SUUR joins separate subsets of PcG, HP1 and B-type lamin targets in *Drosophila*

Alexey V. Pindyurin[1], Celine Moorman[2], Elzo de Wit[2], Stepan N. Belyakin[1], Elena S. Belyaeva[1], George K. Christophides[3], Fotis C. Kafatos[3], Bas van Steensel[2] and Igor F. Zhimulev[1,*]

[1]Institute of Cytology and Genetics of Siberian Division, Russian Academy of Sciences, Novosibirsk 630090, Russia
[2]Netherlands Cancer Institute, 1066 CX Amsterdam, The Netherlands
[3]Imperial College London, London, SW7 2AZ, UK
*Author for correspondence (e-mail: zhimulev@bionet.nsc.ru)

Accepted 17 April 2007
Journal of Cell Science 120, 2344-2351 Published by The Company of Biologists 2007
doi:10.1242/jcs.006007

Summary

***Drosophila melanogaster Suppressor of Under-Replication* (*SuUR*) gene encodes a protein that modulates replicative properties of heterochromatin in endocycles of polytene cells. The *SuUR* mutation abolishes underreplication of intercalary heterochromatin and results in partial underreplication of pericentric heterochromatin. We performed a genome-wide mapping of SUUR target genes in non-polytenic *Drosophila* Kc cells by using the DamID approach. We show that SUUR preferentially binds genes that are transcriptionally silent and late-replicated. Distinct subsets of SUUR targets are associated with PcG proteins (Pc and Esc; Polycomb and Extra sexcombs), heterochromatic proteins [HP1 and SU(VAR)3-9] and B-type lamin. The SUUR binding profile negatively correlates with the DNA polytenization levels of salivary gland polytene chromosomes. Finally, SUUR target genes are repressed in *Drosophila* embryos and gradually activated later in development. Together these results suggest that SUUR is a ubiquitous marker of heterochromatin in different cell types.**

Supplementary material available online at
http://jcs.biologists.org/cgi/content/full/120/14/2344/DC1

Key words: Heterochromatin, Nuclear lamina, Repression, SuUR, Underreplication

Introduction

Two chromatin domains can be distinguished in eukaryotic chromosomes: euchromatin and heterochromatin. In *Drosophila*, euchromatin makes up the biggest part of the chromosomes and contains the majority of genes, whereas heterochromatin constitutes about 30% of the genome and localizes around centromeres (pericentric heterochromatin; PH), at telomeres and at dispersed sites along the chromosomes (intercalary heterochromatin; IH) (Zhimulev and Belyaeva, 2003). Characteristic features of heterochromatin domains include high compaction of DNA, reduced chromatin accessibility, special histone modifications and presence of specific non-histone silencing proteins, low level of meiotic recombination, late replication (LR) in the S phase, condensed appearance throughout the cell cycle and association with the nuclear lamina (reviewed by Craig, 2005).

IH regions correspond mainly to genetically silenced gene clusters (Belyakin et al., 2005) and frequently bind Polycomb group (PcG) proteins (Zhimulev et al., 2003b). PcG proteins are required to maintain the repressed transcriptional state of important developmental regulators such as homeotic genes. They act together as components of multiprotein complexes (Saurin et al., 2001; Tie et al., 2001; Czermin et al., 2002; Muller et al., 2002; Tie et al., 2003) that assemble at Polycomb response elements (PREs) and induce a specific higher-order chromatin structure to maintain transcriptional repression of neighbouring genes (reviewed by Ringrose and Paro, 2004; Brock and Fisher, 2005). When inserted in the genome via transposition, PREs can cause long-range variable inactivation of reporter genes (Pirrotta and Rastelli, 1994; Zink and Paro, 1995). Many PcG target genes are thought to exist in the *Drosophila* genome, as PcG proteins bind to approximately 240 sites of polytene chromosomes (Zhimulev et al., 2003b). Recently, this has been confirmed at the genomic level by the identification of a large number of PcG target loci (Negre et al., 2006; Schwartz et al., 2006; Tolhuis et al., 2006).

PH is composed mainly of repetitive sequences and transposable elements, which form transcriptionally repressive domains that can suppress the expression of juxtaposed euchromatic genes (position effect variegation; PEV) (reviewed by Zhimulev and Belyaeva, 2003). It contains relatively few genes, which require a heterochromatic environment for their activity (reviewed by Yasuhara and Wakimoto, 2006). HP1 and SU(VAR)3-9 are two characterized *Drosophila* PH non-histone proteins necessary for the repression of euchromatic genes subjected to PEV (reviewed by Craig, 2005; Huisinga et al., 2006). These two proteins interact directly with each other and potentially propagate the formation of heterochromatic structure (Schotta et al., 2002). SU(VAR)3-9 plays a primary role in this process by methylating lysine 9 of histone H3 (Czermin et al., 2001; Schotta et al., 2002). HP1 binds the methylated nucleosome through its conserved chromodomain (Jacobs and Khorasanizadeh, 2002) and recruits SU(VAR)3-9, which in turn methylates the next nucleosome. In *Drosophila* Kc cells, 218 target loci of HP1 and 127 target loci of SU(VAR)3-9 have

been mapped (Greil et al., 2003). Comparison of the chromosomal distribution of HP1 and SU(VAR)3-9 targets reveals that colocalization of these proteins is not universal and occurs predominantly in pericentric regions (Greil et al., 2003).

In the salivary gland polytene chromosomes of *Drosophila melanogaster*, IH and PH regions are characterized by DNA underreplication (UR), chromosomal breaks and frequent nonhomologous (ectopic) contacts (reviewed by Zhimulev, 1998). The only known factor affecting the UR of IH regions is the Suppressor of Under-Replication (SUUR) protein (Belyaeva et al., 1998). In wild-type flies, SUUR is located in late replicating IH sites and in PH (Makunin et al., 2002). Mutation of *SuUR* leads to earlier completion of DNA replication and suppression of UR, whereas extra transgenic copies of $SuUR^{+}$ cause UR in numerous LR sites that normally are fully polytenized (Zhimulev et al., 2003b). We have previously utilized the unique ability of SUUR to modulate UR in polytene chromosomes to identify a large set of genes that are localized in UR regions (Belyakin et al., 2005). In total, 1036 genes arranged in clusters located in 52 UR chromosomal regions have been identified. The levels of UR depend on the relative gene position within these clusters: genes closer to the centre of these regions show increased levels of UR compared with genes at the borders of the regions. Furthermore, UR regions greatly overlap with LR regions in the chromosomes of both polytenic salivary glands and non-polytenic Kc cells and encompass genes with similar expression profiles, which are repressed in embryos (Belyakin et al., 2005).

In this report, we use DamID (van Steensel et al., 2001) to identify SUUR targets in the *Drosophila* Kc cell line which is of embryonic origin. This cell line was chosen in our study for two main reasons. First, *SuUR* mRNA is maternally loaded into the embryo, indicating that its product plays some role at early stages of development (Belyaeva et al., 1998; Makunin et al., 2002); *SuUR* transcripts were detected in Kc cells by non-quantitative reverse transcription-PCR analysis (reviewed by Kolesnikova et al., 2006). Second, there is an accumulating amount of information about gene expression, replication timing and genome-wide distribution of various histone modifications and chromatin proteins in Kc cells (Schübeler et al., 2002; Greil et al., 2003; Schübeler et al., 2004; Tolhuis et al., 2006), which allows us to perform comparative studies.

We detected 3001 SUUR target genes, which correspond to 27% of the 11268 probes that we examined (~70% of all *Drosophila* genes). These genes are mostly transcriptionally inactive in Kc cells, late-replicated, depleted of active chromatin histone modifications (H3-di-meK4, H3-tri-meK4, H3-Ac, H4-Ac, H3-di-K79) and enriched in the Polycomb repression nucleosome marker H3-tri-meK27. Different subsets of SUUR targets are associated with the PcG proteins Pc and Esc (Polycomb and Extra sexcombs, respectively), the heterochromatic proteins HP1 and SU(VAR)3-9 and B-type lamin (Lam). The negative correlation between the SUUR binding profile and DNA polytenization levels in salivary gland chromosomes confirms SUUR involvement in UR. Similar to genes located in UR regions of polytene chromosomes, SUUR target genes are predominantly repressed in embryos.

Results

In vivo mapping of SUUR gene targets

We mapped in vivo target genes of SUUR in cultured *Drosophila* Kc cells (embryonic origin) using the DamID chromatin profiling technique (van Steensel and Henikoff, 2000; van Steensel et al., 2001). This approach involves expression of the *Escherichia coli* DNA adenine methyltransferase (Dam) fused to a chromatin-binding protein of interest, thereby driving targeted Dam-mediated methylation of DNA sequences in the vicinity of the chromatin protein binding and/or recruitment sites. We generated both N- and C-terminal fusions of full-length *SuUR* with Dam (Dam-SUUR and SUUR-Dam, respectively) through a Myc epitope-tag linker, using previously described plasmid vectors (van Steensel and Henikoff, 2000). The fusion proteins were under the transcriptional control of the *hsp70* promoter, which drives very low expression levels in the absence of heat-shock induction (Greil et al., 2006). Under these conditions Dam fusion proteins were not detected by immunofluorescence microscopy or western blotting, indicating that only trace amounts might be present in the cell (van Steensel and Henikoff, 2000). However, upon heat shock, both fusion proteins (Dam-SUUR and SUUR-Dam) showed strong expression, with similar subnuclear localization: they co-localized with DAPI-bright regions in the nuclei of Kc cells, with the maximum concentration detected in chromocentres (Fig. 1). No signal was detected in the cytoplasm.

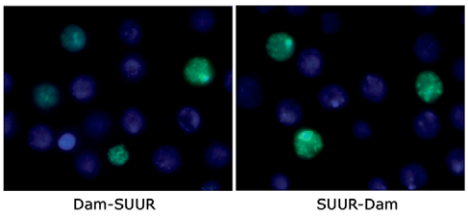

Fig. 1. Localization of transfected Dam-SUUR and SUUR-Dam fusion proteins after heat shock induction in Kc cells. The average transfection efficiency was 15-30%. FITC-conjugated donkey-anti-mouse IgG was used as a secondary antibody for detection of the 9E10 mouse monoclonal anti-Myc antibody. DNA was stained with DAPI.

To determine SUUR binding sites, genomic DNA was isolated from Kc cells transfected with the Dam fusions 24 hours earlier, in the absence of heat-shock induction, as described previously (van Steensel and Henikoff, 2000). Methylated DNA fragments were selectively amplified and then labelled with a fluorescent dye and hybridized to a microarray as described previously (Greil et al., 2003). Genomic DNA from cells transfected with a construct expressing Dam alone was subjected to the same procedures, labelled with a different fluorescent dye and used as a reference. For each SUUR fusion we performed four independent replicates (in total eight experiments). We used spotted cDNA microarrays containing 11459 probes from the *Drosophila* gene collections 1 and 2 (Rubin et al., 2000; Stapleton et al., 2002) that represent >70% of the predicted *Drosophila* genes. The data obtained from these eight experiments (Array Express, E-MEXP-863) were highly reproducible (Pearson's correlation coefficients between 0.48–0.94) and were combined to generate the DamID chromatin profile of SUUR. Low-quality and redundant probes were removed, and hybridization data from 11268 probes were used for downstream analysis. The normalized average ratio of

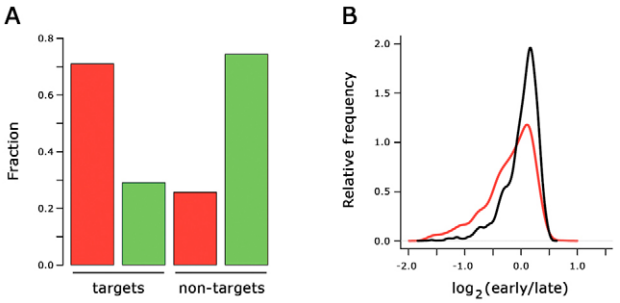

Fig. 2. SUUR target genes are mostly repressed and late replicated. Gene activity and replication timing data were taken from Schübeler et al. (Schübeler et al., 2002). Information was available for 1034 SUUR target and 4322 non-target genes. (A) Transcriptional activity of SUUR target and non-target genes. Active and repressed gene fractions are shown as green and red bars, respectively. (B) Density plot of $\log_2$-transformed replication timing of SUUR target (red) and non-target (black) genes.

SUUR-Dam or Dam-SUUR (SUUR~Dam) versus Dam was used as a measure of SUUR binding. Using a previously described error model (Greil et al., 2003), we identified 3001 probes as putative targets of SUUR.

SUUR targets were evenly distributed along the *Drosophila* chromosomes (see Fig. S1 in supplementary material). Among them we detected all genes of the Bithorax and Antennapedia complexes that were present on the microarray (*Ahcy89E*, *Abd-B*, *Glut3*, *Ubx*, *Antp* and *lab*). These two major homeotic gene complexes are well known to be PcG-dependent and are located in two typical IH regions, 89E and 84AB (Rastelli et al., 1993; Moshkin et al., 2001). This indicates that SUUR maintains these two major homeotic gene complexes in an underreplicated state. It is noteworthy that SUUR target genes were often situated next to each other (Fig. S2 in supplementary material), suggesting that SUUR may affect genes organized in gene clusters.

SUUR target genes are mostly inactive and late-replicated

It has been previously shown that SUUR is associated with regions of LR and UR in salivary gland polytene chromosomes, where predominantly silent genes reside (reviewed by Zhimulev et al., 2003a). To examine the expression stage and replication timing of the SUUR target genes in Kc cells, we compared our data with a previously reported genome-wide mRNA expression and replication timing dataset (Schübeler et al., 2002). Information was available for approximately one-third of the SUUR targets (1034 genes). The results showed that only ~30% of the SUUR target genes were expressed in Kc cells, whereas the opposite was the case for SUUR non-target genes (Fig. 2A), indicating that SUUR targets are significantly enriched for inactive genes ($P=4\times10^{-161}$, hypergeometric test). SUUR targets are also replicated later in S phase compared to non-targets ($P=4\times10^{-56}$, Wilcoxon rank sum test; Fig. 2B).

Next, we compared the SUUR binding profile with profiles of histone modifications that are associated with active chromatin, i.e. H3-di-meK4, H3-tri-meK4, H3-Ac, H4-Ac and H3-di-K79, mapped previously by Schübeler et al. (Schübeler et al., 2004). Again, information was available for the same 1034 SUUR targets. SUUR target genes were significantly depleted in all these histone modification profiles in comparison to non-target genes ($P=9\times10^{-168}$, $P=8\times10^{-152}$, $P=2\times10^{-180}$, $P=1\times10^{-159}$, $P=1\times10^{-249}$, respectively, Wilcoxon rank sum test; Fig. 3A-E). By contrast, SUUR target genes were specifically associated with a repressive histone modification,

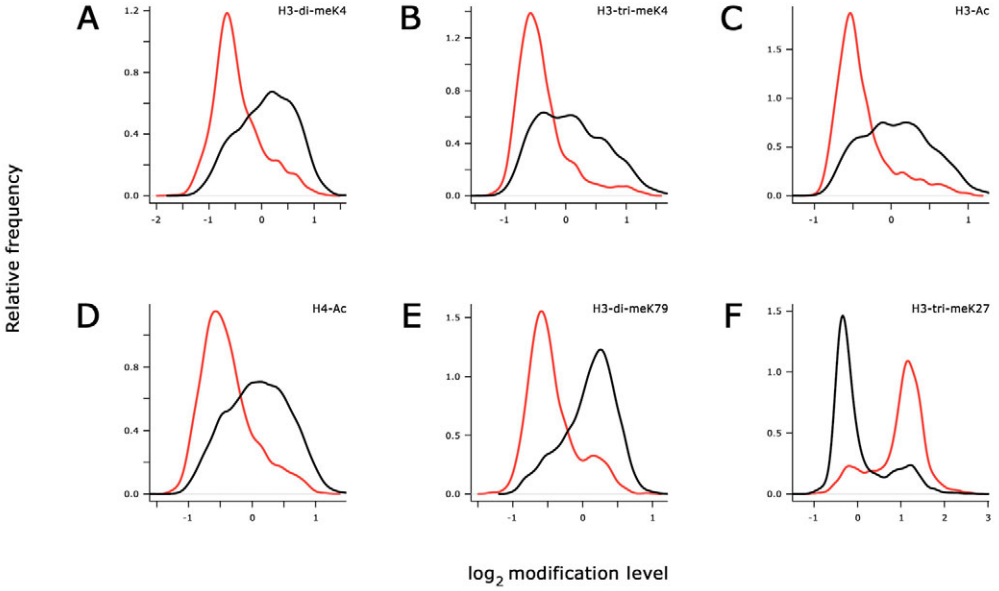

Fig. 3. SUUR target genes are depleted with active histone marks and enriched with repressive histone mark. Density plots of histone modification of SUUR target (red) and non-target (black) genes are shown for H3-di-meK4 (A), H3-tri-meK4 (B), H3-Ac (C), H4-Ac (D), H3-di-meK79 (E), and H3-tri-meK27 (F). Histone modification data for 1034 SUUR targets and 4322 non-targets (A-E) were taken from Schübeler et al. (Schübeler et al., 2004) and for 2497 SUUR targets and 7083 non-targets (F) from Tolhuis et al. (Tolhuis et al., 2006).

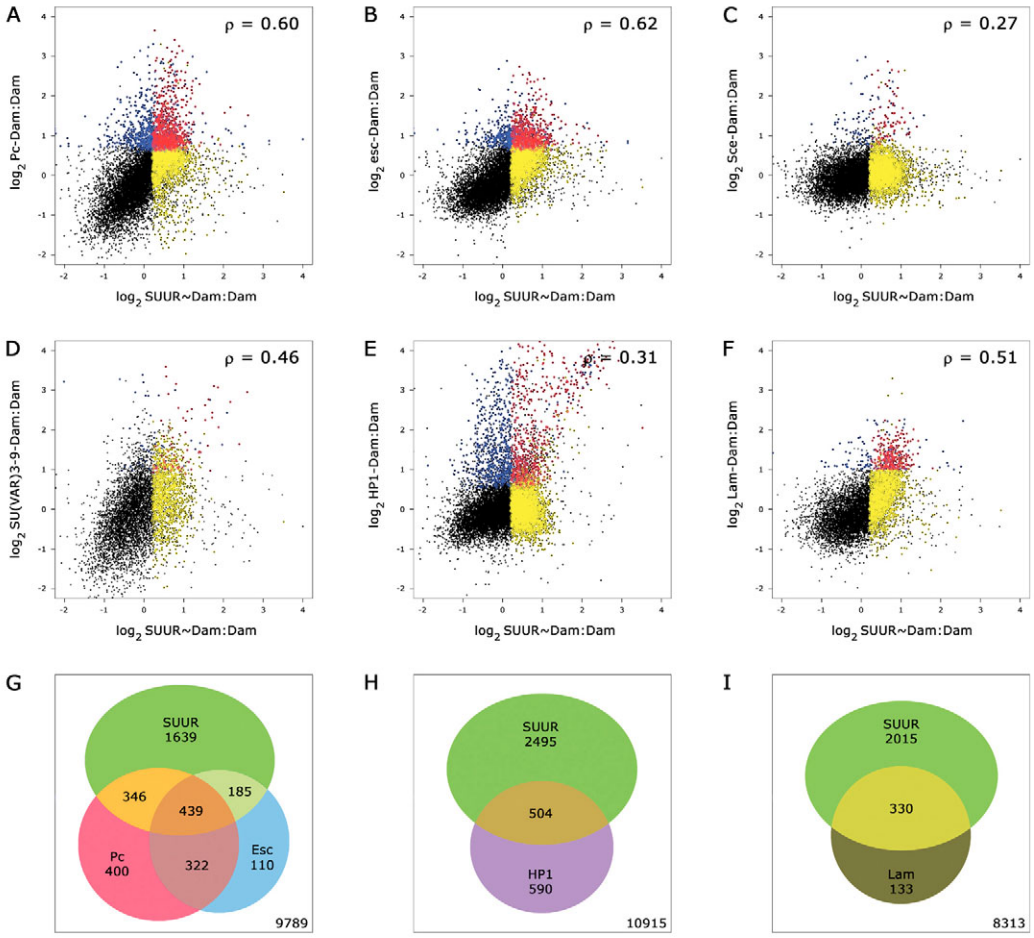

Fig. 4. Degree of similarity between the binding profiles of SUUR and Pc, Esc, Sce, SU(VAR)3-9, HP1 or Lam. (A-F) Bivariate scatter plots of $\log_2$-transformed binding ratios are shown for: SUUR versus Pc (A), SUUR versus Esc (B), SUUR versus Sce (C), SUUR versus SU(VAR)3-9 (D), SUUR versus HP1 (E) and SUUR versus Lam (F). Spearman's rank correlation coefficient is denoted at the right top corner of each plot. SUUR targets are shown in yellow, targets of other proteins in blue, and co-targets in red. (G-I) SUUR shares target genes with PcG proteins, HP1 and Lam. Venn diagrams show the overlap between target genes of SUUR and Pc, Esc (G), HP1 (H) or Lam (I). The number of investigated probes is indicated at the bottom right corner of each diagram. Pc and Esc datasets were taken from Tolhuis et al. (Tolhuis et al., 2006) and the Lam dataset from Pickersgill et al. (Pickersgill et al., 2006).

H3-tri-meK27 ($P \ll 2.2\times10^{-16}$, Wilcoxon rank sum test, Fig. 3F), shown to be associated with PcG-mediated repression (Tolhuis et al., 2006). These results again suggest a link between PcG silencing and SUUR binding. Together these data strongly support the hypothesis that SUUR is mainly associated with late-replicated genes that are transcriptionally silent.

High similarity between SUUR and PcG proteins binding profiles

Immunofluorescent staining of polytene chromosomes has previously revealed a high percentage colocalization between SUUR and PcG proteins: 67% of the SUUR binding sites overlap with PcG binding sites (Zhimulev et al., 2003b), suggesting that these proteins bind to common sequences in the *Drosophila* genome. To examine this hypothesis, we took advantage of the recently reported datasets of genome-wide binding profiles of three PcG proteins in Kc cells: Pc, Esc and Sce (Tolhuis et al., 2006). SUUR binding profile was highly similar to those of Pc and Esc and less similar to that of Sce protein (Fig. 4A-C). The Spearman's rank correlation coefficients were 0.60, 0.62 and 0.27, respectively. Notably, 52% of Pc (785 genes) and 59% of Esc (624 genes) targets overlap with SUUR targets. Of these, 439 genes were co-targets of all three proteins (Fig. 4G).

Modest similarity between SUUR and PH proteins binding profiles

Since SUUR localizes in PH of polytene chromosomes (Makunin et al., 2002) and influences the heterochromatin-dependent PEV silencing (Belyaeva et al., 2003), we were interested in its colocalization with other heterochromatic proteins. DamID chromatin profiles of HP1 and SU(VAR)3-9 in Kc cells have been previously generated by Greil et al. (Greil et al., 2003) using microarrays containing over 6200 cDNA and genomic fragments. For comparison with the SUUR binding profile, we used the published dataset for SU(VAR)3-9 and generated a new HP1 dataset using the *Drosophila* gene collections 1 and 2 microarrays. We performed three

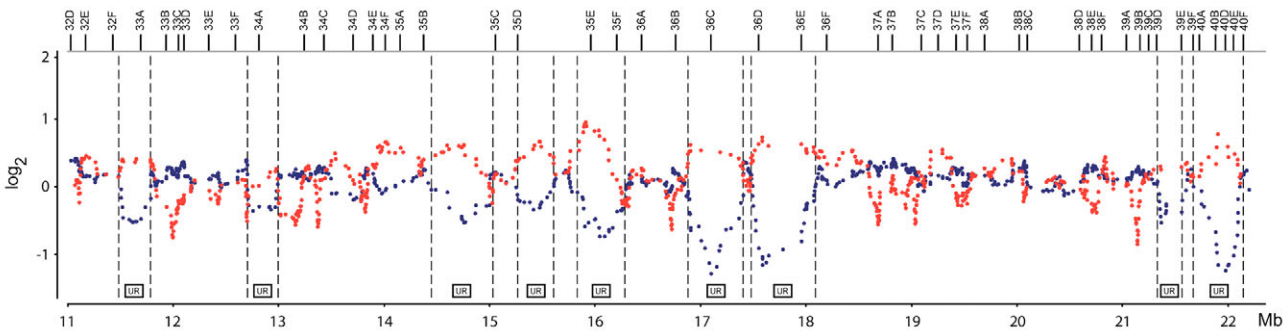

Fig. 5. Correlation between the SUUR binding profile in Kc cells and the UR profile of salivary gland polytene chromosomes. An 11.3 Mb fragment of chromosomal arm 2L is shown. Cytological regions and chromosome nucleotide positions are indicated above and below the graph. A running mean algorithm (a sliding window of 10 genes, one gene per step) was applied to the UR and SUUR binding data. The UR running mean is represented by blue dots and the SUUR binding running mean by red dots. UR regions are confined by vertical dotted lines and labelled in rectangles. UR data were taken from Belyakin et al. (Belyakin et al., 2005).

independent DamID experiments for HP1 (Array Express, E-MEXP-864) and identified 1094 targets (see Fig. S3 in supplementary material). The results were highly consistent to those published previously (Greil et al., 2003).

The comparison of the SUUR DamID profile with those of HP1 and SU(VAR)3-9 detected a modest but apparent similarity, with Spearman's rank correlation coefficients 0.31 and 0.46, respectively (Fig. 4D,E); 46% of the HP1 (504 genes) and 61% of the SU(VAR)3-9 (69 genes) targets were at the same time SUUR targets (Fig. 4H and Fig. S4 in supplementary material). It has been previously shown that a part of HP1 targets are genes or transposable elements flanked by repetitive genomic sequences (de Wit et al., 2005). SUUR and HP1 targets identified in our study were also frequently flanked by such repeats (see Fig. S5 in supplementary material).

High similarity between SUUR and B-type lamin binding profiles

B-type lamin is an essential component of the inner nuclear membrane encoded by the ubiquitously expressed *Dm0* (also known as *Lam*) gene. Genes that interact in vivo with the nuclear lamina (Lam targets) have been identified recently in Kc cells (Pickersgill et al., 2006). Comparison of SUUR and Lam binding profiles demonstrated a high level of similarity, with Spearman's rank correlation coefficient 0.51 (Fig. 4F); 71% of Lam (330 genes) targets were at the same time SUUR targets (Fig. 4I). In addition, we found that 92% (48 out of 52) of Lam target gene clusters (Pickersgill et al., 2006) correspond to previously mapped IH regions of polytene chromosomes (Zhimulev et al., 2003b). Moreover, 24 Lam target gene clusters coincide with chromosomal regions whose levels of UR in polytene chromosomes depend on the dose of *SuUR* gene (Belyakin et al., 2005). Notably, 118 genes were simultaneously bound by SUUR, Lam and Pc, whereas the overlap between targets of SUUR, Lam and HP1 was limited to 18 genes (see Fig. S6 in supplementary material). These data suggest that SUUR is associated with almost separate subsets of Pc, HP1 and Lam targets.

SUUR binding profile in Kc cells negatively correlates with DNA polytenization level in polytene chromosomes

In salivary gland polytene chromosomes, 52 genomic regions whose replication level is affected by SUUR have been previously detected (Belyakin et al., 2005). These regions correspond to cytologically defined UR regions. We observed moderate negative correlation between the levels of DNA polytenization and SUUR~Dam-mediated DNA methylation: the less DNA present in the genomic locus of the polytene chromosome, the more methylation is detected in Kc cells (Fig. 5). The Pearson's correlation coefficients were –0.368, –0.478, –0.418, –0.320, –0.352 for the X, 2L, 2R, 3L and 3R chromosomal arms, respectively. This is the first direct observation of correlation between SUUR localization and UR at the gene level, and confirms earlier results showing dependence of UR levels, e.g. along the Bithorax complex, upon *SuUR* gene dose (Moshkin et al., 2001; Zhimulev et al., 2003b).

Developmental expression of SUUR target genes

We investigated the expression profiles of the SUUR target

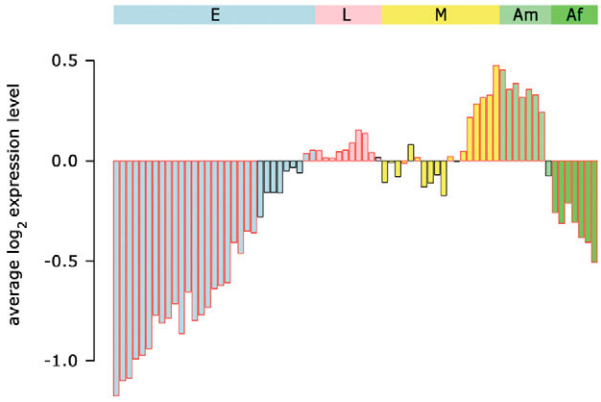

Fig. 6. SUUR is associated with genes that are repressed in embryos. Average expression of 744 SUUR target genes during *Drosophila* development: E, embryos; L, larvae; M, metamorphosis; Am, adult male; and Af, adult female. Bars with red borders indicate developmental stages with significant difference in expression between SUUR targets and non-targets ($P<0.01$, Wilcoxon rank sum test, Bonferroni corrected). Gene expression data were taken from Arbeitman et al. (Arbeitman et al., 2002).

genes throughout *Drosophila* development utilizing a previously published dataset (Arbeitman et al., 2002). The results revealed that on average SUUR targets are repressed in embryos, but are gradually activated later in development, reaching their expression peak in late pupae and adult males, but adult females show no activity of SUUR targets (Fig. 6 and Fig. S7 in supplementary material). This pattern substantially resembles our previously reported data on the transcriptional activity of genes located in UR regions of polytene chromosomes (Belyakin et al., 2005).

Discussion

SUUR preferentially binds to genes that are transcriptionally silent and late replicated in the non-polytenic cells

To date, our knowledge about SUUR protein localization in the *Drosophila* chromosomes has derived mainly from low-resolution immunostainings of polytene chromosomes (Makunin et al., 2002). It has been known that SUUR possesses in vitro DNA-binding activity for sequences of two putative heterochromatic boundary regions underreplicated in the salivary gland polytene chromosomes (Tchurikov et al., 2004). However, no consensus DNA binding motif between these boundary regions has been detected, suggesting that SUUR may bind to these regions in a structure-specific manner (Tchurikov et al., 2004). More recently, putative SUUR genomic DNA targets have been identified by taking advantage of SUUR's effect on the replication levels of salivary gland polytene chromosomes (Belyakin et al., 2005). However, it has been unclear whether direct binding or indirect influence of SUUR in these regions causes this effect. Here, our genome-wide mapping of SUUR target genes on non-polytenic chromosomes of Kc cells is the first attempt to identify DNA sequences associated with the SUUR protein. As both N- and C-terminal Dam fusions with SUUR generated very similar results and there was a very extensive overlap between SUUR binding and UR regions in salivary gland polytene chromosomes, we believe that we mapped native SUUR binding sites in the chromosomes of Kc cells. Using cDNA microarrays that encompass 70% of the predicted *Drosophila* genes, we identified ~3000 genes as putative SUUR targets. Therefore it could be estimated that the total number of SUUR target genes are over 4000.

We show for the first time that SUUR binds preferentially to genes that are transcriptionally silent and late-replicated in non-polytenic cells. Most likely SUUR is not directly involved in transcriptional repression of its targets, or at least it is not essential for their silencing, since *SuUR* mutants are viable and do not display any phenotype that would link to incorrect expression of SUUR targets (Belyaeva et al., 1998).

SUUR is associated with IH regions localizing at the nuclear envelope

Recent evidence suggests that localization of chromatin loci at the inner nuclear membrane is a regulated process with profound consequences on their activity. LR foci and inactive genes are often more prominent along the nuclear periphery (reviewed by McNairn and Gilbert, 2003; Misteli, 2004). Nuclear lamina proteins are essential for viability of cells and whole organisms, and play an important role in the positioning of genes and chromosomes at the nuclear envelope (reviewed by Taddei et al., 2004), however, detailed molecular mechanisms responsible for this positioning are still unclear.

Our work reveals a positive correlation between the binding profiles of SUUR and Lam, and extensive overlap of their targets. Most of the Lam target gene clusters (Pickersgill et al., 2006) are in IH regions that bind SUUR in polytene chromosomes (Zhimulev et al., 2003b). Interestingly, IH regions have been found in close association with the nuclear lamina of polytene nuclei in *Drosophila* third instar larvae (Hochstrasser and Sedat, 1987). The moderate negative correlation between SUUR-mediated DNA polytenization levels in salivary gland polytene chromosomes (Belyakin et al., 2005) and SUUR binding in Kc cells detected here indicates that the more SUUR is bound in an IH region the less DNA is present. Thus, IH regions localizing at the inner nuclear membrane are in association with SUUR in salivary gland polytenic cells and cells of embryonic origin.

SUUR is ubiquitous marker of heterochromatin regions

Genome-wide analysis of SUUR targets shows that these could be divided into at least four almost non-overlapping subsets. The first and second subsets are associated with the well-known silencing proteins Pc and HP1, respectively. The third subset is composed of genes from IH regions localizing at the inner nuclear membrane (B-type lamin targets). The rest of the SUUR targets (fourth subset) might also be associated with other proteins of silent chromosomal domains. Hence, we consider SUUR as an ubiquitous marker of heterochromatin regions in *Drosophila*. As the microarray platform used in this study is void of transposable elements and repetitive sequences, which are prominent targets of SU(VAR)3-9 and HP1 (Greil et al., 2003), the obtained results concern IH rather than PH regions. It is possible that SUUR associates with different types of silencing mechanisms by recognizing the repressed chromatin structure, but it is yet unclear how binding of SUUR to certain chromatin sites leads to their LR and – as a consequence – UR in polytene cells. The current hypothesis assumes that this could be achieved by SUUR interfering with the DNA replication machinery.

Previously, we demonstrated that IH regions, whose level of UR in salivary gland polytene chromosomes is affected by SUUR, harbour transcriptional territories (Belyakin et al., 2005), i.e. groups of 10-30 adjacent and similarly expressed genes (Spellman and Rubin, 2002). The expression pattern of these territories during *Drosophila* development (Belyakin et al., 2005) largely coincides with the expression pattern of SUUR targets in Kc cells. The high degree of similarity between the binding profiles of SUUR and Pc in Kc cells along with the high colocalization level of these proteins in IH regions of salivary gland polytene chromosomes (Zhimulev et al., 2003b) indicate similar organization of repressed chromatin domains in different cell types. Taken together, our findings suggest that SUUR has a universally temporal and spatial pattern of binding to genomic DNA throughout *Drosophila* development.

Materials and Methods

Constructions of pDam-Myc-SUUR and pSUUR-Myc-Dam plasmids

A DNA fragment containing the 5′ end of the *SuUR* cDNA (from –3 to +267, where +1 is the first nucleotide of the translation start codon) was amplified by PCR from the f40 cDNA clone (Makunin et al., 2002) using the oligonucleotide primers SbA

(forward; 5′-ttccgcggatcATGtatcactttgtatccg-3′) and an4 (reverse; 5′-ttaag-cttaccggtgagcagttgc-3′). After digestion with *Sac*II and *Age*I (which recognize the sequences underlined in the primers), the DNA fragment was inserted back into corresponding sites of the f40 clone to produce the pBS-S2-SUUR plasmid.

A DNA fragment containing the 3′ end of the *SuUR* cDNA (from +2483 to +2887, right up to the stop codon) was amplified by PCR from the f40 clone using the primers an5 (forward; 5′-aaggatccgcatgcgattgaactc-3′) and SbS (reverse; 5′-aaggtaccccgcggacttgaacagttccaatc-3′). The product was digested with *Sph*I and *Kpn*I and cloned into the corresponding sites of pBS-S2-SUUR to produce the pBS-S2-SUUR-S2 plasmid, which was sequenced to confirm the absence of mutations resulting from errors during PCR.

To obtain the pDam-Myc-SUUR and pSUUR-Myc-Dam plasmids, a *Sac*II-*Sac*II fragment bearing the full *SuUR* open reading frame was excised from pBS-S2-SUUR-S2 and cloned into the pNDamMyc and pCMycDam vectors (van Steensel and Henikoff, 2000), respectively.

DamID

DNA adenine methyltransferase identification (DamID) experiments were performed on Kc167 cells according to previously reported protocols (Greil et al., 2003). Four independent experiments (two with reversed dye orientation) for both Dam-SUUR and SUUR-Dam fusion protein were performed. To map HP1 we used the previously described pDamHP1 plasmid (van Steensel and Henikoff, 2000) and performed three independent experiments (one with reversed dye orientation). To locate the SUUR fusion proteins in Kc cells after heat shock induction, cells were stained with the 9E10 mouse monoclonal anti-Myc antibody (Abcam, UK). *Drosophila* Gene Collections Releases 1 and 2 (Rubin et al., 2000; Stapleton et al., 2002) cDNA microarrays were produced in the European Molecular Biology Laboratory (Heidelberg, Germany) and contained 11459 full-length cDNAs. Pre-hybridizations were carried out at 42°C in a buffer containing 5×SSC, 0.1% SDS and 1% BSA for 2 hours. Hybridizations were performed overnight at 42°C in a buffer containing 25% formamide, 2.5×SSC, 0.1% SDS and 12.5% chemical block mix KREA*block*™ (KREATECH Biotechnology B.V., Amsterdam, The Netherlands). The first three washes were carried out at 42°C (few seconds in 5×SSC/0.1% SDS, 30 seconds in 2×SSC/0.1% SDS and 5 minutes in 1×SSC) and were followed by two washes at room temperature (2 minutes in 0.2×SSC and 20 seconds in 0.05×SSC). Microarray scanning and analyses were performed as previously reported (Greil et al., 2003; de Wit et al., 2005).

Computational analyses

Computational analyses were performed with the R statistical programming environment (http://www.r-project.org). Expression, histone modification and replication timing data were taken from Schübeler et al. (Schübeler et al., 2002; Schübeler et al., 2004), Pc, Esc, Sce and H3-tri-meK27 data from Tolhuis et al. (Tolhuis et al., 2006), SU(VAR)3-9 data from Greil et al. (Greil et al., 2003) and Lam data from Pickersgill et al. (Pickersgill et al., 2006). Analysis of flanking repeats was performed as in de Wit et al. (de Wit et al., 2005).

Accession numbers

The microarray data have been submitted to Array Express under accession numbers E-MEXP-863 and E-MEXP-864.

We thank F. Greil and L. Guelen for helpful suggestions and the Genomics Core Facility (GeneCore) of the European Molecular Biology Laboratory (Heidelberg, Germany) for providing the *Drosophila* cDNA microarrays. This work was funded by an Interdisciplinary Integration Project of SB RAS #45, a Molecular and Cellular Biology RAS Program #10.1 and a Leading Scientific Schools #942.2006.4. A.V.P. was supported by a Russian Foundation for Basic Research Grant #06-04-49305a and by an EMBO Short-Term Fellowship; S.N.B. was supported by a Young Scientists Program of SB RAS. Work in the B.v.S. laboratory was supported by a Human Frontier Program Grant and a European Young Investigator Award.

References

Arbeitman, M. N., Furlong, E. E., Imam, F., Johnson, E., Null, B. H., Baker, B. S., Krasnow, M. A., Scott, M. P., Davis, R. W. and White, K. P. (2002). Gene expression during the life cycle of Drosophila melanogaster. *Science* **297**, 2270-2275.

Belyaeva, E. S., Zhimulev, I. F., Volkova, E. I., Alekseyenko, A. A., Moshkin, Y. M. and Koryakov, D. E. (1998). Su(UR)ES: a gene suppressing DNA underreplication in intercalary and pericentric heterochromatin of Drosophila melanogaster polytene chromosomes. *Proc. Natl. Acad. Sci. USA* **95**, 7532-7537.

Belyaeva, E. S., Boldyreva, L. V., Volkova, E. I., Nanayev, R. A., Alekseyenko, A. A. and Zhimulev, I. F. (2003). Effect of the Suppressor of Underreplication (SuUR) gene on position-effect variegation silencing in Drosophila melanogaster. *Genetics* **165**, 1209-1220.

Belyakin, S. N., Christophides, G. K., Alekseyenko, A. A., Kriventseva, E. V., Belyaeva, E. S., Nanayev, R. A., Makunin, I. V., Kafatos, F. C. and Zhimulev, I. F. (2005). Genomic analysis of Drosophila chromosome underreplication reveals a link between replication control and transcriptional territories. *Proc. Natl. Acad. Sci. USA* **102**, 8269-8274.

Brock, H. W. and Fisher, C. L. (2005). Maintenance of gene expression patterns. *Dev. Dyn.* **232**, 633-655.

Craig, J. M. (2005). Heterochromatin – many flavours, common themes. *BioEssays* **27**, 17-28.

Czermin, B., Schotta, G., Hulsmann, B. B., Brehm, A., Becker, P. B., Reuter, G. and Imhof, A. (2001). Physical and functional association of SU(VAR)3-9 and HDAC1 in Drosophila. *EMBO Rep.* **2**, 915-919.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A. and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**, 185-196.

de Wit, E., Greil, F. and van Steensel, B. (2005). Genome-wide HP1 binding in Drosophila: developmental plasticity and genomic targeting signals. *Genome Res.* **15**, 1265-1273.

Greil, F., van der Kraan, I., Delrow, J., Smothers, J. F., de Wit, E., Bussemaker, H. J., van Driel, R., Henikoff, S. and van Steensel, B. (2003). Distinct HP1 and Su(var)3-9 complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. *Genes Dev.* **17**, 2825-2838.

Greil, F., Moorman, C. and van Steensel, B. (2006). DamID: mapping of in vivo protein-genome interactions using tethered DNA adenine methyltransferase. *Meth. Enzymol.* **410**, 342-359.

Hochstrasser, M. and Sedat, J. W. (1987). Three-dimensional organization of Drosophila melanogaster interphase nuclei. II. Chromosome spatial organization and gene regulation. *J. Cell Biol.* **104**, 1471-1483.

Huisinga, K. L., Brower-Toland, B. and Elgin, S. C. (2006). The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* **115**, 110-122.

Jacobs, S. A. and Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* **295**, 2080-2083.

Kolesnikova, T. D., Andreeva, E. N., Pindyurin, A. V., Ananko, N. G., Belyakin, S. N., Shloma, V. V., Yurlova, A. A., Makunin, I. V., Pokholkova, G. V., Volkova, E. I. et al. (2006). Contribution of the SuUR gene to the organization of epigenetically repressed regions of Drosophila melanogaster chromosomes. *Russ. J. Genet.* **42**, 1013-1028.

Makunin, I. V., Volkova, E. I., Belyaeva, E. S., Nabirochkina, E. N., Pirrotta, V. and Zhimulev, I. F. (2002). The Drosophila suppressor of underreplication protein binds to late-replicating regions of polytene chromosomes. *Genetics* **160**, 1023-1034.

McNairn, A. J. and Gilbert, D. M. (2003). Epigenomic replication: linking epigenetics to DNA replication. *BioEssays* **25**, 647-656.

Misteli, T. (2004). Spatial positioning; a new dimension in genome function. *Cell* **119**, 153-156.

Moshkin, Y. M., Alekseyenko, A. A., Semeshin, V. F., Spierer, A., Spierer, P., Makarevich, G. F., Belyaeva, E. S. and Zhimulev, I. F. (2001). The bithorax complex of Drosophila melanogaster: Underreplication and morphology in polytene chromosomes. *Proc. Natl. Acad. Sci. USA* **98**, 570-574.

Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E. and Simon, J. A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111**, 197-208.

Negre, N., Hennetin, J., Sun, L. V., Lavrov, S., Bellis, M., White, K. P. and Cavalli, G. (2006). Chromosomal distribution of PcG proteins during Drosophila development. *PLoS Biol.* **4**, e170.

Pickersgill, H., Kalverda, B., de Wit, E., Talhout, W., Fornerod, M. and van Steensel, B. (2006). Characterization of the Drosophila melanogaster genome at the nuclear lamina. *Nat. Genet.* **38**, 1005-1014.

Pirrotta, V. and Rastelli, L. (1994). White gene expression, repressive chromatin domains and homeotic gene regulation in Drosophila. *BioEssays* **16**, 549-556.

Rastelli, L., Chan, C. S. and Pirrotta, V. (1993). Related chromosome binding sites for zeste, suppressors of zeste and Polycomb group proteins in Drosophila and their dependence on Enhancer of zeste function. *EMBO J.* **12**, 1513-1522.

Ringrose, L. and Paro, R. (2004). Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu. Rev. Genet.* **38**, 413-443.

Rubin, G. M., Hong, L., Brokstein, P., Evans-Holm, M., Frise, E., Stapleton, M. and Harvey, D. A. (2000). A Drosophila complementary DNA resource. *Science* **287**, 2222-2224.

Saurin, A. J., Shao, Z., Erdjument-Bromage, H., Tempst, P. and Kingston, R. E. (2001). A Drosophila Polycomb group complex includes Zeste and dTAFII proteins. *Nature* **412**, 655-660.

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.

Schübeler, D., Scalzo, D., Kooperberg, C., van Steensel, B., Delrow, J. and Groudine, M. (2002). Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing. *Nat. Genet.* **32**, 438-442.

Schübeler, D., MacAlpine, D. M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D. E., O'Neill, L. P., Turner, B. M., Delrow, J. et al. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* **18**, 1263-1271.

Schwartz, Y. B., Kahn, T. G., Nix, D. A., Li, X. Y., Bourgon, R., Biggin, M. and Pirrotta, V. (2006). Genome-wide analysis of Polycomb targets in Drosophila melanogaster. *Nat. Genet.* **38**, 700-705.

Spellman, P. T. and Rubin, G. M. (2002). Evidence for large domains of similarly expressed genes in the Drosophila genome. *J. Biol.* **1**, 5.

Stapleton, M., Liao, G., Brokstein, P., Hong, L., Carninci, P., Shiraki, T., Hayashizaki, Y., Champe, M., Pacleb, J., Wan, K. et al. (2002). The Drosophila gene collection: identification of putative full-length cDNAs for 70% of D. melanogaster genes. *Genome Res.* **12**, 1294-1300.

Taddei, A., Hediger, F., Neumann, F. R. and Gasser, S. M. (2004). The function of nuclear architecture: a genetic approach. *Annu. Rev. Genet.* **38**, 305-345.

Tchurikov, N. A., Kretova, O. V., Chernov, B. K., Golova, Y. B., Zhimulev, I. F. and Zykov, I. A. (2004). SuUR protein binds to the boundary regions separating forum domains in Drosophila melanogaster. *J. Biol. Chem.* **279**, 11705-11710.

Tie, F., Furuyama, T., Prasad-Sinha, J., Jane, E. and Harte, P. J. (2001). The Drosophila Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. *Development* **128**, 275-286.

Tie, F., Prasad-Sinha, J., Birve, A., Rasmuson-Lestander, A. and Harte, P. J. (2003). A 1-megadalton ESC/E(Z) complex from Drosophila that contains polycomblike and RPD3. *Mol. Cell. Biol.* **23**, 3352-3362.

Tolhuis, B., Muijrers, I., de Wit, E., Teunissen, H., Talhout, W., van Steensel, B. and van Lohuizen, M. (2006). Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in Drosophila melanogaster. *Nat. Genet.* **38**, 694-699.

van Steensel, B. and Henikoff, S. (2000). Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat. Biotechnol.* **18**, 424-428.

van Steensel, B., Delrow, J. and Henikoff, S. (2001). Chromatin profiling using targeted DNA adenine methyltransferase. *Nat. Genet.* **27**, 304-308.

Yasuhara, J. C. and Wakimoto, B. T. (2006). Oxymoron no more: the expanding world of heterochromatic genes. *Trends Genet.* **22**, 330-338.

Zhimulev, I. F. (1998). Polytene chromosomes, heterochromatin, and position effect variegation. *Adv. Genet.* **37**, 1-566.

Zhimulev, I. F. and Belyaeva, E. S. (2003). Intercalary heterochromatin and genetic silencing. *BioEssays* **25**, 1040-1051.

Zhimulev, I. F., Belyaeva, E. S., Makunin, I. V., Pirrotta, V., Semeshin, V. F., Alekseyenko, A. A., Belyakin, S. N., Volkova, E. I., Koryakov, D. E., Andreyeva, E. N. et al. (2003a). Intercalary heterochromatin in Drosophila melanogaster polytene chromosomes and the problem of genetic silencing. *Genetica* **117**, 259-270.

Zhimulev, I. F., Belyaeva, E. S., Makunin, I. V., Pirrotta, V., Volkova, E. I., Alekseyenko, A. A., Andreyeva, E. N., Makarevich, G. F., Boldyreva, L. V., Nanayev, R. A. et al. (2003b). Influence of the SuUR gene on intercalary heterochromatin in Drosophila melanogaster polytene chromosomes. *Chromosoma* **111**, 377-398.

Zink, D. and Paro, R. (1995). Drosophila Polycomb-group regulated chromatin inhibits the accessibility of a trans-activator to its target DNA. *EMBO J.* **14**, 5660-5671.