DOT1L suppresses nuclear RNAi originating from enhancer elements in *Caenorhabditis elegans*

Ruben Esse,[1] Ekaterina Gushchanskaia,[1] Avery Lord,[1] and Alla Grishok[1,2]

[1]Department of Biochemistry, Boston University School of Medicine, Boston, MA 02118, USA

[2]Correspondence: agrishok@bu.edu (A.G.)

Abstract

Methylation of histone H3 lysine 79 by DOT1L is associated with actively transcribed genes. DOT-1.1, the *Caenorhabditis elegans* DOT1L homologue, cooperates with the chromatin-binding protein ZFP-1 (AF10 homologue) to negatively modulate transcription of highly and widely expressed target genes. Also, *zfp-1* deficiency is linked to reduced RNA interference response. In addition, in MLL-rearranged leukemia, H3K79 methylation at homeobox genes antagonizes repressive chromatin marks. Here, we connect these various manifestations of DOT1L function. We demonstrate that DOT1L suppresses transcription at enhancers, thereby preventing double-stranded RNA formation and ectopic histone H3 lysine 9 methylation. Moreover, we show that lethality of *dot-1.1* mutant worms is suppressed by mutations in the RNAi pathway, suggesting ectopic activation of endogenous RNAi in the absence of DOT1L. Since DOT1L is conserved and similarly prevents H3K9 methylation in *C. elegans* and mammals, our observations point to a possible link between nuclear RNAi and H3K9 methylation in higher organisms.

Introduction

Screens for chromatin-binding factors involved in RNAi in *C. elegans* have identified the zinc finger protein ZFP-1 as a putative mediator of dsRNA-induced silencing in the nucleus (Grishok et al. 2005; Kim et al. 2005; Dudley et al. 2002). Interestingly, its mammalian homologue, AF10, is a frequent fusion partner of MLL in chimeric proteins causing leukemia (Meyer et al. 2018), and its interaction with H3K79 methyltransferase DOT1L is critical for DOT1L recruitment to the *HOXA9* and *MEIS1* genes and their subsequent hyperactivation (Okada et al. 2005). In the attempt to identify proteins interacting with ZFP-1 that could explain its role in RNAi, we have previously identified a *C. elegans* homologue of DOT1L, which we named DOT-1.1, as the major ZFP-1 binding partner (Cecere et al. 2013).

Although DOT1L is the only H3K79 methyltransferase in mammals and H3K79 methylation is present on all actively transcribed genes, inhibition of DOT1L methyltransferase activity does not result in dramatic changes in gene expression in cultured cells (Zhu et al. 2018). However, expression of specific genes, such as *HOXA9* and *MEIS1*, is strongly dependent on DOT1L, especially in leukemias induced by MLL-fusion proteins (Okada et al. 2005). The importance of H3K79 methylation in antagonizing heterochromatin formation at the *HOXA* cluster has recently been established (Chen et al. 2015), resembling pivotal studies in yeast showing that Dot1 prevents the spreading of factors associated with heterochromatin into active regions (Katan-Khaykovich and Struhl 2005; Ng et al. 2002). Importantly, knockout of DOT1L in mice results in embryonic lethality, underscoring its important developmental function. In *Drosophila*, both DOT1L (*Grappa*) and AF10 (*Alhambra*) mutants showed

developmental defects (Mohan et al. 2010; Shanower et al. 2005). In *C. elegans*, we observed specific developmental abnormalities in a *zfp-1* loss-of-function mutant, such as defects in neuronal migration (Kennedy and Grishok 2014), and implicated ZFP-1 in lifespan control (Mansisidor et al. 2011). Whether or not the anti-silencing effect of DOT1L contributes to its developmental role is presently not known.

We previously established that, in *C. elegans*, ZFP-1 and DOT-1.1 co-localize to promoters of highly and widely expressed genes and negatively modulate their expression during development (Cecere et al. 2013). This mechanism is likely conserved and applicable to the majority of ubiquitously expressed genes in other systems. Here, we demonstrate the role of the ZFP-1/DOT-1.1 complex in enhancer regulation. Known epigenetic signatures typical of active enhancers include histone H3 lysine 27 acetylation (H3K27ac) and histone H3 lysine 4 monomethylation (H3K4me1), whereas high level of H3K27me3 and absence of H3K27ac are found on enhancers poised for later tissue-specific expression (Bonn et al. 2012). Enhancer elements are themselves transcription units, generating non-coding enhancer RNAs (eRNAs) (Li et al. 2016), and, most recently, regulation of enhancer transcription was shown to be very similar to that of protein-coding genes (Henriques et al. 2018). Indeed, we show that DOT-1.1 negatively modulates enhancer RNA transcription, similarly to the negative modulation of active genes described earlier (Cecere et al. 2013). Moreover, we find an antagonistic relationship between H3K79 methylation and H3K9 methylation at enhancers, a mechanism consistent with studies in leukemic cells (Chen et al. 2015). Finally, we demonstrate that suppression of anti-sense transcription by ZFP-1/DOT-1.1 prevents dsRNA formation and induction of nuclear RNAi via the Dicer/RDE-4/RDE-1

pathway, which likely promotes H3K9 methylation at enhancers in the absence of DOT-1.1. The elevation of endogenous dsRNA level is most likely responsible for the apparent deficiency of *zfp-1* mutant worms in responding to exogenous dsRNA. Based on these compelling observations, we propose that negative modulation of bidirectional transcription at enhancers by DOT1L does not allow dsRNA levels to reach the threshold required for inducing heterochromatin formation.

Results

ZFP-1 and DOT-1.1 are enriched at enhancers

Recently, the *C. elegans* genome was organized into 20 domains of different structure and activity based on chromatin modification signatures (Evans et al. 2016). We intersected coordinates of such domains with ZFP-1 ChIP-seq (third larval stage, L3) and DOT-1.1 ChIP-chip (embryo) peaks to interrogate in which genomic regions the ZFP-1/DOT-1.1 complex is most dominant. More than 40% of promoter regions (domains 1 and 8 in embryos and domain 1 at L3) are enriched in ZFP-1/DOT-1.1 (Fig. 1A), in line with our previous observations of promoter-proximal binding of the complex (Cecere et al. 2013). More interestingly, similar levels of enrichment are observed for domains corresponding to predicted intronic (domain 8 at L3) and intergenic enhancers (domain 9 in both embryos and L3 animals) (Fig. 1A,C). These domains have chromatin signatures characteristic of enhancers, such as H3K4me1 and H3K27ac. In addition, both promoters and enhancers are often devoid of nucleosomes and hence amenable to be identified by techniques assessing open chromatin regions. A recent study has identified such regions in *C. elegans* embryos and L3 animals by ATAC-seq. In this

study, approximately 5000 distal non-coding regions displayed dynamic changes in chromatin accessibility between the two developmental stages. Also, several of these putative enhancers revealed enhancer activity in transgenic reporter assays (Daugherty et al. 2017). Interestingly, overlap analysis shows that both intragenic and distal ATAC-seq peaks are enriched in DOT-1.1 and ZFP-1 in embryos and L3 animals, respectively (Fig. 1B,C). These observations raise the exciting possibility that, in addition to the previously disclosed role of ZFP-1/DOT-1.1 in the negative modulation of transcription via promoter-proximal binding, this complex may also control transcription through enhancers.

The ZFP-1/DOT-1.1 complex regulates distal and intragenic enhancer transcription

In light of the observation that enhancer elements in *C. elegans* are enriched in ZFP-1/DOT-1.1, we have hypothesized that this complex could control enhancer transcription. Distal enhancers often produce non-coding transcripts (eRNAs) which are typically unstable and not represented in the steady-state transcript population (Li et al. 2016). This technical limitation may be circumvented by using techniques designed to assess nascent transcription, such as the global run-on sequencing (GRO-seq) method (Hah et al. 2013; Lam et al. 2013). GRO-seq data previously obtained in our lab were instrumental in providing insight into the modulatory role of ZFP-1/DOT-1.1 in transcription of genes targeted by the complex at promoter-proximal regions (Cecere et al. 2013). We re-analyzed our published data to gain a comprehensive understanding of the effect of ZFP-1/DOT1L on the regulatory genomic regions and chromatin domains revealed by recent studies (Evans et al. 2016; Daugherty et al. 2017). Nascent transcription is not limited to gene-directed transcription and is observed also at non-

coding regions of the genome, including enhancers (Supplemental Fig. S1A,B). We quantified GRO-seq coverage at genomic domains characterized by unique chromatin signatures and functional annotation (Evans et al. 2016) in wild-type (WT) and *zfp-1* mutant (*zfp-1(ok554)*) L3 animals. In this mutant, ZFP-1 is C-terminally truncated and does not interact with DOT-1.1, leading to decreased abundance of DOT-1.1 at chromatin (Cecere et al. 2013). Interestingly, both intragenic and distal eRNAs were globally increased in *zfp-1* mutant animals in comparison with WT larvae (Fig. 1D and Supplemental Fig. S1B). Analysis of GRO-seq data at distal ATAC-seq peaks also revealed a global increase in transcription *zfp-1* mutant animals compared with WT larvae (Fig. 1E). These observations indicate that, in addition to negatively modulating target gene transcription via promoter-proximal binding, the ZFP-1/DOT-1.1 complex also regulates production of RNAs derived from both intragenic and distal enhancers.

Genes with enhancer signature bound by ZFP-1/DOT-1.1 display bidirectional transcription upon loss of ZFP-1

Intragenic enhancer transcription has been suggested to interfere with host gene transcription (Cinghu et al. 2017). In light of the elevation of intragenic enhancer transcription observed in the *zfp-1* mutant, we have analyzed cumulative GRO-seq coverage along gene bodies. Genes bound by ZFP-1/DOT-1.1 at the promoters were not considered, since we have previously observed that these show increased GRO-seq coverage in the *zfp-1* mutant compared with WT (Cecere et al. 2013). The remaining genes in the assembly were considered to be either bound or not bound by ZFP-1 at the coding regions. Both sense transcription and anti-sense transcription were analyzed. Interestingly, transcriptional changes in the *zfp-1* mutant compared with WT

at target genes depended on the presence of enhancer chromatin signature. Genes targeted by ZFP-1 and lacking enhancer signature displayed an overall decrease in sense transcription (Fig. 2A, top panel) and no significant change in anti-sense transcription (Fig. 2A, middle panel). However, those with enhancer signature showed a significant increase in anti-sense transcription (Fig. 2A, middle panel, and Fig. 2B), whereas sense transcription did not change significantly (Fig. 2A, top panel, and Fig. 2B).

We then compared changes in sense and anti-sense transcription assessed by GRO-seq with mRNA expression assessed by microarray and, interestingly, genes targeted by ZFP-1 at the coding region did not display an overall change in mRNA levels between *zfp-1* mutant worms and WT animals (Fig. 2A, bottom panel). These observations suggest a role for the ZFP-1/DOT-1.1 complex in preventing anti-sense transcription at target genes with intragenic enhancer signatures.

Pervasive bidirectional transcription may result in double-stranded RNA (dsRNA) formation from complementary base pairing of sense and anti-sense transcripts (Piatek et al. 2016). We hypothesized that the increase in anti-sense transcription observed in the *zfp-1* mutant could result in formation of dsRNA. To address this hypothesis, we looked directly at the amount of dsRNA in *zfp-1* mutant worms by immunostaining with a dsRNA-specific antibody previously employed to visualize dsRNA *in vivo* (Kaneko et al. 2011; Saldi et al. 2014). In agreement with our hypothesis, we have observed an accumulation of nuclear dsRNA foci in *zfp-1* mutant worms, when compared with WT (Fig. 2C).

Lethality of dot-1.1 knockout is suppressed by Dicer/RDE-4/RDE-1 pathway mutations

Our attempts to generate *dot-1.1* knockout mutants in the WT background by CRISPR/Cas9 technology were not successful. Motivated by the implication of DOT1L in apoptotic cell death (Nguyen et al. 2011; Feng et al. 2010), we have resorted to the *ced-3(n1286)* mutant strain, which is defective in the pro-apoptotic caspase CED-3 (homologue of mammalian caspase-3) (Xue et al. 1996), and have successfully generated a viable *dot-1.1(knu339)* null mutant in this mutant background (Fig. 3A). Upon outcrossing the *ced-3(n1286)* allele, we found that a few surviving adult *dot-1.1* null animals produced only developmentally arrested embryos and larvae. The arrested larvae displayed numerous vacuoles typically seen in worms undergoing necrotic cell death (Xu et al. 2001). Notably, the most pronounced effect was seen in the intestinal cells that do not normally undergo programmed cell death (Fig. 3B). Therefore, DOT-1.1 does not simply suppress the initiation of programmed cell death in specific cells that are destined to die.

We then hypothesized that the *dot-1.1(knu339)* lethality phenotype could be related to the increase in dsRNA observed in *zfp-1* mutant worms. In *C. elegans* and other organisms, exogenous dsRNA is subject to processing by the Dicer complex, which cleaves it into small interfering RNAs (siRNAs) (Grishok 2005). In *C. elegans*, the first discovered *R*NAi-*de*ficient mutants included *rde-1* and *rde-4* (Tabara et al. 1999). The product of the *rde-1* gene is the Argonaute protein RDE-1, which binds primary siRNAs produced by Dicer cleavage (Yigit et al. 2006), and the RDE-4 protein, which binds long dsRNA and assists Dicer in the processing reaction (Tabara et al. 2002; Parker 2006). The RDE-4 and RDE-1 proteins are also important for the anti-viral response (Wilkins et al. 2005) and thought to have a limited role in endogenous RNAi. To determine if the

RDE-4/RDE-1 pathway responds to the ectopic dsRNA accumulation seen in the *zfp-1* and *dot-1.1* mutants, we crossed the double *dot-1.1; ced-3* mutant strain with loss-of-function mutants for *rde-4* and *rde-1*. Excitingly, the double *dot-1.1(knu339); rde-4(ne301)* and *dot-1.1(knu339); rde-1(ne219)* mutant worms were homozygous viable and did not require the presence of the *ced-3* mutant allele for the viability. In a control experiment, we crossed the double *dot-1.1; ced-3* mutant strain with a mutant unrelated to the Dicer/RDE-4/RDE-1 pathway and the lethality of *dot-1.1(knu339)* was not rescued (Fig. 3C). These clear genetic results indicate that the lethality phenotype associated with *dot-1.1* deletion is due to generation of ectopic small RNAs from excessive nuclear dsRNA.

Excessive dsRNA resulting from loss of ZFP-1/DOT-1.1 competes for the Dicer complex

Genetic analyzes of RNAi in *C. elegans* initially identified genes regulating silencing in response to exogenous dsRNA (exo-dsRNA), some of which were later found to participate in endogenous RNAi as well (Grishok 2005; Lee 2006). Importantly, other mutants in the endogenous RNAi ERI (Enhancer RNAi) pathway were found in screens selecting for the increased response to exo-dsRNA (Simmer et al. 2002; Kennedy et al. 2004; Duchaine et al. 2006; Yigit et al. 2006). In the ERI pathway, the RNAi-initiating dsRNA is generated by the RNA-dependent RNA polymerase RRF-3 at the select endogenous loci, mostly representing duplicated genes (Gent et al. 2010; Fischer et al. 2013). This dsRNA is subject to Dicer processing and results in the generation of primary siRNAs bound by the ERGO-1 Argonaute (Vasale et al. 2010; Yigit et al. 2006). The ERI-1 protein and RRF-3 both interact with Dicer and act upstream of ERGO-1 in the ERI pathway. Thus, the ERI/ERGO-1 and the exo-RNAi/RDE-1 pathways compete

for the Dicer complex. We hypothesized that, upon loss of DOT-1.1 or ZFP-1, another pool of endogenous dsRNA (in addition to the RRF-3-produced) is competing with exo-RNAi, and efficiency of the latter is reduced as a result of this competition. To test this hypothesis, we used an *elt-2::*GFP*/lacZ* transgene-based (Fukushige et al. 1999) model. The *elt-2::*GFP*/lacZ* transgenic strain contains a repetitive transgenic array prone to silencing through the Dicer/RDE-4/RDE-1 pathway (Grishok et al. 2005). This transgene is exclusively expressed in the nuclei of intestinal cells and shows partial silencing at 16°C in the WT background and complete silencing in ERI pathway mutants (Grishok et al. 2005), such as *eri-1(mg366)* (Fig. 4A). On the contrary, in *zfp-1(ok554)* (not shown) or *dot-1.1(gk520244)* partial loss-of-function mutant backgrounds, it is readily apparent that transgenic expression is increased compared to WT (Fig. 4A). Importantly, the level of GFP expression in the double mutant *dot-1.1; eri-1* remains low, similar to that of *eri-1* alone (Fig. 4A). This result indicates that the effect of *dot-1.1* mutation (i.e. transgene de-silencing) is suppressed by the loss of ERI-1 working together with Dicer. In other words: DOT-1.1 acts upstream of ERI-1/Dicer. This result is not consistent with the long-standing belief that ZFP-1 induces RNAi in the nucleus in response to siRNAs produced by Dicer. However, it completely supports our hypothesis about the mechanism of the apparent RNAi-promoting function of ZFP-1 and DOT-1.1. In summary, our results demonstrate that, upon loss of ZFP-1/DOT-1.1, in addition to the dsRNA generated through the ERI pathway and the dsRNA produced from transgenes, the dsRNA produced from endogenous loci enters in the competition for the Dicer complex (Fig. 4B and Fig. 6). Our data strongly suggest that this competition occurs in the nucleus.

Loss of DOT-1.1 results in ectopic heterochromatin at distal and intragenic enhancers

DOT1L activity has been shown to restrict the spreading of factors associated with heterochromatin in yeast and leukemia cell lines (Ng et al. 2002; Katan-Khaykovich and Struhl 2005; Chen et al. 2015). On the other hand, siRNAs direct Argonaute proteins to complementary nascent transcripts to initiate H3K9 methylation, resulting in heterochromatin formation (Ni et al. 2014; Grewal and Elgin 2007). We reasoned that these two distinct phenomena could be put together in the context of the regulation of enhancer transcription by ZFP-1/DOT-1.1. Consistently, histone H3 lysine 9 dimethylation (H3K9me2), a hallmark of heterochromatin, is increased in both intragenic and distal enhancer domains in *dot-1.1; ced-3* animals compared to *ced-3* animals (Fig. 5A and Supplemental S2A). In addition, genes containing enhancer signatures have increased H3K9me2 in *dot-1.1; ced-3* animals (Fig. 5B and Supplemental S2B).

Overall, our results demonstrate that DOT1L activity prevents both dsRNA production and heterochromatin formation at enhancer elements in *C. elegans*.

Discussion

In the past few years, genomic and transcriptomic data from high-throughput sequencing studies have been shedding light on the importance of non-coding genomic regulatory regions, including enhancers. These are typically demarcated by certain characteristics, such as prevalence of the H3K4me1 and H3K27ac marks and open chromatin configuration, and these features have been extensively used to predict their coordinates in a variety of organisms. Importantly, exciting new observations point to the remarkable similarity between enhancers and promoters in terms of chromatin

architecture and transcriptional profile, challenging traditional binary views of enhancers and promoters (Rennie et al. 2017; Henriques et al. 2018; Mikhaylichenko et al. 2018). As these notions become dimmer, an interesting view gaining momentum is that some enhancers can act as weak promoters, while, conversely, bidirectional promoters often act as strong enhancers. This plasticity of enhancer and promoter activities is manifested in the levels and directionality of transcription (Mikhaylichenko et al. 2018). In the face of this shifting paradigm, it is reasonable to believe that the commonalities between enhancers and promoters may extend to how their transcriptional output is controlled by the epigenetic machinery of the cell, including molecular writers such as histone methyltransferases and histone acetyltransferases. We had previously demonstrated that ZFP-1/DOT-1.1 exert a negative modulatory role on transcription of genes with promoter-proximal binding of the complex (Cecere et al. 2013). These are mostly highly and widely expressed genes not subject to spatiotemporal regulation. Here, we present tantalizing evidence that the DOT1L complex also modulates enhancer-directed transcription in a mechanism that likely partakes on the regulation of genes subject to tissue and time-specific control of expression, ensuring the fidelity of developmental programs. Of note, DOT1L has been implicated in processes as diverse as embryonic and postnatal hematopoiesis, proliferation of mouse embryonic stem cells, induced and natural reprogramming, cardiac development and chondrogenesis (McLean et al. 2014), and the regulation of these processes is greatly contributed by enhancer activity. Therefore, it is possible that DOT1L assists promoter-enhancer cooperativity at key developmental transitions. Also, perturbation of enhancer activity is increasingly being recognized as an important player in malignant transformation (Sur

and Taipale 2016). Recently, chimeric oncoproteins MLL-AF9 and MLL-AF4 were found to bind specific subsets of non-overlapping active distal enhancers in acute myeloid leukemia cell lines. Interestingly, in this study, MLL-AF9-bound enhancers displayed higher levels of H3K79me2 than enhancers bound by the MLL protein alone (Prange et al. 2017). This observation can be explained by the fact that AF9, similarly to AF10, is a frequent binding partner of DOT1L and drives its localization to chromatin (Kuntimaddi et al. 2015). Therefore, it is possible that, in MLL-driven leukemias, inappropriate recruitment of DOT1L to distal regulatory elements perturbs their transcriptional output. In addition, homeobox loci with an indisputable role in malignant transformation, such as *HOXA9* and *MEIS1*, display both coding and non-coding transcription, including anti-sense transcripts (Popovic et al. 2008; Sessa et al. 2006), and DOT1L activity may play a yet unappreciated role in the control of such transcripts. In fact, our observations suggest that DOT1L controls anti-sense transcription, particularly at genes characterized by presence of enhancer signature. Interestingly, a sequence conserved in vertebrate *Hox* gene introns was reported to function as an enhancer element in *Drosophila* (Keegan et al. 1997; Haerry and Gehring 1996). In addition, the c-*myc* (Nepveu and Marcu 1986) and N-*myc* (Krystal et al. 1990) oncogenic loci display anti-sense transcription, and exciting new findings point to the implication of DOT1L in cancers underlain by c-Myc and N-Myc activation, particularly breast cancer (Cho et al. 2015) and neuroblastoma (Wong et al. 2017). Also, the contribution of anti-sense transcription to malignant transformation has been gaining attention in the past few years (Wenric et al. 2017; Balbin et al. 2015). Therefore, the enticing implications of our

observations on the *C. elegans* DOT1L homologue call for the role of anti-sense and non-coding transcription in cancer to be appreciated in the context of DOT1L activity. Our analysis of GRO-seq data revealed an increase in anti-sense transcription upon loss of ZFP-1/DOT-1.1, particularly at genes with enhancer signatures, implying overlapping bidirectional transcription and formation of sense-antisense dsRNA structures. This inference was reinforced by the clear observation of dsRNA foci in the nucleus of *zfp-1* mutant worms, indicating accumulation of dsRNA. Formation of dsRNA is a common feature of multiple gene suppression phenomena known collectively as RNAi, which may be endogenous or exogenous. Importantly, production of both endogenous (endo-siRNAs) and exogenous (exo-siRNAs) small interfering RNAs requires the activity of DCR-1/Dicer. Our observations demonstrate that endogenous dsRNA resulting from loss of ZFP-1/DOT-1.1 competes with exogenously introduced dsRNA for DCR-1/RDE-4/RDE-1-mediated processing into small RNAs. This competition can explain the implication of ZFP-1 in RNAi (Grishok et al. 2005; Kim et al. 2005; Dudley et al. 2002). Also, lethality of *dot-1.1* deletion is suppressed by Dicer/RDE-4/RDE-1 pathway mutations, showing that endogenous small RNA generation is involved in the lethality phenotype. It is well known that exo-dsRNA induces apoptosis in a variety of mammalian systems (Estornes et al. 2012; Yamane et al. 2006; Kalai et al. 2002). Our observations uncover an interesting phenomenon whereby ectopically formed endogenous dsRNA, and not exo-dsRNA, causes cell death. To the best of our knowledge, this finding is unprecedented and lays ground for further research.

A further implication of our observations lies at the crosstalk between H3K79 methylation and H3K9 methylation. Knowledge of this crosstalk was pioneered by work in yeast (Katan-Khaykovich and Struhl 2005; Ng et al. 2002) and later placed in the context of MLL-driven leukemias (Chen et al. 2015). We have observed that H3K9me2, a hallmark of heterochromatin, is elevated at enhancer regions upon loss of ZFP-1/DOT-1.1. Therefore, DOT1L may prevent the encroachment of heterochromatin to enhancers.

In summary, the compelling observations disclosed here bring together various aspects of DOT1L function (Fig. 6). In addition to controlling transcription through promoters, as previously described (Cecere et al. 2013), DOT1L activity controls enhancer-directed transcription. This finding may attest to the novel concept of the fluidity of enhancer/promoter states. Furthermore, we have linked DOT1L-mediated transcriptional control to dsRNA formation and small RNA generation, as well as to heterochromatin formation, specifically at enhancer regions. Whereas this study connects DOT1L and the RNAi factors required for production of siRNAs from dsRNA genetically, we describe the identification of small RNAs matching enhancers in another publication (Gushchanskaia et al. 2018).

Materials and methods

Strains

Strains were maintained at 20°C under standard conditions (Brenner 1974). Bristol N2 was the WT strain used. The following mutant strains were obtained from the Caenorhabditis Genetics Center (CGC): VC40220, containing *dot-1.1(gk520244)* I,

RB774 – *zfp-1(ok554)* III, WM49 – *rde-4(ne301)* III, VC20787, containing *mml-1(gk402844)* III, GR1373 – *eri-1(mg366)* IV, MT3002 – *ced-3(n1286)* IV, WM27 – *rde-1(ne219)* V. The RB and VC strains were provided by the *C. elegans* Gene Knockout Project at the Oklahoma Medical Research Foundation and by the *C. elegans* Reverse Genetics Core Facility at the University of British Columbia, respectively, which were part of the International *C. elegans* Gene Knockout Consortium. Mutations of interest were outcrossed four times and the following derivative strains were obtained: AGK760 – *mml-1(gk402844)* III, AGK768 – *dot-1.1(gk520244)* I. The *rtIs11 (elt-2P::GFP/LacZ)* X transgenic strain was a gift from Dr. Anne Hart. The *dot-1.1* null mutant (strain COP1302 – *dot-1.1 [knu339 - (pNU1092 - KO loxP::hygR::loxP)]* I; *ced-3 (n1286)* IV), in which exons one to four in the *dot1.1* gene locus were deleted, was generated using a proprietary protocol by Knudra Transgenics. The following derivative strains were produced in this study:

AGK779 – *dot-1.1(gk520244)* I; *rtIs11 (elt-2P::GFP/LacZ)* X

AGK780 – *eri-1(mg366)* IV; *rtIs11 (elt-2P::GFP/LacZ)* X

AGK781 – *dot-1.1(gk520244)* I; *eri-1(mg366)* IV; *rtIs11 (elt-2P::GFP/LacZ)* X

AGK782 – *dot-1.1 [knu339 - (pNU1092 - KO loxP::hygR::loxP)]* I; *rde-4(ne301)* III

AGK783 – *dot-1.1 [knu339 - (pNU1092 - KO loxP::hygR::loxP)]* I; *rde-1(ne219)* V

Microscopy and immunofluorescence

The *elt-2P::GFP/LacZ* transgenic strains were grown at 16°C for at least two generations without starvation before imaging. Live animals were mounted on 2% agarose pads and examined using the Zeiss AxioImager Z1 instrument. The same

exposure times were used for capturing images and comparing GFP expression in different strains.

For immunofluorescence, young adult worms were handpicked and placed in M9 solution supplemented with 0.01% Tween and 25 mM sodium azide on a slide. Worms were dissected using scalpel and fixed with 2% paraformaldehyde for 1 h at room temperature and post-fixed in ice-cold methanol for 5 minutes. Control worms were treated with RNase A prior to fixation at 37°C for 30 minutes to confirm the antibody specificity, and then washed 3 times with M9. After fixation, worms were transferred to glass tubes and blocked with 1% bovin serum albumin (BSA) in M9 supplemented with 0.05% Tween-20. Blocked samples were then incubated with the J2 antibody (1:1000) (a generous gift of Dr. Elke Mühlberger, National Emerging Infectious Diseases Laboratories) overnight at 4°C, washed 3 times in washing buffer (BSA 0.25%, Tween-20 0.05% in M9) and then incubated with secondary fluorophore-conjugated (AF 568) antibody (1:500) for 2 h in the dark. After 3 washes, the samples were transferred to a slide and mounted with DAPI-containing solution (Invitrogen, Prolong Diamond Antifade Mountant with DAPI). The images were taken at 63X on a Zeiss AxioImager Z1 instrument and processed with ImageJ (version 1.51k).

Chromatin immunoprecipitation and library preparation

Hypochlorite-synchronized L1 stage WT or mutant animals were grown for 36 hours post-hatching at 20ºC on standard NGM plates with OP-50 *E. coli*. Synchronized L3 larvae were then washed off the plates using isotonic M9 solution. Fixation was performed for 30 min at 20°C in crosslinking solution (2% paraformaldehyde in M9 buffer), and excess formaldehyde was quenched with 0.1 M Tris-HCl pH 7.5. The pellet

was then washed twice with M9 buffer at 4°C. The worm pellet was resuspended in 1 mL ice-cold RIPA buffer (10 mM Tris-HCl pH 8.0, 0.1% SDS, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM EDTA, 0.15 M NaCl) supplemented with Halt™ Protease Inhibitor Cocktail (Thermo Fisher Scientific). The suspension was then transferred to a Covaris milliTUBE 1 mL with AFA® fiber and shearing was performed using the following settings in SonoLab7: Peak Incident Power (PIP) = 240W; Duty Factor = 20%; Cycles per burst (cpb) = 200; time = 480 seconds. After shearing, the crude extract was spun for 10 min at 10,000 x *g* and the supernatant was collected. Protein concentration was determined by the Bradford method. Extract corresponding to 1 mg protein was diluted to 500 µl with ice-cold RIPA buffer supplemented with Halt™ Protease Inhibitor Cocktail (Thermo Fisher Scientific). Antibody specific for H3K9me2 (Abcam, ab1220, 5 µg) was added and, after incubation at 4°C for 1 h, 50 µl of Dynabeads™ Protein G suspension (Thermo Fisher Scientific) were added and the mixture was incubated at 4°C for 1 h on rotation. Beads were recovered with a Dynal® Magnetic Particle Concentrator (Invitrogen) and washed at 4°C with: 0.5 mL buffer TSE-150 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 150 mM NaCl), 2 X 0.5 mL buffer TSE-500 (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 500 mM NaCl), 0.5 mL buffer TSE-1M (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.1, 1 M NaCl), 0.5 mL buffer LiCl (250 mM LiCl, 1% NP40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.1), 2 X 0.5 mL buffer TE (10 mM Tris pH 8.0, 1 mM EDTA). Antibody-bound chromatin was recovered in 2 X 0.2 mL elution buffer (1% SDS, 0.1 M NaHCO3) by shaking at room temperature for 15 min. Elution buffer was added to input samples (5%) to a total volume of 0.4 mL. Input and ChIP

chromatin samples were subsequently processed in parallel. After 20 µl of 5 M NaCl was added to each sample, cross-links were reversed overnight at 65°C. 10 µl of 0.5 M EDTA, 20 µl of 1 M Tris-HCl pH 6.5 and 0.5 µl of 20 mg/mL Proteinase K (Thermo Fisher Scientific) were added and each sample was incubated for 1 h at 42°C. DNA was recovered using the QIAquick PCR Purification Kit (Qiagen) in accordance with the manufacturer's instructions. The immunoprecipitated DNA was either quantified by qPCR to calculate the percentage of immunoprecipitation relative to the input, or subjected to library preparation and high-throughput sequencing. Immunoprecipitated DNA and input libraries were prepared using TruSeq ChIP Library Preparation kit following the manufacturer's instructions (Illumina) and checked for concentration and quality on DNA chips with a Bioanalyzer (Agilent). 75-bp single read sequences were generated on the NextSeq 500 sequencer according to manufacturer’s instructions (Illumina).

RNA extraction and expression profiling

Synchronized L3 larvae were washed off the plates using isotonic M9 solution. RNA was isolated in triplicate for each mutant strain using TRIzol (Invitrogen) followed by miRNeasy mini column (Qiagen). RNA integrity was confirmed using a 2100 Bioanalyzer (Agilent Technologies). Samples were labeled and hybridized to Affymetrix GeneChip™ *C. elegans* Gene 1.0 ST arrays. Raw Affymetrix CEL files were normalized to produce gene-level expression values using the Robust Multiarray Average (RMA) (Irizarry et al. 2012) in Affymetrix Expression Console (version 1.4.1.46). The default probesets defined by Affymetrix were used to assess array quality using the Area Under the [Receiver Operating Characteristics] Curve (AUC) metric. Principal Component

Analysis (PCA) was performed using the *prcomp* R function with expression values that had been normalized across all samples to a mean of zero and a standard deviation of one. Differential expression was assessed using the moderated (empirical Bayesian) ANOVA and *t* test implemented in the *limma* package (version 3.14.4) (i.e., creating simple linear models with *lmFit*, followed by empirical Bayesian adjustment with *eBayes*). Correction for multiple hypothesis testing was accomplished using the Benjamini-Hochberg false discovery rate (FDR) (Benjamini and Hochberg 1995). All microarray analyses were performed using the R environment for statistical computing (version 2.15.1).

Analysis of chromatin domains and ATAC-seq peaks

Coordinates of chromatin domains were obtained from Evans et al., 2016. DOT-1.1 ChIP-chip and ZFP-1 ChIP-seq peak coordinates were obtained from modENCODE (modENCODE_2970 and modENCODE_6213, respectively). A chromatin domain region was called bound by DOT-1.1 or by ZFP-1 if the center base pair of at least one peak was located within the coordinates of the region. Putative enhancer regions were obtained by combining coordinates of domains 8, 9 and 10. Enhancer domains at least 1500 bp distal to any annotated transcription start site or transcription termination site were considered distal enhancer domains. Enhancer domains intersecting coordinates of genes < 15kb by at least 50 bp were considered intragenic enhancer domains. ATAC-seq peak coordinates (Daugherty et al. 2017) were downloaded from the NCBI GEO database (GSE89608). Distal and intragenic ATAC-seq peaks were obtained as for enhancer domains. Intersections of genomic intervals were performed in R using the valr package (A. Riemondy et al. 2017).

Analysis of GRO-seq data

GRO-seq data was obtained from the NCBI GEO database (GSE47132) and the reads were mapped to the *C. elegans* genome (ce10 assembly) using ChIPdig, a software application to analyze ChIP-seq data (Esse and Grishok 2017). Then, reads matching ribosomal RNA loci were removed, as described before (Cecere et al. 2013). Read counting in regions (either genes, regions or genomic bins) was performed with package GenomicAlignments (Lawrence et al. 2013); only reads with mapping quality 20 or higher were included in subsequent analyses. Regions without reads across the sample set were removed. Counts were then normalized using the TMM method, which takes RNA composition bias into account (Robinson and Oshlack 2010), using the edgeR package (Robinson et al. 2010). Coverage was expressed as RPKM (reads per kilobase per million mapped) and log2-transformed.

Analysis of ChIP-seq data

Analysis of ChIP-seq data was performed as described for the analysis of GRO-seq data, with the exception that, after normalization of read counts, coverage was expressed as the log2-transformed ratio of the RPKM value in the immunoprecipitated DNA sample divided by the RPKM value in the non-immunoprecipitated (input) DNA. Only regions in which the normalized count value in the immunoprecipitated DNA sample was higher than that in the corresponding input DNA in at least one sample in the set were considered.

Acknowledgements

This research was financially supported by the National Institutes of Health R01 GM107056 grant (awarded to A.G.). *C. elegans* strains were provided by the *C. elegans* Gene Knockout Project at the Oklahoma Medical Research Foundation, which is part of the International *C. elegans* Gene Knockout Consortium. We thank the BU Microarray and Sequencing Resource Core Facility (CTSA grant UL1-TR001430) for performing the expression profiling by microarray and sequencing the ChIP-seq libraries. We would also like to acknowledge Yunfan (Frank) Liang for assistance in immunostaining, and members of the labs of Daniel Cifuentes and Nelson Lau for discussions.

References

A. Riemondy K, M. Sheridan R, Gillen A, Yu Y, G. Bennett C, R. Hesselberth J. 2017. valr: Reproducible genome interval analysis in R. *F1000Research* **6**: 1025. https://f1000research.com/articles/6-1025/v1.

Balbin OA, Malik R, Dhanasekaran SM, Prensner JR, Cao X, Wu Y, Robinson D, Wang R, Chen G, Beer DG, et al. 2015. The landscape of antisense gene expression in human cancers. *Genome Res* **25**: 1068–1079.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* **57**: 289–300. http://www.jstor.org/stable/2346101.

Bonn S, Zinzen RP, Girardot C, Gustafson EH, Perez-Gonzalez A, Delhomme N, Ghavi-Helm Y, Wilczyåski B, Riddell A, Furlong EEM. 2012. Tissue-specific

analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet* **44**: 148–156.

Brenner S. 1974. The genetics of Caenorhabditis elegans. *Genetics* **77**: 71–94.

Cecere G, Hoersch S, Jensen MB, Dixit S, Grishok A. 2013. The ZFP-1(AF10)/DOT-1 Complex Opposes H2B Ubiquitination to Reduce Pol II Transcription. *Mol Cell* **50**: 894–907.

Chen CW, Koche RP, Sinha AU, Deshpande AJ, Zhu N, Eng R, Doench JG, Xu H, Chu SH, Qi J, et al. 2015. DOT1L inhibits SIRT1-mediated epigenetic silencing to maintain leukemic gene expression in MLL-rearranged leukemia. *Nat Med* **21**: 335–343.

Cho MH, Park JH, Choi HJ, Park MK, Won HY, Park YJ, Lee CH, Oh SH, Song YS, Kim HS, et al. 2015. DOT1L cooperates with the c-Myc-p300 complex to epigenetically derepress CDH1 transcription factors in breast cancer progression. *Nat Commun* **6**.

Cinghu S, Yang P, Kosak JP, Conway AE, Kumar D, Oldfield AJ, Adelman K, Jothi R. 2017. Intragenic Enhancers Attenuate Host Gene Expression. *Mol Cell* **68**: 104–117.e6.

Daugherty AC, Yeo RW, Buenrostro JD, Greenleaf WJ, Kundaje A, Brunet A. 2017. Chromatin accessibility dynamics reveal novel functional enhancers in C. elegans. *Genome Res* **27**: 2096–2107.

Duchaine TF, Wohlschlegel JA, Kennedy S, Bei Y, Conte D, Pang KM, Brownell DR, Harding S, Mitani S, Ruvkun G, et al. 2006. Functional proteomics reveals the

biochemical niche of C. elegans DCR-1 in multiple small-RNA-mediated pathways. *Cell* **124**: 343–354.

Dudley NR, Labbe JC, Goldstein B. 2002. Using RNA interference to identify genes required for RNA interference. *Proc Natl Acad Sci U S A* **99**: 4191–4196. http://www.ncbi.nlm.nih.gov/pubmed/11904378.

Esse R, Grishok A. 2017. ChIPdig: a comprehensive user-friendly tool for mining multi-sample ChIP-seq data. *bioRxiv*. http://biorxiv.org/content/early/2017/11/16/220079.abstract.

Estornes Y, Toscano F, Virard F, Jacquemin G, Pierrot A, Vanbervliet B, Bonnin M, Lalaoui N, Mercier-Gouy P, Pachéco Y, et al. 2012. DsRNA induces apoptosis through an atypical death complex associating TLR3 to caspase-8. *Cell Death Differ* **19**: 1482–1494.

Evans KJ, Huang N, Stempor P, Chesney MA, Down TA, Ahringer J. 2016. Stable Caenorhabditis elegans chromatin domains separate broadly expressed and developmentally regulated genes. *Proc Natl Acad Sci* **113**: E7020–E7029. http://www.ncbi.nlm.nih.gov/pubmed/27791097%255Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5111720.

Feng Y, Yang Y, Ortega MM, Copeland JN, Zhang M, Jacob JB, Fields TA, Vivian JL, Fields PE. 2010. Early mammalian erythropoiesis requires the Dot1L methyltransferase. *Blood* **116**: 4483–4491.

Fischer SEJ, Pan Q, Breen PC, Qi Y, Shi Z, Zhang C, Ruvkun G. 2013. Multiple small

RNA pathways regulate the silencing of repeated and foreign genes in C. elegans. *Genes Dev* **27**: 2678–2695.

Fukushige T, Hendzel MJ, Bazett-Jones DP, McGhee JD. 1999. Direct visualization of the elt-2 gut-specific GATA factor binding to a target promoter inside the living Caenorhabditis elegans embryo. *Proc Natl Acad Sci U S A* **96**: 11883–11888. http://www.ncbi.nlm.nih.gov/pubmed/10518545.

Gent JI, Lamm AT, Pavelec DM, Maniar JM, Parameswaran P, Tao L, Kennedy S, Fire AZ. 2010. Distinct Phases of siRNA Synthesis in an Endogenous RNAi Pathway in C. elegans Soma. *Mol Cell* **37**: 679–689.

Grewal SIS, Elgin SCR. 2007. Transcription and RNA interference in the formation of heterochromatin. *Nature* **447**: 399–406.

Grishok A. 2005. RNAi mechanisms in Caenorhabditis elegans. *FEBS Lett* **579**: 5932–5939.

Grishok A, Sinskey JL, Sharp PA. 2005. Transcriptional silencing of a transgene by RNAi in the soma of C. elegans. *Genes Dev* **19**: 683–696.

Gushchanskaia E, Esse R, Ma Q, Lau N, Grishok A. 2018. The interplay between small RNA pathways shapes chromatin landscape in C. elegans. *bioRxiv*. http://biorxiv.org/content/early/2018/05/11/320713.abstract.

Haerry TE, Gehring WJ. 1996. Intron of the mouse Hoxa-7 gene contains conserved homeodomain binding sites that can function as an enhancer element in Drosophila. *Proc Natl Acad Sci U S A* **93**: 13884–13889.

Hah N, Murakami S, Nagari A, Danko CG, Lee Kraus W. 2013. Enhancer transcripts mark active estrogen receptor binding sites. *Genome Res* **23**: 1210–1223.

Henriques T, Scruggs BS, Inouye MO, Muse GW, Williams LH, Burkholder AB, Lavender CA, Fargo DC, Adelman K. 2018. Widespread transcriptional pausing and elongation control at enhancers. *Genes Dev* **32**: 26–41.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. 2012. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. In *Selected Works of Terry Speed*, pp. 601–616.

Kalai M, Van Loo G, Vanden Berghe T, Meeus A, Burm W, Saelens X, Vandenabeele P. 2002. Tipping the balance between necrosis and apoptosis in human and murine cells treated with interferon and dsRNA. *Cell Death Differ* **9**: 981–994.

Kaneko H, Dridi S, Tarallo V, Gelfand BD, Fowler BJ, Cho WG, Kleinman ME, Ponicsan SL, Hauswirth WW, Chiodo VA, et al. 2011. DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. *Nature* **471**: 325–332.

Katan-Khaykovich Y, Struhl K. 2005. Heterochromatin formation involves changes in histone modifications over multiple cell generations. *EMBO J* **24**: 2138–2149.

Keegan LP, Haerry TE, Crotty DA, Packer AI, Wolgemuth DJ, Gehring WJ. 1997. A sequence conserved in vertebrate Hox gene introns functions as an enhancer regulated by posterior homeotic genes in Drosophila imaginal discs. *Mech Dev* **63**: 145–157.

Kennedy LM, Grishok A. 2014. Neuronal migration is regulated by endogenous RNAi

and chromatin-binding factor ZFP-1/AF10 in Caenorhabditis elegans. *Genetics* **197**: 207–220.

Kennedy S, Wang D, Ruvkun G. 2004. A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. *Nature* **427**: 645–649.

Kim JK, Gabel HW, Kamath RS, Tewari M, Pasquinelli A, Rual JF, Kennedy S, Dybbs M, Bertin N, Kaplan JM, et al. 2005. Functional genomic analysis of RNA interference in C. elegans. *Science (80-)* **308**: 1164–1167. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15790806.

Krystal GW, Armstrong BC, Battey JF. 1990. N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. *Mol Cell Biol* **10**: 4180–4191. http://mcb.asm.org/lookup/doi/10.1128/MCB.10.8.4180.

Kuntimaddi A, Achille NJ, Thorpe J, Lokken AA, Singh R, Hemenway CS, Adli M, Zeleznik-Le NJ, Bushweller JH. 2015. Degree of Recruitment of DOT1L to MLL-AF9 Defines Level of H3K79 Di- and Tri-methylation on Target Genes and Transformation Potential. *Cell Rep* **11**: 808–820.

Lam MTY, Cho H, Lesch HP, Gosselin D, Heinz S, Tanaka-Oishi Y, Benner C, Kaikkonen MU, Kim AS, Kosaka M, et al. 2013. Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* **498**: 511–515.

Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT,

Carey VJ. 2013. Software for Computing and Annotating Genomic Ranges. *PLoS Comput Biol* **9**.

Lee RC. 2006. Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. *RNA* **12**: 589–597. http://www.rnajournal.org/cgi/doi/10.1261/rna.2231506.

Li W, Notani D, Rosenfeld MG. 2016. Enhancers as non-coding RNA transcription units: Recent insights and future perspectives. *Nat Rev Genet* **17**: 207–223.

Mansisidor AR, Cecere G, Hoersch S, Jensen MB, Kawli T, Kennedy LM, Chavez V, Tan MW, Lieb JD, Grishok A. 2011. A conserved PHD finger protein and endogenous RNAi modulate insulin signaling in caenorhabditis elegans. *PLoS Genet* **7**.

McLean CM, Karemaker ID, van Leeuwen F. 2014. The emerging roles of DOT1L in leukemia and normal development. *Leukemia* **28**: 2131–2138.

Meyer C, Burmeister T, Gröger D, Tsaur G, Fechina L, Renneville A, Sutton R, Venn NC, Emerenciano M, Pombo-De-Oliveira MS, et al. 2018. The MLL recombinome of acute leukemias in 2017. *Leukemia* **32**: 273–284.

Mikhaylichenko O, Bondarenko V, Harnett D, Schor IE, Males M, Viales RR, Furlong EEM. 2018. The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev* **32**: 42–57.

Mohan M, Herz H-M, Takahashi Y-H, Lin C, Lai KC, Zhang Y, Washburn MP, Florens L, Shilatifard A. 2010. Linking H3K79 trimethylation to Wnt signaling through a novel

Dot1-containing complex (DotCom). *Genes Dev* **24**: 574–589.

Nepveu A, Marcu KB. 1986. Intragenic pausing and anti-sense transcription within the murine c-myc locus. *Embo J* **5**: 2859–2865. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3024965.

Ng HH, Feng Q, Wang H, Erdjument-Bromage H, Tempst P, Zhang Y, Struhl K. 2002. Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev* **16**: 1518–1527.

Nguyen AT, Xiao B, Neppl RL, Kallin EM, Li J, Chen T, Wang D-Z, Xiao X, Zhang Y. 2011. DOT1L regulates dystrophin expression and is critical for cardiac function. *Genes Dev* **25**: 263–274.

Ni JZ, Chen E, Gu SG. 2014. Complex coding of endogenous siRNA, transcriptional silencing and H3K9 methylation on native targets of germline nuclear RNAi in C. elegans. *BMC Genomics* **15**.

Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, Su L, Xu G, Zhang Y. 2005. hDOT1L links histone methylation to leukemogenesis. *Cell* **121**: 167–178.

Parker GS. 2006. RDE-4 preferentially binds long dsRNA and its dimerization is necessary for cleavage of dsRNA to siRNA. *RNA* **12**: 807–818. http://www.rnajournal.org/cgi/doi/10.1261/rna.2338706.

Piatek MJ, Henderson V, Zynad HS, Werner A. 2016. Natural antisense transcription from a comparative perspective. *Genomics* **108**: 56–63.

Popovic R, Erfurth F, Zeleznik-Le N. 2008. Transcriptional complexity of the HOXA9 locus. *Blood Cells, Mol Dis* **40**: 156–159.

Prange KHM, Mandoli A, Kuznetsova T, Wang SY, Sotoca AM, Marneth AE, Van Der Reijden BA, Stunnenberg HG, Martens JHA. 2017. MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia. *Oncogene* **36**: 3346–3356.

Rennie S, Dalby M, Lloret-Llinares M, Bakoulis S, Vaagenso CD, Jensen TH, Andersson R. 2017. Transcription start site analysis reveals widespread divergent transcription in D. melanogaster and core promoter encoded enhancer activities. *bioRxiv* 221952. https://www.biorxiv.org/content/early/2017/11/18/221952.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139–40. http://www.ncbi.nlm.nih.gov/pubmed/19910308%255Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2796818.

Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**.

Saldi TK, Ash PEA, Wilson G, Gonzales P, Garrido-lecca A, Christine M, Dostal V, Gendron TF, Stein LD, Blumenthal T, et al. 2014. TDP- 1 , the Caenorhabditis elegans ortholog of TDP- 43 , limits the accumulation of double-stranded RNA. *EMBO J* **33**: 2947–2966.

Sessa L, Breiling A, Lavorgna G, Silvestri L, Casari G, Orlando V. 2006. Noncoding RNA synthesis and loss of Polycomb group repression accompanies the colinear activation of the human HOXA cluster. *RNA* **13**: 223–239. http://www.rnajournal.org/cgi/doi/10.1261/rna.266707.

Shanower GA, Muller M, Blanton JL, Honti V, Gyurkovics H, Schedl P. 2005. Characterization of the grappa gene, the Drosophila histone H3 lysine 79 methyltransferase. *Genetics* **169**: 173–184.

Simmer F, Tijsterman M, Parrish S, Koushika SP, Nonet ML, Fire A, Ahringer J, Plasterk RHA. 2002. Loss of the putative RNA-directed RNA polymerase RRF-3 makes C. Elegans hypersensitive to RNAi. *Curr Biol* **12**: 1317–1319.

Sur I, Taipale J. 2016. The role of enhancers in cancer. *Nat Rev Cancer* **16**: 483–493.

Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC. 1999. The rde-1 gene, RNA interference, and transposon silencing in C. elegans. *Cell* **99**: 123–132.

Tabara H, Yigit E, Siomi H, Mello CC. 2002. The dsRNA binding protein RDE-4 interacts with RDE-1, DCR-1, and a DExH-Box helicase to direct RNAi in C. elegans. *Cell* **109**: 861–871.

Vasale JJ, Gu W, Thivierge C, Batista PJ, Claycomb JM, Youngman EM, Duchaine TF, Mello CC, Conte D. 2010. Sequential rounds of RNA-dependent RNA transcription drive endogenous small-RNA biogenesis in the ERGO-1/Argonaute pathway. *Proc Natl Acad Sci* **107**: 3582–3587.

http://www.pnas.org/cgi/doi/10.1073/pnas.0911908107.

Wenric S, ElGuendi S, Caberg J-H, Bezzaou W, Fasquelle C, Charloteaux B, Karim L, Hennuy B, Frères P, Collignon J, et al. 2017. Transcriptome-wide analysis of natural antisense transcripts shows their potential role in breast cancer. *Sci Rep* **7**: 17452. http://www.nature.com/articles/s41598-017-17811-2.

Wilkins C, Dishongh R, Moore SC, Whitt MA, Chow M, Machaca K. 2005. RNA interference is an antiviral defence mechanism in Caenorhabditis elegans. *Nature* **436**: 1044–1047.

Wong M, Tee AEL, Milazzo G, Bell JL, Poulos RC, Atmadibrata B, Sun Y, Jing D, Ho N, Ling D, et al. 2017. The histone methyltransferase DOT1L promotes neuroblastoma by regulating gene transcription. *Cancer Res* **77**: 2522–2533.

Xu K, Tavernarakis N, Driscoll M. 2001. Necrotic cell death in C. elegans requires the function of calreticulin and regulators of Ca2+ release from the endoplasmic reticulum. *Neuron* **31**: 957–971.

Xue D, Shaham S, Horvitz HR. 1996. The Caenorhabditis elegans cell-death protein CED-3 is a cysteine protease with substrate specificities similar to those of the human CPP32 protease. *Genes Dev* **10**: 1073–1083.

Yamane D, Kato K, Tohya Y, Akashi H. 2006. The double-stranded RNA-induced apoptosis pathway is involved in the cytopathogenicity of cytopathogenic Bovine viral diarrhea virus. *J Gen Virol* **87**: 2961–2970.

Yigit E, Batista PJ, Bei Y, Pang KM, Chen CCG, Tolia NH, Joshua-Tor L, Mitani S,

Simard MJ, Mello CC. 2006. Analysis of the C. elegans Argonaute Family Reveals that Distinct Argonautes Act Sequentially during RNAi. *Cell* **127**: 747–757.

Zhu B, Chen S, Wang H, Yin C, Han C, Peng C, Liu Z, Wan L, Zhang X, Zhang J, et al. 2018. The protective role of DOT1L in UV-induced melanomagenesis. *Nat Commun* **9**.

Figure Legends

Figure 1. ZFP-1/DOT-1.1 negatively modulate enhancer-directed transcription. (*A*) Overlap analysis showing that promoters and enhancer chromatin domains are enriched in ZFP-1/DOT-1.1. DOT-1.1 ChIP-chip and ZFP-1 ChIP-seq peak coordinates were obtained from modENCODE (modENCODE_2970 and modENCODE_6213, respectively). Chromatin domain coordinates were obtained by hidden Markov models (Evans et al. 2016). Data are presented as percentage of regions enriched in DOT-1.1 or ZFP-1 in each of the 20 chromatin domains described for embryos and for third larval stage animals, respectively. A chromatin domain region was called bound by DOT-1.1 or ZFP-1 if the center base pair of at least one peak was located within the coordinates of the region. Domains 1 and 8 at the embryonic stage, as well as domain 1 at the third larval stage (L3), were annotated as promoter regions. Domain 8 at L3 and domain 9 in both embryos and L3 correspond to predicted intronic and intergenic enhancers, respectively (Evans et al. 2016). (*B*) Intragenic and distal ATAC-seq peaks in embryos and L3 animals are enriched in DOT-1.1 and ZFP-1, respectively. *C. elegans* gene coordinates (WS220/ce10) were extracted from the UCSC genome browser and intersected with ATAC-seq peak coordinates (Daugherty et al. 2017). Peaks at least 1500 bp distal to any annotated transcription start site or transcription termination site were considered distal ATAC-seq peaks. Peaks intersecting coordinates of genes < 15kb by at least 50 bp were considered intragenic ATAC-seq peaks. (*C*) UCSC genome browser snapshots showing representative examples of overlap of ZFP-1 with distal enhancer domains and ATAC-seq peaks (left) and with intragenic enhancer domains (right). (*D*) Transcription originating from intragenic and distal enhancers predicted by

chromatin modification signatures (Evans et al. 2016) is elevated in *zfp-1* mutant animals compared with WT larvae (left). GRO-seq coverage was normalized using 250 bp bins spanning the genome. The P-value was determined by Mann-Whitney U test (enhancers versus genomic bins). Representative examples are shown in the UCSC genome browser multi-region snapshot (right). Each highlighted genomic region corresponds to a distal enhancer domain. (*E*) Transcription originating from distal enhancers corresponding to open chromatin regions (Daugherty et al. 2017) is elevated in *zfp-1* mutant animals compared with WT larvae (left). GRO-seq coverage normalization and statistical tests were performed as in (D). Representative examples are shown in the UCSC genome browser multi-region snapshot (right). Each highlighted region corresponds to a distal ATAC-seq peak.

Figure 2. ZFP-1/DOT-1.1 prevent double-stranded RNA formation at target genes with intragenic enhancers. (*A*) Anti-sense transcription at ZFP-1 target genes with intragenic enhancer signature is increased in *zfp-1* mutant animals compared with WT larvae, whereas sense transcription does not change. The corresponding steady-state mRNA population does not overall change significantly between *zfp-1* mutant worms and WT animals. ** and *** denote P-values < 0.01 and < 0.001, respectively, as determined by Mann-Whitney U test (each category of genes versus all genes). Genes with promoter-bound ZFP-1 were not included in this analysis. (*B*) UCSC genome browser multi-region snapshot showing representative examples of increase of anti-sense transcription at ZFP-1 gene targets with enhancer signature in *zfp-1* mutant animals (orange) compared with WT larvae (gray). Each highlighted region corresponds to a gene. (*C*) Double-

stranded RNA (dsRNA) accumulates in *zfp-1* mutant animals. Fixed, isolated intestinal tissue was probed with anti-dsRNA antibody (J2) recognizing stretches of 40 bp or more in a sequence-independent manner. Stained inclusions (red) were detected in intestinal nuclei (blue, DAPI counterstain) and were increased in *zfp-1* mutant animals compared with WT larvae. Scale bar (in red, upper left panel): 50 µm.

Figure 3. Lethality of *dot-1.1* knockout is suppressed by Dicer/RDE-4/RDE-1 pathway mutations. (*A*) (left) Western blot with the anti-DOT-1.1 antibody (Cecere et al. 2013) detects DOT-1.1 in the control strain, but not in *dot-1.1(knu339)*. (right) The *dot-1.1; ced-3* mutant displays global loss of H3K79me2 as assessed by Western blotting, in agreement with the function of DOT-1.1 in catalyzing H3K79 methylation. (*B*) Outcrossing the *ced-3(n1286)* allele from the *dot-1.1; ced-3* mutant results in arrested development, cell death. Note the numerous vacuoles typical of necrotic cell death in *C. elegans*. (*C*) Mutations in *rde-4* and *rde-1*, both affecting primary siRNA production and function, rescue the lethality phenotype associated with knockout of the *dot-1.1* locus.

Figure 4. Ectopic endogenous dsRNA formed upon loss of ZFP-1/DOT-1.1 titrates the Dicer complex away from exogenous dsRNA. (*A*) In *dot-1.1(gk520244)* partial loss-of-function mutant worms, expression of the *elt-2::*GFP/*lacZ* transgene is increased compared with WT animals, indicating reduced efficiency of silencing by exogenously introduced dsRNA. The level of GFP expression in *dot-1.1; eri-1* mutant worms is similar to that observed in *eri-1* mutant worms, showing that DOT-1.1 acts upstream of

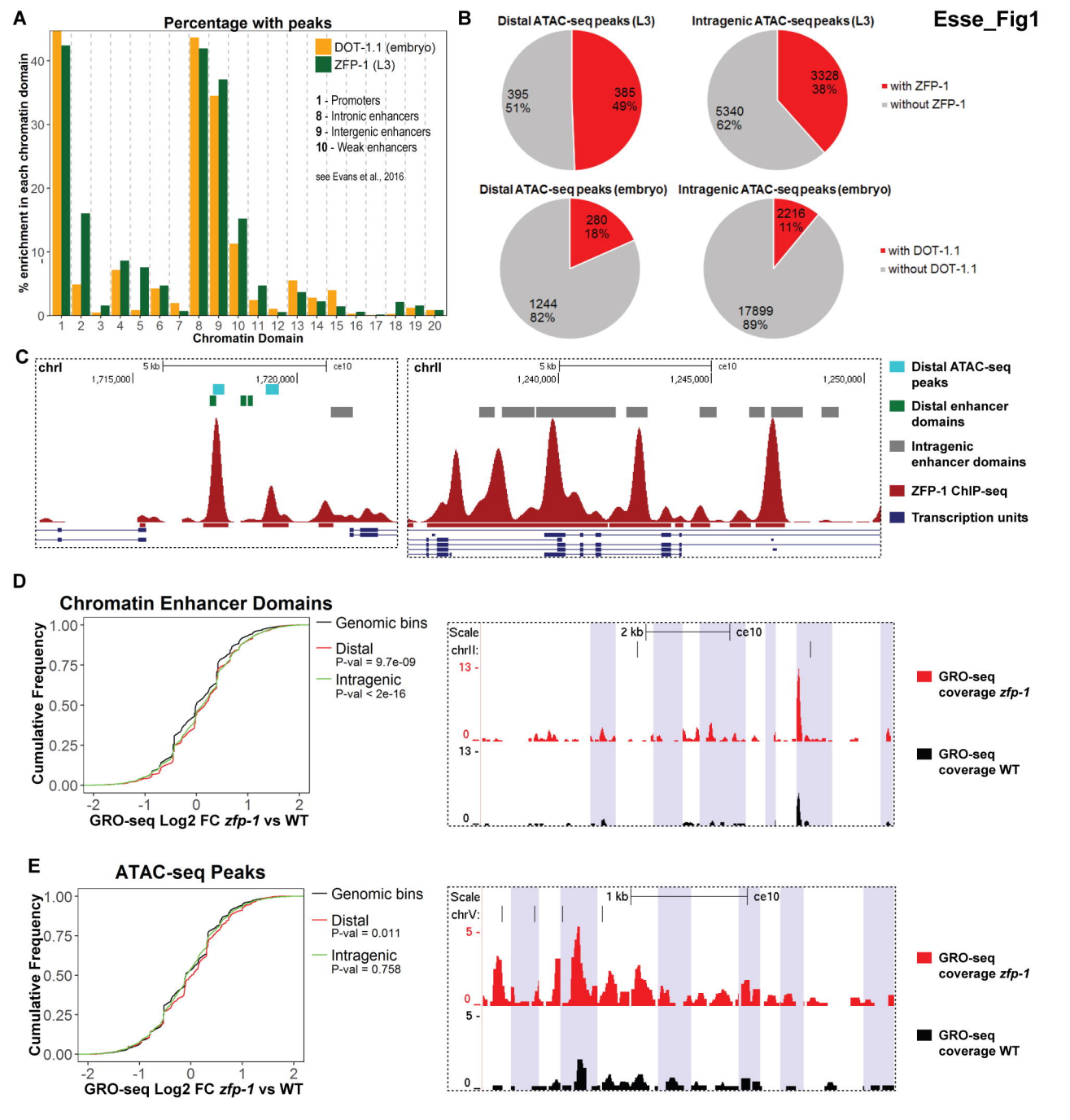
ERI-1/Dicer. Scale bar (in white, upper left panel): 200 µm. (*B*) Different pools of dsRNA compete for the Dicer/RDE-4/RDE-1 pathway, including ectopic dsRNA produced endogenously. The latter is exacerbated upon loss of ZFP-1/DOT-1.1.

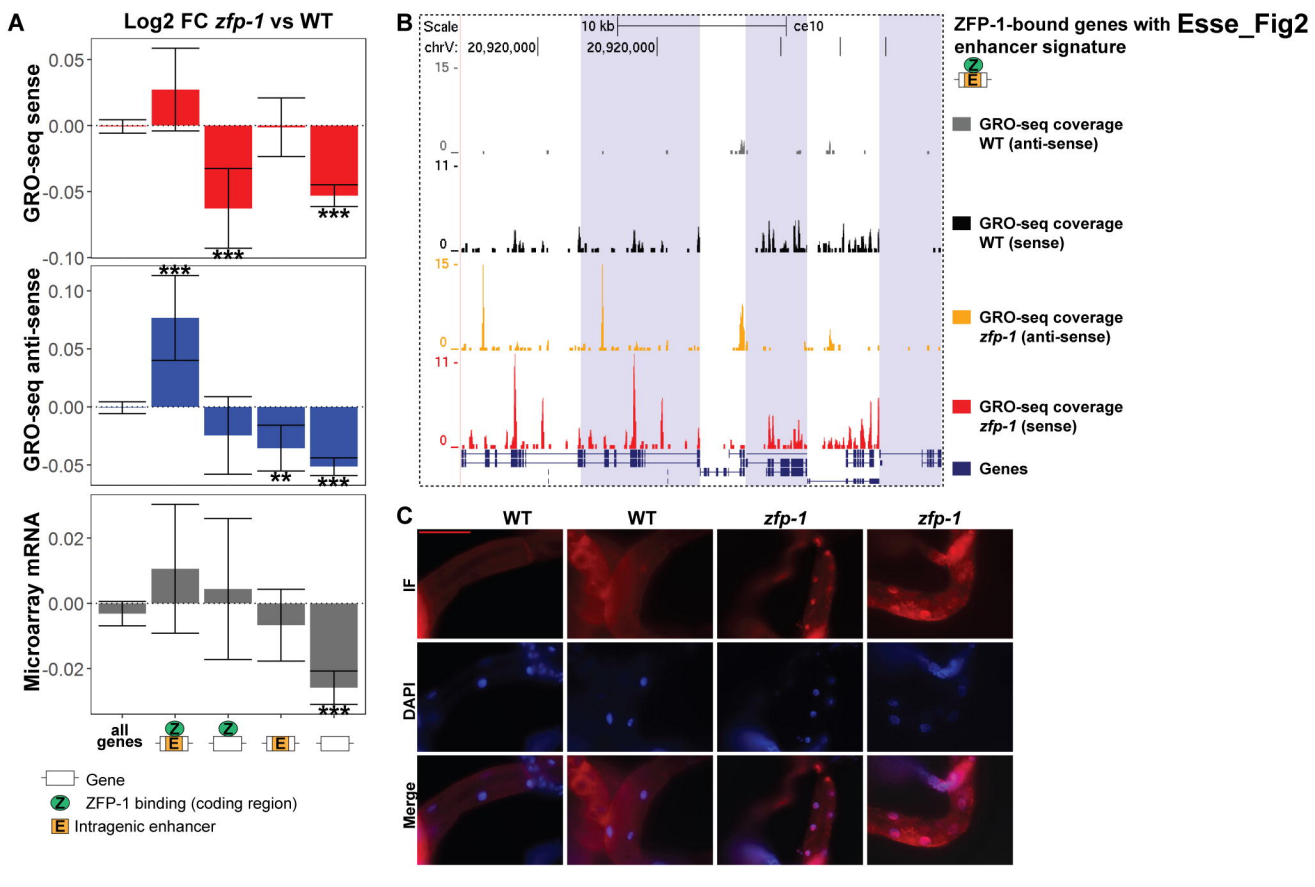
Figure 5. Distal and intragenic enhancers show ectopic heterochromatin upon loss of DOT-1.1. (*A*) Distal and intragenic enhancers predicted by chromatin enhancer signatures display ectopic H3K9me2 in the *dot-1.1; ced-3* mutant compared with *ced-3* animals. The P-value was determined by Mann-Whitney U test (enhancers versus genomic bins). (right) Representative examples are shown in the UCSC genome browser; each highlighted region corresponds to a distal enhancer domain. (*B*) H3K9me2 is elevated at genes with enhancer signature in the *dot-1.1; ced-3* mutant compared with *ced-3* animals. The P-value was determined by Mann-Whitney U test (*dot-1.1; ced-3* versus *ced-3*). (right) Representative examples are shown in the UCSC genome browser; each highlighted region corresponds to an intragenic enhancer domain.

Figure 6. Mechanism of initiation of nuclear RNAi upon loss of DOT1L. Loss of ZFP-1/DOT1L leads to increase in anti-sense transcription, resulting in accumulation of endogenous dsRNA molecules. These are processed by the Dicer complex into endo-siRNAs, ultimately leading to H3K9 methylation and heterochromatin formation. Generation of this new pool of dsRNA attracts the dsRNA-binding RNAi pathway protein RDE-4 and Dicer, thus limiting the efficiency of exogenous dsRNA-induced silencing.

Supplemental Figure S1. Nascent transcription is negatively regulated by the ZFP-1/DOT-1.1 complex. (*A*) Transcription originating from intragenic and distal enhancers predicted by chromatin modification domains 8, 9 and 10 (Evans et al. 2016) comprises approximately 12% of total nascent transcription observed in WT animals, as assessed by GRO-seq (Cecere et al. 2013). (*B*) Nascent transcription is increased in *zfp-1* mutant animals compared with WT larvae not only at promoter-proximal regions and gene bodies represented by domains 1, 2 and 3, as previously reported (Cecere et al. 2013), but also at chromatin enhancer domains 8, 9 and 10. The P-value was determined by Mann-Whitney U test (*zfp-1* versus WT). * and *** denote P-values < 0.05 and < 0.001, respectively.

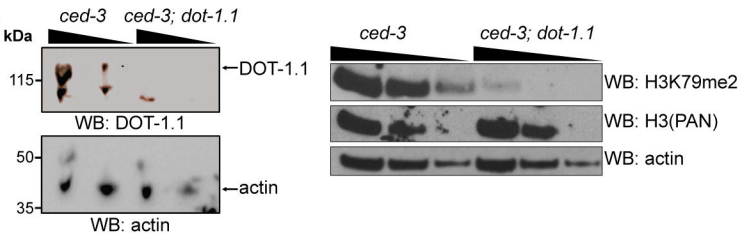
Supplemental Figure S2. ChIP-seq Analysis of H3K9me2 on Independently Obtained Replicas Shows Ectopic Heterochromatin Formation at Enhancers upon Loss of DOT-1.1. (*A*) Ectopic H3K9me2 at enhancer domains in the *dot-1.1; ced-3* mutant compared with *ced-3* animals in two independent experiments. The P-value was determined by Mann-Whitney U test (enhancers versus genomic bins). (*B*) Increased H3K9me2 at genes with enhancer signature in the *dot-1.1; ced-3* mutant compared with *ced-3* animals in the two independent experiments performed. The P-value was determined by Mann-Whitney U test (*dot-1.1; ced-3* versus *ced-3*).


Esse_Fig1
A
Percentage with peaks
DOT-1.1 (embryo)
ZFP-1 (L3)
1 - Promoters
8 - Intronic enhancers
9 - Intergenic enhancers
10 - Weak enhancers
see Evans et al., 2016
% enrichment in each chromatin domain
Chromatin Domain
B
Distal ATAC-seq peaks (L3)
385 49%
395 51%
Intragenic ATAC-seq peaks (L3)
3328 38%
5340 62%
with ZFP-1
without ZFP-1
Distal ATAC-seq peaks (embryo)
280 18%
1244 82%
Intragenic ATAC-seq peaks (embryo)
2216 11%
17899 89%
with DOT-1.1
without DOT-1.1
C
chrI
chrII
Distal ATAC-seq peaks
Distal enhancer domains
Intragenic enhancer domains
ZFP-1 ChIP-seq
Transcription units
D
Chromatin Enhancer Domains
Genomic bins
Distal P-val = 9.7e-09
Intragenic P-val < 2e-16
Cumulative Frequency
GRO-seq Log2 FC zfp-1 vs WT
GRO-seq coverage zfp-1
GRO-seq coverage WT
E
ATAC-seq Peaks
Genomic bins
Distal P-val = 0.011
Intragenic P-val = 0.758
Cumulative Frequency
GRO-seq Log2 FC zfp-1 vs WT
GRO-seq coverage zfp-1
GRO-seq coverage WT

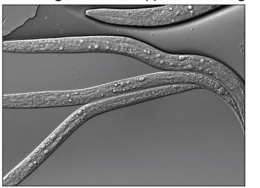

Esse_Fig2
A
Log2 FC zfp-1 vs WT
GRO-seq sense
GRO-seq anti-sense
Microarray mRNA
all genes
Gene
ZFP-1 binding (coding region)
Intragenic enhancer
B
Scale
chrV:
20,920,000
10 kb
ce10
ZFP-1-bound genes with enhancer signature
GRO-seq coverage WT (anti-sense)
GRO-seq coverage WT (sense)
GRO-seq coverage zfp-1 (anti-sense)
GRO-seq coverage zfp-1 (sense)
Genes
C
WT
WT
zfp-1
zfp-1
IF
DAPI
Merge

Esse_Fig3

A



B *dot-1.1(-); ced-3(+)* arrested larva resulting from *ced-3(-)* outcrossing



C

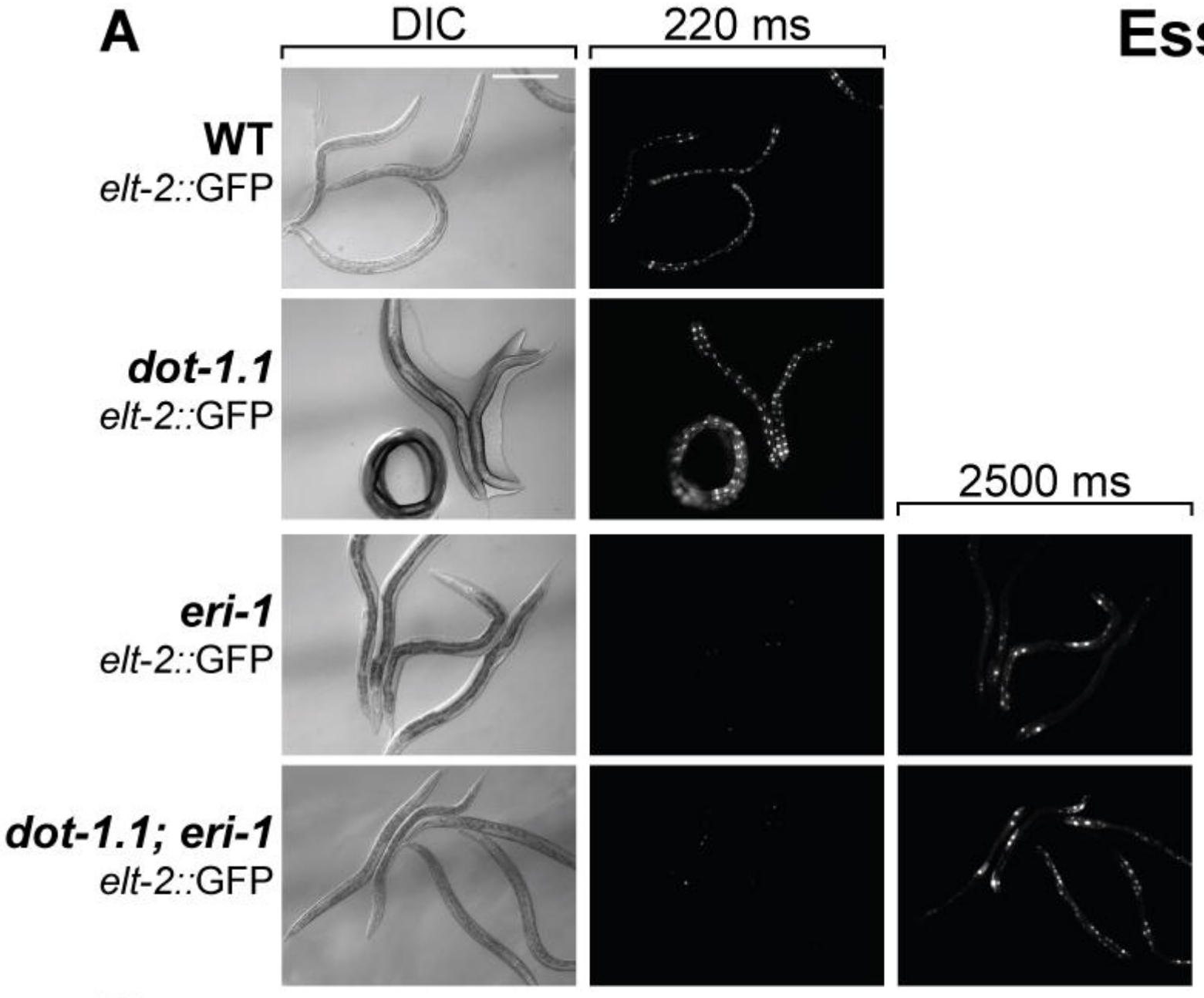
Results of *dot-1.1* outcrosses from *ced-3* background

Mutant crossed	F2 picked (N)	Homozygous viable *dot-1.1* F2 lines			
		WT background	*ced-3* status	Mutant background	*ced-3* status
mml-1(gk402844)	21	2	2/2 (-)	3	3/3 (-)
rde-4(ne301)	40	0	n/a	5	2/5 (-) 3/5 (+)
*rde-1(ne219)**	10	n/a	n/a	4	1/4 (-) 3/4 (+)

* Mutant F2 animals were pre-selected

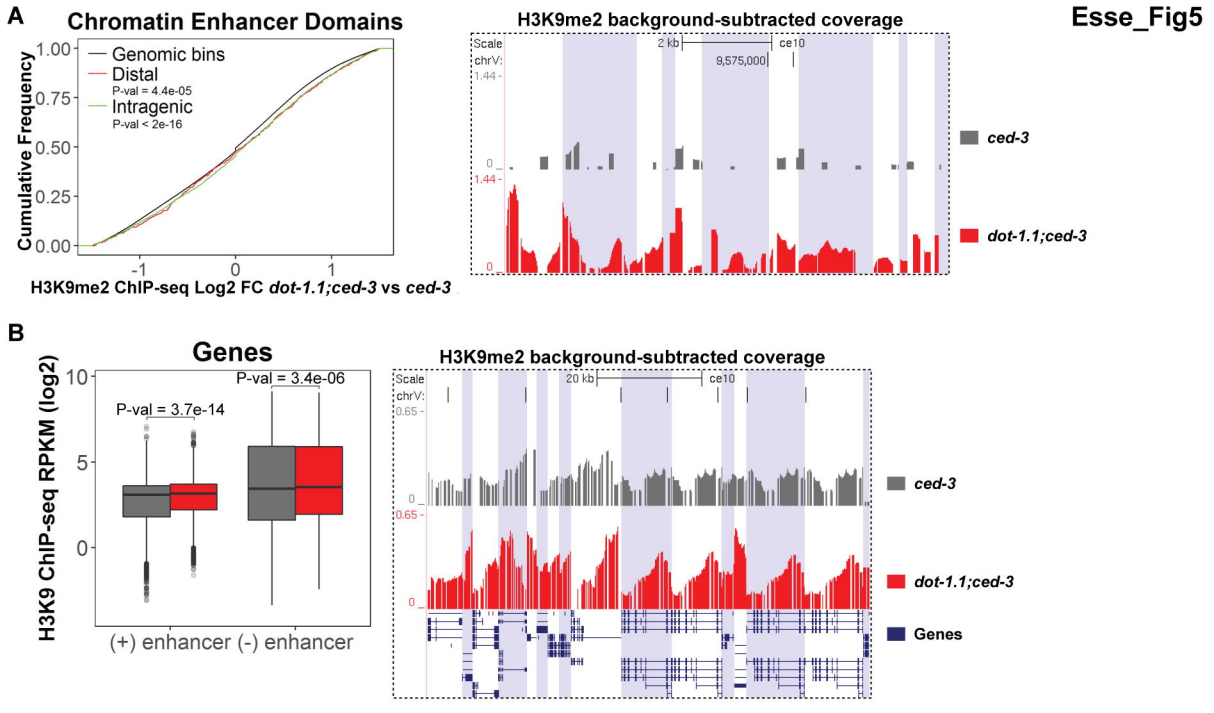
Esse_Fig4

A

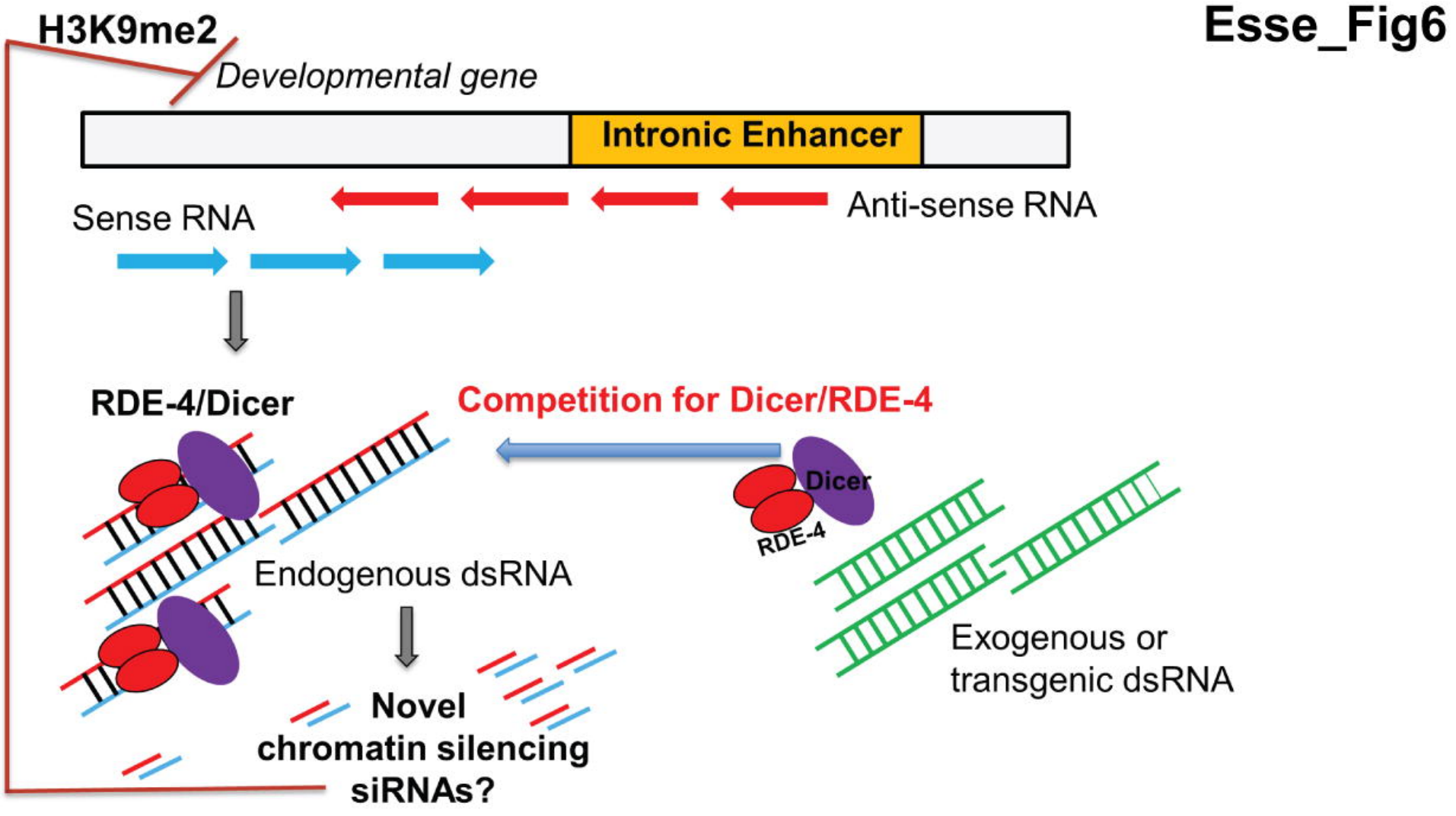


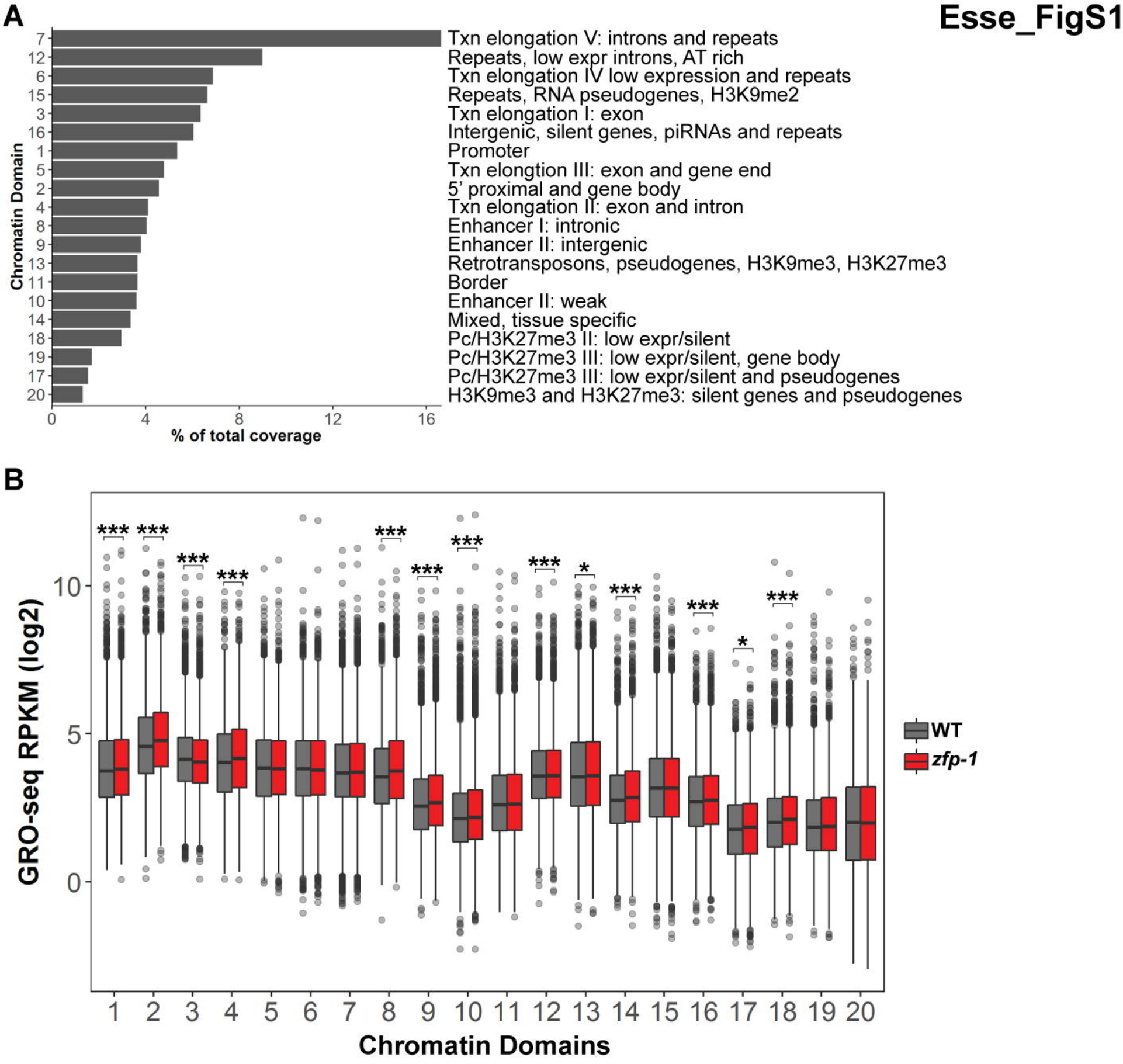
B

***elt-2::* GFP background**	**Competing dsRNA pools**		
	Repetitive transgenes	ERI pathway-produced	Ectopic endogenous
WT	+	+	-
eri-1	+	-	-
dot-1.1	+	+	+
eri-1; dot-1.1	+	-	+

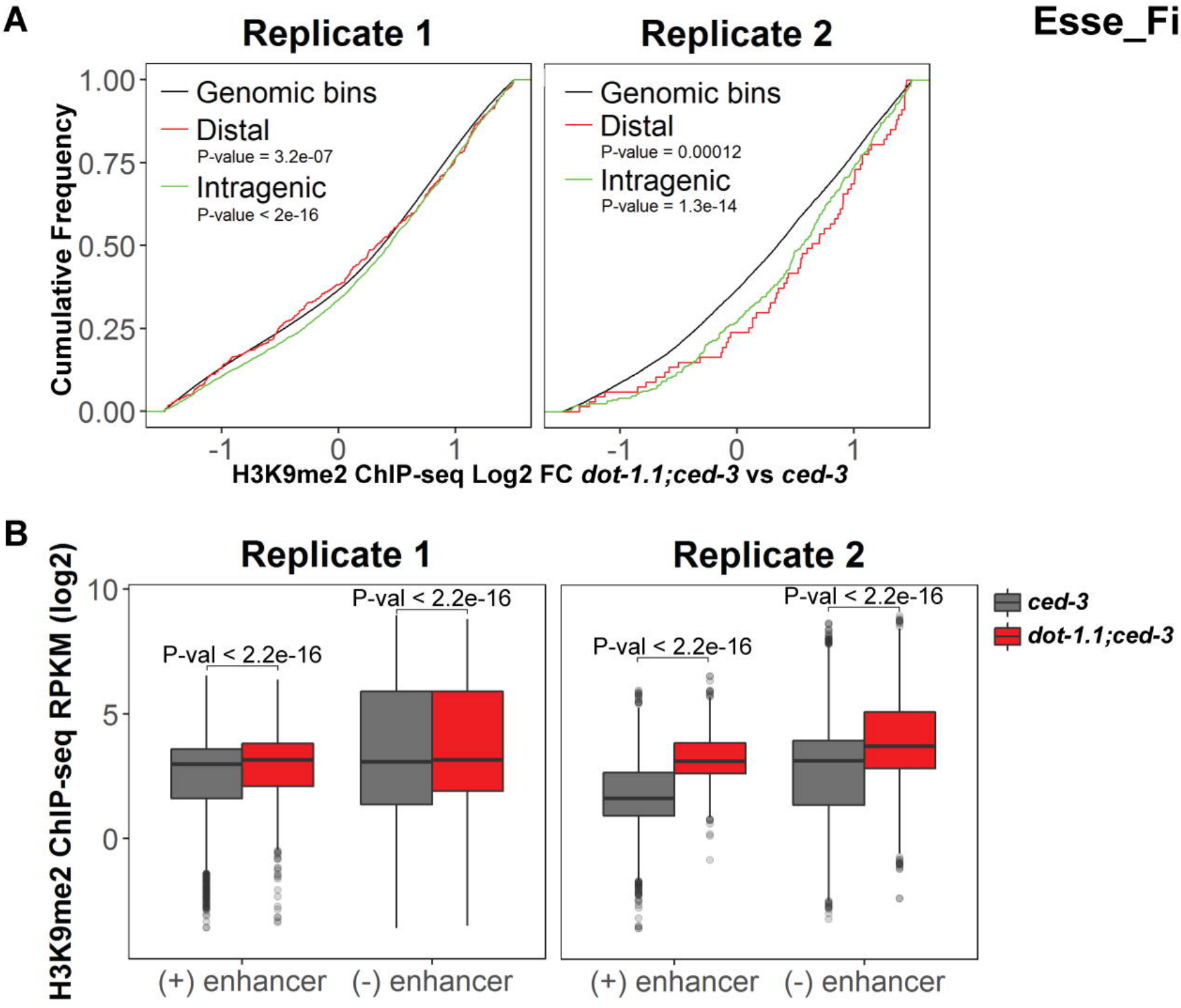


Esse_Fig5
A
Chromatin Enhancer Domains
Cumulative Frequency
1.00
0.75
0.50
0.25
0.00
Genomic bins
Distal
P-val = 4.4e-05
Intragenic
P-val < 2e-16
-1
0
1
H3K9me2 ChIP-seq Log2 FC dot-1.1;ced-3 vs ced-3
H3K9me2 background-subtracted coverage
Scale
chrV:
2 kb
ce10
9,575,000
1.44
0
1.44
0
ced-3
dot-1.1;ced-3
B
Genes
H3K9 ChIP-seq RPKM (log2)
10
5
0
P-val = 3.7e-14
P-val = 3.4e-06
(+) enhancer
(-) enhancer
H3K9me2 background-subtracted coverage
Scale
chrV:
20 kb
ce10
0.65
0
0.65
0
ced-3
dot-1.1;ced-3
Genes


Esse_Fig6
H3K9me2
Developmental gene
Intronic Enhancer
Sense RNA
Anti-sense RNA
RDE-4/Dicer
Competition for Dicer/RDE-4
Dicer
RDE-4
Endogenous dsRNA
Exogenous or
transgenic dsRNA
Novel
chromatin silencing
siRNAs?


Esse_FigS1
A
Chromatin Domain
% of total coverage
0
4
8
12
16
7 Txn elongation V: introns and repeats
12 Repeats, low expr introns, AT rich
6 Txn elongation IV low expression and repeats
15 Repeats, RNA pseudogenes, H3K9me2
3 Txn elongation I: exon
16 Intergenic, silent genes, piRNAs and repeats
1 Promoter
5 Txn elongtion III: exon and gene end
2 5' proximal and gene body
4 Txn elongation II: exon and intron
8 Enhancer I: intronic
9 Enhancer II: intergenic
13 Retrotransposons, pseudogenes, H3K9me3, H3K27me3
11 Border
10 Enhancer II: weak
14 Mixed, tissue specific
18 Pc/H3K27me3 II: low expr/silent
19 Pc/H3K27me3 III: low expr/silent, gene body
17 Pc/H3K27me3 III: low expr/silent and pseudogenes
20 H3K9me3 and H3K27me3: silent genes and pseudogenes
B
GRO-seq RPKM (log2)
10
5
0
Chromatin Domains
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
WT
zfp-1

Esse_FigS2


A
Replicate 1
Replicate 2
Genomic bins
Distal
P-value = 3.2e-07
Intragenic
P-value < 2e-16
Genomic bins
Distal
P-value = 0.00012
Intragenic
P-value = 1.3e-14
Cumulative Frequency
1.00
0.75
0.50
0.25
0.00
-1
0
1
H3K9me2 ChIP-seq Log2 FC dot-1.1;ced-3 vs ced-3
B
Replicate 1
Replicate 2
H3K9me2 ChIP-seq RPKM (log2)
10
5
0
P-val < 2.2e-16
P-val < 2.2e-16
P-val < 2.2e-16
P-val < 2.2e-16
(+) enhancer
(-) enhancer
(+) enhancer
(-) enhancer
ced-3
dot-1.1;ced-3