2006

The dot chromosome of Drosophila: insights into chromatin states and their change over evolutionary time

Nicole Riddle
Washington University in St. Louis

Sarah C.R. Elgin
Washington University in St. Louis, selgin@wustl.edu

Chromosome Research (2006) 14:405–416
DOI: 10.1007/s10577-006-1061-6

The dot chromosome of *Drosophila*: Insights into chromatin states and their change over evolutionary time

Nicole C. Riddle & Sarah C. R. Elgin*
Department of Biology, Washington University, One Brookings Dr, St. Louis, MO 63130, USA; Tel: +1-314-935-5348; Fax: +1-314-935-5125; E-mail: selgin@biology.wustl.edu
**Correspondence*

Key words: chromatin, chromosome evolution, *Drosophila*, F element, heterochromatin

Abstract

Historically, chromatin has been subdivided into heterochromatin, transcriptionally inactive regions that remain densely packaged throughout the cell cycle, and euchromatin, transcriptionally active regions that take on a diffuse appearance as the cell enters interphase. The banded portion of the small fourth chromosome (dot chromosome) of *Drosophila melanogaster* is unusual in exhibiting many characteristics of heterochromatic domains, and at the same time maintaining a gene density typical of euchromatin. Similar to genes embedded in pericentric heterochromatin, many of the dot chromosome genes have adapted to a heterochromatic environment. Little is known about the regulation of these genes and less about their evolution in a chromatin context. Interestingly, most of the genes from the *D. melanogaster* fourth chromosome remain clustered on a small chromosome throughout the genus *Drosophila*; yet the dot chromosome appears euchromatic in some species, such as *D. virilis*. Existing genomic sequence data allow an exploration of the underlying differences in DNA sequence organization between species. Here we review the available data describing the dot chromosome, which derives primarily from *D. melanogaster*. With its unusual and changing nature, the dot chromosome in the genus *Drosophila* provides a unique opportunity for the examination of transitions between chromatin states during evolution.

Introduction

Study of the genus *Drosophila,* and especially *Drosophila melanogaster,* began early in the 20th century. Fruit flies were of great value to early geneticists for two main reasons: visible mutations were easily generated in the fruit fly in large numbers, and the study of their karyotype was greatly facilitated by the large polytene chromosomes present in the salivary glands. Comparative studies of a variety of species soon revealed that chromosome organization and karyotypes exhibited a wide array of forms within the genus *Drosophila*. In 1940 Müller published a nomenclature for the six chromosome segments of *D. melanogaster* that can be applied to the other species in the genus. In this nomenclature each chromosome arm is assigned a letter (A–F), with the fourth chromosome corresponding to the F element (Müller 1940). While the six elements (represented in *D. melanogaster* as the chromosome arms X, 2L, 2R, 3L, 3R, and 4) are present in some form or another in all the *Drosophila* species studied, the chromosome number in the various species differs (see Figure 1). For instance, the genome of *D. melanogaster* is distributed among four chromosome pairs, while in *D. virilis* there are

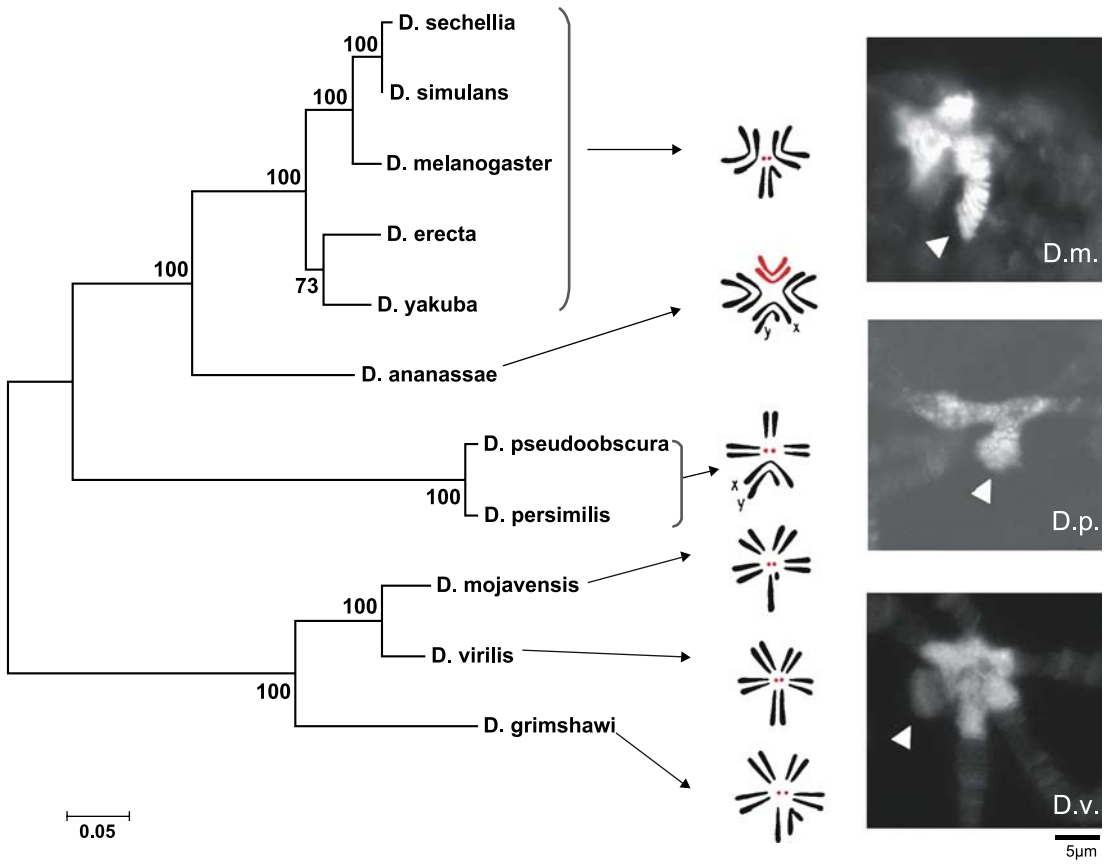

Figure 1. The dot chromosome in the genus *Drosophila*. A neighbor-joining tree based on an amino-acid alignment of POF is shown in the left part of the diagram (Scale: no. of amino acid substitutions per site). Bootstrap support is given based on 1000 replicates. In the middle section, karyotypes are shown from representative species [adapted from Patterson & Stone 1952 and http://insects.eugenes.org/species/ (*D. grimshawi*)]. The pictures on the right show staining of the dot chromosomes of different species with antibodies specific for HP1. D.m. = *D. melanogaster*; D.p. = *D. pseudoobscura*; D.v. = *D. virilis*. Note the negative results for *D. virilis*. The white arrowhead points to the dot chromosome.

six chromosomes, one for each of the six elements defined by Müller.

In *D. melanogaster*, chromosome 4 is the smallest of the autosomes. It corresponds to Müller's F element and is often simply called 'the dot' (Ashburner *et al.* 2005). Interestingly, despite their separate evolutionary histories, many *Drosophila* species have maintained an equivalent of the *D. melanogaster* dot chromosome. Among the 12 *Drosophila* species being sequenced by the *Drosophila* Genome Project, only *D. willistoni* does not possess a recognizable F element. In the majority of the remaining species, the F element is similar to that of *D. melanogaster*, maintained as a small dot chromosome. The only exception in this species group is *D. ananassae*, where the F element is a larger chromosome with two distinct arms (Kikkawa 1938). In this species it appears that the F element has acquired (among other things) an rRNA gene array; however, it remains haplo-sufficient (Kikkawa 1938; see discussion below). Although the dot is the smallest of the *D. melanogaster* chromosomes, it has ignited considerable interest as biologists seek to understand its unique characteristics and behavior, which will be discussed in this review.

The dot chromosome exhibits both heterochromatic and euchromatic features

One of the most unusual aspects of the *D. melanogaster* dot chromosome is its chromatin architec-

Table 1. Heterochromatic and euchromatic characteristics of the banded portion of chromosome 4 in *D. melanogaster*

Heterochromatin	Euchromatin	Chromosome 4
Late replicating	Early replicating	H
No recombination	Recombination	H
Transcriptionally silent	Transcriptionally active	E
HP1 staining	No HP1 staining	H/E
High in H3K9me	Low in H3K9me	H
Low gene density	High gene density	E
High repeat density	Low repeat density	H
Induces PEV	No PEV induction	H/E
Not polytenized	Polytenized	E

E: Euchromatin-like; H: heterochromatin-like.

ture. In contrast to the other autosomes the chromatin of the gene-rich portion of the dot exhibits characteristics of both euchromatin and heterochromatin (see Table 1 for a summary). Heterochromatin and euchromatin represent two different forms of chromatin structure. Heterochromatin is often found associated with centromeres, pericentric regions, telomeres, rRNA gene repeats and, in yeast, with mating type loci. These heterochromatic regions of the genome exhibit little or no recombination, while recombination rates within euchromatin are high. Heterochromatin is also late replicating, while euchromatic regions of the genome replicate earlier in S phase. The DNA sequences found within heterochromatin are characterized by an abundance of repeats of various classes. In contrast to euchromatin, which contains most of the genes, heterochromatin in general is gene-poor, and its structure often confers a transcriptionally silent state on any euchromatic gene placed in proximity by rearrangement or transposition. This silencing effect is thought to occur due to the 'dense' packaging of the DNA into a more regular nucleosome array; this form of packaging may exclude transcription factors and other activating elements from DNA sequences within heterochromatin.

In addition, the biochemical nature of heterochromatin is quite distinct from that of euchromatin. Histones within euchromatin carry so-called 'active' marks, such as acetylation of histones H3 and H4, that convey an open chromatin structure. Nucleosomes within heterochromatic regions contain deacetylated histones, and histone H3 methylated at lysine K9 (H3K9me). Heterochromatin is often associated with heterochromatin protein 1 (HP1), a non-histone chromosomal protein that specifically interacts with H3K9me and imposes – in conjunction with other proteins – a closed/silent chromatin structure (for a review on heterochromatin see e.g. Grewal & Elgin 2002).

When the karyotype of *D. melanogaster* was initially described, dark staining associated with the specific chromatin structure of heterochromatin was not reported for the dot chromosome (Hochman 1976). However, subsequent studies soon showed that one arm, and in fact the majority, of chromosome 4 displayed some characteristics of heterochromatin. A comparison between the appearance of the dot in meitotic chromosome spreads and in polytene chromosome squashes first suggested that chromosome 4 of *D. melanogaster* contains a heterochromatin component other than the centromeric regions. While two small arms are visible in meitotic chromosome spreads, only one of the chromosome arms was found to be amplified in the polytene chromosomes of the salivary gland. This finding seemed to indicate that much of the chromosome is heterochromatic and, similar to the centromeric heterochromatin, does not undergo polytenization (Hochman 1976).

Lack of recombination is another feature of heterochromatin that is exhibited by the entire fourth chromosome of *D. melanogaster*. Early geneticists mapped several visible mutations to the fourth chromosome, some of which appeared to be essential. This finding showed that, despite its peculiar characteristics, the dot chromosome carries a number of important genes. Work with these mutations, especially various mapping efforts, led to the realization that, under most conditions, recombination is suppressed on the entire fourth chromosome. While increased rearing temperature (30°C) or heat shock

could be used to induce recombination on the fourth chromosome, under natural conditions recombination proved to be exceedingly rare, not to say non-existent (reviewed in Hochman 1976, Ashburner *et al*. 2005). More recent work has documented that in more than 58,000 meioses not a single case of recombination on the fourth chromosome has been observed under the standard rearing conditions for *D. melanogaster* (Sandler & Szauter 1978). Numerous population genetics studies focusing on the dot chromosome have confirmed the extremely low recombination rate, which is reflected in a general lack of genetic variation for this chromosome (Jensen *et al*. 2002, Sheldahl *et al*. 2003, Wang *et al*. 2004). However, the detailed analysis of a 200 kb region demonstrated that, while rare, recombination did occur in natural populations and could lead to localized higher levels of polymorphisms, even on the fourth chromosome (Wang *et al*. 2002). Nonetheless, overall the entire chromosome 4 resembles centromeric chromatin in its lack of recombination. If one considers the close association of the fourth chromosome with the chromocenter – the location of heterochromatin in polytene nuclei – one could argue that potentially the whole of chromosome 4 should be considered heterochromatic in this regard. How the 82 known/predicted genes encoded on the dot chromosome are expressed in this environment is a question of considerable interest.

Shared properties of the dot chromosome and the X chromosome

Another phenomenon that is particular to the fourth chromosome of *D. melanogaster* is the occurrence of aneuploidy. Similar to the X chromosome, the dot chromosome can be present within individuals in a dosage other than two. Individuals carrying only one copy of chromosome 4 (haplo-IV) are viable and fertile, with some fertility defects observed in females. These individuals exhibit a *minute* phenotype, but are otherwise normal (Mohr 1932). Trisomy of chromosome 4 (triplo-IV) is frequently observed in the offspring of triploid females and causes minor phenotypic alterations; however, most individuals remain viable and fertile (Sturtevant 1934, 1936). Tetrasomy of chromosome 4 also occurs, and the viability of these individuals is surprisingly high, reaching approximately 70% relative to diploid individuals (Ashburner *et al*. 2005). In addition to complete aneuploidy, somatic elimination of chromosome 4 is also relatively frequent, and leads to mosaic individuals exhibiting a partial *minute* phenotype (Mohr 1932).

Loss of either chromosome 2 or chromosome 3, or of either of their chromosome arms, causes embryonic lethality, as does trisomy. In contrast, loss as well as gain of copies of chromosome 4 is well tolerated in *D. melanogaster*. It is possible that this finding is due to the small size of chromosome 4 – that, by chance, none of the 82 genes on chromosome 4 is haplo-lethal. (There is at least one locus on chromosome 4 that is haplo-insufficient, as it causes the minute phenotype observed in haplo-IV flies.) When small regions of chromosomes 2 and 3 have been investigated, very few segments show a haplo-lethal phenotype, and only one triplo-lethal region has been identified (Lindsley *et al*. 1972). These data suggest that the viability of haplo-IV and triplo-IV individuals might be due to the small size of the aneuploid chromosome. However, another interpretation of the data is possible, as the X chromosome, like the dot chromosome, can exist in copy numbers other than two in viable and fertile flies. A mechanism that can modulate the regulatory state of the chromosome as a whole may be present in both cases.

Additional insights into the behavior of chromosome 4 have been gained from studies of chromosomal translocations involving the fourth chromosome as one partner, including studies of the segregation of the compound chromosomes. When translocations are induced using X-ray treatment, exchanges between chromosome 4 and the X chromosome are found to be overrepresented compared to other possible chromosome combinations (Sandler & Novitski 1955). In addition to a bias in the recovery of translocations, segregation distortion – involving both chromosome 4 itself and its translocation derivatives – is also often observed. This effect can be illustrated by the studies of Sturtevant in the 1930s concerning segregation distortion in *D. melanogaster* trisomic for chromosome 4. Sturtevant discovered that extraneous copies of chromosome 4 do not segregate randomly among daughter cells. Rather, it appears that the segregation of chromosome 4 is influenced by the segregation of the sex chromosomes, in particular the X chromosome (Sturtevant 1934, 1936, Sandler & Novitski 1955). The additional

fourth chromosome is less likely to go to the same spindle pole as the X chromosome. A similar pattern is observed for compound chromosomes derived from translocations involving chromosome 4. In these individuals the compound fourth chromosome can pair with an X chromosome, leading occasionally to non-disjunction of the X chromosomes. Overall, these classical genetic studies indicate a possible connection between chromosome 4 and the X chromosome, based on their ability to interact with each other during meiosis.

Studies of karyotype evolution lend further credence to a possible connection between the X chromosome and chromosome 4 in the genus *Drosophila*. Inversions and rearrangements appear to have been common during the evolution of the F element. For instance, based on altered fluorescent *in-situ* hybridization staining patterns, it has been postulated that at least three whole arm inversion events occurred for the dot chromosome within the melanogaster subgroup of *Drosophila* (Podemski *et al*. 2001). Among species not included in the sequencing efforts, the F elements of *D. neorepleta* and *D. busckii* bear mention. As in the case of *D. ananassae*, the F element of *D. neorepleta* is no longer a dot chromosome, but appears to have acquired additional sequences, mostly heterochromatic (Sturtevant 1946). Based on the mapping of known *D. melanogaster* fourth chromosome genes in *D. busckii*, it appears that the F element in this species has fused with the X chromosome (Krivshenko 1952, 1955, 1959). These examples establish that the evolution of the dot chromosome reflects a number of different mechanisms, such as chromosome fusion and multiple inversions. Despite the peculiar characteristics that set the dot chromosome apart from the other autosomes, such mechanisms acting on chromosome structure over evolutionary time are shared with the other chromosomes.

The biochemistry of the dot chromosome is distinct from that of the other autosomes

In 2001 a further peculiarity of the dot chromosome of *D. melanogaster* was discovered with the characterization of a locus named *painting of fourth* (*pof*). The product of the *pof* locus is a protein which binds exclusively to the fourth chromosome, based on immunostaining of polytene chromosomes. The POF protein shows little similarity to known proteins, the only identifiable motifs being an RpL29 signature motif, an RRM1 RNA binding domain, a nuclear localization signal, and a coiled coil domain (Larsson *et al*. 2001). *pof* is weakly expressed in the embryo, with little or no maternal contribution, and expression levels increase during development. Expression of the gene is stronger in adults than in larvae, and is stronger in males than in females, mainly due to high expression levels in the testes. Detailed analysis of the staining pattern on polytene chromosomes indicates that while the POF antibody stains the polytenized arm of chromosome 4 in a banded pattern, it does not stain chromosome 4 at the base closest to the chromocenter. Notably, the POF-staining bands do not correspond to the DAPI bright bands. Translocation studies have shown that POF will not associate with other genomic regions translocated onto the fourth chromosome centromere. Proper localization of POF to the fourth chromosome appears to require both the centromere of chromosome 4 and a distal portion of the chromosome arm to initiate binding. Nonetheless, binding of POF to a translocated arm of chromosome 4 can occur *in trans* provided that the distal part of the translocated chromosome 4 is paired with an intact copy of the chromosome (Larsson *et al*. 2001). Hence, this binding pattern suggests a spreading mechanism for assembly of POF-associated chromatin that can act *in trans*.

Besides the male-sex-lethal (MSL) complex, which is involved in dosage compensation of the X chromosome, POF is the only other known protein to be associated with just one chromosome in *Drosophila*. The localization of the MSL complex in male flies is guided by two non-coding RNAs, rox1 and rox2. For proper localization at least one of these RNA species is required, and the RNA appears to recruit the MSL complex to the appropriate chromosome (Franke & Baker 1999). In contrast to results from studies of the dosage compensation complex, RNase treatment of polytene chromosomes does not interfere with the detection of POF bound to chromosome 4 (Larsson *et al*. 2001). Despite this finding, POF might provide another link between the X chromosome and the F element. In *D. busckii* (where the F element has fused with the X chromosome) the POF antibody stains the entire X chromosome in male flies, indicating that the fourth might be

derived from the X and that POF might be connected to the dosage compensation complex (Larsson *et al*. 2001).

Recent work on POF has provided further insights into the evolution and behavior of the F element in *Drosophila* (Larsson *et al*. 2004). Genetic analysis of *pof* has demonstrated that the locus is required for female fertility and for proper development in both sexes, indicating that appropriate packaging and, most likely, expression, of fourth chromosome genes plays an important role in the fruit fly. This hypothesis is confirmed by a comparative study of POF association in a variety of *Drosophila* species. In all but one species POF faithfully stains the F element, and in some cases also the X chromosome. Only in *D. willistoni*, where the F element is fused to one of the autosomes, is no staining observed. Thus, it appears that a functional requirement for POF association with the F element is conserved. Conservation of the protein is shown by the observation that the POF protein from *D. ananassae* can bind to the fourth chromosome in a transgenic line of *D. melanogaster* (Larsson *et al*. 2004). These data suggest that both protein function, as well as the recognition site, on the fourth chromosome have been conserved over several million years.

Mapping chromatin structure at high resolution on the dot chromosome

The perception that the dot chromosome is largely heterochromatic has been reinforced by its peculiar behavior with regard to position effect variegation (PEV). PEV refers to the observation that, when a gene that is normally located in euchromatin is translocated/moved close to heterochromatin, the gene will be silenced in a stochastic manner, resulting in a variegated pattern of expression. Within the tissue in which the gene is normally expressed, certain cell lineages remain active while others are silenced, leading to a mottled appearance of a visual marker. The proportion of silenced versus actively expressing cell lineages is dependent on the position of the translocation – the closer the gene is to pericentric heterochromatin, the more frequently the gene is silenced. The first reports of PEV on the fourth chromosome date back several decades and noted that PEV at the fourth chromosome locus *ci* was 'atypical (reviewed in Hochman 1976).

While PEV was initially discovered as the consequence of a large chromosome inversion, it also affects transgene reporters inserted within or near heterochromatic regions of the genome. One such reporter gene that is commonly used is a *P* element carrying a copy of the eye color gene *white*. In a strain of *D. melanogaster* deficient for *white*, the expression level of this reporter gene depends on its position in the genome. If the *P* element is inserted into a euchromatic domain, *white* will be expressed, resulting in a red eye color phenotype. However, if the reporter is inserted in a heterochromatic domain, variegated eye color results, indicating that the *white* gene has been silenced. In this case the severity of silencing depends on the position in relation to heterochromatin. Thus, these reporter lines are of great value in investigating chromatin structure, as they provide a simple readout of local packaging. Initial studies with this reporter showed that a variegating eye phenotype is observed when the *P* element is inserted into the pericentric heterochromatin, the telomeres, or the fourth chromosome, as confirmed by *in-situ* hybridization of the polytene chromosomes (Wallrath & Elgin 1995). Interestingly, among the large number of insertions mapping to chromosome 4, many occurred in its so-called 'euchromatic' arm (Wallrath & Elgin 1995). This finding again indicates that even the portion of the dot chromosome of *D. melanogaster* that is polytene in salivary gland cells shows properties reminiscent of heterochromatin.

The discovery that locations within the polytene (banded) portion of chromosome 4 could induce PEV led to a number of studies to understand the structure of chromosome 4 in greater detail. Analysis of the chromatin structure of the reporter *P* element in lines exhibiting a variegating phenotype (indicating heterochromatin packaging) has shown reduced accessibility to restriction nucleases and a more regular nucleosome array, indicating alternative packaging at the nucleosome level (Wallrath & Elgin 1995, Sun *et al*. 2001). In addition to the lines with variegating transgenes recovered with insertions on the fourth chromosome, lines with a red eye phenotype mapping to the fourth chromosome were also discovered (Sun *et al*. 2000). In assays using *Xba*I, these transcriptionally active transgenes showed greater accessibility than the silenced, variegated inserts, similar to what is observed at euchromatic loci on the arms of chromosomes 2 and 3. The nucleosome

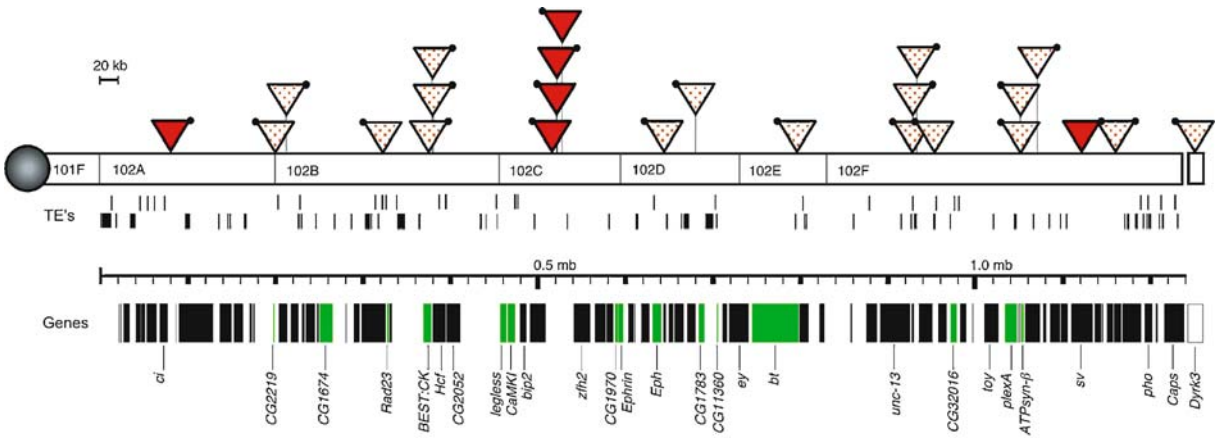

Figure 2. Interspersed chromatin domains on the *D. melanogaster* dot chromosome. Schematic representation of the fourth chromosome of *D. melanogaster*, with the centromere shown on the left. Triangles above the chromosome diagram represent insertion sites of the *P* element *white* reporter construct. Full red triangles mark 'euchromatic' sites where the *white* gene is expressed, and stippled triangles mark heterochromatic sites, where the *white* gene is silenced. Below the chromosome the location of transposable elements (TE) is marked, with the upper row corresponding to the *1360* element. The distribution of genes is shown in the lowest tier, with green representing genes that have been marked by a *P* element insertion. Adapted from Haynes *et al*. (2004).

spacing of these active inserts is also less regular than the pattern seen in variegating inserts in pericentric heterochromatin (Sun *et al*. 2000, 2001). The variegating inserts on the fourth chromosome respond to many *Suppressor of variegation* [*Su(var)*] modifier loci in a manner characteristic of centromeric insertions, but differ from transgenes silenced by insertion into telomeres. Both centromeric and fourth chromosome PEV lines show a loss of silencing in response to *Su(var)2-5* and *Su(var)3-7* mutations, which disrupt the genes coding for HP1 and for a zinc finger protein, respectively (Sun *et al*. 2000). These results suggest that the chromatin environments of pericentric regions and the fourth chromosome are similar with respect to their effects on transgene reporters. However, these two domains are not identical; reporters in pericentric heterochromatin are sensitive to mutations in *Su(var)3-9*, which codes for an H3 methyltransferase, while reporters on the fourth chromosome are not (K. Haynes, personal communication). In addition, antibody staining of H3K9me on polytene chromosomes of *Su(var)3-9* mutants revealed that, while the overall level of H3K9me is reduced, a large amount of H3K9me is still detected on the fourth chromosome (Czermin *et al*. 2002, Schotta *et al*. 2002). Given the high levels of H3K9me observed on the fourth, one can infer that a different H3 methyltransferase must be involved.

The implication of the reporter gene studies is that chromatin domains with euchromatic and heterochromatic characteristics are in close proximity to each other on the fourth chromosome (see Figure 2). The close proximity of the these two types of chromatin can be illustrated by the fact that local deletions or duplications of 5 to 80 kb can result in a transgene reporter on the fourth chromosome switching from a red to a variegated phenotype and vice-versa (Sun *et al*. 2004). Examination of the sequences surrounding the variegating insertions on the fourth chromosome has revealed that, contrary to expectation, most variegating inserts are within 2 kb of an annotated gene. In fact, out of 18 investigated insertion sites, 11 are actually within the transcribed region of a gene (Sun *et al*. 2004). This finding indicates that, while these domains on the fourth chromosome exhibit characteristics of heterochromatin and exert heterochromatic effects on reporters derived from euchromatic genes, they do not appear to block the transcription of endogenous loci. It should be noted, however, that tissue-specific differences may occur, i.e. the chromatin state in the developing eye disc, where the reporter is normally expressed, could differ from the chromatin state in the cell type in which a given fourth chromosome gene is expressed.

However, similar results have been obtained for other so-called heterochromatic genes. In contrast to

popular belief, the pericentric heterochromatin regions of many genomes are not completely devoid of genes (Copenhaver *et al*. 1999, Nagaki *et al*. 2004, Yan *et al*. 2005). The efforts of the heterochromatin sequencing project in *D. melanogaster* have revealed that the pericentromeric regions of the fly genome harbor in excess of 100 genes (Hoskins *et al*. 2002). Few of these genes have been studied in detail, but some have, particularly *rolled* and *light*. Using translocation lines it was discovered that proper *light* expression is dependent on its heterochromatic location. If *light* as well as a number of other heterochromatic genes in close proximity are translocated to euchromatin by a rearrangement, they exhibit a variegated phenotype, indicating that they undergo silencing at this new location (Wakimoto & Hearn 1990). Further evidence for the dependency of heterochromatin genes on their specific chromatin environment comes from experiments with *Su(var)* mutations. *Light* as well as other heterochromatin genes depend on several *Su(var)* loci for proper expression (Hearn *et al*. 1991, Schulze *et al*. 2005). Of particular interest is the finding that HP1 is required for the expression of these genes, as this protein is an integral part of heterochromatin and associates with H3K9me modified nucleosomes. In contrast to the findings for euchromatic reporters, mutations in *Su(var)2-5* (the gene encoding HP1) cause a decrease in expression of heterochromatic genes such as *light*, indicating that the regulation of the two gene classes is quite different. Endogenous heterochromatic loci do not show the same regular nucleosome pattern and decrease in nuclease accessibility that genes translocated into heterochromatin (including the fourth chromosome) exhibit (Sun *et al*. 2001). This has led to the suggestion that the requirement for heterochromatin packaging reflects the organization of regulatory sites, rather than differences in the transcribed region (Eissenberg & Elgin 2000).

DNA sequence organization on the dot chromosome

While observations linking the so-called euchromatic arm of the dot chromosome to a heterochromatin-like structure have been documented as early as the 1930s (e.g. its lack of recombination), only recently have advances in genomics allowed for detailed study of the DNA sequences from this chromosome. The *D. melanogaster* dot chromosome is approximately 4.2 Mb in size, about 3.0 Mb of which consist of repeated sequences surrounding the centromere and making up the short arm (4R; Locke & McDermid 1993). These regions are similar to other heterochromatic regions of the *Drosophila* genome, e.g. the centromeres, in that they contain mainly repeated sequences and are not amplified in polytene chromosomes of the salivary glands. More interesting is the analysis of the banded portion of chromosome 4, the distal 1.2 Mb region which is amplified in polytene chromosomes. As demonstrated by the PEV assays discussed above, this portion of the fourth chromosome also exhibits heterochromatin features. The long arm of chromosome 4 encodes 82 known or predicted genes, and this gene density is similar to that found for the other autosomes. What sets the banded portion of the fourth chromosome apart from the other autosomes is the high frequency of repeated sequences found along this entire arm of chromosome 4 (Locke *et al*. 1999, Bartolome *et al*. 2002). These repeated elements are found at much higher density on the fourth chromosome than in other euchromatic regions, to the extent that the repeat density on the dot chromosome resembles that observed in pericentric regions. Particularly noticeable is the high frequency of the short DINE-1 fragments (Locke *et al*. 1999). In addition, again similar to pericentric regions, repeated elements on the fourth chromosome are found within as well as between genes (Locke *et al*. 1999, Bartolome *et al*. 2002, Hoskins *et al*. 2002).

In the past year, two large-scale sequence comparison studies have been published that specifically focus on chromosome 4. The first sought to identify the sequence characteristics that separate the F element from the remaining chromosomes (Stenberg *et al*. 2005), in particular seeking to identify potential binding sites for POF. In their analysis of the F elements and autosomes from *D. melanogaster*, *D. pseudoobscura* and *D. yakuba*, the investigators compared the frequency of short sequences up to six basepairs using principal components analysis. They identified a nonamer (corresponding to a pair of overlapping hexamers) that matches to a *DINE-1* element and is much more frequent on the F element than on the other chromosomes. It is also possible to distinguish F element exons from exons of genes on

the other chromosomes, indicating that sequence differences exist in exons as well as non-coding sequences. Areas of the fourth chromosome enriched for the nonamer matching to the *DINE-1* element correlated with the banding patterns seen in POF staining, suggesting that the two areas overlap, and that the nonamer element might be involved in recruiting POF protein to the F element. However, the *DINE-1* element and the nonamer cannot be the sole requirement for recruitment of POF, as they also occur at other genomic locations, albeit at lower concentration (Stenberg *et al*. 2005). Interestingly, the *DINE-1* element distribution has been recently compared between *D. melanogaster* and *D. yakuba*. The data suggest that *D. yakuba* has experienced two bouts of *DINE-1* transposition, resulting in two classes of *DINE-1* elements. Of these, only the older class of events shows an association with heterochromatic regions of the genome, while the newer class is evenly distributed along the *D. yakuba* chromosomes. This finding raises the question whether (at least in *D. yakuba*), the subclasses of *DINE-1* play different roles in distinguishing the F element (Yang *et al*. 2005).

A second large-scale sequence comparison study has been conducted using fosmid data from *D. melanogaster* and *D. virilis*. *D. virilis* is of great interest for studies trying to understand the chromatin structure of the F element in the genus *Drosophila*. As in *D. melanogaster*, the F element of *D. virilis* is a small dot-like chromosome. However, the *D. virilis* F element does not appear to be heterochromatic in character. Recombination occurs on the *D. virilis* dot chromosome more frequently than in *D. melanogaster*, and the biochemical nature of the dot chromatin appears to be different from *D. melanogaster*. In polytene chromosome squashes the *D. virilis* dot chromosome arm does not stain with antibodies for HP1 or H3K9me, which are hallmarks of heterochromatin. Thus, it appears that the F element has undergone a change in chromatin state during the approximately 40–60 million years of evolution that separate *D. melanogaster* from *D. virilis* (Slawson *et al*. 2006). Presumably this change in packaging reflects an underlying change in DNA sequence organization.

The sequence comparison includes high-quality finished and annotated sequences from parallel euchromatic regions of *D. virilis* and *D. melanogaster* as well as *ca*. 300 kb of sequence from their respective F elements. Analysis of the genes present on the F element reveals that 27 of 28 genes studied have been maintained on the two dot chromosomes; only one of the dot chromosome genes from *D. melanogaster* is not found on the *D. virilis* F element, and only one gene has been introduced from another chromosome. However, despite the conservation of the genes' location on the dot chromosomes, the gene order is very different on the two chromosomes. Syntenic regions are small, and inversions and other rearrangements appear to have been quite common. Gene density on the F element of both species is similar to that found on each species' other chromosomes. Genes found on either of the dot chromosomes have approximately 2-fold longer introns than genes on the other autosomes. This finding is interesting in that commonly heterochromatin genes, such as those found in pericentric chromatin, have longer introns than euchromatic genes. However, based on immunostaining of polytene chromosomes, the F element of *D. virilis* should be considered euchromatic, while the F element of *D. melanogaster* should be considered heterochromatic. As genes on both chromosomes show the longer intron pattern, this feature might reflect some other characteristic, such as proximity to the chromocenter (Slawson *et al*. 2006).

Another interesting similarity between the F elements of the two species concerns their repeat density. Both F elements contain a large number of repeats with an overall similar density (26% and 23%, for *D. melanogaster* and *D. virilis*, respectively). However, the predominant repeat types differ between the two species. The long arm of the dot chromosome of *D. melanogaster* is depleted of the CA/GT repeat that usually characterizes euchromatic regions. While the F element of *D. virilis* is also depleted of this repeat relative to the species other autosomes, the level of CA/GT repeat on the *D. virilis* F element is approximately 20-fold higher than the level on the *D. melanogaster* dot chromosome. This difference might contribute to the difference in chromatin structure between the euchromatic dot in *D. virilis* and the heterochromatic dot chromosome in *D. melanogaster* (Slawson *et al*. 2006).

Besides the difference in simple sequence repeats, a difference in the transposable elements found on the *D. virilis* and *D. melanogaster* dot chromosomes is observed. In *D. virilis* the DNA transposon class is underrepresented compared to *D. melanogaster*; the

D. virilis dot chromosome has only approximately one-third of the DNA transposons found in *D. melanogaster*. The DNA transposons that are increased in frequency on the *D. melanogaster* dot chromosome include *DINE-1*, which was identified in the study described above as a distinguishing sequence feature of the F element in *D. melanogaster* and *D. yakuba* (Stenberg *et al*. 2005, Slawson *et al*. 2006).

A second DNA transposon identified as overrepresented on the *D. melanogaster* fourth chromosome is the *hoppel* element, also known as *1360* (Slawson *et al*. 2006). While in *D. melanogaster* 4.1% of the examined sequence corresponds to the *1360* element, in *D. virilis* only 0.8% corresponds to *1360* (Slawson *et al*. 2006). This finding is of interest since a previous study found that proximity to a *1360* element had a silencing effect on the *white* reporter gene. Whenever a *P* element carrying a *white* reporter giving a variegating eye phenotype was identified on the fourth chromosome of *D. melanogaster*, in most cases a *1360* element was found to be within 10 kb of the insertion, while no such correlation was found for non-variegating reporters (Haynes *et al*. 2004, Sun *et al*. 2004). Thus, it was suggested that the *1360* element might be a focus for heterochromatin formation. *1360* is highly abundant in the pericentric heterochromatin as well as on the fourth chromosome. The finding that *1360* as well as other DNA transposons are underrepresented on the euchromatic F element of *D. virilis* adds support to this theory.

Future prospects

The sequencing of multiple species in the genus *Drosophila* is providing an important tool for researchers interested in the F element. Such data have been especially useful for studies of karyotype evolution, as well as for studies focusing on the evolution of individual chromosomes. An example of the type of studies that will be facilitated by the forthcoming genomics data is provided by the recent analysis of the evolution of the Y chromosome in *Drosophila* (Carvalho & Clark 2005). Based on a comparison of the genomic locations of Y chromosome genes in *D. pseudoobscura* and *D. melanogaster*, the Y chromosomes in these species appear to be unrelated in origin. Rather, the evidence suggests that the *D. pseudoobscura* Y chromosome is derived from an autosome, the shift potentially facilitated by the fusion of the ancestral X chromosome with one copy of an autosome and the subsequent degeneration of the second autosome into the current Y chromosome. Studies of this nature are only possible with a large amount of sequence information from closely related species. Hopefully the *Drosophila* database will provide opportunities in the near future to test ideas regarding the origin of the F element within the *Drosophila* genus, with particular emphasis on the potential relationship between the F element and the X chromosome.

A second area that will benefit greatly from the various sequencing projects within the genus *Drosophila* are studies focusing on the evolution of chromatin structure. These studies have great potential to shed light on the behavior of the F element, as the presence of a euchromatic dot chromosome in *D. virilis* and a heterochromatic dot chromosome in *D. melanogaster* indicates that the F element has undergone a change in chromatin state at least once within the genus *Drosophila*. At this point, however, we know almost nothing regarding the consequences of changing chromatin states for the genes affected, nor how such a change might be brought about. Some small-scale studies of chromatin structure provide a proof of principle. When staining patterns of HP1 on polytene chromosomes were compared, it was found that the centric heterochromatin staining was conserved among all species included in the study (Fanti *et al*. 2003). In contrast, staining patterns in the chromosome arms are highly variable between species. Interestingly, the localization of HP1 is preserved in hybrid individuals and appears not to be due to sequence repeats based on fluorescent *in situ* hybridization (Fanti *et al*. 2003). A second study of chromatin evolution deals with a heterochromatic gene cluster which includes the *light* gene. While this cluster of genes is found within the centric heterochromatin in *D. melanogaster*, it is found within euchromatin in *D. ananassae*, *D. pseudoobscura*, and *D. virilis*. The genes in *D. melanogaster* have accumulated transposable elements, show increased AT richness, and are longer. In contrast, the same genes in the other species do not show these trends. These findings indicate that changes in the chromatin environment have strong effects on the genes, their structure and potentially their control elements (Yasuhara *et al*. 2005).

In order to come to a true understanding of the various peculiar features of the dot chromosome in *Drosophila*, further work is needed. However, with the completion of the genome sequencing efforts for a total of 12 *Drosophila* species, many new avenues of research are open to scientists interested in the dot chromosome as well as chromosome biology in general. As we have illustrated with the examples above, the dot chromosome in *Drosophila* offers unique opportunities to study the evolution of chromosomes and karyotype. It also provides a great system for studies of chromatin structure and the regulation of genes in various chromatin environments. Comparative studies are a very powerful tool, and we hope they will allow us eventually to come to an understanding of the evolutionary dynamics of chromatin structure changes, and the effects of these changes on individual genes.

Acknowledgements

We thank Carolyn Craig for the images of polytene chromosomes shown in Figure 1, and the members of the Elgin laboratory for critical review of the manuscript. We also thank Drs Scott Hawley (Stowers Institute, Kansas City, USA), John Locke (University of Alberta, Canada), and Jan Larsson (Umea University, Sweden) for providing comments on the manuscript. The authors were supported by National Institutes of Health grant GM073190 to S.C.R.E.

References

Ashburner M, Golic KG, Hawley RS (2005) *Drosophila: A Laboratory Handbook*. Cold Spring Harbor: Cold Spring Harbor Press.

Bartolome C, Maside X, Charlesworth B (2002) On the abundance and distribution of transposable elements in the genome of *Drosophila melanogaster*. *Mol Biol Evol* **19**: 926–937.

Carvalho AB, Clark AG (2005) Y chromosome of *D. pseudoobscura* is not homologous to the ancestral *Drosophila* Y. *Science* **307**: 108–110.

Copenhaver GP, Nickel K, Kuromori T *et al*. (1999) Genetic definition and sequence analysis of *Arabidopsis* centromeres. *Science* **286**: 2468–2474.

Czermin B, Melfi R, McCabe D *et al*. (2002) *Drosophila* enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell* **111**: 185–196.

Eissenberg JC, Elgin SC (2000) The HP1 protein family: getting a grip on chromatin. *Curr Opin Genet Dev* **10**: 204–210.

Fanti L, Berloco M, Piacentini L, Pimpinelli S (2003) Chromosomal distribution of heterochromatin protein 1 (HP1) in *Drosophila*: a cytological map of euchromatic HP1 binding sites. *Genetica* **117**: 135–147.

Franke A, Baker BS (1999) The rox1 and rox2 RNAs are essential components of the compensasome, which mediates dosage compensation in *Drosophila*. *Mol Cell* **4**: 117–122.

Grewal SI, Elgin SC (2002) Heterochromatin: new possibilities for the inheritance of structure. *Curr Opin Genet Dev* **12**: 178–187.

Haynes KA, Leibovitch BA, Rangwala SH, Craig C, Elgin SC (2004) Analyzing heterochromatin formation using chromosome 4 of *Drosophila melanogaster*. *Cold Spring Harbor Symp Quant Biol* **69**: 267–272.

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*. New York: Academic Press, pp. 903–928.

Hoskins RA, Smith CD, Carlson JW *et al*. (2002) Heterochromatic sequences in a *Drosophila* whole-genome shotgun assembly. *Genome Biol* **3**: Research0085.

Jensen MA, Charlesworth B, Kreitman M (2002) Patterns of genetic variation at a chromosome 4 locus of *Drosophila melanogaster* and *D. simulans*. *Genetics* **160**: 493–507.

Kikkawa H (1938) Studies on the genetics and cytology of *Drosophila ananassae*. *Genetica* **20**: 458–516.

Krivshenko J (1952) A cytogenetic study of the Y chromosome in *Drosophila busckii*. *Genetics* **37**: 500–518.

Krivshenko J (1955) A cytogenetic study of the X chromosome of *Drosophila busckii* and its relation to phylogeny. *Genetics* **41**: 1071–1079.

Krivshenko J (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.

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, Makitalo M (2004) *Painting of fourth* in genus *Drosophila* suggests autosome-specific gene regulation. *Proc Natl Acad Sci USA* **101**: 9728–9733.

Lindsley DL, Sandler L, Baker BS *et al*. (1972) Segmental aneuploidy and the gross structure of the *Drosophila* genome. *Genetics* **71**: 157–184.

Locke J, McDermid HE (1993) Analysis of *Drosophila* chromosome 4 using pulsed field gel electrophoresis. *Chromosoma* **102**: 718–723.

Locke J, Podemski L, Roy K, Pilgrim D, Hodgetts R (1999) 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.

Mohr OL (1932) Genetical and cytological proof of somatic elimination of the fourth chromosome in *Drosophila melanogaster*. *Genetics* **17**: 60–80.

Müller HJ (1940) Bearings of the *Drosophila* work on systematics.

In: Huxley J, eds. *The New Systematics*. Oxford: Clarendon Press, pp. 185–268.

Nagaki K, Cheng Z, Ouyang S *et al*. (2004) Sequencing of a rice centromere uncovers active genes. *Nat Genet* **36**: 138–145.

Podemski L, Ferrer C, Locke J (2001) Whole arm inversions of chromosome 4 in *Drosophila* species. *Chromosoma* **110**: 305–312.

Sandler L, Novitski E (1955) 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.

Schotta G, Ebert A, Krauss V *et al*. (2002) Central role of *Drosophila* SU(VAR)3-9 in histone H3-K9 methylation and heterochromatic gene silencing. *EMBO J* **21**: 1121–1131.

Schulze SR, Sinclair DA, Fitzpatrick KA, Honda BM (2005) A genetic and molecular characterization of two proximal heterochromatic genes on chromosome 3 of *Drosophila melanogaster*. *Genetics* **169**: 2165–2177.

Sheldahl LA, Weinreich DM, Rand DM (2003) Recombination, dominance and selection on amino acid polymorphism in the *Drosophila* genome: contrasting patterns on the X and fourth chromosomes. *Genetics* **165**: 1195–1208.

Slawson EE, Shaffer CD, Malone CD *et al*. (2006) Comparison of dot chromosome sequences from D. melanogaster and D. virilis reveals an enrichment of DNA transposon sequences in heterochromatic domains. *Genome Biol* (In press).

Stenberg P, Pettersson F, Saura AO, Berglund A, Larsson J (2005) Sequence signature analysis of chromosome identity in three *Drosophila* species. *BMC Bioinformatics* **6**: 158.

Sturtevant AH (1934) Preferential segregation of the fourth chromosomes in *Drosophila melanogaster*. *Genetics* **20**: 515–518.

Sturtevant AH (1936) Preferential segregation in triplo-IV females in *Drosophila melanogaster*. *Genetics* **21**: 444–466.

Sturtevant AH (1946) On the dot chromosome of *Drosophila repleta* and *D. hydei*. *Genetics* **31**: 259–268.

Sun FL, Cuaycong MH, Craig CA *et al*. (2000) The fourth chromosome of *Drosophila melanogaster*: interspersed euchromatic and heterochromatic domains. *Proc Natl Acad Sci USA* **97**: 5340–5345.

Sun FL, Cuaycong MH, Elgin SC (2001) Long-range nucleosome ordering is associated with gene silencing in *Drosophila melanogaster* pericentric heterochromatin. *Mol Cell Biol* **21**: 2867–2879.

Sun FL, Haynes K, Simpson CL *et al*. (2004) cis-Acting determinants of heterochromatin formation on *Drosophila melanogaster* chromosome four. *Mol Cell Biol* **24**: 8210–8220.

Wakimoto BT, Hearn MG (1990) The effects of chromosome rearrangements on the expression of heterochromatic genes in chromosome 2L of *Drosophila melanogaster*. *Genetics* **125**: 141–154.

Wallrath LL, Elgin SC (1995) Position effect variegation in *Drosophila* is associated with an altered chromatin structure. *Genes Dev* **9**: 1263–1277.

Wang W, Thornton K, Berry A, Long M (2002) Nucleotide variation along the *Drosophila melanogaster* fourth chromosome. *Science* **295**: 134–137.

Wang W, Thornton K, Emerson JJ, Long M (2004) Nucleotide variation and recombination along the fourth chromosome in *Drosophila simulans*. *Genetics* **166**: 1783–1794.

Yan H, Jin W, Nagaki K *et al*. (2005) Transcription and histone modifications in the recombination-free region spanning a rice centromere. *Plant Cell* **17**: 3227–3238.

Yang HP, Hung TL, You TL, Yang TH (2005) Genome-wide comparative analysis of the highly abundant transposable element *DINE-1* suggests a recent transpositional burst in *Drosophila yakuba*. *Genetics*

Yasuhara JC, DeCrease CH, Wakimoto BT (2005) Evolution of heterochromatic genes of *Drosophila*. *Proc Natl Acad Sci USA* **102**: 10958–10963.