vector for controls.

RNAi embryos: Adult worms were collected from plates and washed with M9 buffer (22mM KH2PO4, 42mM Na2HPO4, 86mM NaCl, 1 mM MgSO4), bleached with sodium hypochlorite (5% bleach with 1.0N NaOH) to isolate embryos. Embryos were placed on NGM (Nematode Growth Medium) plates without food overnight. The synchronized hatched L1s were transferred to plates with OP50 bacteria and grown for 36 hours until the L3 larval stage. L3s were washed with M9 buffer 3 times and transferred to induced RNAi plates (NGM+1mM IPTG+1mM Ampicillin) pre-seeded with bacteria expressing the desired dsRNA. The worms were grown on RNAi plates for 36 hours, after which the gravid adults were washed with M9 buffer and separated from any extruded embryos by filtration through a 40μm mesh cell strainer (Fisher Scientific, #22363547). The adults were bleached as described above to collect *in utero* embryos for analysis.

RNAi Adults (S6B Fig): L3 larvae prepared as above were fed dsRNA-expressing bacteria for 55 hours instead of 36 hours and directly processed for total RNA purification and analysis by qRT-PCR.

Immunofluorescence

Intact embryos were fixed in 2.5%PFA/ethanol [95] or methanol/acetone [96] for all immuno-fluorescence except for those probed with monoclonal antibody H5, which was fixed in metha-nol/formaldehyde [95]. Primary antibodies used were: anti-H3K4me2 (CMA30) 1:1000 END Millipore], anti-P-granules [OIC1D4, 1:5 [96, 97]]), anti-H3K36me3 [(CMA333), 1:1000 [95]], anti-Ser2p RNA pol II CTD (H5, 1:500, Covance MMS-129R), anti-GFP (1:1000, Novus NB600-308), anti-AMA-1 (1:10,000, Novus 38520002), and anti-FLAG (M2, 1:1000, Sigma F1804). Secondary antibodies used were Alexa Fluor 488-conjugated donkey anti-mouse (1:500, Invitrogen R37114) and Alexa Fluor 594-conjugated goat anti-rabbit (1:500, Invitrogen R37117). Samples were mounted in ProLong Gold anti-fade reagent (Life technologies, P36934) and observed under a fluorescence microscope (Leica DMRXA; Hamamatsu Photon-ics, Hamamatsu, Japan) with Simple PCI software (Hamamatsu Photonics). Image J was used for quantification of raw immunofluorescence intensity [98].

Immunoprecipitation assays

Embryos were collected as described above, washed, and resuspended in 3X volume of ice-cold Hypotonc Triton-X buffer [20 mM Tris–HCl (pH 7.4), 10 mM KCl, 10 mM MgCl2, 2 mM EDTA, 10% glycerol, 1% Triton X-100, 2.5 mM β-glycerophosphate, 1 mM NaF, 1 mM DTT, and Complete protease inhibitors; [99]]. Resuspended embryos were frozen in liquid nitrogen and ground into a fine powder using a mortar and pestle and thawed on ice for 10 min. The suspension was sonicated for 2 min at high setting using a Bioruptor sonicator (Diagenode Inc., Denville, NJ, USA). The salt concentration was then adjusted to either 150mM or 350mM NaCl, and incubated for 30 min with rotation at 4°C. After an additional 2 min sonication, the lysate was centrifuged for 15 min at 13,000g. The supernatant was transferred to new tubes, and 1ml of each lysate supernatant was pre-cleared by incubation with 60 μl of either Protein A (Life Technologies, 10002D) or Protein G Dynabead (Life Technologies, 10004D) for 30 min with rotation at 4°C. 100μl of each pre-cleared lysate was saved as input sample, and the remaining 900ul was used for immunoprecipitation. Either anti-FLAG (Sigma, F1804) or anti-GFP (Novus biological, NB600-308) for SIG-7 IP and anti-AMA-1 (Novus biological, 38520002) for Pol II (AMA-1) IP were added to the lysate (10ug of antibody/2.5mg of lysates)

and incubated for 12 hours at 4°C. 60µl of either Protein A or Protein G Dynabeads were added directly to the lysate/antibody mix and incubated at 4°C for 3 hours. Beads were separated from solution using a magnetic bar, washed 2 times for 5 min in Hypotonic Triton-X buffer, and washed twice more with 500mM NaCl hypotonic Triton-X buffer for 10 min at room temperature. For final elution, beads were incubated with 150µl of 2X SDS-PAGE sample buffer for 15 min at room temperature. The final eluates were further analyzed by SDS-PAGE and Western blot.

Protein isolation and western blot analysis

RNAi-treated embryos were resuspended in 4X volume of RIPA buffer (Thermo Scientific, #89901) and 2X volume of glass beads (Sigma, G8772), and homogenized using a Mini Bead-beater-16 (BioSpec, Bartlesville, OK, USA) for 1 min 3 times at 4°C. The homogenized embryos were incubated at 4°C on a rotator for 30 min and further processed in a Bioruptor sonicator (Diagenode Inc., Denville, NJ, USA) at high setting for 10 min to fragment chromatin. The final lysates were centrifuged at 13,000g for 10 min at 4°C, and the supernatants were collected. The protein concentration was determined using the Bradford reagent (Biorad, #500–0006). Supernatants were mixed with an equal volume of 2X SDS-PAGE sample buffer and denatured for 5 min at 95°C, and equal protein amounts were loaded and run on a 4–20% precast SDS-PAGE gel (Biorad, #456–1094) and transferred to PVDF membrane. Transferred proteins were blocked in 5% milk PBST for 1 hour, incubated with primary antibody overnight, and washed 3 times with 1X PBST for 10 min each. After incubation with secondary antibody for 2 hours at room temperature, the blot was washed 3 times with 1X PBST for 10 min each. The washed membrane was incubated with chemiluminescence reagent (Thermo Scientific, #34087) for 5 min, and protein bands were visualized with autoradiography film (Genesee Scientific, #30–100). The primary antibodies used are the following: anti-FLAG (Sigma, 1:2000, F1804), anti-Actin (Millipore, 1:10,000, MAB1501), H14 (Covance, 1:3000, MMS-134R), H5 (Covance, 1:3000, MMS-129R), anti-AMA-1 (Novus Biological, 1:5000, 38520002), 8WG16 (Covance, 1:1000, MMS-126R), anti-H3K36me3 (Abcam, 1:1000, Ab9050), anti-H3K79me2 (Abcam, 1:1000, Ab3594), anti-H3K4me3 (Abcam, 1:1000, Ab8580) and anti-H3 (Abcam, 1:20,000, Ab1792). The following secondary antibodies were used: Goat Anti-Rabbit IgG, HRP-conjugated (Millipore, 1:2500, 12–348), Goat Anti-Mouse IgG, Peroxidase-conjugated (Millipore, 1:2500, AP124P), AffiniPure Goat Anti-Mouse IgM, Peroxidase-conjugated (Jackson ImmunoResearch Laboratories, Inc, 1:5000, 115-035-075).

RNA purification and qRT-PCR

Embryos collected after RNAi exposure were washed with M9, and pelleted embryos were resuspended in Trizol (50µl of embryos/300µl of Trizol, Invitrogen), snap frozen in liquid nitrogen, and subjected to 3 freeze/thaw cycles. 62µl of chloroform was added and mixed thoroughly by shaking 10 times and spun down for 15 min at 4°C. Nucleic acids were precipitated with 0.3M acetic acid in 100% isopropanol and resuspended in 100µl of nuclease-free water. Total RNA was purified using RNeasy kit (Qiagen, Valencia, CA, USA) as per the manufacturers' instructions. cDNA was synthesized from 1µg of purified total RNA using iScript select cDNA synthesis kit (Bio-Rad, 170–8896). 50ng of cDNA was used for qPCR using SsoFast reagent (Bio-Rad, 172–5201) on CFX96 Real-Time system (C1000 Thermal Cycler, Bio-Rad). The transcript levels of genes analyzed were first normalized to 18S RNA for each sample, and the normalized transcript levels of either *sig-7(RNAi)* or *ama-1(RNAi)* experiments were then compared to the transcript levels of L4440 controls to generate ΔΔCT plots of relative

transcript levels. The averages of two technical replicates from two biological samples were plotted with standard deviation.

Library preparation and RNA sequencing

Total RNAs were purified as described for qRT-PCR and sent to Axeq Asia (Seoul, Korea) for transcriptome sequencing. 1μg of total RNA was used as starting material, and a sequencing library was prepared using TruSeq Stranded Total RNA Sample Prep Kit after treatment with Ribo-Zero (Human/Mouse/Rat) for rDNA depletion. Library QC was performed using Tapestation D1000 Screen Tape (Agilent) and quantified using KAPA Library Quantification Kit (for Illumina platforms). Clusters were generated by HiSeq PE (Paired-End) Cluster Kit v3 cBot, and sequencing was done on a HiSeq2000 with 100bp paired-ends using TruSeq SBS v3-HS kit reagents.

Analysis of RNA-seq

RNA-seq reads were quality-checked using FastQC version 0.5.2 to ensure per-base sequence quality, per sequence quality scores, per base sequence content, per base GC content, per sequence GC content, per base N content, sequence length distribution, sequence duplication levels, kmer-content, and that over-represented sequences were within accepted norms. FastQ Quality Trimmer version 1.0 was used to trim reads with less than optimal quality scores. The DE analysis protocol outlined in Trapnell et. al. was used to perform the DE analysis [100]. The quality filtered reads were mapped to *C. elegans* (ce10) reference genome using TopHat2 version 0.5. TopHat2 internally uses Bowtie2 to map the reads. Mapping results were used to identify splice junctions between exons. Cufflinks version 2.1.1 was used to assemble transcripts and estimate their abundance. The transcript assembly outputs from Cufflinks were merged into a unified list of transcripts using Cuffmerge. Cuffdiff version 2.0.2 was then used to quantify gene and transcript expression levels and test them for significant differences. Default parameters were utilized for all steps. This analysis was done in part by the Emory Integrated Genomics Core (EIGC), which is subsidized by the Emory University School of Medicine and is one of the Emory Integrated Core Facilities. For Fig 7B, exon and intron coordinates were obtained from WS2220. 99,830 exons with length ≥ 50bp and 69,762 introns of ≥ 50bp were obtained. Custom scripts were used to calculate the average read coverage of exons and introns per gene. For S12 Fig, all RNA-seq samples were scaled to 10 million mapped reads. The averaged values from two biological replicates were plotted using the same pipeline employed for the metagene analysis of RNA Pol II CHIP-seq data, described below.

Library preparation from ChIP material for sequencing

Chromatin immunoprecipitation (ChIP) was done as described in Ercan *et al.* with the following modifications [101]. 1) Worms were grown on 15cm peptone rich plates seeded with NA22 bacteria. 2) Samples were sonicated using a Bioruptor sonicator at high setting for 40min (40sec on/ 20sec off). After collection of immunoprecipitated DNA, DNA libraries were prepared as described in [101, 102]. DNA libraries were sent to Axeq Asia (Seoul, Korea) for sequencing. Library QC was done using BioAnalyzer High sensitive DNA chip (Agilent). Clusters were generated by HiSeq PE (Paired-End) Cluster Kit v3 cBot, and sequencing was performed on HiSeq2000 with 100bp paired-ends using TruSeq SBS v3-HS kit reagents.

Analysis of ChIP-seq data

Gene set definitions used were as published in [54]. Briefly, ubiquitous genes are defined as genes expressed in 4 different tissue-specific SAGE data sets: germline, neuronal, muscle, and gut [103, 104]. The germline-enriched category was defined by [55], although spermatogenesis-specific genes were removed. Germline-specific genes were defined as expressed in the germline either in Reinke *et al.* or in Wang *et al.*, and were intersected with the strict maternal gene class in Baugh *et al.*; any genes expressed in any of the somatic SAGE expression data of Meissner 2009 were removed [55, 103–105]. Soma-specific (any) genes were defined as expressed in any of the somatic SAGE data sets of Meissner et al. but not in the germline SAGE data set of Wang *et al.* or the germline-enriched set of Reinke *et al.* [55, 103, 104]. Embryo-expressed is the "strict embryonic" class defined in Baugh 2004 [105]. chrX genes are all the X-linked genes in WS220. Silent genes are mostly serpentine receptors and were defined in [106, 107]. The hyper-geometric distribution was used to calculate the significance of the enrichment or depletion of any of the gene sets among the mis-regulated genes in Fig 4B. AMA-1 ChIP-seq data were mapped to WS220 using bowtie [108]. MACS2 was used to obtain peak calls for each replicate of L4440 and *sig-7(RNAi)* [109]. The broad peak option was found to produce the most appropriate peak calls and a significance cutoff of q = 0.05. The peak calls were mapped to WS220 gene annotations. If AMA-1 peak calls of both replicates of a condition overlapped with a gene body, the gene was called bound by AMA-1. Meta-gene profiles were produced using custom R scripts. Genes were aligned at their Transcription Start Site (TSS) and Transcription End Site (TES), and signal over the gene bodies was averaged in 50 bp windows. The 95% confidence interval of the mean is shown with error bars. To normalize reads between samples AMA-1 peak regions from both conditions were removed. The remaining read coverage was scaled genome-wide so the total number of reads was 2 million reads.

Protein sequence alignment

Protein sequences of SIG-7 homologues were obtained from the NCBI protein database. The alignment of homologues was generated with ClustalW2 [110]. The conserved protein domain/motif search was done using ScanProsite web-based tool [111, 112]. The accession numbers for proteins used for alignment are the following: SIG-7(CAB03088.2), PPIL-4 (NP_624311.1), CG5808(AAF56342.1), AtCyp59(NP_175776.2), Rct-1(CAB52803.1) and KIN241(CAC35733.1).

Supporting Information

S1 Fig. Pleiotropic defects in *sig-7* mutants. A,B) Homozygous *sig-7(cc629)* adults exhibit a highly penetrant protruding vulva (Pvul, arrow in A) and lower penetrance multiple vulvae (Muv, arrows B). C) Adult hermaphrodites are sterile due to failure to switch from spermatogenesis to oogenesis (masculinization of germline or Mog phenotype), as shown in a DAPI stained whole mount ovary. The dotted line and arrow shows the direction of meiotic progression. The excessive sperm accumulation is outlined. Other phenotypes observed in the *cc629* allele include molting and seam cell defects and enhancement of germline tumors (not shown). D) RNAi depletion of *sig-7* from hermaphrodites starting at the L3 stage resulted in ~95% embryonic lethality (examples of terminal, arrested embryos shown) among their progeny, with rare survivors exhibiting the defects described above. Scale bar = 20um.
(TIF)

S2 Fig. The *sig-7* gene, protein, mutant alleles, and orthologs in other species. A) The CEOP1492 operon (blue), the *sig-7* gene with exons (red boxes) and introns (black lines), the

position of the GFP::3XFLAG tag, and the genomic region used in the rescuing transgene construct (dotted line with arrowheads). B) The *n5037* deletion and *cc629* splice acceptor mutant alleles are illustrated. C) The domain structure of SIG-7 is illustrated; all three domains are also found in SIG-7 orthologs listed in (D). No homolog has been identified in budding yeast.
(TIF)

S3 Fig. Sequence alignment of SIG-7 orthologs. Protein sequences were aligned using ClustalW2 (110). The degree of conservation is noted with different symbols in the bottom row of each alignment. The cyclophilin-type peptidyl-prolyl cis-trans isomerase domain (PPI) and RNA Recognition Motif (RRM) were identified using the ScanProsite web-based tool and are indicated with black solid lines [111, 112]. This alignment shows the high degree of conservation among orthologs within the PPI and RRM domains.
(TIF)

S4 Fig. SIG-7 is a ubiquitously expressed nuclear protein. The *sig-7(n5037)* allele was rescued with a SIG-7::GFP::3XFLAG fusion protein. The GFP expression patterns in live animals are shown. A-D) SIG-7::GFP::3XFLAG is nuclear and observed in all tissues and at all stages, including early embryos (A), late embryos (B), adult somatic cells (C; arrows and arrowheads), and germ cells (C, D; outlined with dotted lines). Small speckles in A,C,D are auto-fluorescent gut granules. The protein is associated with DNA in all germ cells except in transcriptionally inactive diakinesis-stage oocytes, where it becomes dispersed in the nucleoplasm (insets in D). The protein is also associated with DNA in male germ cells at all stages, with faint nuclear signal observed in mature sperm (not shown).
(TIF)

S5 Fig. SIG-7 interacts with RNA Pol II *in vivo*. Lysates from embryos expressing SIG-7::GFP::3XFLAG were divided into equal aliquots and incubated with antibodies specific for RNA Pol II (α-AMA-1; Lanes 2 and 3), GFP (Lane 4), and the FLAG epitope (Lane 5), and the immunoprecipitates were analyzed by western blots probed with anti-Pol II Ser-5P (H14; top), anti-FLAG (anti-SIG-7; middle) and anti-Actin (bottom). The lysate in Lane 2 was incubated with 100ug of RNase A for 30min at RT prior to the IP procedure.
(TIF)

S6 Fig. SIG-7 is required for germline transcription. The abundance of germline-specific transcripts was measured by qRT-PCR in RNAi embryos (from adult animals fed dsRNA 36hrs post-L3, A) and adults (fed 55 hrs post-L3; see Materials and Method, B). RNA levels were normalized to 18S RNA levels and plotted relative to L4440 control in each experiment. Error bars = S.D. from two technical replicates each of two biological replicates.
(TIFF)

S7 Fig. Comparison of embryonic stages present in control vs *sig-7(RNAi)*. Embryos collected from RNAi experiments were fixed and stained with DAPI. The approximate numbers of nuclei in the embryos in a field were binned as indicated. (N = total number of embryos quantified)
(TIF)

S8 Fig. *sig-7(RNAi)* predominantly affects zygotic gene expression. Of the 1522 down-regulated genes and 389 upregulated in *sig-7(RNAi)* embryos (at least two fold expression difference with q value < 0.05), 607 and 205 genes, respectively, were among those classified by Baugh et al. as Strictly maternal, Strictly Embryonic, or Maternal/Embryonic [105]. The percentages of up- or down-regulated genes that fall into these gene classes are indicated.
(TIF)

S9 Fig. SIG-7 is required for efficient splicing of nascent transcripts. A) Examples of zygotically expressed genes with splicing defects. The number of aligned reads generated by Tophat are indicated on the left. The y-axis of the *sig-7(RNAi)* reads, which are reduced compared to L4440 controls, is scaled to the exon reads in L4440. An "outron" is the RNA segment removed by trans-splicing of nascent transcript; its sequence corresponds to that between the TSS (Transcription Start Site) and SL1 (Splice Leader 1 acceptor site) of trans-spliced genes. Exons (blue boxes) and introns (black solid lines) are shown under each RNA-seq profile. The relative levels of introns and outrons (indicated with asterisks) compared to exons are significantly higher in *sig-7(RNAi)* compared to L4440 control, reflecting persistence of primary transcripts. B) Average log2 read coverage per gene for exons and introns in *sig-7(RNAi)* vs L4440 is shown. Genes up- and downregulated in *sig-7(RNAi)* compared to L4440 by cuffdiff analysis are shown in red and green, respectively. Exon levels change in the manner expected for mis-regulated genes, while intron levels remain relatively unchanged.
(TIF)

S10 Fig. *sig-7(RNAi)*-dependent changes in RNA Pol II occupancy correlate with expression changes observed by RNA-seq. Genome-wide RNA Pol II (anti-AMA-1 antibody) ChIP-seq was performed in L4440 control RNAi and *sig-7(RNAi)* embryos, and the read density profiles were compared with the RNA-seq results from the experiments described in Fig 4. The comparison shows an excellent correlation between the loss of RNA Pol II from genes with a decrease in RNA abundance, indicating that the changes observed with *sig-7(RNAi)* are associated with a transcription defect.
(TIF)

S11 Fig. *sig-7(RNAi)*-dependent changes in RNA Pol II occupancy among different gene classes are consistent with defects observed by RNA-seq. Genes with RNA Pol II occupancy changes in *sig-7(RNAi)* embryos relative to control L4440 embryos were classified and compared as in Fig 4. A) Genes expressed in somatic lineages, embryo-expressed genes, and X-linked genes are overrepresented among genes showing decreased RNA Pol II occupancy in *sig-7(RNAi)* embryos. B) Genes expressed in the germline show an increase in RNA Pol II occupancy in *sig-7(RNAi)* embryos.
(TIF)

S12 Fig. Metagene analyses of RNA-seq reads from *sig-7(RNAi)* embryos and L4440 control embryos. Metagene displays of RNA-seq reads, plotted from 1kb upstream and downstream of the annotated Transcription Start Sites (TSS) and Transcription End Sites (TES; actually PolyA addition site), for different classes of genes in L4440 RNAi control (blue) and *sig-7(RNAi)* (red). A) Genes were categorized as in Figs 4 and 6. B) Graphs showing scaled reads for all genes with 2-fold lower or higher read counts. All RNA-seq samples were scaled to 10 million mapped reads, and the graphs illustrate combined results from 2 biological replicates. Error bars indicate the 95% confidence interval of the mean signal, indicated by the circles. Reads were normalized as indicated in Materials and Methods. Genes expressed in the germline showed a large 3' of the annotated TES. This is likely from RNAs corresponding to downstream genes in operons, which predominate for germline expressed genes.
(TIF)

S1 Table. RNA-seq analysis of conserved splicing factors. The log2 fold change after *sig-7 (RNAi)* in RNA reads corresponding to *C. elegans* homologs of known splicing factors is

shown.
(TIFF)

S2 Table. ChIP-seq and RNA-seq analyses data.
(XLS)

Acknowledgments

Author Contributions

Conceived and designed the experiments: JHA WGK.

Performed the experiments: JHA.

Wrote the paper: JHA WGK SS AR.

Performed the RNA-seq analysis: JHA AR. Performed the ChIP-seq data analyses: AR. Helped guide the data analyses: SS.

References

1. Voss TC, Hager GL. Dynamic regulation of transcriptional states by chromatin and transcription factors. Nature reviews Genetics. 2014; 15(2):69–81. doi: 10.1038/nrg3623 PMID: 24342920.

2. Heidemann M, Hintermair C, Voss K, Eick D. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. Biochimica et biophysica acta. 2013; 1829(1):55–62. doi: 10.1016/j.bbagrm.2012.08.013 PMID: 22982363.

3. Liu X, Bushnell DA, Kornberg RD. RNA polymerase II transcription: structure and mechanism. Biochimica et biophysica acta. 2013; 1829(1):2–8. doi: 10.1016/j.bbagrm.2012.09.003 PMID: 23000482; PubMed Central PMCID: PMC4244541.

4. Holoch D, Moazed D. RNA-mediated epigenetic regulation of gene expression. Nature Reviews Genetics. 2015; 16:71–84. Epub 2 January 2015. doi: 10.1038/nrg3863 PMID: 25554358

5. Teves SS, Weber CM, Henikoff S. Transcribing through the nucleosome. Trends in biochemical sciences. 2014; 39(12):577–86. doi: 10.1016/j.tibs.2014.10.004 PMID: 25455758.

6. Kulaeva OI, Hsieh FK, Chang HW, Luse DS, Studitsky VM. Mechanism of transcription through a nucleosome by RNA polymerase II. Biochimica et biophysica acta. 2013; 1829(1):76–83. doi: 10.1016/j.bbagrm.2012.08.015 PMID: 22982194; PubMed Central PMCID: PMC3535581.

7. Medvedeva YA, Khamis AM, Kulakovskiy IV, Ba-Alawi W, Bhuyan MSI, Kawaji H, et al. Effects of cytosine methylation on transcription factor binding sites. BMC Genomics. 2013; 15(119):1–12. doi: 10.1186/1471-2164-15-119

8. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nature genetics. 2007; 39(1):61–9. doi: 10.1038/ng1929 PMID: 17128275

9. Murakami K, Elmlund H, Kalisman N, Bushnell DA, Adams CM, Azubel M, et al. Architecture of an RNA polymerase II transcription pre-initiation complex. Science. 2013; 342(6159):1238724. doi: 10.1126/science.1238724 PMID: 24072820; PubMed Central PMCID: PMC4039082.

10. Cheung AC, Sainsbury S, Cramer P. Structural basis of initial RNA polymerase II transcription. The EMBO journal. 2011; 30(23):4755–63. doi: 10.1038/emboj.2011.396 PMID: 22056778; PubMed Central PMCID: PMC3243610.

11. Fuda NJ, Ardehali MB, Lis JT. Defining mechanisms that regulate RNA polymerase II transcription in vivo. Nature. 2009; 461(7261):186–92. doi: 10.1038/nature08449 PMID: 19741698; PubMed Central PMCID: PMC2833331.

12. Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nature reviews Genetics. 2012; 13(10):720–31. doi: 10.1038/nrg3293 PMID: 22986266; PubMed Central PMCID: PMC3552498.

13. Selth LA, Sigurdsson S, Svejstrup JQ. Transcript Elongation by RNA Polymerase II. Annual review of biochemistry. 2010; 79:271–93. doi: 10.1146/annurev.biochem.78.062807.091425 PMID: 20367031.

14. Kohoutek J. P-TEFb- the final frontier. Cell division. 2009; 4:19. doi: 10.1186/1747-1028-4-19 PMID: 19723344; PubMed Central PMCID: PMC2748068.

15. Kwak H, Lis JT. Control of transcriptional elongation. Annual review of genetics. 2013; 47:483–508. doi: 10.1146/annurev-genet-110711-155440 PMID: 24050178; PubMed Central PMCID: PMC3974797.

16. Sikorski TW, Buratowski S. The basal initiation machinery: beyond the general transcription factors. Current opinion in cell biology. 2009; 21(3):344–51. doi: 10.1016/j.ceb.2009.03.006 PMID: 19411170

17. Bataille AR, Jeronimo C, Jacques PE, Laramee L, Fortin ME, Forest A, et al. A universal RNA poly-merase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isom-erase enzymes along genes. Molecular cell. 2012; 45(2):158–70. doi: 10.1016/j.molcel.2011.11.024 PMID: 22284676.

18. Hirose Y, Ohkuma Y. Phosphorylation of the C-terminal Domain of RNA Polymerase II Plays Central Roles in the Intergrated Events of Eukaryotic Gene Expression. J Biochem. 2007; 141:601–8. doi: 10.1093/jb/mvm090 PMID: 17405796

19. Mayer A, Heidemann M, Lidschreiber M, Schreieck A, Sun M, Hintermair C, et al. CTD tyrosine phos-phorylation impairs termination factor recruitment to RNA polymerase II. Science. 2012; 336 (6089):1723–5. doi: 10.1126/science.1219651 PMID: 22745433.

20. Glover-Cutter K, Larochelle S, Erickson B, Zhang C, Shokat K, Fisher RP, et al. TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. Molecular and cellular biology. 2009; 29(20):5455–64. doi: 10.1128/MCB.00637-09 PMID: 19667075; PubMed Central PMCID: PMC2756882.

21. Kim M, Suh H, Cho EJ, Buratowski S. Phosphorylation of the yeast Rpb1 C-terminal domain at serines 2, 5, and 7. The Journal of biological chemistry. 2009; 284(39):26421–6. doi: 10.1074/jbc.M109.028993 PMID: 19679665; PubMed Central PMCID: PMC2785330.

22. Lin PS, Tremeau-Bravard A, Dahmus ME. The repetitive C-terminal domain of RNA polymerase II: multiple conformational states drive the transcription cycle. Chemical record. 2003; 3(4):235–45. doi: 10.1002/tcr.10063 PMID: 14595832.

23. Kim JB, Sharp PA. Positive Transcription Elongation Factor b Phosphorylates hSPT5 and RNA Poly-merase II Carboxyl-terminal Domain Independently of Cyclin-dependent Kinase-activating Kinase. The Journal of biological chemistry. 2001; 276(15):12317–23. doi: 10.1074/jbc.M010908200 PMID: 11145967

24. Liu Y, Warfield L, Zhang C, Luo J, Allen J, Lang WH, et al. Phosphorylation of the transcription elonga-tion factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. Molecular and cellu-lar biology. 2009; 29(17):4852–63. doi: 10.1128/MCB.00609-09 PMID: 19581288; PubMed Central PMCID: PMC2725703.

25. Ping Y-H, Rana TM. DSIF and NELF Interact with RNA Polymerase II Elongation Complex and HIV-1 Tat Stimulates P-TEFb-mediated Phosphorylation of RNA Polymerase II and DSIF during Transcrip-tion Elongation. The Journal of biological chemistry. 2001; 276(16):12951–8. doi: 10.1074/jbc.M006130200 PMID: 11112772

26. Mortillaro MJ, Blencowe BJ, Wei X, Nakayasu H, Du L, Warren SL, et al. A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. Proceedings of the National Academy of Sciences. 1996; 93(16):8253–7.

27. Yuryev A, Patturajan M, Litingtung Y, Joshi RV, Gentile C, Gebara M, et al. The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. Proceedings of the National Academy of Sciences. 1996; 93(14):6975–80.

28. Gu B, Eick D, Bensaude O. CTD serine-2 plays a critical role in splicing and termination factor recruit-ment to RNA polymerase II in vivo. Nucleic acids research. 2012. doi: 10.1093/nar/gks1327

29. Lin S, Coutinho-Mansfield G, Wang D, Pandit S, Fu XD. The splicing factor SC35 has an active role in transcriptional elongation. Nat Struct Mol Biol. 2008; 15(8):819–26. doi: 10.1038/nsmb.1461 PMID: 18641664; PubMed Central PMCID: PMC2574591.

30. Koga M, Satoh T, Takasaki I, Kawamura Y, Yoshida M, Kaida D. U2 snRNP Is Required for Expres-sion of the 3′ End of Genes. PloS one. 2014; 9(5):e98015. doi: 10.1371/journal.pone.0098015 PMID: 24845214

31. Brès V, Gomes N, Pickle L, Jones KA. A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. Genes & development. 2005; 19(10):1211–26. doi: 10.1101/gad.1291705

32. Fong YW, Zhou Q. Stimulatory effect of splicing factors on transcriptional elongation. Nature. 2001; 414(6866):929–33. PMID: 11780068

33. Krishnamurthy S, Ghazy MA, Moore C, Hampsey M. Functional interaction of the Ess1 prolyl isomerase with components of the RNA polymerase II initiation and termination machineries. Molecular and cellular biology. 2009; 29(11):2925–34. doi: 10.1128/MCB.01655-08 PMID: 19332564; PubMed Central PMCID: PMC2682014.

34. Ma Z, Atencio D, Barnes C, DeFiglio H, Hanes SD. Multiple roles for the Ess1 prolyl isomerase in the RNA polymerase II transcription cycle. Molecular and cellular biology. 2012; 32(17):3594–607. doi: 10.1128/MCB.00672-12 PMID: 22778132; PubMed Central PMCID: PMC3422015.

35. Kops O, Zhou XX, Lu KP. Pin1 modulates the dephosphorylation of the RNA polymerase II C-terminal domain by yeast Fcp1. FEBS letters. 2002; 513:305–11. Epub Jan 23 2001. PMID: 11904169

36. Liou YC, Zhou XZ, Lu KP. Prolyl isomerase Pin1 as a molecular switch to determine the fate of phosphoproteins. Trends in biochemical sciences. 2011; 36(10):501–14. doi: 10.1016/j.tibs.2011.07.001 PMID: 21852138; PubMed Central PMCID: PMC3185210.

37. Werner-Allen JW, Lee CJ, Liu P, Nicely NI, Wang S, Greenleaf AL, et al. cis-Proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72. The Journal of biological chemistry. 2011; 286(7):5717–26. doi: 10.1074/jbc.M110.197129 PMID: 21159777; PubMed Central PMCID: PMC3037684.

38. Kubicek K, Cerna H, Holub P, Pasulka J, Hrossova D, Loehr F, et al. Serine phosphorylation and proline isomerization in RNAP II CTD control recruitment of Nrd1. Genes & development. 2012; 26 (17):1891–6. doi: 10.1101/gad.192781.112 PMID: 22892239; PubMed Central PMCID: PMC3435493.

39. Hanes SD. Prolyl isomerases in gene transcription. Biochimica et biophysica acta. 2014. doi: 10.1016/j.bbagen.2014.10.028 PMID: 25450176.

40. Krzywicka A, Beisson J, Jerka-Dziadosz M, Klotz C. KIN241: a gene involved in cell morphogenesis in Paramecium tetraurelia reveals a novel protein family of cyclophilin–RNA interacting proteins (CRIPs) conserved from fission yeast to man. Molecular microbiology. 2001; 42(1):257–67. doi: 10.1046/j.1365-2958.2001.02634.x PMID: 11679083

41. Gullerova M, Barta A, Lorkovic ZJ. AtCyp59 is multidomain cyclophilin from *Arabidopsis thaliana* that interacts with SR proteins and the C-terminal domain of the RNA polymerase II. Rna. 2006; 12 (4):631–43. PMID: 16497658

42. Bannikova O, Zywicki M, Marquez Y, Skrahina T, Kalyna M, Barta A. Identification of RNA targets for the nuclear multidomain cyclophilin atCyp59 and their effect on PPlase activity. Nucleic acids research. 2013; 41(3):1783–96. Epub Dec 16 2012. doi: 10.1093/nar/gks1252 PMID: 23248006

43. Gullerova M, Barta A, Lorkovic ZJ. Rct1, a Nuclear RNA Recognition Motif-Containing Cyclophilin, Regulates Phosphorylation of the RNA Polymerase II C-Terminal Domain. Molecular and cellular biology. 2007; 27(10):3601–11. doi: 10.1128/MCB.02187-06 PMID: 17339332

44. Skrahina T. The role of Schizosaccharomyces pombe cyclophilin Rct1 in RNA polymerase II transcription. PhD Dissertation, uniwien. 2009.

45. Aroian RV, Levy AD, Koga M, Ohshima Y, Kramer JM, Sternberg PW. Splicing in *Caenorhabditis elegans* Does Not Require an AG at the 3' Splice Acceptor Site. Molecular and cellular biology. 1993; 13 (1):626–37. PMID: 8417357

46. Sarov M, Murray JI, Schanze K, Pozniakovski A, Niu W, Angermann K, et al. A genome-scale resource for in vivo tag-based protein function exploration in C. elegans. Cell. 2012; 150(4):855–66. doi: 10.1016/j.cell.2012.08.001 PMID: 22901814; PubMed Central PMCID: PMC3979301.

47. Aravind L, Watanabe H, Lipman DJ, Koonin EV. Lineage-specific loss and divergence of functionally linked genes in eukaryotes. Proceedings of the National Academy of Sciences of the United States of America. 2000; 97(21):11319–24. doi: 10.1073/pnas.200346997 PMID: 11016957; PubMed Central PMCID: PMC17198.

48. Kelly WG, Schaner CE, Dernburg AF, Lee M-H, Kim SK, Villeneuve AM, et al. X-chromosome silencing in the germline of C. elegans. Development. 2002; 129(2):479–92. PMID: 11807039

49. Nance J, Lee J-Y, Goldstein B. Gastrulation in *C. elegans*. Wrombook. 2005:1–13. doi: 10.1895/wormbook.1.23.1

50. Powell-Coffman JA, Knight J, Wood WB. Onset of C. elegans Gastrulation Is Blocked by Inhibition of Embryonic Transcription with an RNA polymerase Antisense RNA. Developmental biology. 1996; 178:472–83. PMID: 8812143

51. Baugh LR. Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. Development. 2003; 130(5):889–900. doi: 10.1242/dev.00302 PMID: 12538516

52. Broitman-Maduro G, Owraghi M, Hung WWK, Kuntz S, Sternberg PW, Maduro MF. The NK-2 class homeodomain factor CEH-51 and the T-box factor TBX-35 have overlapping function in *C. elegans* mesoderm development. Development. 2009; 136(16):2735–46. doi: 10.1242/dev.038307 PMID: 19605496

53. Robertson SM, Shetty P, Lin R. Identification of lineage-specific zygotic transcripts in early *Caenorhabditis elegans* embryos. Developmental biology. 2004; 276(2):493–507. doi: 10.1016/j.ydbio.2004.09.015 PMID: 15581881.

54. Rechtsteiner A, Ercan S, Takasaki T, Phippen TM, Egelhofer TA, Wang W, et al. The Histone H3K36 Methyltransferase MES-4 Acts Epigenetically to Transmit the Memory of Germline Gene Expression to Progeny. PLoS genetics. 2010; 6(9):e1001091. doi: 10.1371/journal.pgen.1001091 PMID: 20824077

55. Reinke V, Gil IS, Ward S, Kazmer K. Genome-wide germline-enriched and sex-biased expression profiles in Caenorhabditis elegans. Development. 2004; 131(2):311–23. doi: 10.1242/dev.00914 PMID: 14668411

56. Bashirullah A, Cooperstock RL, Lipshitz HD. Spatial and Temporal Control of RNA Stability. PNAS. 2001; 98(13):7025–8. doi: 10.1073/pnas.111145698 PMID: 11416182

57. Bitar M, Boroni M, Macedo AM, Machado CR, Franco GR. The Spliced Leader Trans-Splicing Mechanism in Different Organisms: Molecular Details and Possible Biological Roles. Frontiers in Genetics. 2013; 4. doi: 10.3389/fgene.2013.00199

58. Blumenthal T, Thomas J. *Cis* and *trans* mRNA splicing in *C. elegans*. Trends in Genetics. 1988; 4 (11):305–8. doi: 10.1016/0168-9525(88)90107-2 PMID: 3070853

59. Thomas J, Lea K, Zucker-Aperison E, Blumenthal T. The spliceosomal snRNAs of *Caenorhabditis elegans*. Nucleic acids research. 1990; 18(9):2633–42. PubMed Central PMCID: PMCPMC330746. PMID: 2339054

60. Allen MA, Hillier LW, Waterston RH, Blumenthal T. A global analysis of *C. elegans trans*-splicing. Genome research. 2011; 21:255–64. doi: 10.1101/gr.113811.110. PubMed Central PMCID: PMCPMC3032929. PMID: 21177958

61. Chen RA, Down TA, Stempor P, Chen QB, Egelhofer TA, Hillier LW, et al. The landscape of RNA polymerase II transcription initiation in C. elegans reveals promoter and enhancer architectures. Genome research. 2013; 23(8):1339–47. doi: 10.1101/gr.153668.112 PMID: 23550086; PubMed Central PMCID: PMC3730107.

62. Saito TL, Hashimoto S, Gu SG, Morton JJ, Stadler M, Blumenthal T, et al. The transcription start site landscape of C. elegans. Genome research. 2013; 23(8):1348–61. doi: 10.1101/gr.151571.112 PMID: 23636945; PubMed Central PMCID: PMC3730108.

63. Kruesi WS, Core LJ, Waters CT, Lis JT, Meyer BJ. Condensin controls recruitment of RNA polymerase II to achieve nematode X-chromosome dosage compensation. Proudfoot N, editor2013 2013-06-18 18: 46:53.

64. Blumenthal T. Trans-splicing and operons in *C. elegans*. WormBook: the online review of C elegans biology. 2012:1–11. doi: 10.1895/wormbook.1.5.2

65. Morton JJ, Blumenthal T. RNA processing in C. elegans. Methods in cell biology. 2011; 106:187–217. doi: 10.1016/B978-0-12-544172-8.00007-4 PMID: 22118278.

66. Zahler AM. Pre-mRNA splicing and its regulation in *Caenorhabditis elegans*. WormBook: the online review of C elegans biology. 2012. doi: 10.1895/wormbook.1.31.2

67. Khodor YL, Rodriguez J, Abruzzi KC, Tang C-HA, Marr MT, Rosbash M. Nascent-seq indicates widespread cotranscriptional pre-mRNA splicing in Drosophila. Genes & development. 2011; 25 (23):2502–12. doi: 10.1101/gad.178962.111

68. Singh J, Padgett RA. Rates of in situ transcription and splicing in large human genes. Nature structural & molecular biology. 2009; 16(11):1128–33. http://www.nature.com/nsmb/journal/v16/n11/suppinfo/nsmb.1666_S1.html.

69. Pandya-Jones A, Black DL. Co-transcriptional splicing of constitutive and alternative exons. Rna. 2009; 15(10):1896–908. doi: 10.1261/rna.1714509 PMID: 19656867

70. Das R, Yu J, Zhang Z, Gygi MP, Krainer AR, Gygi SP, et al. SR Proteins Function in Coupling RNAP II Transcription to Pre-mRNA Splicing. Molecular cell. 2007; 26(6):867–81. doi: 10.1016/j.molcel.2007.05.036 PMID: 17588520

71. Tilgner H, Knowles DG, Johnson R, Davis CA, Chakrabortty S, Djebali S, et al. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. Genome research. 2012; 22(9):1616–25. doi: 10.1101/gr.134445.111 PMID: 22955974

72. Das R, Dufu K, Romney B, Feldt M, Elenko M, Reed R. Functional coupling of RNAP II transcription to spliceosome assembly. Genes & development. 2006; 20:1100–9. doi: 10.1101/gad.1397406

73. Lin S, Coutinho-Mansfield G, Wang D, Pandit S, Fu X-D. The splicing factor SC35 has an active role in transcriptional elongation. Nature structural & molecular biology. 2008; 15(8):819–26. http://www.nature.com/nsmb/journal/v15/n8/suppinfo/nsmb.1461_S1.html.

74. Custodio N, Carvalho C, Condado I, Antoniou M, Blencome BJ, Carmo-Fonseca M. In vivo recruitment of exon junction complex proteins to transcription sites in mammalian cell nuclei. Rna. 2004; 10 (4):622–33. doi: 10.1261/rna.5258504 PMID: 15037772

75. Bowman EA, Bowman CR, Ahn JH, Kelly WG. Phosphorylation of RNA polymerase II is independent of P-TEFb in the *C. elegans* germline. Development. 2013; 140(17):3703–13. doi: 10.1242/dev.095778 PMID: 23903194

76. Bowman EA, Kelly WG. RNA polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation. Nucleus. 2014; 5(3):1–13. doi: 10.4161/nucl.29347

77. Egloff S, Dienstbier M, Murphy S. Updating the RNA polymerase CTD code: adding gene-specific layers. Trends in genetics: TIG. 2012; 28(7):333–41. doi: 10.1016/j.tig.2012.03.007 PMID: 22622228.

78. Mayer A, Lidschreiber M, Siebert M, Leike K, Soding J, Cramer P. Uniform transitions of the general RNA polymerase II transcription complex. Nature structural & molecular biology. 2010; 17(10):1272–8. doi: 10.1038/nsmb.1903 PMID: 20818391.

79. Hampsey M, Reinberg D. Tails of Intrigue: Phosphorylation of RNA Polymerase II Mediates Histone Methylation. Cell Cycle. 2003; 113:429–32.

80. Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NCT, et al. Active genes are tri-methylated at K4 of histone H3. Nature. 2002; 419(6905):407–11. http://www.nature.com/nature/journal/v419/n6905/suppinfo/nature01080_S1.html. PMID: 12353038

81. Ng HH, Robert F, Young RA, Struhl K. Targeted Recruitment of Set1 Histone Methylase by Elongating Pol II Provides a Localized Mark and Memory of Recent Transcriptional Activity. Molecular cell. 2003; 11:709–19. PMID: 12667453

82. Zhou Q, Li T, Price DH. RNA polymerase II elongation control. Annual review of biochemistry. 2012; 81:119–43. doi: 10.1146/annurev-biochem-052610-095910 PMID: 22404626; PubMed Central PMCID: PMC4273853.

83. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. Nature genetics. 2008; 40(7):897–903. doi: 10.1038/ng.154. PMC2769248. PMID: 18552846

84. Steger DJ, Lefterova MI, Ying L, Stonestrom AJ, Schupp M, Zhuo D, et al. DOT1L/KMT4 Recruitment and H3K79 Methylation Are Ubiquitously Coupled with Gene Transcription in Mammalian Cells. Molecular and cellular biology. 2008; 28(8):2825–39. doi: 10.1128/MCB.02076-07. PMC2293113. PMID: 18285465

85. Crisucci EM, Arndt KM. The Roles of the Paf1 Complex and Associated Histone Modifications in Regulating Gene Expression. Genetics research international. 2011; 2011:15. doi: 10.4061/2011/707641

86. de Almeida SF, Grosso AR, Koch F, Fenouil R, Carvalho S, Andrade J, et al. Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. Nature structural & molecular biology. 2011; 18(9):977–83. http://www.nature.com/nsmb/journal/v18/n9/abs/nsmb.2123.html—supplementary-information.

87. Moehle EA, Braberg H, Krogan NJ, Guthrie C. Adventures in time and space. RNA Biology. 2014; 11 (4):313–9. doi: 10.4161/rna.28646 PMID: 24717535

88. Brzyżek G, Świeżewski S. Mutual interdependence of splicing and transcription elongation. Transcription. 2015; 6(2):37–9. doi: 10.1080/21541264.2015.1040146 PMID: 25890997

89. Khodor YL, Menet JS, Tolan M, Rosbash M. Cotranscriptional splicing efficiency differs dramatically between Drosophila and mouse. Rna. 2012; 18(12):2174–86. doi: 10.1261/rna.034090.112 PMID: 23097425

90. Brody Y, Neufeld N, Bieberstein N, Causse SZ, Bohnlein EM, Neugebauer KM, et al. The in vivo kinetics of RNA polymerase II elongation during co-transcriptional splicing. PLoS biology. 2011; 9(1): e1000573. doi: 10.1371/journal.pbio.1000573 PMID: 21264352; PubMed Central PMCID: PMC3019111.

91. Tardiff DF, Lacadie SA, Rosbash M. A Genome-Wide Analysis Indicates that Yeast Pre-mRNA Splicing Is Predominantly Posttranscriptional. Molecular cell. 2006; 24(6):917–29. doi: 10.1016/j.molcel.2006.12.002 PMID: 17189193

92. Koga M, Hayashi M, Kaida D. Splicing inhibition decreases phosphorylation level of Ser2 in Pol II CTD. Nucleic acids research. 2015; 43(17):8258–67. doi: 10.1093/nar/gkv740 PMID: 26202968

93. Spingola M, Grate L, Haussler D, Ares MJ. Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. Rna. 1999; 5:221–34. PubMed Central PMCID: PMCPMC1369754. PMID: 10024174

94. Frokjaer-Jensen C, Wayne Davis M, Hopkins CE, Newman BJ, Thummel JM, Olesen S-P, et al. Single-copy insertion of transgenes in Caenorhabditis elegans. Nature genetics. 2008; 40(11):1375–83. http://www.nature.com/ng/journal/v40/n11/suppinfo/ng.248_S1.html. doi: 10.1038/ng.248 PMID: 18953339

95. Furuhashi H, Takasaki T, Rechtsteiner A, Li T, Kimura H, Checchi PM, et al. Transgenerational epigenetic regulation of *C. elegans* primordial germ cells. Epigenetics & Chromatin. 2010; 3(15):1–21. doi: 10.1186/1756-8935-3-15

96. Strome S, Wood WB. Generation of asymmetry and segregation of germ-line granules in early C. elegans embryos. Cell. 1983; 35(1):15–25. doi: 10.1016/0092-8674(83)90203-9 PMID: 6684994

97. Li T, Kelly WG. A Role for Set1/MLL-Related Components in Epigenetic Regulation of the *Caenorhabditis elegans* Germ Line. PLoS genetics. 2011; 7(3):e1001349. doi: 10.1371/journal.pgen.1001349 PMID: 21455483

98. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Meth. 2012; 9(7):671–5.

99. Klenova E, Chernukhin I, Inoue T, Shamsuddin S, Norton J. Immunoprecipation techniques for the analysis of transcription factor complexes. Methods. 2002; 26(3):254–9. http://dx.doi.org/10.1016/S1046-2023(02)00029-4. PMID: 12054881

100. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols. 2012; 7(3):562–78. doi: 10.1038/nprot.2012.016. PMC3334321. PMID: 22383036

101. Ercan S, Whittle C, Lieb J. Chromatin immunoprecipitation from *C. elegans* embryos. Protocol Exchange. 2007.

102. Bowman S, Simon M, Deaton A, Tolstorukov M, Borowsky M, Kingston R. Multiplexed Illumina sequencing libraries from picogram quantities of DNA. BMC Genomics. 2013; 14(1):466. doi: 10.1186/1471-2164-14-466

103. Meissner B, Warner A, Wong K, Dube N, Lorch A, McKay SJ, et al. An Integrated Strategy to Study Muscle Development and Myofilament Structure in *Caenorhabditis elegans*. PLoS genetics. 2009; 5(6):e1000537. doi: 10.1371/journal.pgen.1000537 PMID: 19557190

104. Wang X, Zhao Y, Wong K, Ehlers P, Kohara Y, Jones SJ, et al. Identification of genes expressed in the hermaphrodite germ line of C. elegans using SAGE. BMC Genomics. 2009; 10:213-. doi: 10.1186/1471-2164-10-213. PMC2686737. PMID: 19426519

105. Baugh LR, Hill AA, Slonim DK, Brown EL, Hunter CP. Composition and dynamics of the Caenorhabditis elegans early embryonic transcriptome. Development. 2003; 130(5):889–900. doi: 10.1242/dev.00302 PMID: 12538516

106. Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J. Differential chromatin marking of introns and expressed exons by H3K36me3. Nature genetics. 2009; 41(3):376–81. doi: 10.1038/ng.322. PMC2648722. PMID: 19182803

107. Kolasinska-Zwierz P, Down T, Latorre I, Liu T, Liu XS, Ahringer J. Differential chromatin marking of introns and expressed exons by H3K36me3. Nature genetics. 2009; 41(3):376–81. doi: 10.1038/ng.322 PMID: 19182803; PubMed Central PMCID: PMC2648722.

108. Langmead B, Trapnell C, Pop M, Salzberg S. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology. 2009; 10(3):R25. doi: 10.1186/gb-2009-10-3-r25 PMID: 19261174

109. Zhang Y, Liu T, Meyer C, Eeckhoute J, Johnson D, Bernstein B, et al. Model-based Analysis of ChIP-Seq (MACS). Genome biology. 2008; 9(9):R137. doi: 10.1186/gb-2008-9-9-r137 PMID: 18798982

110. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23(21):2947–8. doi: 10.1093/bioinformatics/btm404 PMID: 17846036

111. de Castro E, Sigrist CJ, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic acids research. 2006; 34(Web Server issue):W362–5. doi: 10.1093/nar/gkl124 PMID: 16845026; PubMed Central PMCID: PMC1538847.

112. Sigrist CJ, de Castro E, Cerutti L, Cuche BA, Hulo N, Bridge A, et al. New and continuing developments at PROSITE. Nucleic acids research. 2013; 41(Database issue):D344–7. doi: 10.1093/nar/gks1067 PMID: 23161676; PubMed Central PMCID: PMC3531220.