CMLS, Cell. Mol. Life Sci. 54 (1998) 80–93
1420-682X/98/010080-14 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

SET domain proteins modulate chromatin domains in eu- and heterochromatin

T. Jenuwein[a,*], G. Laible[a], R. Dorn[b] and G. Reuter[b]

[a]Research Institute of Molecular Pathology (IMP), Dr. Bohrgasse 7, A-1030 Vienna (Austria), Fax +43 1 798 7153
[b]Institute of Genetics, Martin-Luther University of Halle, Domplatz 1, D-06108 Halle (Germany)

Abstract. The SET domain is a 130-amino acid, evolutionarily conserved sequence motif present in chromosomal proteins that function in modulating gene activities from yeast to mammals. Initially identified as members of the *Polycomb*- and *trithorax*-group (Pc-G and trx-G) gene families, which are required to maintain expression boundaries of homeotic selector (HOM-C) genes, SET domain proteins are also involved in position-effect-variegation (PEV), telomeric and centromeric gene silencing, and possibly in determining chromosome architecture. These observations implicate SET domain proteins as multifunctional chromatin regulators with activities in both eu- and heterochromatin – a role consistent with their modular structure, which combines the SET domain with additional sequence motifs of either a cysteine-rich region/zinc-finger type or the chromo domain. Multiple functions for chromatin regulators are not restricted to the SET protein family, since many trx-G (but only very few Pc-G) genes are also modifiers of PEV. Together, these data establish a model in which the modulation of chromatin domains is mechanistically linked with the regulation of key developmental loci (e.g. HOM-C).

Key words. SET domain; Pc-G and trx-G genes; PEV in *Drosophila*; telomeric silencing; Su(var) and E(var) genes; chromatin domains.

Introduction

Gene expression in eukaryotes is modulated by positional information and higher-order chromatin. Permissive domains (euchromatin) are separated from more restricted, developmentally regulated regions (facultative heterochromatin) and from the transcriptionally inactive telomeres and centromeres (constitutive heterochromatin) [1] (reviewed in ref. 2). Whereas the majority of the genome in unicellular eukaryotes is transcriptionally competent, more than 60% of the mammalian genome is permanently silenced, and only ~3% of its DNA encodes structural genes (fig. 1). This dramatic increase of a repressive chromatin environment is mediated by diverse silencing mechanisms, including histone deacetylation (D. Allis; B. Turner; both this issue), DNA methylation and the establishment of restrictive chromatin domains by specialised 'chromatin regulators'. Such a multilayered repression system has been proposed to be required for the controlled definition of complex developmental programmes by preventing inappropriate gene expression [3]. In agreement with this notion, deregulation of chromatin regulators in *Saccharomyces cerevisiae* (*SIR4*) [4], *Drosophila* (Pc-G and trx-G) (reviewed in refs 5–7) and of several mammalian Pc-G and trx-G homologues (reviewed in refs 8, 9) results in perturbation of life-span, altered anteroposterior patterning (homeotic transformations) and the onset of leukemia (M. van Lohuizen, this issue).
Although basic chromatin components (histones) participate in chromatin-dependent gene activity (reviewed in ref. 10), genetic screens on telomeric silencing in *S. cerevisiae* [11] (L. Pillus, this issue), centromeric silencing in *Schizosaccharomyces pombe* [12, 13] and on posi-

* Corresponding author.

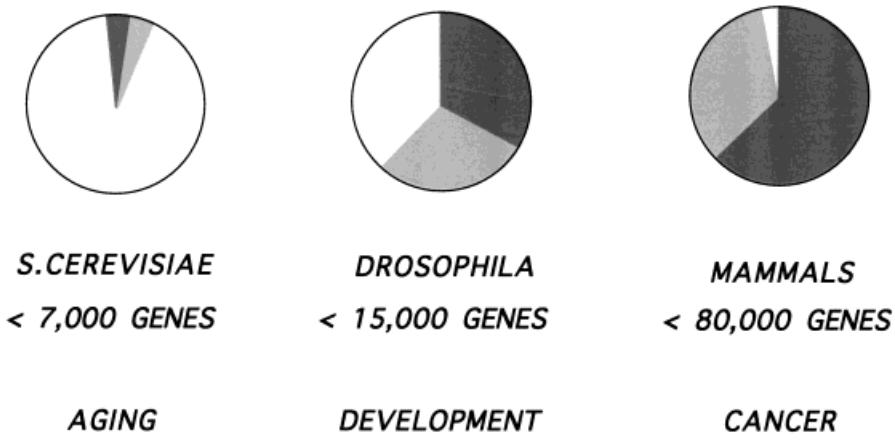

Figure 1. Increased gene silencing correlates with genome complexity. The ~10-fold increase in gene number from unicellular (*S. cerevisiae*) eukaryotes to higher vertebrates (mammals) is accompanied by an ~300-fold expansion of genome sizes, including mostly repetitive and noncoding (illustrated by grey shading) sequences. Whereas the majority of the *S. cerevisiae* genome is transcriptionally competent (indicated by the large white segment), more than 60% of the mammalian genome is permanently silenced (heterochromatic; shown in dark shading). Therefore, complex developmental programmes of multicellular organisms are particularly sensitive to perturbation of gene-silencing mechanisms.

tion-effect-variegation (PEV) in *Drosophila* (reviewed in refs 14, 15) have identified distinct classes of genes that are involved in the regional organisation of chromatin domains. For example, ~120 loci have been described that enhance [=E(var) genes] or suppress [=Su(var) genes] activity of heterochromatin-associated marker genes in *Drosophila*, thus implicating E(VAR) proteins in the establishment of active and SU(VAR) proteins in the organisation of repressive chromatin domains. Interestingly, several of the few characterised E(var) and Su(var) genes [14] share sequence motifs with either activators (trx-G) or repressors (Pc-G) of the homeotic selector genes (HOM-C cluster), suggesting a similar mechanism(s) for the regulation of gene activity at hetero- (PEV) and euchromatic positions (HOM-C). Whereas the chromo domain (J. Eissenberg, this issue) is found in the repressive SU(VAR)2-5 (HP1) [16], PC [17] and SU(VAR)3-9 [18] proteins, other protein modules are even exchanged between antagonistic chromatin regulators.

Such a 'promiscuous' sequence motif is the 130-amino acid SET domain which is present in the strongest PEV suppressor gene *Su(var)3-9* [18], in the Pc-G gene *Enhancer of zeste* (*E[z]*) [19] and in the activating trx-G gene *trithorax*. In addition to these founding *Drosophila* members of the SET domain gene family, the SET domain currently characterises >40 gene products ranging from yeast to man [20]. Other prominent members include the *S. cerevisiae SET1* gene which affects mating-type switching and telomeric silencing [21], the *S. pombe Clr4*$^{+}$ gene which is involved in centromere function [22], and the human *trx* homologue *HRX* [23–25] which is implicated in the onset of translocation-induced leukaemia.

In this review, we focus on the recent extension of the SET domain gene family by yeast and mammalian homologues, resulting in the demonstration of conserved mechanism(s) of gene silencing and implications in controlling chromatin structure and cellular proliferation/ differentiation. We will not discuss the molecular nature and function of *Drosophila* Pc-G and trx-G genes in governing HOM-C expression (recently reviewed in refs 5–7), but instead provide a comprehensive analysis of the potential of Pc-G and trx-G genes in regulating position-effect-variegation, which indicates that most trx-G but only very few Pc-G genes are modifiers of PEV. Together, these data establish a model in which the modulation of chromatin domains is mechanistically linked with the regulation of key developmental loci (e.g. HOM-C).

The SET protein family

The diverse implications and the modular structure of SET domain proteins obscures possible mechanisms of their mode of action. However, because the SET domain has been proposed to represent a novel interacton motif (see below), we used its primary amino acid sequence to align the currently characterised SET domain proteins. Ignoring other sequence similiarities, this alignment allows classification into four different subgroups (fig. 2).

Enhancer of zeste [E(Z)]: subgroup I

The E(z) subgroup I comprises the *Drosophila* Pc-G gene *E(z)* and its mammalian *EZH2* [26, 27], *Ezh1* [27] and plant *CLF* [28] homologues. *E(z)* restricts expres-

sion boundaries of HOM-C [29, 30] but has also been shown to be involved in repression of the early acting segmentation and gap genes [31–33]. Although *E(z)* may during distinct developmental stages and at specific promoter contexts display an activating function [34], the demonstration of gene silencing by *E(z)* homologues in the vicinity of *Drosophila* heterochromatin or at *S. cerevisiae* telomeres [27], and the repression of the floral homeotic gene *agamous* by *CLF* in *Arabidopsis thaliana* [28] clearly classifies *E(z)*-related gene products as repressive chromosomal proteins. Interestingly, DNA methylation seems to be required for the *CLF*-dependent repression, a functional link not yet described in mammals. In addition to its role in chromatin-dependent gene regulation, some *E(z)* mutants induce a general decondensation of polytene chromosomes [35] that is also reflected by increased chromosome breakage and cell cycle defects [36]. Involvement in the control of cell proliferation is further indicated by the puzzling interaction of EZH2 with the gene product of the *Vav* proto-oncogene [26] – an observation that may suggest a possible role in signal transduction. Moreover, a maternal-effect sterile mutant has recently been described as a *Caenorhabditis elegans* homologue of *E(z)* (R. Holdeman, S. Nehrt and S. Strome, unpublished; GenBank accession number AF011893). Mutations in the cysteine-rich cluster preceding the SET domain of E(Z) abolish polytene binding [37], and a temperature-sensitive (ts) *E(z)* null allele reduces immunostaining of several PC-G proteins [35] and of TRX [38]. Moreover, recruitment of PC-G protein complexes to ectopic heterochromatic positions has been shown to depend on *E*(*z*) [39, 40]. Although similar studies with ts-mutants of other Pc-G genes are still lacking, these observations indicate a central role for *E*(*z*)-related proteins in chromosome architecture and the nucleation of repressive protein complexes in eukaryotic chromatin.

Trithorax (TRX): subgroup II

Subgroup II connects the *Drosophila* trx-G gene *trx*, its human homologue *HRX* [23–25], an open reading frame in *C. elegans* and *S. cerevisiae SET1* [21]. trx-G genes antagonise the repressive function of Pc-G genes (for review, see refs 5–7), and several trx-G genes encode components – like BRAHMA [41] and ISWI [42] – of multimeric protein complexes that remodel inactive chromatin templates (reviewed in ref. 43). Consistent with its antagonistic function, TRX colocalises with many PC-G protein binding sites at polytene chromosomes [44] but maps to only a few sites also preferably stained by E(Z) [37]. Specific targets for TRX have been identified [38, 45], in agreement with the presence of a putative cysteine-rich DNA binding motif [20, 46], suggesting that TRX could recruit activating protein complexes. Additional cysteine-rich regions include several PHD-fingers (reviewed in ref. 47), mutations of which are embryonic lethal [20]. Searches for the mixed lineage leukaemia (*MLL*) or acute lymphoblastic leukaemia (*ALL*) gene identified *ALL-1*/*HRX* as the human *trx* homologue which is frequently disrupted by 11q23 translocations in acute leukaemias [23–25]. The breakpoints are clustered in the central region of *HRX*, and the pathological transcripts combine the amino-terminal half of HRX, containing a repression domain [48] and several A/T hooks of the HMGI/Y type, with diverse ALL-1 fusion (AF) partners. These observations suggest that HRX may also be involved in gene repression. However, since unfused HRX amino-terminal transcripts do not induce perturbed proliferation [49], the aberrant AF partners appear to be the causative transforming principle.

At least five SET domain genes have been identified through the genome sequencing project in *S. cerevisiae*. Among these, *SET1* is most closely related to *trx*, and disruption of *SET1* results in loss of silencing at telomeric positions and at the mating-type loci [21]. *set1Δ* mutant cells display morphological abnormalities and growth and sporulation defects, indicating regulation of many targets that may also include transcriptional activation. Indeed, several genes show reduced expression in *set1Δ* mutants [21], and a 12-bp 'SET1-consensus' sequence has been derived by competitive band-shift experiments (C. Nislow, L. Pillus, personal communication). Together, the observations suggest pleiotropic functions for the trx-subgroup proteins in both gene repression and activation, and the definition of specific target sequences further indicates a possible direct interaction with DNA.

Absent small or homeotic 1 (ASH1): subgroup III

Subgroup III contains the trx-G gene *absent small or homeotic (ash1)* [50] and *S. cerevisiae SET2*. The encoded gene products differ from other SET domain proteins because in ASH1 and SET2, the SET domain is not localised at the very carboxy-terminus. In addition to the SET domain, ASH1 contains a PHD finger (like TRX), and the modular structure of SET2 resembles that of E(Z), since a cysteine-rich region immediately precedes the SET domain. ASH1 colocalises with TRX at many polytene-binding sites [50], and a ts-*ash1* mutant reduces immunostaining of TRX [38]. Polytene chromosomes in *ash1* mutants appear more fragile [50], a phenotype reminiscent of *E(z)* mu-

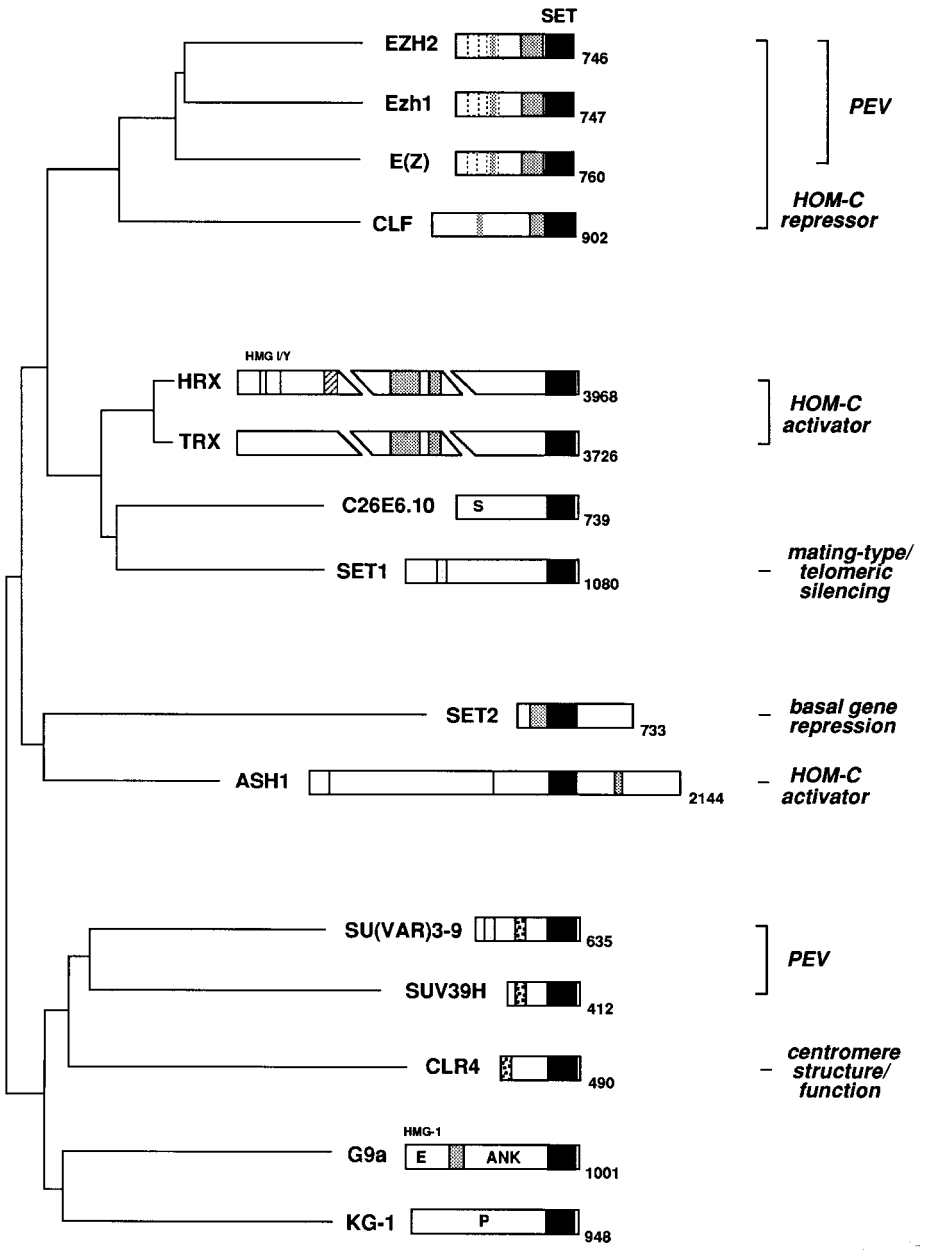

Figure 2. The SET protein family. SET domain proteins currently comprise >40 members, ranging from fungi to mammals. Ignoring other sequence similarities, only the SET domains of functionally characterised family members were compared using the PILEUP and PLOT programs of the GCG software package, resulting in the distinction of four subgroups as indicated. An apparent intron sequence (position indicated by arrowhead in fig. 3) in the human KG-1 (Nomura et al., unpublished; GenBank accession number D31891) SET domain has been removed prior to the alignment. The 130-amino acid SET domain is shown as a black box. Hatched bars are cysteine-rich regions in the E(Z) family and in *S. cerevisiae* SET2 (GenBank accession number Z49444), or encode a putative DNA binding motif and PHD fingers in *Drosophila* TRX, human HRX or a PHD finger in *Drosophila* ASH-1. Mammalian TRX homologues also contain a cysteine-rich region (cross-hatched bar) shared with DNA methyl transferases [106]. The chromo domain in *Drosophila* SU(VAR)3-9 and in its human SUV39H (G. Laible et al., unpublished) or *S. pombe* CLR4 (P. Lord and R. Allshire, personal communication) homologue 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; 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.

tants [35]. A truncated *ash1* mutant that deletes sequences carboxy-terminal of the middle of the PHD finger almost abolishes polytene binding [50], again suggesting that zinc-finger-type structures in SET domain proteins are important motifs for chromatin association. A mutation in the cysteine-rich cluster of

SET2 relieves basal repression of the GAL4 promoter, but *set2Δ* mutants display no apparent defect in telomeric or mating-type silencing (T. Hesman and M. Johnston, personal communication). Thus, in contrast to *SET1*, *SET2* appears to be specialized for gene repression at euchromatic positions where it may antagonise the activating function of nucleosome remodelling machines [51].

Suppressor of variegation SU(VAR)3-9: subgroup IV

Subgroup IV comprises the *Drosophila* PEV suppressor gene *Su(var)3-9* [18], its human (*SUV39H*) and *S. pombe* (*Clr4*$^+$) homologues, the human, ankyrin repeat-containing protein G9a [52] and an uncharacterised open reading frame, KG-1. In particular, the *Su(var)3-9* related proteins deserve special attention because they combine the two most characteristic motifs of chromatin regulators: the chromo [53, 54] and SET domains. *Su(var)3-9* is the strongest Su(var) gene described to date [18] and displays dose-dependent (triplo-enhancer/haplo-suppressor) effects in modulating expression of heterochromatin-associated marker genes, thus categorizing *Su(var)3-9* as a key regulator in the organisation of repressive chromatin domains. Consistent with such a function, recent immunodetection of SU(VAR)3-9 (S. Kuhfittig et al., unpublished) or of its mammalian homologue (L. Aagaard et al., unpublished) indicates preferred association with condensed regions in both hetero- and euchromatin. The *S. pombe Clr4*$^+$ (for cryptic locus regulator) gene has been isolated in a suppressor screen for mating-type silencing in fission yeast [12]. Subsequently it was shown that *Clr4*$^+$ is also a modifier of centromeric position effects [22], a phenomenon closely resembling telomeric silencing in *S. cerevisiae*. Mutations in *Clr4*$^+$ impair centromere function by disrupting interaction with the integral centromere component SWI6 [22, 55] – another chromo domain containing protein. The molecular nature of CLR4 (P. Lord and R. Allshire, personal communication) most closely resembles that of human SUV39H, suggesting that the GTP-binding motif in the amino-terminal third of the fly protein (see fig. 2) may provide additional functions to *Su(var)3-9*-related gene products.

The presence of the SET domain in both repressing and activating chromatin regulators, and the dual role of e.g *E(z)*, *HRX* and *SET1* in the control of gene regulation, are reminiscent of the bifunctional activities displayed by the *S. cerevisiae* repressor/activator protein RAP1 [56]. Similarly, although the chromo domain was initially found in repressing proteins, like HP1 and PC, it is also present in activating gene products that have putative ATPase (CHD1) [57, 58] or histone acetyltransferase (MOF) [59] activities. Chromo domain-swap experiments [39] and the resolution of its three-dimensional (3D) structure [60] suggest that the chromo domain is a protein-specific interaction motif which may also function as an adaptor in the recruitment of specialised transcription factors [61]. Finally, HP1 null mutants display an apparent defect in chromosome segregation [62]. Thus, in analogy to these examples and based on its evolutionary conservation, we propose that the SET domain represents an ancient protein module which is intrinsically connected to different networks governing chromosome architecture and transcriptional regulation. Its combination with a cysteine-rich region [E(Z) subgroup], DNA binding motifs and PHD fingers (TRX and ASH1 subgroups) or with the chromo domain [SU(VAR)3-9 subgroup] may then impart distinct target specificities upon the various SET domain proteins.

The SET domain

Despite these observations, the molecular function of the SET domain has remained elusive. Because of the antagonistic relationship between *E(z)* and *trx* [19, 38, 63], it has been suggested that the SET domain represents a surface for the assembly of either activating or repressing chromatin complexes, dependent on interactions with accessory TRX-G or PC-G proteins. Although protein interactions have been shown for the *Drosophila* PC [64] and the corresponding mammalian M33/BMI-1 complex [65], no interactions have been reported for SET domain proteins. The best demonstration of a direct function of the SET domain is the complementation of a *set1Δ*-induced silencing defect by the isolated SET domain of *SET1* [21]. This rescue is neutralised by a point mutation that alters a conserved glycine residue ($G_{955} \rightarrow S$), initially shown to be embryonic lethal (*trx*^{Z11}) in the corresponding position of the TRX SET domain [20] (see fig. 3). Other SET domain lesions include the hypomorphic ($N_{1385} \rightarrow I$) mutation in ASH1 [50], several uncharacterised mutations in *E(z)* (L. Carrington and R. Jones, personal communication) and three *Su(var)3-9* alleles (see fig. 3) that suppress PEV (V. Krauss and G.R., unpublished).

A DNA or chromatin binding activity for the SET domain can most likely be excluded, since point mutations in the cysteine-rich cluster in E(Z) [37] or a truncation of the PHD finger in ASH1 [50] disrupt polytene association, despite the presence of unaltered

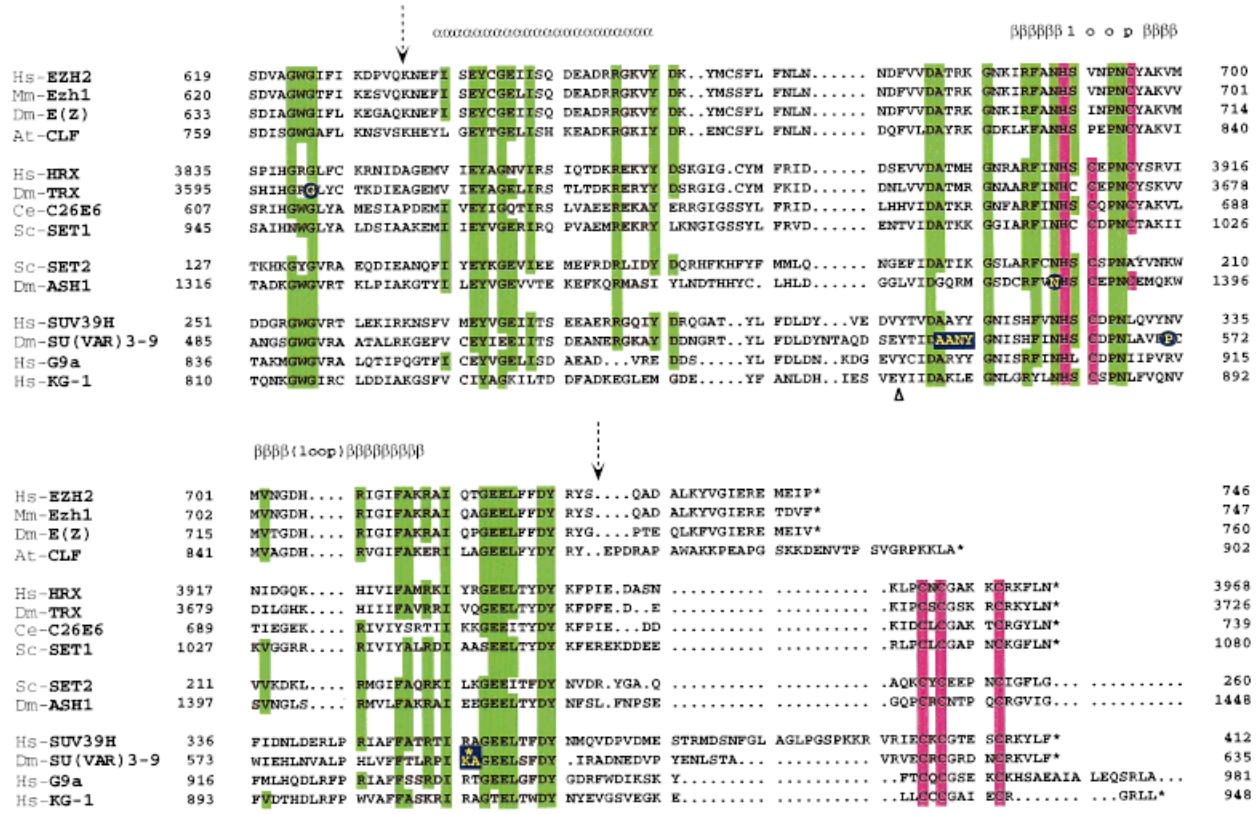

Figure 3. Conserved subregions of the SET domain. Amino acid sequences of the SET domains of the 15 representative SET protein family members shown in figure 2 were aligned using the PILEUP program of the GCG software package. An apparent intron sequence (position indicated by arrowhead) in the human KG-1 SET domain has been removed prior to the alignment. Amino acids that are present in all four subgroups are indicated in green, and conserved cysteine/histidine residues are highlighted by pink colour. Secondary structure algorithms [66] predict an α-helical stretch (region 1) and a strand-loop-strand domain (region 2), both of which coincide with regions of high sequence identity. Based on these overlapping conservations, regions 1 and 2 define a 'core SET domain' (demarcated by arrows). Additional conserved stretches outside this core are the amino-terminal GWG motif and a carboxy-terminal tail [absent in the E(Z) subgroup] which contains three invariably spaced cysteine residues. Also shown are positions of SET domain point mutations that are embryonic lethal in TRX ($G_{3601} \rightarrow S$; *trx*^{Z11}) [20] or hypomorphic in ASH1 ($N_{1385} \rightarrow I$; *ash*^{10}) [50]. The three indicated SU(VAR)3-9 SET domain mutations include an in-frame deletion ($\Delta_{549}AANY_{552}$; *Su(var)3-9*^{15}), a point mutation ($P_{571} \rightarrow S$; *Su(var)3-9*^{13}) and a P element-induced truncation ($_{592}KAGEE \rightarrow HDEIT*$; *Su(var)3-9*^{14}), all of which were identified as PEV suppressor alleles (V. Krauss and G.R., unpublished).

SET domains in the respective proteins. Further, ectopic expression of the isolated EZH2 or SUV39H SET domains (directed to the nucleus by a nuclear localization signal) do not appear to bind to mammalian interor metaphase chromatin (L. Aagaard and T.J., unpublished). The last result also suggests that the SET domain may only be functional in combination with other protein modules.

Alignment of the SET domains of the 15 representative members shown in figure 2 reveals two regions of high sequence identity (fig. 3), both of which coincide with ordered structures predicted by secondary structure algorithms [66]. Region 1 comprises a putative α-helical stretch of $\sim$20 amino acids that is invariant in all members analysed. Region 2 defines $\sim$50 amino acids with an inner domain of a predicted strand-loop-strand structure and two to three conserved histidine/cysteine residues. Since the spacing of the histidine/cysteine residues varies among the four subgroups, region 2 appears to be a good candidate to impart distinct specificities upon the respective SET domains. This interpretation is supported by the clustering of most of the currently known SET domain mutations in this region (see fig. 3). Additional conserved stretches outside this 'core SET domain' (as defined by regions 1 and 2) are the amino-terminal GWG motif, mutations in which have been proposed to disrupt a predicted turn [20], and a carboxy-terminal tail [absent in the E(Z) subgroup] which contains three invariably spaced cysteine residues. Two-dimensional nuclear magnetic resonance (2D NMR) analyses, similar to the studies that led to the definition of the chromo domain as an archetypal

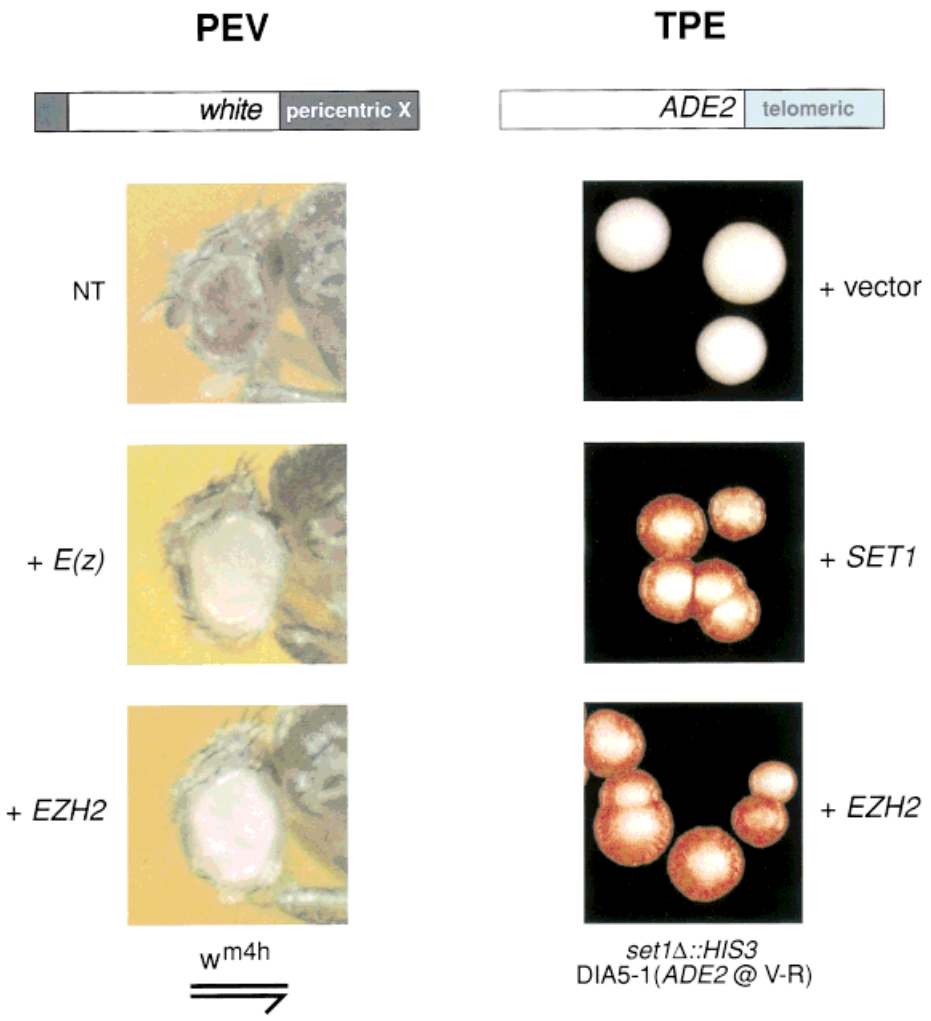

Figure 4. *E(z)*-related genes stabilise repressive domains in *Drosophila* and *S. cerevisiae* chromatin (left panel). Extra gene copies of *Drosophila E(z)* or of human *EZH2* enhance PEV in *Drosophila*. Transgene-dependent repression of pericentromeric w^{m4h} is detected by increased proportions of unpigmented areas in the eyes. NT, nontransgenic offspring. (right panel) Human *EZH2* restores telomeric silencing in *set1Δ* mutant cells. Repression of telomeric *ADE2* by plasmid-borne *EZH2* is reflected by the appearance of red-sectored colonies. See text and ref. [27] for details.

protein-interaction motif [60], are required to help elucidating the molecular function (protein interaction?) of the SET domain.

Conserved gene silencing in eukaryotic chromatin

The apparent central role of E(Z) in nucleating repressive protein complexes and the high evolutionary conservation between fly and mammalian *E(z)* homologues [27] suggest crucial function(s) in restricting chromatin-dependent gene activities in eukaryotes. Evidence for an involvement in establishing repressive chromatin domains has now been provided by analysing the potential of *E(z)* genes to modify position-dependent gene expression in *Drosophila* and *S. cerevisae*. Several indicator strains have been described in *Drosophila* that allow the analysis of chromosomal position effects. In one of the more widely used PEV strains, an inversion (w^{m4h}) places the *white* gene adjacent to pericentric X heterochromatin, resulting in variegated patterns of gene expression that can be easily detected as red (active transcriptional state) or white (repressed transcriptional state) patches in the *Drosophila* eye (fig. 4, left panel). Despite cytogenetic differences between chromatin of yeast and multicellular eukaryotes [67], telomeric silencing or telomeric-position-effect (TPE) closely resembles PEV in *Drosophila*. For example, insertion of marker genes (e.g. *ADE2*) at subtelomeric regions results in stochastic silencing of *ADE2* in many, but not all, cells within a given population [11]. If *ADE2* is repressed, cells accumulate a chromogenic

intermediate providing a colour assay to detect silenced (red-sectored) or active (white) transcriptional states. Whereas *ADE2* is repressed in the wild-type *SET1* strain, disruption of *SET1* results in loss of telomeric silencing [21] and the appearance of predominantly white colonies (fig. 4, right panel).

Using both of these model systems as in vivo assays for the modulation of chromatin domains, extra gene copies of *Drosophila E(z)* or of human *EZH2* in transgenic flies increase the proportion of unpigmented areas in the eyes, therefore indicating repression of the heterochromatin-associated *white* marker gene and classifying *E(z)* and *EZH2* as dose-dependent enhancers of PEV (fig. 4). Further, overexpression of *EZH2* in *set1Δ* mutant cells results in the appearance of red-sectored colonies, demonstrating restoration of telomeric silencing in *S. cerevisae* by a mammalian protein. Together, these data provide a functional link between the stable maintenance of gene repression and the organisation of inactive chromatin domains [27]. Moreover, since EZH2 and SET1 share 41% amino acid identity within the SET domain but otherwise do not display significant sequence similarities, these results also implicate an important role for the SET domain in participating in an evolutionarily conserved mechanism(s) that regulates transcriptional states in eukaryotic chromatin. However, consistent with the modular nature of SET domain proteins outlined above and in agreement with the more distant relationship of the SET domain in *Su(var)3-9*-related gene products with respect to EZH2 and SET1 (see fig. 2), human *SUV39H* did not rescue a *set1Δ*-dependent telomeric silencing defect (A. Wolf and T.J., unpublished).

Similar to the *EZH2*-mediated complementation of telomeric silencing, *Drosophila ORC2* has been shown to rescue an *orc2Δ*-dependent mating-type silencing defect in *S. cerevisiae* [68]. Loss-of-function mutations in the *Drosophila* and *S. cerevisiae* homologues of the human histone deacetylase (*HD-1*) gene result – in contrast to the expected increase in gene activation (D. Allis; B. Turner; both this issue) – in enhanced gene silencing [69]. These intriguing parallels of conserved mechanisms in directing regional gene activities in eukaryotic chromatin are further extended by the recent discoveries of mammalian *SIR2* [70] and SAS [71] (L. Pillus, this issue) homologues and by the surprising potential of a murine H19-imprinting element to function as a silencer in *Drosophila* [72].

Chromatin domains: many trx-G genes are E(var)s but Pc-G genes differ from Su(var)s

Several TRX-G proteins, like BRAHMA, ISWI, SNR1 and GAGA are components or cofactors of 'remodelling machines' that increase nucleosome mobility in vitro (reviewed in ref. 43). Although an antagonistic compaction of the chromatin structure has been postulated as a possible mechamism for PC-G and SU(VAR)-dependent gene silencing (discussed in ref. 73), direct interactions of 'locking complexes' with histones/nucleosomes (with the exception of the *S. cerevisiae* silencing factors SIR3 and SIR4 [74]) have not yet been reported. In fact, the complex functions of PC-G and SU(VAR) proteins in chromosome architecture and transcriptional regulation most likely require higher-order chromatin – a template almost inaccessible with in vitro assays.

Earlier studies on position-effect-variegation in *Drosophila* indicated that several Pc-G and trx-G genes are also modifiers of PEV [75]. However, while these observations have frequently been cited to imply a function for Pc-G and trx-G genes in organising repressive/active chromatin domains, most of the data have been conflicting, since the modifier effect on PEV did not co-map with the homeotic phenotype (see below). In addition, the dependence of PEV modification upon Pc-G/Su(var) genes differs between telomeric and pericentric PEV marker strains [76]. Gene expression of the pericentromeric *white-roughest*, *yellow* or *scute* PEV alleles appears to variegate according to the spreading of heterochromatin along the chromosome (''*cis*-silencing''), similar to the extension of inactive telomeric domains in *S. cerevisiae* [77, 78]. In contrast, ectopic blocks of heterochromatin, transposed to euchromatic positions (like in the $brown^{Dominant}$ [bw^D] mutation) modulate gene expression by ''*trans*-inactivating'' the paired wild-type copy across homologous chromosomes, resulting in a change of the nuclear localisation/ compartmentalisation of the heterochromatinised region [79, 80]. Importantly, bw^D variegation and pericentric PEV are both dependent on Su(var) gene function, indicating that repressive chromatin domains in these two systems may be established by similar mechanisms which have been explained by the ''pairing-looping model'' [81]. Despite these functional overlaps in repressing gene activity, significant differences are found in the dependance of bw^D variegation upon E(var) mutations. Variegation of *brown* appears to be influenced by the strength of repression, and usually hypomorphic E(var) mutations are less efficient in enhancing PEV of *brown* (G. R., unpublished). In particular, bw^D, which results in an almost complete inactivation of *brown*, is insensitive to E(var) gene function (G. Sass and S. Henikoff, personal communication). By contrast, pericentromeric PEV alleles (including *brown*) respond to trx-G/E(var) mutations (see below), suggesting that marker genes in the vicinity of centromeric heterochromatin appear more accessible to transcriptional activation.

Using position-effect-variegation of the pericentromeric w^{m4h} marker gene, heterozygous loss-of-function mutations (null alleles as judged by the homeotic phenotype)

Table 1. Drosophila Pc-G and trx-G genes that have been tested for PEV modification.

Gene	Protein (# amino acids)	Mammalian homologues	Protein domains	Modifier of PEV (w^{m4h}) extra gene copies (E = enhancer of PEV)	loss of function (Su = suppressor of PEV)	References
Polycomb (Pc)	390	*M33*, *MPc2*	chromo	nt	-	
Polycomblike (Pcl)	857	human EST	PHD fingers	nt	-	G.R., unpublished
polyhomeotic (ph)	1589	*Mph1*, *Rae28*	zinc finger, SAM	nt	-	Sinclair et al., in press [82]
Posterior sex combs (Psc)	1603	*Bmi1*, *mel18*	RING finger	nt	-	
Suppressor-2 of zeste [Su(z)2]	1365			nt	(Su-tel)	Wallrath et al., 1995 [76]
Sex comb on midleg (Scm)	877	human EST	zinc finger, SAM	nt	(Su-tel)	
Additional sex combs (Asx)	1668	human EST	cys-cluster	nt	E	Sinclair et al., in press [82]
extra sex combs (esc)	425	*eed*	WD-40 repeats	nt	-	
Enhancer of zeste [E(z)]	760	*Ezh1*, *Ezh2*	cys-rich, SET	E	Su (weak)	Laible et al., 1997 [27] M. Gatti (pers. com.)
Enhancer of Polycomb [E(Pc)]	2023	human EST		nt	Su	G.R., unpublished; Sinclair et al., in press [82]
trithorax (trx)	3726	*MII*, *HRX*	PHD fingers, DNA-binding motif, SET	nt	-*	G.R. and R.D., unpublished
brahma (brm) [SWI2-related]	1638	*mbrm*, *BRG1*	bromo, ATPase/helicase motifs	nt	E	
female sterile homeotic (fs[1]h)	1118	human RING3	bromo	nt	nt	
trithorax-like (Trl) [GAGA]	519		zinc finger, BTB	nt	E	Farkas et al., 1994 [90]
E(var)3-93D [mod(mdg4)]	401		BTB	nt	E	Dom et al., 1993 [89]
absent small or homeotic 1 (ash1)	2144		PHD finger, SET	nt	-	G.R., unpublished
ISWI	1027	hSNF2L	ATPase motifs	nt	nt	
snf5-related 1 (snr1)	370	human EST	acidic region	nt	Su (weak)	G.R., unpublished
zeste (z)	575		HTH-DNA binding motif, heptad repeats	nt	E (recessive)	Judd, 1995 [107]

The uncharacterised Pc-G genes *Sex comb extra*, *super sex combs* and *pleiohomeotic* do not modify PEV [82]. The uncharacterised trx-G genes *devenir*, *kismet*, *verthandi*, *moira* and *sallimus* are all enhancers of PEV, whereas *ash2*, *osa* and *skuld* show no effect (G.R., unpublished). The data cited in ref. 75 have been critically re-examined (G.R., unpublished) [82]. With the exception of *Asx*, the modifier effect on PEV did not co-map with the homeotic phenotype, indicating that second-site factors were accumulated. Whereas the enhancer effect of *Asx* is consistent with anterior transformations displayed by several *Asx* alleles [83], the weak suppressor effect of *snr1* is exceptional for trx-G genes. Very few Pc-G genes [*E(z)* and *E(Pc)*] are also suppressors of PEV, and *E(z)* is the only Pc-G gene for which a dose-dependent (triplo-enhancer/haplo-suppressor) modifier effect on PEV has been shown. nt, not tested; – , no effect on PEV; (Su-tel), no effect on pericentric PEV, but suppresses telomeric position effects. * Enhances *E(var)3-93D* mutations [89]. Part of this table describing the molecular nature of PC-G and TRX-G proteins has been adopted from previous reviews [5, 7]. Additional references for recently isolated genes are: *MPc2*, M. van Lohuizen, cited in ref. 9; *Mph1* [65]; *Scm* [108]; *Asx* (D. A. Sinclair, T. Milne, J. Shellard, J. Hodgson and H. W. Brock, unpublished); *eed* [109]; *EZH2* [26]; *Ezh1* and *EZH2* [27]; *E(Pc)* (K. Stankunas, Randazzo F. and Brock H. W., unpublished). The originally identified 65-amino acid SAM domain [110] has also been named SPM or SEP domain.

of all currently characterised Pc-G and trx-G genes have been carefully re-examined for their effects on PEV (G.R., unpublished) [82]. Of 13 Pc-G genes analysed, only two [*E(z)* and *E(Pc)*] are suppressors of PEV (table 1), indicating a specialised function of the encoded proteins in stabilising a repressive chromatin domain. Moreover, *E(z)* is the only Pc-G gene for which a dose-dependent (triplo-enhancer/haplo-suppressor) modifier effect has been shown [27], in agreement with a spreading or pairing-looping model for *cis*-silencing. *Su(z)2* and *Scm* do not modify pericentric PEV, but suppress less stringent telomeric position effects [76]. Finally, although the E(var) effect of *Asx* is unexpected, several *Asx* alleles have been shown to induce anterior rather than posterior transformations [83]. Together, these data indicate that most Pc-G genes do not modify PEV, consistent with the absence of PC-G proteins from constitutive heterochromatin (reviewed in refs 5–7). In a converse analysis, only one out of the six currently characterised Su(var) genes [16, 18, 84–87] affects homeotic transformations; overall, $<1\%$ of Su(var) genes appear to function in HOM-C control (G.R. and R.D., unpublished).

In contrast to the Pc-G genes, the majority (9/15) of trx-G genes are enhancers of PEV (table 1), and four of the seven E(var) genes analysed to date [69, 88–92] display homeotic transformations. Genetic data indicate that $>6\%$ of E(var) genes can be categorised as trx-G genes (G.R. and R.D., unpublished). Indeed, GAGA factor is encoded by *E(var)3-70F* (identical with *Trl* [90]), and the E(VAR)3-93D protein shares a structural motif (BTB domain [89]) with several transcription factors.

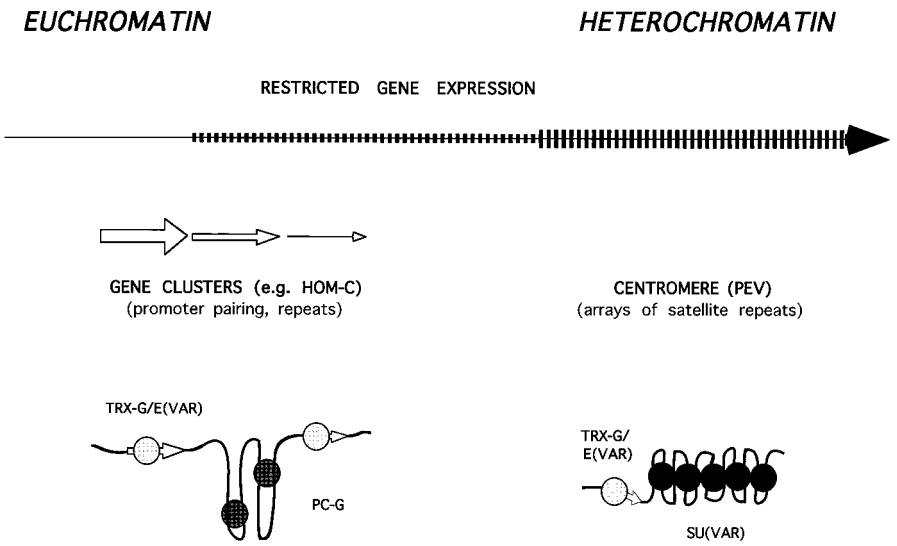

Figure 5. Repeat-induced model for the organisation of repressive chromatin domains by PC-G and SU(VAR) proteins. The basic concepts for the establishment of a restrictive chromatin environment have been adopted from the pairing-looping model [81]. Repressive chromatin domains are modulated by density of repeat sequences, regulated promoter pairing in gene clusters, transcriptional activity and selective distribution of PC-G and SU(VAR) proteins (see text for detailed discussion). SU(VAR) protein complexes (dark circles) are structural components of constitutive heterochromatin (e.g. at the transcriptionally silent centromere) with a presumed high affinity for repetitive arrays (e.g. satellite repeats). By contrast, most PC-G protein complexes (grey circles) are excluded from constitutive heterochromatin and facilitate co-operative interactions and promoter pairing of linked loci at euchromatic positions, thereby inducing a more condensed (looped) chromatin structure. Activating TRX-G/E(VAR) protein complexes (light-shaded circles) and transcriptional 'on states' (arrows) antagonise Pc-G dependent modulation of facultative heterochromatin (at e.g. the HOM-C cluster) and counteract propagation of SU(VAR)-specific silencing complexes in the vicinity of constitutive heterochromatin (PEV). Although these operational simplifications illustrate the main functions, more dynamic transitions in the organization of eu- and heterochromatic regions are suggested by sequence and functional similarities shared between a few Pc-G and Su(var) genes (see text).

E(var)3-93D is allelic to a modifier gene [*mod(mdg4)*] whose loss-of-function converts an insulator into a bidirectional silencer [93]. The convergence between trx-G and E(var) gene function most probably reflects association of the encoded proteins with either nucleosome remodelling (TRX-G) or activating transcriptional complexes [E(VAR)], both of which would counteract the spreading of repressive chromatin domains. In addition to the documented competition between transcriptional activation and gene silencing [94], some TRX-G/E(VAR) proteins may even be 'passenger components' of constitutive heterochromatin – as shown for the GAGA factor which associates with telomeric $(AAGAG)_n$ repeats.

Selective and promiscuous functions of chromatin regulators in eu- and heterochromatin

Based on the comprehensive PEV analysis discussed above, a dynamic model for the organisation of chromatin domains involves selective as well as promiscuous functions of chromatin regulators (fig. 5). Activating TRX-G/ E(VAR) proteins function at both euchromatic positions and in the vicinity of heterochromatin, probably by sensing promoter activity and by associating with transcription complexes. Their involvement in modulating chromatin domains would be a consequence of increased transcriptional activity, which antagonises the establishment/propagation of a repressive chromatin structure.

PC-G proteins are largely excluded from constitutive heterochromatin and resemble sequence non-specific transcriptional repressors [95] with functions in restricting expression patterns of 'complex loci' (e.g the homeotic gene cluster). Gene clusters could potentially promote an altered chromatin structure (e.g. by promoter pairing), reminiscent of transgene repeats that have been shown to induce variegated gene expression [96]. Moreover, co-suppression by presumed homologous gene pairing depends on *Polycomb* [97]. According to this view, and in agreement with shared sequence motifs (e.g. the chromo and SET domains), some PC-G gene products would be attenuated SU(VAR) proteins displaying affinity for paired or looped regions, thereby inducing a more condensed chromatin structure. Possible paired targets also comprise so-called Pc-G response elements (PREs) (reviewed in ref. 98), some of which may contain repeat sequences. For most PC-G proteins, however, restriction of Pc-G-mediated repression to more euchromatic positions is probably a consequence for the requirement of co-operative interactions which may even include

association with transcription factors and/or promoter complexes. Dependent on the ratio of activators (TRX-G) and repressors (PC-G) in these presumptive protein complexes, transcriptional 'on or off states' of the linked genes would then antagonise or favour the stabilisation of a repressed chromatin structure. Taken together, PC-G proteins appear to provide dual functions both in transcriptional control and chromatin structure that are consistent with the developmentally regulated establishment of facultative heterochromatin at complex gene loci (e.g. the HOM-C cluster) (discussed in ref. 2).

In contrast to several TRX-G/ E(VAR) and PC-G gene products, SU(VAR) proteins have not yet been carefully investigated (but see two-hybrid interactions for HP1 homologues reported in ref. 61) for a direct interference with transcriptional regulation. Distribution of the characterised chromosomal SU(VAR)2-5 (HP1) [99], SU(-VAR)3-6 (MODULO) [86], SU(VAR)3-7 [100] and SU(VAR)3-9 (S. Kuhfittig et al., unpublished) proteins indicates association with condensed chromatin regions. In particular, the chromo domain con taining HP1 and SU(VAR)3-9 proteins are enriched at the chromocenter, a localization paralleled by the SWI6 [55] and presumably also CLR4 (see above) homologues at *S. pombe* centromeres. Two-hybrid interactions (A. Fischer, M. O'Grady and G.R., unpublished) predict heteromeric complexes between HP1 and SU(VAR)3-9, and the human SU(VAR)3-9 homologue homodimerises in vitro (G.L. and T.J., unpublished). A direct interaction between SU(VAR)3-7 and HP1 has recently been reported [100]. It is not known whether-any of these interactions are mutually exclusive. Possibly, arrays of heterochromatin-specific satellite repeats may provide preferred binding sites for SU(VAR) proteins which – in the absence of transcriptional activity at regions like the chromocenter – establish stable complexes that propagate repressive higher order chromatin. This speculative interpretation is supported by analogy to the methyl-CpG binding protein MeCP2, whose function/localization requires a high density of methylated CpGs [101]. From the available data, chromosomal SU(VAR) proteins can be defined as high-affinity structural components (rather than transcriptional regulators) of constitutive heterochromatin, and their loss of function would directly compromise the establishment/propagation of a repressive chromatin domain.

Although these operational simplifications are useful to illustrate the main functions of the different classes of chromatin regulators, more dynamic transitions in the organisation of eu- and heterochromatic regions are required to control programmed development [102] (B. Turner, this issue). Dose-dependent (i.e. triplo-enhancer/haplo-suppressor) chromatin regulators, like *HP1* [103], *E(z)* [27], *Su(var)3-7* [85] and *Su(var)3-9* [18] (also inferred by its intriguing combination of the chromo and SET domains) are good candidates for such developmental switches, which probably link networks controlling chromatin structure and transcriptional regulation. Moreover, multiple regulatory levels are further indicated by the existence of an antagonising class of dose-dependent (i.e. triplo-suppressor/haplo-enhancer) PEV genes [104], most of which have not yet been characterised at the molecular level. In an extension to the repeat/pairing-looping dependent model [81] outlined above, such multifunctional chromatin regulators may even participate in X inactivation or imprinting (where the special topology of an RNA moiety could provide paired regions) – a hypothesis further stimulated by the involvement of the chromo domain-like containing MOF protein in regulating dosage compensation in *Drosophila* [59]. Finally, the approximately tenfold increase in gene number from unicellular eukaryotes to higher vertebrates (see fig. 1) is accompanied by an ~300-fold expansion of genome sizes, including mostly repetitive and noncoding sequences. According to recent observations which implicate silencing mechanisms in a genome-wide defence against transposon-induced expansions of repeat sequences (discussed in ref. 105) (M. Matzke, this issue), specialised members of the SET protein family may also be part of a multilayered repressive system that is required to maintain the integrity of the genome.

Acknowledgements. We are particularly indebted to Hugh Brock for sharing data on PEV modification by Pc-G genes prior to publication, to Steven Henikoff for helpful comments on PEV mechanisms, to Wendy Bickmore for advice on mammalian heterochromatin, and to Philipp Selenko for the secondary structure alignments of the SET domain. We would like to thank Robin Allshire, Mark Johnston, Lorraine Pillus, Rick Jones, Veiko Krauss, Andreas Fischer and Louise Aagaard for their agreement to cite work in progress, and Louise Aagard, Philipp Selenko and Dónal O'Carroll for critical reading of the manuscript. Research in T.J.'s laboratory was funded by the IMP and by the Austrian Industrial Research Promotion Fund. G.R. and R.D. were supported by grants from the Deutsche Forschungsgemeinschaft.

1 Heitz E. (1929) Heterochromatin, Chromocentren, Chromomeren. Ber. Dtsch. Bot. Ges. **47:** 274–284
2 Bickmore W. and Craig J. (1997) Chromosome Bands: Patterns in the Genome, pp. 133–162, Springer, Heidelberg
3 Bird A. (1995) Gene number, noise reduction and biological complexity. Trends Genet. **11:** 94–100
4 Kennedy B. K., Austriaco N. R. Jr., Zhang J. and Guarente L. (1995) Mutation in the silencing gene *SIR4* can delay aging in *S. cerevisiae*. Cell **80:** 485–496
5 Kennison J. A. (1995) The Polycomb and trithorax group proteins of *Drosophila*: trans-regulators of homeotic gene function. Annu. Rev. Genet. **29:** 289–303
6 Simon J. (1995) Locking in stable states of gene expression: transcriptional control during *Drosophila* development. Curr. Opin. Cell Biol. **7:** 376–385
7 Paro R. and Harte P. J. (1996) The role of Polycomb group and trithorax group complexes in the maintenance of determined cell states. In: Epigenetic Mechanisms of Gene Regulation, pp. 507–528, Russo V. E. A., Martienssen R. A. and Riggs A. D. (eds), CSHL, New York

8 Schumacher A. and Magnuson T. (1997) Murine *Polycomb*- and *trithorax*-group genes regulate homeotic pathways and beyond. Trends Genet. **13:** 167–170

9 Gould A. (1997) Functions of mammalian *Polycomb* group and *trithorax* group related genes. Curr. Opin. Genet. Dev. **7:** 488–494

10 Thompson J. S., Hecht A. and Grunstein M. (1993) Histones and the regulation of heterochromatin in yeast. In: DNA and Chromosomes, pp. 247–256, Symp. Quant. Biol. 58, CSHL, New York

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

12 Thon G., Cohen A. and Klar A. J. (1994) Three additional linkage groups that repress transcription and meiotic recombination in the mating-type region of *Schizosaccharomyces pombe*. Genetics **138:** 29–38

13 Allshire R. C., Nimmo E. R., Ekwall K., Javerzat J. P. and Cranston G. (1995) Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. Genes Dev. **9:** 218–233

14 Reuter G. and Spierer P. (1992) Position effect variegation and chromatin proteins. BioEssays **14:** 605–612

15 Weiler K. S. and Wakimoto B. T. (1995) Heterochromatin and gene expression in *Drosophila*. Ann. Rev. Genet. **29:** 577–605

16 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

17 Paro R. and Hogness D. S. (1991) The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila*. Proc. Natl. Acad. Sci. USA **88:** 263–267

18 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

19 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

20 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

21 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 **8:** 2421–2436

22 Ekwall K., Nimmo E. R., Javerzat J. P., Borgstrom B., Egel R., Cranston G. et al. (1996) Mutations in the fission yeast silencing factors clr4+ and rik+ disrupt the localization of the chromo domain protein Swi6p and impair centromere function. J. Cell. Sci. **109:** 2637–2648

23 Tkachuk D. C., Kohler S. and Cleary M. L. (1992) Involvement of a homolog of *Drosophila trithorax* by 11q23 chromosomal translocations in acute leukaemias. Cell **71:** 691–700

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

25 Djabali M., Selleri L., Parry P., Bower M., Young B. D. and Evans G. A. (1992) A *trithorax*-like gene is interrupted by 11q23 translocations in acute leukaemia. Nature Genet. **2:** 113–118

26 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

27 Laible G., Wolf A., Dorn R., Reuter G., Nislow C., Lebersorger A. et al. (1997) Mammalian homologues of the *Polycomb*-group gene *Enhancer of zeste* mediate gene silencing in *Drosophila* heterochromatin and at *S. cerevisiae* telomeres. EMBO J. **16:** 3219–3232

28 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

29 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

30 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

31 Moazed D. and O'Farrell P. H. (1992) Maintenance of the engrailed expression pattern by *Polycomb* group genes in *Drosophila*. Development **116:** 805–810

32 McKeon J., Slade E., Sinclair D. A., Cheng N., Couling M. and Brock H. W. (1994) Mutations in some *Polycomb* group genes of *Drosophila* interfere with regulation of segmentation genes. Mol. Gen. Genet. **244:** 474–483

33 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

34 LaJeunesse D. and Shearn A. (1996) *E(z)*: a *Polycomb* group gene or a *trithorax* group gene? Development **122:** 2189–2197

35 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

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

37 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

38 Kuzin B., Tillib S., Sedkov Y., Mizrokhi L. and Mazo A. (1994) The *Drosophila trithorax* gene encodes a chromosomal protein and directly regulates the region-specific homeotic gene *fork head*. Genes Dev. **8:** 2478–2490

39 Platero J. S., Harnett T. and Eissenberg J. C. (1995) Functional analysis of the chromo domain of HP-1. EMBO J. **14:** 3977–3986

40 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

41 Tamkun J. W., Deuring R., Scott M. P., Kissinger M., Pattatucci A. M., Kaufman T. C. et al. (1992) brahma: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/ SWI2. Cell **68:** 561–572

42 Tsukiyama T., Daniel C., Tamkun J. W. and Wu C. (1995) ISWI, a member of the SWI2/ SNF2 ATPase family, encodes the 140 kDa subunit of the nucleosome remodeling factor. Cell **83:** 1021–1026

43 Tsukiyama T. and Wu C. (1997) Chromatin remodeling and transcription. Curr. Opin. Genet. Dev. **7:** 182–191

44 Chinwalla V., Jane E. P. and Harte P. J. (1995) The *Drosophila* trithorax protein binds to specific chromosomal sites and is co-localized with Polycomb at many sites. EMBO J. **14:** 2056–2065

45 Chang Y. L., King B. O., O'Connor M., Mazo A. and Huang D. H. (1995) Functional reconstruction of trans-regulation of the Ultrabithorax promoter by the products of two antagonistic genes, *trithorax* and *Polycomb*. Mol. Cell. Biol. **15:** 6601–6612

46 Mazo A. M., Huang D. H., Mozer B. A. and David, I. B. (1990) The *trithorax* gene, a transacting regulator of the bithorax complex in *Drosophila*, encodes a protein with zinc-binding domains. Proc. Natl. Acad. Sci. USA **87:** 2112–2116

47 Aasland R., Gibson T. J. and Stewart A. F. (1995) The PHD finger: implications for chromatin-mediated transcriptional regulation. Trends Biochem. Sci. **20:** 56–59

48 Le Zeleznik N. J., Harden A. M. and Rowley J. D. (1994) 11q23 translocations split the 'AT-hook' cruciform DNA-

binding region and the transcriptional repression domain from the activation domain of the mixed-lineage leukaemia (*MLL*) gene. Proc. Natl. Acad. Sci. USA **91:** 10610–10614
49 Lavau C., Szilvassy S. J., Slany R. and Cleary M. L. (1997) Immortalization and leukemic transformation of a myelomonocytic precursor by retrovirally transduced HRX-ENL. EMBO J. **16:** 4226–4237
50 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
51 Cairns B. R., Lorch Y., Li Y., Zhang M., Lacomis L., Erdjument-Bromage H. et al. (1996) RSC, an essential, abundant chromatin-remodeling complex. Cell **87:** 1249–1260
52 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
53 Aasland R. and Stewart A. F. (1995) The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. Nucleic Acids Res. **23:** 3163–3173
54 Koonin E. V., Zhou S. and Lucchesi J. C. (1995) The chromo superfamily: new members, duplication of the chromo domain and possible role in delivering transcription regulators to chromatin. Nucleic Acids Res. **23:** 4229–4233
55 Ekwall K., Javerzat J. P., Lorentz A., Schmidt H., Cranston G. and Allshire R. (1995) The chromo domain protein Swi**6:** a key component at fission yest centromeres. Science **269:** 1429–1431
56 Shore D. (1994) RAP1: a protean regulator in yeast. Trends Genet. **10:** 408–412
57 Delmas V., Stokes D. G. and Perry R. P. (1993) A mammalian DNA-binding protein that contains a chromodomain and an SNF2/ SWI2-like helicase domain. Proc. Natl. Acad. Sci. USA **90:** 2414–2418
58 Stokes D. G., Tartof K. D. and Perry R. P. (1996) CHD1 is concentrated in interbands and puffed regions of *Drosophila* polytene chromosomes. Proc. Natl. Acad. Sci. USA **93:** 7137–7142
59 Hilfiker A., Hilfiker-Kleiner D., Pannuti A. and Lucchesi J. C. (1997) *mof*, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*. EMBO J. **16:** 2054–2060
60 Ball L. J., Murzina N. V., Broadhurst R. W., Raine A. R. C., Archer S. J., Stott F. J. et al. (1997) Structure of the chromatin binding (chromo) domain from mouse modifier protein 1. EMBO J. **16:** 2473–2481
61 Le Douarin B., Nielsen A. L., Garnier J.-M., Ichinose H., Jeanmougin F., Losson R. et al. (1996) A possible involvement of TIF1α and TIF1β in the epigenetic control of transcription by nuclear receptors. EMBO J. **15:** 6701–6715
62 Raff J. W., Kellum R. and Alberts B. (1994) The Drosophila GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle. EMBO J. **13:** 5977–5983
63 Campbell R. B., Sinclair D. A., Couling M. and Brock H. W. (1995) Genetic interactions and dosage effects of *Polycomb* group genes of *Drosophila*. Mol. Gen. Genet. **246:** 291–300
64 Franke A., DeCamillis M., Zink D., Cheng N., Brock H. W. and Paro R. (1992) Polycomb and polyhomeotic are constituents of a multimeric protein complex in chromatin of *Drosophila melanogaster*. EMBO J. **11:** 2941–2950
65 Alkema M. J., Bronk M., Verhoeven E., Otte A., van't Veer L. J., Berns A. et al. (1997) Identification of bmi1-interacting proteins as constituents of a multimeric mammalian Polycomb complex. Genes Dev. **11:** 226–240
66 Rost B. and Sander C. (1993) Prediction of protein structure at better than 70% accuracy. J. Mol. Biol. **232:** 584–599
67 Pillus L. and Grunstein M. (1995) Chromatin structure and epigenetic regulation in yeast. In: Chromatin Structure and Gene Expression, pp. 123–146, Elgin S. C. R. (ed.), IRL Press, New York
68 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
69 DeRubertis F., Kadosh D., Henchoz S., Pauli S., Reuter G., Struhl G. et al. (1996) The histone deacetylase RPD3 counteracts genomic silencing in *Drosophila* and yeast. Nature **384:** 589–591
70 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
71 Reifsnyder C., Lowell J., Clarke A. and Pillus L. (1996) Yeast *SAS* silencing genes and human genes associated with AML and HIV-1 Tat interactions are homologous with acetyltransferases. Nat. Genet. **14:** 42–49
72 Lyko F., Brenton J. D., Surani M. A. and Paro R. (1997) An imprinting element from the mouse *H19* locus funcitons as a silencer in *Drosophila*. Nature Genet. **16:** 171–173
73 Kingston R. E., Bunker C. A and Imbalzano A. N. (1996) Repression and activation by multiprotein complexes that alter chromatin structure. Genes Dev. **10:** 905–920
74 Hecht A., Laroche T., Strahl-Bolsinger S., Gasser S. M. and Grunstein M. (1995) HistoneH3 and H4 N-termini interact with *SIR3* and *SIR4* proteins: a molecular model for the formation of heterochromatin in yeast. Cell **80:** 583–592
75 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
76 Wallrath L. L. and Elgin S. C. R. (1995) Position eff ect variegation in *Drosophila* is associated with an altered chromatin structure. Genes Dev. **9:** 1263–1277
77 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
78 Hecht A., Strahl-Bolsinger S. and Grunstein M. (1996) Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. Nature **383:** 92–95
79 Dernburg A. F., Broman K. W., Fung J. C., Marshall W. F., Philips J., Agard D. A. et al. (1996) Pertubation of nuclear architecture by long-distance chromosome interactions. Cell **85:** 745–759
80 Csink A. K. and Henikoff S. (1996) Genetic modification of heterochromatic association and nuclear organisation in *Drosophila*. Nature **381:** 529–531
81 Henikoff S. (1996) Position-effect-variegation in *Drosophila*: recent progress. In: Epigenetic Mechanisms of Gene Regulation, pp. 319–334, Russo V. E. A., Martienssen R. A. and Riggs A. D. (eds), CSHL, New York
82 Sinclair D. A., Clegg N. J., Antonchuk J., Milne T. A., Stankunas K., Ruse C. et al., (1997) *Enhancer of Polycomb* is a suppressor of position-effect-variegation in *Drosophila melanogaster*. Genetics, in press
83 Sinclair D. A., Campbell R. B., Nicholls F., Slade E. and Brock H. W. (1992) Genetic analysis of the *additional sex combs* locus of *Drosophila melanogaster*. Genetics **130:** 817–825
84 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
85 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
86 Garzino V., Pereira A., Laurenti P., Graba Y., Levis R. W., LeParco Y. et al. (1992) Cell lineage-specific expression of *modulo*, a dose-dependent modifier of variegation in *Drosophila*. EMBO J. **11:** 4471–4479
87 Baksa K., Morawietz H., Dombradi V., Axton M., Taubert H., Szabo G. et al. (1993) Mutations in the protein phosphatase 1 gene at 87B can differentially affect suppression of position-effect-variegation and mitosis in *Drosophila melanogaster*. Genetics **135:** 117–125

88 Dorn R., Szidonya J., Korge G., Sehnert M., Taubert H., Archoukieh I. et al. (1993) P transposon-induced dominant enhancer mutations of position-effect-variegation in *Drosophila melanogaster*. Genetics **133:** 279–290
89 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
90 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
91 Seum C., Spierer A., Pauli D., Szidonya J., Reuter G. and Spierer P. (1996) Position-effect-variegation in *Drosophila* depends on dose of the gene encoding the E2F transcriptional activator and cell cycle regulator. Development **122:** 1949–1956
92 Henchoz S., DeRubertis F., Pauli D. and Spierer P. (1996) The dose of a putative ubiquitin-specific protease affects position-effect-variegation in *Drosophila melanogaster*. Mol. Cell. Biol. **16:** 5717–5725
93 Gerasimova T. I., Gdula D. A., Gerasimov D. V., Simonova O. and Corces V. G. (1995) A *Drosophila* protein that imparts directionality on a chromatin insulator in an enhancer of position-effect-variegation. Cell **82:** 587-597
94 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
95 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
96 Dorer D. R. and Henikoff S. (1994) Expansions of transgene repeats cause heterochromatin formation and gene silencing in *Drosophila*. Cell **77:** 993–1002
97 Pal-Bhadra M., Bhadra U. and Birchler J. A. (1997) Cosuppression in *Drosophila*: gene silencing of *Alcohol dehydrogenase* by white-Adh transgenes is Polycomb dependent. Cell **90:** 479–490
98 Pirrotta V. (1997) Pc-G complexes and chromatin silencing. Curr. Opin. Genet. Dev. **7:** 249–258
99 Powers J. A. and Eissenberg J. C. (1993) Overlapping domains of the heterochromatin-associated protein HP1 mediate nuclear localization and heterochromatin binding. J. Cell. Biol. **120:** 291–299
100 Cleard F., Delattre M. and Spierer P. (1997) SU(VAR)3-7, a *Drosophila* heterochromatin-associated protein and companion of HP1 in the genomic silencing of position-effect-variegation. EMBO J. **17:** 5280–5288
101 Nan X., Campoy F. J and Bird A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell **88:** 471–481
102 Lu B. Y., Bishop C. P. and Eissenberg J. C. (1996) Developmental timing and tissue- specificity of heterochromatin-mediated silencing. EMBO J. **15:** 1323–1332
103 Eissenberg J. C., Morris G. D., Reuter G. and Hartnett T. (1992) The heterochromatin-associated protein HP1 is an essential protein in *Drosophila* with dosage-dependent effects on position-effect variegation. Genetics **131:** 345–352
104 Wustmann G., Szidonya J., Taubert H. and Reuter G. (1989) The genetics of position-effect modifying loci in *Drosophila melanogaster*. Mol. Gen. Genet. **217:** 520–527
105 Henikoff S. and Matzke M. M. (1997) Exploring and explaining epigenetic effects. Trends Genet. **13:** 293–295
106 Ma Q., Alder H., Nelson K. K., Chatterjee D., Gu Y., Nakamura T. 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
107 Judd B. H. (1995) Mutations of zeste that mediate transvection are recessive enhancers of position-effect-variegation in *Drosophila melanogaster*. Genetics **141:** 245–253
108 Bornemann D., Miller E. and Simon J. (1996) The *Drosophila Polycomb* group gene *Sex comb on migleg* (*Scm*) encodes a zinc finger protein with similarity to polyhomeotic protein. Development **122:** 1621–1630
109 Schumacher A., Faust C. and Magnuson T. (1996) Positional cloning of a global regulator of anterior-posterior patterning in mice. Nature **383:** 250–253
110 Pontig C. P. (1995) SAM: a novel motif in yeast sterile and *Drosophila* polyhomeotic proteins. Protein Sci. **4:** 1928–1930