Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal

Targeting and regulatory mechanisms in chromosome-wide gene regulation using *Drosophila melanogaster* as a model organism

Ana Margarida Lopes Adónis Figueiredo

Mestrado em Biologia Molecular e Genética

2010

Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal

Targeting and regulatory mechanisms in chromosome-wide gene regulation using *Drosophila melanogaster* as a model organism

Ana Margarida Lopes Adónis Figueiredo

Mestrado em Biologia Molecular e Genética

2010

Dissertacão orientada por:

Orientador externo: Prof. Dr. Jan Larsson, Department of Molecular Biology, Umeå University, Sweden

Orientador interno: Prof. Dr. Rui Gomes, Departamento de Biologia Vegetal, Faculdade de Ciências da Universidade de Lisboa

Resumo

O cromossoma 4 de *Drosophila melanogaster* é muito heterocromático, sendo rico em DNA satélite e DNA de elementos transposáveis. Apesar das suas propriedades heterocromáticas, o cromossoma 4 tem uma densidade génica e expressão semelhantes às dos principais braços cromossómicos. A expressão da maioria dos genes do cromossoma 4 é regulada por pelo menos duas diferentes proteinas: Painting of Fourth (POF) que estimula o output transcricional e Heterochromatin Protein 1 (HP1) que reprime o output transcricional. POF e HP1 ligam-se interdependentemente ao cromossoma 4 e a ligação e propagação de POF depende da heterocromatina. POF e HP1 colocalizam em exões de genes activos do cromossoma 4 e adicionalmente a HP1 liga-se aos promotores dos genes activos. HP1 é recrutada pela metilação de H3K9 a qual é realizada pelas HKMTs SU(VAR)3-9 e SETDB1. Experiências de imunocitoquímica demonstraram que SU(VAR)3-9 é responsável pela metilação de H3K9 principalmente na região centromérica enquanto que SETDB1 produz H3K9 no cromossoma 4, o que resulta no recrutamento de HP1 para essas regiões.

Com o objectivo de perceber o papel destas HKMTs no recrutamento de HP1, realizámos ChIP-on-chip e experiências de imunocitoquímica em mutantes *Su(var)3-9, Setdb1* e *Pof* e analisámos a ligação de HP1. Descobrimos que SETDB1 e POF são essenciais no recrutamento de HP1 para exões de genes activos do cromossoma 4 e que SU(VAR)3-9 é essencial para recrutar HP1 para as regiões centroméricas de todos os cromossomas. Adicionalmente, a ligação de HP1 a promotores em genes do cromossoma 4 assim como dos outros cromossomas é independente de SETDB1 e de SU(VAR)3-9. Propomos que a ligação de HP1 aos promotores ocorre de um modo independente da metilação de H3K9 ou que existe outra HKMT responsável pela produção de H3K9me especificamente nos promotores.

Palavras-chave: HP1, POF, H3K9me, SU(VAR)3-9, SETDB1

Abstract

The fourth chromosome of *Drosophila melanogaster* is largely heterochromatic being enriched in satellite repeats and in DNA from transposable elements. Despite its heterochromatic properties it has a gene density and expression similar to the major chromosome arms. The expression of almost all chromosome 4 genes is fine-tuned by at least two different proteins: Painting of Fourth (POF) that stimulates transcription output and Heterochromatin Protein 1 (HP1) that represses transcription output. POF and HP1 bind interdependently to the 4th chromosome and the binding and spreading of POF depends on heterochromatin. POF and HP1 colocalize in exons of active genes on the 4th chromosome. In addition, HP1 binds to promoters of the active genes. HP1 is recruited by H3K9 methylation which is performed by the HKMTs SU(VAR)3-9 and SETDB1. Immunostaining experiments have shown that SU(VAR)3-9 is mainly responsible for H3K9 methylation in the chromocentre whereas SETDB1 produces H3K9 methylation on the 4th chromosome and that this results in HP1 recruitment to these regions.

With the aim of understanding the role of these HKMTs in the recruitment of HP1, we have performed ChIP-on-chip and immunostaining experiments in *Su(var)3-9, Setdb1* and *Pof* mutants and analysed the binding of HP1. We have found that SETDB1 and POF are essential for the recruitment of HP1 to the exons of active genes of the 4th and that SU(VAR)3-9 is essential for the recruitment of HP1 to centromeric regions of all chromosomes. Furthermore, HP1 binding at promoters in the genes on the 4th chromosome and in other chromosomes is independent on SETDB1 and on SU(VAR)3-9. We propose that HP1 binding to promoters occurs in a H3K9me independent manner or that there is another HKMT that produces H3K9me specifically at promoters.

Key-words: HP1, POF, H3K9me, SU(VAR)3-9, SETDB1

Abbreviations

POF	Painting of Fourth
HP1	Heterochromatin Protein 1
HKMT	Histone Lysine Methyltransferase
H3K9me1,me2,me3	Histone 3 Lysine 9 mono-, di- and trimethylation
ChIP	Chromatin Immunoprecipitation
SU(VAR)3-9	Supressor of variegation 3-9
2R	Right arm of the chromosome 2
2L	Left arm of the chromosome 2
3R	Right arm of the chromosome 3
3L	Left arm of the chromosome 3
4^{th}	Chromosome 4
bp	Base pairs
kb	Kilobasepairs

Contents

Introduction

Epigenetics

Epigenetics can be defined as the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence [1]. Epigenetics and epigenetic changes are fundamental for our understanding of development and evolution [2].

One of the most important concepts in Epigenetics is that the DNA, molecule responsible for carrying genetic information, is the same in all somatic cells of an organism [2]. Nuclear transplantation experiments done by Laskey and Gurdon in 1970 showed for the first time that the DNA of a somatic cell nucleus was competent to direct embryogenesis when introduced into an enucleated egg [3]. The developmental programme, responsible for generating an adult organism from a fertilized egg, is not the result of alterations in the germ-line DNA sequence but rather alterations in the chromatin, also called epigenetic marks, which establish different patterns of gene expression from the same genome [2]. As all genetic information, epigenetic marks are also inheritable from mother to daughter cell and in some cases also from generation to generation [4].

Chromatin structure and remodeling

In eukaryotes DNA exists associated with proteins as a complex called chromatin [2]. The fundamental repeating unit of chromatin is the nucleosome, which is constituted by an octamer of histones (one H3/H4 tetramer and two sets of H2A/H2B dimers) around which 147 bp of DNA is wrapped [2]. The octamer of histones is stabilized by the linker histone H1 and repeated every 200 bp throughout the genome [2]. Histones are highly conserved basic proteins constituted by a globular domain and a flexible amino-terminal tail (N'-tail) which protrude from the surface of the nucleosome [2]. Chromatin is a dynamic complex which can exist in different states that range from a decondensed structure prone to transcription named euchromatin, to a transcriptionally inactive highly compact structure called heterochromatin [2]. Constitutive heterochromatin is found mainly at telomeres and centromeres, which contain satellite repeats, transposable elements and some single copy DNA, and is silenced in all cell types and

throughout the cell cycle [2]. Although this is the classical view, it has been shown that there are expressed genes located also within the constitutive heterochromatin [5]. Facultative heterochromatin exists for instance in tissue specific genes, developmental genes, and stress induced genes, and can be reverted to euchromatin [2]. Chromatin needs to be remodeled during the cell cycle so that the extended DNA molecules of the interphase can be highly compacted in a metaphase chromosome, which is relevant for chromosome pairing, recombination and segregation [2]. The less compact DNA fiber of 11nm "beads-on-a string" is compacted into a 30nm fiber, further organized into looped chromatin domains of 300-700 nm and finally into the 1.5 μm metaphase chromosome [2]. Chromatin can be remodeled not only at the entire chromosome level, but at specific DNA regions, for instance when the transcription of one particular gene needs to be activated/repressed [2].

There are covalent and non-covalent mechanisms involved in chromatin dynamics [2]. The former involves DNA methylation at cytosine residues and posttranslational histone modifications [2]. DNA methylation is performed by DNA methyltransferases (DNMT) and is present in constitutive heterochromatin and on CpG islands of genes where it mainly causes gene repression [2]. DNA methylation occurs in varying degrees in all eukaryotes except yeast and for instance in *Drosophila*, DNA methylation has largely been lost as an epigenetic mechanism [2]. Posttranslational N'-tail histone modifications as methylation, acetylation, phosphorylation, ADP ribosylation, biotinylation, ubiquitylation and sumoylation, have the capacity to alter the chemical properties of the chromatin which results in its compaction or decompaction, and can also be involved in the targeting of other proteins relevant to the remodeling process [2]. Histone phosphorylation for instance at H3Serine10, performed by JIL-1 kinase in *Drosophila*, has been shown to act on the maintenance of euchromatic domains while counteracting heterochromatization and gene silencing [6]. Histone acetylation, for instance on H4K16, is usually associated with gene expression and is performed by histone acetyltransferases (HAT) which are believed to be recruited by transcription activators, whereas repressors recruit histone deacetylases (HDAC) [2]. Histone methylation is performed by protein arginine methyltransferases (PRMT) and by histone lysine methyltransferases (HKMT), and removed by histone demethylases (HDM) [2]. HKMTs add mono, di and tri methyl (me1, me2, me3) groups to lysines mainly present on N'-tails of histones H3 and H4 [2]. Regarding which histone and N'-tail residue is modified, histone methylation can have different effects in chromatin. H3K9me and H3K27me have been associated with heterochromatin formation [2]. On the other hand, H3K4me3 and H3K36me have been linked to gene expression [2]. Other non-covalent mechanisms of chromatin remodeling involve the substitution of core histones by

histone variants (*e.g.* H3.3, H2AZ H2AX, macroH2A, CENP-A), ATP dependent chromatin remodeling complexes that mobilize (SNF2 / ISWI) or alter (Brahma / SWI/SNF) nucleosome structures or act at the whole chromosome level (cohesin and condensin complexes and topoisomerase II) and the targeting role of small non-coding RNAs (*e.g. Xist* RNA in inactivation of the human X chromosome) [2]. A factor that also seems to influence gene expression is the repositioning of the locus relative to nuclear compartments (nuclear lamina, nuclear pore, transcriptional and splicing factories) and other genomic loci (clustering of silenced or active genes in trans) [7].

Gene transcription is the result of the cooperation between *cis*-acting DNA sequences (promoters, enhancers, silencers, locus control regions, insulators) and bound trans-acting factors (transcription factors, (co)activators, (co)repressors, RNA polymerase and associated factors) which is only possible if the chromatin is in a decondensed, open state [2].

H3K9 methylation

H3K9 methylation is an epigenetic mark essential for heterochromatin formation and gene repression mainly because it recruits Heterochromatin Protein 1 (HP1) [2]. The H3K9 methylation is performed by the HKMTs SU(VAR)3-9, SETDB1 and dG9a in *Drosophila*, which are orthologous to the human SUV39H1, SETDB1 and EHMT2/G9a respectively [8, 9]. Su(var)3-9 and SETDB1 have nonoverlapping and independent functions, suggesting that they have distinct biological roles, although it´s believed that they act synergistically to accomplish the global H3K9 methylation which is mainly seen in the chromocentre, on the 4th chromosome and in some regions in euchromatic arms and telomeres [8, 11, 10].

Su(var)3-9 is a non-essential protein and contains a chromo domain, which allows its association with methylated N'-tails of histones and a SET domain, which has H3-K9 methyltransferase activity [8, 9]. Additionally, SU(VAR)3-9 association with heterochromatin is mediated by the direct interaction of its N'-tail with the chromo-shadow domain of HP1 and the association of both proteins to chromatin is interdependent [9]. This supports a model for heterochromatin spreading in which Su(var)3-9 is simultaneously recruited to chromatin by H3K9 methylation and by HP1, produces H3K9 methylation in adjacent nucleosomes which in turn recruits more HP1 [2]. SU(VAR)3-9 localizes mainly in the chromocentre where it is responsible for H3K9me2 and H3K9me3 [9, 8, 11]. It has been shown that SU(VAR)3-9 can be posttranslational phosphorylated by JIL-1 which renders it inactive [6].

SETDB1 is an essential protein composed of two Tudor domains (recognize dimethylated arginines), a methyl CpG binding domain (MBD), and a Pre-SET/SET domain. In addition, it has a histone deacetylase-interacting domain [8]. SETDB1 can thus be recruited by methylated arginines and methylated DNA, produce H3K9 methylation and recruit HP1 and HDAC, which collectively potentiate chromatin compactation and gene repression [2]. In somatic cells in *D. melanogaster* SETDB1 mainly associates to the 4th chromosome where it is responsible for H3K9me2, but also localizes to the euchromatic arms where it produces H3K9me1 and H3K9me2 marks [8]. SETDB1 seems to have important developmental roles, being localized in pericentric heterochromatin in germ cells, including oocytes, where it is the only HKMT that produce H3K9me3 [12]. In later oogenesis SU(VAR)3-9 replaces this function. The 4th chromosome is enriched in developmental genes which may support the specificity of SETDB1 to this chromosome [12].

HP1

HP1 is an essential 206 amino acids protein involved in heterochromatin formation and spreading [13]. HP1 consists of a N'-chromo domain separated from a C'-chromo shadow domain by a hinge region [13]. The C'-chromo shadow domain mediates protein-protein interactions and is thus responsible for both HP1 homodimerization and HP1 binding to other proteins [13]. HP1 seems to interact with proteins involved in heterochromatin formation, transcriptional regulation, telomere maintenance, DNA replication and repair, chromosome segregation and nuclear organization [13]. The N'-chromo domain is responsible to the binding to the N'-tail of histone H3 when it is di- and trimethylated on lysine 9 [13]. A second potential target for HP1 is H3K27me [14]. It has been suggested that one HP1 dimer could bind, by its two monomeric chromodomains, to Lys9 and Lys27 on a N'-tail of a single H3 or to Lys9 on N'-tails of two adjacent histone H3 proteins [14]. The binding to two neighbouring H3 would generate a crosslink between them and render them less mobile [13]. At the same time, HP1 recruits more SU(VAR)3-9 which will propagate the epigenetic mark. On the other hand, it has been suggested that HP1 can be bound to H3 histone-fold, in an independent manner of N'-tail methylation [15, 16]. It has been reported that the hinge region of HP1 interacts with H1 contributing to a more condensed state of chromatin, and directs HP1 to its correct chromatin target sites by interacting with RNA and H3K9me [16, 17]. HP1 localizes at the chromocentre, telomeres, chromosome 4 and approximately 200 euchromatic sites on the polytene

chromosomes [8]. Additionally to its role in heterochromatin formation and spreading, it has been proposed that HP1 can be involved in the expression of genes embedded in heterochromatin and of genes in euchromatin [4]. It has been shown that HP1 is associated with transcripts of more than 100 euchromatic genes and that it is necessary for the expression of these genes [18]. Furthermore, HP1 associates with heterogeneous nuclear ribonucleoproteins involved in RNA processing which have roles in heterochromatin formation [18]. An additional proposed function of HP1 is in movement of chromatin domains to the nuclear membrane vicinity by direct HP1 association with lamin proteins which may help direct the large scale repression of these domains [13]. HP1 function has a high degree of versatility, which is supported by the fact that many eukaryotes express two or more paralogs of HP1 [13]. In *Drosophila melanogaster* there are five HP1 paralogs: the originally discovered HP1A present in pericentric heterochromatin and on the 4th chromosome, which bind H3K9me; HP1C that localizes primarly to euchromatin, binds H3K9me but also colocalizes with histone modifications typically associated with transcriptionally active chromatin; HP1D which doesn´t bind H3K9me and is believed to repress mobile elements in the female germline; HP1B and HP1E which have not yet been characterized, but HP1E is hypothesized to have functions similar to HP1D, although in the male germline whereas HP1B has been shown to localize to both euchromatin and heterochromatin [13, 19].

Drosophila **as a model organism**

Drosophila melanogaster has been an important model organism especially in the fields of genetics and developmental biology. *Drosophila melanogaster* has a small genome of ~180Mb where 1/3 is heterochromatic (centromeric regions, the Y chromosome and the 4th chromosome). The genome has been entirely sequenced and there are a lot of available tools when it comes to its genetics. One of *Drosophila´*s main advantage is the giant polytene chromosomes which provide a visual map of the genome. The polytene chromosomes can be seen in different tissues of the 3rd instar larvae and the salivary glands are the best example since there the euchromatic arms have undergone 10 rounds of replication (endoreplication) without proceeding through mitosis, generating around 1000 copies of each arm and maintaining all the chromatin strands paired in a perfect synapsis [2]. The replication is not uniform, and for instance repetitive elements are less replicated and satellite repeats are not replicated at all [2]. This is why the heterochromatic regions of the 4th-chromosome and the

entire Y chromosome are not visible in polytene chromosomes. In a polytene chromosome microscope preparation it is possible to see five enlarged long chromosome arms (X, 2R, 2L, 3R, 3L) fused at the chromocentre made up of pericentric heterochromatin, and a short 4th chromosome arm emerging from the chromocentre. In the polytenized region it is possible to see bands that correspond to heterochromatin, and interbands that consist of decondensed chromatin where the genes reside. Relevant achievements to the epigenetics field made in *Drosophila* are the discovery of Polycomb and Trithorax systems which maintain the repressed and active expression patterns of Homeotic genes, the discovery of for instance *Su(var)2-5* (gene that encodes HP1) and *Su(var)3-9* as suppressors of Position-Effect-Variegation (PEV) and thus involved in heterochromatin formation/spreading, and the X chromosome dosage compensation system [2].

Regulation of gene expression is mainly considered to work at the single gene level (e.g. tissue specific genes) or on clusters of genes (e.g. ribosomal genes). However, regulation of gene expression also takes place at a chromosome-wide level. Chromosome-wide gene regulation has so far only been reported for the X chromosome dosage compensation system and on the 4th chromosome of *Drosophila melanogaster* [20]. X chromosome dosage compensation system equalizes the transcriptional output of the two X chromosomes in females with that of the single X chromosome in the males [20]. Strategies differ between organisms: in mammals, one X chromosome in females is completely inactivated and forms the Barr body, in *C. elegans* expression of both X chromosomes in the hermaphrodites (XX) are downregulated and in *Drosophila* the transcription of the male X chromosome is doubled [20]. Dosage compensation in *Drosophila* is achieved by the binding of two non-coding RNAs, *rox1* and *rox2*, together with five MSL (male specific lethal) proteins to the full extent of the male X chromosome [20]. The binding of the protein Painting of Fourth (POF) to the 4th chromosome of *Drosophila* is the second chromosome-wide gene regulatory system discovered and its study is relevant to understand whole chromosome targeting mechanisms [20].

Chromosome 4

The chromosome 4 of *Drosophila melanogaster* is an atypical chromosome mainly because it is only 5Mb in length, is highly heterochromatic having a higher A/T content compared to the other chromosomes, a 3x higher frequency of transposable elements compared to the entire euchromatic part of the genome, a 25% of overall repeat density,

whereas the euchromatic chromosome arms only have around 6% [21, 22]. The 4th chromosome doesn´t undergo meiotic recombination, and is late replicating in phase S, similar to heterochromatin [21]. Only 1.23 Mb of this chromosome is polytenized, and it is constituted by 92 genes [20]. The other 3-4 Mb is the pericentric heterochromatin which is enriched in satellite repeats [20]. Also the polytenized region has heterochromatic properties, being enriched in repetitive DNA from transposable elements interspersed with unique sequences and enriched in HP1 binding and H3K9me [23, 24, 21]. Although its heterochromatic features, the 4th has a gene density and average gene expression level comparable to the other chromosomes [25]. The 4th chromosome shares more similarities to the X chromosome than to the other autosomes, which raises the question if the 4th was part of the X in the past [23, 26, 24]. Both have a major left arm and a tiny right arm. Similar to the X, which is decorated by chromosome-specific MSL proteins, the 4th is decorated by the chromosome-specific protein POF. Haplo and triplo 4th flies are viable and fertile, which is true also for the X but not for the other autosomes. Exons on the 4th and on the X differ in sequence composition compared to the other chromosomes and in triplo-4th the number of X chromosome nondisjunction is increased in meiosis suggesting pairing between the 4th and the X chromsomes. In the *Drosophila* species the chromosome 4 correspondent is called F element [23, 26, 24]. Supporting the evolutionary relation between X and 4th is the fact that in the related species *Drosophila busckii* the F-element is fused with the X, and POF binds to the entire male X instead of MSL proteins [23]. In species evolutionary closer to *D. melanogaster*, the F-element appears as an independent autososome, but for instance in *D. ananassae* POF binds to both the male X and the F-element in both sexes and colocalizes with MSL on the male X [26]. The conservation of POF binding to the F-element through evolution suggests that POF association has a selective advantage *in vivo* [26].

POF and chromosome 4 gene regulation

The *Pof* gene is localized on the right arm of the 2nd chromosome and although it is not essential for survival of diplo-4 flies, it is essential for the survival of haplo-4th individuals [23, 20]. This suggests that *Pof* is involved in dosage compensation like mechanism of chromosome 4 genes [20, 27]. POF is a 495 aminoacids nuclear protein that has a RNA recognition motif (RRM) [23]. POF binds specifically to the 4th chromosome in both males and females, although it´s predominant in males (20-30% higher) [23]. So far POF has been shown to bind to the 4th in

Drosophila salivary glands, ovary nurse, follicle, S2 cells and is predominant in testes [26]. POF binding initiates near the centromeric region and spreads in *cis* to the distal part of the chromosome [23]. The evidence that POF doesn´t spread in *cis* to another chromosome arm translocated to the distal part of the 4th supports the view that POF binding depends on sequence signatures specific of the 4th chromosome [23]. Additionally, POF binding can spread in *trans* between two paired 4th chromosomal regions [23]. The binding of POF depends on the amount and compactation of nearby heterochromatin [20]. Immunostaining experiments have shown that POF binds within interbands of the 4th polytene chromosome and that it colocalizes with HP1 binding [20]. Expression microarrays experiments have shown that POF and HP1 binding to the 4th correlates with transcription [20, 28]. However, HP1 knockdown causes increase in 4th chromosomal gene expression whereas *Pof* mutants have decreased gene expression from the 4th, suggesting that POF stimulates and HP1 represses transcription output [20]. POF and HP1 binding to the 4th is interdependent [20, 28]. Immunostainings have shown that in *Pof* mutants, H3K9me2 and HP1 are reduced on the 4th [11]. It has been hypothesized that POF and HP1 are part of balancing mechanism that fine-tunes gene expression output from the 4th [20, 28]. Chromatin Immunoprecipitation followed by tiling arrays (ChIP-on-chip) experiments have confirmed that POF and HP1 bind to active genes on the 4th, and have shown that HP1 binds additionally to the promoters of these genes [28]. It has been shown that POF interacts *in vivo* with SETDB1 and that in *Setdb1* mutants POF localization to the 4th is decreased and 73% of chromosome 4th genes show decreased expression [11]. It´s believed that SETDB1 functions cooperatively with HP1 and POF for the epigenetic regulation of the 4th chromosome of Drosophila [11]. ChIP-on-chip experiments have also shown that POF has a bias towards the 3´terminal of the genes and towards exons [28]. It is not known yet if POF requires RNA for its stability/complex formation or if POF associates directly to chromatin, since in the ChIP-on-chip technique POF is immunoprecipitated with everything that is crosslinked to it (RNAs, DNA. proteins) [29].

POF increases the amount of transcripts from the 4th which could be explained by a potential role in RNA splicing, protection or export [29]. RIP-chip experiments have shown that POF associates with nascent RNA transcripts from the 4th during transcription elongation when nascent RNAs are still attached to the RNA PolII and facilitates the export of the RNAs to the nuclear pore which has been proposed to be mediated by a relocalisation of the chromatin domain closer to a nuclear pore [29].

Aims

The main goal of this thesis is to increase the understanding of the role of the HKMTs SETDB1 and SU(VAR)3-9 in the recruitment of HP1 to the 4th chromosome of *Drosophila melanogaster*. Since HP1 localizes in promoters and exons of active genes on the 4th whereas POF only binds to exons, and since SETDB1 seems to be a 4th chromosome specific HKMT, we hypothesized that SETDB1 could be mainly responsible for H3K9me on the exons of the 4th and SU(VAR)3-9 would act on the promoters of the 4th.

Since chromosome 4 is a good model to study chromosome-wide gene regulation, we hope that the results from this work will help to elucidate in a general way how HP1 chromosome targeting functions.

In the present work we aim to answer these questions:

- Is SETDB1 responsible for HP1 recruitment to exons of active genes on the 4th?
- Is SU(VAR)3-9 responsible for HP1 recruitment to promoters of active genes on the 4th?

Material and Methods

Fly stocks and genetic crosses

In order to obtain third instar larvae, 6 adult females and 6 adult males were crossed in each vial, containing mashed potato and yeast in agar medium as food and maintained at 21°C. There were 20 vials for each genotype: wildtype (Oregon-R), *Pof* $^{D119/D119}$, $\frac{\text{Su(var)3-9}^{evo}}{\text{Su(var)3-9}^{06}}$ (flycross: $\frac{\text{Su(var)3-9}^{evo}}{\text{Su(var)3-9}^{evo}}$ x $\frac{\text{Su(var)3-9}^{06}}{\text{TM6, Tb}}$) , $\frac{\text{Setdb1}^{10.1\text{-}1aex}}{\text{Setdb1}^{10.1\text{-}1aex}}$ (flycross: $\frac{\text{Setdb1}^{10.1\text{-}1aex}}{\text{CyO,GFP}}$ x . $\frac{\text{Setdb1}^{10.1\text{-}1aex}}{\text{CyO,GFP}}$). The Su*(var)3-9^{evo} / Su(var)3-9^{06}* and the *Setdb1$^{10.1\text{-}1aex}$* were selected at 1st instar larvae as non-GFP larvae. The flies were changed to new vials every 2 days until the required amount of larvae was reached.

Sallivary gland dissection, fixation and sonication

Third instar larvae were dissected in 1 × PBS and 125 salivary gland pairs harvested, fixed for 2 minutes in freshly made 2% formaldehyde in 1x PBST (0.1% Triton X-100, 1 x PBS) and stored in 1 ml of 1 x PBS until sonication. To stop the fixation, glycine was added to a final concentration of 0.125M. The salivary glands were then washed one time in 1 ml of 1mM PMSF, 1 x Protease inhibitor cocktail (Complete, EDTA-free - Roche) in 1 x PBS and 2 times in 1 ml of 1mM PMSF, 1x Protease inhibitor cocktail (Complete, EDTA-free - Roche) in 1x TBS. In order to lyse the salivary gland cells, 250 µl of SDS Lysis Buffer (1% SDS, 10mM EDTA, 50mM Tris, pH 8.1) was added followed by homogenizing the glands using a *STIR 20* (InterMed, *Labassco*). To shear the chromatin, 60µl of glass beads (SIGMA) were added to the lysate which was then sonicated by 4 cycles of 30 seconds at level 5 followed by 30 seconds of resting, using a *Bioruptor* (Diagenode) at 4°C. The sonicated samples were centrifuged for 10 minutes at 13000rpm, 4°C and the supernatants were recovered. We obtained around 250µl of sonicated sample for each 125 salivary glands dissected. The sonicated samples were kept at -80°C until the ChIP procedure.

Chromatin Immunoprecipitation (ChIP)

We mixed two vials of sonicated sample from the same genotype in order to have 500µl of final volume of sonicated sample, corresponding to 250 salivary glands. From this volume, 150µl were used for the ChIP DNA (Immunoprecipitated DNA) and another 150µl for the input DNA (non-immunoprecipitated DNA). In this way both the ChIP DNA and the input DNA samples had the same starting material. The protocol was done in parallel for both the ChIP DNA and the input DNA samples.

To reduce non-specific background, 150µl of sonicated sample corresponding to 75 salivary glands pairs was diluted 10 times in ChIP Dilution Buffer (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM Tris-HCl, pH8.1, 167mM NaCl) containing 1x Protease inhibitor cocktail, added to 60µl of magnetic beads conjugated with Protein A (Dynabeads Protein A, Invitrogen) and incubated during 30 minutes at 4°C with rotation. The magnetic beads were previously equilibrated in 150µl of ChIP Dilution Buffer for 10 minutes at room temperature with rotation. The supernatant was recovered and the pellet discarded using the Magnetic Particle Concentrator (Dynal MPC). These steps were done for both the ChIP DNA and the input DNA samples. After this, the precleared input DNA sonicated sample remained at 4°C with rotation until the reversal of the crosslinking.

The immunoprecipitating antibody was added to the precleared ChIP DNA sonicated sample and it was incubated overnight at 4°C with rotation. The following antibodies were used: 3µl of anti-HP1 (291C, Covance), 4µl of anti-POF (285N-7.0), 5µl of anti-SU(VAR)3-9 (rabbit) and 4µl of anti-HA ab9110. In the day after, in order to precipitate the antibody complexes, the samples were added to previously equilibrated magnetic beads and incubated for 1h at 4°C with rotation. The supernatants were discarded and the immune complex [chromatin-antibody-protein A magnetic beads] was washed for 3-5 minutes at 4°C with rotation, in 1ml of the following immune complex washing buffers: one time in Low Salt Buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, pH 8.1, 150mM NaCl); one time in High Salt Buffer (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris-HCl, pH 8.1, 500mM NaCl); one time in LiCl Buffer (0.25M LiCl, 1% IGEPAL-CA630, 1% deoxycholic acid, 1mM EDTA, 10mM Tris, pH 8.1) and two times in TE Buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). To elute the chromatin from the complex antibody-protein A magnetic beads, 250µl of freshly made Elution Buffer (1% SDS,

0.1M NaHCO3) was added to the beads and incubated for 15min at room temperature with rotation. The elution was repeated one more time and the supernatant fractions were recovered and combined.

To reverse the cross-linking and recover protein free DNA, NaCl to a final concentration of 200mM was added to the 500µl eluate ChIP DNA sample and to the 250µl DNA input sample and the samples were incubated at 65°C for 4h. The samples were stored at -20°C overnight. On the day after, the samples were incubated with 10µl of 0.5M EDTA, 20µl of 1M Tris-HCl pH 6.5 and 1µl of 20mg/ml proteinase K for 1h at 45°C. The DNA was then recovered by phenol/chloroform extraction and ethanol precipitation: two times one volume of phenol/chloroform followed by agitation, 3 minutes of centrifugation at 13000rpm at room temperature and recovery of the upper phase; one time one volume of chloroform followed by agitation, 3 minutes of centrifugation at 13000rpm at room temperature and recovery of the upper phase; one time three volumes of ethanol 99,5% and 1.5µl of glycogen followed by agitation and 1h of incubation at -80°C; 15 minutes of centrifugation at 13000rpm at 4°C; removal of the supernatant and washing the pellet two times in 1ml of cold 75% ethanol followed by 5 minutes of centrifugation at 13000rpm at 4°C; air drying the pellet followed by resuspension in 24µl of MQ H_2O.

After spectrophotometric measurement of DNA concentration (NanoDrop 1000, Thermo Scientific), the samples were stored at -20°C. In order to proceed with the ChIP-on-chip, the quality of the DNA samples obtained after ChIP was analysed by real-time PCR, the DNA amplified and further analyzed in tiling arrays.

Real-Time PCR

To check if the ratio profiles of ChIP DNA:input DNA before and after amplification were the same, to check the quality of the DNA before and after amplification and to check if the expected enriched regions of the genome for the immunoprecipitated protein were enriched on the immunoprecipitated DNA pool, real-time PCR was performed. For negative controls, primer pairs to amplify a region inside the *Spt-I* gene (chromosome 2R) or the *rox1* gene (chromosome X) were chosen. For the promoter and exon on chromosome 4, primer pairs to amplify the regions 226kb and 220kb respectively, of the gene *CG31998* were chosen. For each reaction, 12.5µl of IQ SYBR Green Supermix (Bio Rad) was added to 8.5µl of mQH_2O, 1µl of primer

sense, 1µl of primer antisense and 2µl of DNA. The real-time PCR conditions were: cycle 1 (95°C for 3min), 35x cycle 2 (94°C for 15sec, 58°C for 30sec, 72°C for 20sec), cycle 3 (95°C for 1min), cycle 4 (55°C for 1min) and 80x cycle 5 (55°C for 10sec). To measure the amounts of the specific DNA fragments, its threshold cycle values were compared to a five-point standard curve of input DNA which concentrations ranged from 0.02 to 1.92%. The specificity of the primers was assessed by melting curve analysis.

Primer pairs sequence:

CG31998 gene:

226kb (promoter)
Sense: 5´-TATGACGACGCAAATGACGGA-3´
Antisense: 5´-TGAATGGGTGGGATTATGACA-3´
220kb (exon)
Sense: 5´-GGCTGCCAATAGTACAAAC-3´
Antisense: 5´-ATTCTTCGCAGCAAATCTCAAA-3´

roX1 gene:

Sense: 5´-TACCGCTCTCTTTCGGGACTTG-3´
Antisense: 5´-TCCATCACTCTCTATCGGGCTG-3´

CG4016 gene:

Sense: 5´-TTGTTCAACGAAATCGGCAGTA-3´
Antisense: 5´-TCACGGTTATGAGCAGTAGGGT-3´

DNA Amplification and Purification

In order to proceed with tiling Arrays after the ChIP protocol, it was necessary to have 7-10µg of DNA. The GenomePlex Complete Whole Genome Amplification (WGA) Kit (SIGMA) was used to amplify 50ng of the ChIP DNA / input DNA. Since the unamplified DNA had been sonicated, the Fragmentation step was skipped although 1µl of 10x Fragmentation Buffer was added to the DNA to a final volume of 10µl before starting the Library Preparation. The amplified

DNA was purified using QIAquick PCR Purification Kit (Quiagen) and the final concentration measured spectrophotometrically (NanoDrop 1000, Thermo Scientific).

Tiling Arrays Analysis

For tiling array analysis, the amplified DNA samples were fragmented, labeled and hybridized to an Affymetrix *Drosophila* Genome 2.0 array (B.E.A, NOVUM, Karolinska Institutet). The signal intensity data was analyzed with the Affymetrix Tiling Analysis Software (v.1.1.02) using 250bp bandwidth as parameter for smoothing: The smoothing is done by a sliding window comprising approximately 10 probes (each probe is 25bp) that scans the genome and calculates the median score of all probes within de 250bp and displays it as the value for the middle probe. As a control we used the input DNA. The protein binding profiles obtained are visualized in Integrated Genome Browser (IGB 6.2.2) software. Each peak of POF or HP1 enrichment represents $\frac{\text{ChIP DNA}}{\text{input DNA}}$ in log2 scale.

Immunostaining of polytene chromosomes

Third instar larvae were dissected in 1 x PBST and the salivary gland pairs were harvested and incubated in fixation solution 1 (3.7% formaldehyde in 1 x PBST) during 10-60 seconds. The salivary gland pairs were then incubated in a drop of fixation solution 2 (50% acetic acid, 1% formaldehyde in MQ H_2O) for 2-4min. The tissue was squashed, the quality of the polytene chromosomes assessed in a light microscope (Carl Zeiss, 64038, Germany) and frozen on dry ice before removal of the coverslip and long-term storage in ethanol 99.5% at -20°C. The polytene chromosome preparations were incubated for 30 minutes in blocking solution [1 x blocking reagent (Roche), 1 x washing solution without Tween 20 (10mM maleic acid, 9mM NaCl, 8mM NaOH in MQ H_2O)] at room temperature with agitation. The chromosome preparations were then incubated with 23µl of the primary antibodies mixture [anti-HP1 (291C, Covance, 1:400), anti-POF (67µ7.0, 1:75), anti-SU(VAR)3-9 (1:300)] diluted in the blocking solution overnight in a wet chamber at 4°C. The preparations were washed 2 times for 10min in washing solution (10mM maleic acid, 9mM NaCl, 8mM NaOH, 0.3% Tween 20 in MQ H_2O) and then incubated for 30 minutes in blocking solution. The chromosome preparations were

incubated with 23µl of the secondary antibodies mixture [Goat-α-IgY – Alexa Fluor 488 (1:300), goat-α-rabbit – Alexa Fluor 555 (1:300), DAPI (1µg/ml)] diluted in the blocking solution for 2h in a wet chamber at room temperature. The slides were washed two times for 10 minutes in washing solution and then mounted in 8.5µl of Vectashield (VECTOR). The immunostainings of the polytene chromosomes were analysed in the UV-light microscope (Zeiss Axiophot microscope equipped with a KAPPA DX20C CCD camera). Immunostainings of wildtype chromosomes and mutant chromosomes were done in parallel and the nuclei with good cytology were chosen based on DAPI staining and photographed. All settings were identical for each specific antibody. Images were assembled and merged using Adobe Photoshop.

Results

HP1 and POF binding is not completely overlapping

To verify the genome-wide distribution of POF and HP1 we performed an immunostaining of polytene chromosomes. In **figure 1** is shown a double immunostaining of a single nucleus of wt salivary glands. HP1 binds strongly to both the chromocentre, region where centromeres of all the polytene chromosomes meet, and to the 4th chromosome while POF specifically decorates the 4th chromosome. The staining also shows that HP1 and POF colocalize on the 4th chromosome at this resolution.

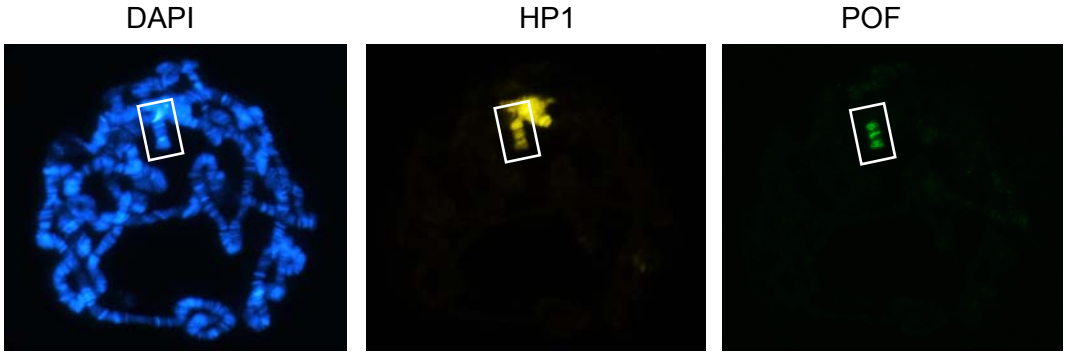


Figure 1: HP1 localizes in the chromocentre and on the 4th chromosome.
HP1 and POF immunostainings on wt polytene chromosomes from third instar larvae salivary glands. DNA is stained with DAPI (blue), HP1 with an anti-HP1 antibody (yellow) and POF with an anti-POF antibody (green). HP1 decorates the chromocentre and the 4th chromosome whereas POF just localize on the 4th. The 4th chromosome is indicated with boxes.

The resolution of immunostainings of polytene chromosomes is considered to be in the range 20-50kb. To map POF and HP1 at a higher resolution to determine the binding at individual gene level we performed ChIP with antibodies against POF and HP1 on wt salivary glands from third instar larvae to check if the binding profile was consistent to what has been previously reported [28]. To confirm if the preferentially bound regions for POF and HP1 were enriched on the immunoprecipitated DNA pool obtained after ChIP and to check the quality of the DNA, we performed real-time PCR. To assess the enrichment of certain DNA sequences after ChIP, in real-time PCR, we used input DNA to produce the standard curve. The input DNA is a non-immunoprecipitated DNA which consists in the same amount of sonicated salivary glands as the ChIP DNA. The ChIP protocol was done in parallel for the ChIP DNA sample and

the corresponding input DNA sample. The enrichment of ChIP DNA is thus calculated as a percentage of input DNA.

We used two set of primers to amplify one exon (220kb region) and the promoter (226kb region) of the gene *CG31998* which is present on the 4th chromosome and which is known to bind POF and HP1 [20]. As a negative control, we used primers to amplify the gene *roX1* which is present on the X chromosome and is not a target for HP1 or POF. As shown in **figure 2**, after the ChIP we had between 5 and 6 fold enrichment on DNA from the 4th gene *CG31998*, compared to the negative control *roX1* gene. We also confirmed that the HP1 immunoprecipitated DNA pool was enriched in sequences of both the exon and the promoter of *CG31998* (**figure 2A**) whereas the POF immunoprecipitated DNA pool was enriched in sequences of only the exon of this 4th chromosome gene (**figure 2B**).

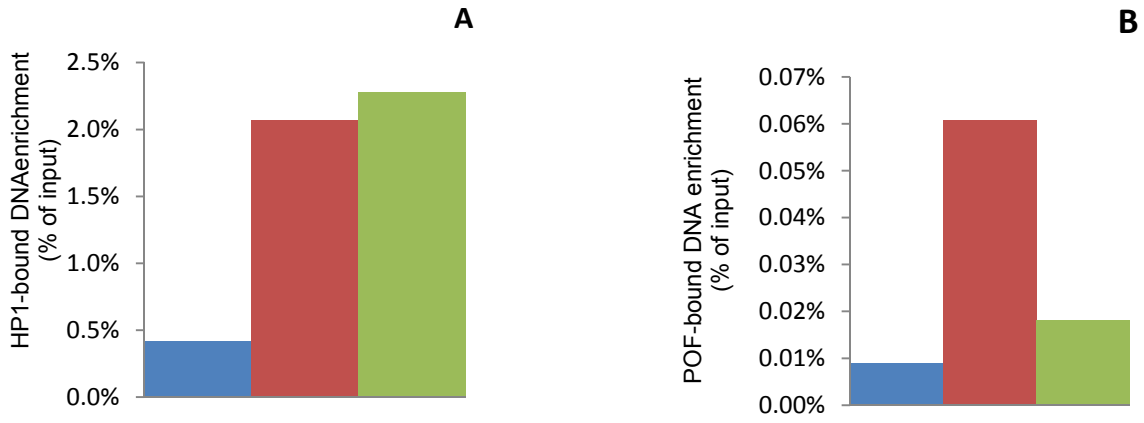


Figure 2: HP1 and POF are differently enriched on exons and promoters on the 4th chromosome.
HP1 and POF enrichment determined by qPCR analysis of ChIP experiments. HP1 (A) and POF (B) enrichment on the negative control X chromosome gene *roX1* (blue) and in one exon (red) and in the promoter (green) of the 4th chromosome gene *CG31998*, measured by real-time PCR after ChIP. (A) HP1 is enriched in both the exon and the promoter of the *CG31998* gene. (B) POF is enriched only within the exon of the *CG31998* gene.

We amplified the immunoprecipitated DNA, in order to have 7-10μg to proceed with the affymetrix tiling array analysis, and analyzed the quality of the amplified DNA by real-time PCR. After checking the melting curves and threshold cycles we have confirmed that the ratio $\frac{\text{ChIP DNA}}{\text{input DNA}}$ before and after amplification was the same. Since the ChIP protocol and the amplification step seemed to be robust, the samples were further analyzed in tiling arrays.

The ChIP-chip tiling array data for HP1 and POF binding to the genome of *Drosophila melanogaster* were analysed using the Tiling Analysis Software (TAS) from Affymetrix which allows the signal of each 25 bp probe to be matched to the right position in the genome. Every

ChIP sample was analysed against the corresponding input DNA in order to get enrichment profiles that correspond to $\frac{\text{ChIP DNA}}{\text{input DNA}}$. We used the Integrated Genome Browser (IGB, Affymetrix) to visualize the binding profiles of the proteins to the DNA. Each peak corresponds to the mean enrichment of 10 probes sitting next to each other on the genome that is represented in the middle probe. This 250bp window slides from one probe to the next in an overlapping manner and generate mean peaks that together form the binding profiles seen in **figure 3**.

In the results shown in **figure 3** we have confirmed that POF and HP1 colocalize on exons of active genes from the 4th chromosome and that HP1 binds additionally to promoters of these genes [28]. We have chosen the *Ank* gene as an example and the binding profile of POF and HP1 to it is shown in **figure 3**. The profile is similar to almost all active genes on the 4th chromosome. In order to see if the epigenetic marks responsible for HP1 recruitment, H3K9me2 and H3K9me3, colocalized with HP1, we have downloaded the ChIP-on-chip data from ModEncode for the H3K9me3 and H3K9me2 in whole third instar larvae, and analyzed it together with our results for POF and HP1. In this comparison the HP1 profile correlates with both H3K9me2 and H3K9me3 in the gene body. However, in the promoter peak of the *Ank* gene, HP1 correlates with H3K9me3 but not with H3K9me2.

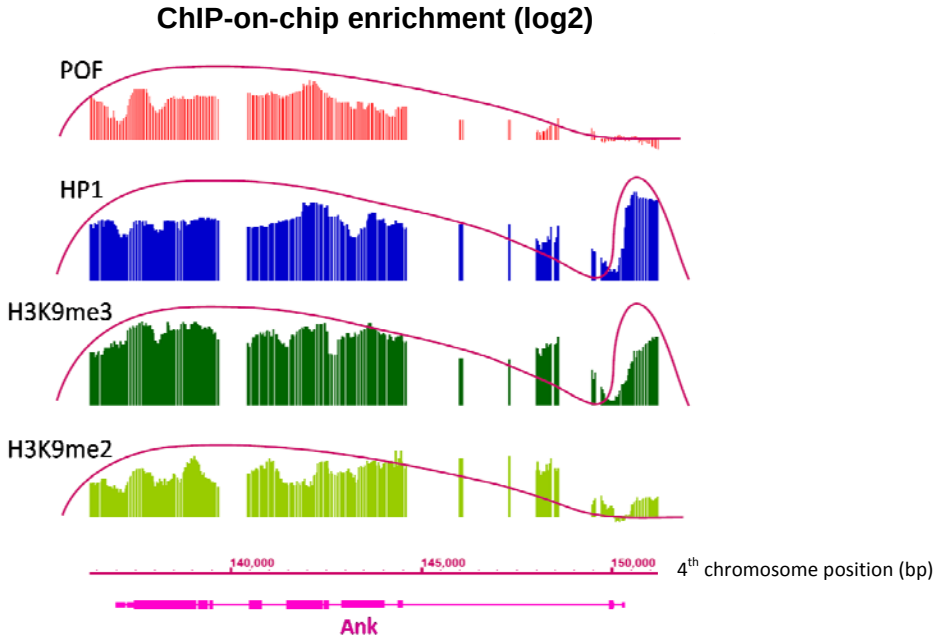


Figure 3: POF and HP1 binds within genes and HP1 binds additionally to promoters.
ChIP-on-chip binding profiles of POF and HP1 on one gene of the 4th chromosome, obtained from third instar larvae salivary glands. For comparison the downloaded ModEncode data for H3K9me2 and H3K9me3 on whole third instar larvae are shown. POF only binds within the *Ank* gene whereas HP1 in addition binds to the promoter of the *Ank* gene (peak on the right). HP1 seems to correlate better with H3K9me3 than with H3K9me2. The line in pink is a schematic view of the typical binding profile to 4th chromosome active genes, based on previous analysis of mean enrichments [28]. Note that the example *Ank* gene is transcribed from the right to the left.

SU(VAR)3-9 is required for the recruitment of HP1 to the chromocentre

Since H3K9 methylation in *Drosophila* is performed by the two proteins SETDB1 and SU(VAR)3-9, we decided to check on immunostainings of polytene chromosomes what happened to HP1 and POF localization in the absence of these HKMTs. The results from these immunostaings are shown in **figure 4.** In *Su(var)3-9* mutants HP1 decreases in the chromocentre region. On the contrary, in *Setdb1* mutants both HP1 and POF bindings decrease on the 4th chromosome. Additionally, we found that POF remains bound to the 4th chromosome in *Su(var)3-9* mutants and that the binding is slightly increased in comparison to wt, which seems also to be true for HP1.

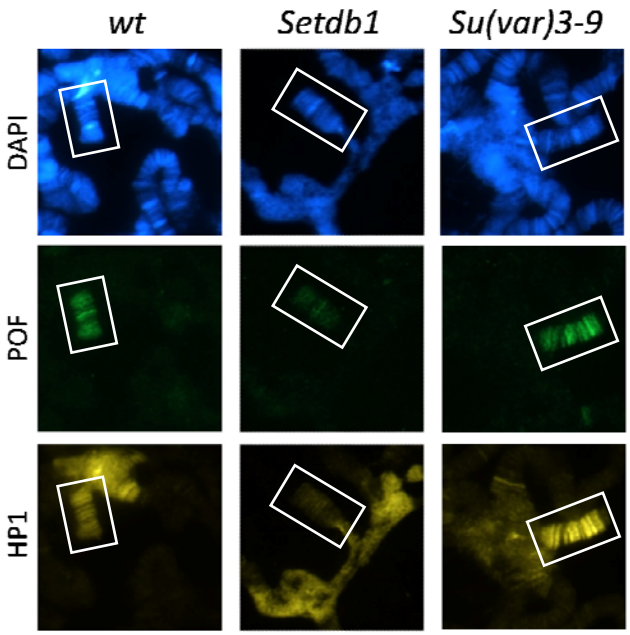


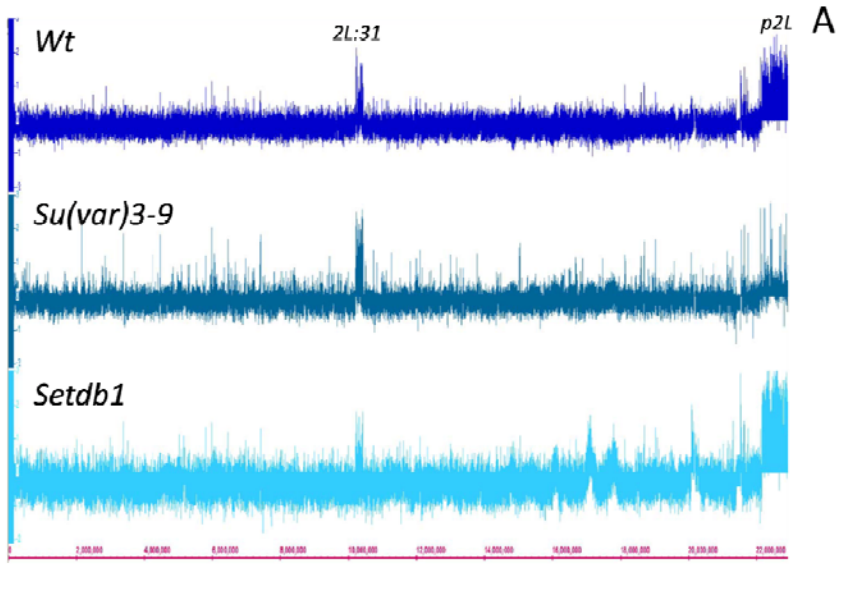
Figure 4: Setdb1 and Su(var)3-9 affect HP1 and POF localization on chromosomes.
HP1 and POF immunostainings in wt, *Setdb1* mutants and *Su(var)3-9* mutants polytene chromosomes from third instar larvae salivary glands. DNA is stained with DAPI (blue), HP1 with an anti-HP1 antibody (yellow) and POF with an anti-POF antibody (green). In *Setdb1* mutants HP1 and POF signals are decreased on the 4th chromosome. In *Su(var)3-9* mutants HP1 signal is decreased in the chomocentre, whereas on the 4th chromosome both HP1 and POF seem to be slightly increased. The 4th chromosome is indicated with boxes.

Since SETDB1 has been proposed to mediate H3K9me2 on the 4th chromosome and to mediate both H3K9me2 and H3K9me3 on chromosome arms, whereas SU(VAR)3-9 mediates H3K9me2 and H3K9me3 in the chromocentre [8, 9, 11], we hypothesized that on the 4th chromosome, SETDB1 would act on exons, producing H3K9me2, and SU(VAR)3-9 would recruit HP1 to promoters by the production of H3K9me3. In order to test this hypothesis, we have performed ChIP-on-chip on third instar larvae salivary glands to check HP1 binding on wt, *Setdb1* mutants and *Su(var)3-9* mutants.

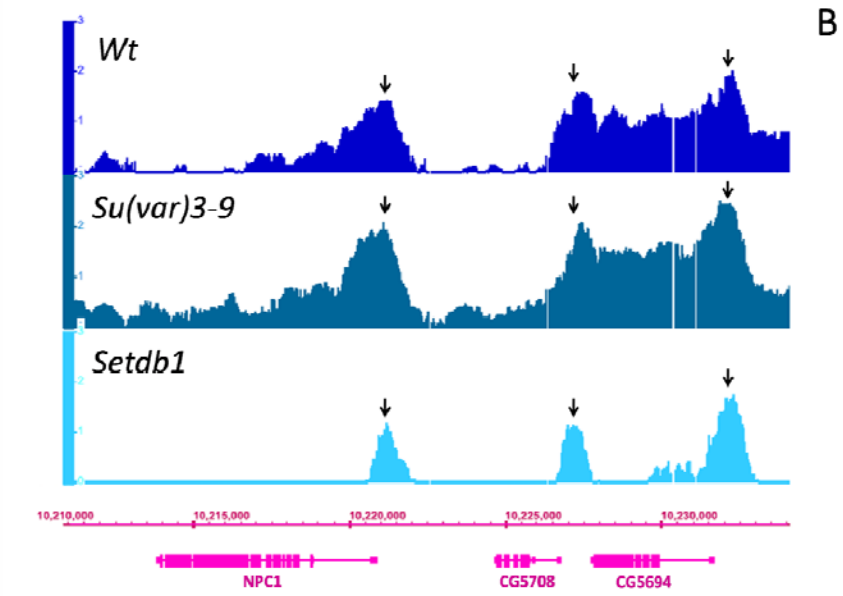
We started by analyzing the binding profile of HP1 to the chromosome 2L – **figure 5A**, since we knew that HP1 binds preferentially to the cytological region 31 of this chromosome (*2L:31*), additionally to the centromeric region (*p2L*). We found that in *Su(var)3-9* mutants HP1 binding to the centromeric region is decreased **(figure 5A)**. This reduction of HP1 in the centromeric region in *Su(var)3-9* mutants is evident when this region is analyzed at high resolution **(figure 5C)**. On the contrary, in *Setdb1* mutants HP1 binding to the centromeric p2L region is slightly increased **(figure 5A and C)**, The promoter peak of HP1 in region 31 of the

chromosome 2L is retained in both mutants and wildtype although a reduction of HP1 in the gene body is seen in *Setdb1* mutants **(figure 5B)**.

HP1 enrichment on chromosome arm 2L (log2)



HP1 enrichment in the chromosome region 2L:31 (log2)



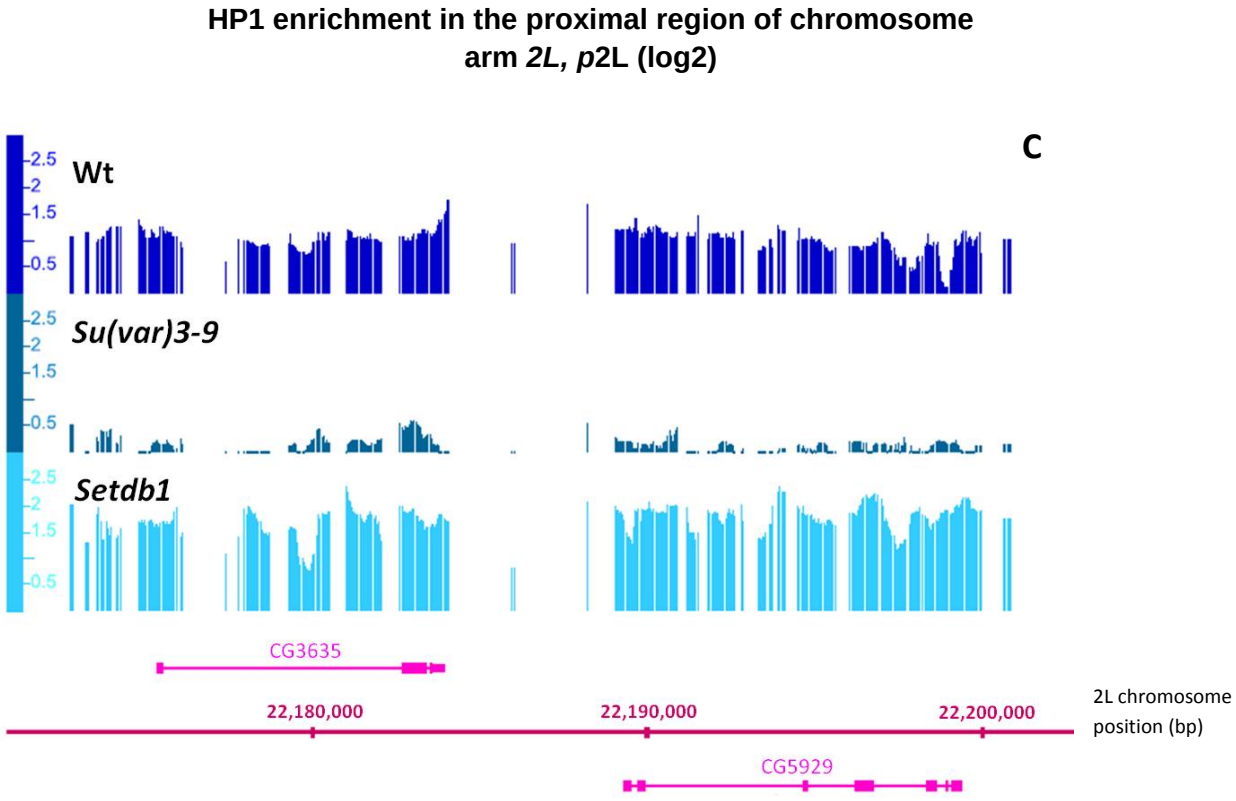


Figure 5: HP1 binding to centromeric regions depends on SU(VAR)3-9.
ChIP-on-chip binding profiles of HP1 on the 2L chromosome in *wt*, *Setdb1* mutant and *Su(var)3-9* mutant third instar larvae salivary glands. Genes below the chromosome position line are transcribed from right to left whereas genes above the chromosome position arm are transcribed from left to right. (A) View of the entire 2L chromosome arm shows HP1 enrichment in the 31 region (*2L:31*) and in the centromeric region (*p2L*) in wt. HP1 binding is reduced in the *p2L* in *Su(var)3-9* mutants and slightly increased in the *p2L* in *Setdb1* mutants. (B) Three example genes within *2L:31* show similar HP1 binding to promoters in wt, *Setdb1* and *Su(var)3-9* mutants. HP1 promoter peaks are indicated by arrows. (C) In *Su(var)3-9* mutants HP1 bindning in the *p2L* region is decreased while the corresponding binding is slightly increased in *Setdb1* mutants. Note that similar profiles are seen in the other major chromosome arms. 2L was chosen as an example chromosome.

SETDB1 and POF are required for the HP1 recruitment to 4th gene bodies

On the 4th chromosomes HP1 binds within the gene body and to a promoter peak of active genes *(figure 3)*. When we analyzed HP1 binding profile to the 4th chromosome genes in the HKMT mutants we found surprisingly that in *Setdb1* mutants HP1 was released from the gene body of all active genes on the 4th that bind HP1 in wt, but remained at promoters **(figure 6)**. Additionally, similar to our results from the immunostainings, HP1 was slightly increased on

the 4th chromosome in *Su(var)3-9* mutants. In **figure 6** these results are shown for one example gene located on the 4th chromosome.

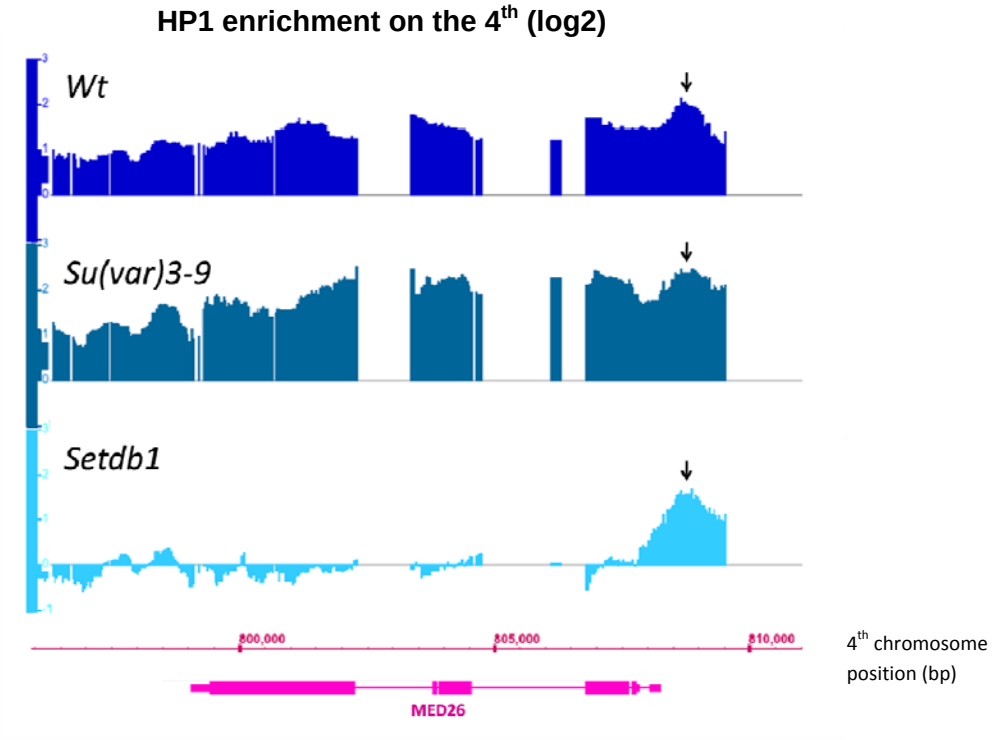


Figure 6: HP1 binding to gene bodies on the 4th depends on SETDB1.
ChIP-on-chip binding profile of HP1 to one example gene on the 4th chromosome in *wt*, *Setdb1* mutants and *Su(var)3-9* mutants third instar larvae salivary glands. The example gene *MED26* is transcribed from right to left. In *Setdb1* mutants HP1 is lost from the gene body but remains at the promoter. HP1 binding in *Su(var)3-9* is slightly increased. HP1 promoter peaks are indicated by arrows. Note that the profile shown for *MED26* gene is similar to the profiles of HP1 binding on other active genes located on the 4th chromosome.

Since it has been suggested that HP1 and POF are interdependent and involved in a balancing mechanism that regulates gene expression on the 4th chromosome [20, 28], we decided to perform ChIP-on-chip to analyze HP1 binding to the 4th chromosome in *Pof* mutants. Our ChIP-on chip results showed that in *Pof* mutants HP1 is released from the gene body of chromosome 4th genes but remains at promoters **(figure 7)**. It is important to note that the HP1 binding profile on all active genes from the 4th is very similar between *Setdb1* mutants (**figure 6**) and *Pof* mutants (**figure 7**)

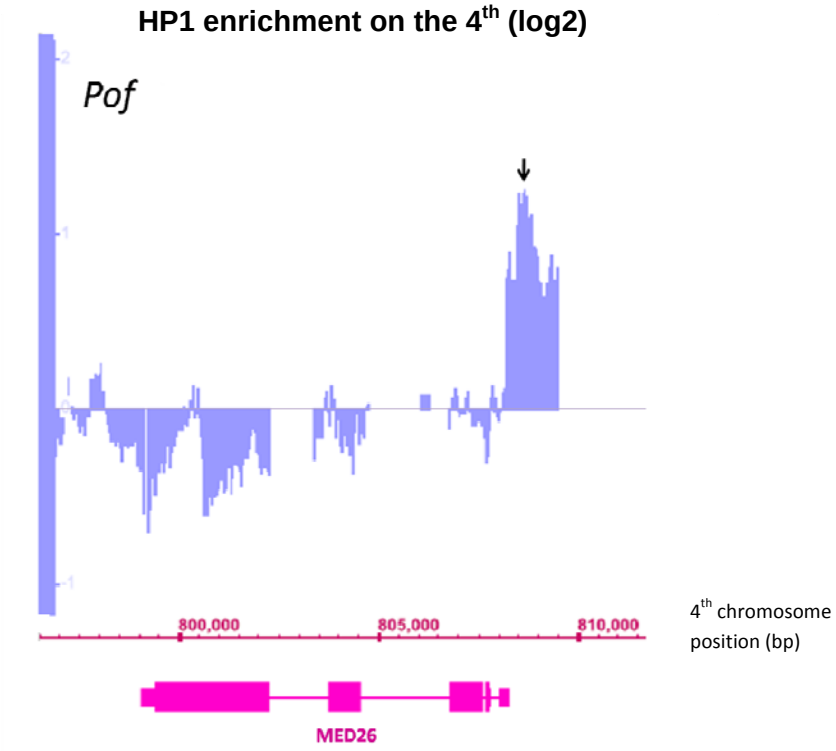


Figure 7: HP1 binding to gene bodies on the 4th depends on POF.
ChIP-on-chip binding profile of HP1 to one example gene on the 4th chromosome in *Pof* mutants third instar larvae salivary glands. The example gene *MED26* is transcribed from right to left. HP1 is lost from the gene body but remains in the promoters. HP1 promoter peak is indicated by an arrow. Note that the profile shown for *MED26* gene is similar to the profiles of HP1 binding on other active genes located on the 4th chromosome.

POF binding depends on HP1

Since HP1 and POF are interdependent we also wanted to see what happens to POF binding in the absence of HP1. We therefore performed ChIP-on-chip to analyze POF binding to the 4th chromosome of wt, *Setdb1* and *Su(var)3-9* mutants third instar larvae salivary glands. The binding profile of POF to a 50kb region of the 4th chromosome is shown in **figure 8**. On 4th chromosome active genes, POF is decreased in *Setdb1* mutants compared to wt and slightly increased in *Su(var)3-9* mutants compared to wt.

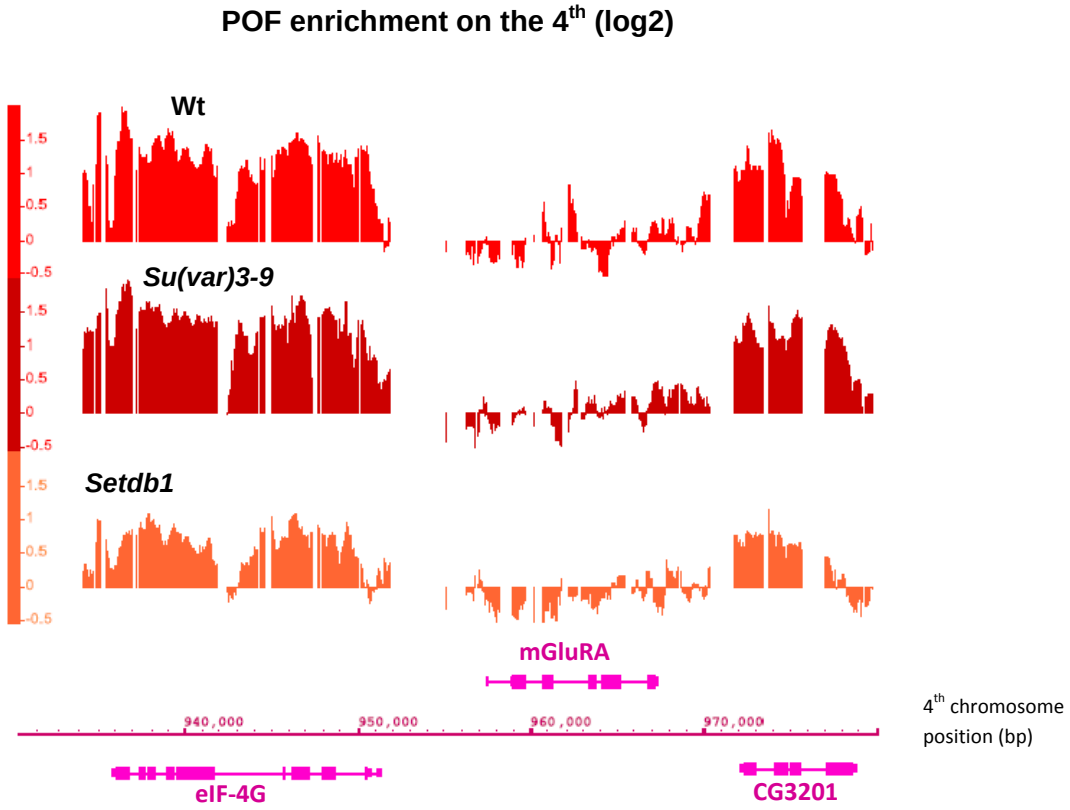


Figure 8: POF binding to genes on the 4th decreases in *Setdb1* mutants.
ChIP-on-chip binding profiles of POF at three example genes on the 4th chromosome in wt, *Su(var)3-9* and *Setdb1* mutants third instar larvae salivary glands. Genes below the chromosome position line are transcribed from right to left whereas genes above the chromosome position arm are transcribed from left to right. POF binding to genes on the 4th chromosome decreases in *Setdb1* mutants and is slightly increased in *Su(var)3-9* mutants. Note that the binding profile of POF is similar for all active genes on the 4th chromosome.

Discussion

It is known that HP1 is recruited by Histone-3 Lysine-9 methylation (H3K9me) which is done in *Drosophila melanogaster* by the following HKMTs: SU(VAR)3-9 and SETDB1. It has been suggested that SETDB1 is specific for the 4th chromosome whereas SU(VAR)3-9 acts mainly on the chromocentre [8, 9, 11]. HP1 colocalizes with POF on gene bodies of active genes on the 4th chromosome and in addition it binds to promoters [28]. The difference in binding profile between HP1 and POF and the methylation of H3K9 by different HKMTs would suggest that HP1 has two separable targets on the 4th chromosome, which may reflect differences in its function.

To further elucidate this mechanisms, we have performed chromatin immunoprecipitation followed by tiling array analysis (ChIP-on-chip) on salivary glands from wt, *Pof* mutants, *Setdb1* mutants and *Su(var)3-9* mutants third instar larvae, with antibodies against POF and HP1. We have also performed immunostainings with antibodies against HP1 and POF on polytene chromosomes from wt, *Su(var)3-9* mutants and *Setdb1* mutants third instar larvae in order to have a general view of the localization of this proteins on chromosomes which could then be compared to highly detailed protein binding profiles to specific genes obtained by ChIP-on-chip.

When comparing the binding profile to the 4th chromosome of HP1 with those from H3K9me2 and H3K9me3 (**figure 3**) we have noticed, for the *ank* gene for instance, that the HP1 promoter peak correlates better with H3K9me3 than with H3K9me2. This evidence brought us to the question if HP1 at promoters is recruited by H3K9me3. Although the literature is conflicting, SETDB1 seems to produce H3K9me2 on the 4th chromosome and both H3K9me2 and H3K9me on chromosome arms, whereas Su(var)3-9 produce H3K9me2 and H3K9me3 on the chromocentre [8, 9, 11,]. Our hypothesis was that on the 4th SETDB1 would act on exons, producing H3K9me2, and SU(VAR)3-9 would recruit HP1 to promoters by the production of H3K9me3.

To test this hypothesis we performed ChIP-on-chip experiments. When analyzing HP1 binding to active genes on the 4th chromosome in *Setdb1* mutants (**figure 6**) and *Pof* mutants (**figure 7**), we found surprisingly that HP1 was released from the exons although it was still bound to promoters. Based on these results we suggest that both SETDB1 and POF are required for the recruitment of HP1 to the exons of active genes on the 4thchromosome (**figure**

9). Additionally, since we see no effect of HP1 promoter peak in *Su(var)3-9* mutants we exclude the possibility that SU(VAR)3-9 may be involved in the recruitment of HP1 to the promoters of 4th chromosome genes. The fact that HP1 and POF are interdependent and for instance in *Pof* mutants HP1 is decreased from the 4th chromosome in immunostaining experiments [20] support the requirement of POF bound to the chromosomes in order to recruit HP1. SETDB1 seems to be the HKMT responsible for H3K9 methylation, maybe di-methylation, on the exons of the 4th. The similarity in HP1 binding profiles to the example gene *MED26* for *Setdb1* and *Pof* mutants, shown in **figures 6** and **7**, supports the view that SETDB1 acts cooperatively with POF and HP1 to regulate gene expression on the 4th [11]. It has previously been suggested that SETDB1 interacts with POF [11], which may suggest a mechanism where POF recruits SETDB1 to the exons on the 4th, which in turn produce H3K9me on the 4th which will then recruit HP1. It will be interesting to know what the functional roles of HP1 at promoters and at exons are. We speculate that HP1 at promoters is associated with decreased transcription output, since in *Setdb1* mutants transcription output from the 4th is decreased [11] and in HP1 knockdown there is an increase in gene expression [20]. The binding of HP1 to promoters seen at the 4th chromosome in *Setdb1* and *Su(var)3-9* mutants (**figures 6** and **7**) was also seen for chromosome 2L (**figure 5B**) and for all the other chromosomes of *Drosophila* (2R, 3R, 3L, X). We suggest that HP1 in promoters is recruited by a mechanism independent of H3K9me or that there is another HKMT responsible for H3K9 methylation on promoters (**figure 9**). A mechanism independent of H3K9 methylation could be the binding of HP1 to the core histone H3 [15, 16]. Additionally HP1 could theoretically be bound in promoters to proteins or methylated proteins other than histones, by its chromo and/or chromo-shadow domains. The remaining HKMT in *Drosophila* that was not studied here is the G9a, which is a candidate for H3K9 methylation, although there´s still little knowledge about this protein. Alternatively, another HKMT not yet discovered in *Drosophila* could be responsible for this promoter H3K9 methylation. In humans there are many more HKMTs [8,9] than in *Drosophila*. It will be interesting to further analyse H3K9me3 and H3K9me2 ChIP-on-chip profile in *Setdb1* mutants in order to conclude if HP1 recruitment to promoters in independent on H3K9 methylation or if there is another HKMT involved in H3K9 methylation at promoters.

When we analyzed the HP1 binding in centromeric regions of the chromosome 2L (**figure 5C**) as well as all the other chromosomes (2R, 3R, 3L, X, 4), we noticed that HP1 is released in *Su(var)3-9* mutants. This suggests that SU(VAR)3-9 is the HKMT responsible for H3K9 methylation at the centromeric regions and subsequent recruitment of HP1 (**figure 9**).

These results are in agreement with previous studies [8, 9, 11]. Taken together these results may suggest that SU(VAR)3-9 is a HKMT specific for constitutive heterochromatin.

Additionally, in immunostaining experiments HP1 binding to chromosome 4 (**figure 4**) is slightly increased on *Su(var)3-9* mutants. These results are in agreement with the slight increase of HP1 in *Su(var)3-9* mutants on 4th chromosome active genes (**figure 6**) and 31 region of the 2L chromosome (**figure 5B**) and *Setdb1* mutants centromeric region of the 2L chromosome (**Figure 5C**). We propose that HP1 is produced at a relatively constant level, but when not recruited to its normal targets it redistributes partially in other chromosomal sites where it usually also binds, the 4th chromosome on *Su(var)3-9* mutants or the chromocentre in *Setdb1* mutants. The mechanism behind HP1 redistribution could be the binding to other HP1 molecules by its chromo-shadow domain, forming homodimers.

Analysis of POF binding to the 4th chromosome genes (**figure 8**) showed that it increases and decreases slightly in *Su(var)3-9* mutants and *Setdb1* mutants, respectively. We suggest that this is due to the fact that POF is partially recruited by HP1, which is coherent with the interdependence between POF and HP1.

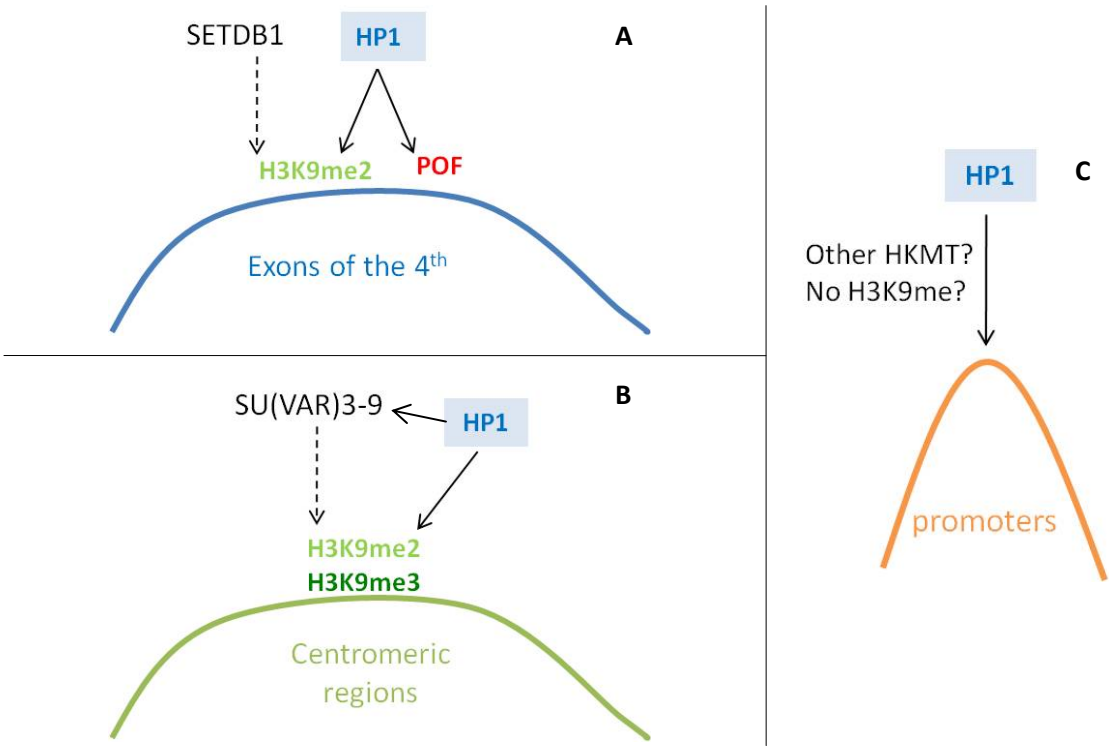


Figure 9: Proposed model for HP1 recruitment to exons of active genes on the 4th (**A**), centromeric regions of all the chromosomes (**B**) and promoters of some genes in all chromosomes (**C**). (**A**) HP1 is recruited by POF and by H3K9me2, done by SETDB1 at exons of the 4th. (**B**) HP1 is recruited by H3K9me2,3 done by SU(VAR)3-9 at centromeric regions of all chromosomes. (**C**) HP1 is recruited by a mechanism independent of H3K9me or by H3K9me done by a HKMT other than SETDB1 and SU(VAR)3-9.

Conclusions

- HP1 recruitment to exons of active genes on the 4th chromosome depends on both SETDB1 and POF
- HP1 recruitment to centromeric regions of all chromosomes depends on SU(VAR)3-9.
- HP1 recruitment to promoters of some genes on all chromosomes is independent on both SETDB1 and SU(VAR)3-9.
- POF recruitment to the 4th chromosome depends partly on SETDB1.

References

[1] Riggs A. D. and Porter T. N. 1996. Overview of epigenetic mechanisms *in* Epigenetic mechanisms of gene regulation (Ed V.E.A. Russo et al). pp. 29-45. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.

[2] Allis C. D. (edt), Jenuwein T. (edt), Reinberg D. (edt). 2007. Epigenetics. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, New York.

[3] Laskey R. A. and Gurdon J. B. 1970. Genetics content of adult somatic cells tested by nuclear transplantation from cultured cells. Nature. 228: 1332-1334.

[4] Probst A. V., Dunleavy E. and Almouzni G. 2009. Epigenetic inheritance during the cell cycle. Nature Reviews Molecular Cell Biology. 10:192-206

[5] Dimitri P., Caizzi R., Giordano E., Accardo M. C., Lattanzi G. and Biamonti G. 2009. Constitutive heterochromatin: a surprising variety of expressed sequences. Chromosoma. 118:419–435

[6] Boeke J., Regnard C., Cai W., Johansen J., Johansen K. M.,. Becker P. B. and Imhof A. 2010. Phosphorylation of SU(VAR)3–9 by the chromosomal kinase JIL-1. PLoS One. 5:e10042.

[7] Lanctôt C., Cheutin T., Cremer M., Cavalli G. and Cremer T. 2007. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nature Reviews Genetics. 8: 104-115

[8] Seum C., Reo E., Peng H., Rauscher F. J., Spierer P. and Bontron S. 2007. *Drosophila* SETDB1 is required for chromosome 4 silencing. PLoS Genetics. 3: e76

[9] Schotta G., Ebert A., Krauss V., Fischer A., Hoffmann J., Rea S., Jenuwein, T., Dorn R. and Reuter G. 2002. Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. EMBO J. 21: 1121-1131

[10] Brower-Toland B., Riddle N. C., Jiang H., Huisinga K. L. and Elgin S. C. R. 2009. Multiple SET methyltransferases are required to maintain normal heterochromatin domains in the genome of *Drosophila melanogaster*. Genetics. 181: 1303–1319

[11] Tzeng T. Y., Lee C. H., Chan L. W. and Shen C. K. 2007. Epigenetic regulation of the *Drosophila* chromosome 4 by the histone H3K9 methyltransferase dSETDB1. PNAS. 104: 12691-12696

[12] Yoon J., Lee K., Park J. S., Yu K., Paik S. and Kang Y. 2008. dSETDB1 and SU(VAR)3–9 Sequentially function during germline-stem cell differentiation in *Drosophila melanogaster*. PLoS One. 3: E2234

[13] Vermaak D. and Malik H. S. 2009. Multiple roles for Heterochromatin Protein 1 genes in *Drosophila.* Annual Review of Genetics. 43: 467-492

[14] Jacobs S. A. and Khorasanizadeh S. 2002. Structure of HP1 chromodomain bound to a lysine 9 methylated histone H3 tail. Science. 295:2080-2083

[15] Billur M., Bartunik H. D. and Singh P. B. 2009. The essential function of HP1β: a case of the tail wagging the dog? Trends in Biochemical Sciences. 35: 115-123.

[16] Nielsen A. L., Oulad-Abdelghani M., Ortiz J. A., Remboutsika E., Chambon P. and Losson R. 2001. Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. Molecular Cell. 7: 729-739

[17] Muchardt C., Guilleme M., Seeler J. S., Trouche D., Dejean A. and Yaniv M. 2002. Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1α. EMBO Reports. 3: 975-981

[18] Piacentini L., Fanti L., Negri R., Del Vescovo V., Fatica A., Altieri F. and Pimpinelli S. 2009. Heterochromatin Protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in *Drosophila*. PLoS Genetics. 5:e1000670

[19] Joan Font-Burgada, David Rossell, Herbert Auer, et al. 2008. *Drosophila* HP1c isoform interacts with the zinc-finger proteins WOC and Relative-of-WOC to regulate gene expression. Genes and Development. 22: 3007-3023

[20] Johansson A-M., Stenberg P., Bernhardsson C. and Larsson J. 2007. Painting of fourth and chromosome-wide regulation of the 4th chromosome in *Drosophila melanogaster*. EMBO J. 26: 2307-2316

[21] Riddle N. C., Shaffer C. D. and Elgin S. C. 2009. A lot about a little dot – lessons learned from *Drosophila melanogaster*. Cell Biology. 87: 229-241

[22] Kaminker J. S., Bergman C. M., Kronmiller B., Carlson J., Svirskas R. et al. 2002. The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. Genome Biology. 3: RESEARCH0084

[23] Larsson J., Chen J. D., Rasheva V., Rasmuson Lestander A. and Pirrotta V. 2001. Painting of fourth, a chromosome-specific protein in *Drosophila*. PNAS. 98: 6273-6278

[24] Stenberg P., Pettersson F., Saura A. O., Berglund A. and Larsson J. 2005. Sequence signature analysis of chromosome identity in three *Drosophila* species. BMC Bioinformatics. 6: 158

[25] Haddrill P. R., Waldron F. M. and Charlesworth B. 2008. Elevated levels of expression associated with regions of the *Drosophila* genome that lack crossing over. Biology Letters. 4: 758-761

[26] Larsson J., Svensson M. J., Stenberg P. and Mäkitalo M. 2004. Painting of fourth in genus *Drosophila* suggests autosome-specific gene regulation. PNAS. 101: 9728-9733

[27] Stenberg P., Lundberg L. E., Johansson A-M., Rydén P., Svensson M. J. and Larsson J. 2009. Buffering of segmental and chromosomal aneuploidies in *Drosophila melanogaster*. PLoS Genetics. 5:e1000465

[28] Johansson A-M., Stenberg P., Pettersson F. and Larsson J. 2007. POF and HP1 bind expressed exons, suggesting a balancing mechanism for gene regulation. PLoS Genetics. 3:e209

[29] Johansson A-M., Stenberg P., Allgardsson A. and Larsson J. 2010. POF regulates the 4th chromosome through association to nascent RNAs. *Manuscript*

Acknowledgements

I want to thank those who made this master thesis possible:

To Jan Larsson, my supervisor, for the opportunity, for all the support, for the great knowledge, for stimulating freedom of thinking and for the kindness. To Anna-Mia for letting me work in her project and for the help with the ChIP. To the other elements in JL group, Lina, Anders and Karin. To Per and Philgie in the group next door. I would like to thank you all for being so friendly since the beginning, for being always available to help, for the great scientific discussions and knowledge and for creating a funny and relaxed environment to work in. To the Flyfloor people for the fikas.

To Sven and Lilly Lawskis Foundation for Natural Science Research for the master stipend and to ERASMUS Programme for the opportunity of studying abroad.

Ao Professor Rui Gomes, meu orientador interno, pela disponibilidade e prontidão.

To all the people that I met during the erasmus, for changing my view of life. Especially to my friends Catarina, Pedro, Rui, Melanie, Courtney, Khurram, Coralie, Thierry, Seth, Raone, Paulo, Martin, Susanne, Diana and Katjana. For all the parties, travels, lunches, dinners, fikas, games, sports. For making this year in Umeå extraordinary.

Como com esta tese termino o período de estudante, quero também agradecer às pessoas que tornaram possível a minha chegada até aqui:

Á Rita, Nuno e Sofia, no ITQB, por terem sido os primeiros a me mostrarem o mundo da ciência e pela amizade.

À Dona Emília pela matemática e pela psicologia. Pela amizade e apoio.

Ao Hugo pelo apoio e amizade.

Á Sílvia pelas gargalhadas para lá do IMM e por me encorajar a mudar. Á Joana Pinto e Melissa pela loucura e amizade, por tudo o que partilhámos. À Inês Tenente pelo incentivo à perfeição.

Aos amigos para lá de BCM: Pequita, Tânia, Melissa, Daniel, Sofia, Marina. Pelas gargalhadas, choros e stresses partilhados. Pela paixão à biologia e às rambóias. Por me fazerem acreditar em mim e me transformarem.

Às CA"I" Lúcia, Joana, Clara e Filipa, por estarem sempre lá. Por todos os almoços e lanches calóricos. Por sermos tão diferentes e tão iguais.

Á Pequita e Tânia por estarem sempre lá, por ouvirem, pela generosidade e pelo optimismo.

Á Bárbara por ser a minha amiga de sempre, por tudo.

Principalmente aos meus pais por serem o meu abrigo. O vosso amor e apoio não podiam ser maiores. Obrigada por me fazerem acreditar que tudo é possível. Ao meu irmão por me fazer rir.

Á restante família por todo o apoio.