The development of multicellular organisms relies on cells and tissues establishing unique gene expression programmes. To achieve this, signalling pathways converge on DNA-binding transcription factors, which guide the binding and activity of RNA polymerase II (Pol II) at gene promoters. However, through studying these mechanisms, it has become apparent that chromatin and histone post-translational modification can also profoundly influence transcription and gene expression. The importance of chromatin-modifying complexes in controlling development is exemplified by the Polycomb repressive system. Polycomb genes were initially discovered in *Drosophila melanogaster*, where they are required for repression of homeotic (Hox) genes and thus for body plan specification. Polycomb genes were subsequently shown to have mammalian orthologues that have important roles in controlling gene expression throughout development.

The importance of chromatin-modifying complexes in controlling development is exemplified by the Polycomb repressive system. Polycomb genes were initially discovered in *Drosophila melanogaster*, where they are required for repression of homeotic (Hox) genes and thus for body plan specification. Polycomb genes were subsequently shown to have mammalian orthologues that have important roles in controlling gene expression throughout development.

Since these initial discoveries, biochemical analyses have shown that Polycomb proteins typically assemble into one of two large multiprotein complexes that post-translationally modify histones. Polycomb repressive complex 1 (PRC1) monoubiquitylates histone H2A at Lys119 (H2AK119ub1), whereas PRC2 monomethylates, dimethylates and trimethylates histone H3 at Lys27 (H3K27me1, H3K27me2 and H3K27me3, respectively). Importantly, PRC1 and PRC2 tend to spatially converge on the same sites in the genome to form Polycomb chromatin domains, which are uniquely enriched in H2AK119ub1, H3K27me3 and Polycomb complexes. Polycomb chromatin domains are then thought to counteract transcription, although the mechanisms that enable this remain incompletely defined.

Although early genetic and biochemical investigations of Polycomb proteins were focused on *D. melanogaster*, detailed studies over the past decade have revealed that the Polycomb system has expanded and diverged in mammals. Detailed studies are now uncovering the molecular mechanisms that enable mammalian PRC1 and PRC2 to identify their target sites in the genome, communicate through feedback mechanisms to create Polycomb chromatin domains and control transcription to regulate gene expression. In this Review, we discuss and contextualize the emerging principles that define how this fascinating chromatin-based system regulates gene expression in mammals.
and describe their enzymatic activities. We examine how Polycomb repressive complexes identify their target sites in the genome and communicate with each other to initiate the formation of Polycomb chromatin domains. We then discuss the regulation of Polycomb chromatin domain formation and ultimately how this is integrated with and regulates gene expression. Finally, we conclude by considering future avenues of research that will provide further mechanistic understanding of the principles of Polycomb biology in health and disease.

**The wide repertoire of mammalian Polycomb repressive complexes**

The rapid progress in understanding mammalian Polycomb biology has relied on detailed and systematic biochemical interrogation of Polycomb repressive complexes and their histone-modifying activities. This work has shown that Polycomb complexes comprise catalytic cores, which bind auxiliary proteins to create distinct PRC1 and PRC2 assemblies. In this section we introduce the composition, structure and enzymatic activity of the different Polycomb complexes as a prerequisite for examining how these complexes underpin the functions of mammalian Polycomb complexes in gene regulation.

**PRC1 and its E3 ubiquitin ligase activity.** The catalytic core of PRC1 is composed of RING1B or its paralogue RING1A, and one of six Polycomb group RING finger (PCGF) proteins (PCGF1, PCGF2, PCGF3, PCGF4, PCGF5 or PCGF6; Fig. 1a). RING1 and PCGF proteins have similar domain architecture, including an amino-terminal (N-terminal) RING domain and a carboxy-terminal RING finger and WD40-associated ubiquitin-like (RAWUL) domain. RING1 and the PCGF protein dimerize through their RING domains, which facilitates their interaction with an E2 conjugating enzyme to enable histone ubiquitylation. Their RAWUL domains bind to a range of auxiliary subunits, which regulate the catalytic activity of PRC1 and target it to specific sites in the genome. Importantly, the PCGF components dictate which auxiliary subunits are incorporated into specific PRC1 complexes. These complexes are generally categorized as either canonical PRC1 (cPRC1) or variant PRC1 (vPRC1; Fig. 1a), a nomenclature that reflects the order in which these complexes were initially identified. cPRC1 complexes, which were the first to be isolated, assemble around either PCGF2 or PCGF4, and include one of five chromodomain-containing paralogs (CBX2, CBX4, CBX6, CBX7 or CBX8) and a Polyhomeotic (PHC) subunit (PHC1, PHC2 or PHC3). By contrast, vPRC1 complexes, which were identified later, can assemble around any of the six PCGF proteins (PCGF1–PCGF6), and include RING1 and YY1-binding protein (RYBP), or its paralogue YAF2, and various additional subunits depending on the PCGF component present in the complex.

A key function of PRC1 complexes is to monoubiquitylate histone H2A. Given that E3 ubiquitin ligases are notoriously promiscuous, and that histones have a high density of potential Lys acceptor residues, it was initially puzzling how mammalian PRC1 could preferentially target Lys119 of H2A. The solution to this problem was provided by a landmark study detailing the atomic structure of a RING1B–PCGF4 RING domain dimer bound to an E2 ubiquitin conjugating enzyme and the nucleosome. This structure revealed extensive contacts between PRC1 and the nucleosome acidic patch, which is a binding surface used by many nucleosome-interacting factors. Furthermore, the E2 enzyme contacts the N terminus of RING1B and DNA at the nucleosome dyad, thereby orienting itself to transfer ubiquitin to H2AK119 and explaining the substrate specificity of PRC1.

Although all six mammalian PCGF proteins have highly similar RING domains, the ubiquitin ligase activity of the PRC1 complexes they form differ considerably. Notably, vPRC1 complexes are much more active on nucleosome substrates than their cPRC1 counterparts. Although these catalytic differences stem in part from the PCGF component directly influencing the activity of the core complex, incorporation of the vPRC1-specific auxiliary subunit RYBP also dramatically stimulates E3 ligase activity. By contrast, incorporation of CBX proteins into cPRC1 complexes has a far less pronounced stimulatory effect, and cPRC1 complexes are thus far less active than vPRC1 complexes.

**PRC2 and its methyltransferase activity.** PRC2 complexes form around a tetrameric core complex consisting of Enhancer of Zeste 2 (EZH2) or its paralogue EZH1, embryonic ectoderm development (EED), Suppressors of Zeste 12 (SUZ12), and retinoblastoma-binding protein 4 (RBBP4) or RBBP7 (Fig. 1b). Biochemical and structural analyses of PRC2 have largely focused on EZH2-containing complexes, which contribute more prominently to H3K27 methylation in most contexts and are essential for early development. Cryo-electron microscopy structures have revealed that PRC2 forms two distinct lobes: a catalytic lobe and a targeting and regulatory lobe, which are bridged by SUZ12 (REFS 14–16). Within the catalytic lobe, EED and the VEFS (VRN2–EMF2–FIS2–SUZ12) domain of SUZ12 interact with EZH1 or EZH2, causing its SET (Su(var)3-9, Enhancer of Zeste and Trithorax) domain to transition from an autoinhibitory conformation into an active one that supports methyltransferase activity.

Also important for the catalytic activity of PRC2 is the CXC domain of EZH2, which engages nucleosomal DNA, and a patch of residues on the surface of EZH2, which interacts with the N-terminal portion of the H3 tail. Together, these interactions position the N-terminal tail of H3 such that Lys27 is adjacent to the active site of PRC2, to enable its methylation.

The targeting and regulatory lobe of PRC2 includes the N-terminal portion of SUZ12, which binds to RBBP4 or RBBP7 and interacts with various additional auxiliary factors in a mutually exclusive manner, giving rise to PRC2 assemblies with distinct biochemical properties (Fig. 1b). PRC2.1 complexes contain a Polycomb-like (PCL) protein (PCL1, PCL2 or PCL3) and PRC2-associated LCOR isoform 1 (PALI1), PALI2 or elongin BC and PRC2-associated protein (EPOP).

---

**References**

---

---
By contrast, PRC2.2 complexes contain Jumonji and (A+T)-rich interaction domain-containing protein 2 (JARID2) and adipocyte enhancer-binding protein 2 (AEBP2). Importantly, PRC2.1-specific and PRC2.2-specific subunits also affect PRC2 methyltransferase activity. For example, within PRC2.2, AEBP2 contains a patch of basic amino acids that contribute to nucleosome binding and stimulate H3K27 methylation.

Similarly, incorporation of either EPOP or PALI1 into PRC2.1 supports its methyltransferase activity through an undefined mechanism. Interestingly, EZH inhibitory protein (EZHIP), a protein that is expressed predominantly in germ cells, was recently shown to interact with the core PRC2 complex and inhibit its methyltransferase activity while also blocking interactions with PRC2.1-specific and PRC2.2-specific subunits.

**Fig. 1** | A diverse repertoire of mammalian Polycomb repressive complexes. a | The core of Polycomb repressive complex 1 (PRC1) comprises a RING1A or a RING1B protein (RING1A/B) and one of six Polycomb group RING finger (PCGF) proteins (PCGF1–PCGF6). RING1 and PCGF proteins dimerize through their RING domains to form the catalytic core of PRC1, while the RING finger and WD40-associated ubiquitin-like (RAWUL) domains of both proteins interact with a range of auxiliary subunits, giving rise to biochemically distinct PRC1 complexes. Canonical PRC1 complexes (top) assemble around PCGF2 or PCGF4, and include a chromobox protein (CBX2, CBX4, CBX6, CBX7 or CBX8) and a Polyhomeotic (PHC) protein (PHC1, PHC2 or PHC3). In some cases, canonical PRC1 complexes also contain an SCM protein (sex comb on midleg-like protein 1 (SCML1), SCML2 or sex comb on midleg homologue 1 (SCMH1)). By contrast, variant PRC1 complexes (bottom) can assemble around all six PCGFs and include a chromobox protein (CBX2, CBX4, CBX6, CBX7 or CBX8) and a Polyhomeotic (PHC) protein (PHC1, PHC2 or PHC3). In some cases, canonical PRC1 complexes also contain an SCM protein (sex comb on midleg-like protein 1 (SCML1), SCML2 or sex comb on midleg homologue 1 (SCMH1)). By contrast, variant PRC1 complexes (bottom) can assemble around all six PCGFs and contain RING1 and YY1-binding protein (RYBP) or YY1-associated factor 2 (YAF2). The identity of the PCGF protein dictates the incorporation of other auxiliary subunits, resulting in a number of distinct variant PRC1 complexes.

b | The catalytic lobe of PRC2 is formed by a SET (Su(var)3-9, Enhancer of Zeste and Trithorax)-domain containing Enhancer of Zeste 1 (EZH1) or EZH2 protein, together with embryonic ectoderm development (EED) and the VEFS (VRN2–EMF2–FIS2–SUZ12) domain of Suppressor of Zeste 12 (SUZ12). The amino-terminal part of SUZ12 forms a distinct regulatory lobe that interacts with retinoblastoma-binding protein 4 (RBBP4) or RBBP7 and with other, auxilary subunits that give rise to distinct PRC2.1 (top) and PRC2.2 (bottom) complexes. PRC2.1 complexes contain a Polycomb-like (PCL) subunit (PCL1/2/3) and elongin BC and PRC2-associated protein (EPOP) or PRC2-associated LCOR isoform 1 (PALI1) or PALI2, whereas PRC2.2 complexes contain adipocyte enhancer-binding protein 2 (AEBP2) and Jumonji and (A+T)-rich interaction domain-containing protein 2 (JARID2). Protein domains are italicized. AUTS2, autism susceptibility protein 2; BCOR, BCL-6 co-repressor; CK2, casein kinase 2; DP1, dimerization partner 1 (also known as transcription factor Dp-1); FBRS, fibrosin; HDAC, histone deacetylase; KDM2B, lysine-specific demethylase 2B; L3MBTL2, lethal(3)malignant brain tumour-like protein 2; MAX, MYC-associated factor X; MGA, MAX gene-associated protein; SKP1, S-phase kinase-associated protein 1; USP7, ubiquitin carboxy-terminal hydrolase 7.
This function suggests that context-specific regulators of PRC2 activity have important roles during development.

**Target site identification**

Despite an intrinsic capacity to engage with nucleosomes, genome-wide studies have shown that Polycomb complexes are enriched at gene promoters and other gene regulatory elements, indicating that more specific mechanisms must dictate the observed binding patterns. In the following sections, we examine emerging evidence that primary mechanisms of Polycomb complex targeting in mammals rely on a complement of sequence-specific DNA-binding factors, RNA-dependent mechanisms and binding to CpG islands (CGIs).

**Target site recognition through sequence-specific DNA-binding factors.** In *D. melanogaster*, targeting of Polycomb complexes is thought to rely primarily on sequence-specific DNA-binding factors. However, many of these factors lack clear mammalian orthologues, or if they are present in mammals, do not target Polycomb...
complexes to chromatin. This difference has brought into question whether these mechanisms are also important for mammalian Polycrom target site identification. In beginning to address this important question, it has recently been shown that the PCGF6-vPRC1 complex, which is absent in D. melanogaster, stably incorporates auxiliary subunits that function as sequence-specific DNA-binding domains ([Fig. 2a]). For example, the subunit MAX gene-associated protein (MGA) contains a zinc-finger-CXXC domain that binds specifically to non-methylated CpG dinucleotides. In PRC1, the Polyclb-1 (PCL1), PCL2 or PCL3 subunit contains a winged-helix domain that binds non-methylated CpG dinucleotides in certain sequence contexts.

Other mammalian Polycrom repressive complexes do not stably incorporate proteins with sequence-specific DNA-binding domains. Nevertheless, certain Polycrom complexes can engage in more transient interactions with DNA-binding factors, and this engagement appears to be important for target site recognition in certain contexts. For example, PCGF3-vPRC1 recognizes some of its target sites in embryonic stem cells (ESCs) by interacting with the transcription factors upstream stimulatory factor 1 (USF1) and USF2 ([Ref. 26]). Other DNA-binding factors, including REST, RUNX1 and SNAIL1, have been proposed to contribute to PRC1 or PRC2 targeting in certain cell types. Despite these reports, in most cases direct interactions between DNA-binding factors and Polycrom complexes have not been unequivocally demonstrated, and they often account for only a subset of Polycrom complex binding events in a limited number of developmental contexts. Therefore, it appears unlikely that sequence-specific DNA-binding factors alone can explain how Polycrom complexes recognize their target sites in mammals.

**Target site identification through chromatin-associated RNA.** As an alternative to DNA-binding transcription factors, it was suggested that targeting of Polycrom complexes could be achieved through long non-coding RNAs (lncRNAs) that associate with defined sites in chromatin. Mammalian X chromosome inactivation, which is mediated by the lncRNA X inactive specific transcript (XIST), involves Polycrom repressive complexes. During X chromosome inactivation, XIST is expressed from one of the two female X chromosomes and spreads in cis to drive monoallelic chromosomal gene silencing. Detailed characterization of this process has revealed that the PCGF3/5-vPRC1 complex interacts with the RNA-binding protein heterogeneous nuclear ribonucleoprotein K (hnRNPK), which recognizes XIST, leading to PCGF3/5-vPRC1 enrichment on the inactive X chromosome ([Ref. 27] ([Fig. 2b]). However, this example is somewhat atypical for Polycrom complex targeting in that XIST leads to enrichment of PCGF3/5-vPRC1 across the entire inactivated X chromosome, whereas on autosomes, Polycrom complex enrichment is typically more restricted to gene regulatory elements.

On the basis of early observations made in XIST studies, numerous other lncRNAs have been proposed to target Polycrom complexes to sites in autosomes. Notably, the lncRNA HOX transcript antisense RNA (HOTAIR) was reported to function as a trans-acting repressor of the HOXD locus by physically interacting with and recruiting PRC2 ([Ref. 28]). However, recent work has demonstrated that transcription repression by HOTAIR is independent of PRC2, and that PRC2 recruitment to the HOXD locus occurs primarily in response to gene silencing, not through direct targeting by HOTAIR. These findings highlight the need for caution when interpreting how lncRNAs influence gene expression and the recruitment of Polycrom complexes ([Ref. 29]). However, consistent with a potential role for lncRNAs in targeting the Polycrom complexes to chromatin, the lncRNAs Airn and Kcnq1ot1 promote monoallelic expression of autosomal genes in mouse tropheoblast stem cells through a mechanism that appears to be highly analogous to X chromosome inactivation ([Ref. 30]).

In addition to lncRNAs, recent studies have suggested that RNA–DNA hybrid structures known as R-loops may also support Polycrom target site recognition. R-loops were found to occur at a number of Polycrom target genes, and their enzymatic resolution perturbed Polycrom complex targeting at some sites. By contrast, in other contexts, R-loops were reported to negatively affect PRC2 binding and activity. Although there
appears to be a connection between Polycomb complexes and R-loops, the underlying biochemical mechanisms are unknown, and the extent to which R-loops mediate Polycomb complex targeting to chromatin remains to be elucidated.

**Targeting to CGIs.** Polycomb complexes appear to recognize some of their target sites using sequence-specific DNA-binding factors or through associating with IncRNAs. However, these mechanisms alone cannot explain all Polycomb complex binding observed in mammalian genomes, suggesting that additional targeting mechanisms exist. Early genome-wide studies revealed that Polycomb complexes associate primarily with CGIs\(^\text{105,106}\). CGIs are short (1–2-kb) regions of CpG-rich DNA that are associated with approximately 70% of mammalian gene promoters\(^\text{89,90}\). In CGIs, CpG dinucleotides remain free of DNA methylation, whereas elsewhere in the genome they are heavily methylated and function to maintain heterochromatin and repress the expression of parasitic DNA elements\(^\text{106,107}\). CGIs are absent in non-vertebrate model organisms, including *D. melanogaster*, which lacks DNA methylation. This led to speculation that CGIs may be used by Polycomb complexes to identify target gene promoters and other regulatory elements in mammals\(^\text{108,109}\).

Molecular evidence that CGI recognition may be important for Polycomb complex targeting first came from biochemical purifications of mammalian PRC1 complexes, in which the PCGF1-vPRC1 complex was shown to stably integrate lysine-specific demethylase 2B (KDM2B)\(^\text{108,109}\). KDM2B contains a zinc-finger-CXXC domain that specifically binds to non-methylated CpG dinucleotides\(^\text{108,109}\). Consistent with this specificity, KDM2B can target the PCGF1-vPRC1 complex to CGIs\(^\text{107,108}\) (Fig. 2c). PCL proteins in PRC2.1 complexes were also recently shown to bind to non-methylated CpG-containing DNA through a winged-helix domain that has a preference for certain sequences and is also influenced by the helical properties of DNA\(^\text{113,114}\) (Fig. 2c). Consistent with in vitro evidence that PRC2 may function in some cases as a dimer\(^\text{113}\), it was recently proposed that a PRC2.1 dimer could stabilize binding to CpG-rich DNA, possibly by enabling simultaneous engagement with multiple non-methylated CpG sites\(^\text{114}\). Although the precise contribution of dimerization to PRC2.1 function in vivo remains to be elucidated, the winged-helix domains of PCL proteins are crucial for PRC2.1 targeting, indicating that CGI binding is also central to target site identification by PRC2\(^\text{REF.115}\).

**The Polycomb chromatin domain**

Once Polycomb complexes engage with target sites, this primary targeting must then be converted into the formation of repressive Polycomb chromatin domains, which can extend up to tens of kilobases and have extremely high levels of H2AK119ub1, H3K27me3 and Polycomb complex occupancy. In the following sections, we describe an intricate series of feedback and communication mechanisms that link the activity of Polycomb complexes on chromatin to enable the formation and maintenance of Polycomb chromatin domains.

**H2AK119ub1 is recognized by PRC2.2 and vPRC1.** The complex regulation of mammalian gene promoters and the diversity of Polycomb targeting mechanisms have made it challenging to study the formation of Polycomb chromatin domains. To circumvent this difficulty, tethering experiments have been used, in which specific Polycomb complexes can be artificially nucleated at an ectopic site in the genome and their capacity to communicate with other Polycomb components and to form Polycomb chromatin domains can be examined. Such experiments revealed that tethering of PRC1 was sufficient to drive de novo enrichment of PRC2, accumulation of H3K27me3 and the formation of a Polycomb chromatin domain that spreads several kilobases from the primary engagement site\(^\text{116,117}\). Interestingly, this activity was exclusive to vPRC1 complexes and relied on their capacity to ubiquitylate H2AK119\(^\text{REF.59}\) (Fig. 3a). Subsequently, a biochemical link between H2AK119ub1 and PRC2 was narrowed down to the JARID2 subunit of PRC2.2, which can directly bind to H2AK119ub1 through a ubiquitin binding motif\(^\text{118}\); this interaction also stimulates the methyltransferase activity of PRC2\(^\text{REF.19}\) (Fig. 3b). The structural basis of this interaction was recently elucidated in a cryo-electron microscopy model of PRC2.2 bound to an H2AK119ub1-containing mononucleosome, which showed JARID2 interacts with monoubiquitylated H2A on one face of the nucleosome through its ubiquitin-binding motif\(^\text{118}\). Interestingly, AEBP2 interacts with a second monoubiquitylated H2A on the opposite face of the nucleosome through tandem C2H2 zinc-fingers. This indicates that auxiliary subunits in PRC2.2 have a specialized role in binding to H2AK119ub1-containing nucleosomes to trimethylate H3K27, which is consistent with earlier biochemical work\(^\text{119,120}\). The importance of H2AK119ub1-mediated vPRC1 communication with PRC2 in the formation of Polycomb chromatin domains has also been shown through detailed genetic perturbation studies in mouse ESCs. Disruption of vPRC1 complexes caused major reductions in PRC2 occupancy and H3K27 trimethylation in Polycomb chromatin domains, and this relied on their E3 ligase activity, consistent with ectopic tethering experiments and recent work in the developing mouse and zebrafish embryo\(^\text{32,30,72,107,115,119–125}\).

In addition to its influence on PRC2.2, H2AK119ub1 also reinforces chromatin binding and modification by vPRC1 through a zinc-finger in RYBP, which can directly bind H2AK119ub1\(^\text{32,117,126}\) (Fig. 3a). This interaction forms the basis of a feedback mechanism that supports spreading of H2AK119ub1 onto neighbouring nucleosomes in a manner that appears to be aided by histone H1\(^\text{REF.56}\). Together, these data demonstrate that once Polycomb target sites are identified through the primary targeting of vPRC1, feedback mechanisms inherent to vPRC1 complexes and H2AK119ub1\(^\text{REF.58,60,72,107,115,119–125}\), are essential for the formation and spreading of Polycomb chromatin domains.

**H3K27me3 stimulates PRC2 activity and is recognized by cPRC1.** Although H3K27me3 and the occupancy of PRC2.2 in Polycomb chromatin domains relies on
PRC1 and H2AK119ub1, primary targeting modalities also enable PRC2 to identify target sites and catalyse H3K27 trimethylation. PRC2 can then bind the product of its own catalysis (H3K27me3) through a WD40-repeat domain in EED, causing allosteric activation of its methyltransferase activity\(^\text{39,127,128}\) (Fig. 3c). PRC2 activation creates a feedback mechanism that enables spreading of H3K27me3 along chromatin\(^\text{129}\). A recently elucidated cryo-electron microscopy structure of PRC2 bound to a dinucleosome revealed that EED engages with one H3K27me3-containing nucleosome\(^\text{43}\) and positions the SET domain of EZH2 to methylate an unmodified H3 tail on an adjacent nucleosome\(^\text{39}\). The geometry of this interaction also revealed that a short length of linker DNA between nucleosomes is important for optimal methyltransferase activity, which is consistent with evidence that PRC2 prefers more compact chromatin substrates\(^\text{126}\).

H3K27me3 is also recognized by a chromodomain in the CBX subunit of cPRC1 complexes, suggesting that PRC2 can communicate with cPRC1 through H3K27me3\(^\text{6,8,131,132}\) (Fig. 3d). In agreement with this

---

**Fig. 3 | Formation and spreading of Polycomb chromatin domains.** Primary targeting of Polycomb repressive complex 1 (PRC1) and PRC2 is followed by feedback and communication mechanisms that enable the formation and spreading of Polycomb chromatin domains, which are characterized by elevated Polycomb complex occupancy and high levels of monoubiquitylated histone H2A Lys119 (H2AK119ub1) and trimethylated histone H3 Lys27 (H3K27me3). a | H2AK119ub1 is recognized by the RING1 and YY1-binding protein (RYBP) subunit (or by the YY1-associated factor 2 (YAF2) subunit; not shown) of variant PRC1 (vPRC1) complexes. This creates a feedback mechanism, supported by histone H1, that reinforces vPRC1 binding and amplifies H2AK119 ubiquitylation, thereby enabling spreading of vPRC1 and H2AK119ub1 away from the primary vPRC1 targeting site. b | PRC1 complexes ubiquitylate H2AK119, which is then recognized by the Jumonji and (A+T)-rich interaction domain-containing 2 (JARID2) and adipocyte enhancer-binding protein 2 (AEBP2) subunits of PRC2.2. This causes elevated PRC2 occupancy and stimulates the trimethylation of H3K27. Therefore, H2AK119ub1 facilitates communication between PRC1 and PRC2 in Polycomb chromatin domains. c | H3K27me3 is recognized by the embryonic ectoderm development (EED) subunit of PRC2, which allosterically activates its methyltransferase activity. H3K27me3 creates a feedback mechanism that reinforces PRC2 binding and amplifies H3K27 trimethylation, thereby enabling spreading of PRC2 and H3K27me3 away from the primary targeting site. A short length of linker DNA between nucleosomes supports the methyltransferase activity of PRC2. d | H3K27me3 is also recognized by the chromobox (CBX) subunit (CBX2, CBX4, CBX6, CBX7 or CBX8) of canonical PRC1 (cPRC1) complexes. Although cPRC1 complexes are less catalytically active than vPRC1 complexes (as indicated by the dotted arrow), in some contexts their activity may lead to H2AK119 ubiquitylation. Therefore, H3K27me3 can facilitate communication between PRC2 and PRC1 in Polycomb chromatin domains.
The capacity for liquid–liquid phase separation resides in the condensates and their relevance to Polycomb functions is therefore required.

### cPRC1 supports the organization of Polycomb chromatin domains in the nucleus

Although cPRC1 complexes contribute minimally to H2AK119ub1 and stabilize cPRC1 occupancy during the formation of Polycomb chromatin domains.

**cPRC1 supports the organization of Polycomb chromatin domains in the nucleus.** Although cPRC1 complexes contribute minimally to H2AK119ub1 and stabilize cPRC1 occupancy during the formation of Polycomb chromatin domains. However, consistent with the fact that cPRC1 complexes have extremely low E3 ligase activity in vitro, cPRC2 and cPRC1 contribute only minimally to H2AK119ub1 levels in vivo. Furthermore, ectopic tethering of PRC2 or cPRC1 to chromatin in ESCs did not lead to elevated H2AK119ub1 levels, indicating that these complexes are incapable of efficiently driving de novo Polycomb chromatin domain formation in this context. Therefore, in contrast to vPRC1 complexes, which communicate effectively with PRC2.2 to drive H3K27 trimethylation, PRC2 appears to work much less efficiently through cPRC1 to reinforce H2AK119 ubiquitylation. Nevertheless, overexpression of the cPRC1 subunit CBX7 was recently shown to support H3K27me3-dependent maintenance of H2AK119ub1 at an artificial target gene promoter, suggesting that in some circumstances PRC2 may support this process more efficiently. In summary, whereas primary targeting modalities and H2AK119ub1 recognition are important for PRC2 nucleation at target sites, H3K27me3 deposition is necessary to reinforce PRC2 binding, support spreading of PRC2 and H3K27me3, and stabilize cPRC1 occupancy during the formation of Polycomb chromatin domains.

**The maintenance of Polycomb chromatin domains following DNA replication.** The integrity of Polycomb chromatin domains is challenged by passage of the replication machinery, which displaces parental nucleosomes from DNA and necessitates the incorporation of new, unmodified histones and the reincorporation of modified ‘parental’ histones into the replicated daughter strands. Communication and feedback mechanisms inherent to Polycomb complexes could in theory support copying and epigenetic maintenance of Polycomb chromatin domains, but this would rely on modified parental nucleosomes being deposited in the same location on daughter strands. Elegant experiments examining the kinetics of histone deposition during DNA replication have recently shown that H3K27me3-containing nucleosomes are incorporated into daughter DNA strands almost precisely where they originated from in the parental DNA, providing a potential conduit for maintaining the identity of Polycomb chromatin domains.

Interestingly, however, a considerable lag exists between DNA replication and the establishment of H3K27me3 on the newly incorporated, unmodified histone H3, despite reports that Polycomb repressive complexes can travel with the replication machinery. Furthermore, H3K27me3 profiles can be re-established de novo following complete erasure of H3K27me3 and work in D. melanogaster has shown that long-term maintenance of H3K27me3 in Polycomb chromatin domains relies on the underlying DNA sequence. Together, these observations have brought into question the extent to which epigenetic maintenance through feedback mechanisms in Polycomb chromatin domains contributes to their re-establishment following DNA replication. It will be important to understand in more detail the kinetics of re-establishment of Polycomb chromatin domains following replication, including whether parental H2AK119ub1 and H3K27me3 are also faithfully redeposited. Understanding these basic principles will be essential for defining whether epigenetic maintenance is important for the function of Polycomb chromatin domains in actively dividing cells.
Controlling Polycomb chromatin domains

It is often posited that chromatin-modifying complexes are recruited to defined target sites in the genome, where they modify histones to drive gene activation or repression. However, these simple instructive models appear for the most part to be incompatible with our evolving understanding of where and how Polycomb chromatin domains form and function. If instructive processes dictated the formation of Polycomb chromatin domains (Fig. 5a), one would expect that all sites in the genome where primary targeting mechanisms allow Polycomb complexes to engage would acquire repressive Polycomb chromatin domains. However, genome-wide studies have shown that only a small subset of sites where Polycomb complex binding can be detected ultimately go on to form Polycomb chromatin domains that have high levels of H2AK119ub1, H3K27me3 and Polycomb complex occupancy. For example, the PCGF1-vPRC1 complex is detected at most CGIs in the genome, albeit at low levels. However, despite the inherent de novo Polycomb chromatin domain-forming activity of PCGF1-vPRC1 (REF. 107), only ~20% of CGIs in a given cell type will acquire a domain, indicating that this process must be highly regulated. In this section we discuss the emerging principles that regulate Polycomb chromatin domain formation and how this is related to transcription.

Transcription regulates susceptibility to Polycomb chromatin domain formation. A unifying feature of target sites that form Polycomb chromatin domains is extremely low or non-detectable levels of transcription.
raising the possibility that transcription itself limits the activity of Polycomb complexes. In agreement with this notion, treating cells with transcription inhibitors can stabilize PRC2 binding and increase H3K27me3 levels at thousands of previously active CGI-associated gene promoters. A similar effect was noted when a polyadenylation signal, which promotes transcription termination, was inserted adjacent to an active CGI-associated promoter. These findings are consistent with experiments showing that synthetic CpG-rich non-methylated DNA can recruit Polycomb complexes and lead to the apparent formation of Polycomb chromatin domains, but only when these sites lack binding of transcription activators or are not close to enhancer elements. Furthermore, experiments examining the kinetics of Polycomb chromatin domain formation during cellular-state transitions demonstrated that this primarily occurs following transcription cessation. Together, these observations indicate that despite Polycomb complexes broadly engaging with regulatory elements through targeting modalities such as CGI recognition, transcription potently counteracts the formation of Polycomb chromatin domains.

**Nascent RNA regulates the activity of PRC2.** The realization that gene activity appears to be incompatible with Polycomb chromatin domain formation has led to a concerted effort to understand the mechanisms that regulate this process. Although some lncRNAs appear to target Polycomb complexes to specific sites, an overwhelming body of evidence now also indicates that PRC2 binds promiscuously to nascent RNA. Several studies have reported that nascent RNA competes for PRC2 binding and displaces it from chromatin, while also inhibiting its methyltransferase activity, both of which would antagonize the formation of Polycomb chromatin domains.
view and proposed that interaction with nascent RNA, either directly by PRC2 (REF.226) or through the pre-mRNA splicing factor RBFOX2 (REF.197), supports PRC2 occupancy at target sites.

To reconcile these seemingly contradictory effects, an ‘RNA-bridging model’ has been proposed204. In this model, nascent transcripts from very lowly transcribed genes help bring PRC2 into proximity with promoter elements to support chromatin binding and trimethylation of H3K27. However, at more highly transcribed genes, elevated levels of nascent transcripts would compete with PRC2 for chromatin binding and decrease H3K27 trimethylation. Although the molecular details of this model remain to be fully tested, it appears that PRC2 is highly attuned to sensing transcript levels and using this information to regulate the formation of Polycomb chromatin domains.

**The influence of transcription-associated chromatin features.** In addition to producing nascent RNA, transcription also profoundly influences the chromatin environment at transcribed genes. For example, during transcription initiation, histone methyltransferases trimethylate H3K4 at gene promoters, whereas elongation is associated with H3K36 trimethylation in gene bodies97,196. Importantly, in vitro studies have demonstrated that the methyltransferase activity of PRC2 is inhibited by the presence of H3K4me3, H3K36me2 or H3K36me3 (REFS.15,197–202), and disruption of H3K4 methyltransferase activity in vivo can lead to enhanced Polycomb chromatin domain formation203. In addition, histone demethylases act as part of transcription co-activator complexes to remove H3K27 methylation205–209, and transcription-associated H3K27, H4K16 and H4K20 acetylation counteracts PRC2 activity42,197–202. Similarly, the H2AK119ub1-specific deubiquitylases ubiquitin carboxy-terminal hydrolase 16 (USP16) and 2A-DUB (also known as MYSM1) have been shown to associate with promoters of actively transcribed genes and counteract H2A.K119 ubiquitylation196,207.

In addition to histone modification-dependent effects, the chromatin remodelling ATPase BRG1 can actively evict PRC1 from promoters of transcribed target genes118–220. This activity was reported to affect RYBP-containing vPRC1 complexes, suggesting that promoter-associated chromatin remodelling activities can displace vPRC1 complexes and limit H2A.K119 ubiquitylation221. Together, these findings suggest that although primary targeting activities may enable Polycomb complexes to engage with a wide range of gene promoters, transcription-associated histone modifications and chromatin remodelling activities at regulatory elements can limit the catalysis and feedback mechanisms required for efficient formation of Polycomb chromatin domains.

**Target sampling and a responsive model of Polycomb chromatin domain formation.** Our growing understanding of Polycomb complex targeting and the mechanisms that control Polycomb chromatin domain formation appears to be consistent with a dynamic sampling mode of action that is responsive to transcription222 (FIG. 5d).

According to this model, DNA-binding activities associated with Polycomb complexes would provide a mechanism to dynamically engage with, or sample, potential target sites. Central to this process would be PCGF1-vPRC1 and PRC2.1, which can recognize non-methylated DNA in CGI elements that are associated with most mammalian gene promoters198–212. Sampling could allow Polycomb complexes to constantly interrogate a wide range of potential target sites, with only those that are lowly or not transcribed being susceptible to Polycomb-mediated histone modification, which is necessary to initiate the feedback and communication mechanisms required for efficient formation and spreading of Polycomb chromatin domains.

A responsive mode of Polycomb function based on target site sampling would in theory provide a generic mechanism for the formation of Polycomb chromatin domains at non-transcribed or lowly transcribed genes, without the need to evolve complex cell type-specific targeting mechanisms. Therefore, the Polycomb system could function to protect inactive genes against low-level or inappropriate transcription activation signals and help to maintain gene repression in many, if not all, cell types. This model is also consistent with early genetic experiments in fruit flies and mammals in which Polycomb proteins were observed to maintain, rather than initiate, gene repression213,214. Nevertheless, more work is required to test the validity of a sampling-based model for the formation and function of Polycomb chromatin domains in different biological contexts.

**Polycomb-mediated gene regulation**

After receptive target sites establish Polycomb chromatin domains, these then have an important role in controlling gene expression. In this section we examine the mechanisms of Polycomb chromatin domain-mediated gene repression, discuss atypical cases of gene activation by Polycomb complexes and consider how CGIs might utilize Polycomb complexes to create bistable chromatin states that shape gene expression during cell differentiation.

**Transcription repression by Polycomb chromatin domains.** Early models posited that the Polycomb system may repress transcription through cPRC1 binding to H3K27me3 and compacting chromatin to limit access to gene regulatory elements197,198,199,210,215,227. Although this mechanism may contribute to Polycomb-mediated gene repression in some contexts, mice deficient in cPRC1 components often have milder phenotypes than mice in which vPRC1 components are perturbed, and cPRC1-deficient ESCs exhibit relatively few gene expression changes117,120,121,148,220,228,229. Furthermore, chromatin accessibility does not always change when Polycomb systems are disrupted211,212. Together, these findings indicate that, at least during early development and in ESCs, Polycomb chromatin domain features distinct from H3K27me3 and cPRC1 must contribute centrally to transcription repression.

In line with this possibility, recent studies have pointed to a central role for H2AK119ub1 and vPRC1 in gene repression. This is evident in ESCs, where disruption of PRC1 resulted in increased expression of...
thousands of Polycomb target genes, effects that were almost entirely dependent on PCGF1/3/5/6-vPRC1 complexes. Importantly, disruption of the catalytic activity of PRC1 recapitulated these gene expression defects despite vPRC1 binding being retained at target sites, suggesting that gene repression by PRC1 in this context is mediated by H2AK119 ubiquitylation. H2AK119ub1 having a central role in gene repression is consistent with PRC1 binding to chromatin being highly dynamic and displaying extremely low target site occupancy. Although the precise mechanism or mechanisms by which H2AK119ub1 affect transcription remain to be determined, it does not appear to rely on PRC2 occupancy. Furthermore, although H2AK119ub1-reader proteins, including RYBP and the chromatin-remodelling factor RSF1, have been linked with transcription repression, their activities appear insufficient to account for PRC1-mediated gene repression. An alternative possibility is that the addition of a bulky ubiquitin moiety to H2A could more directly counteract the function of the transcription machinery. Consistent with this possibility, acute disruption of PRC1 and loss of H2AK119ub1 caused rapid new binding of Pol II and elevated the frequency of transcription bursts, suggesting PRC1 and H2AK119ub1 affect transcription initiation, which is in agreement with previous reports that PRC1 might antagonize the assembly or activity of the transcription pre-initiation complex. Other reports have suggested that PRC1 limits transcription by constraining a post-initiation, promoter-proximal ‘poised’ form of Pol II. More detailed biochemical work is required to define the mechanisms by which H2AK119ub1 influences transcription.

Consistent with the catalytic activity of PRC1 being required for Polycomb-mediated gene repression in ESCs, an essential role for H2AK119ub1 in Polycomb target gene repression was recently demonstrated by examination of the early stages of neuronal cell fate restriction. However, this requirement was less pronounced at later stages of neuronal differentiation, when a shift towards H2AK119ub1-independent gene repression was observed. These findings suggest that communication and feedback between PRC1 and PRC2, which support the formation of Polycomb chromatin domains, appear in some contexts to synergize the repressive activities of PRC1 and PRC2 at target genes (centre), thereby providing a robust barrier against inappropriate gene expression.

In some contexts, Polycomb complexes can activate genes. Canonical PRC1 (cPRC1) mediates the formation of chromatin topologies that can bring poised enhancers and their target promoters into close proximity (left). Once activation signals are received at the enhancer through transcription factor (TF) binding, rapid induction of transcription can ensue (right). PCGF2/4, Polycomb group RING finger 2 or 4; Pol II, RNA polymerase II.

---

**Fig. 6 | Mechanisms of Polycomb-mediated gene regulation.**

**a** Repression

PRC1-mediated repression

PRC2-mediated repression

Synergistic repression

**b** Poising for activation?

H2AK119ub1 H3K27me3

Poising for activation?

H2AK119ub1 H3K27me3

H2AK119ub1

H3K27me3

H2AK119ub1

H3K27me3

H2AK119ub1

H3K27me3

H2AK119ub1

H3K27me3

H2AK119ub1

H3K27me3

H2AK119ub1

H3K27me3

H2AK119ub1

H3K27me3

H2AK119ub1

H3K27me3

H2AK119ub1

H3K27me3
H2AK119ub1 could possibly confer a form of repression in stem cells that can be overcome in response to appropriate developmental gene expression cues. By contrast, as cells commit to particular lineages and subsets of genes will no longer be expressed, the Polycomb system may use additional mechanisms, perhaps involving structural changes to chromatin, that enable long-term and robust maintenance of repression. This possibility has been supported by studies examining the contribution of the catalytic activity of PRC1 to mouse development which reported that PRC1 uses a combination of H2AK119ub1-dependent and H2AK119ub1-independent mechanisms to control gene expression. However, it was subsequently shown that the PRC1 catalytic mutations used in these studies do not fully eliminate H2AK119 ubiquitylation. Nevertheless, in some cases in D. melanogaster, the repression of Hox genes and other Polycomb target genes can be maintained by PRC1 independently of H2AK119 ubiquitylation. Clearly more work is required to define the contribution of H2AK119ub1-dependent and H2AK119ub1-independent mechanisms to Polycomb-mediated gene repression in different cell types and developmental stages.

PRC2 also has important roles in Polycomb-mediated gene repression, yet its disruption in mammals typically has less severe effects on gene expression and development than removal of PRC1. Studies in D. melanogaster have suggested that PRC2-mediated repression relies on H3K27 methylation, as a histone H3 Lys27-substitution mutation that renders this residue refractory to methylation largely phenocopies the gene expression and developmental defects observed when PRC2 is disrupted. Conversely, a hyperactive form of EZH2, which causes increased levels of H3K27me3, drives inappropriate repression of target genes. Consistent with these findings, chemical inhibitors that limit the methyltransferase activity of PRC2 and cause a reduction in H3K27 methylation are associated with derepression of Polycomb target genes in mammals, as is replication-mediated dilution of H3K27me3.

Recently, an evolutionarily conserved bromo adjacent homology (BAH) domain in BAHDC1 was shown to specifically recognize H3K27me3 and contribute to gene repression, possibly by recruiting histone deacetylases, which are known to counteract gene expression. This effect suggests that pervasive H2AK119 ubiquitylation functions like a rheostat to control the transcriptional potential of the genome.

PRC2 also acts pervasively throughout the genome: approximately 70% of H3K27 residues undergo monomethylation or dimethylation and up to 15% undergo trimethylation. Although the mechanisms that control H3K27 methylation outside Polycomb chromatin domains remain poorly understood, this pervasive activity relies primarily on the core complex and not on PRC2.1-specific and PRC2.2-specific subunits. Importantly, there is evidence that these alternative pools of H3K27me can broadly antagonize transcription, possibly by counteracting H3K27 acetylation and limiting the activity of enhancers. Therefore, it is clear that Polycomb complexes can also influence gene expression by functioning outside Polycomb chromatin domains, and it has recently been suggested that these activities are particularly relevant in germ cells and in very early developmental stages.

**Gene activation by Polycomb complexes.** Although Polycomb complexes appear to primarily support transcription repression, there is increasing evidence that in some contexts Polycomb complexes can potentiate...
gene expression. For example, a PCGF5-vPRC1 complex that contains AUTS2 and casein kinase 2 and inefficiently ubiquitylates H2AK119 was shown to potentiate transcription by recruiting the histone acetyltransferase p300 (Ref. 279). Alternatively, the H2AK119ub1-specific reader ZRF1 (also known as DNAJC2) was linked to activation of Polycomb-repressed genes during differentiation280. These effects are in agreement with earlier reports of PRC1 supporting gene expression in resting B cells and T cells280.

It has also been suggested in some instances that Polycomb chromatin domains can support gene expression by forming chromatin topologies that aid the induction of gene expression75,58,284–286. Specifically, it has been proposed that, by supporting long-range chromatin interactions, cPRC1 can bring into close proximity enhancers that are poised for activation and their target promoters, thereby enabling rapid induction of transcription in response to activation signals during cell differentiation58,284,285 (Fig. 6b). This idea is consistent with observations that cPRC1 is associated with enhancer function in breast cancer cells and with activation of genes during cardiac differentiation286,287. The precise mechanisms and extent to which Polycomb complexes potentiate gene expression is an exciting area of active investigation.

Chromatin bistability and transitions in gene expression.

During cellular differentiation, Polycomb target genes can be activated to support new cellular functions288,289. This activation is accompanied by loss of the Polycomb chromatin domain and acquisition of a new type of chromatin domain formed through the function of Trithorax group chromatin-modifying complexes, which include myeloid/lymphoid or mixed-lineage leukaemia protein 1 (MLL1; also known as KMT2A)–MLL2 (KMT2D) and SETD1A–SETD1B290. Interestingly, some Trithorax group complexes can also sample CGIs through

---

**Fig. 7 | CpG islands and chromatin bistability.** a | Illustration of how chromatin bistability could form at CpG islands (CGIs). When transcription activation signals are low or absent, communication and feedback between Polycomb repressive complex 1 (PRC1) and PRC2 could drive the formation of repressive Polycomb chromatin domains (enriched in monoubiquitylated histone H2A Lys119 (H2AK119ub1) and trimethylated histone H3 Lys27 (H3K27me3)) that antagonize Trithorax (Trx) complexes and RNA polymerase II (Pol II) activity (left). When activation signals are high and persistent, communication and feedback between Trithorax complexes and Pol II could drive the formation of transcription-permissive Trithorax chromatin domains (enriched in H3K4me3) that antagonize PRC1 and PRC2 (right). The capacity of both Polycomb and Trithorax systems to sample CGIs coupled with the feedback and antagonistic mechanisms inherent to the formation of each chromatin state would provide the opportunity to switch between predominantly Polycomb or predominantly Trithorax chromatin states as gene activation signals increase or decrease (middle). We speculate that this mode of gene regulation could help to shape gene expression transitions (see part b) and also provide a chromatin-encoded hysteresis of the current transcriptional state of the gene in the face of inherently stochastic and pulsatile transcription initiation induced from single gene promoters.

b | If transitions between gene expression states (for example, gene induction during cellular differentiation) scaled linearly with activation signal, then one would predict a graded expression output (left). However, if CGIs help to create bistable chromatin at gene promoters, this could shape binary, switch-like, gene expression transitions, in which Polycomb chromatin domains constrain activation signals until appropriate activation thresholds are reached, at which point transcription initiation would precipitate a rapid switch into a Trithorax chromatin state and potentiate transcription. In such a system, one might predict CGIs could help to convert graded gene activation signals into binary, switch-like gene expression outputs through chromatin bistability. This mechanism could be particularly useful in supporting decisive gene expression transitions during development. Adapted with permission from Ref. 290, Elsevier.
zinc-finger-CXXC DNA-binding domains, but in contrast to the Polycomb complexes, they initiate the formation and spreading of transcription-permissive, H3K4me3-containing Trithorax chromatin domains through communication and feedback mechanisms linked to active transcription. Antagonism between the Polycomb and Trithorax complexes helps to ensure that at most gene promoters either a transcriptionally repressive or a transcriptionally permissive chromatin state predominates.

On the basis of these feedback mechanisms and antagonism between the Polycomb and Trithorax systems, CGIs have been proposed to support the formation of bistable chromatin states at gene promoters to regulate gene expression. In defined cell types, the capacity to form either Polycomb or Trithorax chromatin domains could provide memory of gene expression programmes and buffer against low-level or spurious transcription activation signals that could be detrimental to cell identity. However, the convergence of these systems on CGIs might also shape gene expression transitions during differentiation. For example, Polycomb chromatin domains maintained by feedback mechanisms at inactive genes could constrain gene activation signals until an appropriate induction threshold is reached. By contrast, following productive initiation of transcription, formation of permissive Trithorax chromatin domains could potentiate transcription while also blocking Polycomb activity. A rapid switch between chromatin states might influence transcription by converting graded gene activation signals into binary gene expression outputs.

In support of this possibility, the Polycomb system has been shown to impart switch-like effects on gene expression in Arabidopsis thaliana and, there is some evidence that this gene expression switch also occurs in mammals. We envisage that switch-like transitions in the expression of master regulators of cell fate could help to ensure decisive cell fate decisions, and Polycomb and Trithorax complexes are known to have important roles in regulating these types of genes during mammalian development. Furthermore, once individual genes acquire defined chromatin states, these states could provide hysteresis of their current transcription status to help maintain cell identify in the face of inherently stochastic and pulsatile transcription signals.

Recent mathematical models incorporating the known activities of Polycomb and Trithorax systems have provided theoretical evidence to support the existence of bistable chromatin states at gene regulatory elements. Importantly, interrogation of these models also suggests that bivalent chromatin states, which are characterized by the combined presence of H3K4me3 and H3K27me3 and were previously proposed to define an alternative poised transcriptional state, are unlikely to be stable and may instead correspond to the transition between Polycomb and Trithorax chromatin states. Although many of the underlying biochemical activities of Polycomb and Trithorax systems are consistent with the possibility that bistable chromatin states could be created at their target sites, a major challenge for our research field will be to directly test these models and their relevance to the kinetics of gene expression during cell differentiation.

**Conclusion and future perspective**

Detailed biochemical and structural characterization has revealed that the Polycomb system comprises a diverse array of multiprotein complexes that are highly specialized in how they engage with nucleosomes, catalyse histone modifications and function through communication and feedback mechanisms to form Polycomb chromatin domains. Our understanding of the mechanics that underpin these processes has advanced rapidly in recent years, but it still remains poorly understood at the mechanistic level how, once formed, Polycomb chromatin domains influence the transcription machinery to enable gene repression. This is a central unresolved question in our field. Given the complexity of gene regulatory elements in vivo, we anticipate that addressing this important question will require a reductionist approach leveraging in vitro transcription reactions on reconstituted chromatin templates and single-molecule approaches that can capture the behaviour of the transcription machinery.

Building on these biochemical underpinnings, the molecular study of how Polycomb systems function in vivo has indicated they are more dynamic than originally anticipated and, somewhat counterintuitively, appear to respond to transcription in forming Polycomb chromatin domains. This is forcing us to rethink how the Polycomb system both influences and is related to gene transcription. In this Review, we have discussed some emerging molecular principles that could explain these dynamic behaviours. However, the static and ensemble experimental approaches that our field has relied on until now are largely unsuited to exploring the relevance of these concepts. Instead, this will require live, single-cell and quantitative measurements. These approaches will be of particular relevance in determining whether the Polycomb system enables chromatin bistability, how it shapes the kinetics of gene-expression transitions during cell differentiation and development, and how the functions of Polycomb complexes are disrupted in human disease.

Finally, much of our understanding of the mammalian Polycomb system has come from studies in stem cells and in early developmental contexts. Although the ideas and models discussed in this Review draw primarily on this body of work, it is becoming increasingly clear that the mechanisms that guide Polycomb function in specific tissues and at later developmental stages are likely to differ considerably. Therefore, moving forwards, a central objective remains to uncover the complement of mechanisms and principles that define mammalian Polycomb-mediated gene regulation in diverse cellular and developmental contexts.

As our field ventures into testing the relevance of new and emerging concepts in Polycomb biology, the coming years will inevitably be a fast-paced and exciting period in the ongoing quest to uncover the molecular principles that define gene regulation by Polycomb repressive complexes.

Published online 16 August 2021
The role of the nucleosome acidic patch in modulating higher order chromatin structure. J. R. Soc. Interface 10, 20121002 (2013).

Rose, N. R. et al. RYBP stimulates PRC1 to shape chromatin-bounding polycomb repressive complexes. elife 5, e18591 (2016).

Taherihosseini, A. M., Huang, O. W. & Cochrane, A. G. BMI-1-RING1A ubiquitylation modulates H2A ubiquitylation in Drosophila. Cell. Mol. Biol. (Noisy) 62, 17–29 (2016).

Zhang, J. et al. RYBP/AF2-PRC1 complexes and histone H1-dependent chromatin compaction mediate propagation of H2A ubiquitylation during cell division. Cell. Mol. Biol. 62, 17–29 (2016).

Muller, J. et al. Histone methyltransferase activity of MLL in polycomb group silencing. Mol. Cell 74, 727–739 (2019).

Cao, R. et al. SUZ12 is required for both the structural and catalytic activity of polycomb repressive complex 2. Curr. Opin. Struct. Biol. 67, 135–144 (2020).

Glancy, E., Ciferrì, C. & Bracken, A. P. Structural basis for PRC2 engagement with nucleosomes. Curr. Opin. Struct. Biol. 67, 135–144 (2020).

Brosoun, A. et al. Polymerase repressive complex 2 structure with inhibitor reveals a mechanism of activation and drug resistance. Nat. Commun. 11, 1384 (2020).

Cao, R. & Zhang, Y. SUZ12 is required for both the histone methylation activity and the silencing function of the EZH2-EZH2 complex. Mol. Cell 15, 57–67 (2004).

Jiao, L. et al. Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. Science 350, 4a4385 (2018).

Antonny, J. et al. Recruitment of disease-associated mutations and putative mechanisms of autoinhibition revealed by X-ray crystallographic analysis of the EZH2-SET domain. PLoS ONE 8, e64147 (2013).

Justin, N. et al. Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. Nat. Struct. Mol. Biol. 26, 671–680 (2019).

Finogov, A. et al. Structural basis for PRC2 restoring of active histone modifications marks near silenced genes. Nat. Struct. Mol. Biol. 26, 671–680 (2019).

Pospel, S., Kasnath, V. & Nogales, E. Cryo-EM structures of PRC2 complexes simultaneously engaged with two histone H3K27me3 marks. Nat. Struct. Mol. Biol. 25, 154–162 (2018).

Chen, S., Jiao, L., Shubbar, M., Yang, X. & Liu, X. Unique structural platforms of Suz12 dictate distinct functions of polycomb binding. Mol. Cell 69, 840–852 e845 (2018).

Beringer, M. et al. EPOP functions linearly along elongin BC and USP7 to modulate the chromatin landscape. Mol. Cell 66, 659–672 (2017).

Zhang, Z. et al. PRC2 complexes with JARID2, MTF2, and e49/PRC1 modulate ES cell pluripotency and somatic cell reprogramming. Stem Cell 29, 229–240 (2011).

Samul, S. et al. JARID2 translocation via the polymerase repressive complex 2 regulates H3K73me3 deposition during cell differentiation. Mol. Cell 57, 769–783 (2015).

Lee, C. H. et al. Distinct stimulatory mechanisms regulate the catalytic activity of polycomb repressive complex 2. Mol. Cell 70, 435–448 e435 (2018).

Conway, E. et al. Family of vertebrate-specific cohesins recruits the PRC1 core complex and mediates PRC2 substrate activities. Mol. Cell 70, 408–421 e408 (2018).

Pajot, F. et al. Molecular heterogeneity and Cxcr6 alterations in posterior fossa group A (PFA) ependymomas. Acta Neuropathol. 136, 211–226 (2018).

Puni, A. et al. CATACOMB: an endogenous inducible gene that antagonizes H3K27 methylation by regulating the catalytic activity of polycomb repressive complex 2. Mol. Cell 70, 435–448 e435 (2018).

Mendenhall, E. M. et al. GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLoS Genet. 6, e1001244 (2010).

Ngan, C. Y. et al. Chromatin interaction analyses elucidate the roles of PRC2 in mouse development. Nat. Genet. 52, 266–272 (2020).

Loubiere, V., Papadopoulos, G. L., Szabo, Q., Martinez, A. M. & Cavalli, A. C. Developmental regulation of developmental gene expression characterized by PRC1-dependent chromatin looping. Sci. Adv. 6, eaax4401 (2020).

Chan, H. L. et al. PRC2 complexes associate with enhancers and promote oncogenic transcriptional programs in cancer through multiple mechanisms. Nat. Rev. Mol. Cell Biol. 20, 937–957 (2019).

Bernstein, B. E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 166, 464–476 (2016).

Ku, M. et al. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. Mol. Cell 70, 937–947 (2018).

Tanay, A., O’Donnell, A. H., Damelin, M. & Bestor, T. H. Hyperconserved Cpg domains underlie Polycomb-binding sites. Nat. Genet. 104, 5521–5526 (2020).

Schlesinger, Y. et al. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. Nat. Genet. 50, 252–256 (2018).

Bauer, M., Trupka, J. & Ringrose, L. The quest for mammalian Polycomb targets: are there any? Chromosoma 125, 471–496 (2016).

Kass, J. A. & Brown, J. L. Polycomb group response elements in insects and vertebrates. Adv. Genet. 81, 83–118 (2015).

Vella, P., Barozzi, I., Cuomo, A., Bonald, T. & Pasini, D. Yin Yang 1 extends the Myc-related transcription factors network in embryonic stem cells. Nat. Cell Biol. 16, 400–416 (2014).
81. Almeida, M. et al. PGC5/FS-PRC1 initiates Polycomb recruitment in X chromosome inactivation. Science 356, 1081–1084 (2017).
82. Pintacuda, G. et al. hnRNPK recruits PGC5/FS-PRC1 to the Xist RNA to establish polycomb-mediated silencing. Mol. Cell 68, 955–969 e910 (2017).
83. Chu, C. et al. Discovery of Xist RNA binding proteins. Cell 161, 404–416 (2015).
84. Zylcz, J. J. et al. The implication of early chromatin changes in X chromosome inactivation. Cell 176, 182–197 e1235 (2019).
85. Tsai, M. C. et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science 368, 689–697 (2020).
86. Somarowthu, S. et al. HOTAIR is dispensable for HOTAIR-mediated transcriptional repression. EMBO J. 36, 981–996 (2017).
87. Brodskoff, N. Noncoding RNA and Polycromb recruitment. RNA 19, 429–442 (2013).
88. Davidovich, C. & Cech, T. R. The recruitment of chromatin modifiers by long noncoding RNAs: lessons from PRC2. RNA 21, 2007–2022 (2015).
89. Pandey, R. R. et al. Kcnq1ot1 antisense noncoding RNA distinguishes two distinct classes of promoters. Mol. Cell Biol. 25, 456–464 (2015).
90. Arrigoni, R. et al. The Polycomb- associated protein BAP3 recruits polycomb repressive complexes PRC1 and PRC2. Nat. Chem. Biol. 15, 1566–1571 (2019).
91. Kalb, R. et al. Histone H2A mono-ubiquitination promotes histone H2A mono-ubiquitination. Nat. Struct. Mol. Biol. 21, 569–571 (2014).
92. Chu, C. et al. Polycomb repressive complex 1 recruits H2A ubiquitination in mouse preimplantation embryos. Mol. Cell 77, 1256–1278 e1267 (2020).
93. Cooper, S. et al. Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment. Cell Rep. 7, 1458–1467 (2016).
94. Cooper, S. et al. Jarid2 binds mono-ubiquitylated H2A lysis 119 to mediate cross talk between Polycomb complexes PRC1 and PRC2. Nat. Commun. 7, 15661 (2016).
95. Healy, E. et al. PRC2.1 and PRC2.2 Synergize to regulate Polycomb repressive activity in elongating transcription. Nature 523, 429–442 (2013).
96. Kasinath, V. et al. JARID2 and AEBP2 regulate PRC2 chromatin recruitment and polycomb domain formation. FEBS Lett. 583, 1002–1010 (2018).
97. He, J. et al. Kdm2b maintains murine embryonic stem cell identity. Cell Stem Cell 15, 970–980 e974 (2019).
98. Sanchez, C. et al. Proteomics analysis of Ring1B/Rnf2 complexes reveals critical role of CBX2. Science 366, 2884–2895 (2020).
99. Salomone, S. V. et al. Polycomb recruitment plays a central role in disseminating histone modifications. Nature 494, 84–94 (2013).
100. Davidovich, C. et al. H3K27me3 in mouse preimplantation embryos. Science 342, 3500–3510 (2013).
101. Ren, X., Vincenz, C. & Kerppola, T. K. Changes in the distributions and dynamics of polycomb repressive complexes during embryonic stem cell differentiation. Genes Dev. 26, 2884–2895 (2012).
102. Satijn, D. P. et al. RING1 is associated with the polycomb group protein complex and as a transcriptional repressor. Mol. Cell Biol. 17, 4105–4113 (2001).
103. Kim, C. A., Gingerly, M., Pilipa, R. M. & Bowie, J. U. The SAM domain of polycomb forms a helical sandwich. Nature 457, 452–457 (2009).
104. Boyle, S. et al. A central role for canonical PRC1 in shaping the 3D nuclear landscape. Genes Dev. 34, 531–543 (2020).
105. Isorno, K. et al. SAM domain polymerization links subnuclear clustering of PRC1 to gene silencing. Dev. Cell 26, 567–576 (2014).
106. Kundu, S. et al. Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. Dev. Cell 35, 432–446 (2017).
107. Wani, A. H. et al. Chromatin topology is coupled to Polycomb group protein subnuclear organization. Nat. Cell Biol. 23, 1029–1039 (2016).
108. Belousov, A. et al. Liquid chromatin Hi- C characterizes compartmentalized chromatin interaction dynamics. Nat. Genet. 55, 367–378 (2021).
109. Huang, M. K. et al. Distinct single particle tracking of PRC1 reveals a highly dynamic system with low target site occupancy. Nat. Commun. 12, 887 (2021).
110. Youmans, D. T., Schmidt, J. C. & Cech, T. R. Live-cell imaging reveals the dynamics of PRC2 and recruitment to chromatin by SL212 associated sub complexes. Nat. Genet. 22, 794–805 (2018).
111. Chen, Z. Y. et al. Live-cell single-molecule tracking reveals co-recognition of H2AK119me3 and DNA targets polycomb Cbx7-PRC1 to chromatin. eLife 9, e6766 (2020).
112. Vandenbunder, B. et al. PRC1 components exhibit different binding kinetics in Polycomb bodies. Biol. Cell 106, 111–125 (2014).
113. Cuadrado, A. et al. Specific contributions of cohesin-SA1 and cohesin-SA2 to TADs and polycomb domains in embryonic stem cells. Cell Rep. 27, 5050–5061 e5054 (2019).
114. Liu, J. A., Colignton, D., Samowo, H., Nabert, B. & Lee, J. T. Balancing cohesin eviction and retention prevents aberrant chromosomal interactions. Polycomb-mediated repression, and X-inactivation. Nat. Genet. 51, 1970–1976 (2019).
115. Pys, A. J. et al. Phase separation of polycomb-repressive complex 1 is governed by a charged disordered region of Fbxl10. Dev. Cell 55, 793–813 (2019).
116. Tatasevich, R. et al. Nuclear condensates of the Polycomb protein chromobox2 (CBX2) assemble through phase separation. J. Biol. Chem. 294, 1451–1465 (2019).
117. Eftendis, J. M., Kapoor, M. & Brangwynne, C. P. Epigenetic memory as a time integral over prior
The chromatin signaling pathway: diverse mechanisms of recruitment of histone-modifying enzymes and varied biological outcomes. 
Mol. Cell. 40, 689–701 (2010).

Klee, S. C. et al. Domain model explains propagation of histone modification and its inheritance of cellular identity. 
Nature. 528, 86–90 (2015).
genes in mouse ES cells. Nat. Cell Biol. 9, 1428–1435 (2007).

240. Brookes, E. et al. Polycomb associates genome-wide with a specific RNA polymerase II variant, and regulates the expression of ES cell-specific genes. J. Cell Sci. 120, 157–170 (2012).

241. Tsuiboi, M. et al. Ubiquitination-independent repression of RNA polymerase II during nuclear factor restation in the developing mouse neocortex. Dev. Cell 47, 758–772 e755 (2018).

242. Penczek, A. R, Kull, R., Frink, K. & Muller, J. Transcriptional repression by PRC1 in the absence of H2A mono ubiquitination. Genes Dev. 29, 1487–1492 (2015).

243. Pastoe, E. et al. Polycomb function during oogenesis is required for mouse embryonic development. Genes Dev. 26, 920–952 (2012).

244. Pasini, D., Braut, A., Jensen, M. R., Lazzerini Denchi, E. & Hein, K. Suz12 is essential for mouse development and for E2H2 histone methyla
transferase activity. EMBO J. 25, 4061–4071 (2006).

245. Endoh, M. et al. Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. Development 135, 1513–1524 (2008).

246. Cohen, I. et al. PRC1 preserves epidermal tissue integrity as an independently of PRC2. Genes Dev. 35, 55–60 (2019).

247. Chiarchia, F., Rossi, A., Jammula, S., Zanotti, M. & Paolocci, D. Polycomb regulates genistein and restricts secretory lineage commitment. EMBO J. 35, 2301–2314 (2016).

248. Chavira-Martinez et al. Polycomb complex PRC1 preserves histone stem cell identity by sustaining Wt1-betacatenin transcriptional activity. Cell Stem Cell 18, 91–110 (2021).

249. Pengelly, A. R., Copur, O., Jackle, H., Herzig, A. & Muller, J. A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. Dev. Cell 300–315 (2018).

250. Stepanik, V. A. & Harte, P. J. A mutation in the E(Z) Gene of Drosophila melanogaster acts as a dosage compensation factor and causes inappropriate silencing of active Polycomb target genes. Dev. Biol. 364, 249–258 (2012).

251. Bizgiri, Y. et al. EZH2 inhibitor efficacy in non- Hodgkin’s lymphoma does not require suppression of H3K27 demethylase. Chem. Biol. 24, 1–14 (2017).

252. Jadhav, U. et al. Replicational dilution of H3K27me3 in mammalian cells and the role of poised promoters. Mol. Cell 78, 141–151 e145 (2020).

253. Fan, H. et al. The polycomb domain within mammalian BAH1 connects H3K27me3 to Polycomb gene silencing. Nucleic Acids Res. 49, 4641–4655 (2020).

254. Zhao, D. et al. The BAH domain of BAHD1 is a histone H3K27me3 reader. Protein Cell 7, 222–226 (2016).

255. Xu, S. et al. Mutations in H3K27me3 reader BAHD1 represses Polycomb-dependent transcriptional repression in erythropoiesis. Blood 137, 155–167 (2021).

256. Abadie, S. et al. Polycomb repressive complex 2 methylates elongin A to regulate transcription. Mol. Cell. Biol. 88, 872–884 e876 (2017).

257. Ha, A. et al. PRC2 directly methylates CAT4 and represses its transcriptional activity. Genes Dev. 26, 37–42 (2012).

258. Zepeda-Martinez, J. A. et al. Parallel PRC2s/pRC1 and vPRC1 pathways silences lineage-specific genes and maintain self-renewal in mouse embryonic stem cells. Adv. Sci. 6, eaae6992 (2020).

259. Cohen, I. et al. Polycomb complexes redundantly maintain epidermal stem cell identity during development. Genes Dev. 35, 354–366 (2021).

260. Lee, M. et al. Polycomb complexes act redundantly to repress genomic repeats and genes. Genes Dev. 24, 265–276 (2010).

261. Inoue, A., Jiang, L., Lu, F., Suzuki, T. & Zhang, Y. Polycomb H3K27me3 controls methylation-independent imprinting. Nature 547, 419–424 (2017).

262. Beguelin, W. & Harte, P. J. Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. Development 135, 1513–1524 (2008).

263. O’Higlin, A. et al. MicroRNA regulation of Cbx7 mediates a switch of Polycomb orthologs during ESC differentiation. Cell Stem Cell 10, 35–46 (2012).

264. Matsui, S. I., Seon, B. K. & Sandberg, A. A. Disappearance of Polycomb and Polycomb-related protein A24 in mitosis: implications for molecular basis of condensation differentiation. Proc. Natl Acad. Sci. USA 76, 6586–6590 (1979).

265. Pena-Llopis, S. et al. BAP1 loss defines a new class of rRNA deficiency in the developing mouse neocortex. Dev. Cell 94, 751–759 (2020).

266. Scheuermann, J. C. et al. Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature 465, 243–247 (2010).

267. Tsubok, M. et al. Polycomb repression by counteracting widespread H2AK119ub1 deposition and chromatin condensation. Mol. Cell. Biol. 81, 1–16 (2021).

268. Fursova, N. A. et al. BAP1 constrains pervasive H2AK119ub1 to control the transcriptional potential of the genome. Genes Dev. 35, 749–770 (2021).

269. Ferrari, K. J. et al. Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. Mol. Cell. Biol. 53, 69–72 (2014).

270. Bryzovsky, M., Gu, Y. Z. & Staudt, L. M. Calibrating ChIP-seq with nuclear internal standards to measure histone modification density and activity genome-wide. Nat. Comms. 10, 3489 (2019).

271. Hoffjeld, J. W. et al. Non-core subunits of the PRC2 complex are collectively required for its target-specific epigenetic activity. Genes Dev. 26, 222–226 (2012).

272. Peters, A. A. et al. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Cell 181, 577–590 (2020).

273. Lee, H. G., Kahn, T. G., Simcox, A., Schwartz, Y. B. & Pirrota, G. Genome-wide activities of Polycomb complexes control pervasive transcription. Genome Res. 25, 1107–1115 (2015).

274. Du, Z. et al. Polycomb group proteins regulate chromatin architecture in mouse oocytes and early embryos. Mol. Cell 65, 1066–1079 (2021).

275. Gao, Z. et al. An AUTS2-Polycomb complex activates gene expression in the CNS. Nature 516, 349–354 (2014).

276. Frangini, A. et al. The aurora B kinase and the polycomb ring [B]2C combine to regulate active promoters in quiescent lymphocytes. Mol. Cell. Biol. 31, 147–161 (2021).

277. Crepec, C., Palau, A., Malinverni, R., Valero, V. & Buschbeck, M. A Cbx8-containing polycomb complex facilitates the transition to gene silencing during ES cell differentiation. PLoS Genet. 10, e1004851 (2014).

278. Macauara, S. et al. Polycomb directs timely activation of germline genes in spermatogenesis. Genes Dev. 31, 1693–1707 (2017).

279. Ichthy, H. et al. Classical activation of polycomb-repressed genes by ZF1. Nature 468, 1124–1128 (2010).

280. Cruz-Molina, S. et al. PRC2 facilitates the regulatory topology required for poised enhancer function during pluripotent stem cell differentiation. Cell Stem Cell 20, 689–705 e689 (2017).

281. Kondo, T. et al. Polycomb potentiates mes2 activation in midbrain by mediating interaction of the promoter with a tissue-specific enhancer. Dev. Cell 28, 94–101 (2014).

282. Pachano, T. et al. Orphan Cpg islands amply poised enhancer regulatory activity and determine target gene responsiveness. Nat. Genet. 50, 1098–1104 (2018).

283. Morey, L. et al. Polycomb regulates mesoderm cell fate specification in embryonic stem cells through activation and repression mechanisms. Cell Stem Cell 17, 300–315 (2015).

284. Mohn, F. et al. Lineage-specific polycomb targets and the mechanism of their recruitment and potential of neuronal progenitors. Mol. Cell. 30, 755–766 (2008).

285. Xie, R. et al. Polycomb chromatin remodeling mediated by polycomb proteins orchestrates pancreatic differentiation of human embryonic stem cells. Stem Cell Rep. 20, 1233–1257 (2021).

286. Hughes, A. L., Kelley, J. R. & Rose, J. R. Understanding the interplay between CpG island-associated gene promoters and H3K4 methylation.