The EMBO Journal Vol.16 No.11 pp.3219–3232, 1997

Mammalian homologues of the *Polycomb*-group gene *Enhancer of zeste* mediate gene silencing in *Drosophila* heterochromatin and at *S.cerevisiae* telomeres

Gotz Laible, Andrea Wolf, Rainer Dorn[1], Gunter Reuter[1], Corey Nislow[2,3], Angelika Lebersorger, Dan Popkin[4], Lorraine Pillus[2] and Thomas Jenuwein[5]

Research Institute of Molecular Pathology (IMP), Dr Bohrgasse 7, A-1030 Vienna, Austria, [1]Institute of Genetics, Martin Luther University of Halle, Domplatz 1, D-06108 Halle/S, Germany and [2]Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder, Boulder, CO 80309-0347, USA

[3]Present address: Department of Biology, Rutgers University, Camden, NJ 08102, USA
[4]Present address: Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA

[5]Corresponding author

Gene silencing is required to stably maintain distinct patterns of gene expression during eukaryotic development and has been correlated with the induction of chromatin domains that restrict gene activity. We describe the isolation of human (*EZH2*) and mouse (*Ezh1*) homologues of the *Drosophila Polycomb*-group (Pc-G) gene *Enhancer of zeste* [*E(z)*], a crucial regulator of homeotic gene expression implicated in the assembly of repressive protein complexes in chromatin. Mammalian homologues of *E(z)* are encoded by two distinct loci in mouse and man, and the two murine *Ezh* genes display complementary expression profiles during mouse development. The *E(z)* gene family reveals a striking functional conservation in mediating gene repression in eukaryotic chromatin: extra gene copies of human *EZH2* or *Drosophila E(z)* in transgenic flies enhance position effect variegation of the heterochromatin-associated *white* gene, and expression of either human *EZH2* or murine *Ezh1* restores gene repression in *Saccharomyces cerevisiae* mutants that are impaired in telomeric silencing. Together, these data provide a functional link between Pc-G-dependent gene repression and inactive chromatin domains, and indicate that silencing mechanism(s) may be broadly conserved in eukaryotes.

Keywords: mammalian *E(z)* homologues/Pc-G and PEV in *Drosophila*/SET-domain/TPE in *S.cerevisiae*

Introduction

Maintenance of different patterns of gene expression is the underlying mechanism by which cell type identity is inherited in eukaryotes and has been correlated with the organization of chromatin domains that modulate gene activity (Zuckerkandl, 1974). Paradigms for chromatin-controlled regulation of key developmental loci include silencing of the mating type loci in *Saccharomyces cerevisiae* (reviewed in Loo and Rine, 1995) and restriction of expression boundaries within the homeotic gene cluster (*HOM-C*) in *Drosophila* (Wedeen *et al.*, 1986; Simon *et al.*, 1993). Moreover, gene expression in eukaryotes is influenced by structural alterations in chromatin, phenomena referred to collectively as position effects. For example, genes placed in the vicinity of yeast telomeres (Gottschling *et al.*, 1990; Nimmo *et al.*, 1994) and centromeres (Allshire *et al.*, 1994), or at cytologically defined heterochromatic positions in the fly (reviewed in Weiler and Wakimoto, 1995), are severely downregulated. In addition, many genes on the inactive X chromosome and several imprinted genes (for reviews, see Efstradiadis, 1994; Migeon, 1994) are repressed in mammals. Basic chromatin components are known to participate in transcriptional regulation because alterations in dosage (Moore *et al.*, 1979; Han and Grunstein, 1988) and modification (Jeppesen and Turner, 1993) of the core histones affect gene expression. However, more complex levels of control are suggested from genetic screens on position effect variegation (PEV) in *Drosophila* and *S.cerevisiae*. These screens have identified distinct classes of genes that appear centrally involved in the regional organization of chromatin domains and the regulation of chromatin-dependent gene activity.

In contrast to classical transcription factors, many of these chromatin regulators do not seem to bind DNA in a sequence-specific manner, but rather associate with structurally altered chromatin or are recruited by factors already present at target sequences (for reviews, see Loo and Rine, 1995; Orlando and Paro, 1995; Pirrotta, 1995). Models for the function of chromatin regulators have been proposed to explain variegation and the clonal nature of gene expression within a given cell population (Tartof *et al.*, 1984; Henikoff, 1995; Karpen, 1995; Csink and Henikoff, 1996; Dernburg *et al.*, 1996). Consistent with these models, telomeric silencing in yeast and PEV in *Drosophila* display an epigenetic component, and it has been suggested that the epigenetic fine-tuning of chromatin-controlled gene expression patterns facilitates the generation of cell type diversity in eukaryotes (Zuckerkandl, 1974; Eissenberg *et al.*, 1995; Pillus and Grunstein, 1995).

In *Drosophila*, ~120 loci have been described genetically that enhance or suppress position-dependent gene expression, but only ~10% of the corresponding genes have been isolated so far (Reuter and Spierer, 1992). Interestingly, several of the known modifiers of PEV share structural motifs with both activators [*trithorax* or trx-group (trx-G)] and repressors [*Polycomb* or Pc-group (Pc-G)] of chromatin-dependent gene activity within HOM-C. The biological significance of chromatin-mediated gene regulation, particularly for controlling mammalian development, has gained further importance by the functional analysis of murine homologues of *Drosophila Psc* (Pc-G) and *trx*. Both of these mammalian homologues appear to be involved in

specifying expression boundaries within the murine *HOX*-clusters (Alkema *et al.*, 1995; Yu *et al.*, 1995; Akasaka *et al.*, 1996), and their deregulation has been implicated in inducing leukaemia (Gu *et al.*, 1992; Tkachuk *et al.*, 1992; van der Lugt *et al.*, 1992; Alkema *et al.*, 1995).

The structural similarities between gene products that are implicated in controlling gene expression at specific euchromatic locations (*HOM-C*) and in the vicinity of heterochromatin (PEV) point to a common mechanism(s) for the regulation of chromatin-dependent gene activity. However, only a few modifiers of PEV have been shown to induce homeotic transformations (Dorn *et al.*, 1993; Farkas *et al.*, 1994), and a direct structural involvement of Pc-G genes with heterochromatic regions has not yet been demonstrated. Recently, a 130 amino acid, carboxy-terminal region of high sequence similarity has been identified that is shared between *trx* and the Pc-G gene *Enhancer of zeste* [*E(z)*] (Jones and Gelbart, 1993). This carboxy terminus is also conserved in *Su(var)3-9*, a dominant suppressor of PEV in *Drosophila* (Tschiersch *et al.*, 1994). Because of the genetically defined antagonistic relationship between *trx* and *E(z)* (Jones and Gelbart, 1993; LaJeunesse and Shearn, 1996), this novel protein domain, termed SET (Tschiersch *et al.*, 1994), may play a role in the assembly of either activating or repressing chromatin complexes, dependent on interactions with accessory proteins from the trx- or Pc-G. Consistent with this notion, *E(z)* mutants have been shown to reduce immunostaining of some trx- and Pc-G proteins at polytene chromosomes (Rastelli *et al.*, 1993; Platero *et al.*, 1996). Finally, the SET domain is evolutionarily conserved and is present in gene products ranging from yeast to man (Stassen *et al.*, 1995). Based on the above criteria, the SET domain appears to define a new gene family of important chromatin regulators.

Here, we describe the isolation of mammalian homologues of the *Drosophila* Pc-G gene *E(z)* and demonstrate function of the *E(z)* gene family in mediating gene repression at well-defined heterochromatic regions in *Drosophila* and *S.cerevisiae* chromatin. These data provide a functional link between Pc-G-dependent gene silencing and the organization of inactive chromatin domains.

Results

Mammalian E(z) homologues are encoded by two loci

Based on sequence information from the conserved carboxy-terminal SET domain of *Drosophila E(z)*, we screened a human B-cell-specific cDNA library for mammalian homologues (see Materials and methods). Out of 500 000 plaques, 36 primary phages were selected. Five isolates contained partial cDNA fragments of a gene that was recently mapped to human chromosome 17 and which has been named *EZH1* (Abel *et al.*, 1996). By contrast, the majority of the isolates contained novel cDNA sequences from a second human gene, which we designated *EZH2* (*Enhancer of zeste* homologue 2). A full-length human *EZH2* cDNA of 2.6 kb encoding a protein of 746 amino acids was obtained by combining DNA fragments from different subclones. Using the human *EZH2* cDNA as a probe, we next screened a mouse brain cDNA library for murine homologues of *E(z)*. From the mouse brain library, only cDNAs encoding sequences with highest homology to human *EZH1* were found. A 5′ RACE amplification was used to complete the missing 5′ end and, after subcloning, a murine *Ezh1* cDNA that encodes a protein of 747 amino acids was obtained. Recently, a murine cDNA homologous to human *EZH2* has been isolated by others (Hobert *et al.*, 1996a).

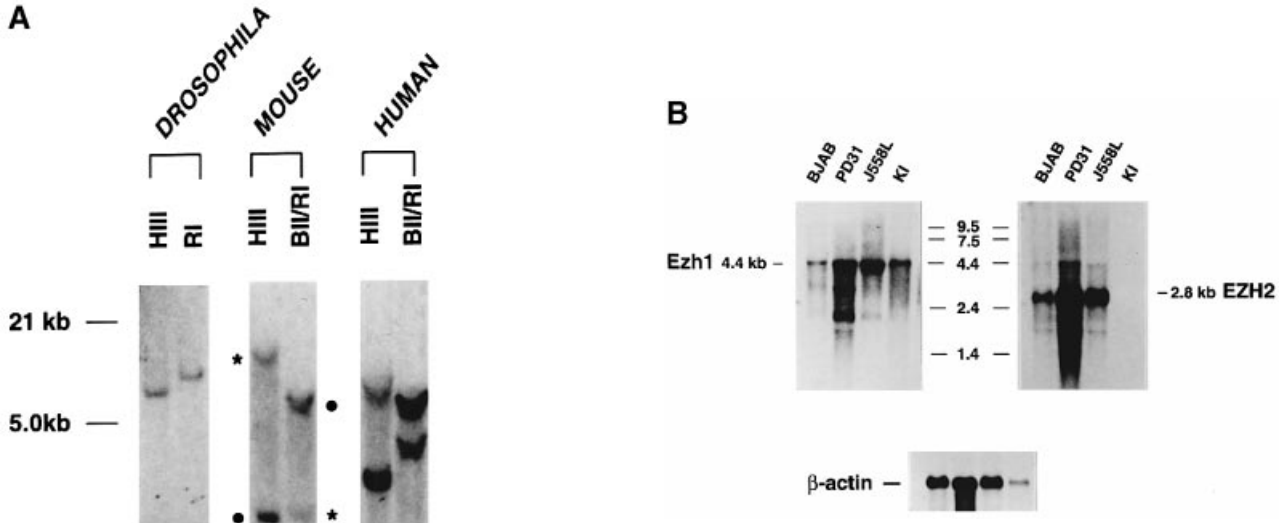
To examine directly the number of mammalian *E(z)* homologues, we hybridized, under reduced stringency, a DNA blot containing *Drosophila*, mouse and human DNA with a mixed DNA probe that was derived from the SET domains of the human *EZH2* and mouse *Ezh1* cDNAs (Figure 1A, top). Whereas only a single fragment is detected in *Drosophila* DNA, 2–4 fragments hybridized with this mixed SET domain probe in both mammalian DNAs. The fragments could either be correlated to predicted sizes from the genomic murine *Ezh1* locus (Figure 1A, bottom) or are exclusively detected by the human *EZH2* SET domain probe (data not shown). Moreover, homology searches (Altschul *et al.*, 1990) against the currently available human expressed sequence tags (ESTs) (Bassett *et al.*, 1995) indicated an almost perfect match with either *EZH2* or *EZH1*. Thus, we conclude that human and mouse *E(z)* homologues are represented by not more than two distinct loci which most likely reflect a gene duplication of a single ancestral locus.

To determine the accurate sizes of the mRNAs encoded by *Ezh1* and *EZH2*, we hybridized an RNA blot containing poly(A)$^+$ RNA from human (BJAB) and mouse (PD31 and J558L) B-cell-specific cell lines and from mouse kidney with DNA probes that were derived from the *Ezh1* and *EZH2* cDNAs (Figure 1B). The *Ezh1* probe detected a specific mRNA of ~4.4 kb which is present in all RNAs analysed and whose size is in good agreement with the 4.0 kb *Ezh1* cDNA that lacks ~200 bp of 3′ untranslated sequences. Rehybridization of the RNA blot with a DNA probe specific for the full-length *EZH2* cDNA identified a mRNA of 2.8 kb which closely matches the size of the *EZH2* cDNA in both the human and mouse cell lines, but not in mouse kidney.

Structural conservation of mammalian and fly E(z) proteins

Human *EZH1* and murine *Ezh1* (G.Mattei, A.Lebersorger and T.Jenuwein, unpublished), and probably also *EZH2*/*Ezh2*, are encoded by orthologous loci in mouse and man, and in both comparisons amino acid identities are 98%. By contrast, inverse comparisons of *EZH2*/*EZH1* or *Ezh2*/*Ezh1* only reach an overall amino acid identity of 67%, suggesting that the presumed gene duplication may have occurred at a relatively early stage after the divergence of vertebrate and non-vertebrate species. Considering the overall cDNA size and sequence comparisons, *EZH2*/*Ezh2* and *E(z)* appear to be most closely related, whereas *EZH1*/*Ezh1* is the more divergent of the mammalian gene pairs. Over the entire length of the 760 amino acids containing *E(z)* protein (Jones and Gelbart, 1993), *EZH2*/*Ezh2* is 61% and *EZH1*/*Ezh1* is 55% identical to *E(z)*. Thus, human and mouse homologues of *Drosophila E(z)* display the highest structural conservation among any of the currently known mammalian and *Drosophila* chromatin regulators.

Alignment of all four mammalian *E(z)* proteins with



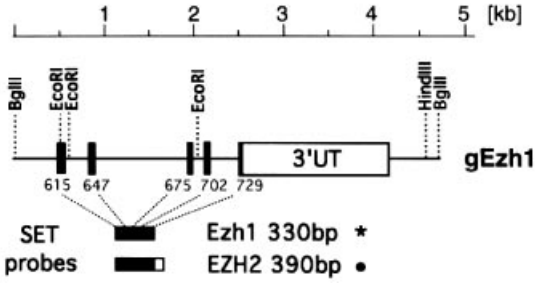


Fig. 1. *Mammalian E(z)* homologues are encoded by two distinct loci. (**A**) DNA blot analysis of *E(z)*-related sequences in 15 μg of digested *Drosophila*, mouse and human DNA. Digests were performed with the restriction enzymes *Hin*dIII (HIII), *Eco*RI (RI) or with a combination of *Bgl*II and *Eco*RI (BII/RI). The DNA blot was hybridized under reduced stringency with a mixed DNA probe derived from the SET domains of the human *EZH2* and murine *Ezh1* cDNAs. Mouse DNA fragments that correspond to predicted sizes from the genomic murine *Ezh1* locus (see the schematic diagram at the bottom) are indicated by an asterisk, and those exclusively detected by the human *EZH2* SET domain probe are highlighted by a solid dot. Position and sizes of DNA marker fragments are indicated at the left. Below, the 3′ portion of the genomic *Ezh1* locus, comprising SET domain exons (black boxes), is shown. Numbering of the exons refers to amino acid positions in *Ezh1*. (**B**) RNA blot analysis of *E(z)*-related transcripts in 5 μg of poly(A)$^+$ RNA isolated from human (BJAB) and mouse (PD31 and J558L) B-cell lines and from mouse kidney (KI). The RNA blot was sequentially hybridized under high stringency with DNA probes derived from nearly full-length murine *Ezh1* cDNA (left panel), full-length human *EZH2* cDNA (right panel) and from murine β-actin sequences (bottom panel). The sizes of the respective transcripts were calculated according to the indicated molecular weight markers.

Drosophila E(z) reveals four regions of high sequence identity (Figure 2). In addition to the most highly conserved SET domain (86% identity), an 115 amino acid, cysteine-rich region immediately precedes the SET domain (68% identity). Most of the cysteine residues, originally noted for *E(z)* (Jones and Gelbart, 1993), are conserved in the mammalian proteins. However, these cysteine residues (a total of 18) are unusually spaced and do not show any apparent homology to other well-defined cysteine-rich regions. In the amino-terminal half, two stretches, one of 66 amino acids (domain I; 66% identity) and one of 114 amino acids (domain II; 56% identity), are highly related between mammalian and fly *E(z)* proteins. Domain II also contains six conserved cysteine residues with unusual spacing and is separated from domain I by a stretch of charged amino acids. Finally, the position and sequence of the nuclear localization signal of the *E(z)* protein (Jones and Gelbart, 1993) is also conserved in all mammalian homologues.

Complementary expression profiles of murine Ezh1 and Ezh2

The differences in the abundance of *Ezh1* and *Ezh2* transcripts in kidney suggested that both loci may be differentially regulated during mouse development. To analyse the temporal and tissue-specific expression profile of the two *Ezh* genes during mouse development, we performed an RNase protection analysis with antisense RNA probes that are specific for the SET domain of *Ezh1* or for the amino-terminal end of *Ezh2*. *Ezh1*-specific transcripts were detected ubiquitously during mouse embryogenesis; however, their relative abundance is gradually upregulated during development and reaches 4- to 5-fold higher levels in kidney, brain and skeletal muscle of adult mice (Figure 3, top panel). A 2- to 3-fold increase in the relative abundance (after correction for differences in the amount of RNA as indicated by the S16 RNA control transcripts) of *Ezh1* transcripts was also detected after retinoic acid-induced *in vitro* differentiation of embryonic stem cells. By contrast, *Ezh1* transcripts remained at a low level in liver, spleen and thymus.

The RNase protection analysis of *Ezh1* transcript levels during mouse embryogenesis and ES cell differentiation demonstrates a specific upregulation of *Ezh1* gene activity at later stages of development. By contrast, *Ezh2* transcript levels are downregulated during progressing development and display a more thymus-restricted expression (Figure 3, middle panel; Hobert *et al.*, 1996a), resulting in complementary expression profiles of *Ezh1* and *Ezh2* that may indicate dosage dependence or distinct functions of the

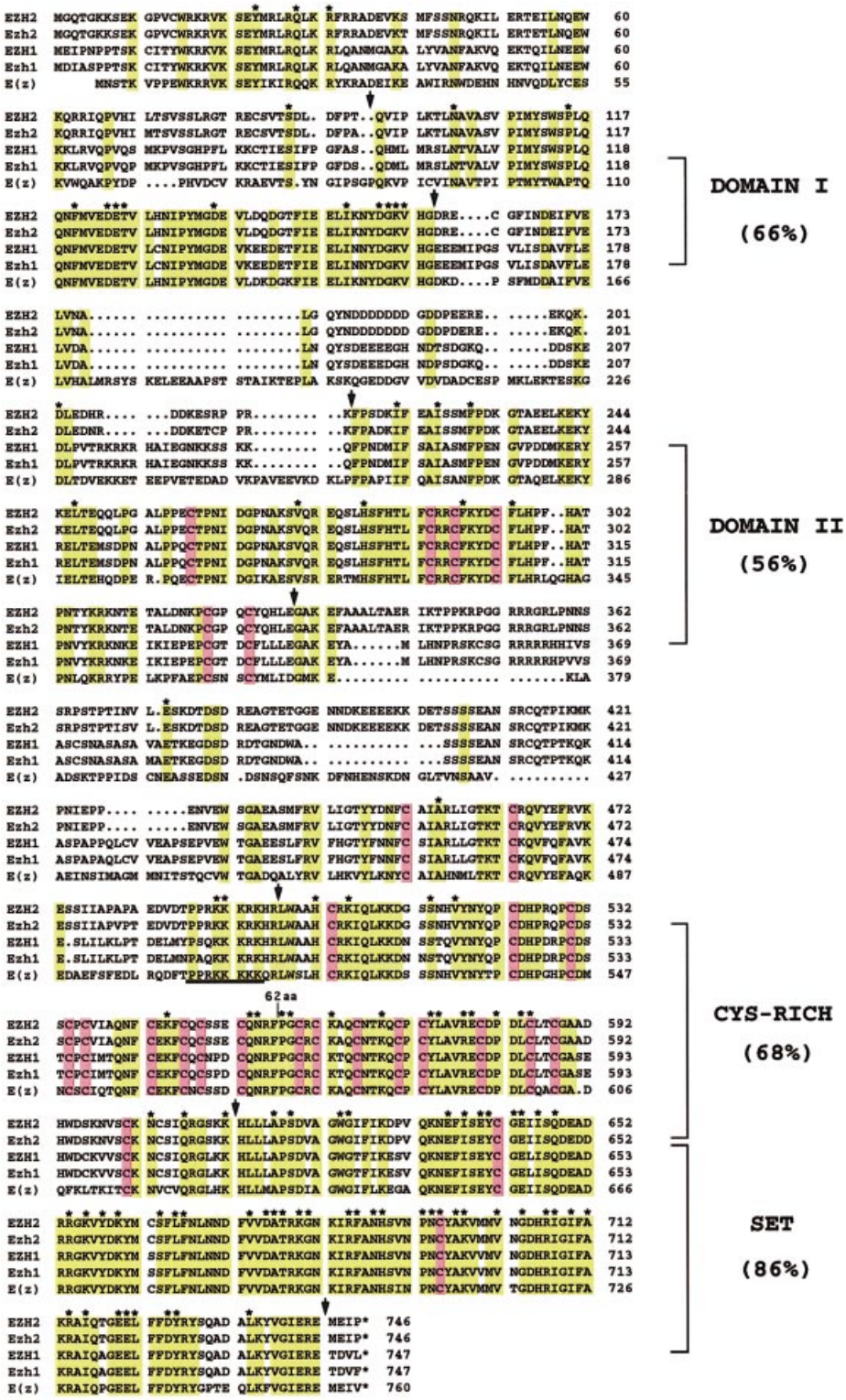


Fig. 2. Conserved domains of mammalian and fly *E(z)* proteins. Amino acid sequences of human (*EZH1* and *EZH2*), mouse (*Ezh1* and *Ezh2*) and *Drosophila E(z)* proteins were aligned using the PILEUP program of the GCG software package. Amino acids that are identical in all five proteins are shown in green, and conserved cysteine residues are highlighted by a pink colour. Four regions of high sequence similarity are indicated to the right, together with the respective amino acid identities calculated from conserved positions in all five proteins within these individual domains. The domain boundaries (indicated by arrowheads) overlap or are in close proximity to exon/intron positions determined for *Ezh1* (data not shown) and *E(z)* (Jones and Gelbart, 1993). The conserved nuclear localization signal noted for *E(z)* (Jones and Gelbart, 1993) is underlined. Full-length sequence information for *EZH1* was kindly provided by Ken Abel (1996), and the *Ezh2* sequence has been reported recently (Hobert *et al.*, 1996a). The full-length *EZH2* protein extends the amino terminus of a truncated *EZH2* variant (Hobert *et al.*, 1996b) by 133 amino acids. Also indicated by an asterisk above the *EZH2* sequence are amino acid positions that are invariant between the *E(z)* gene family and *S.cerevisiae SET1* (1080 amino acids). The position of an additional stretch of 62 amino acids in *SET1* that is unrelated to *E(z)* homologues is shown.

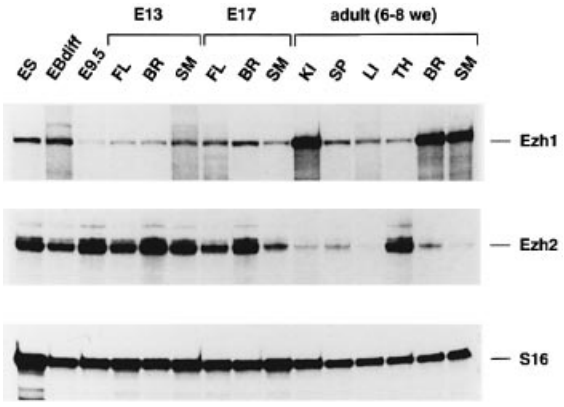


Fig. 3. Expression analysis of murine *Ezh1* and *Ezh2*. RNase protection analysis detecting *Ezh1* and *Ezh2* transcripts during mouse (129/Sv) development. 10 μg of total RNA were hybridized with riboprobes that specifically protect *Ezh1* (top panel) or *Ezh2* (middle panel). As a quantitation control, an RNA probe that protects murine S16 rRNA sequences was included in the hybridizations. Total RNA was prepared from undifferentiated D3 embryonic stem cells (ES), embryoid bodies derived after retinoic acid-induced *in vitro* differentiation of D3 cells (EBdiff), whole embryos at day E9.5 and fetal liver (FL), brain (BR) and skeletal muscle (SM) of day E13 or day E17 embryos. In addition, total RNA was also prepared from adult tissues, including kidney (KI), spleen (SP), liver (LI), thymus (TH), brain (BR) and skeletal muscle (SM).

encoded proteins. Further, observed splicing variants for *EZH2* (data not shown) or *Ezh2* (Hobert *et al.*, 1996a) that remove the amino-terminal half of the conserved cysteine-rich region or part of the SET domain in *EZH1* (Abel *et al.*, 1996) could encode functionally distinct protein isoforms. Alternatively, *Ezh1* and *Ezh2* may provide similar functions, but because of the presence of two loci, transcript levels may be balanced to protect against negative gene dosage effects which, in analogy to other described *Drosophila* chromatin regulators (Reuter *et al.*, 1990; Eissenberg *et al.*, 1992; Tschiersch *et al.*, 1994), could interfere with normal gene function.

EZH2 and E(z) enhance PEV in Drosophila

Although Pc-G-dependent gene silencing has been correlated with the establishment of repressive chromatin domains (see Introduction), most of the Pc-G genes do not affect gene repression in the vicinity of heterochromatin (Grigliatti, 1991; G.Reuter, unpublished). However, *E(z)* mutants have been described recently that reduce immunostaining of several Pc-G proteins at natural euchromatic positions (Rastelli *et al.*, 1993) and prevent the recruitment of Pc-G complexes at ectopic heterochromatic locations (Platero *et al.*, 1996). Together with the extremely high degree of evolutionary conservation between mammalian and fly *E(z)* homologues, these data suggested a crucial role for *E(z)* proteins in regulating regional gene activities in chromatin.

To demonstrate an involvement of *E(z)* genes in controlling the organization of repressive chromatin domains, we analysed the potential of *EZH2* and *E(z)* to modify PEV in *Drosophila*. By P element-mediated germline transformation, we established nine transgenic fly lines that carry the human *EZH2* cDNA under the control of the heat shock promoter hsp70. As a control, we also generated transgenic flies carrying the *E(z)* cDNA (Jones and Gelbart, 1993) (four lines) or a genomic fragment encoding *E(z)* sequences (Jones and Gelbart, 1993) (one line). For all of the lines, the chromosomal assignment of the transgene was mapped genetically (see Table I) and the transgenic lines were crossed into In(1)w^{m4h} and In(1)w^{m4h}; *Su(var)2-1* indicator strains (see Materials and methods). These indicator strains carry an inversion [In(1)w^{m4h}] which places the *white* gene adjacent to pericentric X heterochromatin, resulting in stochastically repressed, variegated patterns of gene expression that can be easily detected as red or white patches in the *Drosophila* eye. Mutations in the suppressor of variegation *Su(var)2-1*, which has been shown to reduce deacetylation of histone H4 (Dorn *et al.*, 1986), derepress heterochromatin-restricted w^{m4h} expression, leading to an almost wild-type red eye phenotype.

Visual inspection of progeny of the respective crosses indicated that six out of the nine *EZH2* lines and two out of the five *E(z)* lines induced an increase in the proportion of unpigmented areas in the eyes, therefore classifying both *EZH2* and *E(z)* as dose-dependent modifiers of PEV (Figure 4). Moreover, after mobilization of the *E(z)* transgene in a negative line (line D), eight sublines with a significant enhancer effect on PEV could be selected (Table I). PEV enhancement was largely independent of heat shock treatment, and subsequent expression analysis by *in situ* hybridization demonstrated a low, basal transcription throughout embryonic development of those lines displaying an enhancer effect (data not shown).

To quantify the degree of PEV enhancement of the individual transgenic lines, eye pigments were extracted from the progenies of the respective crosses and pigment absorbance at 480 nm was measured. The result of these quantitations, summarized in Table I, shows that the strongest enhancer effect of *EZH2* (line E) is similar to *E(z)* and results in an ~10-fold reduction in the concentration of red eye pigments as compared with non-transgenic siblings. Thus, the *Drosophila* Pc-G gene *E(z)*, which has been implicated in restricting gene-specific expression boundaries within the chromatin domains of *HOM-C* (Jones and Gelbart, 1990; Phillips and Shearn, 1990), is also capable of stabilizing a repressive transcriptional state in the vicinity of constitutive heterochromatin. Moreover, this potential of *E(z)* to repress heterochromatin-associated gene activity is conserved in its human homologue *EZH2*.

EZH2-mediated enhancement of PEV displays paternal imprinting

In the course of the matings to the indicator strains, we observed that non-transgenic In(1)w^{m4h} and In(1)w^{m4h}; *Su(var)2-1* male progeny, derived from crosses with most of the transgenic *EZH2* or of the transgenic *E(z)* lines [lines indicated (p) in Table I], also exhibited an ~2- to 3-fold reduction in the proportion of red eye pigments as compared with unmated In(1)w^{m4h} and In(1)w^{m4h}; *Su(var)2-1* control flies (Table I), although the transgene was no longer present. This enhancer effect on w^{m4h} gene repression was strictly dependent upon previous paternal transmission of the transgene. Although the molecular basis for this imprinting-like effect is not known, this result suggests that *EZH2* and *E(z)* can participate not only in the induction, but also in the propagation of an altered chromatin structure which represses w^{m4h} gene activity in *trans*, and is

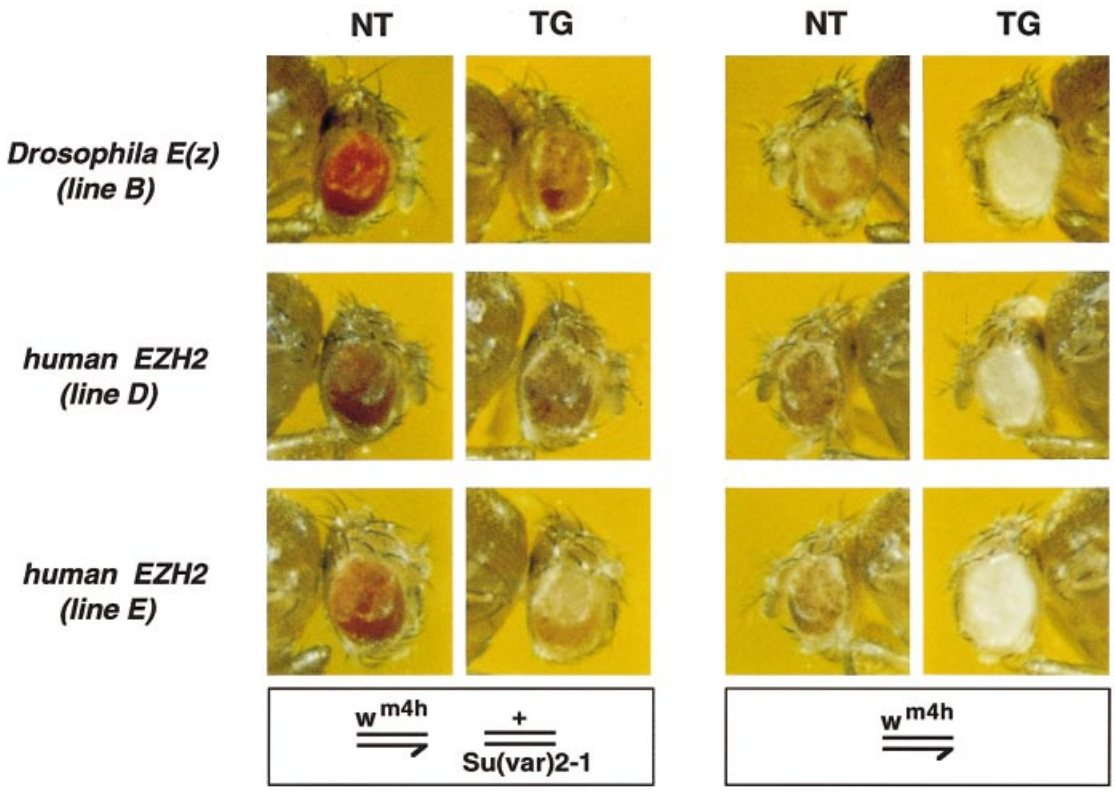


Fig. 4. *EZH2* and *E(z)* enhance PEV in *Drosophila*. Flies transgenic for human *EZH2* (lines D and E) or *Drosophila E(z)* (line B) were crossed to an indicator strain containing an inversion that places the *white* gene adjacent to pericentric X heterochromatin (w^{m4h}). In addition, transgenic flies were also mated to a sensitized indicator strain that carries a strong suppressor mutation [*Su(var)2-1*]. Progeny of the crosses was genotyped according to phenotypic markers, and transgene-dependent repression of w^{m4h} is detected by increased proportions of unpigmented areas in the eyes, thus categorizing *EZH2* and *E(z)* as enhancers of PEV. NT, non-transgenic offspring; TG, transgenic offspring.

reminiscent of transvection phenomena described for the $zeste^1$–*white* interaction (Kalisch and Rasmuson, 1974; Jones and Gelbart, 1990; Phillips and Shearn, 1990). In support of this interpretation, most transgenic *EZH2*/$zeste^1$ heterozygote females display $zeste^1$ gene variegation after heat shock (data not shown). Thus, together with the enhancement of PEV shown above, these data extend the functional conservation between *EZH2* and *E(z)* in regulating repressive transcriptional states in *Drosophila* chromatin.

EZH2 and Ezh1 restore telomeric silencing in S.cerevisiae

As a second assay to examine involvement of mammalian *E(z)* genes in mediating the organization of repressive chromatin domains, we analysed the potential of human *EZH2* and mouse *Ezh1* to restore gene repression in *S.cerevisiae* mutants that are defective in telomeric silencing. Despite cytogenetic differences between chromatin of yeast and multicellular eukaryotes (Pillus and Grunstein, 1995), telomeric silencing or telomeric position effect (TPE) closely resembles PEV in *Drosophila*. For example, insertion of a *URA3* gene at the subtelomeric region of the left arm of chromosome VII results in stochastic silencing of *URA3* expression in many, but not all, cells within a population. If *URA3* is repressed, cells survive selection with the uracil analogue 5-fluoro-orotic acid (5-FOA). The degree of telomeric silencing can be quantitated by comparing the relative growth of serial cell dilutions in permissive versus 5-FOA media, providing a very sensitive assay to detect chromatin-mediated changes in transcriptional states (Gottschling *et al.*, 1990).

Recently, an *S.cerevisiae* SET domain gene (*SET1*) has been identified whose disruption in *URA3*-tel marker strains abolishes the ability of the mutant cells to grow on 5-FOA, indicating complete (>100 000-fold) derepression of telomeric silencing (Nislow *et al.*, submitted; Figure 5A, top panel). Plasmid-borne *SET1* restored growth on 5-FOA, demonstrating that the telomeric silencing defect was due to the *set1* mutation (Figure 5A, middle panel). In addition to the carboxy-terminal SET domain (41% amino acid identity), *SET1* (1080 amino acids) and *EZH2* (746 amino acids) also share several other stretches of moderate sequence similarity that are conserved in the *E(z)*/*EZH* gene family (see Figure 2). Using low-copy (CEN4-*LEU*) vectors carrying galactose-inducible *EZH2* and *Ezh1* cDNAs or a haemagglutinin-tagged *EZH2* variant (to detect protein levels), we introduced mammalian *E(z)* homologues into *set1*Δ indicator strains that contain single (*URA3*-tel; LPY1297) or double (*URA3*-tel and *ADE2*-tel; LPY1727) telomeric reporter genes. Transformants were grown selectively to saturation and plated in serial 10-fold dilutions onto complete, –LEU and –LEU/5-FOA media (Figure 5A, bottom panels).

Surprisingly, on average 50% of both *EZH2* (16/32) and HA-*EZH2* (17/32) primary transformants were capable of growing at low cell densities on 5-FOA, indicating complementation of *set1*Δ and a capacity for human *EZH2*

	GENE/ CONSTRUCT	LINE	CHROMO-SOME	TRANSGENIC	NON TRANSGENIC (MALE OFFSPRING)	ENHANCER OF PEV (p) = paternal imprinting	
				RED PIGMENT (ABSORBANCE 480 nm)			
				$\frac{w^{m4h}}{Su(var)2\text{-}1}$	$\frac{w^{m4h}}{Su(var)2\text{-}1}$		
	none			NA	0.125		
Drosophila E(z)	P[E(z)ry+] (genomic fragment)	B23.1	2	0.009	0.052	+++	(p)
	P[hspE(z)ry+]	A	2	0.128	0.135	-	
		B	3	0.016	0.061	+++	(p)
		C	3	0.111	0.106	-	
		D	3	0.134	0.105	-	
		D-1	3	0.014	0.051	+++	(p)
		D-2	3	0.036	0.080	++	(p)
		$E(z)^5$	3	NA	0.217	Su (weak)	
		$E(z)^1$	3	NA	0.194	Su (weak)	
		$E(z)^{15}$	3	NA	0.122	-	
human EZH2	P[hspEZH2ry+]	A	2	0.055	0.051	++	(p)
		B	3	0.025	0.048	+++	(p)
		D	3	0.057	0.110	++	
		E	2	0.012	0.052	+++	(p)
		F	2	0.045	0.076	++	(p)
		I	2	0.126	0.089	-	
		J	3	0.084	0.077	+	(p)
		L	2	0.148	0.151	-	
		M	2	0.102	0.120	-	

Table I. Quantitation of PEV enhancement and paternal imprinting by human *EZH2* and *Drosophila E(z)*. Offspring of crosses between various fly lines transgenic for human *EZH2* or *Drosophila E(z)* (indicated at left) and w^{m4h} indicator strains that contain an additional suppressor mutation (*Su(var)2-1*) were genotyped according to phenotypic markers. The chromosomal assignment of the respective transgene has been mapped and is shown. The *E(z)* gene has been localized to region 67E on chromosome 3 (Kalisch and Rasmuson, 1974; Jones and Gelbart, 1990). Pooled red eye pigments extracted from the eyes of 10 individual siblings representing the respective genotypes were measured for their absorbance at 480 nm. The values given are the mean of three independent measurements with a SD of <10%. The enhancer effect on PEV (indicating repression of w^{m4h}) varied among the lines and resulted in pronounced (5- to 10-fold; +++), intermediate (2- to 3-fold; ++) and weak (~2-fold; +) reduction in the ratio of the absorbance between w^{m4h}; *Su(var)2-1* (0.125) and transgenic red eye pigments. The *E(z)* transgene of the negative line D was mobilized, resulting in eight sublines (of which only two are shown) with a significant enhancer effect on PEV. By contrast, the $E(z)^5$ null and the antimorphic $E(z)^1$ mutations are weak suppressors (Su) of PEV, whereas the hypomorphic, wild-type like $E(z)^{15}$ mutation has no effect. In addition, most of the lines exhibit a 2- to 3-fold reduction in the proportion of red eye pigments in outcrossed male offspring, although the transgene is no longer present, indicating paternal imprinting (p) of w^{m4h} repression.

to restore telomeric silencing to almost wild-type levels (Table II). Similarly, *Ezh1* (5/12) primary transformants also rescued the *set1*Δ-dependent loss in telomeric silencing, although, when compared with permissive growth, only 5% of these *Ezh1* transformants did form colonies on 5-FOA, suggesting that *Ezh1* has a weaker potential than *EZH2* (30%) or *SET1* (75%) to restore *set1*Δ-dependent telomeric silencing. In control transformations with the empty CEN4-*LEU* vector, seven out of 78 primary colonies exhibited some growth on 5-FOA; however, these likely suppressor transformants did not grow upon further subcloning (see below).

Replica plating onto uracil-deficient medium indicated that *EZH2*/*Ezh1* 5-FOA-resistant colonies are still $URA3^+$ (data not shown). To rule out that rescue of *set1*Δ-dependent telomeric silencing by the mammalian *E(z)* homologues may be the consequence of interference with *URA3* transcription, we next analysed repression of the second telomeric reporter gene (*ADE2*) which had been inserted at the subtelomeric region of the right arm of chromosome V (Singer and Gottschling, 1994). If *ADE2* is repressed, cells accumulate a chromogenic intermediate providing a colour assay to detect silenced (pink to red) or active (white) transcriptional states. Whereas the isogenic wild-type *SET1* strain displays a weak pink colour with sectored patches at the edge of the colonies, disruption of *SET1* results in predominantly white (derepressed) colonies (Figure 5A, right panel). Complementation of *set1*Δ by plasmid-borne *SET1* gives rise to transformants that develop a darker pink colour. Importantly, replica transformants of *EZH2* and HA-*EZH2* 5-FOA-resistant colonies display a red colour on complete medium, whereas most of the 5-FOA-sensitive transformants remain largely white and unsectored. To quantify repression of the *ADE2*-tel marker gene, respective transformants were plated at a density of 200–500 colonies and ~200 colonies/transformant were inspected under a dissecting microscope for the presence of red-sectored patches. The result of this

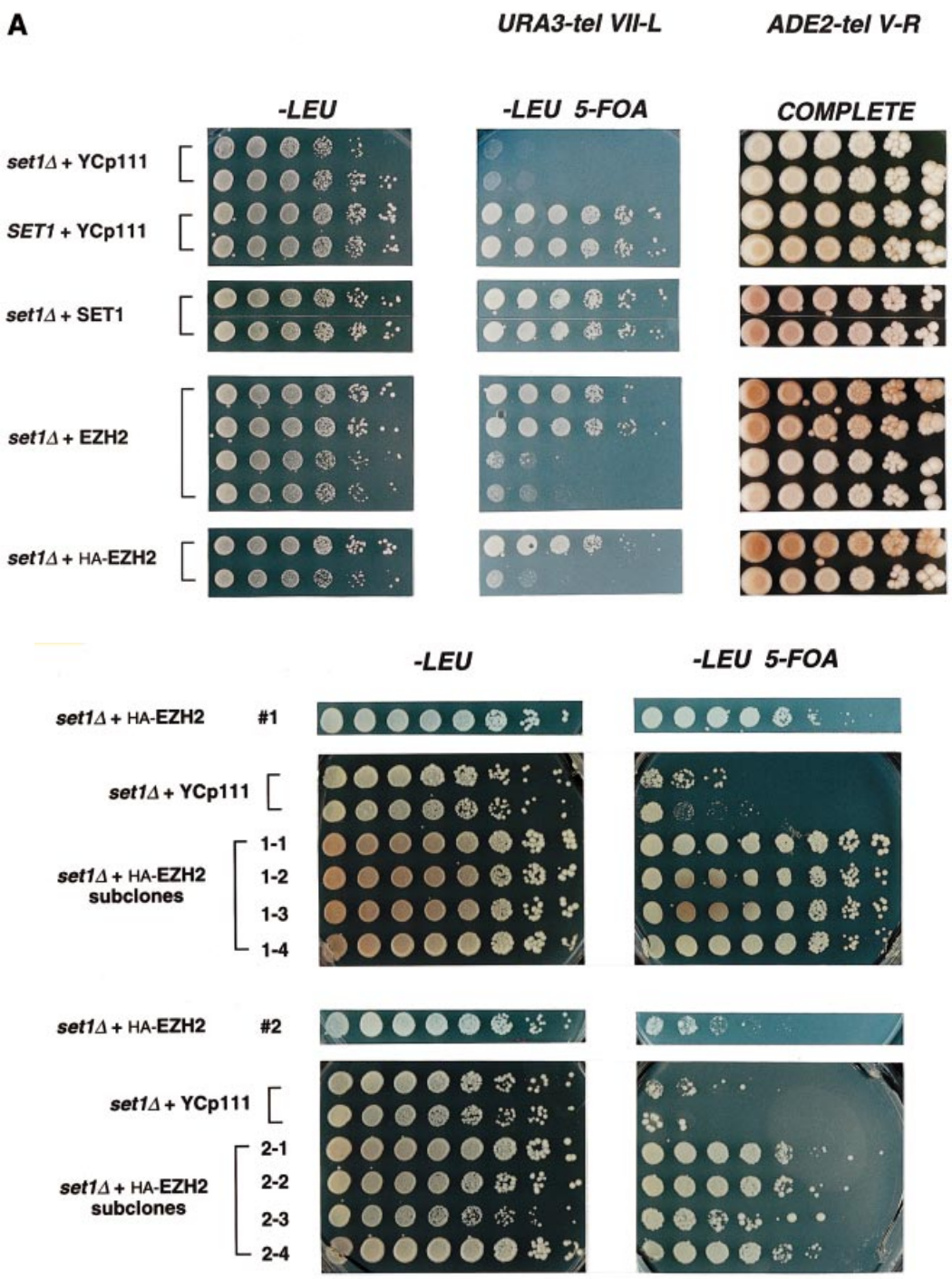


Fig. 5. Mammalian *E(z)* homologues restore telomeric silencing in *S.cerevisiae*. (**A**) Double telomeric marker strains (*URA3*-tel @ VII-L and *ADE2*-tel @ V-R), isogenic for wild-type *SET1* (LPY1638) or a disruption of *SET1* (*set1*Δ::*HIS3*; LPY1727), were transformed with low-copy CEN-*LEU* vectors carrying *SET1*, *EZH2* or an HA-epitope tagged variant of *EZH2*. Primary transformants were picked at random, and serial 10-fold dilutions were replica-plated to test for growth (–LEU), sensitivity to 5-FOA (–LEU/5-FOA) and colour development (complete medium). Complementation of *set1*Δ-dependent derepression of telomeric *URA3* is demonstrated by growth on 5-FOA and repression of telomeric *ADE2* is reflected by the appearance of pink- to red-sectored colonies. The bottom panels show representative *EZH2* and HA-*EZH2* primary transformants. Similar results were also obtained for the murine *Ezh1* homologue. YCp111 are control transformants carrying only the CEN4-*LEU* vector. (**B**) Mammalian *E(z)* homologues induce epigenetic switching from a non-repressed to a silenced state. HA-*EZH2* was introduced into a telomeric *URA3* marker strain in which *SET1* had been disrupted (LPY1297). Primary HA-*EZH2* transformants that are either 5-FOA resistant (#1) or 5-FOA sensitive (#2) were streaked for single colonies, and four randomly picked subclones for each parental transformant were analysed in serial 5-fold dilutions for growth on 5-FOA as described above.

quantitation is summarized in Table II and indicates that *EZH2* also restores *set1*Δ-dependent silencing of a telomere-proximal *ADE2* gene to wild-type (*SET1*) levels.

EZH2 and Ezh1 induce epigenetic switching of an active to a repressed transcriptional state

In contrast to *SET1*, *EZH2* and *Ezh1* complement a *set1*Δ-dependent silencing defect on average in only every other

		URA3-tel VII-L			ADE2-tel V-R
STRAIN	GENE	# COLONIES 5-FOA^R/total	% 5-FOA ^{R}	SWITCHING SUBCLONES 5-FOA^R/total	% SECTORED COLONIES
SET1	none	14/14	100	NA	91
	SET1	12/12	100	NA	ND
	HA-EZH2	12/12	100	NA	ND
set1Δ	none	7/78	0.001	0/8	3
	SET1	24/24	75	NA	90
	EZH2	16/32	30	3/8	92
	HA-EZH2	17/32	30	7/8	ND
	Ezh1	5/12	5	2/8	ND

Table II. Quantitation of *EZH2*- and *Ezh1*-mediated complementation of *set1*Δ-dependent telomeric silencing. The indicated genes or the CEN4-*LEU* vector control (none) were introduced into isogenic *SET1* or *set1*Δ telomeric marker (*URA3* @ VII-L and *ADE2* @ V-R) strains (see Figure 5 and Materials and methods). Repression of telomeric *URA3* is indicated by the absolute number of primary transformants that were able to grow on 5-FOA medium compared with the total number of primary transformants tested (left lane: 5-FOA^R/total). The percentage of 5-FOA resistance reflects the approximate ratio of colony numbers in permissive (–LEU) versus –LEU/5-FOA media and, in contrast to the described 50% intrinsic limit of 5-FOA resistance (Gottschling *et al.*, 1990), was set arbitrarily to 100% in the wild-type *SET1* strain. Also shown is the absolute number of subclones from initially 5-FOA-sensitive primary transformants that were able to switch to 5-FOA resistance. Four subclones for each of two independent primary transformants of the respective genotype were tested (middle lane: 5-FOA^R/total). Repression of telomeric *ADE2* was quantitated by plating 200–500 colonies on complete medium and counting (under a dissecting microscope) the number of pink- to red-sectored colonies.

transformant (see Table II), despite the presence of similar HA-*EZH2* protein levels in silenced versus non-silenced colonies (data not shown). Galactose induction did not increase this ratio of gene silencing (data not shown), indicating that there is not a simple quantitative relationship between the amount of *EZH2* or *Ezh1* and the extent of silencing. Rather, this observation suggests that the mammalian *E(z)* proteins participate in an apparent epigenetic mechanism of chromatin-controlled gene repression. We further examined this epigenetic component by analysing the stability of *EZH2*/*Ezh1*-dependent gene repression. For each of two independent silenced or non-silenced primary colonies expressing either *EZH2*, HA-*EZH2* or *Ezh1*, four subclones were grown selectively to saturation and analysed in a 5-FOA dilution experiment as described above. The results, shown representatively for HA-*EZH2* subclones in Figure 5B, indicated that telomeric *URA3* repression was stably maintained in all of the subclones tested. In contrast, the non-silent state was metastable and subclones from initially 5-FOA-sensitive colonies could switch to a repressed state (Figure 5B, bottom panel, and Table II). This epigenetic switching to *URA3* repression was clearly dependent upon the presence of the mammalian *E(z)* homologues, since neither subclones from a putative suppressor mutation nor control subclones that only contained an empty CEN4-*LEU* vector promoted growth on 5-FOA. Together, these data illustrate the involvement of *E(z)* genes in inducing repressive chromatin structures and demonstrate function of a mammalian protein to regulate telomeric silencing in *S.cerevisiae* chromatin.

Discussion

Our analysis of mammalian homologues of the *Drosophila* Pc-G gene *E(z)* demonstrates a surprising and unprecedented functional conservation of gene repression in eukaryotic chromatin: extra gene copies of human *EZH2* enhance PEV in *Drosophila* and *EZH2* or murine *Ezh1* restore TPE in yeast mutants that are null for *SET1*, an *S.cerevisiae* SET domain gene. Together, these data provide a functional link between the stable maintenance of gene repression mediated by a Pc-G gene and the organization of inactive chromatin domains. Moreover, since *E(z)* homologues and *SET1* share the SET domain, our results also implicate an important functional role of the SET domain in participating in an evolutionarily conserved mechanism(s) that regulates repressive chromatin states in eukaryotic chromatin.

E(z) homologues mediate gene repression in euchromatin and heterochromatin

E(z) was initially identified as a mutation that increased *white* gene repression through genetic interactions with the transcription factor *zeste* (Kalisch and Rasmuson, 1974; Jones and Gelbart, 1990; Phillips and Shearn, 1990). Because *E(z)* restricts expression boundaries of homeotic selector genes (Jones and Gelbart, 1990; Phillips and Shearn, 1990), it has been classified as a Pc-G gene (but see also LaJeunesse and Shearn, 1996). More recently, it has been shown that *E(z)* is also involved in the repression of some of the early acting segmentation/gap genes (Moazed and O'Farrell, 1992; Pelegri and Lehmann, 1994). Together, these observations could indicate recruitment of *E(z)* protein by bound transcription factors, resulting in the downregulation of targeted promoters in euchromatin. Consistent with this interpretation, LexA fusions encoding several *Drosophila* Pc-G proteins have been shown to reduce LexA-dependent reporter activity in transient transfection studies (Bunker and Kingston, 1994).

E(z) null alleles, on the other hand, have been shown to reduce immunostaining of several Pc-G proteins in

polytene chromosomes (Rastelli *et al.*, 1993; Platero *et al.*, 1996) and to result in a general decondensation of chromatin structure (Rastelli *et al.*, 1993) that is reflected by increased chromosome breakage (Gatti and Baker, 1989). Here, we show that extra gene copies of either *Drosophila E(z)* or human *EZH2* result in the enhancement of heterochromatin-induced gene inactivation in position effect rearrangements (Figure 4 and Table I). Further, the $E(z)^5$ null and the $E(z)^1$ antimorphic mutants (only one functional gene copy) are weak suppressors of PEV (Table I). This haplo-suppressor/triplo-enhancer dosage effect of *E(z)* (and most likely also of *EZH2*) is reminiscent of gene products that have been postulated to be key regulators in the organization of chromatin domains (Reuter and Spierer, 1992). Only four *Drosophila* chromatin regulators that belong to the class of haplo-suppressors/triplo-enhancers have been described to date and include *Su(var)2-5* (which encodes the heterochromatin-associated protein *HP-1*) (Eissenberg *et al.*, 1990), *Su(var)3-7* (Reuter *et al.*, 1990), *Su(var)3-9* (Tschiersch *et al.*, 1994), and now *E(z)*. By contrast, most of the Pc-G genes, including *Polycomb* itself, do not affect w^{m4h} expression (Grigliatti, 1991; G.Reuter, unpublished).

Possibly, some of the conserved protein domains of *E(z)* (see Figure 2) may be surfaces for recruitment to selected promoters in euchromatin, whereas others could be involved in remodelling repressive chromatin structures in heterochromatin. Such dual roles in the regulation of both euchromatin and heterochromatin have been demonstrated for the positive chromatin regulators *trx*-like (which encodes the nucleosome-disrupting GAGA transcriptional activator; Farkas *et al.*, 1994) and for the putative transcriptional regulator *E(var)3-93D* (Dorn *et al.*, 1993). Moreover, hierarchies of heterochromatin appear to exist—and may also relate to different levels of euchromatic organization—since some suppressors of PEV relieve *white* gene repression in pericentric but not telomeric heterochromatin in *Drosophila* (Wallrath and Elgin, 1995).

Immunostaining of several Pc-G proteins on polytene chromosomes indicated localization at cytological positions that correspond to some modifiers of PEV (DeCamillis *et al.*, 1992; Rastelli *et al.*, 1993), suggesting that Pc-G genes would only have an indirect effect on heterochromatin, particularly since Pc-G proteins do not appear to decorate heterochromatic regions. However, recent immunodetection of *E(z)* protein at polytene chromosomes demonstrated a complex staining pattern that in addition to discrete locations also includes a more general distribution and masked positions within several Pc-G complexes (Carrington and Jones, 1996). Moreover, recruitment of Pc-G complexes to ectopic heterochromatic positions has been shown to depend on *E(z)* (Platero *et al.* 1996). In fact, since human *EZH2* both enhances PEV in *Drosophila* and restores telomeric silencing in *S.cerevisiae* (see below), we favour the alternative view that *E(z)*-related proteins display multiple functions that may also include a more direct structural involvement in the reorganization of repressive chromatin domains.

Conserved gene silencing from yeast to man

In *S.cerevisiae*, several silencing factors, namely the Silent Information Regulators *SIR2–4*, have been shown to be crucial for the maintenance of repressive chromosomal domains at both the silent mating type loci (*HM*-loci) and at the telomeres, whereas other modifying proteins have more locus-specific functions (Aparicio *et al.*, 1991). Variegating gene expression at yeast telomeres is thought to result from competition between transcriptional activators and silencing complexes (Aparicio and Gottschling, 1994), and may act via a spreading mechanism (Renauld *et al.*, 1993; Hecht *et al.*, 1996) reminiscent of PEV in *Drosophila* pericentric heterochromatin. In addition, *SIR3* and *SIR4* appear to interact with histones, resulting in the establishment of repressive multimeric chromatin complexes (Hecht *et al.*, 1995) that may be prototypes of restrictive chromatin domains mediated by *Drosophila* Pc-G genes. These intriguing parallels between maintenance mechanisms for repressive transcriptional states in single-celled and multicellular organisms have been extended by the discovery of a *SIR2* gene family that is conserved from bacteria to mammals (Brachmann *et al.*, 1995).

The potential of human *EZH2* or murine *Ezh1* to restore silencing in *S.cerevisiae* mutants is specific for loss of the endogenous *SET1* gene, since *EZH2* cannot suppress *sir2*, *sir3* or *sir4* single mutants (data not shown). Further, overexpression of *EZH2* or *SET1* in telomeric marker strains that are wild type for the *SIR* genes does not interfere with TPE (see Table II). These data demonstrate that the *S.cerevisiae SET1* and mammalian *E(z)* proteins do not bypass a requirement for *SIR* function, but possibly modulate interactions with accessory components of yeast chromatin. In agreement with this interpretation, the *EZH2* SET domain appears to represent a protein interaction surface, and preliminary data from a structure–function analysis of *EZH2* indicate that deletion of the SET domain abolishes the ability of *EZH2* to rescue telomeric silencing in *set1*Δ mutants (G.Laible, A.Wolf and T.Jenuwein, unpublished observation). Interestingly, complementation of *set1*Δ-dependent telomeric silencing by the mammalian *E(z)* proteins (Figure 5 and Table II) displays an epigenetic component that is very reminiscent of the *EZH2*- or *E(z)*-mediated paternal imprinting of PEV in *Drosophila* (Table I). Together, these results suggest that *E(z)*-related proteins may function both in the establishment and propagation of repressed chromatin domains. Recently, a *Drosophila* homologue of the *S.cerevisiae ORC2* gene was observed to rescue mating-type silencing in a similar epigenetic fashion (Ehrenhofer-Murray *et al.*, 1995).

The SET protein family of chromatin regulators

Based on our findings of functional conservation between mammalian (*EZH2* and *Ezh1*) and an *S.cerevisiae* (*SET1*) SET domain protein, we suggest that the SET domain represents an ancient protein module that is intrinsically involved in chromatin-controlled gene regulation. Figure 6 summarizes the relationship of representative SET protein family members according to sequence identities within the 130 amino acids comprising the SET domain (but ignoring other structural similarities) that results in a classification of SET domain proteins into four different subgroups. The SET domain of *E(z)*-related genes (subgroup I) appears most similar to that of subgroup II genes which include the *trithorax* homologues *trx*/*HRX*, an open reading frame (C26E6.10) in *Caenorhabditis elegans* and *S.cerevisiae SET1*. The SET domain genes *S.cerevisiae*

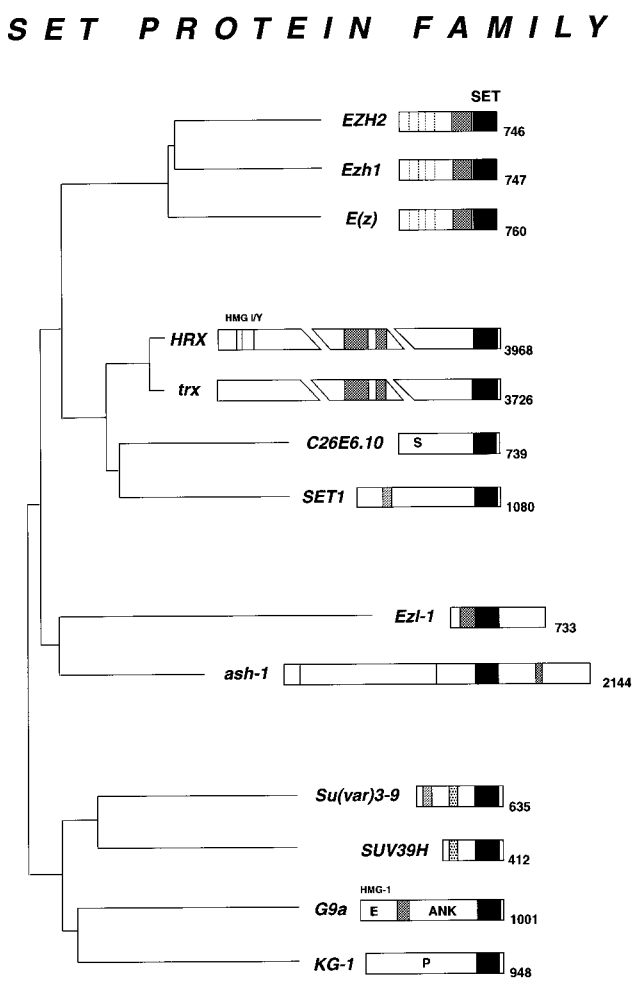


Fig. 6. SET domain proteins fall into four distinct subgroups. SET domain proteins currently comprise ~30 members, ranging from fungi to mammals. Shown here is the relationship of representative family members based on sequence identity within the 130 amino acid carboxy-terminal SET domain. Ignoring other structural similarities, only the SET domains of the indicated proteins were compared using the PILEUP and PLOT programs of the GCG software package, resulting in the distinction of four subgroups as indicated. An apparent intron sequence in the human KG-1 SET domain (Nomura *et al.*, unpublished; GenBank accession number D31891) has been removed prior to the alignment. The SET domain is indicated as a black box. Dashed bars represent novel conserved domains between human *EZH2*, murine *Ezh1* and *Drosophila E(z)* proteins. Hatched bars are cysteine-rich regions in the *E(z)* gene family and in *S.cerevisiae Ezl-1* (GenBank accession number Z49444) or encode putative zinc fingers in *Drosophila trx* (Stassen *et al.*, 1995), human *HRX* (Tkachuk *et al.*, 1992; Ma *et al.*, 1993) or a PHD finger in *Drosophila ash-1* (Tripoulas *et al.*, 1996). The chromo box in *Drosophila Su(var)3-9* (Tschiersch *et al.*, 1994) and in its human homologue *SUV39H* (G.Laible and T.Jenuwein, unpublished) is shown as a stippled bar, and a putative GTP binding motif in *Su(var)3-9* and in *S.cerevisiae SET1* is indicated by a light-shaded box. A/T hooks in *HRX* and *ash-1* are indicated by a vertical line. Ankyrin repeats (ANK) and an acidic stretch (E) resembling HMG-1 are shown for human G9a (Milner and Campbell, 1993); P, proline-rich region. A serine-rich region (S) is indicated for the *C.elegans* open reading frame C26E6.10 (GenBank accession number U13875). Proteins are shown roughly to scale, and numbers refer to amino acid positions.

Ezl-1 and *Drosophila ash-1* (which seems to be related to *trx*; Tripoulas *et al.*, 1996) appear to be unique, because in both of these genes the SET domain is not located at the carboxy terminus (subgroup III). Finally, SET domains of subgroup IV genes comprise those present in *Su(var)3-9* related proteins and human G9a and KG-1.

Possibly, additional structural motifs, like the cysteine-rich cluster (Jones and Gelbart, 1993) in *E(z)* homologues, putative zinc fingers and A/T-hooks in *trx*/*HRX* (Tkachuk *et al.*, 1992; Ma *et al.*, 1993; Stassen *et al.*, 1995), a PHD-finger in *ash-1* (Tripoulas *et al.*, 1996), the chromo box in *Su(var)3-9* (Tschiersch *et al.*, 1994) or ankyrin repeats in G9a (Milner and Campbell, 1993), may impart distinct target specificities to SET domain proteins as they function in chromatin-dependent gene regulation. Moreover, differences in the primary structure of the SET domain itself may also provide distinct specificities to individual SET domains. In agreement with this interpretation and consistent with the more distant relationship of the SET domain in *Su(var)3-9* with respect to *EZH2* and *SET1*, the human *Su(var)3-9* homologue (*SUV39H*) did not rescue *set1*Δ-dependent telomeric silencing in *S.cerevisiae* (data not shown).

Although we are currently only beginning to understand the molecular functions of SET domain genes, our analysis of mammalian *E(z)* proteins provides a direct demonstration for conserved gene silencing in eukaryotes and suggests that the SET domain, like the 50 amino acids comprising the chromo box (Messmer *et al.*, 1992; Platero *et al.*, 1995), may be of primary importance in organizing chromatin domains that regulate gene expression in eukaryotic chromatin.

Materials and methods

Molecular cloning of mammalian E(z) homologues

A DNA probe encoding the conserved SET domain of *E(z)* (Jones and Gelbart, 1993) was amplified by the polymerase chain reaction (PCR) from 1 μg of *Drosophila melanogaster* DNA (Clontech) using the two primers *E(z)*-1910 (5′-actgaattCGGCTGGGGCATCTTTCTTAAGG) and *E(z)*-2280 (5′-actctagACAATTTCCATTTCACGCTCTATG). A human B-cell-specific (BJAB) λgt10 cDNA library was kindly provided by M.Busslinger. A total of 5×10^5 plaque-forming units (p.f.u.) were plated and transferred to GeneScreen membranes as described (Sambrook *et al.*, 1989). The membranes were pre-hybridized for 30 min at 50°C in Church buffer, and hybridized O/N at 50°C with the *E(z)* SET domain DNA probe. Following reduced stringency washes for 10 min in 2× SSC, 1% SDS at room temperature (RT) and 10 min at 50°C, positive phages were identified by autoradiography. Single, well-isolated plaques were obtained after two rounds of rescreening.

Mouse cDNA clones were isolated from a Balb/c brain λgt11 cDNA library (Clontech) by screening, under reduced stringency as above, with a 390 bp *Eco*RI–*Xho*I cDNA fragment that encodes the *EZH2* SET domain. Mouse genomic clones were isolated from a C57Bl/6 cosmid library (Stratagene) by screening with a 498 bp PCR-amplified DNA probe that encodes amino acids 242–408 of *Ezh1*. A membrane representing ~3×10^4 cosmid colonies (kindly provided by D.Barlow) was hybridized O/N at 65°C in Church buffer and washed under high stringency (10 min washes in 2× SSC, 1% SDS at RT and at 65°C, followed by a 10 min wash in 0.2× SSC, 1% SDS at 50°C). Positive clones were identified on frozen masterplates and rescreened until single colonies could be identified.

Sequence analysis

The cDNA inserts of recombinant phages or DNA fragments from the genomic cosmid clone were subcloned into the polylinker of pBluescript KS (Stratagene) and sequenced on an automated sequencer (Applied Biosystems) using the dideoxy method. The complete sequence of independent isolates was determined by primer walking. Sequence data were analysed using the University of Wisconsin Genetics Computer Group sequence analysis software package (Devereux *et al.*, 1984). Homology searches were performed with the Blast (Altschul *et al.*, 1990) network service and also included sequence comparisons with mammalian ESTs (Bassett *et al.*, 1995).

Full-length cDNAs and 5′ RACE

A full-length human *EZH2* cDNA of 2.6 kb was obtained after combining a 0.7 kb *Sal*I–*Bgl*II DNA fragment comprising 5′ untranslated and amino-terminal sequences from clone λ13 with a Bluescript-based subclone of λ37. Missing 5′ sequences of murine *Ezh1* were generated by RACE amplification (Marathon cDNA amplification kit; Clontech) of 1 μg poly(A)$^+$ RNA from the murine B-cell line J558L. First-strand cDNA synthesis was primed with a lock-docking oligo(dT) primer. After second-strand synthesis, the cDNA was amplified using a 5′ adapter oligo and the gene-specific primer *Ezh1*-200 (5′-CATGGACGCA-ATGGCGCTGAAGATCATGTC). 5′ RACE amplification was performed with a nested primer pair. The resultant 0.6 kb PCR product was trimmed with *Not*I and *Sca*I, and combined in several subcloning steps with a Bluescript-based 3.4 kb subclone of λm31. The complete human *EZH2* and murine *Ezh1* cDNAs were sequenced over their entire length to confirm the authenticity of the encoded sequences.

DNA blot analysis

Mouse genomic DNA was prepared from kidney, brain and heart of young 129/Sv males. Human male placenta DNA (Sigma) and *Drosophila melanogaster* DNA (Clontech) was purchased from the indicated suppliers. 15 μg of genomic DNA was digested O/N, fractionated in a 0.8% agarose gel, transferred to a GeneScreen membrane and cross-linked under UV light. Following hybridization O/N at 42°C in a buffer containing 20% formamide, 5× SSC, 1× Denhardt's, 20 mM $NaPO_4$ (pH 6.7), 0.1 mg denatured salmon sperm DNA (Sigma) per millilitre, the membrane was washed at reduced stringency for 10 min at RT in 2× SSC, 0.1% SDS and for 30 min at 65°C in 1× SSC, 0.1% SDS. DNA probes encoding the carboxy-terminal SET domains of mammalian *E(z)* homologues were a 390 bp *Eco*RI–*Xho*I fragment from the *EZH2* cDNA or a 330 bp *Eco*RI–*Xho*I fragment from a modified *Ezh1* cDNA subclone in which nucleotides 4 bp downstream of the stop codon have been converted to an *Xho*I site.

In vitro differentiation of embryonic stem cells

A total of 1.5×10^6 feeder-independent mouse (129/Sv) embryonic stem cells (D3) were plated onto gelatin-coated bacterial dishes (150 cm^2) in high-glucose Dulbecco's modified Eagle's medium (DMEM), supplemented with 15% fetal calf serum (Boehringer Mannheim), 2 mM L-glutamine (Gibco), 1% non-essential amino acids (Gibco), 0.1 mM β-mercaptoethanol and 1% conditioned medium containing leukaemia inhibiting factor (LIF). LIF-conditioned medium was harvested from LIFcα-5 producer cells (Smith *et al.*, 1988). Differentiation was induced with 0.1 μM all-*trans* retinoic acid (RA) (Sigma). Following induction for 2 days, floating embryoid bodies (EB) were removed and cultivated (in the absence of RA) for another 5 days.

RNA isolation and analysis

Total RNA was isolated by isopycnic centrifugation as described (Sambrook *et al.*, 1989). Poly(A)$^+$ RNA was purified using a PolyATrack kit (Promega). For RNA blot analysis, 5 μg of poly(A)$^+$ RNA were fractionated on a formaldehyde–agarose gel, transferred to a GeneScreen plus membrane in 10× SSC and baked for 2 h at 80°C. The membrane was sequentially hybridized under stringent Church conditions with DNA probes that comprised a 1.9 kb PCR-amplified fragment encoding amino acids 173–747 (plus 180 bp of 3′ untranslated sequences) of *Ezh1*, the 2.6 kb full-length *EZH2* cDNA and a 0.3 kb DNA fragment that is specific for a murine β-actin cDNA (Alonso *et al.*, 1986). RNase protection analysis was carried out as described (Sambrook *et al.*, 1989). 10 μg of total RNA were incubated at 60°C with a riboprobe that specifically protects 320 bp of the *Ezh1* SET domain or at 58°C with a riboprobe that protects 270 bp of *Ezh2* amino-terminal sequences. As a control, a riboprobe that protects 156 bp of murine S16 rRNA sequences (Urbanek *et al.*, 1994) was included in the hybridizations. The murine *Ezh1* SET domain riboprobe was derived from a 370 bp PCR-amplified fragment (encoding amino acids 637–743 of *Ezh1*) that has been subcloned into pGEM-3Zf (Promega). After linearization with *Pvu*II, antisense RNA was internally labelled by trancription with SP6 RNA polymerase in the presence of [^{32}P]GTP and the full-length riboprobe of 470 nucleotides was purified from a denaturing urea–polyacrylamide gel. Following a similar strategy, an *Ezh2*-specific riboprobe was generated that protects sequences encoding amino acids 4–92 of *Ezh2*.

Epitope tag and expression plasmids

A *Not*I site just downstream of the start codon in *EZH2* was inserted by replacing a 0.7 kb *Bam*HI–*Bgl*II DNA fragment with a PCR product that was generated with a 3′ primer and the modifying oligonucleotide 5′-ataggattcATAATCATGggcggccgcACTGGGAAGAAATCTGAG-AAGGG (start codon underlined). Overlapping oligonucleotides encoding an in-frame triple haemagglutinin (HA) tag preceding six histidines were ligated, cleaved with *Not*I and inserted into the *Not*I-digested pEZH2 derivative. The resulting epitope-tagged variant (HA-*EZH2*) extends the amino terminus by 44 amino acids.

For *Drosophila* expression plasmids, *EZH2* and *E(z)* coding sequences were converted by PCR amplification into *Not*I fragments and inserted into the *Not*I site of the P element, rosy$^+$ vector pNHT4 (Schneuwly *et al.*, 1987) which directs expression from the hsp70 promoter. The sequence of the primers for human *EZH2* was upstream 5′-ctaagcggccgcAGGCAGTGGAGCCCCGGC and downstream 5′-attagcggccgCTGGTACAAAACACTTTGC. The *Not*I primer pair for *Drosophila E(z)* was 5′-ataagaatgcggccgcgtcgacTATAAAGTCCCTCG-AAGGCATTATGAATAGC and 5′-ataagaatgcggccgcgtcgacGATTTT-CATAAGTAGACTCGTTATC (start and stop codons underlined). The genomic *E(z)* sequences are expressed from the natural promoter and the respective construct has been described (Jones and Gelbart, 1993).

For yeast expression plasmids, a 0.7 kb *Bam*HI–*Eco*RI DNA fragment encoding the Gal1-10 promoters from Gal-CLB5 (Schwob and Nasmyth, 1993) was inserted into the polylinker of the low-copy CEN4-*LEU* plasmid YCplac111 (Gietz and Sugino, 1988), resulting in plasmid YCplac111-G1/10. A 2.4 kb *Sal*I–*Xho*I DNA fragment encoding *EZH2* or a 2.5 kb *Sal*I–*Xho*I DNA fragment encoding HA-*EZH2* were transferred into the *Sal*I site of YCplac111-G1/10 in an orientation that allows transcription of the cDNAs by the Gal1 promoter. *Ezh1* coding sequences were converted by PCR amplification into a 2.3 kb *Sal*I fragment and inserted into the *Sal*I site of YCplac111-G1/10 in the same orientation as *EZH2* or HA-*EZH2*. The *Ezh1*-specific primer pair was 5′-tatcccgggtcgaCCATGAGGAAAATGGATATAGCAAGTCCCC and 5′-tatctcgagtcgactAGGGCTAGACGTCCGTTTCCCTCTCG (start and stop codons underlined). The *SET1* control plasmid contained a 3.8 kb *Asp*718–*Hin*dIII genomic fragment inserted into the polylinker of the CEN6-*LEU* plasmid pRS315 (Nislow *et al.*, submitted). For all of the fly and yeast expression plasmids, the 5′ and 3′ boundaries of the respective cDNAs were confirmed by sequencing.

Drosophila genetics

All *Drosophila* stocks were maintained under standard conditions. A description of chromosomes and mutations is found in Lindsley and Zimm (1992). P element-mediated germline transformation was carried out by standard methods (Rubin and Spradling, 1982), using a ry^{506} fly strain as recipient.

The enhancer effect of the transgenes on w^{m4h} PEV was quantified by measuring the absorbance (at 480 nm) of red eye pigments in w^{m4h}/Y; T(2;3)ap^{Xa} Su(var)2-1^{01} and w^{m4h}/Y; Cy/+;Sb/+ male offspring after crossing w^{m4h}; Cy/T(2;3)ap^{Xa} Su(var)2-1^{01}/Sb females to +/Y or w^{m4h}/Y males carrying the respective transgenes heterozygous to CyO or TM3, ry^{RK} Sb e balancer chromosomes. By contrast, the loss-of-function mutation *E(z)*^{5} (Phillips and Shearn, 1990) and the antimorphic mutation *E(z)*^{1} (Kalisch and Rasmuson, 1974) increased the concentration of w^{m4h} red eye pigments in both Cy/+; Sb/*E(z)* and +/T(2;3)ap^{Xa} Su(var)2-1^{01}/*E(z)* genotypes, whereas the temperature-sensitive *E(z)*^{12} allele (MY939) and the hypomorphic *E(z)*^{15} mutation (NU808) (Phillips and Shearn, 1990) had no effect at 25°C.

Modification of the *zeste*^{1}*–white* eye colour by extra copies of *EZH2* or *E(z)* was analysed in transgenic female and male offspring heterozygous for Dp(1;1)z^{59d15}. The antimorphic *E(z)*^{1} mutation (Kalisch and Rasmuson, 1974) has been used for comparison in control matings.

Remobilization of the P[hspE(z)ry$^+$] transgene in line D was activated in males carrying the TM3, ry^{RK} Sb e P[(ry$^+$)Δ2-3] (99B) transposase source. After a cross to w^{m4h}; Cy/T(2;3)ap^{Xa} Su(var)2-1^{01}/Sb females, exceptional w^{m4h}/Y; +/T(2;3)ap^{Xa} Su(var)2-1^{01}/P[hspE(z)ry$^+$] male offspring with a significant enhancer effect were selected. In all proven exceptions (eight out of 2329 F_1 males), red eye pigment measurements from w^{m4h}/Y; +/T(2;3)ap^{Xa} Su(var)2-1^{01}/P[hspE(z)ry$^+$] males indicated a significant reduction in the proportion of red eye pigments.

Yeast transformations and 5-FOA dilution assays

Saccharomyces cerevisiae cells were grown at 30°C in complete (YPD) or selective media. One gram of 5-FOA (Diagnostic Chemicals) per litre of medium was added to determine resistance to 5-FOA. Yeast transformations were performed by the lithium–acetate protocol. The yeast strains used in this study are UCC1001 [*MAT***a** *his*3-Δ200 *trp*1-Δ1 *leu*2-Δ1 *ura*3-52 *ade*2-101 *lys*2-801 *adh*4::*URA3* (*URA3* @ VII-L), Renauld *et al.*, 1993], LPY1297 (isogenic to UCC1001, except *set1*Δ::*HIS3*), LPY1683 (*MAT***a** *his*3-Δ200 *trp*1-Δ63 *leu*2-Δ1 *ura*3-52

ade2-101 *lys2*-801 *adh4*::*URA3* (*URA3* @ VII-L) DIA5-1 (*ADE2* @ V-R); a derivative of UCC3505, Singer and Gottschling, 1994) and LPY1727 (isogenic to LPY1683, except *set1*Δ::*HIS3*).

5-FOA dilution assays were carried out by growing transformed cells picked from single colonies in –LEU/2% raffinose media for 3–4 days at 30°C until the culture reached saturation. Cells were concentrated by centrifugation, counted in an automatic cell counter (CASY1) and adjusted to a concentration of 200 000 cells/μl. The cell suspension was added to a 96-well microtitre plate and serial 1:10 or 1:5 dilutions were made by using an eight-channel pipettor. Dilutions were transferred onto YPD and selective media plates with a stainless steel pin replicator. Colonies were inspected after 2–4 days of incubation at 30°C.

Accession numbers

The murine *Ezh1* (accession number U60453) and human *EZH2* (accession number U61145) cDNA sequences have been deposited in GenBank.

Acknowledgements

We would like to thank Meinrad Busslinger for the human BJAB cDNA library, Rick Jones and Allen Shearn for the *E(z)* cDNA and for *(E)z* mutant fly strains, Dan Gottschling for *SIR* mutant yeast strains, Ken Abel for DNA sequence information on human *EZH1* prior to publication, Gotthold Schaffner for oligonucleotide synthesis, Robert Kurzbauer and Ivan Botto for sequence analysis, and Meinrad Busslinger and Hartmut Beug for helpful comments and critical reading of the manuscript. This work was supported by the IMP and, in part, by the Austrian Industrial Research Promotion Fund. G.R. and R.D. were supported by grants from the Deutsche Forschungsgemeinschaft (DFG-Re911/1-3 and DFG-Do407/1-2). C.N. was a postdoctoral fellow of the American Cancer Society and research in L.P.'s laboratory was supported by grants from the National Science Foundation (MCB9257685) and the Pew Charitable Trusts (T89-03508-044).

References

Abel,K. *et al.* (1996) Characterization of *EZH1*, a human homolog of *Drosophila Enhancer of zeste* near *BRCA1*. *Genomics*, **37**, 161–171.

Akasaka,T., Kanno,M., Balling,R., Mieza,M.A., Taniguchi,M. and Koseki,H. (1996) A role for *mel-18*, a *Polycomb* group-related vertebrate gene, during the antero-posterior specification of the axial skeleton. *Development*, **122**, 1513–1522.

Alkema,M.J., van der Lugt,N.M., Bobeldijk,R.C., Berns,A. and van Lohuizen,M. (1995) Transformation of axial skeleton due to overexpression of *bmi-1* in transgenic mice. *Nature*, **374**, 724–727.

Allshire,R.C., Javerzat,J.-P., Readhead,N.J. and Cranston,G. (1994) Position effect variegation at fission yeast centromeres. *Cell*, **76**, 157–169.

Alonso,S., Minty,A., Bourlet,Y. and Buckingham,M. (1986) Comparison of three actin coding sequences in the mouse; evolutionary relationships between the actin genes of warm-blooded vertebrates. *J. Mol. Evol.*, **23**, 11–22.

Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.

Aparicio,O.M. and Gottschling,D.E. (1994) Overcoming telomeric silencing: A trans-activator competes to establish gene expression in a cell cycle-dependent way. *Genes Dev.*, **8**, 1133–1146.

Aparicio,O.M., Billington,B.L. and Gottschling,D.E. (1991) Modifiers of position effect are shared between telomeric and mating-type loci in *S.cerevisiae*. *Cell*, **66**, 1279–1287.

Basset,D.R.,Jr, Boguski,M.S., Spencer,F., Reeves,R., Goebl,M. and Hieter.P. (1995) Comparative genomics, genome cross-referencing and XREFdb. *Trends Genet.*, **11**, 372–373.

Brachmann,B.C., Sherman,J.M., Devine,S.E., Cameron,E.E., Pillus,L. and Boeke,J.D. (1995) The *SIR2* gene family, conserved from bacteria to humans, functions in silencing, cell cycle progression, and chromosome stability. *Genes Dev.*, **9**, 2888–2902.

Bunker,C.A. and Kingston,R.E. (1994) Transcriptional repression by *Drosophila* and mammalian *Polycomb* group proteins in transfected mammalian cells. *Mol. Cell. Biol.*, **14**, 1721–1732.

Carrington,E.A. and Jones,R.S. (1996) The *Drosophila Enhancer of zeste* gene encodes a chromosomal protein: examination of wild-type and mutant protein distribution. *Development*, **122**, 4073–4083.

Csink,A.K. and Henikoff,S. (1996) Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature*, **381**, 529–531.

DeCamillis,M., Cheng,N.S., Pierre,D. and Brock,H.W. (1992) The *polyhomeotic* gene of *Drosophila* encodes a chromatin protein that shares polytene chromosome-binding sites with *Polycomb*. *Genes Dev.*, **6**, 223–232.

Dernburg,A.F., Broman,K.W., Fung,J.C., Marshall,W.F., Philips,J., Agard,D.A. and Sedat,J.W. (1996) Pertubation of nuclear architecture by long-distance chromosome interactions. *Cell*, **85**, 745–759.

Devereux,J., Haeberli,P. and Smithies,O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.*, **12**, 387–395.

Dorn,R., Heymann,S., Lindigkeit,R. and Reuter,R. (1986) Suppressor mutation of position effect variegation loci in *Drosophila melanogaster* affecting chromatin properties. *Chromosoma*, **93**, 398–403.

Dorn,R., Krauss,V., Reuter,G. and Saumweber,H. (1993) The enhancer of position-effect variegation of *Drosophila*, *E(var)3-93D*, codes for a chromatin protein containing a conserved domain common to several transcriptional regulators. *Proc. Natl Acad. Sci. USA*, **90**, 11376–11380.

Efstradiadis,A. (1994) Parental imprinting of autosomal mammalian genes. *Curr. Opin. Genet. Dev.*, **4**, 265–280.

Ehrenhofer-Murray,A.E., Gossen,M., Pak,D.T.S., Botchan,M.R. and Rine,J. (1995) Separation of origin recognition complex functions by cross-species complementation. *Science*, **270**, 1671–1674.

Eissenberg,J.C., James,T.C., Foster-Harnett,D.M., Harnett,T., Ngan,T. and Elgin,S.C.R. (1990) Mutation in a heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA*, **87**, 9923–9927.

Eissenberg,J.C., Morris,G.D., Reuter,G. and Hartnett,T. (1992) The heterochromatin-associated protein *HP-1* is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. *Genetics*, **131**, 345–352.

Eissenberg,J.C., Elgin,S.C.R. and Paro,R. (1995) Epigenetic regulation in *Drosophila*: a conspiracy of silence. In Elgin,S.C.R. (ed.), *Chromatin Structure and Gene Expression*. IRL Press, New York, pp. 147–171.

Farkas,G., Gausz,J., Galloni,M., Reuter,G., Gyurkovics,H. and Karch,F. (1994) The *Trithorax-like* gene encodes the *Drosophila* GAGA factor. *Nature*, **371**, 806–808.

Gatti,M. and Baker,B.S. (1989) Genes controlling essential cell-cycle functions in *Drosophila melanogaster*. *Genes Dev.*, **3**, 438–453.

Gietz,R.D. and Sugino,A. (1988) New yeast-*E. coli* shuttle vectors constructed with *in vitro* mutagenized yeast genes lacking six-base pair restriction sites. *Gene*, **74**, 527–534.

Gottschling,D.E., Aparicio,O.M., Billington,B.L and Zakian,V.A. (1990) Position effect at *S.cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell*, **63**, 751–762.

Grigliatti,T.A. (1991) Position effect variegation—an assay for nonhistone chromosomal proteins and chromatin assembly and modifying factors. *Methods Cell Biol.*, **35**, 587–627.

Gu,Y., Nakamura,T., Alder,H., Prasad,R., Canaani,O., Cimino,G. *et al.* (1992) The t(4;11) chromosome translocation of human acute leukemias fuses the *ALL-1* gene, related to *Drosophila trithorax*, to the AF-4 gene. *Cell*, **71**, 701–708.

Han,M. and Grunstein,M. (1988) Histone loss activates downstream promoters *in vivo*. *Cell*, **55**, 1137–1145.

Hecht,A., Laroche,T., Strahl-Bolsinger,S., Gasser,S.M. and Grunstein,M. (1995) Histone H3 and H4 N-termini interact with *SIR3* and *SIR4* proteins: A molecular model for the formation of heterochromatin in yeast. *Cell*, **80**, 583–592.

Hecht,A., Strahl-Bolsinger,S. and Grunstein,M. (1996) Spreading of transcriptional repressor *SIR3* from telomeric heterochromatin. *Nature*, **383**, 92–95.

Henikoff,S. (1994) A reconsideration of the mechanism of position effect. *Genetics*, **138**, 1–5.

Hobert,O., Sures,I., Ciossek,T., Fuchs,M. and Ullrich,A. (1996a) Isolation and developmental expression analysis of *Enx-1*, a novel mouse *Polycomb* gene. *Mech. Dev.*, **55**, 171–184.

Hobert,O., Jallal,B. and Ullrich,A. (1996b) Interaction of Vav with *ENX-1*, a putative transcriptional regulator of homeobox gene expression. *Mol. Cell. Biol.*, **16**, 3066–3073.

Jeppesen,P. and Turner,B.M. (1993) The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell*, **74**, 281–289.

Jones,R.S. and Gelbart,W.M. (1990) Genetic analysis of the *Enhancer of zeste* locus and its role in gene regulation in *Drosophila melanogaster*. *Genetics*, **126**, 185–199.

Jones,R.S. and Gelbart,W.M. (1993) The *Drosophila Polycomb*-group gene *Enhancer of zeste* contains a region with sequence similarity to *trithorax*. *Mol. Cell. Biol.*, **13**, 6357–6366.
Kalisch,W.E. and Rasmuson,B. (1974) Changes in *zeste* phenotype induced by autosomal mutations in *Drosophila melanogaster*. *Hereditas*, **78**, 97–104.
Karpen,G.H. (1994) Position-effect variegation and the new biology of heterochromatin. *Curr. Opin. Genet. Dev.*, **4**, 281–291.
LaJeunesse,D. and Shearn,A. (1996) *E(z)*: a *Polycomb* group gene or a *trithorax* group gene? *Development*, **122**, 2189–2197.
Lindsley,D.L. and Zimm,G.G. (1992). *The Genome of Drosophila melanogaster*. Academic Press, San Diego, CA.
Loo,S. and Rine,J. (1995) Silencing and heritable domains of gene expression. *Annu. Rev. Cell Biol. Dev.*, **11**, 519–548.
Ma,Q. *et al.* (1993) Analysis of the murine *All-1* gene reveals conserved domains with human *ALL-1* and identifies a motif shared with DNA methyltransferases. *Proc. Natl Acad. Sci. USA*, **90**, 6350–6354.
Messmer,S., Franke,A. and Paro,R. (1992) Analysis of the functional role of the *Polycomb* chromo domain in *Drosophila melanogaster*. *Genes Dev.*, **6**, 1241–1254.
Migeon,B. (1994) X-chromosome inactivation: molecular mechanisms and genetic consequences. *Trends Genet.*, **10**, 230–235.
Milner,C.M. and Campbell,R.D. (1993) The G9a gene in the human major histocompatibility complex encodes a novel protein containing ankyrin-like repeats. *Biochem. J.*, **290**, 811–818.
Moazed,D. and O'Farrell,P.H. (1992) Maintenance of the engrailed expression pattern by *Polycomb* group genes in *Drosophila*. *Development*, **116**, 805–810.
Moore,G.D., Procunier,G.D., Cross,D.P. and Grigliatti,T.A. (1979) Histone gene deficiencies and position effect variegation in *Drosophila*. *Nature*, **282**, 312–314.
Nimmo,E.R., Cranston,G. and Allshire,R.C. (1994) Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes. *EMBO J.*, **13**, 3801–3811.
Nislow,C., Ray,E. and Pillus,P. (1997) *SET1*, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol. Biol. Cell*, submitted.
Orlando,V. and Paro,R. (1995) Chromatin multiprotein complexes involved in the maintenance of transcription patterns. *Curr. Opin. Genet. Dev.*, **5**, 174–179.
Pelegri,F. and Lehmann,R. (1994) A role of *Polycomb* group genes in the regulation of gap gene expression in *Drosophila*. *Genetics*, **136**, 1341–1353.
Phillips,M.D. and Shearn,A. (1990) Mutations in *polycombeotic*, a *Drosophila Polycomb*-group gene, cause a wide range of maternal and zygotic phenotypes. *Genetics*, **125**, 91–101.
Pillus,L. and Grunstein,M. (1995) Chromatin structure and epigenetic regulation in yeast. In Elgin,S.C.R. (ed.), *Chromatin Structure and Gene Expression*. IRL Press, New York, pp. 123–146.
Pirrotta,V. (1995) Chromatin complexes regulating gene expression in *Drosophila*. *Curr. Opin. Genet. Dev.*, **5**, 466–472.
Platero,J.S., Harnett,T. and Eissenberg,J.C. (1995) Functional analysis of the chromo domain of *HP-1*. *EMBO J.*, **14**, 3977–3986.
Platero,J.S., Sharp,E.J., Adler,P.N. and Eissenberg,J.C. (1996) In vivo assay for protein-protein interactions using *Drosophila* chromosomes. *Chromosoma*, **104**, 393–404.
Rastelli,L., Chan,C.S. and Pirrotta,V. (1993) Related chromosome binding sites for *zeste*, suppressors of *zeste* and *Polycomb* group proteins in *Drosophila* and their dependence on *Enhancer of zeste* function. *EMBO J.*, **12**, 1513–1522.
Renauld,H., Aparicio,O.M., Zierath,P.D., Billington,B.L., Chhablani,S.K. and Gottschling,D.E. (1993) Silent domains are assembled continuously from the telomere and are defined by promoter distance and strength, and by *SIR3* dosage. *Genes Dev.*, **7**, 1133–1145.
Reuter,G. and Spierer,P. (1992) Position effect variegation and chromatin proteins. *BioEssays*, **14**, 605–612.
Reuter,G., Giarre,M., Farah,J., Gausz,J., Spierer,A. and Spierer,P. (1990) Dependence of position-effect variegation in *Drosophila* on dose of a gene encoding an unusual zinc-finger protein. *Nature*, **344**, 219–223.
Rubin,G.M. and Spradling,A.C. (1982) Genetic transformation of *Drosophila* with transposable element vectors. *Science*, **218**, 348–353.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Schneuwly,S., Klemenz,R. and Gehring,W.J. (1987) Redesigning the body plan of *Drosophila* by ectopic expression of the homeotic gene *Antennapedia*. *Nature*, **325**, 816–818.
Schwob,E. and Nasmyth,K. (1993) *CLB5* and *CLB6*, a new pair of B cyclins involved in DNA replication in *Saccharomyces cerevisiae*. *Genes Dev.*, **7**, 1160–1175.
Simon,J., Chiang,A., Bender,W., Shimell,M.J. and O'Connor,M. (1993) Elements of the *Drosophila Bithorax* complex that mediate repression by *Polycomb* group genes. *Dev. Biol.*, **158**, 131–144.
Singer,M.S. and Gottschling,D.E. (1994) *TLC1*: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science*, **266**, 404–409.
Smith,A.G., Heath,J.K., Donaldson,D.D., Wong,G.G., Moreau,J., Stahl,M. and Rogers,D. (1988) Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature*, **336**, 688–690.
Stassen,M.J., Bailey,D., Nelson,S., Chinwalla,V. and Harte,P.J. (1995) The *Drosophila trithorax* protein contains a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. *Mech. Dev.*, **52**, 209–223.
Tartof,K.D., Hobbs,C. and Jones,M. (1984) A structural basis for variegating position effects. *Cell*, **37**, 869–878.
Tkachuk,D.C., Kohler,S. and Cleary,M.L. (1992) Involvement of a homolog of *Drosophila trithorax* by 11q23 chromosomal translocations in acute leukemias. *Cell*, **71**, 691–700.
Tripoulas,N., LaJeunesse,D., Gildea,J. and Shearn,A. (1996) The *Drosophila ash1* gene product, which is localized at specific sites on polytene chromosomes contains a SET domain and a PHD finger. *Genetics*, **143**, 913–928.
Tschiersch,B., Hofmann,A., Krauss,V., Dorn,R., Korge,G. and Reuter,G. (1994) The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.*, **13**, 3822–3831.
Urbanek,P., Wang,Z.Q., Fetka,I., Wagner,E.F. and Busslinger,M. (1994) Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking *Pax5/BSAP*. *Cell*, **79**, 901–912.
van der Lugt,N.M. *et al*. (1994) Posterior transformation, neurological abnormalities, and severe hematopoietic defects in mice with a targeted deletion of the *bmi-1* proto-oncogene. *Genes Dev.*, **8**, 757–769.
Wallrath,L.L. and Elgin,S.C.R. (1995) Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev.*, **9**, 1263–1277.
Wedeen,C., Harding,K. and Levine,M. (1986) Spatial regulation of *Antennapedia* and *Bithorax* gene expression by the *Polycomb* locus in *Drosophila*. *Cell*, **44**, 739–748.
Weiler,K.S. and Wakimoto,B.T. (1995) Heterochromatin and gene expression in *Drosophila*. *Annu. Rev. Genet.*, **29**, 577–605.
Yu,B.D., Hess,J.L., Horning,S.E., Brown,G.A.J. and Korsmeyer,S.J. (1995) Altered *HOX* expression and segmental identity in *Mll*-mutant mice. *Nature*, **378**, 505–508.
Zuckerkandl,E. (1974) A possible role of 'inert' heterochromatin in cell differentiation. *Biochimie*, **56**, 937–954.

Received on November 22, 1996; revised on February 28, 1997

Note added in proof

While this manuscript was in press, the *curly leaf* mutant was shown to encode a plant member of the *E(z)* gene family that is involved in regulating homeotic gene expression in *Arabidopsis*, thereby extending conserved mechanism(s) controlling gene activities in eukaryotic chromatin (Goodrich,J., Puangsomlee,P., Martin,M., Long,D., Meyerowitz,E.M. and Coupland,G. (1997) A *Polycomb*-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature*, **386**, 44–51.