Mol Biol Rep (2012) 39:5575–5588
DOI 10.1007/s11033-011-1362-5

Identification and characterization of Polycomb group genes in the silkworm, *Bombyx mori*

Zhiqing Li · Tsuneyuki Tatsuke · Kosuke Sakashita · Li Zhu · Jian Xu · Hiroaki Mon · Jae Man Lee · Takahiro Kusakabe

Received: 24 July 2011 / Accepted: 12 December 2011 / Published online: 21 December 2011

Abstract Polycomb group (PcG) proteins are involved in chromatin modifications for maintaining gene repression that play important roles in the regulation of gene expression, tumorigenesis, chromosome X-inactivation, and genomic imprinting in *Drosophila melanogaster*, mammals, and even plants. To characterize the orthologs of PcG genes in the silkworm, *Bombyx mori*, 13 candidates were identified from the updated silkworm genome sequence by using the fruit fly PcG genes as queries. Comparison of the silkworm PcG proteins with those from other insect species revealed that the insect PcG proteins shared high sequence similarity. High-level expressions of all the silkworm PcG genes were maintained through day 2 to day 7 of embryogenesis, and tissue microarray data on day 3 of the fifth instar larvae showed that their expression levels were relatively low in somatic tissues, except for *Enhancer of zeste* (*E(Z)*). In addition, knockdown of each PRC2 component, such as *E(Z)*, *Extra sex combs (ESC)*, and *Suppressor of zeste 12 (SU(Z)12)*, considerably decreased the global levels of H3K27me3 but not of H3K27me2. Taken together, these results suggest that insect PcG proteins are highly conserved during evolution and might play similar roles in embryogenesis.

Keywords *Bombyx mori* · Polycomb group genes · Comparison analysis · Expression patterns · RNA interference

Introduction

Polycomb group (PcG) proteins were originally identified as repressors to regulate *Hox* gene expression in fruit fly (*D. melanogaster*) [1, 2], and are then found to distribute widely in many other species, including mammals [3] and plants [4]. To date, PcG proteins have been implicated in various biological processes, such as the cell cycle program, tumorigenesis, chromosome X-inactivation, genomic imprinting, and so on [5–8].

In general, there are three kinds of mutually cooperative PcG complexes: Polycomb repressive complex 1 (PRC1), Polycomb repressive complex 2 (PRC2), and Pleiohomeotic repressive complex (PhoRC) [9–11]. The proposed model for a transcriptional repression mechanism by PcG proteins is as follows: the DNA binding protein presented in Pho-RC recognizes Polycomb responsive elements (PREs) in the target genes and then recruits the PRC2 complex to these loci. The PRC2 complex catalyzes the tri-methylation at lysine 27 of histone 3 (H3K27me3). Subsequently, the PRC1 complex recognizes H3K27me3 and mediates transcriptional repression [3, 12]. The detailed repression mechanism might be more complicated and varied among different species. Recently, genome-wide identification of PcG target genes in fruit fly and mammals have provided the global profiles of PcG binding sites [13–15]. However, how PcG proteins are recruited to these targets remains largely unknown.

Fruit fly PhoRC core components include Pleiohomeotic (Pho) and SCM-related gene containing four MBT domains (Sfmbt) [11, 16]. It is reported that Pho is the only

Electronic supplementary material The online version of this article (doi:10.1007/s11033-011-1362-5) contains supplementary material, which is available to authorized users.

Z. Li · T. Tatsuke · K. Sakashita · L. Zhu · J. Xu · H. Mon · J. M. Lee · T. Kusakabe (✉)
Laboratory of Silkworm Science, Kyushu University Graduate School of Bioresource and Bioenvironmental Sciences, Hakozaki 6-10-1, Fukuoka 812-8581, Japan
e-mail: kusakabe@agr.kyushu-u.ac.jp

sequence-specific DNA-binding protein essential for PcG targeting and silencing [11]. Yin and Yang 1 (YY1), the mammalian ortholog of Pho, has also been involved in the recruitment of PRC2 to the target locus [17], whereas Sfmbt can interact with Pho and contribute to its binding [16]. Indeed, genome-wide chromatin immunoprecipitation (ChIP) assays showed that 50% of regions were co-localized by both Pho and Sfmbt proteins in fruit fly larvae [6]. Fruit fly PRC2 contains three major subunits: enhancer of zeste (E(z)), Extra sex combs (Esc), and Suppressor of zeste 12 (Su(z)12) [9]. The PRC2 complex catalyzes the methylation of H3K27 [18], which is an epigenetic mark for maintaining the repression state of PcG target genes. E(z) is a catalytic subunit with the SET domain [9], and the other two proteins are reported to be essential for its catalytic activity [19]. Biochemically purified fruit fly PRC1 is composed of Polycomb (Pc), Polyhomeotic (Ph), Sex combs extra (Sce), and Posterior sex combs (Psc) [10]. The PRC1 complex can recognize chromatin marked with H3K27me3 through the CHROMO domain of Pc protein, and finally results in transcriptional silencing [18]. In addition, Ring1b, associated with Bmi1, both of which are mammalian orthologs of Sce and Psc, respectively, catalyzes ubiquitination of histone H2A at lysine 119 (H2AK119ub), and this modification is also involved in transcriptional repression [20, 21].

Mammalian PcG complexes have almost all of the known fruit fly core orthologs, many of which contain different paralogs [3]. In contrast, few PcG proteins have been found in plants. In *Arabidopsis*, all three PRC2 components exist, while other genes have not yet been identified [4, 22]. The loss of the plant PcG gene causes early flowering and mild homeotic transformation phenotypes in flowers [23]. This implicates the PcG genes in plants in repressing flowering *Hox* gene expression.

To extend the investigation of the function of PcG in other species, we have chosen the silkworm, *B. mori*, which possesses a large *Hox* gene family [24]. Furthermore, the silkworm is known to be a model organism for Lepidoptera with holocentric chromosomes [25]. Compared with monocentric organisms, holocentric organisms would have to tolerate chromosome breakage or rearrangement, so as to maintain the integrity of their genome. Therefore, it is worth investigating whether the PcG proteins from holocentric organisms play similar roles in gene regulation demonstrated in fruit fly and mammals, and also how this epigenetic state can spread to the daughter cells in holocentric organisms. In the present study, the molecular composition of PcG genes in silkworm has been characterized by an in silico silkworm genome search [26] using fruit fly PcG genes as queries. As a result, 13 PcG genes were identified, including *Sex combs on midleg* (*SCM*), *Polycomb-like* (*PCL*), *Enhancer of Polycomb* (*E(PC)*), and *Additional sex combs* (*ASX*). Meanwhile, we also identified orthologous genes from *Aedes aegypti*, *Tribolium castaneum*, and *Apis mellifera* based on their genome data [27–29]. To obtain insight into their roles in silkworm embryonic and postembryonic development, their embryonic stage or tissue-specific expression was analyzed by reverse transcription polymerase chain reaction (RT-PCR) or the microarray data previously reported by Xia et al. [30], respectively. Moreover, knockdown of the PRC2 components by RNA interference (RNAi) led to a global decrease in H3K27me3 but not in H3K27me2.

Materials and methods

Biological materials

The silkworm strain p50T, stocked in our laboratory, was reared under the standard condition. To get a full view of PcG gene expression patterns in the silkworm embryogenesis, total RNAs were isolated from silkworm eggs at the stages indicated and the samples were stored in liquid nitrogen until use.

The silkworm normal BmN4 cell line (a gift from Dr. Chisa Aoki, Kyushu University Graduate School) and a transgenetic BmN4 cell line overexpressing *Caenorhabditis elegans* SID-1 (BmN4-SID1) with an ability to uptake double stranded RNA (dsRNA) into cells [31], were cultured at 27°C in IPL-41 medium (Sigma) supplemented with 10% fetal bovine serum (FBS) (Gibco).

Identification of the silkworm Polycomb group genes

The reported fruit fly PcG proteins were downloaded from NCBI (http://www.ncbi.nlm.nih.gov/). We used these sequences as our queries to perform BLASTP search against the updated silkworm genome sequence [26] and identified 13 candidates. These candidate genes were further verified by domain prediction using SMART (http://smart.embl-heidelberg.de/smart/set_mode.cgi) and PROSITE (http://expasy.org/prosite/). The identification for each gene is listed in Table 1.

Sequence alignment and phylogenetic analysis

To get a comparative analysis among different insect species including *D. melanogaster*, *A. aegypti*, *T. castaneum*, and *A. mellifera*, the annotated genes in these species were obtained from NCBI. On the other hand, we also tried to identify their orthologous genes from the public genome data. Except for *PHO* and *PCL* orthologs in *A. aegypti*, all the PcG genes were found in *T. castaneum* and *A. mellifera*

Table 1 Summary of the silkworm PcG genes

PcG	Gene	Gene ID	Protein length (aa)	Scaffold	Chr.	Domain	EST	Probe	Accession no. of fruit fly orthologs
PRC1	*BmPC*	BGIBMGA006904	302	nscaf2860	10	CHROMO	0	sw14515	NP_524199.1
	BmPH	BGIBMGA000509/10	1024	nscaf1690	1	ZF_FCS/SAM	8	sw14299	CAA45211.1
	BmSCE	BGIBMGA006985	377	nscaf2865	17	ZF_RING	14	sw15420	AAF56737.1
	BmPSC	BGIBMGA006194	1082	nscaf2847	4	ZF_RING	5	sw06581	NP_523725.2
PRC2	*BmE(Z)*	BGIBMGA014476	732	nscaf2809	UN	SET	4	sw15702	NP_001137932.1
	BmESC	BGIBMGA006325	411	nscaf2852	6	WD40	5	sw12637	NP_477431.1
	BmSU(Z)12	BGIBMGA011842	746	nscaf3031	11	ZF_C2H2	1	sw14823	NP_652059.1
PRP	*BmPHO*	BGIBMGA011178/9	388	nscaf3026	23	ZF_C2H2	0	sw08208	AAM18016.1
	BmSFMBT	BGIBMGA010551	709	nscaf2993	12	MBT/SAM	3	sw00111	NP_723786.1
	BmSCM	BGIBMGA011089	641	nscaf3015	23	ZF_FCS/MBT/SAM	1	sw13820	AAF54419.2
	BmPCL	BGIBMGA007840	779	nscaf2888	15	ZF_PHD	3	sw00723	NP_476672.1
	BmE(PC)	BGIBMGA009662/3	1043	nscaf2964	2	NO	0	sw02453	AAC64271.1
	BmASX	BGIBMGA006166	994	nscaf2847	4	NO	5	sw07964	NP_725398.1

as well as *B. mori*. The accession numbers are presented in Table S1. Sequences were aligned using ClustalX [32]. The identity of amino acid sequence for each protein was determined using Vector NTI Advance 11 (Invitrogen). Neighbor-joining (NJ) trees were constructed by MEGA 4 under 1,000 replicates [33].

Embryonic transcript analysis

Total RNAs isolated from the silkworm embryos were subjected to reverse transcription using the ReveTra Ace cDNA synthesis kit according to manufacturer's instructions (TOYOBO). The silkworm *glyceraldehyde-3-phosphate dehydrogenase* (*BmGAPDH*) gene was used as an internal control. Gene-specific primers for PcG genes are listed in Table S2. 10 μl of amplified products were separated by electrophoresis on 1% agarose gel and stained with ethidium bromide. All the cDNAs amplified were further cloned into pLits vector and their nucleotide sequences were confirmed by DNA sequencing.

Larvae tissues microarray analysis

By using the genome-wide microarray data with 22,987 70-mer probes, transcriptional levels of PcG genes were analyzed in ten silkworm tissues, such as ovary, testis, head, integument, anterior/median silk gland (A/MSG), posterior silk gland (PSG), midgut, fat body, malpighian tubule, and hemocyte from day 3 of the fifth instar larvae. The microarray data was downloaded from the website (http://silkworm.swu.edu.cn/silkdb/). According to the previous standard [30], the signal intensity of one gene exceeded a value of 400 signal intensity units after subtracting the background was considered to express in tissue. The probes for each PcG gene used in this research are shown in Table 1. The signal intensity values were analyzed by Cluster software and visualized using Java TreeView program [34].

RNA interference

Double stranded RNAs specific for *BmE(Z)*, *BmESC* and *BmSU(Z)12* were produced as follows: DNA fragments for each gene were amplified from total RNA of the silkworm BmN4 cells by RT-PCR using the primers in Table S2 and cloned into a *Stu*I site of pLits vector to be possessed T7 RNA polymerase promoters in both terminuses after inserting the target cDNAs. After cloning and sequencing, these plasmids were used as templates to amplify the cDNAs by a T7 promoter primer (Table S2). The PCR products were purified and then the dsRNAs were synthesized in the reaction solution according to our previous study [35]. Meanwhile, the dsRNA for *enhanced green fluorescent protein* (*EGFP*) gene was also synthesized using the same procedure and used as a control in this study. All the dsRNAs were quantified by NanoDrop (Thermo Scientific).

For RNAi experiment, the BmN4-SID1 cells were previously cultured in 24-well or 6-well plates at a cell density of 0.5×10^5 or 2.0×10^5, respectively, in IPL-41 medium with 10% FBS. Each dsRNA for *EGFP*, *BmE(Z)*, *BmESC* and *BmSU(Z)12* was added to the medium with a final concentration of 0.5 μg/ml. Seven days after incubation with dsRNAs, the cells were harvested to extract RNA and histone proteins.

Histone purification

Histones of cells were purified using the high-salt extraction method as previously described [36]. Briefly, cells

were lysed in extraction buffer (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol) in the presence of protease inhibitor, phosphatase inhibitor, histone deacetylase inhibitor and 0.2% NP40 for 10 min on ice. After centrifugation at 6,500×*g* for 5 min, the pellets were resuspended in extraction buffer (without NP-40 but containing other inhibitors) and centrifuged at 6,500×*g* for 5 min. The nuclei pellets resuspended in no-salt buffer (3 mM EDTA, 0.2 mM EGTA) were incubated on rotator at 4°C for 30 min. Removing the supernatant by centrifugation at 6,500×*g* for 5 min and the pellets containing histones were further dissolved in high-salt buffer (50 mM Tris–Cl pH 8.0, 2.5 M NaCl, 0.05% NP40). After incubation on rotator at 4°C for 30 min and centrifugation at 16,000×*g* for 10 min, the supernatant was recovered as histone solution and used for Western blotting.

Western blotting

Purified histones were resolved on a 15% SDS–PAGE and transferred onto polyvinylidene difluoride (PVDF) membrane (Millipore). The membranes were blocked in 5% skim milk for 1 h and incubated in a primary antibody against H3 (H0164, Sigma), H3K27me2 (07-452, Upstate), H3K27me3 (17-622, Upstate), H3K27me1 (39377, Active motif), or H3K27ac (39133, Active motif) for 1 h. After washing, the membranes were incubated with rabbit anti-mouse IgG conjugated with alkaline phosphatase (Sigma) as a secondary antibody and visualized using CDP-Star chemiluminescent substrate (Tropix).

Immunofluorescence staining

Cultured silkworm cells were fixed with 3.7% formaldehyde in phosphate buffered saline (PBS) for 10 min and permeabilized with 0.1% Triton X-100 in PBS for 5 min. The fixed cells were blocked in 5% skim milk for 1 h and incubated with primary antibodies against H3K27me3 or H3K27me2 for 1 h. After extensive wash, the cells were incubated with a Cy3-conjugated goat anti-rabbit IgG (GE Healthcare) for 1 h, and the nuclei were stained by 4′, 6-diamidino-2-phenylindole (DAPI) (Invitrogen) for 10 min. The cells were washed and mounted with DABCO (Sigma). Images were acquired using Biozero BZ-8000 microscope (KEYENCE).

Results

Identification of Polycomb group orthologs in silkworm

To identify silkworm PcG genes, we referred to fruit fly PcG protein sequences and performed BLASTP analysis against the silkworm genome database [26], and finally identified 13 candidates. These genes were further analyzed by comparing the functional domains conserved in the known PcG genes. Based on the results, these genes were classified into three groups, as were the fruit fly PcG complexes. Core components of PRC1 and PRC2 biochemically purified in fruit fly were also presented in silkworm and named silkworm PRC1 and PRC2. In a very early report, Scm was considered a component of PRC1 [10]. Recent evidence, however, indicates that it should form a new complex [37, 38]. Therefore, we have grouped this protein together with Pho, Sfmbt, Pcl, E(Pc), and Asx into Polycomb-related proteins (PRP).

Genomic information on all the silkworm PcG genes is summarized in Table 1, and the predicted domain architecture for each protein is shown in Fig. 1 (the detail is as follows). Importantly, in this study we have revised some predicted silkworm PcG genes, such as *BmPH*, *BmPHO*, and *BmE(PC)*. The predicted genes registered in the public database lack some structural domains compared with other insect orthologs. To identify the correct nucleotide sequences of these genes, we rescanned the genome sequence and reassembled them. As shown in Table 1, *BmPH*, *BmPHO*, and *BmE(PC)* should be composed of BGIBMGA000509 and BGIBMGA000510, BGIBMGA011178 and BGIBMGA011179, and BGIBMGA009662 and BGIBMGA009663, respectively. To confirm this prediction, we designed the primers spanning these reassembled genes, amplified specific fragments, and determined their DNA sequences.

Unlike the genes listed above, we could not find the homologs of *Suppressor of zeste 2* (*SU(Z)2*), *Extra sex combs-like* (*ESCL*), and *Pleiohomeotic-like* (*PHOL*) in the silkworm genome.

Comparative analysis of insect Polycomb group genes

Polycomb repressive complex 1: Pc, Ph, Sce, and Psc

Pc protein contains a CHROMO domain that can recognize H3K27me3 and is essential for maintaining the silencing state of *Hox* genes during development [2, 18]. As shown in Fig. 2a, Pc amino acid sequences from different insect species have high sequence similarity, especially in the CHROMO domain. Three β sheets and one α helix are presented in this domain based on the previous crystallographic analysis [39, 40]. Also, the insect Pc proteins share three conserved caging aromatic residues that interact with H3K27me3 except for a small difference at the first residue: the Phe residue in silkworm is identical to that in mammals [41]. Moreover, the phylogeny of insect PC genes reveals a corresponding relationship with the classical taxonomic divergence of these insect species [42].

Fig. 1 Domain prediction for each silkworm PcG protein. The structures presented in silkworm PcG proteins are consistent with fruit fly orthologs

Ph can interact with Pc and Scm (another SAM protein in PcG complexes) by its SAM domain [43]. As shown in Fig. 2b, five insects each have a conserved SAM domain, with each domain consisting of five α helixes [44]. In addition to the SAM domain, Ph contains a ZF_FCS domain (Fig. 1), and this domain presented in sop-2, a *C. elegans* PcG protein, is shown to have a binding ability to RNA; similar RNA binding activity is confirmed in mouse Ph homolog Rae28 [45].

The identical C3HC4-type ZF_RING domains found in insect Sce proteins suggest that their conserved functions are probably involved in protein–protein interactions (Fig. 2c) [46].

Similar to Sce, Psc proteins from different insects contain a C3HC4-type ZF_RING domain with high identity (Fig. 2d). As mentioned above, mammalian Sce and Psc orthologs are involved in H2AK119ub.

Polycomb repressive complex 2: E(z), Esc, and Su(z)12

The conserved SET domain of E(z) provides histone methyltransferase (HMTase) activity for H3K27. Based on the structure resolved in the plant SET domain by crystallization [47], the secondary structures of insect SET domains are predicted to contain twelve β sheets and seven amino acid residues necessary for binding to the lysine substrate (Fig. 3a).

Esc has important roles in the propagation of *Hox* gene silencing by forming a complex with E(z) through its WD repeats [48]. Structural analysis shows five WD40 repeats in insect Esc proteins (Fig. 3b). Amino acid sequence alignments of these domains exhibit that the third and fourth WD40 repeats have higher identities than the others.

As shown in Fig. 3c, insect Su(z)12 proteins contain a classical ZF_C2H2 structure that is involved in the regulation of gene expression [46]. The two repeats of Cys and His residues are conserved in these insects.

Polycomb-related proteins: Pho, Sfmbt, Scm, Pcl, E(Pc), and Asx

In fruit fly, PcG complex-mediated silencing is carried out by multiple DNA sequences that are called PREs. The exceptional DNA binding ability of Pho is believed to be crucial for the recognition and recruitment of PcG complexes to the majority of target genes [11, 49]. The structure of Pho includes four ZF_C2H2 domains likely serving as a platform for interacting with specific DNA sequences. All insect Pho proteins studied here have displayed highly conserved structures, especially in the second and third ZF_C2H2 domains, which share 100% identities (Fig. 4a).

Interaction between Sfmbt and Pho provides a bridge between the PREs and PRC complexes [16]. Further structure analysis revealed that the Sfmbt contains four MBT repeats and one SAM domain (Fig. 4b). The four

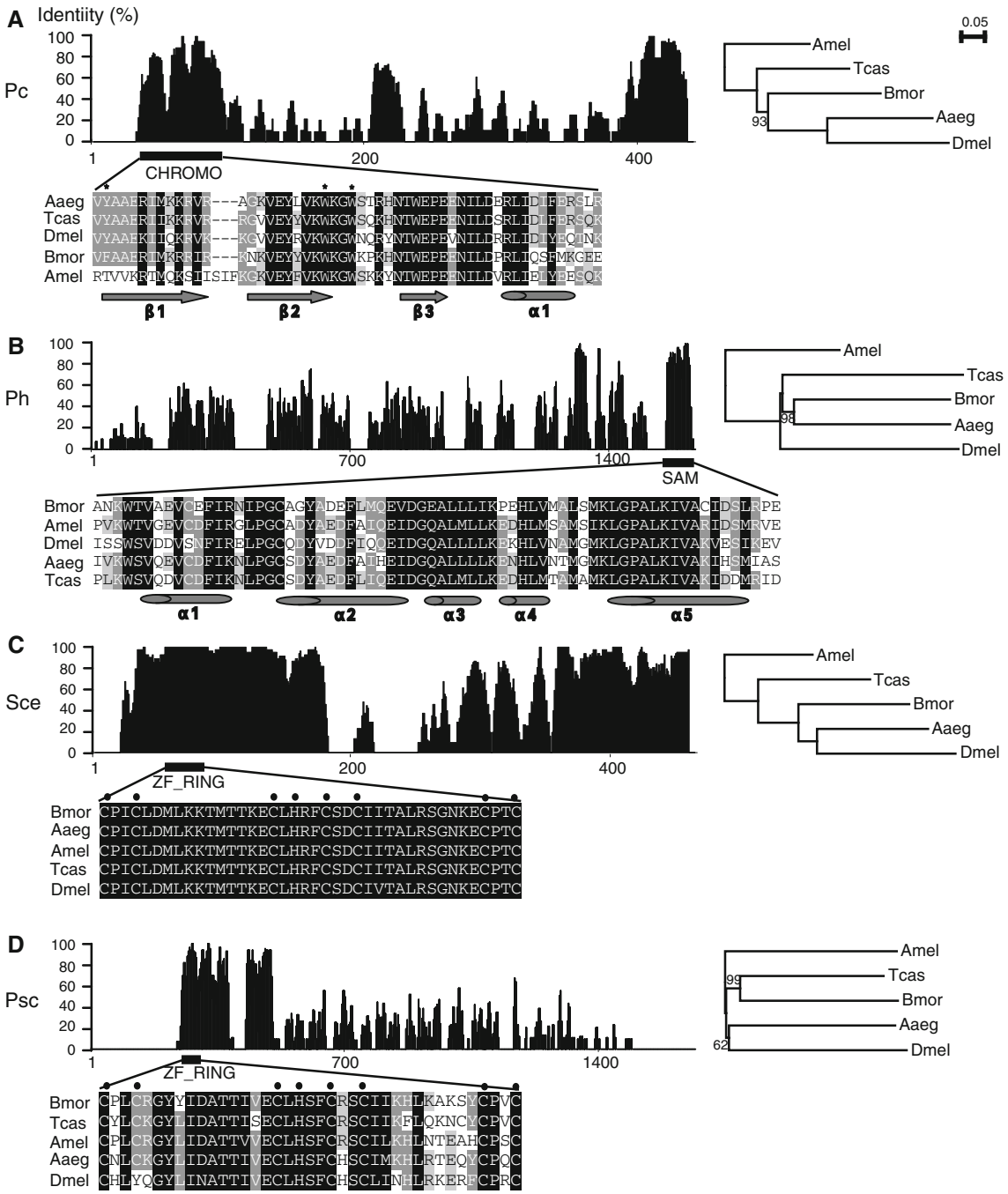


Fig. 2 Sequence alignment and phylogenetic relationship of insect PRC1 complex. The percentage identities at each amino acid position of insect Pc (**a**), Ph (**b**) Sce (**c**) and Psc (**d**) were columned and a corresponding NJ tree for each gene was constructed using the full-length alignment. The numbers on the branches indicated bootstrap confidence values for 1,000 replications (100% was not shown). The conserved domains in each protein were further analyzed. The *asterisks* above the CHROMO domain of Pc represented the caging aromatic residues that were involved in the binding with histone methyl-lysine and the typical residues of C3HC4 structures in Sce and Psc were presented above the sequence alignment by *solid circle*. Secondary structure elements were illustrated under the sequence alignment of CHROMO and SAM domains. Bmor, Dmel, Aaeg, Tcas, and Amel represented *B. mori*, *D. melanogaster*, *A. aegypti*, *T. castaneum*, and *A. mellifera* respectively

MBT repeats from five insect species share 67, 73, 61, and 67% identities and 78, 87, 73, and 77% similarities, respectively.

Scm comprises multiple functional domains, such as two ZF_FCS, two MBT, and one SAM domains (Figs. 2, 4c). These conserved domains presented in Scm suggest the

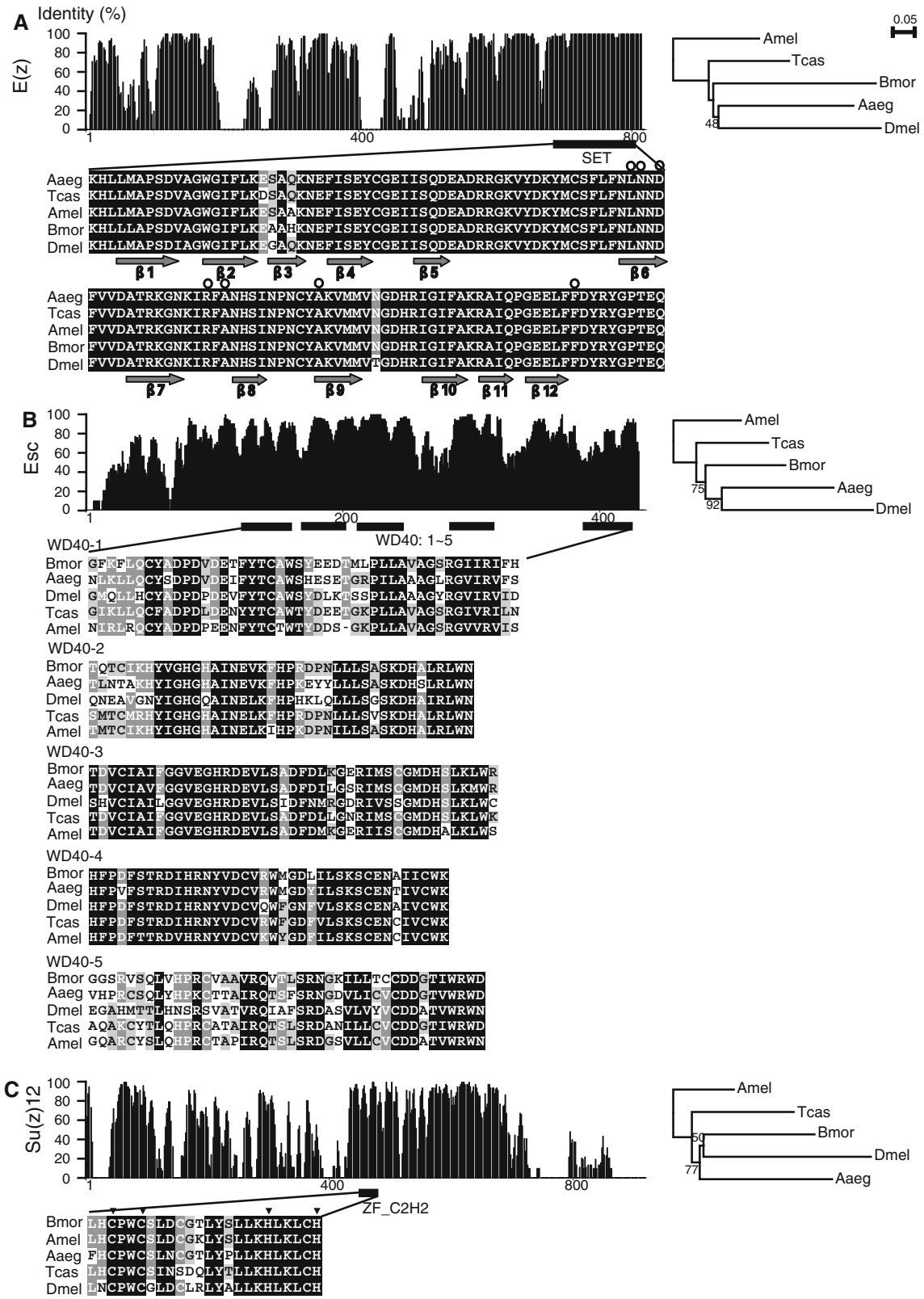
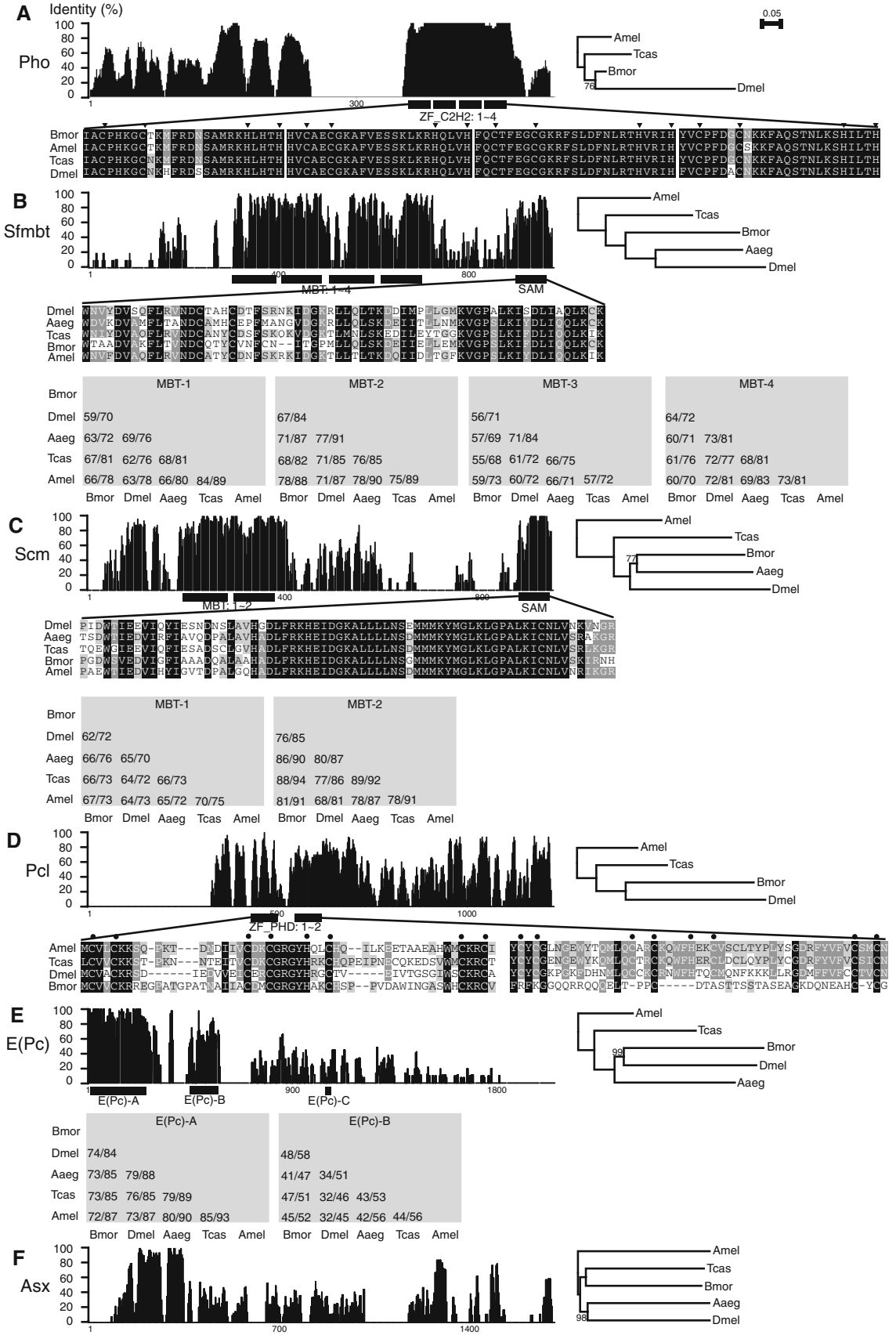


Fig. 3 Sequence alignment and phylogenetic relationship of insect PRC2 complex. The percentage identities and NJ trees of insect E(z) (**a**), Esc (**b**) and Su(z)12 (**c**) were shown. The SET domain of E(z) displayed the predicated twelve β sheets and binding sites with histone lysine (*open circle*) under and above sequence alignment respectively. *Black arrowhead* above the sequence of Su(z)12 represented the residues of C2H2 structure


A
Identity (%)
Pho
0.05
Amel
Tcas
Bmor
Dmel
ZF_C2H2: 1~4
B
Sfmbt
Aaeg
MBT: 1~4
SAM
MBT-1
MBT-2
MBT-3
MBT-4
C
Scm
MBT: 1~2
D
Pcl
ZF_PHD: 1~2
E
E(Pc)
E(Pc)-A
E(Pc)-B
E(Pc)-C
F
Asx

◀**Fig. 4** Sequence alignment and phylogenetic relationship of insect PRP group. The percentage identities and NJ trees of insect Pho (**a**), Sfmbt (**b**), Scm (**c**), Pcl (**d**), E(Pc) (**e**) and Asx (**f**) were shown. *Black arrowhead* and *solid circle* above the sequences of Pho and Pcl represented C2H2 residues and C4HC3 residues, respectively. Pair comparisons among MBT domains from five insects were shown in percent identity/similarity. E(Pc)-A and E(Pc)-B regions in E(Pc) were also compared except for E(Pc)-C region due to its low similarity

various and important roles in PcG-mediated repression in insect species.

Pcl can interact with the PRC2 complex and enhance its HMTase activity. The lack of Pcl, to some extent, reduces the levels of H3K27me3 [50]. However, the ZF_PHD domains of insect Pcl proteins are moderately conserved (Fig. 4d).

To date, the function of E(Pc) is not clear. A comparison of amino acid sequences reveals the conserved N-terminus of insect E(Pc) proteins but not any known domain structures. Previous studies for fruit fly E(Pc), however, have defined three conserved regions, called E(Pc)-A, E(Pc)-B, and E(Pc)-C [51], as shown in Fig. 4e. Except for the E(Pc)-A region with higher identities, the other two regions are diverse, especially in the small E(Pc)-C region.

Fruit fly Asx is reported to form a complex named Polycomb repressive deubiquitinase (PR-DUB) with calypso protein at PcG target genes [52]. Although there are no known functional domains in Asx (Fig. 4f), it exhibits a very high sequence similarity in their N-terminus among different species.

Developmental expression profiles of Polycomb group mRNAs

PcG proteins have been identified primordially as a group of regulators for *Hox* gene expression, which is required for the correct initiation of segmental determination during early embryonic development in fruit fly [1, 2]. To observe the expression patterns of silkworm PcG genes during embryogenesis, total RNA at each time point indicated in Fig. 5 was used as a template for RT-PCR. As shown in Fig. 5, PcG gene expressions were detected throughout the embryonic stage, which likely suggested their function in the development of silkworm embryos. Significantly, almost all these genes had a tendency to reduce the expressions at the end of embryogenesis, just before the hatching, whereas *BmE(Z)* and *BmASX* had a stable expression level.

Furthermore, silkworm microarray data available from the public database allowed us to analyze the tissue-specific expression profiles of PcG genes [30]. The oligonucleotide probes representing the entire silkworm PcG genes are listed in Table 1, and a heat map is created based on signal intensity (Fig. 6). Overall, the expression levels of PcG genes were relatively low in various tissues on day 3 of the fifth instar larvae. However, it was interesting that *BmE(Z)*

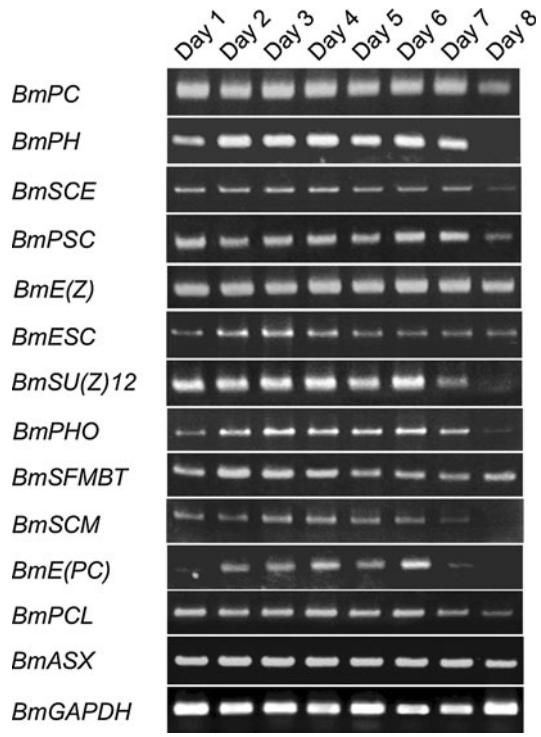


Fig. 5 Developmental expression patterns of silkworm PcG genes during embryogenesis. RT-PCR was performed to detect the expression patterns of PcG genes using specific primers and *BmGAPDH* gene was used as internal control at each time point

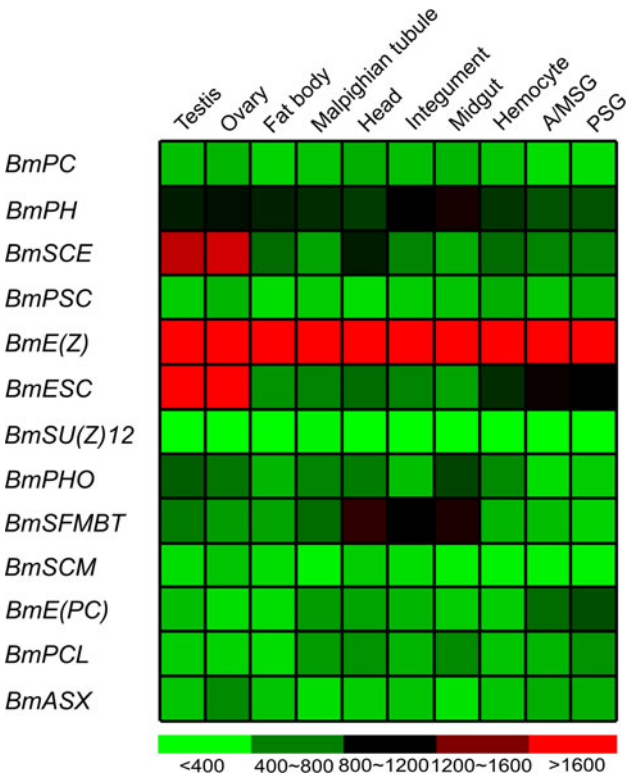


Fig. 6 Multiple organic expression patterns of silkworm PcG genes on day 3 of the fifth instar larvae. The signal intensities representing the expression levels were shown by different *color grades*

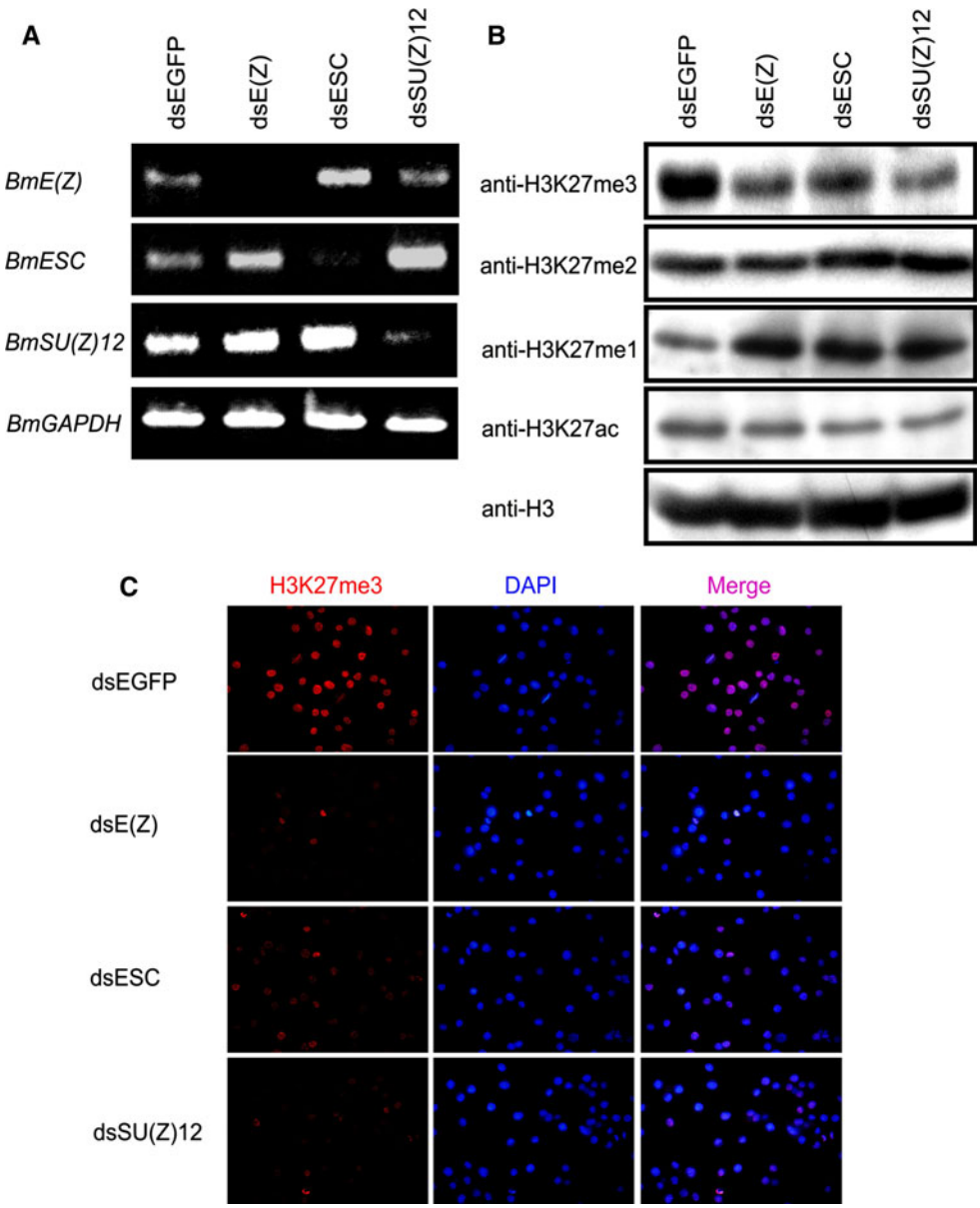


Fig. 7 Knockdown of the PRC2 components affected the methylation profiles of H3K27. **a** RT-PCR analysis of PRC2 knockdown and control cells. *BmGAPDH* was used as control for normalization in the semi-quantitative RT-PCR. **b** Western blotting analysis of histones purified from the knockdown and control cells with antibodies specific for H3K27me3, H3K27me2, H3K27me1, and H3K27ac. Antibody against H3 was used as a loading control. **c** Knockdown and control cells were processed for immunofluorescence staining with H3K27me3 antibody and the nuclei were counterstained with DAPI

was highly expressed in all the tissues. Moreover, *BmSCE* and *BmESC* exhibited higher expression in reproductive organs, the ovary and testis, than the other tissues examined.

Knockdown of the PRC2 components affecting the methylation profiles of H3K27

We now analyze whether the silkworm PRC2 components can catalyze the methylation on the H3K27 site. To do this, we performed the knockdown of *BmE(Z)*, *BmESC*, and *BmSU(Z)12* by specific dsRNA for each gene to study their effects on the methylation patterns of H3K27 in the silkworm BmN4-SID1 cells. As demonstrated in Fig. 7, the transcripts of *BmE(Z)*, *BmESC*, and *BmSU(Z)12* mRNAs were drastically reduced after RNAi induction (Fig. 7a). To evaluate their effects on the methylation of H3K27 in vivo, histones purified from the PRC2 knockdown cells were subjected to Western blot analysis using the modification-specific antibodies against H3K27me3, H3K27me2, H3K27me1, and H3K27ac. Knockdown of *BmE(Z)* significantly decreased H3K27me3, whereas no changes were observed in H3K27me2 compared to the control (Fig. 7b). Interestingly, the H3K27me1 was greatly increased in the knockdown cells. Meanwhile, we also detected the acetylation of H3K27, it was unchanged. Also, we obtained the same results by the knockdown of the other two components, *BmESC* and *BmSU(Z)12*. In agreement with the results of Western blot analysis, immunofluorescence staining showed that H3K27me3 signals were decreased in knockdown cells (Fig. 7c).

H3K27me3 maintenance through mitosis

Transcription repression mediated by PcG complexes is associated with the methylation of H3K27 on the loci of target genes, and H3K27me3 is also a mark for facultative heterochromatin [53, 54], which is involved in the changes in chromatin structure and compaction. Although H3K27me3 has been reported to localize mainly on the pericentric heterochromatin at interphase and mitosis in mammalian cells [55], to our knowledge this methylation state has not been analyzed definitely during different phases of mitosis in holocentric organisms. To examine whether H3K27me3 is maintained through the mitotic stage, we performed immunostaining using a specific antibody for H3K27me3 in the silkworm BmN4 cells and observed the modification profile in different cell phases, including prophase, metaphase, anaphase, and telophase. As shown in Fig. 8, H3K27me3 appeared across the mitotic chromatin, although only a very weak signal was observed in metaphase and anaphase. Presumably, the highly condensed chromatin in these phases prevents the access by the antibody, resulting in a weaker signal than other phases. To test this hypothesis, we performed the knockdown of the *BmCDC27* gene, a homolog that is essential for the metaphase to anaphase transition [56]. Western blot analysis using the H3K27me3 antibody revealed no difference between the metaphase-arrested cells and control cells (data not shown).

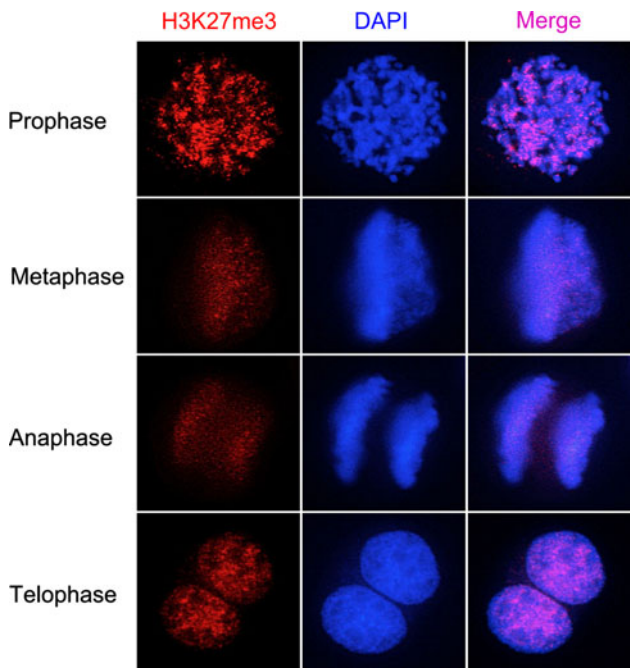


Fig. 8 H3K27me3 pattern on chromatin during mitosis. Immunofluorescence staining of mitotic chromatin in the silkworm BmN4 cells was performed by using the H3K27me3 antibody and the nuclei were counterstained with DAPI

Discussion

PcG complexes have been implicated as important chromatin modifiers, the functions of which have extended their initial identification as regulators for homeotic transformation. Owing to their conserved evolution among various organisms, a large number of species are subjected to targets in order to study the regulatory mechanism of PcG complexes. In this study, we identified the PcG genes from the Lepidopteran model insect silkworm. In our analysis, 13 PcG orthologous genes were found in the silkworm genome and divided into three groups: PRC1, PRC2, and PRP. However, the homologs of three fruit fly genes—*SU(Z)2*, *ESCL*, and *PHOL*—were not identified in silkworm. Also, three other insects—*A. aegypti*, *T. castaneum*, and *A. mellifera*—have no apparent homologs of these three genes, except for the *SU(Z)2* homolog in *A. aegypti* [57]. Previous studies reveal that Su(z)2, Escl, and Phol have roles similar to those of Psc, Esc, and Pho, respectively, and suggest the redundant functions of these genes in fruit fly [57–61]. Further, a comparative analysis of PcG proteins from five insect species showed that these PcG proteins occupied structurally identical domains. It is noteworthy that PcG proteins have various and important domains, such as CHROMO, MBT, SET, SAM, ZF, and so on, that have been reported to be involved in protein–protein, protein–DNA, and protein–RNA interactions. A large variety of structures found in PcG proteins would assign PcG complexes to execute multiple functions during development.

In fruit fly, PcG complexes have crucial roles in early embryogenesis [1, 2]. Similarly, silkworm PcG genes were widely expressed during the entire embryonic period, and the transcripts were reduced to low levels along with the accomplishment of embryogenesis. Importantly, before the shortening stage (around day 4 in this study), a key stage for segmental determination, PcG genes have reached high and stable expression levels, which indicated a regulatory role of PcG in this process. The further experiments, however, need to uncover the relationship between PcG regulation and segmental determination. After organ generation was complete (day 7), the transcripts began to decrease. These expression patterns were comparable to those of fruit fly. Moreover, we analyzed the expression profiles in postembryonic development based on microarray data. Although the majority of PcG genes gave weak expression signals, *BmE(Z)* had high expression levels in all tissues tested. Other than its catalytic activity on the methylation of H3K27, BmE(z) might have other unknown

functions, that needs further study. In addition, *BmSCE* and *BmESC* were more highly expressed in reproductive organs than in other tissues, likely suggesting that they play important roles in the development or maturation of these organs.

We further explored the function of PcG genes in silkworm by using RNAi. Down-regulation of the three components of the PRC2 complex greatly decreased the global levels of H3K27me3, but not of H3K27me2, and increased H3K27me1 levels, consistent with a previous report, in which the *SU(Z)12* gene was knocked down [62], revealing that the silkworm PRC2 complex contributed to the tri-methylation of the H3K27 site as it does in mammals. However, key questions remain to be answered: how does the PRC2 complex catalyze the tri-methylation of H3K27, or what substrates are required for PRC2 activity? A recent report shows different H3K27 marks at the same target locus from various cell lines with different transcription levels, implying that the H3K27 modifications have distinctive correlations with the chromatin state [63]. That study also led to the speculation that H3K27me3 is established predominantly at repressed genes in inactive loci or domains, while H3K27me2 is established at inactive genes in potentially active loci. Thus our data, together with these studies, suggest that, at least on global levels, the silkworm PRC2 complex is involved in the establishment of the H3K27me3 state to form an inactive locus. Potentially, H3K27me1 may be a substrate for the PRC2 complex because of the significant increase in H3K27me1 after the knockdown of the PRC2 components.

Taken together, in the present study, we have identified 13 PcG orthologs in silkworm and characterized their expression profiles and functions. Our data also suggests a conserved mechanism involved by PcG complexes between holocentric organisms and monocentric organisms. These results will provide fundamental knowledge useful for further investigations of PcG functions in silkworm.

Acknowledgments This work was supported in part by a grant from the Ministry of Agriculture, Forestry and Fisheries of Japan (Integrated research project for plant, insect and animal using genome technology INSECT-1201), and KAKENHI no. 22248003 and 22248004 from the Japan Society for the Promotion of Science.

References

1. Struhl G (1981) A gene product required for correct initiation of segmental determination in Drosophila. Nature 293(5827):36–41
2. Lewis EB (1978) A gene complex controlling segmentation in Drosophila. Nature 276(5688):565–570
3. Morey L, Helin K (2010) Polycomb group protein-mediated repression of transcription. Trends Biochem Sci 35(6):323–332
4. Kohler C, Hennig L (2010) Regulation of cell identity by plant Polycomb and trithorax group proteins. Curr Opin Genet Dev 20(5):541–547
5. Bracken AP, Helin K (2009) Polycomb group proteins: navigators of lineage pathways led astray in cancer. Nat Rev Cancer 9(11):773–784
6. Oktaba K, Gutierrez L, Gagneur J, Girardot C, Sengupta AK, Furlong EE, Muller J (2008) Dynamic regulation by Polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. Dev Cell 15(6):877–889
7. Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Wang H, de la Cruz CC, Otte AP, Panning B, Zhang Y (2003) Role of histone H3 lysine 27 methylation in X inactivation. Science 300(5616):131–135
8. Mager J, Montgomery ND, de Villena FP, Magnuson T (2003) Genome imprinting regulated by the mouse Polycomb group protein Eed. Nat Genet 33(4):502–507
9. Muller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE, Simon JA (2002) Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell 111(2):197–208
10. Shao Z, Raible F, Mollaaghababa R, Guyon JR, Wu CT, Bender W, Kingston RE (1999) Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell 98(1):37–46
11. Brown JL, Mucci D, Whiteley M, Dirksen ML, Kassis JA (1998) The Drosophila Polycomb group gene pleiohomeotic encodes a DNA binding protein with homology to the transcription factor YY1. Mol Cell 1(7):1057–1064
12. Simon JA, Kingston RE (2009) Mechanisms of polycomb gene silencing: knowns and unknowns. Natl Rev Mol Cell Biol 10(10):697–708
13. Tolhuis B, de Wit E, Muijrers I, Teunissen H, Talhout W, van Steensel B, van Lohuizen M (2006) Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. Nat Genet 38(6):694–699
14. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K (2006) Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev 20(9):1123–1136
15. Schwartz YB, Kahn TG, Nix DA, Li XY, Bourgon R, Biggin M, Pirrotta V (2006) Genome-wide analysis of Polycomb targets in *Drosophila melanogaster*. Nat Genet 38(6):700–705
16. Klymenko T, Papp B, Fischle W, Kocher T, Schelder M, Fritsch C, Wild B, Wilm M, Muller J (2006) A Polycomb group protein complex with sequence-specific DNA-binding and selective methyl-lysine-binding activities. Genes Dev 20(9):1110–1122
17. Atchison L, Ghias A, Wilkinson F, Bonini N, Atchison ML (2003) Transcription factor YY1 functions as a PcG protein in vivo. EMBO J 22(6):1347–1358
18. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298(5595):1039–1043
19. Ketel CS, Andersen EF, Vargas ML, Suh J, Strome S, Simon JA (2005) Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. Mol Cell Biol 25(16):6857–6868
20. Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y (2004) Role of histone H2A ubiquitination in Polycomb silencing. Nature 431(7010):873–878
21. Kallin EM, Cao R, Jothi R, Xia K, Cui K, Zhao K, Zhang Y (2009) Genome-wide uH2A localization analysis highlights Bmi1-dependent deposition of the mark at repressed genes. PLoS Genet 5(6):e1000506
22. Rodrigues JC, Luo M, Berger F, Koltunow AM (2010) Polycomb group gene function in sexual and asexual seed development in angiosperms. Sex Plant Reprod 23(2):123–133

23. Goodrich J, Puangsomlee P, Martin M, Long D, Meyerowitz EM, Coupland G (1997) A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature 386(6620):44–51
24. Chai CL, Zhang Z, Huang FF, Wang XY, Yu QY, Liu BB, Tian T, Xia QY, Lu C, Xiang ZH (2008) A genomewide survey of homeobox genes and identification of novel structure of the Hox cluster in the silkworm, *Bombyx mori*. Insect Biochem Mol Biol 38(12):1111–1120
25. Murakami A, Imai HT (1974) Cytological evidence for holocentric chromosomes of the silkworms, *Bombyx mori* and *B. mandarina* (Bombycidae, Lepidoptera). Chromosoma 47(2): 167–178
26. International Silkworm Genome Consortium (2008) The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. Insect Biochem Mol Biol 38(12):1036–1045
27. Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, Beeman RW et al (2008) The genome of the model beetle and pest *Tribolium castaneum*. Nature 452(7190):949–955
28. Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu ZJ, Loftus B, Xi Z et al (2007) Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science 316(5832):1718–1723
29. Honeybee Genome Sequencing Consortium (2006) Insights into social insects from the genome of the honeybee *Apis mellifera*. Nature 443(7114):931–949
30. Xia Q, Cheng D, Duan J, Wang G, Cheng T, Zha X, Liu C, Zhao P, Dai F, Zhang Z, He N, Zhang L, Xiang Z (2007) Microarray-based gene expression profiles in multiple tissues of the domesticated silkworm, *Bombyx mori*. Genome Biol 8(8):R162
31. Mon H, Kobayashi I, Ohkubo S, Tomita S, Lee JM, Sezutsu H, Tamura T, Kusakabe T (2012) Effective RNA interference in cultured silkworm cells mediated by overexpression of *Caenorhabditis elegans* SID-1. RNA Biol 9(1) (in press)
32. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25(24):4876–4882
33. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24(8):1596–1599
34. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95(25):14863–14868
35. Tatsuke T, Sakashita K, Masaki Y, Lee JM, Kawaguchi Y, Kusakabe T (2010) The telomere-specific non-LTR retrotransposons SART1 and TRAS1 are suppressed by Piwi subfamily proteins in the silkworm, *Bombyx mori*. Cell Mol Biol Lett 15(1):118–133
36. Shechter D, Dormann HL, Allis CD, Hake SB (2007) Extraction, purification and analysis of histones. Nat Protoc 2(6):1445–1457
37. Wang L, Jahren N, Miller EL, Ketel CS, Mallin DR, Simon JA (2010) Comparative analysis of chromatin binding by Sex Comb on Midleg (SCM) and other polycomb group repressors at a Drosophila Hox gene. Mol Cell Biol 30(11):2584–2593
38. Peterson AJ, Mallin DR, Francis NJ, Ketel CS, Stamm J, Voeller RK, Kingston RE, Simon JA (2004) Requirement for sex comb on midleg protein interactions in Drosophila polycomb group repression. Genetics 167(3):1225–1239
39. Min J, Zhang Y, Xu RM (2003) Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev 17(15):1823–1828
40. Fischle W, Wang Y, Jacobs SA, Kim Y, Allis CD, Khorasanizadeh S (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev 17(15):1870–1881
41. Bernstein E, Duncan EM, Masui O, Gil J, Heard E, Allis CD (2006) Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol Cell Biol 26(7):2560–2569
42. Zdobnov EM, Bork P (2007) Quantification of insect genome divergence. Trends Genet 23(1):16–20
43. Kim CA, Sawaya MR, Cascio D, Kim W, Bowie JU (2005) Structural organization of a Sex-comb-on-midleg/polyhomeotic copolymer. J Biol Chem 280(30):27769–27775
44. Kim CA, Gingery M, Pilpa RM, Bowie JU (2002) The SAM domain of polyhomeotic forms a helical polymer. Nat Struct Biol 9(6):453–457
45. Zhang H, Christoforou A, Aravind L, Emmons SW, van den Heuvel S, Haber DA (2004) The *C. elegans* Polycomb gene SOP-2 encodes an RNA binding protein. Mol Cell 14(6):841–847
46. Matthews JM, Sunde M (2002) Zinc fingers—folds for many occasions. IUBMB Life 54(6):351–355
47. Trievel RC, Beach BM, Dirk LM, Houtz RL, Hurley JH (2002) Structure and catalytic mechanism of a SET domain protein methyltransferase. Cell 111(1):91–103
48. Tie F, Stratton CA, Kurzhals RL, Harte PJ (2007) The N terminus of Drosophila ESC binds directly to histone H3 and is required for E(Z)-dependent trimethylation of H3 lysine 27. Mol Cell Biol 27(6):2014–2026
49. Fritsch C, Brown JL, Kassis JA, Muller J (1999) The DNA-binding polycomb group protein pleiohomeotic mediates silencing of a Drosophila homeotic gene. Development 126(17): 3905–3913
50. Nekrasov M, Klymenko T, Fraterman S, Papp B, Oktaba K, Kocher T, Cohen A, Stunnenberg HG, Wilm M, Muller J (2007) Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. EMBO J 26(18):4078–4088
51. Stankunas K, Berger J, Ruse C, Sinclair DA, Randazzo F, Brock HW (1998) The enhancer of polycomb gene of Drosophila encodes a chromatin protein conserved in yeast and mammals. Development 125(20):4055–4066
52. Scheuermann JC, de Ayala Alonso AG, Oktaba K, Ly-Hartig N, McGinty RK, Fraterman S, Wilm M, Muir TW, Muller J (2010) Histone H2A deubiquitinase activity of the Polycomb repressive complex PR-DUB. Nature 465(7295):243–247
53. Chadwick BP, Willard HF (2004) Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. Proc Natl Acad Sci USA 101(50):17450–17455
54. Silva J, Mak W, Zvetkova I, Appanah R, Nesterova TB, Webster Z, Peters AH, Jenuwein T, Otte AP, Brockdorff N (2003) Establishment of histone h3 methylation on the inactive X chromosome requires transient recruitment of Eed-Enx1 polycomb group complexes. Dev Cell 4(4):481–495
55. Peters AH, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, Martens JH, Jenuwein T (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell 12(6):1577–1589
56. Tugendreich S, Tomkiel J, Earnshaw W, Hieter P (1995) CDC27Hs colocalizes with CDC16Hs to the centrosome and mitotic spindle and is essential for the metaphase to anaphase transition. Cell 81(2):261–268
57. Lo SM, Ahuja NK, Francis NJ (2009) Polycomb group protein Suppressor 2 of zeste is a functional homolog of Posterior Sex Combs. Mol Cell Biol 29(2):515–525
58. Li X, Han Y, Xi R (2010) Polycomb group genes Psc and Su(z)2 restrict follicle stem cell self-renewal and extrusion by controlling canonical and noncanonical Wnt signaling. Genes Dev 24(9):933–946
59. Kurzhals RL, Tie F, Stratton CA, Harte PJ (2008) Drosophila ESC-like can substitute for ESC and becomes required for Polycomb silencing if ESC is absent. Dev Biol 313(1):293–306

60. Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS (2004) Hierarchical recruitment of polycomb group silencing complexes. Mol Cell 14(5):637–646
61. Brown JL, Fritsch C, Mueller J, Kassis JA (2003) The Drosophila pho-like gene encodes a YY1-related DNA binding protein that is redundant with pleiohomeotic in homeotic gene silencing. Development 130(2):285–294
62. Cao R, Zhang Y (2004) SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Mol Cell 15(1):57–67
63. Kim YW, Kim A (2011) Characterization of histone H3K27 modifications in the beta-globin locus. Biochem Biophys Res Commun 405(2):210–215