Trithorax and Polycomb group-dependent regulation: a tale of opposing activities

Sarah J. Geisler¹ and Renato Paro¹,²,*

ABSTRACT

Intricate layers of regulation determine the unique gene expression profiles of a given cell and, therefore, underlie the immense phenotypic diversity observed among cell types. Understanding the mechanisms that govern which genes are expressed and which genes are silenced is a fundamental focus in biology. The Polycomb and Trithorax group chromatin proteins play important roles promoting the stable and heritable repression and activation of gene expression, respectively. These proteins, which are conserved across metazoans, modulate post-translational modifications on histone tails and regulate nucleosomal structures. Here, we review recent advances that have shed light on the mechanisms by which these two classes of proteins act to maintain epigenetic memory and allow dynamic switches in gene expression during development.

KEY WORDS: Chromatin, Epigenetics, Polycomb group, Transcriptional memory, Trithorax group

Introduction

Just as our understanding of personal identity is rooted in our memories, cellular identity is likewise grounded in transcriptional memory. It is essential for cells to remember gene expression states in order to maintain the cell fate decisions made throughout the course of development. Seminal work in Drosophila melanogaster has shown that an exquisitely regulated cascade of transiently expressed transcription factors establishes initial gene expression patterns. Subsequently, these expression states are maintained or remembered by epigenetic cellular memory systems. Polycomb group (PcG) proteins and Trithorax group (TrxG) proteins are vital for these stable and heritable gene expression patterns (Steffen and Ringrose, 2014). PcG proteins generally maintain gene repression, whereas TrxG proteins maintain the active expression state of their target genes. Although, on the surface, this might seem straightforward, a complex interplay between these two counteracting systems is more often observed. Indeed, a number of genes must maintain the ability to respond to new signals and change their activation state, and mechanisms that allow switching between PcG and TrxG action are thus built into the system. In this way, stable gene expression states that still allow for dynamic changes can be achieved.

In recent years, a number of exciting studies have provided valuable insight into this antagonistic relationship. Given the fundamental importance of TrxG and PcG regulation during embryogenesis and growth, many of these studies have been carried out in a developmentally relevant context and potentially shed light on the mechanisms governing cell fate choices. Moreover, with the repertoire of TrxG and PcG targets continually expanding, understanding the interplay between these systems is not only informative for developmental biology but is also of interest for furthering our understanding of general gene expression regulation. Likewise, these topics converge upon hot areas of research in chromatin biology, such as the roles of long noncoding RNAs (lncRNAs) in regulating gene expression and how bivalent domains influence promoters. Additionally, with abrogation in the balance between TrxG-dependent activation and PcG-dependent repression implicated in disease states, it is important to have a better grasp on how these counteracting systems interrelate.

In this Review, we highlight the recent advances made in understanding the dynamic interplay between PcG and TrxG activities. While a brief overview of the PcG and TrxG systems will be provided, more detailed descriptions of these proteins can be found in a number of recent reviews (Di Croce and Helin, 2013; Grossniklaus and Paro, 2014; Kingston and Tamkun, 2014; Schwartz and Pirrotta, 2013; Simon and Kingston, 2013; Steffen and Ringrose, 2014). Our efforts will instead concentrate on how TrxG-dependent gene activation is achieved by overcoming PcG silencing at epigenetic switches in a variety of metazoan organisms and contexts.

An overview of PcG and TrxG proteins

PcG and TrxG proteins were initially isolated in Drosophila as factors involved in maintaining the expression patterns of HOX genes (Grossniklaus and Paro, 2014; Steffen and Ringrose, 2014), which encode transcription factors that are important determinants of patterning. The spatially restricted expression of HOX genes across metazoans underlies the specification of the anterior-posterior body axis. Early studies using homeotic transformations in flies as readouts for the aberrant expression of HOX genes enabled identification of the trans-acting factors required for mediating transcriptional memory. Indeed, the gene encoding the founding member and namesake of the PcG proteins, Polycomb, was named based on the phenotype of its heterozygous mutant, which displayed additional sex combs (Grossniklaus and Paro, 2014; Lewis, 1978). Within this framework, two classes of counteracting regulatory systems were discovered: those that were required for maintaining the active state of a gene belonged to the TrxG class, and those necessary for maintaining repression fell into the PcG class. This paradigm of HOX gene regulation has since provided the foundation for our understanding of PcG and TrxG function in metazoans (Grossniklaus and Paro, 2014; Steffen and Ringrose, 2014).

The roles of PcG and TrxG complexes have arguably been studied at the functional level most extensively in flies. Thus, for simplicity, our discussion of PcG/TrxG components will be from a Drosophila standpoint. Homologous complexes have been

¹Department of Biosystems Science and Engineering, ETH Zurich, Mattenstrasse 26, Basel 4058, Switzerland. ²Faculty of Science, University of Basel, Klingelbergstrasse 50, Basel 4056, Switzerland.

*Author for correspondence (renato.paro@bsse.ethz.ch)

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.
identified in vertebrate and mammalian systems (Table 1); however, the repertoire of factors is more complex, with a number of variant complexes resulting from apparent amplifications within the pathways (Di Croce and Helin, 2013; Schwartz and Pirrotta, 2013).

**PcG complexes**

PcG proteins form distinct complexes (Table 1) containing subunits that harbor specific biochemical activities (Di Croce and Helin, 2013; Grossniklaus and Paro, 2014; Schwartz and Pirrotta, 2013). Notably Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2) are conserved amongst metazoans and monoubiquitylate and di- and trimethylate specific lysine residues on H2A and H3, respectively (Table 1). A third grouping of PcG complexes that contains Pleiohomeotic (Pho) is also present in *Drosophila*.

PRC1 is composed of the core components Polycomb (Pc), Polyhomeotic (Ph), Posterior sex combs (Psc) and Sex combs extra (See, also known as dRing1) (Table 1) (Di Croce and Helin, 2013; Schwartz and Pirrotta, 2013). Pc can bind the H3K27me3 histone modification through its chromodomains, and this is thought to be important for anchoring the complex to chromatin. Together, Psc and See form a heterodimer, which enhances the E3 ubiquitin ligase activity of the complex. This activity is provided by the See subunit of PRC1, which monoubiquitylates H2A*K118 (K119 in mammals) specifically. This ubiquitylation event is thought to restrict RNA polymerase II (Pol II) elongation, but was also shown to recruit PRC2 members (Blackledge et al., 2014). The Ph subunit contains a sterile alpha motif (SAM) domain, which facilitates self-association. This multimerization between Ph SAM domains in turn promotes clustering that is thought to enhance the binding of PcG proteins (Isozo et al., 2013). PRC1-mediated events are also thought to compact chromatin to limit the access of activating factors and the Psc subunit in particular has been linked to this function.

By contrast, the PRC2 core complex is composed of Enhancer of zeste [E(z)], Suppressor of zeste 12 [Su(z)12], Extra sex combs (Esc) and p55 (Nur55 or Caf1) (Di Croce and Helin, 2013; Schwartz and Pirrotta, 2013). E(z) contains a SET domain, and is the PRC2 subunit responsible for the deposition of H3K27 methylation. Indeed, the SET domain, a very common protein domain housing histone methyltransferase (HMT) activity, was named in part after E(z), what extent these Pho-containing complexes functionally relate to the PcG core components are illustrated with a focus on those with characterized functions. Common synonyms for complex components are provided in parentheses.

**Table 1. PcG complex components**

<table>
<thead>
<tr>
<th>Subunit</th>
<th>Complex</th>
<th>Mammalian homologs</th>
<th>Molecular function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pc</td>
<td>PRC1</td>
<td>CBX2, CBX4, CBX6, CBX7, CBX9</td>
<td>Binds H2K27me3 via its chromodomain</td>
</tr>
<tr>
<td>Psc</td>
<td>PRC1</td>
<td>BM11 (PCGF4), MEL18 (PCGF2)</td>
<td>Zinc finger domain-containing protein that binds DNA and compacts chromatin</td>
</tr>
<tr>
<td>Ph</td>
<td>PRC1</td>
<td>PHC1 (EDR1), PHC2 (EDR2), PHC3 (EDR3)</td>
<td>Zinc finger SAM domain promotes self-association</td>
</tr>
<tr>
<td>Sce (dRing1)</td>
<td>PRC1</td>
<td>RING1A, RING1B</td>
<td>E3 ubiquitin ligase monoubiquitylates H2A*K118 (K119 in mammals)</td>
</tr>
<tr>
<td>E(z)</td>
<td>PRC2</td>
<td>EZH1, EZH2</td>
<td>Methylates H3K27 via its SET domain</td>
</tr>
<tr>
<td>Su(z)12</td>
<td>PRC2</td>
<td>SUZ12</td>
<td>Enhances E(z) activity via a VEFS-box domain</td>
</tr>
<tr>
<td>Esc</td>
<td>PRC2</td>
<td>EED</td>
<td>Promotes protein-protein interactions via WD repeats and enhances repression via binding to H3K27me3</td>
</tr>
<tr>
<td>p55</td>
<td>PRC2</td>
<td>RBAP46, RBAP48</td>
<td>Physically interacts with Su(z)12 and histones</td>
</tr>
<tr>
<td>Pho</td>
<td>PhoRC</td>
<td>YY1</td>
<td>Sequence-specific DNA binding via a zinc finger motif</td>
</tr>
<tr>
<td>Sfmbt</td>
<td>PhoRC</td>
<td></td>
<td>Binds methylated lysines in H3 and H4</td>
</tr>
</tbody>
</table>

PcG core components are illustrated with a focus on those with characterized functions. Common synonyms for complex components are provided in parentheses.
Rbbp5 and Wds. Notably, the HMT subunits of the COMPASS and COMPASS-like complexes, all of which methylate H3K4, are among those that constitute distinct subunits with non-redundant functions in the cell. Namely, Set1 is responsible for global gene activation, whereas Trithorax (Trx) and Trithorax-related (Trr) appear to target specific genes (Schuettengruber et al., 2011; Shilatifard, 2012). The composition of TrxG complexes therefore dictates their function. For example, a COMPASS-like complex characterized by Trx and the COMPASS-like protein Menin (Mnn1) targets HOX genes, whereas a complex composed of Trr and Utx targets certain hormone-responsive genes (Schuettengruber et al., 2011). The TAC1 and ASH1 complexes appear to play more specific roles in counteracting PcG silencing, and both complexes are coupled to histone acetyltransferase (HAT) activity via CREB-binding protein (CBP; Nejire – FlyBase) (Kingston and Tamkun, 2014; Kockmann et al., 2013; Schuettengruber et al., 2011). The enzymatic activities of TAC1 are provided by Trx and CBP to couple H3K4 methylation and H3K27 acetylation. Ash1, on the other hand, contains HMT activity specific for H3K36 (Schuettengruber et al., 2011). Although not touched upon in this review, a number of ATP-dependent chromatin remodeling complexes such as SWI/SNF, ISWI and various chromodomain helicase (CHD)-containing complexes recognize the histone tail modifications deposited by the aforementioned HMT and HAT complexes to facilitate active transcription (Kingston and Tamkun, 2014; Schuettengruber et al., 2011).

PcG/TrxG response elements: a simple toggle?
In Drosophila, the genomic nucleation sites of PcG- and TrxG-mediated epigenetic memory have been referred to as PcG/TrxG response elements (PRE/TREs) (Grossniklaus and Paro, 2014; Steffen and Ringrose, 2014). These regulatory DNA elements act as bistable, switchable elements that preserve the transcriptional state of their associated genes over cell generations. To date, a handful of bistable, switchable elements that preserve the transcriptional state have been identified, and when probed to either an “on” or “off” state early in development, maintain that state later in development (Steffen and Ringrose, 2014). Despite all that has been learned, a number of fundamental questions still remain.

First, what is the full repertoire of PRE/TREs in the Drosophila and vertebrate genomes? Genomic studies have expanded the repertoire of PcG and TrxG targets far beyond HOX genes and their associated PRE/TREs to include thousands of genomic loci (Steffen and Ringrose, 2014). Given that PRE/TREs have classically been defined by their ability to impart transcriptional memory, it will be important to know how many of these new PcG and TrxG binding sites truly function in this capacity. Likewise, while the PcG and TrxG proteins themselves as well as their targeting to specific genes are highly conserved, the mammalian counterpart of the PRE/TRE has been more difficult to pin down. Indeed, although some PRE/TRE-like elements have been identified, it is unclear if all of the properties of the Drosophila PRE/TRE (e.g. epigenetic memory) are maintained in the mammalian system (Grossniklaus and Paro, 2014; Kassis and Brown, 2013; Steffen and Ringrose, 2014).

A second key question that emerges concerns how PcG and TrxG proteins are recruited to these elements. We currently lack a clear understanding of the hierarchical recruitment of PcG and TrxG proteins to PRE/TREs, and elucidating these recruitment mechanisms is thus an area of active research. Given that Pho contains a DNA-binding domain and has been shown to be important for HOX gene silencing in Drosophila, it was suggested that Pho might initiate the nucleation of PcG proteins at targets (Wang et al., 2004). A recent report suggests, however, that efficient PhoRC recruitment to PREs is achieved not necessarily through Pho binding DNA, but rather via Sfnb and PRC1 (Kahn et al., 2014). Moreover, because PRC1 was thought to bind chromatin through interactions with PRC2-deposited H3K27me3, a model in which PRC2 recruits PRC1 to chromatin has emerged (Cao et al., 2002; Wang et al., 2004). However, this idea has been challenged by observations that PRC1 can bind chromatin independently of PRC2 and H3K27me3 (Tavares et al., 2012). Conversely, more recent reports even suggest that PRC1 can recruit PRC2 (Blackledge et al., 2014; Cooper et al., 2014). Finally, with a number of publications implicating PcG and TrxG proteins in binding RNA, RNA-based mechanisms of PcG and TrxG recruitment have become attractive (Broedermund, 2013; Mondal and Kanduri, 2013).

Finally, what determines the active (TRE) versus repressed (PRE) state? A simplistic model involving mutually exclusive binding of PcG or TrxG proteins to a PRE/TRE is unlikely because, as discussed in more detail below, PcG and TrxG proteins often colocalize at the same targets irrespective of the activation state of the regulated gene (Beisel et al., 2007; Enderle et al., 2011; Kockmann et al., 2013; Papp and Müller, 2006; Schwartz et al., 2010). Some evidence suggests that the repressed (PRE) state is the default and that the active (TRE) state is achieved by counteracting PcG-mediated silencing (Grossniklaus and Paro, 2014; Klymenko and Müller, 2004). By gaining a deeper understanding of the molecular interplay between PcG and TrxG proteins, we can begin to better appreciate the intricate relationship between these
complexes. Thus far, elucidating the complete picture has proved to be challenging. Nevertheless, interesting new insight has recently been gained, as illustrated below.

**Tipping the balance: the role of noncoding transcripts and lncRNAs**

In recent years, chromatin-associated lncRNAs have captivated the attention of many. As with any hot topic, this interest stems in part from the mix of surprise and skepticism surrounding the breadth of the regulatory potential of these RNAs (Palazzo and Lee, 2015; Ponting and Belgard, 2010). As the number of functionally characterized lncRNAs increases, it is evident that this expanding field will offer valuable insights into chromatin biology (Cech and Steitz, 2014; Lee, 2012; Rinn and Chang, 2012; Sabin et al., 2013). Early reports identified links between specific lncRNAs and PcG-dependent silencing, and lncRNAs have thus often been associated with target gene repression (reviewed by Brockdorff, 2013; Mondal and Kanduri, 2013). However, and as we highlight below, a number of recent reports also implicate lncRNAs in activating gene expression (Cabianca et al., 2012; Gomez et al., 2013; Grote et al., 2013; Hamazaki et al., 2015; Herzog et al., 2014; Mulvey et al., 2014; Wang et al., 2011; Yang et al., 2014).

**lncRNA-dependent gene activation via disruption of PcG silencing**

The PRE/TRE of the *Drosophila vestigial* (vg) gene provides a compelling example of a single regulatory element that bidirectionally produces lncRNAs implicated in both activation and repression (Herzog et al., 2014). An element with this characteristic was termed a gene expression alternating RNA (GEAR) box (Fig. 1A). Strand-specific transcription through this regulatory region was shown to influence the balance between activation and repression, whereby developmentally regulated expression of one strand over the other dictates the state of gene expression. For example, at the vg locus the larvally transcribed forward strand was associated with silencing, and the embryonically transcribed reverse strand was associated with activation. The repression facilitated by the forward strand correlated with long-range pairing interactions between PREs. By contrast, the mechanism by which the reverse strand promoted activation appeared to involve inhibitory interactions between the lncRNA and PRC2. More specifically, it was shown that the lncRNA not only inhibits the HMT activity of the E(z) component of PRC2 but also locally displaces the complex from the vg locus (Fig. 1A).

Interestingly, Herzog et al. (2014) also observed that both the forward and reverse lncRNAs bind to and inhibit the activity of PRC2 in vitro but only the reverse strand binds PRC2 in vivo. This curious finding implies that additional factors influence specific lncRNA-PRC2 binding interactions in the cellular context.

Recent observations in mammalian systems open up the possibility that this RNA binding might be a conserved aspect of PRC2 regulation. Indeed, a mammalian homolog of E(z), EZH2, can promiscuously bind RNA, and this RNA binding ability was shown to inhibit EZH2 activity (Cifuentes-Rojas et al., 2014; Kaneko et al., 2014, 2013). It has also been proposed that EED (the mammalian
homolog of Esc) and JARID2 (a conserved PRC2-interacting factor) influence the RNA binding affinity of EZH2 (Cifuentes-Rojas et al., 2014), a finding that might provide clues as to the factors that contribute to RNA specificity at the vg locus in flies. Because JARID2 reduces EZH2 RNA affinity, the composition of the complex bound to the reverse strand could perhaps allow for stronger associations with the IncRNA in flies. Building upon what they learned at the vg GEAR box, Herzog et al. used bioinformatic analyses to support for several hundred candidate GEAR box elements in both the Drosophila and mouse genomes (Herzog et al., 2014). In fact, they found evidence for a GEAR box element at the homologous mouse vg locus, Vgl2, suggesting that GEAR box elements might be a conserved feature of many PRE/TREs. The proposed functional role of the additional GEAR boxes, however, remains to be experimentally validated.

Additionally, while the above-mentioned functional characterization of the vg GEAR box clearly implicates IncRNA regulation in activation, proper mRNA expression at the vg locus in flies also requires the TrxG system. That is to say, mutations in Trx result in abrogated activation (Pérez et al., 2011). It is possible that Trx is required for active transcription of the reverse strand IncRNA, although further studies are required to fully understand how Trx plays into this scenario.

**IncRNA-dependent gene activation through physical association with TrxG components**

While IncRNA interactions with PRC2 have been a prominent topic of investigation, it is becoming clearer that TrxG components can also physically associate with IncRNAs to regulate gene expression (Fig. 1B) (Cabianca et al., 2012; Gomez et al., 2013; Wang et al., 2011; Yang et al., 2014). One of the first and perhaps best-described examples of this comes from studies by the Chang lab looking at the vertebrate HOXA locus, which produces the long intergenic noncoding RNA (lincRNA) HOTTIP (Wang et al., 2011; Yang et al., 2014). They showed that HOTTIP is expressed from the 5′ edge of the HOXA locus and inferred that HOTTIP might be conserved because a similar lincRNA could be found in human, mouse and chick. The pattern of tissue expression of HOTTIP was consistent with its genomic location, as HOX gene expression patterns correlate with their chromosomal position. In line with a role in promoting active transcription, HOTTIP was demonstrated to bind the TrxG protein WDR5 (Fig. 1B). The siRNA-mediated knockdown of HOTTIP revealed a distance-dependent relationship of the lincRNA with active transcription, with those genes nearest to the HOTTIP locus being most strongly affected. Furthermore, because HOTTIP knockdown resulted in reduced recruitment of the TrxG complex components MII1 (KMT2A) and WDR5 as well as reduced H3K4 trimethylation at target genes, the lincRNA was deduced to be important for the recruitment of TrxG complexes (Fig. 1B).

In a follow-up study, Yang et al. characterized the RNA-binding surface of WDR5 (Yang et al., 2014). To assay the importance of RNA binding, the authors identified a mutation that abrogated RNA binding yet retained MLL catalytic activity. When this RNA-binding-deficient form of WDR5 was introduced into mouse embryonic stem cells (ESCs), defects in WDR5 chromatin localization and global H3K4 trimethylation were seen. Additionally, the authors noticed a nuclear-specific decrease in WDR5 protein stability, which suggested that RNA binding might stabilize the protein in the nucleus. To get a better picture of the full landscape of WDR5-RNA binding, WDR5-associated RNAs were sequenced. Roughly 1400 RNAs were found to be associated with wild-type WDR5. Although a number of important IncRNAs implicated in ESC pluripotency and differentiation were identified, the RNA repertoire was not restricted to IncRNAs and also included mRNAs, primary microRNAs (pri-miRNAs) and small nuclear RNAs (snRNAs). Underscoring the functional importance of RNA binding activity, it was further shown that mouse ESCs expressing a RNA-binding-deficient WDR5 were compromised for maintaining their stem cell fate. Another example of a physiologically relevant WDR5-IncRNA interaction came from a study looking into the role of the IncRNA NeST (Gomez et al., 2013). In mouse CD8+ T cells, this IncRNA was suggested to bind WDR5 and upregulate interferon-γ. Moreover, NeST appeared to influence immunological phenotypes such as viral and bacterial infection susceptibility.

In addition to WDR5, another TrxG protein, Ash1l, has been implicated in physical interactions with IncRNAs. Specifically, it was shown that a chromatin-associated IncRNA produced from primate-specific D4Z4 repeats termed DBE-T (for D4Z4 binding element-transcript) binds to Ash1l (Fig. 1B) (Cabianca et al., 2012). The D4Z4 repeats map to 4q35 and have been associated with facioscapulohumeral muscular dystrophy (FSHD), in which repeat copy number reduction results in the loss of silencing of 4q35 genes. The D4Z4 repeat functions similarly to a PRE in that it recruits PcG proteins. To examine DBE-T-associated gene activation, Cabianca et al. (2012) used primary samples from FSHD patients and healthy individuals as well as an engineered cell culture system. In this cell culture model, human chromosome 4 was introduced into Chinese Hamster Ovary (CHO) cells to generate monochromosomal hybrid cells (chr4/CHO). Using this approach, it was found that reduced repeat copy numbers in FSHD samples correlated with DBE-T expression and the derepression of silenced genes. Additionally, RNAi-mediated loss of PcG complexes or pharmacological loss of silencing resulted in DBE-T expression under conditions in which repeat number was within the healthy range. Consistent with derepression, Ash1l was enriched within the region under conditions in which the IncRNA was expressed. The authors went on to show that Ash1l binds to DBE-T in chr4/CHO cells and that, in vitro, the Ash1l SET domain directly binds DBE-T. Importantly, DBE-T knockdown resulted in impaired Ash1l recruitment and decreased H3K36 dimethylation. Altogether, these results imply that expression of DBE-T recruits Ash1l to facilitate H3K36 methylation and subsequent 4q35 gene transcription (Fig. 1B).

Although derepression of 4q35 genes via DBE-T expression is associated with disease, at least one of the genes silenced by the D4Z4 repeats, DUX4, appears to have a normal role in development. It is therefore possible that the DBE-T IncRNA might play a role during normal development to promote 4q35 gene expression.

Other developmentally important IncRNAs have also been linked to TrxG proteins though physical interactions. For instance, in mice the Fendrr IncRNA, which is required for normal heart and body wall development, has been suggested to bind both PRC2 and MLL complexes (Grote et al., 2013). Likewise, in the context of Drosophila sex determination, the R2 antisense IncRNA has been implicated in promoting transcription of the Sex lethal gene via binding interactions with TrxG proteins (Mulvey et al., 2014). Furthermore, the generation of protein-RNA prediction algorithms such as catRAPID has enabled in silico analyses to support potential TrxG protein-IncRNA interaction pairs (Bellucci et al., 2011; Iglesias-Platas et al., 2012).

**Histone modifications and histone-modifying factors that counteract repression**

Post-translational modifications on histone tails can both positively and negatively influence the expression output of associated genes.
Importantly, early studies showed that TrxG-deposited histone modifications can inhibit PRC2 activity and can, hence, counteract gene repression (Klymenko and Müller, 2004; Papp and Müller, 2006). The molecular mechanisms behind this TrxG-dependent antagonism of PcG silencing are now starting to take shape (Schmitges et al., 2011; Yuan et al., 2011). The TrxG component Ash1 (mammalian homolog ASH1L) has emerged as a particularly vital player, and reports from the last half of the decade provide compelling evidence that the anti-repressive activity of Ash1 resides in its ability to catalyze H3K36 methylation (An et al., 2011; Dorighi and Tamkun, 2013; Miyazaki et al., 2013; Yuan et al., 2011). Additionally, recent evidence physically and functionally links Trx with CBP, suggesting that the histone-modifying activities of TrxG proteins might be coupled to counteract PcG silencing (Tie et al., 2014).

Active H3K36 and H3K4 methylation marks antagonize PRC2

Two reports in 2011 contributed greatly to our understanding of how TrxG-deposited histone marks associated with active gene expression might counteract PRC2-mediated silencing (Schmitges et al., 2011; Yuan et al., 2011). These reports provided both in vivo as well as in vitro biochemical evidence for a mutually exclusive relationship between active and repressive marks on histone H3 tails. In essence, they demonstrated that the post-translational modifications left by TrxG proteins could preclude the activity of PcG complexes.

Specifically, Yuan et al. (2011) investigated the relationship between H3K27 and H3K36 methylation. They demonstrated that active H3K36me2 and me3 marks rarely co-immunoprecipitated with repressive H3K27me in purified mononucleosomes. Mass spectrometry revealed that only a very small percentage of histone H3 tails were di- or trimethylated at H3K27 while at the same time being di- or trimethylated at H3K36 in HeLa cells. Taking this further, the authors showed that PRC2 activity in HMT assays was inhibited if the substrates were pre-methylated at the H3K36 position. Together, these findings suggested that not only do activation-associated Ash1-deposited H3K36 methylation marks rarely coexist on the same histone tail with PRC2-deposited H3K27 methylation, but also that they are likely to inhibit PRC2 activity (Fig. 2A).

Further substantiation for active TrxG-deposited histone marks as impediments to PcG silencing came from structural and biochemical studies of the PRC2 component p55 (Nurf55) (Schmitges et al., 2011). In this study, which looked at both H3K36 methylation and H3K4 methylation, the authors saw that p55 binds the very N-terminus of H3. Importantly, K4 was among the residues found to be especially important for the molecular recognition of the H3 tail. Moreover, it was shown that if H3 was methylated (mono, di or tri) on K4, the binding of a p55-Su(z)12 complex was significantly reduced, as measured by fluorescence polarization and isothermal titration calorimetry (Fig. 2B). Similar to Yuan et al. (2011), Schmitges et al. (2011) found that H3K36me2 and me3 inhibited PRC2 activity. Furthermore, the activity of human, mouse and Drosophila PRC2 was inhibited by H3K4 and H3K36 methylation, suggesting that this is a conserved mechanism for inhibiting PcG-mediated silencing. Interestingly, when surveying the inhibition of PRC2 complexes in Arabidopsis, which possesses three different Su(z)12 homologs in contrast to the single Su(z)12 found in mammals and flies, the authors found that Su(z)12 homologs had different levels of sensitivity to active marks; for example, EMBRYONIC FLOWER 2 (EMF2) was sensitive, whereas VERNALIZATION2 (VRN2) was insensitive to the presence of H3K4me3 in HMT assays. This suggests that, in plants, PRC2 inhibition by active marks can be modulated by swapping out the Su(z)12 homolog.

How exactly TrxG-mediated H3K36 methylation inhibits PRC2 remains unclear. From a structural perspective, the inhibitory effect of K4 methylation makes sense because this residue is a clear point of contact. But K36 is located outside the p55-Su(z)12 binding site (Schmitges et al., 2011) so how does it inhibit PRC2? Although the exact inhibitory effect of H3K36 methylation on PRC2 biochemical activity remains to be determined, connections between this histone modification and Ash1, which has long been described as an ‘anti-repressor’, bolster its role in counteracting silencing (Dorighi and Tamkun, 2013; Klymenko and Müller, 2004; Kockmann et al., 2013; Miyazaki et al., 2013; Papp and Müller, 2006; Tanaka et al., 2007; Yuan et al., 2011).

TrxG proteins counteract PcG silencing by promoting H3K36 and H3K4 methylation

The aforementioned biochemical dissections illustrate a role for active methylation marks as physical impediments to PcG activity. It was suspected that the TrxG protein Ash1 might supply the enzymatic activity that deposits H3K36 dimethylation (Yuan et al., 2011) because the Ash1 SET domain closely resembles other bona fide H3K36 methyltransferases and, like other H3K36 methyltransferases, Ash1 preferred nucleosomal H3 over free H3 in HMT assays. Consistent with a role for Ash1 in methylating H3K36, alanine substitution at H3K36 on nucleosomal substrates resulted in loss of methylation by Ash1 in vitro. Likewise, methyllysine analogs of the di- and trimethylated state of H3K36 also blocked Ash1 activity. Over the years, however, conflicting reports on the specificity of Ash1 for different histone tail lysines have created controversy (An et al., 2011; Byrd and Shearn, 2003; Gregory et al., 2007; Tanaka et al., 2007; Yuan et al., 2011). The factors that account for the in vitro and in vivo differences in enzyme specificity for H3K4, H3K9, H3K36 and H4K20 residues are unclear. Regardless, mounting evidence generates a compelling argument that H3K36 methylation could play a prominent role in the mechanism by which Ash1 antagonizes PcG silencing (Fig. 2A).

A recent study in Drosophila that aimed to better characterize the TrxG protein Kismet (Kis) and how it counteracts PcG silencing has provided further insights into Ash1 function (Dorighi and Tamkun, 2013). Kis has been implicated in gene activation, both as a general transcription elongation factor and as a factor involved in counteracting PcG silencing, and Dorighi and Tamkun (2013) sought to determine if these two functions were connected. They found that pharmacological inhibition of transcription elongation had no negative effect on Kis functions that related to antagonizing PcG activity. Explicitly, Ash1 recruitment was unchanged and there was no increase in PcG-deposited H3K27 methylation upon blocking elongation. These findings are in line with another recent mouse ESC-based study, which similarly concluded that Ash1-dependent methylation of H3K36 and antagonism of PcG silencing occur independently of transcription elongation (Miyazaki et al., 2013). Having ruled out a link between the role of Kis in promoting transcription elongation and its involvement in Ash1 function, Dorighi and Tamkun (2013) went on to investigate how Kis might modulate Ash1 function. It had previously been shown, using loss-of-function mutants for Kis, that Kis was essential for the proper recruitment of Trx and Ash1 to chromatin (Srinivasan et al., 2008). Although this mutant background exhibited elevated levels of PRC2-deposited H3K27 methylation, PRC2 chromatin
association remained unchanged and prompted the idea that Kis might antagonize PRC2 activity rather than recruitment. Interestingly, *kis* mutant larvae display reduced levels of H3K36 methylation, suggesting that Kis might antagonize PcG activity by promoting H3K36 methylation (Dorighi and Tamkun, 2013). Given that Kis is important for Ash1 chromatin association, and the fact that Ash1 can catalyze H3K36 methylation, this paints a picture in which Kis promotes Ash1 recruitment and subsequent H3K36 methylation (Fig. 2A) (Dorighi and Tamkun, 2013; Miyazaki et al., 2013; Srinivasan et al., 2008; Tanaka et al., 2007; Yuan et al., 2011). Consistent with this, Dorighi and Tamkun (2013) saw that mutations in Ash1 resulted in a loss of H3K36 dimethylation and increased H3K27me3 on polytene chromosomes. Thus, Ash1-deposited H3K36 methylation most likely inhibits PRC2-catalyzed H3K27 methylation. An interesting line of future inquiry will be to figure out how exactly Kis promotes Ash1 recruitment. Because *in vitro* and *in vivo* physical interactions between Ash1 and Kis have been elusive to capture, the influence that Kis exerts over Ash1 recruitment is unlikely to involve direct physical associations (Dorighi and Tamkun, 2013; Srinivasan et al., 2008).

These findings, which imply that Ash1 plays a key role in counteracting PcG silencing, are in line with a previous report suggesting that Ash1 blocks repressive PcG-deposited H3K27, H3K9 and H4K20 trimethylation in the ‘on’ state in flies (Papp and Müller, 2006). Back in 2006, Papp and Müller saw that, in developing *Drosophila* tissues, both PcG and TrxG complexes were present at the *Ultrabithorax* (*Ubx*) PRE regardless of the expression state of the *Ubx* gene (Papp and Müller, 2006). Ash1, however, was selectively bound downstream of the transcriptional start site only in the ‘on’ state and correlated with a lack of repressive modifications in the promoter and gene body. Consistent with the findings from Dorighi and Tamkun (2013), mutations in *ash1* resulted in aberrant repressive H3K27 methylation in this earlier study. At the time, the authors surveyed H3K4 methylation rather than H3K36 methylation and, intriguingly, they saw that H3K4 methylation was decreased in the *ash1* mutant background. Although this could suggest that Ash1 also methylates H3K4 *in vivo,* this observation might also imply a relationship between H3K4 and H3K36, such that a level of cross-talk exists whereby H3K36 and H3K4 reinforce each other and loss of Ash1-dependent H3K36 methylation also somehow leads to loss of H3K4 methylation (Fig. 2C). However, it was observed that Ash1 must have a much more general role in functioning as a global co-activator of transcription (Kockmann et al., 2013); acting together with another TrxG protein, Female sterile (1) homeotic [*Fs(1)h*], Ash1 was found at many active gene promoters. Hence, in this case the PcG-mediated silencing would need to overcome the general transcription-promoting function of Ash1 and *Fs(1)h* in order to silence genes.

H3K4 methylation was the focus of a recent paper linking CBP and the HMTs Trx and Trr in *Drosophila* (Tie et al., 2014). In this study, it was found that Trx and Trr were monomethyltransferases, and the authors identified specific tyrosine residues that limited additional methylation *in vitro.* Further supporting roles for Trx and Trr as a monomethyltransferase *in vivo* could be seen with inactivating
mutations, trz211 and trr3, where only H3K4me1 and not H3K4me3 levels were reduced. The authors went on to show that Trx and Trt physically associate with CBP. Consistent with previously described functions for CBP in counteracting PcG silencing via H3K27 acetylation, they saw that the perturbations in PcG silencing upon overexpression of trz were dependent on CBP (Fig. 2D). Furthermore, the genome-wide chromatin localization patterns of Trx, Trt, CBP and H3K4me1 largely overlapped. Biochemical evidence provided by in vitro acetylation assays suggested that monomethylation of H3K4 positively influenced CBP-mediated H3K27 acetylation. Pulling it altogether, Trx-catalyzed H3K4 methylation inhibits PRC2 binding (Schmittges et al., 2011) while at the same time through physical and functional associations with CBP promotes H3K27 acetylation to block repressive H3K27 methylation (Tie et al., 2014).

**TrxG and PcG co-occupation: the front lines of the battle**

In the above examples, TrxG-associated activities act to inhibit PRC2 binding, activity, or both at target loci. However, it is unlikely that the above-mentioned mechanisms converge into a single overarching paradigm such that TrxG-associated activities at every locus ultimately limit the recruitment of PcG proteins. Indeed, PRC2-deposited H3K27me3 is canonically thought to recruit PRC1 and, with clear instances of Trx-mediated PRC2 inhibition, one would expect very little co-occupancy of TrxG proteins and PRC1 components. Contrary to this expectation, a number of reports show that Trx, PRC2 and PRC1 can occupy overlapping sets of genomic loci (Beisel et al., 2007; Enderle et al., 2011; Kockmann et al., 2013; Papp and Müller, 2006; Schwartz et al., 2010). Thus, the mechanisms employed by TrxG proteins to counteract PcG-mediated silencing are likely to be complex. Bivalent domains, which are unique and somewhat puzzling features in which both active H3K4me3 and silent H3K27me3 modifications mark a subset of promoters (Voigt et al., 2013), also represent curious conditions indicative of the simultaneous action of both PcG and TrxG proteins in mammalian cells. Although these domains remain controversial, evidence suggests that they do genuinely occur at certain frequencies. With H3K4 and H3K27 methylation being catalyzed by TrxG and PcG proteins, respectively, bivalent domains have the potential to tell us a lot about the interplay between these two counteracting systems.

**PcG and TrxG proteins colocalize on chromatin**

As mentioned above, we have known for nearly a decade that both activation- (TrxG) and silencing- (PhoRC, PRC1, and PRC2) associated complexes appear to be assembled in both the ‘on’ and ‘off’ states at the Drosophila Ubx PRE (Fig. 2E) (Papp and Müller, 2006). However, genomic approaches in a number of Drosophila cell lines suggest that this phenotype is widespread (Beisel et al., 2007; Enderle et al., 2011; Schwartz et al., 2010).

For example, in the embryonically derived Kc and SF4 Drosophila cell lines, exhaustive chromatin immunoprecipitation (ChIP) followed by DNA tiling array (ChIP-chip) analyses of the ANT-C and BX-C loci, which contain several HOX gene clusters, demonstrated extensive overlap of PcG and TrxG proteins (Beisel et al., 2007). Here, the analysis focused on Trx and Pho as well as the PRC1 components Pc, Ph and Psc. Notably, at virtually all PRC1 and Pho binding sites, Trx could also be detected irrespective of gene expression state. Active and repressed genes, however, displayed marked differences in Pho distribution such that repressed genes had discrete promoter-associated Pho peaks whereas the Pho signal was spread across the gene bodies of active genes. Prominent Pho signal could likewise be detected at the heat shock puffs on polytene chromosomes. Polytene chromosomes are, for example, found in Drosophila salivary glands and are formed by repeated rounds of DNA replication in the absence of cell division. This natural amplification allows for low-resolution imaging of the entire genome, and highly transcribed regions (such as heat shock loci under induction) form recognizable puffed structures. Importantly, pho- mutant larvae took longer to re-silence upon recovery from heat shock stress, suggesting a role for Pho in re-silencing. Given that Pho is important for silencing, this association with actively transcribed gene bodies was intriguing. Furthermore, the difference in the distribution of Pho when comparing active and repressed states could be telling us something about switches in gene expression states.

Recently, Pho was demonstrated to genetically and physically interact with Drosophila Spt5 (Harvey et al., 2013), which is a factor that is essential for promoter-proximal pausing and that, upon phosphorylation by P-TEFb (Cdk9/Cyclin T1), is converted into an elongation-promoting factor. Additionally, Pho and Spt5 binding sites substantially overlapped genome-wide in Drosophila S2 cells. Such an interaction provides a potential means for Pho to spread beyond its recognition sequence to associate with broad regions of active transcription. Pho is not the only PcG protein that has been associated with paused genes. Indeed, PRC1 preferentially targets paused promoters (Enderle et al., 2011). Likewise, PcG repression has been suggested to convey tissue-specific control over poised Pol II in the context of Drosophila development (Gaertner et al., 2012).

The extensive colocalization of TrxG and PcG proteins seen by the above ChIP-chip analysis at the ANT-C and BX-C loci was expanded by larger ChIP-chip and ChIP-seq analyses in BG3, D23, Sg4 and S2 Drosophila tissue culture cells (Enderle et al., 2011; Schwartz et al., 2010). Consistently, all of these studies report pervasive overlaps in the targeting of TrxG and PcG complexes.

**Bivalent promoters: inherently balanced between TrxG and PcG activities**

Although the above examples furnish evidence that PcG and TrxG proteins can co-occupy overlapping sets of genomic locations, it should be noted that the experimental approaches have been based on ChIP analyses of ensembles of cells that, due to technical limitations, cannot definitively determine that both complexes are simultaneously present in the same cell, at the same genomic locus. Such arguments have also been applied to question the existence of bivalent domains at which the histone marks deposited by both TrxG and PcG complexes are found. In the case of bivalent promoter domains, efforts to analyze more homogenous cell populations, as well as sequential ChIP approaches, argue that both PcG-deposited H3K27me3 and TrxG-deposited H3K4me3 can coexist (Fig. 2F) (Voigt et al., 2013). Initially discovered in ESCs, these domains have since been detected in non-stem cell populations. In addition, new evidence has identified Mll2 (Kmt2b/d) as the TrxG-associated H3K4me3 methyltransferase dedicated to bivalent promoters (Fig. 2B) (Hu et al., 2013). Curiously, however, Mll2 knockdown had no influence on ESC self-renewal nor did it affect the rapid induction of bivalent genes in response to retinoic acid stimulus (Denissov et al., 2014; Hu et al., 2013). Given that bivalency has been proposed to poise developmental genes for rapid activation (Voigt et al., 2013), this observation was unexpected and calls into question the function of bivalency. Independent evidence for Mll2 being the bivalent domain H3K4 methyltransferase came in 2014 (Denissov et al., 2014), but, in this study, the authors further examined the interplay between the various Set1/Trithorax-type H3K4 methyltransferases.
Like Hu et al. (2013), they determined that Mll2 is required for H3K4me3 at bivalent promoters. Mll2 could be found at both bivalent as well as active promoters, but active gene expression did not require Mll2. Whereas Mll2 was bound at both bivalent and active promoters, Cxxc1, a Set1 complex subunit, was found only at active gene promoters. Consistent with the findings from Hu et al. (2013), Denissov et al. (2014) further showed that the induction of genes with bivalent promoters was not influenced by abrogation of
the bivalent state upon loss of Mll2. The authors instead propose that Mll2 might function as a ‘pioneer trimethyltransferase’ that marks promoters with H3K4me3 and that PcG proteins might keep bivalent genes from being activated by other HMTs such as the Set1 complex (Denissov et al., 2014).

**TrxG and PcG activities out of balance in disease**

Underscoring the importance of their balanced activities, misregulation that disrupts the balance between PcG and TrxG proteins has been linked to disease states such as cancer, neuronal damage and muscular dystrophy (Fig. 3). For example, a hallmark of leukemia is a block in differentiation, and abnormalities in the balance between PcG silencing and TrxG activation – specifically between EZH2 and MLL (Thiel et al., 2013) – could underlie these differentiation defects (Fig. 3A). The literature linking PcG and TrxG proteins to cancer, however, is vast and complicated with perhaps more questions raised than answered (see reviews by Grossniklaus and Paro, 2014; Kingston and Tamkun, 2014; Koppens and van Lohuizen, 2015; Sauvageau and Sauvageau, 2010; Sceifo et al., 2015). Indeed, lending support to many PcG proteins serving as oncogenes, overexpression of the genes encoding PcG proteins is widespread in a multitude of cancer types (Grossniklaus and Paro, 2014; Koppens and van Lohuizen, 2015). Nonetheless, the exact same factors in many instances have been implicated as tumor suppressors in a different context. EZH2 provides a compelling example of this, in which activating mutations result in lymphomas and overexpression is seen in a number of solid tumors yet inactivating mutations promote myeloid malignancies (Hock, 2012; Koppens and van Lohuizen, 2015). Perhaps context-dependent considerations, such as the exact cohort of misregulated targets within the cell and the balance between PcG and TrxG regulation, influence whether these chromatin regulators function as tumor suppressors versus oncogenes (Koppens and van Lohuizen, 2015).

Aberrations in the balance between TrxG and PcG activities are likely to extend beyond the context of cancer. For instance, the damage induced during neuronal necrosis might result from such disruptions (Fig. 3B). In the context of stroke and brain trauma, calcium overload induces neuronal necrosis and, in *Drosophila* and rodent models, it was recently seen that this necrosis is regulated by a series of chromatin-dependent events (Liu et al., 2014). Explicitly, a cascade of events is initiated by phosphorylation of histone H3 serine 28 (H3S28ph) by JIL-1 (the Drosophila homolog of MSK1/2). H3S28ph leads to eviction of PRC1 and Trx activation (Fig. 3B). Importantly, inhibition of these events alleviated necrosis *in vitro* and *in vivo* (Liu et al., 2014).

Also, as illustrated above in our discussion of *DBE-T*-facilitated Ash1 recruitment, aberrant derepression of PcG-silenced genes in the context of FSHD is likely to occur via a disruption in the balance between PcG and TrxG activities (Cabianca et al., 2012). Here, low repeat copy numbers result in the expression of a lncRNA, which in turn recruits Ash1 to illicit activation of 4q35 genes (Fig. 3C). Perhaps other repeat-associated diseases might likewise tip the balance between PcG and TrxG activities towards pathogenic states.

**Perspectives**

Fine-tuning the balance between active and repressed chromatin is one of the most crucial tasks of a cell. However, how can constitutively expressed chromatin factors that are required to maintain differential expression patterns, and that apparently coexist at most regulatory sites, generate such opposing activities? Most of our knowledge on how PcG and TrxG proteins control target gene expression is based on ChIP analyses of bulk material. Hence, short-term dynamic changes will be lost and cell-to-cell variations averaged out. Historically, the regulation of HOX genes was considered the role model with which to study mechanisms of transcriptional memory. But, more recent studies have highlighted that the clientele of PcG and TrxG proteins is much broader and includes target genes with highly fluctuating expression profiles, suggesting that the control these complexes impose over gene expression can be far from static. Even at the single gene level, the dynamics of gene expression, whether continuous or in bursts, need to be considered (Hebenstreit, 2013) if the two opposing activities are to be better understood. Hence, future work will need to assess the balance between PcG silencing and TrxG activation in a much more time-resolved manner at regulatory sites. This entails the development of more sensitive ChIP methods (Lara-Astiaso et al., 2014), the establishment of *in vivo* observation methods and eventually the *in vitro* reconstitution of chromatin-controlled transcriptional processes and even of the process of epigenetic memory. Given the fundamental roles that PcG and TrxG proteins play in epigenetic gene control, together with the dire consequences observed in many diseases if a fine balance between these proteins is not maintained, these additional research challenges will certainly be highly rewarding.

**Acknowledgements**

We thank Federico Comoglio and Alwyn Pereira for suggestions on the manuscript.

**Competing interests**

The authors declare no competing or financial interests.

**Funding**

S.J.G. is supported by an ETH Fellowship and the work of R.P. is supported by grants from the Swiss National Science Foundation, Epigenesys and the ETH Zurich.

**References**


