

# Buffering and the evolution of chromosome-wide gene regulation

Per Stenberg · Jan Larsson

Received: 7 February 2011 / Revised: 15 March 2011 / Accepted: 31 March 2011 / Published online: 20 April 2011  
© The Author(s) 2011. This article is published with open access at Springerlink.com

**Abstract** Copy number variation (CNV) in terms of aneuploidies of both entire chromosomes and chromosomal segments is an important evolutionary driving force, but it is inevitably accompanied by potentially problematic variations in gene doses and genomic instability. Thus, a delicate balance must be maintained between mechanisms that compensate for variations in gene doses (and thus allow such genomic variability) and selection against destabilizing CNVs. In *Drosophila*, three known compensatory mechanisms have evolved: a general segmental aneuploidy-buffering system and two chromosome-specific systems. The two chromosome-specific systems are the male-specific lethal complex, which is important for dosage compensation of the male X chromosome, and Painting of fourth, which stimulates expression of the fourth chromosome. In this review, we discuss the origin and function of buffering and compensation using *Drosophila* as a model.

## Introduction

Generally, we consider humans and most animals to be diploid, i.e., that two copies of nearly all nuclear genes and chromosomal regions are present in all of their cells except gametes. However, many plants and some animals are

polyploids, i.e., their cells contain one or more additional copies of their entire genome. In humans, polyploid embryos are sometimes formed by fertilization of oocytes by more than one sperm or by faulty meiosis that results in diploid gametes. These embryos usually die early, few surviving until birth. Mature humans are never true polyploids, but some of our tissues contain extra sets of chromosomes. For example, the liver contains cells with up to eight copies of all chromosomes (octoploidy). Variations in the number of individual chromosomes (aneuploidy) are more common than polyploidy in humans. Aneuploidies are formed when chromosomes fail to segregate properly during cell division (non-disjunction) and severe syndromes arise when embryos are formed by aneuploid gametes. Classic examples include Down's syndrome (caused by triploidy of chromosome 21) as well as Turner's and Klinefelter's syndromes (caused by aberrant numbers of X chromosomes).

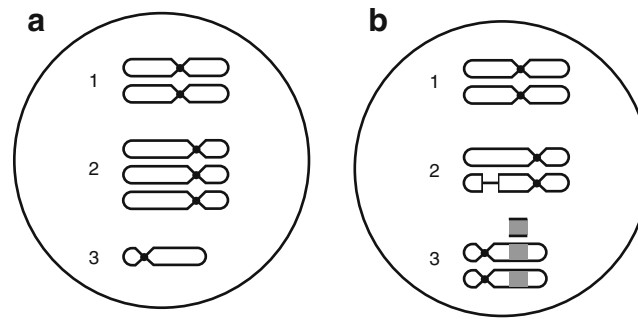
Following the recent revolution in whole genome analysis techniques (e.g., microarrays and next generation sequencing), it has become clear that the copy number of parts of chromosomes (Fig. 1) varies quite extensively in human populations. In fact, if two individuals are randomly picked from any natural population of any species, there are likely to be dozens to hundreds of copy number differences of their functional genes (Schrider and Hahn 2010). Due to resolution limitations of micro-array techniques, most studies have focused on copy number variations (insertions/deletions) (CNVs) of sequences ranging in size from ca. 1-k bases to several mega bases. Coming population studies using sequencing approaches will likely give a more complete understanding of the extent of CNV in nature. However, based on currently available evidence, it has been estimated that about 13% of the human genome varies in copy number

---

Communicated by E. Nigg

P. Stenberg · J. Larsson (✉)  
Department of Molecular Biology, Umeå University,  
90187 Umeå, Sweden  
e-mail: jan.larsson@molbiol.umu.se

P. Stenberg  
Computational Life Science Cluster (CLiC), Umeå University,  
Umeå, Sweden



**Fig. 1** Schematic illustration of CNV (copy number variation). **a** Chromosomal CNV; chromosomes 1, 2, and 3 are in diploid, trisomic, and monosomic states, respectively. **b** Segmental CNV; chromosome 1 represents the wild-type condition, while chromosomes 2 and 3 illustrate

segmental aneuploidies. Chromosome 2 shows a deletion leading to a segmental haploid condition, and chromosome 3 shows a duplication leading to a segmental triploid condition. Note that the duplication (*shaded box*) may be located on any chromosome

and the per-locus mutation rate seems to be higher for CNVs compared to SNPs (Stankiewicz and Lupski 2010).

It is clear that eukaryotic (and hence human) evolution has been significantly shaped by CNVs. Therefore, CNVs presumably do not necessarily have strong negative effects on fitness. On the other hand, changes in numbers of gene copies should generally be proportionally reflected in the RNA levels. Thus, since higher organisms invest substantial effort in maintaining appropriate expression levels for balancing their complex physiological processes, changes in copy number should theoretically be selected against. However, considering the importance of CNVs for evolution, we also expect systems to have evolved that compensate for the differences in gene doses accompanying CNVs, and that these conflicting requirements must be balanced.

### Relationship between gene dose and transcription responses

Until recently, we have had very little insight into the ways that changes in numbers of chromosomes or parts of chromosomes affect gene expression levels, in large part because obtaining accurate measurements of genome-wide expressional effects of CNVs is difficult. Thus, to elucidate the effects of CNVs on gene expression, we need good model systems. Fortunately, we now have techniques to accurately measure global gene expression, and *Drosophila melanogaster* represents a good and well-annotated model system. Therefore, this review mainly focuses on recent discoveries in *Drosophila*.

In *D. melanogaster*, a vast number of well-defined deletions and duplications are available, and new ones can easily be induced in any region of interest. These chromosomal rearrangements can then be observed in isolation or in combination with other rearrangements. Bridges (1935) divided the *D. melanogaster* genome into 102 regions based on chromosomal banding patterns.

These regions have a median size of just over one mega basepairs, and deletions spanning more than one such region (about 1% of the genome) are usually lethal (Lindsley et al. 1972), although when one copy of a single gene is deleted or mutated, one rarely detects any phenotypic effect. The reduction in viability or lethality seen when many genes are present in only one copy may be due to several factors. It could be due to the additive effects of many slightly dosage-sensitive genes, the breakdown of gene networks, or the uncovering of recessive lethal alleles. In general, organisms seem to tolerate extra copies of genomic regions much better than a reduced number of copies (Lindsley et al. 1972).

Few attempts have been made to study the effects of copy number variation on transcription levels in any detail. However, some attempts have been made in human aneuploidies (see for instance FitzPatrick et al. (2002)). Such studies generally examine changes in the expression of all genes with altered copy levels. However, genes that are inactive under the test conditions and genes with expression levels below the detection level will appear to be unaffected by the copy number, complicating attempts to assess the variations and their biological significance. For instance, Gupta et al. (2006) showed that a threefold change in copy number of a genomic region in flies only resulted in a ~1.4-fold change in average mRNA levels, but non-expressed genes were included in this study, thus the results are difficult to interpret. Later, Stenberg et al. (2009) showed that both single-copy and three-copy genomic regions are expressed more closely to wild-type levels than would be expected if mRNA levels correlated perfectly to gene dose. The effect was detectable even after taking inactive genes into account, and haploid regions seem to be more strongly buffered (expressed at 64% of wild type, diploid levels) than regions in three copies (expressed at 146% of wild-type levels). Interestingly, genes that are not ubiquitously expressed in all tissues are more strongly buffered (~70%

of wild-type levels) when in one copy, and ubiquitously expressed genes are more strongly buffered (~138% of wild-type levels) when in three copies. Recently, similar effects were observed by McAnally and Yampolsky (2010). In addition, RNA-Seq and DNA-seq analysis has shown that >43 Mbp of the genome of *Drosophila* S2 cells line is aneuploid (Zhang et al. 2010). The cited authors found that the copy number varies from one to five, and they detected a general buffering of transcript levels. These findings imply that aneuploid regions are buffered at the RNA level. The observations by Stenberg et al. (2009) implied the involvement of a general buffering system acting on the aneuploidy region rather than feedback regulation of individual genes at a gene-by-gene level. Two possible explanations of these findings are that activators or repressors recruited by feedback-regulated genes could locally spread to neighboring genes, thus causing a general buffering effect for aneuploid regions, or currently unknown buffering mechanisms may recognize and target aneuploid regions of the genome.

The findings summarized above show that buffering of aneuploid regions does not restore wild-type expression levels completely. This could be due to the evolutionarily conflicting nature of such systems. On one hand, chromosomal rearrangement is clearly an important driving force of evolution, hence cells and organisms need to be able to cope with them. On the other hand, a perfect buffering system would risk destabilizing the genomic integrity since the selection against aneuploidy would be lowered. Evidently, the organization of genes along chromosomes is far from random. Expressional clustering is common, for a review, see Hurst et al. (2004), and there is also evidence of functional clustering, see for instance Petkov et al. (2005). Many mechanisms, acting at several levels, could be involved in the buffering of aneuploid regions. Clearly,

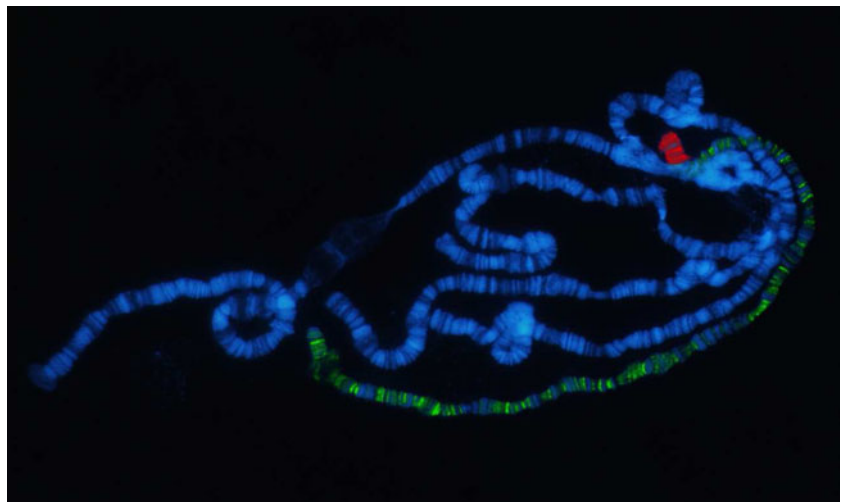
buffering at the RNA level occurs, and a yeast selection study recently showed that mutations inducing ubiquitin-proteasomal degradation reduces the deleterious effects of extra copies of entire chromosomes, suggesting that buffering also acts at the protein turnover level (Torres et al. 2010). Thus, a number of mechanisms working at different levels can probably reduce the deleterious effects of chromosomal rearrangements, and we are just beginning to elucidate these mechanisms and their effects.

There are also two well-studied examples of buffering of entire chromosomes. These are the dosage compensation of the single male X chromosome in *Drosophila* and the compensation of the heterochromatic fourth chromosome (Fig. 2). Both are reviewed below, and we use here the term compensation (rather than buffering) for these systems as their evolution has solved specific dosage and expression problems associated with distinct parts of the genome. In contrast, the buffering system reviewed above is general. However, many mechanisms might still be shared between buffering and the two compensation examples.

### Dosage compensation

In many species, sex determination is associated with heteromorphic sex chromosomes, for example, the X and Y chromosome pair in mammals, which has arisen from autosomes. A Y chromosome can evolve when a gene on a chromosome acquires the ability to determine the male sex. This newly formed chromosome, the proto-Y, will only pass through a single sex. Mutations that favor the male sex are predicted to be accumulated on the proto-Y chromosome, which will eventually suppress recombination, leading to the further accumulation of mutations and mobile elements and eventually loss of genes and degeneration (Charlesworth 1996). The formation of heteromor-

**Fig. 2** Immunostaining of polytene chromosomes in *Drosophila melanogaster* reveals components of the two chromosome-wide targeting and gene regulatory systems. The MSL-C targeting the entire male X chromosome is visualized by an antibody recognizing MSL3 (green). The fourth chromosome is targeted by POF (red) and DNA is counterstained with DAPI (blue)



phic sex chromosomes has occurred independently several times in evolution, and their formation is accompanied by an expression problem that must be solved by the co-evolution of mechanisms that compensate this unbalance. Since most genes on the X chromosome should be expressed at the same amounts in males and females, the gradual degeneration of the Y chromosome will select for systems that compensate for the different doses of the genes located on the X chromosome, as reviewed by Larsson and Meller (2006), Mank (2009) and Vicoso and Bachtrog (2009). This dosage-compensating system should restore the balance between expression from the X chromosome and the autosomes and also compensate for the imbalance of X chromosome between males and females. Here, we focus on *Drosophila*, in which chromosome evolution and dosage compensation have been studied in detail.

The gradual degeneration of the Y chromosome is compatible with an emerging need for dosage compensation at a gene-by-gene level. However, it is not known whether the degeneration of the Y chromosome resulted from a gradual accumulation of slightly deleterious mutations scattered along the chromosome or largely from strongly deleterious mutations of individual genes. These possibilities have different implications for the selective forces promoting dosage compensation, since the former should theoretically favor compensation mechanisms acting on blocks of genes, while the latter should favor mechanisms acting on individual genes (Vicoso and Bachtrog 2009). The required up-regulation of genes in the heterogametic sex together with tolerance of overexpression in the female sex, will determine the levels of expression regulation that are most effective for evolutionary change. Dosage compensation mechanisms can potentially evolve at diverse levels of expression regulation, but in principle, we would expect them to be as sex-specific as possible and only act on active genes. An ideal compensatory system should not cause general de-repression of all X-linked genes. We delineate some of the levels that could potentially be targets for evolution of dosage-compensation mechanisms and discuss the support for them below.

A progressive degeneration of the proto-Y chromosome will inevitably lead to evolutionary pressure at all levels of expression to compensate for the losses of functional gene copies, which should intuitively develop initially (as the aneuploidy emerges) largely on a gene-by-gene basis. An early view in line with this hypothesis was that *cis*-acting elements close to individual genes are important for this response (Baker et al. 1994). For individual genes, e.g., the *white* gene, such *cis*-acting elements have indeed been reported (Qian and Pirrotta 1995). Following the identification of the specific genes on the X chromosome that are dosage compensated in genome-wide investigations in combination with extensive sequence analysis, no unambiguous evidence for *cis*-acting elements for individual

genes on this chromosome has been reported (Alekseyenko et al. 2006; Dahlsveen et al. 2006; Gilfillan et al. 2006; Alekseyenko et al. 2008). However, although no *cis*-acting elements have been identified as a general feature of the dosage-compensated genes, significant overrepresentation of a sequence motif targeted by the DREF (DNA replication element factor) protein has been reported in the 5' region of X chromosomal genes targeted by the male-specific lethal complex (MSL-C) (Legube et al. 2006). The MSL-C is a ribonucleoprotein complex that targets X chromosome genes in males, and its targeting correlates to and can explain most of the observed dosage-compensating effects on the male X chromosome, as reviewed by Straub and Becker (2007), Gelbart and Kuroda (2009), and Hallacli and Akhtar (2009). The complex consists of five proteins, each required for male viability (MSL1, MSL2, MSL3, MLE, and MOF), together with two redundant non-coding RNAs: *roX1* and *roX2*. The MSL-C only forms in males, due to the sex-restricted expression of MSL2, where it binds within the gene body of active X chromosome genes and the acetyltransferase MOF acetylates H4K16. This acetylation becomes strongly enriched in X chromosomal genes of males, and the prevailing models assume that enrichment of H4K16ac leads to decompaction of the chromatin fiber, which then increases gene transcription. Recent data suggest that the MSL-C also has intrinsic properties that constrain the activation potential of MOF, resulting in the required twofold activation of target genes (Prestel et al. 2010b). It should be noted that in *msl2* mutants or in RNAi-mediated knock-downs of *msl2* or *mof*, the male X chromosome expression decrease to ~75% of wild-type expression, suggesting that MSL-C does not mediate the entire twofold activation (Hamada et al. 2005; Deng et al. 2009; Zhang et al. 2010).

The prevailing model of the targeting is that the MSL-C is recruited in a sequence-dependent manner to 150–300 high-affinity MSL recognition elements (MRE) (Alekseyenko et al. 2008). This is followed by a spreading to neighboring genes, which depends on the local MSL complex concentration (Dahlsveen et al. 2006), X chromosome location (Gorchakov et al. 2009), and active transcription (Sass et al. 2003; Larschan et al. 2007). The mechanisms of the MSL-C have been reviewed in detail (Gelbart and Kuroda 2009; Hallacli and Akhtar 2009; Prestel et al. 2010a) and will not be further discussed here. However, an important aspect in this context is the level of expression regulation at which MSL-C acts. It was proposed by Lucchesi (1998) that the MSL-C elevates expression output by facilitating transcription elongation. Genome-wide mappings have shown that the MSL-C and MSL-C-dependent H4K16 acetylation is enriched within gene bodies and biased to their 3' ends (Alekseyenko et al. 2006; Gilfillan et al. 2006; Alekseyenko et al. 2008; Kind et al. 2008; Prestel et al. 2010b). These observations support

the view that MSL-C's main level of action is at transcription elongation, which recently also gained experimental support (Larschan et al. 2011). It should be noted that in principle, the MSL-C may very well act at more than one level.

Although the MSL-C can explain a large extent of the observed dosage compensation in *Drosophila*, and its main level of action is likely to be in transcription elongation, we still need to consider the likelihood that evolutionary pressures have also acted at other levels of expression regulation. We may, for example, consider male X chromosome-specific processes acting on splicing, mRNA release, and post-processing even if no specific evidence for these possibilities have been presented as yet. Furthermore, eukaryotic genomes are spatially organized and sub-nuclear positioning may have important functions for accessibility to transcription machineries and positioning in relation to nuclear pores, the primary channels in the nuclear envelope through which hnRNP molecules (for instance) pass. It has been shown that certain NUP proteins, e.g., NUP98, bind active genes (Capelson et al. 2010; Kalverda et al. 2010). If the binding of NUPs is important for expression output and their binding is sequence-dependent, this would also be a level subject to evolutionary pressure. Although current data are conflicting (see Grimaud and Becker 2009), a biochemical association between NUP proteins and the MSL-C has been suggested, and the nucleoporins Nup153 and Megator have proposed roles in dosage compensation (Mendjan et al. 2006; Vaquerizas et al. 2010).

In principle, evolution of dosage compensation could also act at the level of mRNA stability and decay. Although this possibility has not been investigated in *Drosophila* as yet, in mammals, it has been suggested that autosomal genes are more likely to be subject to nonsense-mediated mRNA decay (NMD) than X chromosomal genes. This is suggested to have a role in the increased X/A expression ratio observed in mammals (Yin et al. 2009). Finally, equalization of expression output from the single male X chromosome to that of the two X chromosomes in females may also involve mechanisms that influence their translation efficiency and the activities of the encoded proteins.

Clearly, most of the observed dosage compensation can be explained by the binding of MSL-C to the gene bodies of X chromosome genes in males and the associated acetylation of H4K16, which most likely facilitates elongation. Nevertheless, as aneuploidies emerge, e.g., via proto-Y chromosome degeneration, evolutionary pressure will act at many levels. Importantly, MSL-C targets gene bodies and new blocks of genes introduced to the X chromosome are relatively quickly adapted by MSL-C targeting (Marin et al. 1996). We speculate that MSL-C targeting may be considered as a secondary adaptation to the monosomic state. In contrast, as aneuploidies emerge as a consequence of Y chromosome degeneration, adaptation to the dose changes seems likely to be based more

on the regulation of individual genes. There may therefore be two major stages in the evolution of dosage compensation mechanisms. The first may be a primary adaptation, in which selection acts at all levels listed above and solves the problems associated with the emerging monosomic condition at a gene-by-gene level. Different genes will have very different requirements for compensation, thus the selection pressure for their compensation will vary. In the next stage, a uniform system may evolve, e.g., the MSL-C that compensates for the changed doses of these emerging set of genes. A likely scenario is that the MSL-C originated from a general buffering system that subsequently evolved X chromosome specificity. This may also explain the notion that upon the addition of (for instance) a “new” X chromosomal arm, MSL-C is recruited, but only when the homolog has degenerated (Bone and Kuroda 1996; Marin et al. 1996). The involvement of more than one system in dosage compensation is also evident from the fact that the MSL-C only supports a 1.35-fold rather than the full, required twofold. Interestingly, combination of this “MSL-C-dependent” effect with the 1.4-fold increase that seems to be general for all aneuploid regions provides very close to the required twofold effect (Stenberg et al. 2009; Prestel et al. 2010a; Zhang et al. 2010). The best evidence for a link between general buffering and the evolution of chromosome-wide gene regulation comes from the studies of POF and the fourth chromosome.

### Autosome-specific gene regulation

For a long time, chromosome-specific targeting was only recognized for sex chromosomes and, accordingly, chromosome-specific gene regulation was considered only to function in dosage compensation. However, in *D. melanogaster*, a second chromosome-wide targeting system was described following the discovery of the protein Painting of fourth (POF) (Larsson et al. 2001). POF specifically targets, or “paints”, the fourth chromosome in *D. melanogaster*, both male and female. Since the fourth chromosome is an autosome, the discovery of POF serves as a starting point to extend the exploration of chromosome-wide regulatory functions to autosomes.

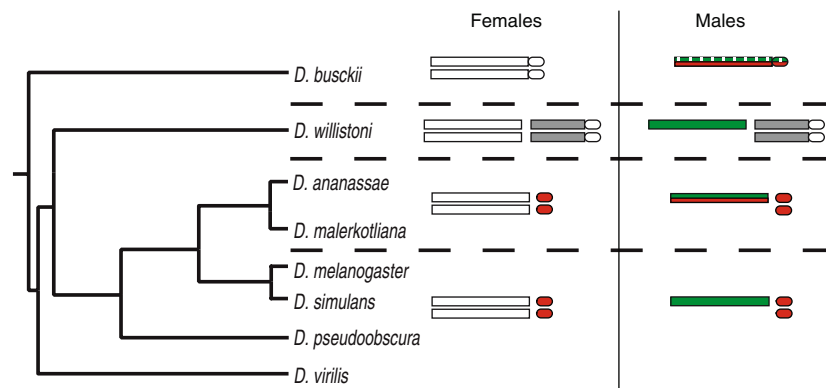
For comparing chromosomes between different *Drosophila* species and to consider the evolution of chromosome-wide gene regulation mechanisms, it is convenient to introduce Muller's naming of chromosomes. In Muller's scheme, the chromosome arms are denoted elements A, B, C, D, E, and F (Muller 1940). By definition, the X chromosome of *D. melanogaster* is named element A, chromosome arms 2L, 2R, 3L, and 3R are named element B, C, D, E, respectively, and the fourth chromosome in *D. melanogaster* is named element F. The major chromosome arms in *D. melanogaster*, i.e., elements A–E, are of similar

length while the fourth chromosome is much shorter and thus also sometimes designated the dot chromosome. The presence of the F element as a unique small dot chromosome is surprisingly well conserved within the genus *Drosophila* (Ashburner et al. 2005). In a few described species, the F element is fused to one of the major chromosome arm. In *Drosophila willistoni*, the F element is fused to element E (Papacit and Juan 1998), and in *Drosophila busckii* and *Scaptodrosophila lebanonensis*, the F element is fused to element A and thus is part of the X chromosome in these two species (Krivshenko 1955; Papacit and Juan 1998). This may actually represent the ancestral state, i.e., the F element may originally have been part of the X chromosome (Fig. 3).

### Evolution of the fourth chromosome

In most *Drosophila* species examined to date, the male X chromosome is targeted by the MSL-C (Bone and Kuroda 1996; Marin et al. 1996) and the F element is targeted by POF in both males and females (Larsson et al. 2004). These two systems represent the two known chromosome-specific targeting and gene regulatory systems in *Drosophila*. The POF protein, similarly to the MSL-C, decorates an entire chromosome by binding to interbands (Fig. 2). There is evidence suggesting a relationship between the X chromosome and the F element. The most striking indication of such a relationship is probably that in *D. busckii*, the F element is located at the base of the X and the Y chromosomes, and there is no free F element (dot chromosome) in this species. Instead, the X chromo-

some consists of the F element and the A element separated by the nucleolus organizer (NO). In *D. busckii*, a euchromatic-banded region in polytene chromosomes is also present on the Y chromosome, suggested to be an F element complement (Krivshenko 1952; Krivshenko 1955; Krivshenko 1959). Further support for the suggested relationship between the F element and the X chromosomes is provided by the minor role in sex determination that has been described for the fourth chromosome. Sexual fate in *Drosophila* is predicted by the ratio of X chromosomes to autosomes, or more specifically, the number of X chromosomes relative to the number of precellular nuclear divisions (Cline and Meyer 1996; Marin and Baker 1998; Erickson and Quintero 2007). In *D. melanogaster*, increased dosage of the fourth chromosome will, in contrast to increased dosage of the other autosomes, shift 2X:3A intersex individuals towards female development, while decreased dosage will shift intersex individuals towards male development (Bridges 1925; Fung and Gowen 1960). Hence, in terms of sex determination, the fourth chromosome behaves more like an X chromosome than an autosome. In flies with three copies of the fourth chromosome, an increased frequency of X chromosome non-disjunction has been observed, suggesting that the fourth chromosome has an intrinsic propensity to pair with the X in meiosis (Sandler and Novitski 1956). Finally, flies with only one copy of the fourth chromosome, i.e., haplo-4 flies, are viable and fertile. This is in marked contrast to the effects of haploidy of the autosomes, but similar to those of the X chromosome. However, a few



**Fig. 3** Phylogenetic relationships of studied species with simplified karyotypes of males and females. The X chromosome is indicated as a *rectangle* and the F element as an *ellipse*. MSL-C targets are indicated in *green* and POF targets in *red* (Bone and Kuroda 1996; Marin et al. 1996; Larsson et al. 2001; Larsson et al. 2004). In *D. willistoni*, the F element is fused to an autosome (*gray rectangle*). The conservation of POF binding to the F element in species such as *D. melanogaster*, *D. pseudoobscura*, and *D. virilis*, strongly indicates that this autosome-specific regulatory system has an important function. The integration

of the F element in the X chromosome of *D. busckii* suggests that it originated as part of the X chromosome and that POF was part of an ancient dosage compensation system. The interrupted *green* on the *D. busckii* male X chromosome indicates that this chromosome is enriched in H4K16ac, but MSL-C has not been detected on it. In *D. ananassae* and *D. malerkotliana*, both POF and MSL-C are present on the male X chromosome (*red/green*), but only POF on the F element in both sexes

cases are known when autosomal deletions are haplosufficient even if they uncover a region that in size and gene number is comparable to the euchromatic part of chromosome 4. In summary, we can conclude that several lines of evidence support a relationship between the F element and the X chromosome, and although the most parsimonious explanation is that the F element originated from the X chromosome, this still remains elusive.

### POF binding in different *Drosophila* species

As well as the binding of the MSL-C, the binding properties of POF have been studied in an evolutionary context, and the findings are summarized in Fig. 3. Initially, POF was found to bind specifically to the fourth chromosome in *D. melanogaster*, and thus identified as an autosome-specific targeting protein (Larsson et al. 2001). This autosome-specific binding is conserved in evolution, as demonstrated by the F element specificity (also detected, for instance, in *Drosophila pseudoobscura* and *Drosophila virilis*). The conservation of the autosome-specific properties of POF indicates that it confers selective advantage, and thus that this autosome-specific targeting is functional. In *D. busckii*, in which the F element is part of the X chromosome, but also in *Drosophila ananassae* and *Drosophila malerkotliana* where the F element is a separate chromosome, POF shows male-specific targeting of the X chromosome. In these three species, POF also colocalizes with H4K16ac, which is indicative of a relationship to X chromosome dosage compensation (Larsson et al. 2004). In *D. busckii*, none of the MSL proteins have so far been detected on the male X chromosome, but enrichment of H4K16ac on the male X chromosome suggests that a similar mechanism is involved (Marin et al. 1996; Larsson et al. 2004). In *D. ananassae*, POF colocalizes perfectly with the MSL-C on the male X chromosome and decorates the F element both in males and females. Notably, the only examined species in the genus *Drosophila* in which POF has not been detected is *D. willistoni*. Here, the F element is part of one of the major autosome arms as it is fused to element E. It is not known how long ago this F+E fusion occurred, but the F element is still retained as a unique block in *D. willistoni*. Taken together, these findings show that the binding of POF to the autosomic F element is conserved, and thus represents an autosome-specific targeting mechanism with an autosome-specific function. Furthermore, the binding to the male X together with the colocalization of POF to the MSL-C and/or the male X chromosome-enrichment of H4K16ac modification suggest a connection to dosage-compensation functions. Thus, it seems likely that POF originated as part of an ancient dosage-compensation mechanism.

### The fourth chromosome of *D. melanogaster*

The fourth chromosome in *D. melanogaster* is an odd chromosome in many respects. It is late replicating (Barigozzi et al. 1966; Zhimulev et al. 2003) and under normal conditions, it exhibits very low levels of meiotic recombination (Hochman 1976; Sandler and Szauter 1978; Wang et al. 2002; Ashburner et al. 2005; Arguello et al. 2010). The banded region seen in polytene chromosomes corresponding to cytogenetic bands 101E–102F appears as a mosaic of unique sequences interspersed with repetitive DNA with a high content of transposable elements (Miklos et al. 1988; Pimpinelli et al. 1995; Locke et al. 1999a; Locke et al. 1999b; Kaminker et al. 2002; Stenberg et al. 2005; Slawson et al. 2006; Leung et al. 2010). The fourth chromosome is also enriched in histone modifications and chromatin-associated proteins that are enriched in heterochromatin. For example, the heterochromatin protein HP1 is enriched on the fourth chromosome, as are specific histone modification markers of heterochromatin, e.g., methylated H3K9 (Eissenberg et al. 1992; Czermin et al. 2002; Schotta et al. 2002). In principle, the entire fourth chromosome can be regarded as the “green-chromatin” according to the definition by van Steensel and coworkers (Filion et al. 2010). In line with these heterochromatic, and thus gene-silencing properties, reporter genes inserted on the fourth chromosome often show a partially silenced, variegated expression (Wallrath and Elgin 1995; Wallrath et al. 1996; Sun et al. 2000). The fourth chromosome consists of a 3–4-Mb proximal part (Locke and McDermid 1993) and a banded, sequenced part that is approximately 1.3 Mb long. Interestingly, despite its heterochromatic properties, the sequenced part of the chromosome includes 92 genes; hence, it has similar gene density to the major chromosome arms. Moreover, the average level of expression of genes on this chromosome is comparable to, or even higher than, that of genes on the other chromosomes (Haddrill et al. 2008) (Johansson, Stenberg and Larsson unpublished). The sequence composition of the fourth chromosome is also abnormal in comparison to the other chromosomes in *D. melanogaster*. In fact, its structure with scattered repetitive elements is reminiscent of the organization of mammalian chromosomes. Thus, the genes located on the fourth chromosome appear to be adapted to function in this hostile, repressive environment.

### Chromatin environment of the fourth chromosome

In line with its enrichment of repeated sequences, the fourth chromosome is targeted by typical heterochromatin histone modifications, e.g., methylated H3K9. At least three H3K9-specific methyl transferases mediate the methylation of H3K9

in *Drosophila*: Su(var)3–9, SETDB1, and G9a. The most thoroughly studied of these transferases is *Su(var)3–9*, which was initially found in genetic screens together with *Su(var)2–5* (encoding HP1) as suppressors of position-effect variegation (Schotta et al. 2003). Methylated H3K9 stabilizes the binding of HP1 to chromatin, and this modification is generally mediated by *Su(var)3–9*. However, in *Su(var)3–9* mutants, levels of pericentric H3K9me and pericentric HP1 are reduced, except on the fourth chromosome where levels of H3K9me appear to be unaltered (Czermin et al. 2002; Schotta et al. 2002; Ebert et al. 2004). This apparent discrepancy has recently been explained by findings that the *Drosophila* *Setdb1* controls H3K9 methylation in the fourth chromosome (Seum et al. 2007; Tzeng et al. 2007). Hence, chromosome 4 is similar to pericentric heterochromatin in its methylated H3K9 enrichment, but this enrichment is mediated by different proteins. These K9-methylated H3 tails are targeted by the N-terminal chromo-domain of HP1 (Bannister et al. 2001; Lachner et al. 2001; Nakayama et al. 2001; Jacobs and Khorasanizadeh 2002). The proposed model for HP1-mediated chromatin condensation postulates that HP1 forms a dimer through its C-terminal shadow domains, and the two chromo-domains in the dimer link two adjacent nucleosomes. The spreading is proposed to be mediated by interaction between HP1 and the histone methyltransferase *Su(var)3–9* (or *Setdb1*), which methylates nearby H3K9 and thus propagates spreading of the methylation and HP1 interaction. Although HP1's interaction with chromatin is mainly considered to act through methylated H3K9, it should be stressed that HP1 has also been shown to interact with high affinity to the H3 histone-fold on oligonucleosomal arrays (Nielsen et al. 2001; Dialynas et al. 2006; Billur et al. 2010). This interaction may be represented by a small, but stable, fraction of HP1 associated with chromatin.

### Chromosome-specific regulation of the fourth chromosome

Despite the repressive nature of the fourth chromosome, the genes located on it are properly expressed, possibly partly because of the activity of the chromosome 4-specific protein POF. In contrast to HP1, which binds the fourth chromosome as well as pericentric heterochromatin, POF is highly specific for the banded part of the fourth chromosome. In this region, POF and HP1 colocalize close to perfectly at the resolution afforded by polytene chromosome staining (Johansson et al. 2007a). The binding of POF and HP1 is interdependent, hence POF binding to the fourth is strongly reduced in *HP1* mutants and HP1 binding is reduced in *Pof* mutants (Johansson et al. 2007a). Although POF binds to the banded, gene rich region of the fourth chromosome, the results suggest that POF requires heter-

chromatic pressure for its targeting. This hypothesis has been corroborated by studies showing that a translocated fourth chromosome will not be targeted by POF unless the proximal heterochromatic region is present, and conditions favor heterochromatin formation (e.g., low temperature and/or absence of a Y chromosome) (Johansson et al. 2007a).

Both POF and HP1 have been mapped on the fourth chromosome to high resolution using the ChIP-chip technique (Johansson et al. 2007b). At ChIP-chip resolution, POF and HP1 colocalization within the transcribed region of chromosome 4 genes is close to perfect. The enrichment levels of these two factors show a linear correlation, and the HP1 and POF binding levels also positively correlate with transcription levels (Johansson et al. 2007b). In addition to the binding within genes, HP1 shows a distinct promoter-binding peak in almost all genes on the fourth chromosome if they are expressed (Johansson et al. 2007a). HP1 targets on the fourth chromosome can thus be divided into two distinct sets: one within the transcribed region and one in the promoters. POF, on the other hand, is only seen within the transcribed gene region. Reduction of *HP1* levels causes increased expression of chromosome 4 genes, whereas loss of *Pof* causes a decrease in chromosome 4 expression (Johansson et al. 2007b). In *D. melanogaster*, haplo-4 flies are viable and fertile. However, the expression of non-ubiquitously expressed genes on the fourth chromosome has been shown to be compensated by POF in these flies, and the lack of this compensation causes lethality (Stenberg et al. 2009). In contrast, there is very little or no compensation for ubiquitously expressed genes. The POF-mediated compensation of chromosome 4 genes is more effective than the observed buffering of segmental aneuploidies. However, these two systems show many striking similarities, e.g., in the differences in their effects on differentially expressed and ubiquitously expressed genes (Stenberg et al. 2009). A balancing model has been proposed suggesting that POF and HP1 stimulate and repress, respectively, gene expression from the fourth chromosome, thus together they provide a buffer that prevents excessive transcription fluctuation. However, if HP1 and POF are interdependent, why does their mutation have opposite effects on transcription output? It is tempting to speculate that the removal of POF causes a reduction in HP1 levels, but that some HP1, for example in the promoter peak, remains and the repressive function dominates. In the opposite situation, when HP1 is reduced, some POF may still be present, and the stimulating function may dominate. An intriguing observation is that silenced transgenes on the fourth chromosome are preferentially located in HP1-rich regions, whereas non-silenced transgenes are preferentially located in regions with low amounts of HP1 (Stenberg et al. 2009). On chromosome 4, high levels of HP1 are observed



in transcribed (and thus active) regions; so why are transgenes silenced? Maybe the HP1 protein will spread into these transgenes, but since these sequences do not have a chromosome 4 origin, POF will not target them. If so, the silencing properties of HP1 would dominate. This possibility has not been tested as yet.

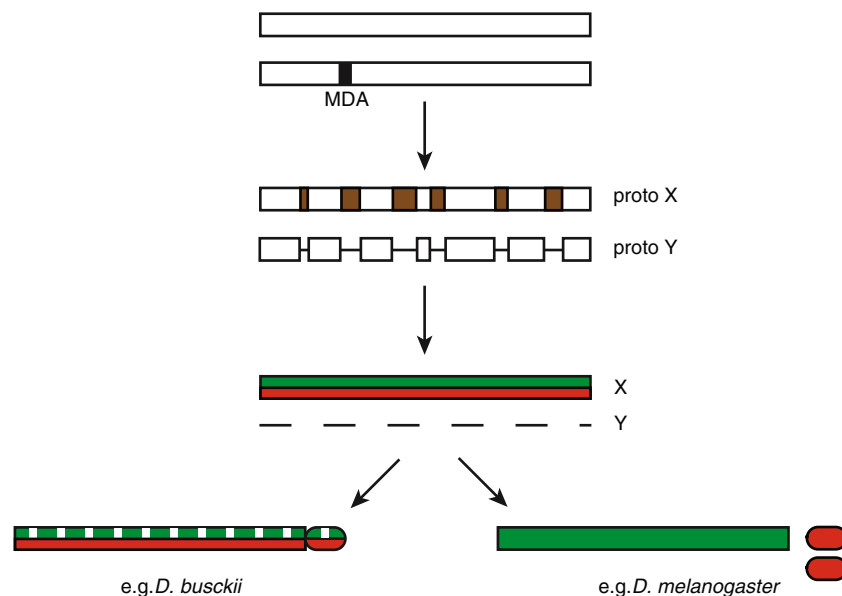
The molecular mechanism of POF-mediated stimulation of gene transcription remains elusive. However, since POF has an RNA-binding domain, it is likely that POF is recruited through nascent RNA from actively transcribing chromosome 4 genes. Whether the stimulation of expression occurs through the promotion of transcription, more efficient splicing, or enhanced export, remains to be determined. The prevailing model for the repressive nature of HP1 is that it targets methylated H3K9, forms a dimer, and links two adjacent nucleosomes. However, it should be stressed that HP1 is not solely linked to gene repression. It is also known to be required for transcription of some genes located within heterochromatic regions (Hearn et al. 1991; Clegg et al. 1998; Lu et al. 2000), it has been shown to be associated with active genes in euchromatic regions, and gene expression analyses suggest that in some cases, HP1 stimulates transcription (Piacentini et al. 2003; Cryderman et al. 2005). It was also recently shown that HP1 associates with transcripts of hundreds of euchromatic *Drosophila* genes and that HP1 positively regulates these genes (Piacentini et al. 2009). A suggested model proposes that HP1's main function is nucleic acid compaction, which can act on both DNA and RNA. Compaction of DNA through

binding to methylated H3K9 will cause repression, while compaction of RNA via interaction with hnRNP proteins may stabilize the messenger and support gene expression (Piacentini et al. 2009).

Why is the binding of POF and HP1 interdependent? So far, no evidence of physical association between POF and HP1 has been reported, but POF and Setdb1 have been shown to interact in vitro (Tzeng et al. 2007). Thus, POF may provide an adaptor system linking histone marks via HP1 and Setdb1 to pre-mRNA, in similar fashion to MRG15 and PTB (Luco et al. 2010; Luco et al. 2011).

### POF in testes

POF is expressed at low levels in most, if not all cells, in both males and females. However, it is expressed at very high levels in testes. A number of appealing hypotheses could explain this strong germline expression. First, the X chromosome is also dosage-compensated in the male germline, but no mechanism for this compensation has been yet been characterized (Bachiller and Sánchez 1986; Rastelli and Kuroda 1998; Gupta et al. 2006). The strong relationship between POF and dosage compensation suggests that POF may have an X chromosome dosage-compensation function in the germline. However, expression microarray analysis has shown that the expression of the X chromosome is not affected in *Pof* mutant testes (Stenberg et al. 2009). Secondly, POF's activities allow



**Fig. 4** Putative origins of the X chromosome, the F element, and the compensatory systems POF (*red*) and MSL-C (*green*). MDA indicates the male determining allele. Upon degeneration of the proto-Y chromosome, the corresponding regions of the proto-X chromosome will require buffering of haploid regions, indicated by brown boxes.

Parts of these general buffering systems later evolved into dosage compensation systems (POF and MSL-C). The interrupted green on *D. busckii* male X chromosome indicates that this chromosome is enriched in H4K16ac, but MSL-C has not been detected on it. The F element is indicated by an ellipse

proper expression of genes in the heterochromatic environment of the fourth chromosome. In the male germline, the genes on the Y chromosome should be expressed, and these genes are also located in a heterochromatic environment. However, expression analysis has shown that in the testes, POF mainly influences expression of chromosome 4 genes, and *Pof* mutants are not male sterile (Stenberg et al. 2009). Third, the fourth chromosome lacks recombination, which is also a hallmark of the *Drosophila* male germline. POF may therefore act as a suppressor of recombination. However, although this possibility has not been fully excluded, *Pof* mutants lack recombination both of the fourth chromosome and in the male germline. Thus, it seems that stimulation of expression of genes located on chromosome 4 is the main function of POF in both germline and somatic cells.

### Origin of MSL- and POF-mediated dosage compensation mechanisms

An obvious question to address is why there is a chromosome-specific gene regulatory mechanism for the fourth chromosome. Possible clues are that POF is required for haplo-4 survival, and somatic elimination of the fourth chromosome is relatively frequent in *D. melanogaster* (Mohr 1932; Ashburner et al. 2005), although the fourth chromosome is normally present in diploid form in both males and females. Thus, POF may be important for sufficient expression of chromosome 4 genes in haplo-4 cells. It seems unlikely that the sole selective advantage of the system is aiding the survival of haplo-4 cells or individuals. However, there is a strong relationship between the X and the fourth chromosomes, and POF may indeed originate from a buffering/dosage compensating system.

As the proto-Y gradually degenerated, initial compensation of the monosomic X regions was probably provided by general buffering mechanisms that evolved due to the need to cope with general dosage differences (Fig. 4). These mechanisms would have acted at different levels of expression regulation with varying levels of compensation precision. During the course of Y degeneration, some of these systems would have had to become specific for the X chromosome to avoid the genomic instability associated with over-efficient general buffering. The resulting X specificity would be advantageous to males but would be accompanied with risks of overexpression in females. Participants in the X chromosome-specific systems that evolved include MSL, and most likely POF, since POF is observed on the male X chromosome in some species (Fig. 3). The F element, as part of the X chromosome, is most likely the ancestral state, and after the split of the ancestral *D. busckii* (Fig. 3), the F element detached from the X.

In *D. busckii*, in which the F element has been retained as part of the X chromosome, both POF and MSL compensate for the difference between sexes in copy number of genes on the X chromosome. In contrast, in other species (e.g., *D. melanogaster*, in which the F element is a unique autosome), the two systems appear to have diverged and now target different chromosomes and meet distinct compensatory needs. However, an intrinsic propensity of the MSL-C to target the F element is still observable in *roX1 roX2* mutants (Meller and Rattner 2002; Deng and Meller 2006). The retention of POF on the F element allows correct expression, despite the heterochromatic nature of this chromosome. This may in turn have caused a functional constraint of the genes located on the fourth chromosome, which may explain the conservation of the F element as a unique dot chromosome in many *Drosophila* species.

### Shaping the genome

Compensatory and buffering systems provide molecular mechanisms that can counter problems associated with changes in gene copy number that emerge during the course of evolution and allow potentially beneficial variations in genome structure. Clearly, CNV is an important evolutionary driving force, and a delicate balance is needed between compensating for variations in gene doses (thus permitting variability) and selection against excessively destabilizing variation. Equally clearly, general buffering systems can be modified in ways that meet compensatory requirements and may provide templates for the evolution of a wide range of specific regulatory mechanisms. Considering the abundance of functional constraints on gene organization, clustering and order, and the conservation of many syntenic regions, it seems likely that several autosome-specific targeting systems in other organisms remain to be discovered.

**Acknowledgements** The authors thank Prof Anssi Saura and Dr Yuri Schwartz for comments on the manuscript.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

### References

- Alekseyenko AA, Larschan E, Lai WR, Park PJ, Kuroda MI (2006) High-resolution ChIP-chip analysis reveals that the *Drosophila* MSL complex selectively identifies active genes on the male X chromosome. *Genes Dev* 20:848–857
- Alekseyenko AA, Peng S, Larschan E, Gorchakov AA, Lee OK, Kharchenko P, McGrath SD, Wang CI, Mardis ER, Park PJ,

- Kuroda MI (2008) A sequence motif within chromatin entry sites directs MSL establishment on the *Drosophila* X chromosome. *Cell* 134:599–609
- Arguello JR, Zhang Y, Kado T, Fan C, Zhao R, Innan H, Wang W, Long M (2010) Recombination yet inefficient selection along the *Drosophila melanogaster* subgroup's fourth chromosome. *Mol Biol Evol* 27:848–861
- Ashburner M, Golic KG, Hawley RS (2005) *Drosophila*: a laboratory handbook. Cold Spring Harbor Laboratory, Cold Spring Harbor
- Bachiller D, Sánchez L (1986) Mutations affecting dosage compensation in *Drosophila melanogaster*: effects in the germline. *Dev Biol* 118:379–384
- Baker BS, Gorman M, Marín I (1994) Dosage compensation in *Drosophila*. *Annu Rev Genet* 28:491–521
- Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410:120–124
- Barigozzi C, Dolfini S, Fraccaro M, Raimondi GR, Tiepolo L (1966) In vitro study of the DNA replication patterns of somatic chromosomes of *Drosophila melanogaster*. *Exp Cell Res* 43:231–234
- Billur M, Bartunik HD, Singh PB (2010) The essential function of HP1 $\beta$ : a case of the tail wagging the dog? *Trends Biochem Sci* 35:115–123
- Bone JR, Kuroda MI (1996) Dosage compensation regulatory proteins and the evolution of sex chromosomes in *Drosophila*. *Genetics* 144:705–713
- Bridges CB (1925) Sex in relation to chromosomes and genes. *Am Nat* 59:127–137
- Bridges CB (1935) Salivary chromosome maps with a key to the banding of the chromosomes of *Drosophila melanogaster*. *J Hered* 26:60–64
- Capelson M, Liang Y, Schulte R, Mair W, Wagner U, Hetzer MW (2010) Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes. *Cell* 140:372–383
- Charlesworth B (1996) The evolution of chromosomal sex determination and dosage compensation. *Curr Biol* 6:149–162
- Clegg NJ, Honda BM, Whitehead IP, Grigliatti TA, Wakimoto B, Brock HW, Lloyd VK, Sinclair DA (1998) Suppressors of position-effect variegation in *Drosophila melanogaster* affect expression of the heterochromatic gene *light* in the absence of a chromosome rearrangement. *Genome* 41:495–503
- Cline TW, Meyer BJ (1996) Vive la difference: male vs females in flies vs worms. *Annu Rev Genet* 30:637–702
- Cryderman DE, Grade SK, Li Y, Fanti L, Pimpinelli S, Wallrath LL (2005) Role of *Drosophila* HP1 in euchromatic gene expression. *Dev Dyn* 232:767–774
- Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V (2002) *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal polycomb sites. *Cell* 111:185–196
- Dahlsveen IK, Gilfillan GD, Shelest VI, Lamm R, Becker PB (2006) Targeting determinants of dosage compensation in *Drosophila*. *PLoS Genet* 2:e5
- Deng X, Koya SK, Kong Y, Meller VH (2009) Coordinated regulation of heterochromatic genes in *Drosophila melanogaster* males. *Genetics* 182:481–491
- Deng X, Meller VH (2006) *roX* RNAs are required for increased expression of X-linked genes in *Drosophila melanogaster* males. *Genetics* 174:1859–1866
- Dialynas GK, Makatsori D, Kourmouli N, Theodoropoulos PA, McLean K, Terjung S, Singh PB, Georgatos SD (2006) Methylation-independent binding to histone H3 and cell cycle-dependent incorporation of HP1 $\beta$  into heterochromatin. *J Biol Chem* 281:14350–14360
- Ebert A, Schotta G, Lein S, Kubicek S, Krauss V, Jenuwein T, Reuter G (2004) *Su(var)* genes regulate the balance between euchromatin and heterochromatin in *Drosophila*. *Genes Dev* 18:2973–2983
- Eissenberg JC, Morris GD, Reuter G, Hartnett T (1992) The heterochromatin-associated protein HP1 is an essential protein in *Drosophila* with dosage-dependent effects on position effect variegation. *Genetics* 131:345–352
- Erickson JW, Quintero JJ (2007) Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in *Drosophila*. *PLoS Biol* 5:e332
- Filion GJ, van Bommel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, van Steensel B (2010) Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* 143:212–224
- FitzPatrick DR, Ramsay J, McGill NI, Shade M, Carothers AD, Hastie ND (2002) Transcriptome analysis of human autosomal trisomy. *Hum Mol Genet* 11:3249–3256
- Fung STC, Gowen JW (1960) Role of autosome-IV in *Drosophila melanogaster* sex balance. *Genetics* 45:988–989
- Gelbart ME, Kuroda MI (2009) *Drosophila* dosage compensation: a complex voyage to the X chromosome. *Development* 136:1399–1410
- Gilfillan GD, Straub T, de Wit E, Greil F, Lamm R, van Steensel B, Becker PB (2006) Chromosome-wide gene-specific targeting of the *Drosophila* dosage compensation complex. *Genes Dev* 20:858–870
- Gorchakov AA, Alekseyenko AA, Kharchenko P, Park PJ, Kuroda MI (2009) Long-range spreading of dosage compensation in *Drosophila* captures transcribed autosomal genes inserted on X. *Genes Dev* 23:2266–2271
- Grimaud C, Becker PB (2009) The dosage compensation complex shapes the conformation of the X chromosome in *Drosophila*. *Genes Dev* 23:2490–2495
- Gupta V, Parisi M, Sturgill D, Nuttall R, Doctolero M, Dudko OK, Malley JD, Eastman PS, Oliver B (2006) Global analysis of X-chromosome dosage compensation. *J Biol* 5:3
- Hadrill PR, Waldron FM, Charlesworth B (2008) Elevated levels of expression associated with regions of the *Drosophila* genome that lack crossing over. *Biol Lett* 4:758–761
- Hallacli E, Akhtar A (2009) X chromosomal regulation in flies: when less is more. *Chromosome Res* 17:603–619
- Hamada FN, Park PJ, Gordadze PR, Kuroda MI (2005) Global regulation of X chromosomal genes by the MSL complex in *Drosophila melanogaster*. *Genes Dev* 19:2289–2294
- Hearn MG, Hedrick A, Grigliatti TA, Wakimoto BT (1991) The effect of modifiers of position-effect variegation on the variegation of heterochromatic genes of *Drosophila melanogaster*. *Genetics* 128:785–797
- Hochman B (1976) The fourth chromosome of *Drosophila melanogaster*. In: Ashburner M, Novitski E (eds) *The genetics and biology of Drosophila*. Academic, New York, pp 903–928
- Hurst LD, Pál C, Lercher MJ (2004) The evolutionary dynamics of eukaryotic gene order. *Nat Rev Genet* 5:299–310
- Jacobs SA, Khorasanizadeh S (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science* 295:2080–2083
- Johansson AM, Stenberg P, Bernhardsson C, Larsson J (2007a) Painting of fourth and chromosome-wide regulation of the 4th chromosome in *Drosophila melanogaster*. *EMBO J* 26:2307–2316
- Johansson AM, Stenberg P, Pettersson F, Larsson J (2007b) POF and HP1 bind expressed exons, suggesting a balancing mechanism for gene regulation. *PLoS Genet* 3:e209
- Kalverda B, Pickersgill H, Shloma VV, Fornerod M (2010) Nucleoporins directly stimulate expression of developmental and cell-cycle genes inside the nucleoplasm. *Cell* 140:360–371
- Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, Patel S, Frise E, Wheeler DA, Lewis SE, Rubin GM, Ashburner M, Celniker SE (2002) The transposable elements of the

- Drosophila melanogaster* euchromatin: a genomics perspective. *Genome Biol* 3:RESEARCH0084
- Kind J, Vaquerizas JM, Gebhardt P, Gentzel M, Luscombe NM, Bertone P, Akhtar A (2008) Genome-wide analysis reveals MOF as a key regulator of dosage compensation and gene expression in *Drosophila*. *Cell* 133:813–828
- Krivshenko JD (1952) A cytogenetic study of the Y chromosome in *Drosophila busckii*. *Genetics* 37:500–518
- Krivshenko JD (1955) A cytogenetic study of the X chromosome of *Drosophila busckii* and its relation to phylogeny. *Proc Natl Acad Sci USA* 41:1071–1079
- Krivshenko JD (1959) New evidence for the homology of the short euchromatic elements of the X and Y chromosomes of *Drosophila busckii* with the microchromosome of *Drosophila melanogaster*. *Genetics* 44:1027–1040
- Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410:116–120
- Larschan E, Alekseyenko AA, Gortchakov AA, Peng S, Li B, Yang P, Workman JL, Park PJ, Kuroda MI (2007) MSL complex is attracted to genes marked by H3K36 trimethylation using a sequence-independent mechanism. *Mol Cell* 28:121–133
- Larschan E, Bishop EP, Kharchenko PV, Core LJ, Lis JT, Park PJ, Kuroda MI (2011) X chromosome dosage compensation via enhanced transcriptional elongation in *Drosophila*. *Nature* 471:115–118
- Larsson J, Meller VH (2006) Dosage compensation, the origin and the afterlife of sex chromosomes. *Chromosome Res* 14:417–431
- Larsson J, Chen JD, Rasheva V, Rasmuson Lestander A, Pirrotta V (2001) Painting of fourth, a chromosome-specific protein in *Drosophila*. *Proc Natl Acad Sci USA* 98:6273–6278
- Larsson J, Svensson MJ, Stenberg P, Mäkitalo M (2004) Painting of fourth in genus *Drosophila* suggests autosome-specific gene regulation. *Proc Natl Acad Sci USA* 101:9728–9733
- Legube G, McWeeney SK, Lercher MJ, Akhtar A (2006) X-chromosome-wide profiling of MSL-1 distribution and dosage compensation in *Drosophila*. *Genes Dev* 20:871–883
- Leung W, Shaffer CD, Cordonnier T, Wong J, Itano MS, Slawson Tempel EE, Kellmann E, Desruisseau DM, Cain C, Carrasquillo R, Chusak TM, Falkowska K, Grim KD, Guan R, Honeybourne J, Khan S, Lo L, McGaha R, Plunkett J, Richner JM, Richt R, Sabin L, Shah A, Sharma A, Singhal S, Song F, Swope C, Wilen CB, Buhler J, Mardis ER, Elgin SC (2010) Evolution of a distinct genomic domain in *Drosophila*: comparative analysis of the dot chromosome in *Drosophila melanogaster* and *Drosophila virilis*. *Genetics* 185:1519–1534
- Lindsley DL, Sandler L, Baker BS, Carpenter AT, Denell RE, Hall JC, Jacobs PA, Miklos GL, Davis BK, Gethmann RC, Hardy RW, Steven AH, Miller M, Nozawa H, Parry DM, Gould-Somero M (1972) Segmental aneuploidy and the genetic gross structure of the *Drosophila* genome. *Genetics* 71:157–184
- Locke J, Howard LT, Aippersbach N, Podemski L, Hodgetts RB (1999a) The characterization of *DINE-1*, a short, interspersed repetitive element present on chromosome and in the centric heterochromatin of *Drosophila melanogaster*. *Chromosoma* 108:356–366
- Locke J, McDermid H (1993) Analysis of *Drosophila* chromosome four by pulse field electrophoresis. *Chromosoma* 102:718–723
- Locke J, Podemski L, Roy K, Pilgrim D, Hodgetts R (1999b) Analysis of two cosmid clones from chromosome 4 of *Drosophila melanogaster* reveals two new genes amid an unusual arrangement of repeated sequences. *Genome Res* 9:137–149
- Lu BY, Emtage PC, Duyf BJ, Hilliker AJ, Eissenberg JC (2000) Heterochromatin protein 1 is required for the normal expression of two heterochromatin genes in *Drosophila*. *Genetics* 155:699–708
- Lucchesi JC (1998) Dosage compensation in flies and worms: the ups and downs of X-chromosome regulation. *Curr Opin Genet Dev* 8:179–184
- Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T (2010) Regulation of alternative splicing by histone modifications. *Science* 327:996–1000
- Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T (2011) Epigenetics in alternative pre-mRNA splicing. *Cell* 144:16–26
- Mank JE (2009) The W, X, Y and Z of sex-chromosome dosage compensation. *Trends Genet* 25:226–233
- Marin I, Baker BS (1998) The evolutionary dynamics of sex determination. *Science* 281:1990–1994
- Marin I, Franke A, Bashaw GJ, Baker BS (1996) The dosage compensation system of *Drosophila* is co-opted by newly evolved X chromosomes. *Nature* 383:160–163
- McAnally AA, Yampolsky LY (2010) Widespread transcriptional autosomal dosage compensation in *Drosophila* correlates with gene expression level. *Genome Biol* 2:44–52
- Meller VH, Rattner BP (2002) The *roX* genes encode redundant male-specific lethal transcripts required for targeting of the MSL complex. *EMBO J* 21:1084–1091
- Mendjan S, Taipale M, Kind J, Holz H, Gebhardt P, Schelder M, Vermeulen M, Buscaino A, Duncan K, Mueller J, Wilm M, Stunnenberg HG, Saumweber H, Akhtar A (2006) Nuclear pore components are involved in the transcriptional regulation of dosage compensation in *Drosophila*. *Mol Cell* 21:811–823
- Miklos GLG, Yamamoto MT, Davies J, Pirrotta V (1988) Microcloning reveals a high frequency of repetitive sequences characteristic of chromosome four and the  $\beta$ -heterochromatin of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 85:2051–2055
- Mohr OL (1932) Genetic and cytological proof of somatic elimination of the fourth chromosome in *Drosophila melanogaster*. *Genetics* 17:60–80
- Muller HJ (1940) Bearings on the *Drosophila* work on systematics. In: Huxley J (ed) *The new systematics*. Clarendon, Oxford, pp 185–268
- Nakayama J, Rice JC, Strahl BD, Allis CD, Grewal SI (2001) Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* 292:110–113
- Nielsen AL, Oulad-Abdelghani M, Ortiz JA, Remboutsika E, Chambon P, Losson R (2001) Heterochromatin formation in mammalian cells: interaction between histones and HP1 proteins. *Mol Cell* 7:729–739
- Papacit M, Juan E (1998) Fate of dot chromosome genes in *Drosophila willistoni* and *Scaptodrosophila lebanonensis* determined by in situ hybridization. *Chromosome Res* 6:49–54
- Petkov PM, Graber JH, Churchill GA, DiPetrillo K, King BL, Paigen K (2005) Evidence of a large-scale functional organization of mammalian chromosomes. *PLoS Genet* 1:e33
- Piacentini L, Fanti L, Berloco M, Perrini B, Pimpinelli S (2003) Heterochromatin protein 1 (HP1) is associated with induced gene expression in *Drosophila* euchromatin. *J Cell Biol* 161:707–714
- Piacentini L, Fanti L, Negri R, Del Vescovo V, Fatica A, Altieri F, Pimpinelli S (2009) Heterochromatin protein 1 (HP1a) positively regulates euchromatic gene expression through RNA transcript association and interaction with hnRNPs in *Drosophila*. *PLoS Genet* 5:e1000670
- Pimpinelli S, Berloco M, Fanti L, Dimitri P, Bonaccorsi S, Marchetti E, Caizzi R, Caggese C, Gatti M (1995) Transposable elements are stable structural components of *Drosophila melanogaster* heterochromatin. *Proc Natl Acad Sci USA* 92:3804–3808
- Prestel M, Feller C, Becker PB (2010a) Dosage compensation and the global re-balancing of aneuploid genomes. *Genome Biol* 11:216
- Prestel M, Feller C, Straub T, Mitlöhner H, Becker PB (2010b) The activation potential of MOF is constrained for dosage compensation. *Mol Cell* 38:815–826

- Qian S, Pirrotta V (1995) Dosage compensation of the *Drosophila white* gene requires both the X chromosome environment and multiple intragenic elements. *Genetics* 139:733–744
- Rastelli L, Kuroda MI (1998) An analysis of maleless and histone H4 acetylation in *Drosophila melanogaster* spermatogenesis. *Mech Dev* 71:107–117
- Sandler L, Novitski E (1956) Evidence for genetic homology between chromosomes I and IV in *Drosophila melanogaster*, with a proposed explanation for the crowding effect in triploids. *Genetics* 41:189–193
- Sandler L, Szauter P (1978) The effect of recombination-defective meiotic mutants on fourth-chromosome crossing over in *Drosophila melanogaster*. *Genetics* 90:699–712
- Sass GL, Pannuti A, Lucchesi JC (2003) Male-specific lethal complex of *Drosophila* targets activated regions of the X chromosome for chromatin remodeling. *Proc Natl Acad Sci USA* 100:8287–8291
- Schotta G, Ebert A, Krauss V, Fischer A, Hoffmann J, Rea S, Jenuwein T, Dorn R, 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
- Schotta G, Ebert A, Dorn R, Reuter G (2003) Position-effect variegation and the genetic dissection of chromatin regulation in *Drosophila*. *Semin Cell Dev Biol* 14:67–75
- Schrider DR, Hahn MW (2010) Gene copy-number polymorphism in nature. *Proc Biol Sci* 277:3213–3221
- Seum C, Reo E, Peng H, Rauscher FJ, Spierer P, Bontron S (2007) *Drosophila* SETDB1 is required for chromosome 4 silencing. *PLoS Genet* 3:e76
- Slawson EE, Shaffer CD, Malone CD, Leung W, Kellmann E, Shevchek RB, Craig CA, Bloom SM, Bogenpohl J, Dee J, Morimoto ET, Myoung J, Nett AS, Ozsolak F, Tittiger ME, Zeug A, Pardue ML, Buhler J, Mardis ER, Elgin SC (2006) Comparison of dot chromosome sequences from *D. melanogaster* and *D. virilis* reveals an enrichment of DNA transposon sequences in heterochromatic domains. *Genome Biol* 7:R15
- Stankiewicz P, Lupski JR (2010) Structural variation in the human genome and its role in disease. *Annu Rev Med* 61:437–455
- Stenberg P, Lundberg LE, Johansson AM, Rydén P, Svensson MJ, Larsson J (2009) Buffering of segmental and chromosomal aneuploidies in *Drosophila melanogaster*. *PLoS Genet* 5:e100302
- Stenberg P, Pettersson F, Saura AO, Berglund A, Larsson J (2005) Sequence analysis of chromosome identity in three *Drosophila* species. *BMC Bioinform* 6:1–17
- Straub T, Becker PB (2007) Dosage compensation: the beginning and end of generalization. *Nat Rev Genet* 8:47–57
- Sun FL, Cuaycong MH, Craig CA, Wallrath LL, Locke J, Elgin SC (2000) The fourth chromosome of *Drosophila melanogaster*: interspersed euchromatic and heterochromatic domains. *Proc Natl Acad Sci USA* 97:5340–5345
- Torres EM, Dephoure N, Panneerselvam A, Tucker CM, Whittaker CA, Gygi SP, Dunham MJ, Amon A (2010) Identification of aneuploidy-tolerating mutations. *Cell* 143:71–83
- Tzeng TY, Lee CH, Chan LW, Shen CK (2007) Epigenetic regulation of the *Drosophila* chromosome 4 by the histone H3K9 methyltransferase dSETDB1. *Proc Natl Acad Sci USA* 104:12691–12696
- Vaquerez JM, Suyama R, Kind J, Miura K, Luscombe NM, Akhtar A (2010) Nuclear pore proteins Nup153 and Megator define transcriptionally active regions in the *Drosophila* genome. *PLoS Genet* 6:e1000846
- Vicoso B, Bachtrog D (2009) Progress and prospects toward our understanding of the evolution of dosage compensation. *Chromosome Res* 17:585–602
- Wallrath LL, Elgin SC (1995) Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev* 9:1263–1277
- Wallrath LL, Guntur VP, Rosman LE, Elgin SC (1996) DNA representation of variegating heterochromatic P-element inserts in diploid and polytene tissues of *Drosophila melanogaster*. *Chromosoma* 104:519–527
- Wang W, Thornton K, Berry A, Long M (2002) Nucleotide variation along the *Drosophila melanogaster* fourth chromosome. *Science* 295:134–137
- Yin S, Deng W, Zheng H, Zhang Z, Hu L, Kong X (2009) Evidence that the nonsense-mediated mRNA decay pathway participates in X chromosome dosage compensation in mammals. *Biochem Biophys Res Commun* 383:378–382
- Zhang Y, Malone JH, Powell SK, Periwal V, Spana E, Macalpine DM, Oliver B (2010) Expression in aneuploid *Drosophila* S2 cells. *PLoS Biol* 8:e1000320
- Zhimulev IF, Belyaeva ES, Makunin IV, Pirrotta V, Volkova EI, Alekseyenko AA, Andreyeva EN, Makarevich GF, Boldyreva LV, Nanayev RA, Demakova OV (2003) Influence of the *SuUR* gene on intercalary heterochromatin in *Drosophila melanogaster* polytene chromosomes. *Chromosoma* 111:377–398