Glimpses of evolution: Heterochromatic histone H3K9 methyltransferases left its marks behind

Veiko Krauss

Department of Genetics, Institute for Biology II, University of Leipzig, Johannisallee 21-23, 04103 Leipzig, Germany (Phone: +49-341 9736751; Fax: +49-341-9736897; E-mail: krauss@rz.uni-leipzig.de)

Abstract

In eukaryotes, histone methylation is an epigenetic mechanism associated with a variety of functions related to gene regulation or genomic stability. Recently analyzed H3K9 methyltransferases (HMTases) as SUV39H1, Clr4p, DIM-5, Su(var)3–9 or SUVH2 are responsible for the establishment of histone H3 lysine 9 methylation (H3K9me), which is intimately connected with heterochromatinization. In this review, available data will be evaluated concerning (1) the phylogenetic distribution of H3K9me as heterochromatin-specific histone modification and its evolutionary stability in relation to other epigenetic marks, (2) known families of H3K9 methyltransferases, (3) their responsibility for the formation of constitutive heterochromatin and (4) the evolution of Su(var)3–9-like and SUVH-like H3K9 methyltransferases. Compilation and parsimony analysis reveal that histone H3K9 methylation is, next to histone deacetylation, the evolutionary most stable heterochromatic mark, which is established by at least two subfamilies of specialized heterochromatic HMTases in almost all studied eukaryotes.

Key words: cladistic analysis, Clr4p, DIM-5, H3K9 methyltransferases, heterochromatin, histone methylation, molecular evolution, Su(var)3–9, SUV39H1, SUVH2

Introduction

Histone methylation is an abundant epigenetic modification of core histones found in all eukaryotic organisms studied. This modification is catalyzed by histone methyltransferases (HMTases), which introduce methyl groups at lysine or arginine residues. Resulting histone methylation states contribute, according to the histone code hypothesis (Strahl & Allis, 2000; Jenuwein & Allis, 2001), to the regulation of chromatin structure and gene expression. The first identified HMTase was the human SET domain protein SUV39H1 which methylates lysine 9 of histone H3 (H3K9me; Rea et al., 2000). Correspondingly, H3K9 might be the most thoroughly studied site of histone methylation. In several eukaryotic model organisms, H3K9 methylation appears to establish transcriptionally inert chromatin (Schotta et al., 2002; Mellone et al., 2003; Peters et al., 2003; Naumann et al. 2005). This role of H3K9 methylation was recently challenged by Vakoc et al. (2005), which found an enrichment of H3K9me in the transcribed regions of active genes. Actually, there exist three, locally and functionally distinct distributions of H3K9me, which (1) define cytologically visible heterochromatic regions, (2) are involved in silencing of euchromatic genes through epigenetic modification of promotors (for e.g. Schultz et al., 2002; Tachibana et al., 2005), and (3) perhaps take part in the repression of illegitimate initiations of transcription inside of active transcription units, similarly as shown for H3K36me (Carrozza et al., 2005; Vakoc et al., 2006). At the same time, in some organisms as in mammals, *Drosophila* and *Arabidopsis*, euchromatic and heterochromatic locations of H3K9me are partially differentiated by distinct levels of methylation, i.e., H3K9 is found to be preferentially monomethylated (me1), dimethylated (me2), or trimethylated (me3) in different regions of the genome (Peters et al., 2003; Ebert et al. 2004; Naumann et al. 2005).

In this review, the phylogenetic distribution of H3K9 methylation states will be discussed together with their subnuclear distribution and their supposed functions in the corresponding

species. This compilation will reveal the phylogenetic stability of histone H3K9 methylation in relation to other epigenetic marks. Second, it will be determined time of origin and phylogenetic distribution of all known H3K9 histone methyltransferase subfamilies. Furthermore, these HMTase subfamilies will be compared in respect to their *in vitro* and *in vivo* activities and to their chromosomal distributions. This will demonstrate that each eukaryotic organism possesses exactly one type of a heterochromatin-associated HMTase. Third, there will be reviewed the assignment of functions to single domains of heterochromatin-specific HMTases. In turn, domain gains or losses during evolution will be reconstructed and will provide interesting hints about functional changes of these HMTs. In summary, the importance of H3K9 methylation and H3K9 methyltransferases will be demonstrated for heterochromatin establishment and maintenance in evolutionary terms.

It should be noted that the concept heterochromatin is used during this review in the classical sense, that is, as cytological visible condensed chromatin throughout the cell cycle (Heitz, 1928). This type of constitutive heterochromatin includes pericentric regions, heterochromatic knobs, nucleolus organizing regions as well as Y and W chromosomes. The telomeric type of heterochromatin will not considered because of its rather complex composition (Andreyeva et al., 2005) and of its small size. Furthermore, I will not refer to facultative heterochromatin because this sex- or developmental-specific type of condensed chromatin appears heterogenously composed (Gilbert et al., 2003; Kohlmaier et al., 2004; Chadwick & Willard, 2004; Bongiorni et al. 2007).

Three epigenetic marks and evolutionary levels of gene silencing

To determine the evolutionary stability of the link between histone H3 lysine 9 methylation (H3K9me) and heterochromatin, data concerning the distribution of H3K9me in euchromatic and heterochromatic chromosomal domains were collected from literature. The compilation of these data into a tree according to commonly supported phylogeny (Figure 1) shows a widespread occurrence of H3K9 methylation in animals, fungi and green plants as well as in amoebozoan (*Dictyostelium*) and in apicomplexan (*Tetrahymena* and *Stylonychia*) protists. It appears that H3K9me2 is the most common methylation state, however, a bias resulting from the preferential search for the H3K9me2 modification cannot be excluded. All three H3K9 methylation states are typically concentrated in heterochromatin. Interestingly, exceptions for each of these states were identified, where H3K9me1 (in the mouse), H3K9me2 (in maize) or H3K9me3 (in *Arabidopsis*, maize and *Chlamydomonas*) are found mainly in euchromatin. Two of these cases occur in recently separated evolutionary lineages (Figure 1), which argues for a young evolutionary origin of preferencially euchromatic distributed H3K9me1 and H3K9me2. The H3K9me3 state seems to occur in the heterochromatin of unikonts and stichotrichous ciliates but (only rarely) in the euchromatin of plants. Together this suggests that all three H3K9 methylation levels are ancient heterochromatin-specific modifications. One complete loss of H3K9me, identified in the *Saccharomyces* lineage, is recent, because the ascomycet relatives *Neurospora* and *Schizosaccharomyces* show H3K9 methylation marks. To summarize, three of five basal groups of eukaryotes, the unikonts (animals, fungi and Amoebozoa), the plants and the chromalveolates (*Tetrahymena* and *Stylonychia*), show heterochromatin-associated H3K9 methylation. The last common ancestor of these groups is certainly located near the root of the eukaryotic tree (Keeling et al., 2005). Therefore, H3K9 methylation of heterochromatic regions might be an ancient feature of most recent eukaryotes.

In contrast, other histone methylation marks within heterochromatin are more or less lineage-specific. A heterochromatic enrichment of H3K27me1, H3K27me2 and H4K20me1 was specifically found in angiosperms (Naumann et al., 2005; Ebert et al., 2006; Fuchs et al., 2006; Shi & Dawe, 2006), whereas only animals use H4K20me3 to establish heterochromatin (Kourmouli et al., 2004; Sanders et al., 2004; Schotta et al., 2004).

Not only histone methylation contributes epigenetic marks which are known to flag heterochromatin. Notably, in telomeric and mating-type heterochromatin of *Saccharomyces* contains no histone methylation at all is found (for review, see Millar & Grunstein, 2006). Instead, heterochromatin-building nucleosomes in budding yeast are constituted by unacetylated histones. Histone hypoacetylation was found inside the heterochromatin of every eukaryote studied. One could ask why hypoacetylation of histones is more strongly conserved than H3K9 methylation during the evolution of heterochromatin. It may cause by the positive charges of the deacetylated sites, which are involved in forming a condensed chromatin structure through their interactions with DNA, with negatively charged proteins of adjacent nucleosomes or through specific binding of heterochromatin proteins (e.g. Sir3, Sir4; Hecht et al., 1995) to the deacetylated histones. That is, this modification may directly and indirectly determine the chromatin structure, while H3K9 methylation may only be able to induce a chromatin condensation indirectly via methyllysine-sensitive binding by chromo or tudor domains of nonhistone chromatin proteins (e.g. Kim et al., 2006). Additionally, methylation of H3K9 requires prior deacetylation of this residue, which is, for example, accomplished by Clr3p in *Schizosaccharomyces* (Nakayama et al., 2001) and by RPD3 (HDAC1) in *Drosophila* (Czermin et al., 2001). It is thus tempting to speculate that hypoacetylation, because of its more direct mechanism, is the older pathway for heterochromatization than H3K9 methylation.

A third heterochromatin-connected, epigenetic modification is cytosine methylation. DNA methylation represents the only chromatin modification for which a means of stable

propagation through cell division has been directly demonstrated (for review, see Goll & Bestor, 2005). The physical link to the DNA sequence and the rarity of active demetylation mechanisms (except during early development in mammals) made cytosine methylation particularly suitable for an epigenetic memory function. On the other hand, cytosine modification renders the DNA sequence more vulnerable to base substitutions (C to T transitions), which becomes evolutionary underlined through the integration of DNA methylation into the pathway of repeat-induced point mutation (RIP) that has occurred in the *Neurospora* lineage (Selker et al., 2003). In addition, the functional consequences of DNA methylation seem to have changed sometimes during evolution, which is illustrated by an apparent hypomethylation of heterochromatin in the crustacean *Asellus aquaticus* (Barzotti, Pelliccia & Rocchi, 2006) and of facultative heterochromatin in mealy bugs (Bongiorni, Cintio & Prantera, 1999). Similarly as H3K9 methylation, DNA methylation may help to condense chromatin only indirectly, for example by recruiting the methyl-binding, architectural chromatin protein MeCP2 (Georgel et al., 2003) or by interacting with H3K9me as shown in *Arabidopsis* (Naumann et al., 2005) and in vertebrates (Lehnertz et al., 2003; Rai et al., 2006).

The mutational costs of cytosine methylation are reflected by the disjointed phylogenetic distribution of high or intermediate levels of this modification mainly in multicellular eukaryotes, which was deduced by the occurrence of enzymes which belong to the effective DNA methyltransferase families Dnmt1 and Dnmt3 (Ponger & Li, 2005; Goll & Bestor, 2005). Regev, Lamb and Jablonka (1998) suggested that DNA methylation is an ancient mechanism for gene silencing and cell memory which was lost in species with low rates of cell turnover. This is consistent with the complete loss of DNA methylation in the *Saccharomyces*, *Schizosaccharomyces*, and *Caenorhabditis* lineages. It can be concluded that histone hypoacetylation is the primary and phylogenetic most stable epigenetic mark of heterochromatin, followed by histone H3K9 methylation and C5-cytosine methylation. On the

other hand, these epigenetic marks are reversely ranked by ontogenetic stability during cell metabolism, division and differentiation (Figure 2).

Subfamilies of H3K9 methyltransferases and their heterochromatic activities

As argued above, histone H3K9 methylation is an important heterochromatic mark with significant phylogenetic and ontogenetic stability. This is especially true for di- and trimethylated H3K9, as demonstrated in Figure 1. Recently, much work has focused in enzymes which set these marks. In this chapter, all known groups of H3K9 methylating enzymes will be considered in respect to their subnuclear localization, *in vivo* and *in vitro* activities and phylogenetic distribution.

H3K9 methyltransferase activity was first described for the human protein SUV39H1 (Rea et al., 2000). The corresponding Su(var)3–9/Clr4p/DIM-5 (Su[var]3–9-like) subfamily of H3K9 methyltransferases represents an orthologous group of proteins, which was identified by reciprocal best BLAST hits (Altschul et al., 1997) based on the similarity of the catalytically important domains SET (Suppressor of variegation, Enhancer of zeste and Trithorax), preSET and postSET (Krauss et al., 2006; and references therein). These enzymes are phylogenetically restricted to unikonts (animals, fungi and Amoebozoa; Figure 3). Su(var)3–9-like HMTases catalyzes *in vitro* the establishment of all three H3K9 methylation levels (Zhang et al., 2003; Eskeland et al., 2004; Table 1). *Drosophila Su(var)3–9* null mutants show a strong reduction of H3K9me2 and H3K9me3 immunofluorescence signals in the heterochromatic chromocenter (Ebert et al., 2004). Similarly, heterochromatic foci of human double null *SUV39H1*$^{-}$*/SUV39H2*$^{-}$ cells failed to show H3K9 trimethylation (Peters et al., 2003). This coincides with the mainly heterochromatic distribution of the corresponding

enzymes (Aagaard et al., 1999; Schotta et al., 2002) and demonstrate that Su(var)3–9-like enzymes are responsible for heterochromatin-specific H3K9 marks in animals. In addition, a gain-of-function mutation of *Drosophila Su(var)3–9* is able to generate ectopic heterochromatin (Ebert et al., 2004). In fungi with small genomes immunocytological analyses are difficult to perform. However, heterochromatic regions of the *Schizosaccharomyces* mating-type locus are specifically methylated in H3K9 by Clr4p, and this methylation is significantly decreased in *clr4*$^{-}$ strains as determined by CHIP analysis (Noma, Allis & Grewal, 2001). In *Neurospora*, the Su(var)3–9 ortholog DIM-5 methylates H3K9 in genomic regions which are subject to repeat-induced point mutation (RIP) (Tamaru & Selker, 2001). The authors discussed that relics of RIP may underlie a heterochromatic transformation. Thus, all HMTases of this group seem to involve in establishment of some type of heterochromatic structures.

The SUVH (Su[var]3-9 homologs) subfamily of H3K9 methyltransferases was initially described by Baumbusch et al. (2001) as a group of ten *Arabidopsis* paralogs and is characterized by a combination of SET, preSET, postSET and YDG domains (Figure 4). SUVH proteins are probably restricted to the Chlorophyta (orthologs exist in the green algae *Volvox* and *Chlamydomonas*, Figure 3). An *in vitro* H3K9 mono- and dimethylase activity was shown for several angiosperm SUVH proteins (Ebbs & Bender, 2006; and references therein), whereas the *Chlamydomonas* ortholog SET3p appear to be an exclusive H3K9 monomethylase (Casas-Mollano et al., 2007) which is *in vivo* responsible for a significant part of the H3K9me1 signal. GFP-tagged SUVH1, SUVH2 and SUVH4 (KYP) proteins localize to heterochromatic foci of *Arabidopsis* interphase nuclei (Naumann et al., 2005; Fischer et al., 2006). Meanwhile, null mutants from *SUVH1, SUVH2, SUVH4, SUVH5* and *SUVH6* were evaluated (Ebbs & Bender, 2006; Fischer et al., 2006; and references therein). Only *SUVH2*$^{-}$ plants show a genome-wide, significant reduction of H3K9 mono- and dimethylation (Naumann et al., 2005). On the other hand, both the tobacco homolog SET1 and *Arabidopsis*

SUVH2 induce through overexpression a significant rise of H3K9me2 and a miniplant phenotype, which is, at least in the case of SUVH2, connected with the generation of ectopic heterochromatin (Shen, 2001; Yu, Dong & Shen, 2004; Naumann et al., 2005). In contrast, the inconspicuous phenotype of *SUVH1, SUVH4, SUVH5* and *SUVH6* mutants is supposed to be caused by the functional redundance of the multiple paralogs. In summary, SUVH methyltransferases are probably the predominant heterochromatic H3K9 HMTases in plants.

The first functional analysis of a member of the G9a/GLP subfamily of H3K9 HMTases was performed by Tachibana et al. (2001) for human G9a. G9a/GLP proteins are characterized by a combination of preSET, SET and postSET domains with Ankyrin repeats. Until now, such proteins could be solely found in animals (Figure 3). Orthologs in the octopus *Euprymna scolopes* (EST) and in the coral *Nematostella vectensis* (genomic sequence) were identified by reciprocal best BLAST hits. Human and *Drosophila* G9a show an *in vitro* methylase activity on H3K9 and H3K27 at free native and recombinant histones (Tachibana et al., 2001; Tachibana et al., 2005; Stabell et al., 2006). *In vivo*, H3K9 was mono- and dimethylated exclusively at euchromatic loci of mammalian cells (Peters et al., 2003; Rice et al. 2003). The endogenous G9a protein is localized specifically in euchromatin of *Drosophila* polytene chromosomes (Stabell et al., 2006) and is found in chromatin-associated and soluble fractions of mammalian nuclei (Loyola et al., 2006). G9a and their mammalian paralog GLP (G9a-like protein) are involved in regulation of euchromatic genes (Tachibana et al., 2005; Stabell et al., 2006), probably as already shown by gene silencing through methylation of H3K9 (Stewart, Li & Wong, 2005). However, an influence of G9a on H3K9 methylation states in heterochromatin cannot be excluded completely because of a partially heterochromatic distribution of overexpressed G9a in mammalian cells (Esteve et al., 2005) and of a weak suppression effect on PEV which was ascribed to a insertion mutation in the 5´UTR of *Drosophila G9a* (Mis, Ner & Grigliatti, 2006).

The human H3K9 HMTase SETDB1 (ESET) was first described by Schultz et al. (2002). The corresponding group of SETDB/ESET proteins are characterized by a SET domain with a specific insert region, a preSET region and a methyl-binding domain (MBD). They show a disrupted phylogenetic distribution: Orthologs are found in animals as cnidarians, flatworms, nematodes, arthropods and annelids, but not in fungi, whereas a orthologous EST was identified in *Acanthamoeba castellanii* (Amoebozoa). Thus, SETDB/ESET HMTases are supposed to originate approximately at the same time as Su(var)3–9-like enzymes in a common ancestor of unikonts (Figure 3). Human SETDB1 was found mainly in cytosolic and non-nucleosomal nuclear fractions of Hela cells (Loyola et al., 2006). Immunofluorescence analysis in mammalian cell nuclei revealed SETDB1 predominantly in euchromatic regions, where it participates in gene silencing (Schultz et al., 2002). Drosophila SETDB1 was found mainly at chromosome 4, but also in euchromatin and at the chromocenter of polytene chromosomes (Seum et al., 2007). A Drosophila SETDB1$^-$ mutant showed a signifikant loss of all three methylation levels in western analysis (Seum et al., 2007). SETDB1 may monomethylate H3K9 of the centromeric heterochromatin in support of the Su(var)3-9 di- and trimethylase, which is consistent with the wild type H3K9me1 distrubution observed in Su(var)3-9$^-$ mutants (Ebert et al., 2004). Without SETDB1 the relative concentrations of the H3K9 methylation levels in the chromo center are unchanged (Seum et al., 2007) but their absolute amounts might be reduced.

Interestingly, recombinant SETDB1 catalyzed *in vitro* only mono- and dimethylation of H3K9, but is able to generate H3K9me3 in interaction with the transcriptional repressor mAM (Wang et al., 2003). In addition, SETDB1/ESET can be artificially tethered by HP1 to heterochromatic sites, where it is supposed to reconstitute the H3K9me3 modification in *SUV39H1*$^-$*/SUV39H2*$^-$ double null cells (Kourmouli et al., 2005). Accordingly, SETDB/ESET protein complexes, in difference to G9a/GLP proteins, are potent H3K9me3 HMTases *in vivo*,

however, wild type localization and the limited methyltransferase activity argues against a significant role in heterochromatin establisment.

Recently, a plant-specific subfamily SUVR (Su[var]3–9-related) of H3K9 methyltransferases was functionally characterized. The *Arabidopsis* SUVR4 protein appears exclusively to transform H3K9me1 to H3K9me2 (Thorstensen et al., 2006). SUVR1, SUVR2 and SUVR4 were found specifically enriched in nucleoli but not in the heterochromatic nucleolus organizing regions, which argues against a role in heterochromatin formation. Additionally to SET and preSET domains, SUVR proteins contain a novel homology region named WIYLD according to characteristic conserved amino acids (Thorstensen et al., 2006). Outside of embryophytes, a SUVR-orthologous gene SET5 were identified in *Chlamydomonas* (Casas-Mollano et al., 2007), therefore, the SUVR HMTase group may be restricted to chlorophytes (Figure 3).

The human RIZ1 (PRDM2) tumor suppressor protein contains a SET domain variant, which was named the PR domain (Huang, Shao & Liu, 1998). It was shown that RIZ1 has H3K9 HMTase activity *in vitro* (Kim, Geng & Huang 2003), however, the subnuclear distribution of RIZ1 is unknown. Reciprocal BLAST analysis show that orthologs of RIZ1 are restricted to tetrapods. In contrast, ESTs coding for PR domain proteins are identified also outside the vertebrates in arthropods (e. g. in *Drosophila*), nematodes and in *Oscarella carmela* (Porifera), which points to a general occurrence of PR domain proteins in animals (Figure 3).

A H3K9 methyltransferase activity was additionally reported for HMTases which appear to show a preference for other histone tail positions, e. g. the Ash1 and E(z) protein family members. A E(z)-containing HMTase complex from *Drosophila* is able to produce H3K9me1, H3K9me2 and H3K9me3 *in vitro* (Czermin et al., 2002). However, immunofluorescence of polytene chromosomes of a *E(z)* null mutant shows that E(z) does not appear to affect H3K9 methylation *in vivo* (Ebert et al., 2004). Therefore, the suppressor effect of *E(z)* mutations on PEV (Laible et al., 1997) results more likely from the H3K27 HMTase activity

of the enzyme, which is found in eu- and in heterochromatin (Ebert et al., 2004). E(z) family proteins are characterized by a combination of a SET, cysteine-rich preSET and a family-specific, N-terminal located domain (called E(z) domain II in Baumbusch et al., 2001). Their phylogenetic distribution is wide (Figure 3), orthologous ESTs were found in unikonts, plants and chromalveolates, that is, in three of the five major groups of eukaryotes (Keeling et al., 2005).

A similar ancient evolutionary origin is suggested for the Ash1 protein family, which is characterized mainly by the type of SET domain and its unusual location in the middle of the polypeptide. Analogous to E(z), orthologous ESTs were identified in unikonts, plants and chromalveolates. According to Beisel et al. (2002), *Drosophila* Ash1 methylates H3K9 *in vitro* and increases the general amount of H3K9me in polytene chromosomes. Byrd and Shearn (2003) suggested a decrease of H3K9 methylation signals in the chromocenter of polytene chromosomes on *Ash1* loss-of-function conditions. However, the exclusive localization of Ash1 at euchromatic sites (Tripoulas et al., 1996) and the appearing inability of the *Arabidopsis* ortholog ATX-1 to methylate H3K9 *in vitro* and *in vivo* (Alvarez-Venegas et al., 2003; Alvarez-Venegas & Avramova, 2005) argues against a significant and/or evolutionary conserved involvement of Ash1-related proteins in H3K9 methylation and heterochromatin establisment. Additionally, the mainly reported HMTase activity of Ash1 is the methylation of H3K4 in animals and plants (Beisel et al., 2002; Alvarez Venegas et al., 2003). H3K4 is hypomethylated in heterochromatic regions of *Drosophila*, mammals and angiosperms (Houben et al., 2003; Vakoc et al., 2005; Ebert et al., 2006). Moreover, the active removal of H3K4 methylation by the *Drosophila* histone demethylase Su(var)3-3 was found to be a prerequisite for H3K9 methylation during the developmental establishment of heterochromatin (Rudolph et al., 2007). This evidence is difficult to reconcile with a HMTase activity of Ash1 orthologs inside of heterochromatin.

To summarize, only Su(var)3–9-like (Suv39/Clr4p/DIM-5) proteins in unikonts and SUVH proteins in plants are H3K9 methyltransferases which are able to establish constitutive heterochromatin. Other groups of H3K9 HMTases, as G9a/GLP, SETDB/ESET, SUVR and probably PR domain proteins, are supposed to be involved in H3K9 methylation outside of heterochromatin, at promotors of inactive genes, inside of active transcription units or outside of nucleosomes (Loyola et al., 2006). Therefore, it appears not appropriate to treat H3K9me as a mark which is sufficient to determine heterochromatin on its own. Promotors which are silenced through H3K9 methylation should thus not be termed "heterochromatized". Well-known properties of constitutive heterochromatin are condensation throughout interphase, high proportion of repetitious sequences, late S-phase replication, sparse transcription, suppression of recombination, spreading ability which results in position-effect variegation (PEV), sequence-unspecific self-interaction, underreplication in polyploid nuclei, a regular spacing of nucleosomes, histone hypoacetylation, histone H3K4 hypomethylation, DNA hypermethylation and H3K9 hypermethylation (for review see e. g. Reuter, Fischer & Hofmann, 2005; Huisinga, Brower-Toland & Elgin, 2006). Not all of these properties are detectable in all organisms, but in most cytological analyzed eukaryotic species, defined heterochromatic blocks exist, often visible, side-by-side to active and silenced euchromatic regions. Such heterochromatic regions showed regularly an enrichment of H3K9me2 or H3K9me3 (Figure 1), together with a mainly heterochromatin-specific H3K9 HMTase of the Su(var)3–9-like or of the SUVH type if this was evaluated. On the other hand, euchromatin-specific H3K9 HMTases were identified at least in animals. Both types of proteins are clearly not able to substitute for each other (Schotta et al., 2002; Peters et al., 2003; Stabell et al., 2006), which is surely caused by specific targeting. For example, the restriction of the Su(var)3–9 localization mainly to heterochromatin is dependent from interactions with Su(var)3-3, RPD3 and HP1, as shown in *Drosophila* (Schotta et al., 2002; Rudolph et al., 2007). Similar interactions are conserved in mammals (Aagaard et al., 1999; Lee et al., 2006). The plant enzyme SUVH2 is also restricted to heterochromatin, which need

to be mediated by more than one domain of the protein as revealed by mutant analysis (Naumann et al., 2005). Similarly, Su(var)3–9 shows a participation of the N-terminus, the chromo and the SET domain on chromosomal localization in Drosophila (Schotta et al., 2002; Figure 4). Interestingly, heterochromatic HMTases are able to broadening their own chromosomal distribution following hyperactivity (*Su[var]3-9*^{ptn} allele in *Drosophila*; Ebert et al., 2004) or overexpression (SUVH2 in *Arabidopsis*; Naumann et al., 2005), which results in formation of ectopic heterochromatin. That is, the lower limit of sufficient targeting interactions can be decreased by increasing the enzymatic activity or the cellular concentration. The association of these H3K9 HMTases with chromatin appears to depend on its enzymatic activity and seems sufficient for heterochromatin formation, which might in turn induce an early arrest of development (*Su[var]3-9*^{ptn} allele in *Drosophila*; Kuhfittig et al., 2001). Another evolutionary conserved hallmark of heterochromatin, HP1, can recruit Su(var)3–9 only by artificial tethering to impose heterochromatin-like regularly spaced nucleosomes and gene silencing in euchromatin of *Drosophila* (Danzer & Wallrath 2004). This is consistent with results from *Xenopus* oocytes which demonstrate that H3K9me and SUV39H1, but not H3K9me and G9a, can recruit HP1 to chromatin (Stewart, Li & Wong, 2005). In contrast to HP1, a substantial population of endogenous SUV39H1 is immobile at heterochromatin, which depends on the SET domain (Krouwels et al., 2005). Thus, heterochromatic H3K9 methyltransferases as Su(var)3–9 and SUVH2 play a central role for heterochromatin formation by both enzymatic activity and as a structural component, which cannot be substituted by other, typically euchromatic distributed H3K9 HMTases.

Evolution of heterochromatic H3K9 methyltransferases

According to the above discussed central role of distinct H3K9 HMTases in heterochromatin, their origin should be strongly connected with the emergence of heterochromatic regions in eukaryotic genomes. Unfortunately, the phylogenetic distribution of visible heterochromatin is rather disjointed. *Saccharomyces* has no pericentromeric heterochromatin and no H3K9 methylation (for review see Millar & Grunstein, 2006). H3K9 methylation exists in *Caenorhabditis* (Reuben & Lin, 2002), however, no visible heterochromatin is found. Also other organisms with holocentric chromosomes do not have pericentromeric heterochromatin (Criniti et al., 2005, and references therein). Moreover, H3K9 HMTases might be lost entirely during evolution, if genomes become profoundly shrinked as in the *Saccharomyces* lineage. On the other hand, these enzymes are found in almost all studied plants (SUVH) and unikonts (Su(var)3–9-like enzymes) (Naumann et al., 2005; Krauss et al., 2006). Additionally, the ciliates *Tetrahymena thermophyla* and *Stylonychia lemnae* perform H3K9 methylation specifically at sequences determined for elimination during development of the macronucleus (Taverna, Coyne & Allis, 2002; Juranek et al., 2005). This process appears functionally analogous to heterochromatinization in multicellular organisms, which is supported by specific interaction of the chromo domain-containing proteins Pdd1 and Pdd3 with the H3K9-dimethylated chromatin in *Tetrahymena*. Thus, a structural and functional molecular ancestor of heterochromatic H3K9 HMTases should have been existed in early eukaryots which diversified later to unikonts, plants and chromalveolates (Figure 3). This is further supported by reciprocal BLAST analysis of heterochromatin protein 1 (HP1), which interacts with H3K9me and Su(var)3-9-like HMTases, and is found in the same groups.

The evolution of heterochromatic H3K9 HMTases, which are characterized by the combination of a SET domain and typical cysteine-rich preSET and postSET regions, was accompanied by the acquisition of different additional domains in different lineages. The

Chlorophyta enzymes (SUVH group, which includes Set3p of *Chlamydomonas reinhardtii*) received a N-terminal YDG (SRA) domain. The molecular function of this domain might consist in interaction with methylated cytosines (Unoki, Nishidate & Nakamura, 2004; Johnson et al., 2007) or in binding of the N-terminal tail of histone H3 (Citterio et al., 2004). The YDG domain of SUVH2 appears to be involved in recruitment of symmetrical (CpG and CpNpG) and asymmetrical (CpNpN) DNA methylation to target sequences, which is a prerequisite for histone methylation by SUVH2 (Naumann et al., 2005).

At about the same time, the Su(var)3-9-like protein of a common ancestor of opisthoconts (unikonts excluding Amoebozoa) acquires a chromo domain, which is supported by the occurrence of chromo domains in the corresponding enzymes of bilaterian animals, *Schizosaccharomyces* (Clr4p) and the zygomycet *Rhizopus oryzae* (Figure 5). Isolated chromo domains of human SUV39H1 are able to bind specifically H3K9me3 (Jacobs, Fischle & Khorasanizadeh, 2004). Substitutions of conserved amino acids of chromo domains impaired the silencing function of Clr4p in *Schizosaccharomyces* (Ivanova et al., 1998) and the H3K9 methylation activity of SUV39H1 *in vitro* (Chin et al., 2006). A comprehensive phylogenetic comparison suggested that Su(var)3–9-like chromo domains might generally bound H3K9me similar as those of HP1 (Krauss et al., 2006). GFP fusion transgenes in *Drosophila* showed that the chromo domain is, together with the N-terminus and the SET domain, essential for pericentromeric localization of Su(var)3–9 (Schotta et al., 2002). Together, these results suggest that the histone H3 tails which become methylated by the SET domain will be immediately bound by the Su(var)3-9 chromo domain. This, in turn, would facilitate the methylation of neighboring histones. Such a mechanism would improve the efficiency of enzymes as Su(var)3-9 or SUV39H1, which appear to be nonprocessive H3K9me3 HMTases, that is, after a transfer of one methyl group the histone tail has to dissociate and rebind to receive the next methyl residue (Eskeland et al., 2004; Chin et al., 2006).

Remarkably, the phylogenetic tree (Figure 5) revealed that the chromo domain was lost independently in at least three lineages during the evolution of fungi and animals. One cause for this loss might be the ability of fungal orthologs as *Neurospora* DIM-5 to processively transfer three methyl residues to a continuously bound histone H3 tail (Zhang et al., 2003). Such effective enzymes might not need a supporting module like the chromo domain. Moreover, an unbalanced, improved catalysis can be fatal, as shown for the *Su(var)3–9*^{ptn} gain-of-function mutation in *Drosophila* (Kuhfittig et al., 2001; Ebert et al., 2004). Thus, an adaptation of molecular functions might have occurred, which connects the loss of the chromo domain with some counter-acting changes in the SET domain of DIM-5 enzymes (Krauss et al., 2006).

The gain and loss of chromo domains are not the only gross structural changes of Su(var)3–9-like proteins during evolution. A common ancestor of dicondylic insects (winged insects and thysanurans) has put the Su(var)3–9 open reading frame, probably by retrotransposition, into an intron of the translation initiation factor subunit gene *eIF2*γ (Krauss et al., 2006). Consequently, *Su(var)3–9* mRNAs become expressed by 3´alternative splicing of the *eIF2*γ gene (Figure 6), using one large *Su(var)3–9* specific exon instead of two or more *eIF2*γ-specific exons. For both involved gene products, this gene fusion became the only source of functional mRNAs (Krauss et al., 2006). The 5´ exons are common to both alternative splice variants and coding for about 80 N-terminal amino acids. This novel N-terminus of Su(var)3–9 cannot fulfill similar functions as in eIF2γ (for e.g. GTP binding) because it is structurally incomplete (Schmitt, Blanquet & Mechulam, 2002). Thus, the fusion of both genes cannot have been a positively selected event and should have occurred only once (Krauss et al., 2006).

In contrast, a re-fission of this kind of gene fusion has took place at least three times during the evolution of insects (Figure 5). In the cases of the pea aphid *Acyrthosiphon pisum* and the human louse *Pediculus humanus*, this reversion may have occurred by a renewed

retrotransposition or a genomic duplication of the *Su(var)3–9* part, because the exon-intron-structure of the *eIF2γ* gene remained largely unchanged (Krauss et al., 2006; data not shown). Notably, in the case of the jewel wasp *Nasonia vitripennis*, the *eIF2γ* part has probably experienced a retrotransposition, because all introns of this gene became lost, which is highly unusual for an animal copy of this gene (Krauss et al., 2005). Taken together, those three independent, but convergent re-fission events support not only the accidental character of the gene fusion event, they let furthermore suspect that the *Su(var)3–9/eIF2γ* fusion imposes a functional burden on the encoded gene products. Indeed, in beetles and butterflies obvious splice artefacts, containing all exons of the fusion, are detectable (Krauss & Reuter, 2000), which might express functionless or antimorphic proteins.

Not only the domain structure of Su(var)3–9 proteins, also the degree of conservation has changed during the evolution of animals. *Caenorhabditis* orthologs show very long tree branch lengths, in difference to the Su(var)3-9 protein of their close relative *Trichinella spiralis* (Figure 5). The corresponding *Caenorhabditis* Su(var)3-9-like proteins may not be able to methylate H3K9 because of a very incomplete conservation of chromo and SET domain (Rea et al., 2000). In *C. elegans*, this correlates with the restriction of H3K9me to male germline X chromosomes and to telomeres (Reuben & Lin, 2002). It would be interesting to see what molecular functions are covered by this unusual ortholog. In contrast, the loss of pericentromeric heterochromatin alone, as evident in all holocentric species, is obviously not correlated with reduced structural conservation of Su(var)3–9 HMTases, which was demonstrated by the successful cloning of orthologs from butterflies (*Scoliopteryx, Bombyx*), hemipterans (*Cercopis, Acyrthosiphon*) and earwigs (*Forficula*) (Krauss & Reuter, 2000; Krauss et al., 2006). This points to a supposed role of Su(var)3-9 in the establishment of the H3K9me3 mark identified in the facultative heterochromatin of the holocentric insect Planococcus (Bongiorni et al., 2007), but also to other roles of Su(var)3–9

as described in euchromatic gene silencing (Ivanova et al., 1998; Hwang, Eissenberg & Worman, 2001), at telomeres (Donaldson, Lui & Karpen, 2002), in chromosome segregation (Peters et al., 2001) and in nuclear organization (Peng and Karpen, 2007).

In conclusion, the phylogenetic distribution of H3K9 methylation within the genome shows that all levels of histone H3-K9 methylation are common indexing marks of cytologically visible, constitutive heterochromatin in most of the evaluated species of eukaryotes. These marks are established by structurally conserved SET domain protein subfamilies, which display a rather diverse, but functionally convergent domain architecture. Their heterochromatin-establishing property corresponds to both H3K9 HMTase activity and molecular interactions with proteins like HP1 and, possibly, RNA molecules (Maison et al., 2002; Pal-Bhadra et al., 2004). These molecular functions are common to different heterochromatic H3K9 HMTase subfamilies as Su(var)3–9/Clr4p/DIM-5 as well as SUVH and differentiate them from other H3K9-methylating enzymes.

Acknowledgements

I would like to thank Dr Gunter Reuter for critical reading of the manuscript. I gratefully acknowledge the sequencing of the yet unpublished genomes of *Coprinus cinereus, Rhizopus oryzae, Nematostella vectensis, Acropora millepora, Hydra magnipapillata, Schistosoma mansoni, Aplysia californica, Pediculus humanus, Nasonia vitripennis, Tribolium castaneum* and *Trichinella spiralis*. Research on evolution of Su(var)3–9-like proteins herein was funded by a grant from the Deutsche Forschungsgemeinschaft.

References

Aagaard, L., G. Laible, P. Selenko, M. Schmid, R. Dorn, G. Schotta, S. Kuhfittig, A. Wolf, A. Lebersorger, P.B. Singh, G. Reuter & T. Jenuwein, 1999. Functional mammalian homologues of the Drosophila PEV modifier *Su(var)3-9* encode centromere-associated proteins which complex with the heterochromatin component M31. EMBO J. 18(7): 1923 38.

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller & D.J. Lipman, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(17): 3389-402.

Alvarez-Venegas, R., S. Pien, M. Sadder, X. Witmer, U. Grossniklaus & Z. Avramova, 2003. ATX-1, an *Arabidopsis* homolog of trithorax, activates flower homeotic genes. Curr Biol. 13(8): 627-37.

Alvarez-Venegas, R. & Z. Avramova, 2005. Methylation patterns of histone H3 Lys 4, Lys 9 and Lys 27 in transcriptionally active and inactive *Arabidopsis* genes and in *atx1* mutants. Nucleic Acids Res. 33(16): 5199-207.

Andreyeva, E.N., E.S. Belyaeva, V.F. Semeshin, G.V. Pokholkova & I.F. Zhimulev, 2005. Three distinct chromatin domains in telomere ends of polytene chromosomes in *Drosophila melanogaster Tel* mutants. J Cell Sci. 118: 5465-77.

Barzotti, R., F. Pelliccia & A. Rocchi, 2006. DNA methylation, histone H3 methylation, and histone H4 acetylation in the genome of a crustacean. Genome 49(1): 87-90.

Baumbusch, L.O., T. Thorstensen, V. Krauss, A. Fischer, K. Naumann, R. Assalkhou, I. Schulz, G. Reuter & R.B. Aalen, 2001. The *Arabidopsis thaliana* genome contains at least 29 active

genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. Nucleic Acids Res. 29(21):4319-33.

Beisel, C., A. Imhof, J. Greene, E. Kremmer & F. Sauer, 2002. Histone methylation by the *Drosophila* epigenetic transcriptional regulator Ash1. Nature 419: 857-62.

Bongiorni, S., O. Cintio & G. Prantera, 1999. The relationship between DNA methylation and chromosome imprinting in the coccid *Planococcus citri*. Genetics 151(4): 1471-8.

Bongiorni, S., B. Pasqualini, M. Taranta, P.B. Singh & G. Prantera, 2007. Epigenetic regulation of facultative heterochromatinisation in Planococcus citri via the Me(3)K9H3-HP1-Me(3)K20H4 pathway. J Cell Sci. 120(6): 1072-80.

Byrd, K.N. & A. Shearn, 2003. ASH1, a *Drosophila* trithorax group protein, is required for methylation of lysine 4 residues on histone H3. Proc Natl Acad Sci USA 100(20): 11535-40.

Carrington, E.A. & R.S. Jones, 1996. The Drosophila *Enhancer of zeste* gene encodes a chromosomal protein: examination of wild-type and mutant protein distribution. Development 122(12): 4073-83.

Carrozza, M.J., B. Li, L. Florens, T. Suganuma, S.K. Swanson, K.K. Lee, W.J. Shia, S. Anderson, J. Yates, M.P. Washburn & J.L. Workman, 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell 123(4): 581-92.

Casas-Mollano, J.A., K. van Dijk, J. Eisenhart & H. Cerutti, 2007. SET3p monomethylates histone H3 on lysine 9 and is required for the silencing of tandemly repeated transgenes in *Chlamydomonas*. Nucleic Acids Res. 35(3): 939-50.

Chadwick, B.P. & H.F. Willard, 2004. Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. Proc Natl Acad Sci U S A 101(50): 17450-5.

Chin, H.G., D. Patnaik, P.O. Esteve, S.E. Jacobsen & S. Pradhan, 2006. Catalytic properties and kinetic mechanism of human recombinant Lys 9 histone H3 methyltransferase SUV39H1: participation of the chromodomain in enzymatic catalysis. Biochemistry 45(10): 3272-84.

Citterio, E., R. Papait, F. Nicassio, M. Vecchi, P. Gomiero, R. Mantovani, P.P. Di Fiore & I.M. Bonapace, 2004. Np95 is a histone-binding protein endowed with ubiquitin ligase activity. Mol Cell Biol. 24(6): 2526-35.

Cowell, I.G., R. Aucott, S.K. Mahadevaiah, P.S. Burgoyne, N. Huskisson, S. Bongiorni, G. Prantera, L. Fanti, S. Pimpinelli, R. Wu, D.M. Gilbert, W. Shi, R. Fundele, H. Morrison, P. Jeppesen & P.B. Singh, 2002. Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. Chromosoma 111(1): 22-36.

Criniti, A., G. Simonazzi, S. Cassanelli, M. Ferrari, D. Bizzaro & G.C. Manicardi, 2005. X-linked heterochromatin distribution in the holocentric chromosomes of the green apple aphid *Aphis pomi*. Genetica 124(1): 93-8.

Czermin, B., G. Schotta, B.B. Hulsmann, A. Brehm, P.B. Becker, G. Reuter & A. Imhof, 2001. Physical and functional association of SU(VAR)3-9 and HDAC1 in *Drosophila*. EMBO Rep. 2(10):915-9.

Czermin, B., R. Melfi, D. McCabe, V. Seitz, A. Imhof & V. Pirrotta, 2002. *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell 111(2): 185-96.

Danzer, J.R. & L.L. Wallrath, 2004. Mechanisms of HP1-mediated gene silencing in *Drosophila.* Development 131(15): 3571-80.

Donaldson, K.M., A. Lui & G.H. Karpen, 2002. Modifiers of terminal deficiency-associated position effect variegation in *Drosophila.* Genetics 160(3): 995-1009.

Ebbs, M.L. & J. Bender, 2006. Locus-specific control of DNA methylation by the *Arabidopsis* SUVH5 histone methyltransferase. Plant Cell 18(5): 1166-76.

Ebert, A., G. Schotta, S. Lein, S. Kubicek, V. Krauss, T. Jenuwein & G. Reuter, 2004. *Su(var)* genes regulate the balance between euchromatin and heterochromatin in *Drosophila*. Genes Dev. 18(23): 2973-83.

Ebert, A., S. Lein, G. Schotta & G. Reuter, 2006. Histone modification and the control of heterochromatic gene silencing in *Drosophila*. Chromosome Res. 14(4): 377-92.

Eskeland, R., B. Czermin, J. Boeke, T. Bonaldi, J.T. Regula & A. Imhof, 2004. The N-terminus of *Drosophila* SU(VAR)3-9 mediates dimerization and regulates its methyltransferase activity. Biochemistry 43(12): 3740-9.

Esteve, P.O., D. Patnaik, H.G. Chin, J. Benner, M.A. Teitell & S. Pradhan, 2005. Functional analysis of the N- and C-terminus of mammalian G9a histone H3 methyltransferase. Nucleic Acids Res. 33(10): 3211-23.

Fischer, A., I. Hofmann, K. Naumann & G. Reuter, 2006. Heterochromatin proteins and the control of heterochromatic gene silencing in *Arabidopsis.* J Plant Physiol. 163(3): 358-68.

Fuchs, J., D. Demidov, A. Houben & I. Schubert, 2006. Chromosomal histone modification patterns - from conservation to diversity. Trends Plant Sci. 11(4):199-208.

Georgel, P.T., R.A. Horowitz-Scherer, N. Adkins, C.L. Woodcock, P.A. Wade & J.C. Hansen, 2003. Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. J Biol Chem. 278(34): 32181-8.

Gilbert, N., S. Boyle, H. Sutherland, J. de Las Heras, J. Allan, T. Jenuwein & W.A. Bickmore, 2003. Formation of facultative heterochromatin in the absence of HP1. EMBO J. 22(20):5540-50.

Goll, M.G. & T.H. Bestor, 2005. Eukaryotic cytosine methyltransferases. Annu Rev Biochem. 74: 481-514.

Hecht, A., T. Laroche, S. Strahl-Bolsinger, S.M. Gasser & M. Grunstein, 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(4): 583-92.

Heitz, E., 1928. Das Heterochromatin der Moose. Jahrb Wiss Bot. 69: 762-818.

Houben, A., D. Demidov, D. Gernand, A. Meister, C.R. Leach & I. Schubert, 2003. Methylation of histone H3 in euchromatin of plant chromosomes depends on basic nuclear DNA content. Plant J. 33(6): 967-73.

Huang, S., G. Shao & L. Liu, 1998. The PR domain of the Rb-binding zinc finger protein RIZ1 is a protein binding interface and is related to the SET domain functioning in chromatin-mediated gene expression. J Biol Chem. 273(26): 15933-9.

Huisinga, K.L., B. Brower-Toland & S.C. Elgin, 2006. The contradictory definitions of heterochromatin: transcription and silencing. Chromosoma 115(2): 110-22.

Hwang, K.K., J.C. Eissenberg & H.J. Worman, 2001. Transcriptional repression of euchromatic genes by *Drosophila* heterochromatin protein 1 and histone modifiers. Proc Natl Acad Sci USA 98(20): 11423-7.

Ivanova, A.V., M.J. Bonaduce, S.V. Ivanov & A.J. Klar, 1998. The chromo and SET domains of the Clr4 protein are essential for silencing in fission yeast. Nat Genet. 19(2): 192-5.

Jacobs, S.A., W. Fischle & S. Khorasanizadeh, 2004. Assays for the determination of structure and dynamics of the interaction of the chromodomain with histone peptides. Methods Enzymol. 376: 131-48.

Jenuwein, T. & C.D. Allis, 2001. Translating the histone code. Science 293: 1074-80.

Johnson, L.M., M. Bostick, X. Zhang, E. Kraft, I. Henderson, J. Callis & S.E. Jacobsen, 2007. The SRA methyl-cytosine-binding domain links DNA and histone methylation. Curr Biol. 17(4): 379-84.

Juranek, S.A., S. Rupprecht, J. Postberg & H.J. Lipps HJ, 2005. snRNA and heterochromatin formation are involved in DNA excision during macronuclear development in stichotrichous ciliates. Eukaryot Cell 4(11): 1934-41.

Kaller, M., U. Euteneuer & W. Nellen, 2006. Differential effects of heterochromatin protein 1 isoforms on mitotic chromosome distribution and growth in *Dictyostelium discoideum*. Eukaryot Cell 5(3): 530-43.

Keeling, P.J., G. Burger, D.G. Durnford, B.F. Lang, R.W. Lee, R.E. Pearlman, A.J. Roger & M.W. Gray, 2005. The tree of eukaryotes. Trends Ecol Evol. 20(12): 670-6.

Kim, J., J. Daniel, A. Espejo, A. Lake, M. Krishna, L. Xia, Y. Zhang & M.T. Bedford, 2006. Tudor, MBT and chromo domains gauge the degree of lysine methylation. EMBO Rep. 7(4): 397-403.

Kim, K.C., L. Geng & S. Huang, 2003. Inactivation of a histone methyltransferase by mutations in human cancers. Cancer Res. 63(22): 7619-23.

Kohlmaier, A., F. Savarese, M. Lachner, J. Martens, T. Jenuwein & A. Wutz, 2004. A chromosomal memory triggered by *Xist* regulates histone methylation in X inactivation. PLoS Biol. 2(7): e171.

Kourmouli, N., P. Jeppesen, S. Mahadevhaiah, P. Burgoyne, R. Wu, D.M. Gilbert, S. Bongiorni, G. Prantera, L. Fanti, S. Pimpinelli, W. Shi, R. Fundele & P.B. Singh, 2004. Heterochromatin and tri-methylated lysine 20 of histone H4 in animals. J Cell Sci. 117(Pt 12): 2491-501.

Kourmouli, N., Y.M. Sun, S. van der Sar, P.B. Singh & J.P. Brown, 2005. Epigenetic regulation of mammalian pericentric heterochromatin *in vivo* by HP1. Biochem Biophys Res Commun. 337(3): 901-7.

Krauss, V. & G. Reuter, 2000. Two genes become one: the genes encoding heterochromatin protein Su(var)3-9 and translation initiation factor subunit eIF-2γ are joined to a dicistronic unit in holometabolic insects. Genetics 156(3): 1157-67.

Krauss, V., M. Pecyna, K. Kurz & H. Sass, 2005. Phylogenetic mapping of intron positions: A case study of translation initiation factor eIF2γ. Mol Biol Evol. 22(1): 74-84.

Krauss, V., A. Fassl, P. Fiebig, I. Patties & H. Sass, 2006. The evolution of the histone methyltransferase gene *Su(var)3-9* in metazoans includes a fusion with and a re-fission from a functionally unrelated gene. BMC Evol Biol. 6: e18.

Krouwels, I.M., K. Wiesmeijer, T.E. Abraham, C. Molenaar, N.P. Verwoerd, H.J. Tanke & R.W. Dirks, 2005. A glue for heterochromatin maintenance: stable SUV39H1 binding to heterochromatin is reinforced by the SET domain. J Cell Biol. 170(4): 537-49.

Kuhfittig, S., J. Szabad, G. Schotta, J. Hoffmann, E. Mathe & G. Reuter, 2001. pitkin(D), a novel gain-of-function enhancer of position-effect variegation, affects chromatin regulation during oogenesis and early embryogenesis in Drosophila. Genetics 157(3): 1227-44.

Laible, G., A. Wolf, R. Dorn, G. Reuter, C. Nislow, A. Lebersorger, D. Popkin, L. Pillus & T. Jenuwein, 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(11): 3219-32.

Lee, M.G., C. Wynder, D.A. Bochar, M.A. Hakimi, N. Cooch & R. Shiekhattar, 2006. Functional interplay between histone demethylase and deacetylase enzymes. Mol Cell Biol. 26(17): 6395-402.

Lehnertz, B., Y. Ueda, A.A. Derijck, U. Braunschweig, L. Perez-Burgos, S. Kubicek, T. Chen, E. Li, T. Jenuwein & A.H. Peters, 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol. 13(14): 1192-200.

Loyola, A., T. Bonaldi, D. Roche, A. Imhof & G. Almouzni, 2006. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. Mol Cell 24(2): 309-16.

Maddison, D. & W. Maddison, 2005. MacClade: analysis of phylogeny and character evolution, Version 4.08. Sinauer associates, Sunderland, Massachusetts, USA.

Maison, C., D. Bailly, A.H. Peters, J.P. Quivy, D. Roche, A. Taddei, M. Lachner, T. Jenuwein & G. Almouzni, 2002. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. Nat Genet. 30(3): 329-34.

Melcher, M., M. Schmid, L. Aagaard, P. Selenko, G. Laible & T. Jenuwein, 2000. Structure-function analysis of SUV39H1 reveals a dominant role in heterochromatin organization, chromosome segregation, and mitotic progression. Mol Cell Biol. 20(10): 3728-41.

Mellone, B.G., L. Ball, N. Suka, M.R. Grunstein, J. F. Partridge & R.C. Allshire, 2003. Centromere silencing and function in fission yeast is governed by the amino terminus of histone H3. Curr Biol. 13(20): 1748 57.

Millar, C.B. & M. Grunstein, 2006. Genome-wide patterns of histone modifications in yeast. Nat Rev Mol Cell Biol. 7(9): 657-66.

Mis, J., S.S. Ner & T.A. Grigliatti, 2006. Identification of three histone methyltransferases in *Drosophila*: *dG9a* is a suppressor of PEV and is required for gene silencing. Mol Genet Genomics 275(6): 513-26.

Mosiolek, M., P. Pasierbek, J. Malarz, M. Mos & A.J. Joachimiak, 2005. *Rumex acetosa* Y chromosomes: constitutive or facultative heterochromatin? Folia Histochem Cytobiol. 43(3): 161-7.

Nakayama, J., J.C. Rice, B.D. Strahl, C.D. Allis & S.I. Grewal, 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. Science 292(5514):110-3.

Naumann, K., A. Fischer, I. Hofmann, V. Krauss, S. Phalke, K. Irmler, G. Hause, A.C. Aurich, R. Dorn, T. Jenuwein & G. Reuter, 2005. Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. EMBO J. 24(7): 1418-29.

Noma, K., C.D. Allis & S.I.S. Grewal, 2001. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science 293: 1150-5.

Pal-Bhadra, M., B.A. Leibovitch, S.G. Gandhi, M. Rao, U. Bhadra, J.A. Birchler & S.C. Elgin, 2004. Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. Science 303: 669-72.

Peng, J.C. & G.H. Karpen, 2007. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. Nat Cell Biol. 9(1): 25-35.

Peters, A.H., D. O'Carroll, H. Scherthan, K. Mechtler, S. Sauer, C. Schofer, K. Weipoltshammer, M. Pagani, M. Lachner, A. Kohlmaier, S. Opravil, M. Doyle, M. Sibilia & T. Jenuwein, 2001. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107(3): 323-37.

Peters, A.H., S. Kubicek, K. Mechtler, R.J. O'Sullivan, A.A. Derijck, L. Perez-Burgos, A. Kohlmaier, S. Opravil, M. Tachibana, Y. Shinkai, J.H. Martens & T. Jenuwein, 2003. Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell 12(6): 1577-89.

Ponger, L. & W.H. Li, 2005. Evolutionary diversification of DNA methyltransferases in eukaryotic genomes. Mol Biol Evol. 22(4): 1119-28.

Rai, K., L.D. Nadauld, S. Chidester, E.J. Manos, S.R. James, A.R. Karpf, B.R. Cairns & D.A. Jones, 2006. Zebra fish Dnmt1 and Suv39h1 regulate organ-specific terminal differentiation during development. Mol Cell Biol. 26(19): 7077-85.

Rea, S., F. Eisenhaber, D. O'Carroll, B.D. Strahl, Z.W. Sun, M. Schmid, S. Opravil, K. Mechtler, C.P. Ponting, C.D. Allis & T. Jenuwein, 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature 406: 593-9.

Regev, A., M.J. Lamb & E. Jablonka, 1998. The role of DNA methylation in invertebrates: developmental regulation or genome defense? Mol Biol Evol. 15(7): 880-91.

Reuben M. & R. Lin, 2002. Germline X chromosomes exhibit contrasting patterns of histone H3 methylation in *Caenorhabditis elegans*. Dev Biol. 245(1): 71-82.

Reuter, G., A. Fischer & I. Hofmann, 2005. Heterochromatin and the control of gene silencing in plants, pp. 106-133 in Plant Epigenetics (Annual Plant Reviews 19), edited by P. Meyer. Blackwell Publishing.

Rice, J.C., S.D. Briggs, B. Ueberheide, C.M. Barber, J. Shabanowitz, D.F. Hunt, Y. Shinkai & C.D. Allis, 2003. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol Cell 12(6): 1591-8.

Rudolph, T., M. Yonezawa, S. Lein, K. Heidrich, S. Kubicek, C. Schäfer, S. Phalke, M. Walther, A. Schmidt, T. Jenuwein & G. Reuter, 2007. Heterochromatin Formation in Drosophila Is Initiated through Active Removal of H3K4 Methylation by the LSD1 Homolog SU(VAR)3-3. Mol Cell. 26(1): 103-15.

Sanders, S.L., M. Portoso, J. Mata, J. Bahler, R.C. Allshire & T. Kouzarides, 2004. Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. Cell 119(5): 603-14.

Schmitt, E., S. Blanquet & Y. Mechulam, 2002. The large subunit of initiation factor aIF2 is a close structural homologue of elongation factors. EMBO J. 21(7): 1821-32.

Schotta, G., A. Ebert, V. Krauss, A. Fischer, J. Hoffmann, S. Rea, T. Jenuwein, R. Dorn & G. Reuter, 2002. Central role of *Drosophila* SU(VAR)3–9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. 21(5): 1121-31.

Schotta, G., M. Lachner, K. Sarma, A. Ebert, R. Sengupta, G. Reuter, D. Reinberg & T. Jenuwein, 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. Genes Dev. 18(11): 1251-62.

Schultz, D.C., K. Ayyanathan, D. Negorev, G.G. Maul & F.J. Rauscher, 2002. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. 16(8): 919 32.

Selker, E.U., N.A. Tountas, S.H. Cross, B.S. Margolin, J.G. Murphy, A.P. Bird & M. Freitag, 2003. The methylated component of the *Neurospora crassa* genome. Nature 422: 893-7.

Seum, C., E. Reo, H. Peng, F.J. Rauscher, P. Spierer & S. Bontron, 2007. Drosophila SETDB1 is required for chromosome 4 silencing. PLoS Genet. 3(5): e76.

Shen, W.H., 2001. NtSET1, a member of a newly identified subgroup of plant SET-domain-containing proteins, is chromatin-associated and its ectopic overexpression inhibits tobacco plant growth. Plant J. 28(4): 371-83.

Shi, J. & R.K. Dawe, 2006. Partitioning of the maize epigenome by the number of methyl groups on histone H3 lysines 9 and 27. Genetics 173(3): 1571-83.

Stabell, M., R. Eskeland, M. Bjorkmo, J. Larsson, R.B. Aalen, A. Imhof & A. Lambertsson, 2006. The *Drosophila G9a* gene encodes a multi-catalytic histone methyltransferase required for normal development. Nucleic Acids Res. 34(16): 4609-21.

Steenkamp, E.T., J. Wright & S.L. Baldauf, 2006. The protistan origins of animals and fungi. Mol Biol Evol. 23(1): 93-106.

Stewart, M.D., J. Li & J. Wong, 2005. Relationship between histone H3 lysine 9 methylation, transcription repression, and heterochromatin protein 1 recruitment. Mol Cell Biol. 25(7): 2525-38.

Strahl, B.D. & C.D. Allis, 2000. The language of covalent histone modifications. Nature 403: 41-5.

Swofford, D.L., 2002. PAUP: Phylogenetic Analysis Using Parsimony (and other methods), Version 4.0b10. Sinauer associates, Sunderland, Massachusetts, USA.

Tachibana, M., K. Sugimoto, T. Fukushima & Y. Shinkai, 2001. Set domain containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. J Biol Chem. 276(27): 25309-17.

Tachibana, M., J. Ueda, M. Fukuda, N. Takeda, T. Ohta, H. Iwanari, T. Sakihama, T. Kodama, T. Hamakubo & Y. Shinkai, 2005. Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. Genes Dev. 19(7): 815-26.

Tamaru, H. & E.U. Selker, 2001. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. Nature 414: 277-83.

Tamaru, H., X. Zhang, D. McMillen, P.B. Singh, J. Nakayama, S.I.S. Grewal, C.D. Allis, X. Cheng & E.U. Selker, 2003. Trimethylated lysine 9 of histone H3 is a mark for DNA methylation in *Neurospora crassa*. Nat Genet. 34(1): 75-9.

Taverna, S.D., R.S. Coyne & C.D. Allis. Methylation of histone h3 at lysine 9 targets programmed DNA elimination in *Tetrahymena*. Cell 110(6): 701-11.

Thorstensen, T., A. Fischer, S.V. Sandvik, S.S. Johnsen, P.E. Grini, G. Reuter & R.B. Aalen, 2006. The *Arabidopsis* SUVR4 protein is a nucleolar histone methyltransferase with preference for monomethylated H3K9. Nucleic Acids Res. 34(19): 5461-70.

Tripoulas, N., D. LaJeunesse, J. Gildea & A. Shearn, 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(2): 913-28.

Unoki, M., T. Nishidate & Y. Nakamura, 2004. ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. Oncogene 23(46): 7601-10.

Vakoc, C.R., S.A. Mandat, B.A. Olenchock & G.A. Blobel, 2005. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol Cell 19(3): 381-91.

Vakoc, C.R., M.M. Sachdeva, H. Wang & G.A. Blobel, 2006. Profile of Histone Lysine Methylation across Transcribed Mammalian Chromatin. Mol Cell Biol. 26(24): 9185-95.

Wang, H., W. An, R. Cao, L. Xia, H. Erdjument-Bromage, B. Chatton, P. Tempst, R.G. Roeder & Y. Zhang, 2003. mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. Mol Cell 12(2): 475-87.

Yu, Y., A. Dong & W.H. Shen, 2004. Molecular characterization of the tobacco SET domain protein NtSET1 unravels its role in histone methylation, chromatin binding, and segregation. Plant J. 40(5): 699 711.

Zhang, X., Z. Yang, S.I. Khan, J.R. Horton, H. Tamaru, E.U. Selker & X. Cheng, 2003. Structural basis for the product specificity of histone lysine methyltransferases. Mol Cell 12(1): 177-85.

Figure legends

Figure 1. Cladogram which represents, based on commonly supported phylogeny, the evolution of histone H3K9 methylation. Data were extracted from Cowell et al. (2002), Taverna, Coyne & Allis (2002), Houben et al. (2003), Peters et al. (2003), Tamaru et al. (2003), Yu, Dong & Shen (2004), Juranek et al. (2005), Mosiolek et al. (2005), Naumann et al. (2005), Reuter, Fischer & Hofmann (2005), Barzotti, Pelliccia & Rocchi (2006), Ebert et al. (2006), Fischer et al. (2006), Kaller, Euteneuer & Nellen (2006), Millar & Grunstein (2006), Shi & Dawe (2006) and Casas-Mollano et al. (2007). General trends identified in two clades are given.

Figure 2. Relative stability of heterochromatic marks during cell metabolism, proliferation and differentiation (horizontal) and evolution (vertical). Note the reverse relationship.

Figure 3. Phylogenetic distribution of H3K9 methyltransferases and HP1. The tree was adapted from Keeling et al. (2005) and Steenkamp, Wright & Baldauf (2006). Orthology was determined using NCBI genome and EST data via reciprocal best BLAST hits (Altschul et al., 1997). The times of emergence of the protein families were estimated by cladistic analysis (MacClade 4.0; Maddison & Maddison, 2005).

Figure 4. Functional motifs of heterochromatic H3K9 methyltransferases. Shown are locally restricted functions defined for *Drosophila* Su(var)3–9 (Schotta et al., 2002; Ebert et al., 2004; Eskeland et al., 2004), human SUV39H1 (Aagaard et al. 1999; Melcher et al., 2000;

Rea et al., 2000), *Schizosaccharomyces* Clr4p (Ivanova et al., 1998; Rea et al., 2000) and *Arabidopsis* SUVH2 (Naumann et al., 2005).

Figure 5. Phylogenetic mapping of unikont heterochromatic H3K9 methyltransferases and their evolutionary transitions. Shown is a maximum parsimony tree (PAUP 4.0b10; Swofford 2002) based on preSET, SET and postSET domains, constrained according to commonly supported organismal phylogeny. Branch lengths are proportional to evolutionary changes (steps). The tree contains 5796 steps, only 84 more than three most parsimonious trees, which contain several long branch artefacts (not shown). Branches of proteins which contain the chromo domain are light gray shadowed (three independent losses are white marked). The *Su(var)3–9/eIF2γ* gene fusion, occurring during the evolution of hexapods, is dark gray underlaid. Three independent losses are light gray marked.

Figure 6. 3´ alternative splicing as identified in the Su(var)3-9/eIF2γ gene fusion of insects (Krauss & Reuter, 2000). This gene structure expresses two mRNA variants (A: Su(var)3-9 and B: eIF2γ). The first exon is common to both mRNAs. The coding parts of the exons are shown in gray.

Tables

Table 1. H3K9 acitivities and subnuclear distributions of histone methyltransferases. Data were extracted from Carrington and Jones (1996), Tripoulas et al. (1996), Beisel et al. (2002), Czermin et al. (2002), Schotta et al. (2002), Schultz et al. (2002), Byrd and Shearn (2003), Kim, Geng & Huang (2003), Peters et al. (2003), Wang et al. (2003), Ebert et al. (2004), Eskeland et al. (2004), Kourmouli et al. (2005), Naumann et al. (2005), Ebbs and Bender (2006), Loyola et al. (2006), Stabell et al. (2006), Thorstensen et al. (2006) and Seum et al. (2007).

Enzyme	H3K9 methylation is						Subnuclear distribution
	Catalyzed *in vitro*			Facilitated *in vivo*			
	me1	me2	me3	me1	me2	me3	
Suv39/Clr4/dim5	x	x	x		x	x	mainly heterochromatic
SUVH	x	x		x	x		heterochromatic
G9a/GLP	x[a]	x[a]	x[a]	x	x		euchromatic, nucleoplasmatic
SETDB/ESET	x	x	x[b]	x	x	x[c]	nucleoplasmatic, euchromatic, weakly heterochromatic
SUVR		x[d]					nucleolar
RIZ1	x[e]						
E(z)	x	x	x				euchromatic[f]
Ash1	x[e]				x[g]		euchromatic

[a] on free histones, no activity on nucleosomal arrays (Stabell et al., 2006)

[b] in association with mAM (Wang et al., 2003)

[c] shown by artificial targeting in *Suv39h*- cells (Kourmouli et al., 2005)

[d] methylates only monomethylated H3K9 (Thorstensen et al., 2006)

[e] degree of methylation was not tested

[f] E(z)-dependent H3K27me was also found in heterochromatin (Ebert et al., 2004)

[g] according to Byrd and Shearn (2003) in heterochromatin

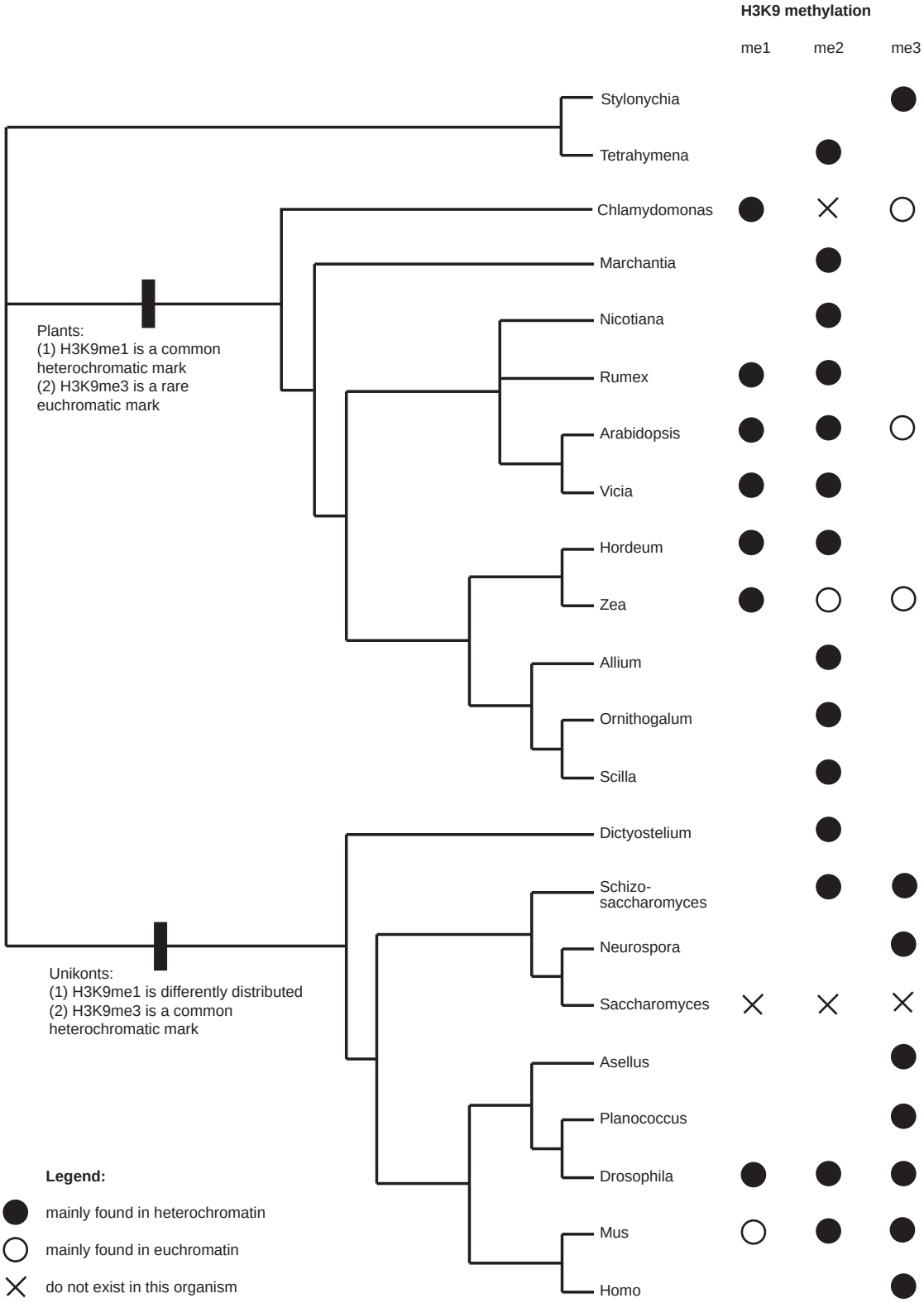

H3K9 methylation
me1
me2
me3
Stylonychia
Tetrahymena
Chlamydomonas
Marchantia
Nicotiana
Rumex
Arabidopsis
Vicia
Hordeum
Zea
Allium
Ornithogalum
Scilla
Dictyostelium
Schizo-
saccharomyces
Neurospora
Saccharomyces
Asellus
Planococcus
Drosophila
Mus
Homo
Plants:
(1) H3K9me1 is a common
heterochromatic mark
(2) H3K9me3 is a rare
euchromatic mark
Unikonts:
(1) H3K9me1 is differently distributed
(2) H3K9me3 is a common
heterochromatic mark
Legend:
mainly found in heterochromatin
mainly found in euchromatin
do not exist in this organism

Fig. 1

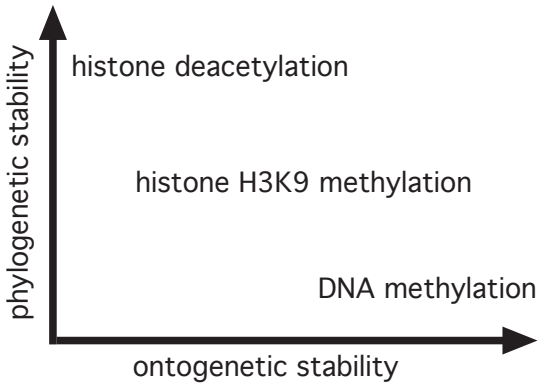

phylogenetic stability
histone deacetylation
histone H3K9 methylation
DNA methylation
ontogenetic stability

Fig. 2

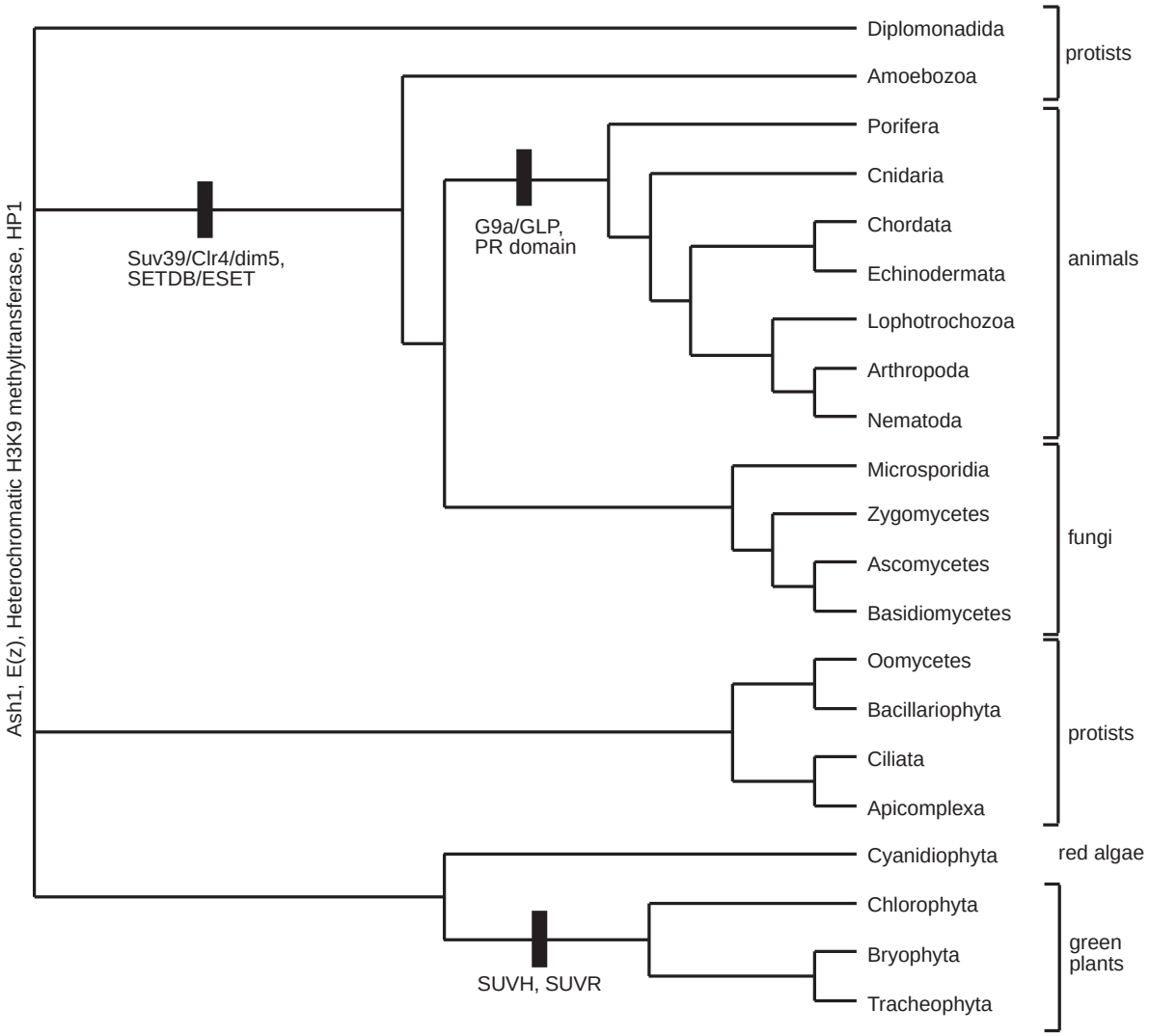
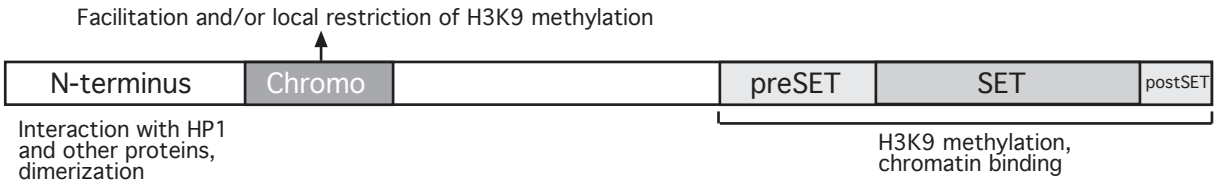
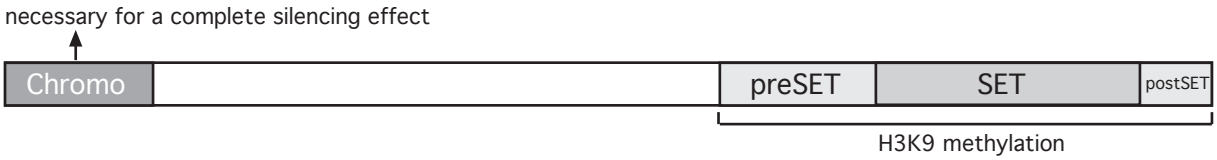

Ash1, E(z), Heterochromatic H3K9 methyltransferase, HP1
Suv39/Clr4/dim5,
SETDB/ESET
G9a/GLP,
PR domain
SUVH, SUVR
Diplomonadida
Amoebozoa
Porifera
Cnidaria
Chordata
Echinodermata
Lophotrochozoa
Arthropoda
Nematoda
Microsporidia
Zygomycetes
Ascomycetes
Basidiomycetes
Oomycetes
Bacillariophyta
Ciliata
Apicomplexa
Cyanidiophyta
Chlorophyta
Bryophyta
Tracheophyta
protists
animals
fungi
protists
red algae
green
plants

Fig. 3

Su(var)3-9, SUV39H1



Clr4p



SUVH2

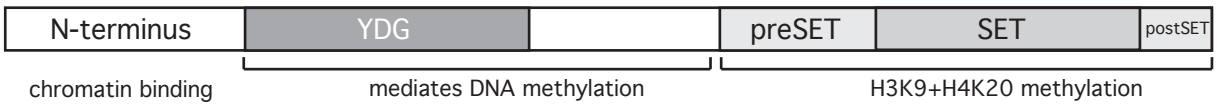


Fig. 4

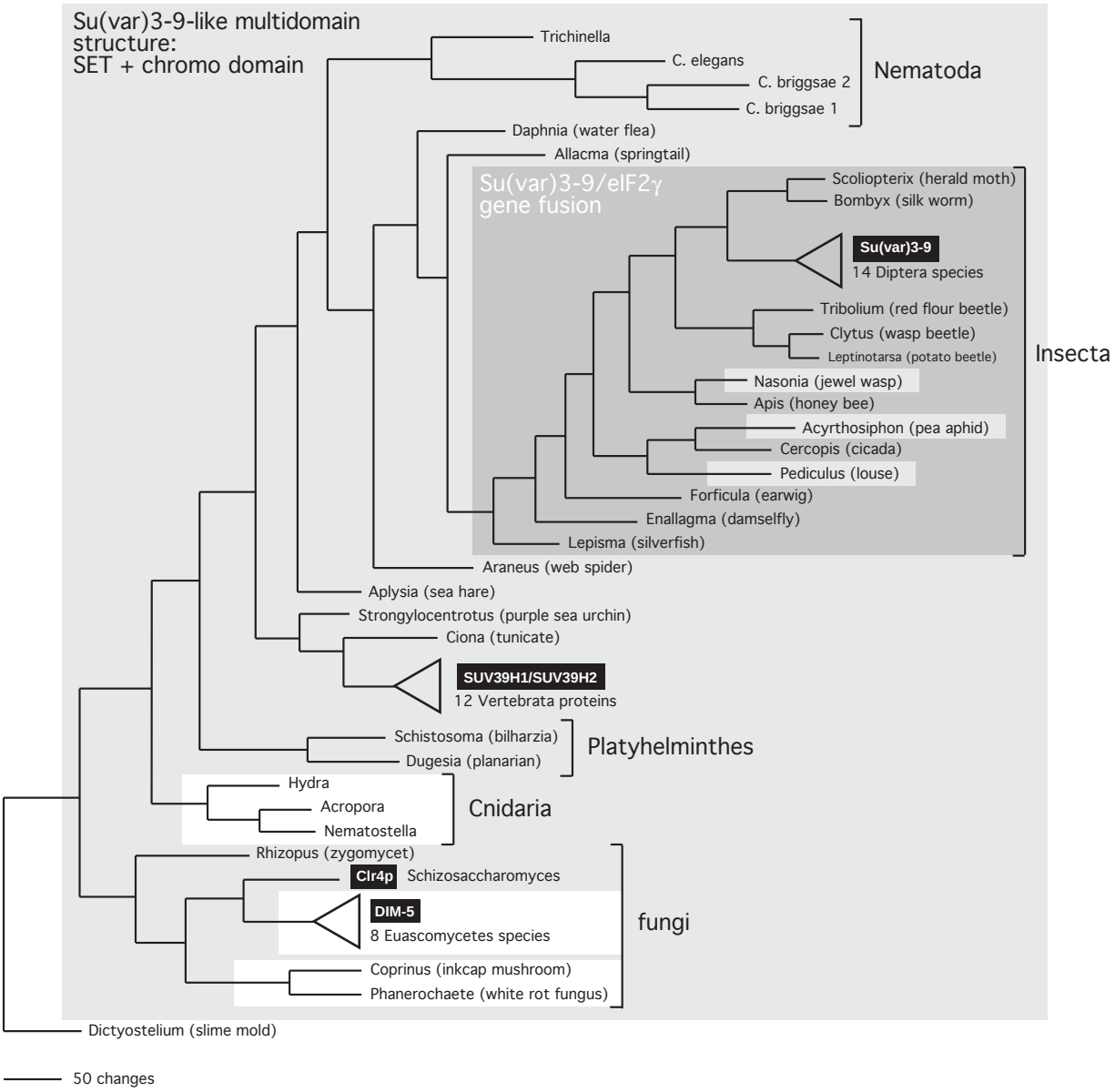

Su(var)3-9-like multidomain structure:
SET + chromo domain
Trichinella
C. elegans
C. briggsae 2
C. briggsae 1
Nematoda
Daphnia (water flea)
Allacma (springtail)
Su(var)3-9/eIF2γ gene fusion
Scoliopterix (herald moth)
Bombyx (silk worm)
Su(var)3-9
14 Diptera species
Tribolium (red flour beetle)
Clytus (wasp beetle)
Leptinotarsa (potato beetle)
Insecta
Nasonia (jewel wasp)
Apis (honey bee)
Acyrthosiphon (pea aphid)
Cercopis (cicada)
Pediculus (louse)
Forficula (earwig)
Enallagma (damselfly)
Lepisma (silverfish)
Araneus (web spider)
Aplysia (sea hare)
Strongylocentrotus (purple sea urchin)
Ciona (tunicate)
SUV39H1/SUV39H2
12 Vertebrata proteins
Schistosoma (bilharzia)
Dugesia (planarian)
Platyhelminthes
Hydra
Acropora
Nematostella
Cnidaria
Rhizopus (zygomycet)
Clr4p
Schizosaccharomyces
DIM-5
8 Euascomycetes species
fungi
Coprinus (inkcap mushroom)
Phanerochaete (white rot fungus)
Dictyostelium (slime mold)
50 changes

Fig. 5

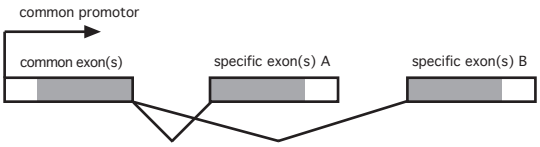

common promotor
common exon(s)
specific exon(s) A
specific exon(s) B

Fig. 6