Heredity (2007) 98, 3–12

www.nature.com/hdy

SHORT REVIEW

Drosophila melanogaster as a model for studying protein-encoding genes that are resident in constitutive heterochromatin

N Corradini[1], F Rossi[1], E Giordano[2], R Caizzi[3], F Verní[1] and P Dimitri[1]

[1]*Laboratorio di Genomica Funzionale e Proteomica di Sistemi modello and Dipartimento di Genetica e Biologia Molecolare 'Charles Darwin', Università 'La Sapienza', Piazzale Aldo Moro 5, 00185 Roma, Italy;* [2]*Dipartimento di Genetica, Biologia Generale e Molecolare Università di Napoli 'FedericoII', Via Mezzocannone 8, 80100 Napoli, Italy;* [3]*Dipartimento di Anatomia patologica e Genetica, Università di Bari, Via Amendola 165/A, 70126 Bari, Italy*

The organization of chromosomes into euchromatin and heterochromatin is one of the most enigmatic aspects of genome evolution. For a long time, heterochromatin was considered to be a genomic wasteland, incompatible with gene expression. However, recent studies – primarily conducted in *Drosophila melanogaster* – have shown that this peculiar genomic component performs important cellular functions and carries essential genes. New research on the molecular organization, function and evolution of heterochromatin has been facilitated by the sequencing and annotation of heterochromatic DNA. About 450 predicted genes have been identified in the heterochromatin of *D. melanogaster*, indicating that the number of active genes is higher than had been suggested by genetic analysis. Most of the essential genes are still unknown at the molecular level, and a detailed functional analysis of the predicted genes is difficult owing to the lack of mutant alleles. Far from being a peculiarity of *Drosophila*, heterochromatic genes have also been found in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Oryza sativa* and *Arabidopsis thaliana*, as well as in humans. The presence of expressed genes in heterochromatin seems paradoxical because they appear to function in an environment that has been considered incompatible with gene expression. In the future, genetic, functional genomic and proteomic analyses will offer powerful approaches with which to explore the functions of heterochromatic genes and to elucidate the mechanisms driving their expression.
Heredity (2007) **98,** 3–12. doi:10.1038/sj.hdy.6800877; published online 1 November 2006

Keywords: constitutive heterochromatin; heterochromatic genes; *Drosophila*

Heterochromatin: general definition

The word 'heterochromatin' is a cytological term that originally referred to chromosome portions that were stained deeply at prophase and retained a compact structure throughout the mitotic cell cycle (Heitz, 1928). Heterochromatin was classified further into facultative and constitutive types (Brown, 1966). Facultative heterochromatin corresponds to silenced euchromatin (chromosome regions, entire chromosomes or even whole genomes), whereas constitutive heterochromatin is found commonly in large blocks near centromeres and telomeres. The constitutive component consists mostly of repetitive DNA sequences and maintains its characteristics on homologous chromosomes.

Facultative heterochromatin

A well-known example of facultative heterochromatin is the inactive X chromosome in the somatic cells of female mammals (reviewed by Plath *et al*, 2002, 2003). The inactive X chromosome becomes heteropycnotic, suggesting that the chromatin in the silent regions is relatively condensed. The essential steps leading to inactivation can be summarized as follows: initiation of the heterochromatinization process in early developmental stages, starting from a specific locus called the X inactivation centre; spreading of heterochromatinization along the entire chromosome; and once established, maintaining the heterochromatic state through subsequent somatic cell divisions.

Heterochromatinization is achieved by changing the chromatin of the X chromosomes from a transcriptionally active state to an inactive state. This involves a cascade of chromatin modifications that inhibit the establishment of transcription complexes. These modifications include methylation of histone H3 lysine 9 and histone H3 lysine 27; hypoacetylation of histones H2A, H3 and H4; decrease of histone H3 lysine 4 methylation; and changes in the time of DNA replication. Such features of the inactive chromatin of X chromosomes seem to be shared by inactive chromatin elsewhere in mammalian genomes.

Constitutive heterochromatin is not a genomic wasteland

Constitutive heterochromatin is a ubiquitous and common component of eukaryotic genomes. It forms about

Correspondence: P Dimitri, Laboratorio di Genomica Funzionale e Proteomica di Sistemi modello and Dipartimento di Genetica e Biologia Molecolare 'Charles Darwin', Università 'La Sapienza', Piazzale Aldo Moro 5, 00185 Roma, Italy. E-mail: patrizio.dimitri@uniroma1.it
Received 20 December 2005; accepted 28 June 2006; published online 1 November 2006

5% of the genome in *Arabidopsis thaliana*, 30% in humans, 30% in *Drosophila melanogaster* and up to 90% in certain nematodes (Moritz and Roth, 1976; Gatti and Pimpinelli, 1992; Arabidopsis Genome Initiative, 2000). Although constitutive heterochromatin is one of the basic components of eukaryotic chromosomes, the reasons for its widespread occurrence are still unclear.

Several unusual properties characterize constitutive heterochromatin in virtually all animal and plant species, which together have led to the traditional view of this material as a 'desert' of genetic functions (reviewed by John, 1988): a strongly reduced level of meiotic recombination; low gene density; repression of the activity of euchromatic genes when nearby, a phenomenon termed position effect variegation; late replication during S phase; enrichment in highly and repetitive DNAs; and transcriptional inertness.

In the past two decades, however, the idea that constitutive heterochromatin is merely a genomic wasteland has been modified in the light of genetic, cytological and molecular studies conducted primarily in the model organism *D. melanogaster*. These studies have shown that constitutive heterochromatin performs important cellular functions and carries essential genes for viability and fertility (reviewed by Gatti and Pimpinelli, 1992; Williams and Robbins, 1992; Weiler and Wakimoto, 1995; Dernburg *et al*, 1996; Elgin, 1996; Karpen *et al*, 1996; Eissenberg and Hilliker, 2000; Henikoff *et al*, 2001; Coulthard *et al*, 2003; Fitzpatrick *et al*, 2005; Dimitri *et al*, 2005a, b). In addition, about 450 predicted genes have recently been identified by the annotation of the heterochromatin sequence (Hoskins *et al*, 2002). Remarkably, the presence of coding genes in heterochromatin, far from being a peculiarity of *Drosophila*, seems to be a conserved trait in the evolution of eukaryotic genomes. Heterochromatic genes have been found in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Oryza sativa* and *A. thaliana*, as well as in humans (Kuln *et al*, 1991; Arabidopsis Genome Initiative, 2000; Horvath *et al*, 2000; Brun *et al*, 2003; Nagaki *et al*, 2004).

Genes resident in the constitutive heterochromatin of *Drosophila*

Single-copy genes encoding essential functions

D. melanogaster is the model organism in which the greatest progress in the study of heterochromatin function has been made as a result of combined genetic, cytological and genomic approaches. Mutatable genes essential for viability and fertility were initially identified in *D. melanogaster* by recessive lethal mutations that were genetically linked to heterochromatin (Hilliker, 1976; Marchant and Holm, 1988). Complementation analysis using rearrangements with cytologically determined breakpoints in the heterochromatin of mitotic chromosomes finally demonstrated that these genes are embedded in heterochromatin and enabled their mapping (Dimitri, 1991; Koryakov *et al*, 2002). Notably, most of the heterochromatic genes thus far detected are located in chromosome regions that fluoresce weakly after staining with 4,6-diamino-2-phenylindole-dihydrochloride (DAPI). These regions harbour clusters of transposable elements and are devoid of highly repetitive satellite DNAs (Lohe *et al*, 1993; Pimpinelli *et al*, 1995). Thus far, at least 32 essential genes are known to map to the mitotic heterochromatin of chromosomes 2 and 3. Only a few of these, however, are sufficiently defined at the molecular level: *RpL5*, *light*, *concertina*, *rolled*, *RpL38*, *Nipped-B*, *Nipped-A*, *Parp* and *RpL15* (Hilliker, 1976; Devlin *et al*, 1990a, b; Parks and Wieschaus, 1991; Biggs *et al*, 1994; Rollins *et al*, 1999; Tulin *et al*, 2002; Myster *et al*, 2004; Marygold *et al*, 2005; Schulze *et al*, 2005).

Predicted genes

The release of the sequence of *D. melanogaster* heterochromatin by the Berkeley Drosophila Genome Project (http://www.fruitfly.org/) and the Drosophila Heterochromatin Genome Project (DHGP; http://www.dhgp.org/index_release_notes.html) has greatly facilitated the study of the molecular organization and function of heterochromatic genes. Initially, 3.8 Mb of about 120 Mb of the *D. melanogaster* euchromatic genome sequence was found to correspond to heterochromatic sequences (Adams *et al*, 2000). More recently, an improved whole-genome shotgun assembly (heterochomatic-WGS3; Hoskins *et al*, 2002) was produced, which includes 20.7 Mb of draft-quality heterochromatic sequence. About 450 predicted genes have been identified by the annotation of the heterochromatin sequence (Hoskins *et al*, 2002), suggesting that the number of active genes in the constitutive heterochromatin of *D. melanogaster* is higher than that defined by genetic analysis. In the WGS3 heterochromatin, 45% of predicted genes have an overlapping expressed sequence tag, whereas 20% are based on full-insert sequences of complementary DNAs (Hoskins *et al*, 2002).

Several studies have concentrated on efforts to map predicted genes to the mitotic heterochromatin of *D. melanogaster* using bacterial artificial chromosomes, complementary DNAs and P-elements (see Figure 1;

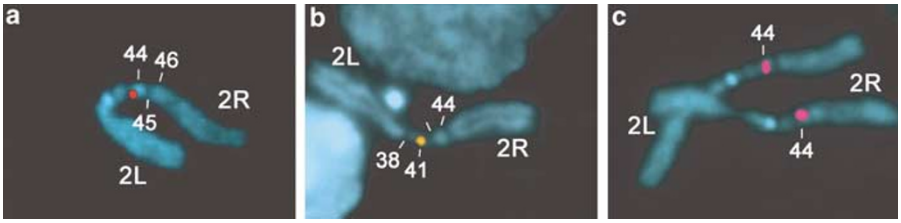

Figure 1 Examples of FISH mapping of three different DNA sequences to the mitotic heterochromatin of chromosome 2. Prometaphase chromosomes of larval brain cells from the *y; cn bw sp* strain were stained with DAPI and pseudo-coloured in blue; the different FISH signals are shown in red, yellow and pink. (**a**) The CG17691 complementary DNA signal (red) maps to the proximal edge of region h44 of 2Rh. (**b**) The 19.74.3 P-element insert maps to h40–h41 of 2Rh. (**c**) The hybridization signal of BACR04P15 is located in h44.

Table 1 Known heterochromatic genes and gene models on chromosome 2 and their functions

Scaffold (kb)	*CG (alleles)*	D. melanogaster *molecular function (GO)*	*Closest BLASTx human hit (P(n))*
AABU1002768 (594)	CG17540 (1)	Pre-mRNA splicing factor	NP_116294.1, RNA-binding motif protein 17 (1e-55)
	CG41117 (1)	Unknown	No significant similarity found
	Cht3 (1)	Chitin binding, chitinase	AAG60019.1, acidic mammalian chitinase precursor (8e-75)
	concertina (14)	GTPase; GTP binding; signal transducer	NP_006563.2, guanine nucleotide-binding protein alpha 13 (6e-100)
	CR40471	Misc. RNAs	
	CG40005 (1)	GTPase	AAA74235.1, guanine nucleotide regulatory protein (7e-23)
	light (47)	Ubiquitin-protein ligase; zinc ion binding	P49754, vacuolar assembly protein VPS41 homologue (S53) (3e-147)
	CG41118 (1)	Unknown	No significant similarity found
	CG17715 (1)	Unknown	AAH32396.1, LOC157378 protein (4e-29)
	CG41121 (1)	Unknown	No significant similarity found
	CG40439 (1)	Unknown	No significant similarity found
	CG41119 (1)	Unknown	No significant similarity found
	CG41120 (1)	Unknown	No significant similarity found
	CG40006 (1)	Serine-type endopeptidase inhibitor; receptor	AAH80647.1, scavenger receptor class B member 1 (2e-48)
	CG17494 (1)	Unknown	NP_009090.2, sarcolemma-associated protein (1e-61)
	CG17493 (1)	Calmodulin binding; calcium ion binding	AAP35920.1, centrin EF hand protein 2 (7e-62)
	CG17490 (1)	Unknown	No significant similarity found
	RpL5 (5)	Ribosomal protein L5	NP_000960.2, ribosomal protein L5 (4e-119)
AABU01002756 (199)	CG12567 (1)	Thiamine diphosphokinase	No significant similarity found
	CG40040 (1)	Unknown	No significant similarity found
	CG40041 (1)	Hormone	AA033390.1, glycoprotein hormone beta subunit (2e-10)
	CG40042 (1)	Carrier; protein transporter	AAR14723.1, mitochondrial inner membrane translocase 23 (8e-36)
AABU01002199 (160)	CG40211 (1)	Unknown	No significant similarity found
	CG40212 (1)	Unknown	No significant similarity found
	CG40216 (1)	Unknown	No significant similarity found
	CG40218 (2)	Kinesin binding	AAP88821.1, craniofacial development protein 1 (2e-20)
	CG40334 (1)	Nucleic acid binding; damaged DNA binding	XP_371082.1, PREDICTED: hypothetical protein FLJ20753 (6e-19)
	CG40220 (1)	Unknown	No significant similarity found
	CG40214 (1)	Nucleic acid binding; damaged DNA binding	XP_371082.1, PREDICTED: hypothetical protein FLJ20753 (4e-29)
AABU01001947 (252)	CG40241 (1)	Unknown	No significant similarity found
	CG12552 (1)	Unknown	No significant similarity found
	CG40498 (1)	Unknown	No significant similarity found
	CG40239 (1)	Unknown	No significant similarity found
	rolled (90)	MAP kinase	AAH17832.1, mitogen-activated protein kinase 1 (2e-168)
	CG40244 (1)	Unknown	CAA77753.1, 40 kDa protein kinase (5e-05)
AABU01002549 (120)	CG41063 (1)	Unknown	No significant similarity found
	CG40190 (1)	Protein kinase	NP_002737.1, mitogen-activated protein kinase 3 (6e-45)
	CG41066 (1)	Unknown	No significant similarity found
	CG40192 (1)	Unknown	No significant similarity found
	CG41065 (1)	Unknown	No significant similarity found
	CG40191 (1)	Unknown	AAD34052.1, CGI-57 protein (2e-20)
AABU01002750 (143)	CG17691 (1)	3 methyl-2-oxobutanoate dehydrogenase	CAI15049.1, branched chain keto acid dehydrogenase E1, beta polypeptide (maple syrup urine disease) (1e-135)
	CG40068 (1)	Nucleic acid binding; translation factor	AAM14617.1, mitochondrial translation-initiation factor 2 (4e-32)
	CG40388 (1)	Unknown	No significant similarity found
	CG40390 (1)	Unknown	CAI12944.1, KARP-1-binding protein (KAB) (0.005)
	CG40067 (1)	Unknown	No significant similarity found

Table 1 (Continued)

Scaffold (kb)	*CG (alleles)*	D. melanogaster *molecular function (GO)*	*Closest BLASTx human hit (P(n))*
AABU01002748 (134)	CG40084 (1)	Unknown	NP_951058.1, cyclin M2 isoform 2 (4e-117)
	CG40081 (1)	Unknown	No significant similarity found
	CG40080 (1)	ATP binding; protein serine/threonine kinase	NP_114171.2, serine/threonine protein kinase Haspin (3e-69)
	CG40085 (1)	Unknown	No significant similarity found
	CG41098 (1)	Unknown	No significant similarity found
AABU01002711 (342)	CG40130 (1)	Unknown	No significant similarity found
	CG40129 (1)	ATP binding; G-protein-coupled receptor kinase	NP_005151.1, beta adrenergic receptor kinase 2 (0.0)
	CG17665 (2)	Unknown	NP_075391.3, hypothetical protein LOC65123 (0.0)
	CG40131 (1)	Unknown	No significant similarity found
	CG40127 (1)	Unknown	NP_001004333.1, hypothetical protein LOC440400 (1e-11)
	CG40128 (1)	Unknown	No significant similarity found
	CG40133 (1)	Unknown	No significant similarity found
	CG17683 (1)	Oxidoreductase; ferredoxin hydrogenase	NP_071938.1, nuclear prelamin A recognition factor-like (5e-113)
AABU01002740 (252)	CG17514 (1)	Translation activator; kinase regulator	NP_006827.1, GCN1 general control of amino-acid synthesis 1-like 1 (0.0)
	CG40103 (1)	Unknown	No significant similarity found
	CG40163 (1)	Unknown	No significant similarity found
	CG10837 (*eIF-4B*) (8)	RNA binding; translation initiation factor	AAH73154.1, eukaryotic translation initiation factor 4B (6e-32)
	CG40105 (1)	Specific RNA pol II transcription factor	NP_002136.1, homeo box B2 (1e-09)
AABU01002769 (2465)	*RpL38* (6)	Ribosomal protein 38	NP_000990.1, ribosomal protein L38 (1e-24)
	CG40293 (1)	ATP binding; protein serine/threonine kinase	DAA01797.1, TPA: STE20-related adaptor protein (1e-41) AAG48269.1, breast cancer antigen NY-BR-96 (1e-41)
	CG40290 (1)	Unknown	No significant similarity found
	p120ctn (12)	Adherens junction protein	AAB97957.1, arm-repeat protein NPRAP/neurojungin (2e-131)
	CG17486 (1)	Ligase; asparagine synthase	NP_061921.1, hypothetical protein LOC54529 (8e-73)
	CG17478 (1)	Unknown	No significant similarity found
	CG17883 (1)	Unknown	NP_653229.1, TBC1 domain family, member 20 (1e-66)
	Nipped-B (8)	Transcriptional activator; sister chromatid cohesion	NP_597677.2, delangin isoform A (SCC2 homologue) (0.0)
	CG17706 (2)	Unknown	No significant similarity found
	CG17082 (1)	Unknown	NP_277050.2, Rho GTPase activating protein 18 (7e-23)
	CG12547 (1)	Unknown	CAH73013.1, novel NHL repeat domain-containing protein (8e-127)
	CG17528 (1)	Microtubule binding; ATP binding; protein serine/threonine kinase	NP_004725.1, doublecortin and CaM kinase-like 1 (5e-131)
	CG40285 (1)	Unknown	NP_689529.1, hypothetical protein LOC120534 (1e-15)
AE003788 (585)	CR30260	TRNA	—
	CR30505	TRNA	—
	CG33492 (1)	Ionotropic glutamate receptor	CAH73378.1, glutamate receptor, ionotropic, delta 1 (5e-08)
	TpnC41C (1)	Calcium ion binding	AAH08437.1, Calmodulin 2 (1e-33)
AE003787 (294)	CG3107 (1)	Metalloendopeptidase	NP_055704.2, metalloprotease 1 (0.0)
	gustavus (4)	Oocyte anterior/posterior axis determination	BAD92796.1, SPRY domain-containing SOCS box protein SSB-1 variant (3e-115)
	CG3136 (2)	DNA binding; protein homodimerization	BAA34722.1, cAMP-dependent transcription factor ATF-6 beta (4e-18)
	Nipped-A (47)	Transcription regulator; receptor signalling	NP_003487.1, TRRAP protein (0.0)
	d4 (1)	Transcription factor; ubiquitin-protein ligase	AAP35364.1, requiem, apoptosis response zinc-finger gene (1e-51)
	Ogt (1)	Transferase activity, transferring glycosyl groups	NP_858059.1, O-linked GlcNAc transferase, isoform 2 (0.0)
	CG10465 (2)	Voltage-gated potassium channel; protein binding	NP_114160.1, potassium channel tetramerization domain containing 10 (1e-100)
	CG10395 (1)	HIT Zn-finger protein domain	NP_112578.1, high mobility group AT-hook 1-like 4 (4e-20)
	CG30441 (1)	Unknown	AAP50265.1, intraflagellar transport protein 20-like protein (8e-16)

Table 1 (Continued)

Scaffold (kb)	*CG (alleles)*	D. melanogaster *molecular function (GO)*	*Closest BLASTx human hit (P(n))*
	CG10396 (1)	Cytochrome *c* oxidase	P13073, Cytochrome *c* oxidase polypeptide IV (7e-24)
	CG33347 (1)	Unknown	No significant similarity found
	CG10417 (1)	Protein serine/threonine phosphatase	015355, protein phosphatase 2C gamma isoform (8e-67)
AE003786 (280)	CG30437 (2)	Laccase; copper ion binding	AAQ24260.1, MY 01 protein (0.001)
	CG32838 (1)	Another laccase; 75% identical to CG30437	AAQ24260.1, MY 01 protein (5e-04)
	CG30440 (1)	Guanyl-nucleotide exchange factor	AAA52172.1, DBL transforming protein (1e-74)
	CG30438 (1)	Transferase activity, transferring glycosyl groups	AAC51187.1, ceramide UDPgalactosyltransferase (2e-61)
	TpnC4 (1)	Calcium ion binding	AAH08437.1, Calmodulin 2 (7e-25)
	CG17510 (1)	Unknown	Q9Y3D6, tetratricopeptide repeat protein 11 (Fis 1 homologue) (1e-16)
	CG17508 (1)	Unknown	NP_543011.1, hypothetical protein LOC116151 (1e-23)
	CG11665 (1)	Monocarboxylic acid transporter	NP_963860.1, monocarboxylate transporter 13 (7e-36)

Only BLASTX hits with $E<10^{-3}$ were selected.

Hoskins *et al*, 2002; Corradini *et al*, 2003; Yasuhara *et al*, 2003; R Rossi *et al*, unpublished). For example, about a hundred predicted genes have been assigned to specific regions of the mitotic heterochromatin map of chromosome 2, in which genetic analyses detected 17 essential genes (Table 1). At least two factors might account for the high number of predicted genes: first, heterochromatin might contain an excess of non-essential coding genes. Indeed, CG40293, *p120* and CG17486 of 2Rh (Table 1) were found to be non-essential (Myster *et al*, 2004). Second, seemingly different predicted genes might be portions of the same gene, or some predictions might be incorrect.

Studying the function of heterochromatic predicted genes by RNA interference-mediated inactivation

RNA interference (RNAi) in *Drosophila* cell cultures has proven to be an extremely powerful method for disrupting the expression of genes and for studying their functions (Somma *et al*, 2002). Phenotypic analyses of RNAi cells can thus be useful for studying the function of heterochromatic predicted genes that lack mutant alleles (Figure 2a). In collaboration with DHGP, we have undertaken RNAi studies of about 100 predicted genes that have recently been mapped to the heterochromatin of chromosome 2 (Hoskins *et al*, 2002; Corradini *et al*, 2003; Yasuhara *et al*, 2003; Rossi *et al*, submitted). Clearly, the phenotypic analyses of RNAi cells define only those genes required at different steps of cell division and behaviour (such as cytokinesis, chromosome segregation, chromosome condensation, DNA repair and metabolism and cell–cell interactions) that are likely to give rise to obvious cytological defects.

RNAi inactivation of heterochromatic genes can also be performed *in vivo* with suitable vectors. One such vector is SympUAST, which was designed to generate simultaneous transcription of both strands of a given DNA insert by means of two flanking convergent arrays of the Gal4-responsive UAS regulatory elements (Figure 2b; Giordano *et al*, 2002). The advantage of this vector is that it can accommodate DNA fragments longer than those used in vectors that transcribe inverted repeats and at the same time ensure high stability in the genome. In turn, there will be stability at the phenotypic level owing to the absence of recombinogenic inverted repeats. These transgenes are able to repress gene activity in transformed adult flies. An advantage of this procedure is that, according to the expression pattern of the Gal4 line used as a driver, RNAi can be induced ubiquitously or in selected tissues at specific developmental stages, and that the silencing effect is cell-autonomous.

Structure, function and evolution of genes located in heterochromatin

Organization and function of heterochromatic genes

A difference between heterochromatic and euchromatic genes is likely to be due to their molecular structure. Genes located in the heterochromatin of *D. melanogaster* are on average significantly larger than euchromatic genes owing to the occurrence of long introns enriched in transposable element-related DNA sequences (Devlin *et al*, 1990a, b; Tulin *et al*, 2002; Dimitri *et al*, 2003). The

example of the Y-chromosome fertility factors of *D. melanogaster* is marked as they contain up to 4 Mb of DNA (Gatti and Pimpinelli, 1992) and carry transposable-element-rich mega-introns that can account for 1 or 2 Mb (Kurek *et al*, 2000; Carvalho *et al*, 2001). Although some of the predicted genes mapped to the heterochromatin of chromosomes 2 and 3 are also large, this does not seem to be a general rule. Several hypotheses can be used to explain these observations. First, during evolution, older genes in heterochromatin might have increased in size by being targets for transposable-element insertions. In this case, it would follow that short predicted genes in heterochromatin were younger, having recently 'moved' to heterochromatin. Second, genes in heterochromatin might be targeted differently by transposable elements, with some genes being more refractory than others. Third, there might be selective pressure to maintain some genes of short size in heterochromatin owing to particular functional properties. Highly expressed genes tend to have substantially shorter introns than those expressed at low levels (Castillo-Davis *et al*, 2002). In this regard, it is worth noting that although *RpL38*, *RpL5* and *RpL15* – three highly expressed ribosomal protein-coding genes – are located in heterochromatin, they are all short and carry short introns (Marygold *et al*, 2005; Schulze *et al*, 2005).

The expression of heterochromatic essential genes such as *light*, *rolled*, *RpL5*, *RpL38*, *RpL15*, *Nipped-B* and *Parp* is present throughout all developmental stages (Biggs *et al*, 1994; Rollins *et al*, 1999; Tulin *et al*, 2002; Marygold *et al*, 2005; Schulze *et al*, 2005). The same is true for fourteen predicted genes from chromosome 2 heterochromatin that we have recently analysed (F Rossi, N Corradini, R Moschetti, R Caizzi and P Dimitri, unpublished data). The Y-chromosome fertility factors are an exception, as they have tissue- and sex-specific limited expression. In general, on the basis of molecular and bioinformatic analyses, predicted genes and known genes found in heterochromatin do not have molecular functions that would distinguish them from genes located in euchromatin (Hoskins *et al*, 2002; FlyBase, 2006). In other words, heterochromatin does not seem to have a distinctive proteome.

Gene expression in heterochromatic domains

The presence of expressed genes in the heterochromatin of evolutionarily distant organisms seems paradoxical: how can protein-coding genes work properly in an environment that is thought to be incompatible with

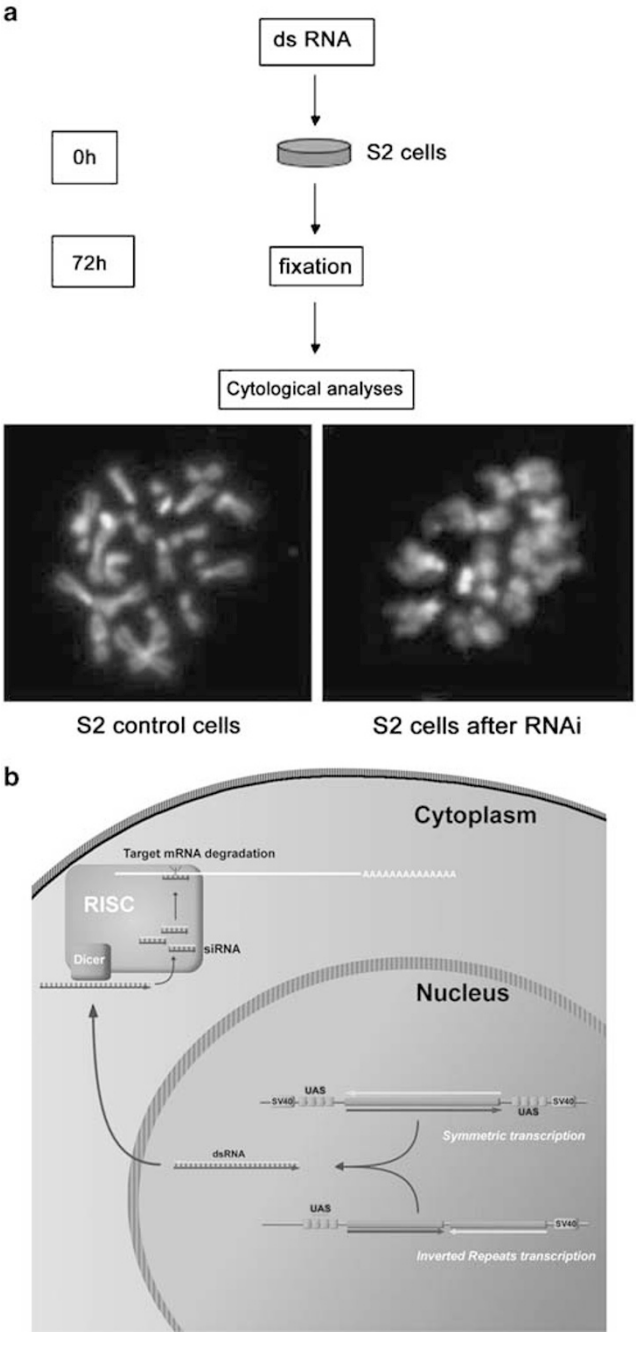

Figure 2 Knocking down heterochromatic gene function by RNAi. (**a**) RNAi in S2 cells. At 72 h after double-stranded RNA (dsRNA) injection, cells are fixed with methanol:acetic acid and chromosomes are prepared by air drying. Chromosomes of control and treated cells are stained by DAPI, and cytological analysis is performed to reveal apparent mitotic defects. At the bottom of the figure, an aberrant chromosome condensation morphology after RNAi of a 2Rh gene model is shown (right panel). This aberrant phenotype clearly differs from that of the nontreated control cells (left panel). (**b**) *In vivo* RNAi using dsRNA-producing transgenes coupled to the GAL4/UAS expression system. In symmetrical transcription, dsRNA synthesis can be induced by cloning, as in the SympUAST vector, a single gene fragment between two convergent arrays of the GAL4-responsive UAS sequences. This strategy permits the use of a long DNA fragment and results in a very stable plasmid both in the bacterial host during the cloning procedure and in the *Drosophila* genome after transgenesis. dsRNA can also be efficiently produced *in vivo* by mono-directional transcription of inverted repeats (IR) of a given DNA fragment under the control of a single UAS regulatory region. In this case, the inverted repeats should be separated by a short DNA spacer (100–200 bp) to avoid any rearrangement due to the recombinogenic potential of contiguous IR in the bacterial host and *Drosophila* genome. The synthesized dsRNA is then processed by the Dicer RNase, which is a component of the RNAi silencing complex (RISC), located in the cytoplasm. The Dicer RNase cleaves the exported dsRNA into smaller, 21-nucleotide small inhibitory RNAs (siRNAs). The siRNAs are used by the RISC complex as a template for destroying the homologous mRNA (in white). Through this mechanism, specific messages can be degraded in a cell-autonomous manner and in specific tissues according to the expression pattern of the GAL4 enhancer trap line used.

gene expression? Genes located in constitutive heterochromatin might have regulatory requirements that are different from those in euchromatin. In fact, heterochromatic genes are repressed when moved to euchromatin by chromosomal rearrangements, indicating that they are dependent on a native heterochromatic location for correct expression (Wakimoto and Hearn, 1990; Eberl *et al*, 1993; see below). Therefore, these genes cannot simply be considered as euchromatin-like active sequences embedded in a repetitive genomic compartment. It would be more correct to consider euchromatin and constitutive heterochromatin as two different chromatin regions, in both of which gene expression can occur, possibly depending on the formation of differential multiprotein complexes. Chromosomal proteins required for the establishment of the heterochromatic state, such as heterochromatin protein 1 (HP1) and others, might also be involved in the control of gene expression in heterochromatin (Weiler and Wakimoto, 1995). Experimental evidence consistent with this proposal has been reported in *D. melanogaster*. First, genetic experiments suggested that modifiers of position effect variegation, such as suppressor of variegation (Su(var)) gene products, can interact to guarantee the proper expression of the *light* gene in its normal heterochromatic location (Clegg *et al*, 1998). Second, the amount of mRNA of *light* and *rolled* heterochromatic genes was found to be reduced about 2.5-fold in HP1 mutant larvae (Lu *et al*, 2000). To obtain more insight into the roles of HP1 and SU(VAR)3–9 – the gene encoding a histone methyltransferase – large-scale mapping of their target genes was recently carried out in *Drosophila* embryonic Kc cells (Greil *et al*, 2003). The results revealed that HP1 and SU(VAR)3–9 bind together to genes and transposable elements in constitutive heterochromatin. More recently, whole-genome and computational approaches were used to study HP1 targeting of genomic sites (de Wit *et al*, 2005). The HP1 protein was found to bind to the heterochromatic *rolled* gene, and the binding is prominent in both unique and repetitive portions of the genomic region of the gene. Together, these results corroborate the view that HP1, SU(VAR)3–9 and possibly other protein complexes are required for the proper functioning of genes resident in heterochromatin. It is tempting to speculate that such protein complexes, possibly by recognizing specific patterns of histone modifications, might be involved in the insulation of gene expression in constitutive heterochromatin.

One of the peculiar features of the heterochromatic DNA is its late replication in S phase, unlike the early replication of the bulk of euchromatin sequences. One could speculate that such a dichotomy might also occur at the transcriptional level: for example, heterochromatic and euchromatic genes might differ in their timing of expression during the cell cycle. Notably, cell-cycle-dependent changes in the chromosomal localization of heterochromatin-binding proteins have been observed in *Drosophila* (Platero *et al*, 1998), which might be related to the differential expression of heterochromatic genes.

The role of RNAi in the formation of heterochromatin in evolutionarily distant organisms has been documented and discussed at length (Hall *et al*, 2002; Reinhart and Bartel, 2002; Volpe *et al*, 2002; Lehnertz *et al*, 2003; Fukagawa *et al*, 2004; Pal-Bhadra *et al*, 2004; Sun *et al*, 2004; Verdel *et al*, 2004; Bernstein and Allis, 2005). However, it is still unclear whether RNAi is also involved in regulating the activity of genes located in constitutive heterochromatin. Studying the effects of mutations in the RNAi genes, such as *piwi* and *argonaute*, on the transcription of *light* and *rolled* genes might help to define the functional relevance of RNAi machinery for heterochromatic gene expression.

D. melanogaster heterochromatic genes in other species

Recently, Yasuhara *et al* (2005) studied a group of genes of 2L heterochromatin, including the *light* gene, and found that they are euchromatic in both *D. pseudoobscura* and *D. virilis*. In particular, 7 out of 11 genes found in *D. melanogaster* 2L heterochromatin have orthologues in *D. pseudoobscura* that are clustered on the corresponding chromosome arm. The authors propose that, ancestrally, *light* and its neighbouring genes were typical euchromatic genes found proximally in the euchromatin; alternatively, they might have been relocated closer to the centromere by chromosome rearrangement. Subsequently, an infiltration of pericentromeric heterochromatin into proximal euchromatin might have occurred. Interestingly, Yasuhara *et al* (2005) also found that the promoters of the heterochromatic genes they studied did not differ from euchromatic promoters. Thus, evolutionary adaptation to heterochromatin has been achieved without changing the basic promoter type. At present, it is unclear whether the expression pattern of the euchromatic *light* gene in *D. pseudoobscura* and *D. virilis* is different from that of its heterochromatic orthologue in *D. melanogaster*. These results, together with the finding that orthologous of other essential heterochromatic genes of *D. melanogaster*, such as *rolled*, *Nipped-A*, *Nipped-B*, *RpL15*, *RpL38* and *Parp* are euchromatic in different species, including yeast, mouse, humans (Biggs *et al*, 1994; Krantz *et al*, 2004; Tonkin *et al*, 2004; Schulze *et al*, 2006; Rossi *et al*, submitted), indicate that during evolution the conservation of a heterochromatic location may not be crucial for the proper expression of a given gene.

An interesting approach for identifying the genomic location of the orthologues of *D. melanogaster* heterochromatic genes in other species was developed by the DHGP and is based on the analysis of the repeat content of orthologous introns and scaffolds (Smith *et al*, 2005). This approach should allow one to estimate whether genes have moved into or out of heterochromatin regions in other species and might provide new insight into the evolution of constitutive heterochromatin.

During evolution, transposable elements might have contributed to the plasticity of heterochromatin by stimulating chromosome rearrangements or gene transfer into this peculiar genomic compartment. Once in heterochromatin, genes would have been targeted by recurrent insertions of transposable elements, thus generating the present structural organization. Transposable element-related sequences in heterochromatin might thus contribute, in ways that are still poorly understood, to many of the structural and functional properties of heterochromatin, including the regulatory evolution of these genes (Pimpinelli *et al*, 1995; Weiler and Wakimoto, 1995; Dimitri, 1997; Dimitri and Junakovic, 1999; Yasuhara *et al*, 2005), similar to documented instances found in euchromatin

(Von Stenberg *et al*, 1992; Miller *et al*, 1999; Kidwell and Lish, 2000; Jordan *et al*, 2003; Brandt *et al*, 2005; Kapitonov and Jurka, 2005). Over evolutionary time, transposable elements can thus be viewed as 'artisans' that have shaped heterochromatin, rather than merely as parasitic sequences (Dimitri *et al*, 2005b). Suggestive models on the evolutionary origin of heterochromatic genes in different species have been recently proposed by Yasuhara and Wakimoto (2006).

Drosophila heterochromatic genes related to human disease genes

The use of *D. melanogaster* for the study of human gene functions is now well documented. A systematic search for human disease-causing genes in *Drosophila* showed that about 75% of human disease genes match unique *Drosophila* sequences (Reiter *et al*, 2001). Interestingly, about 70% of the 101 putative proteins encoded by the heterochromatic predicted genes annotated on chromosome 2 have strong similarity to human proteins (Table 1). Some of the orthologous genes in humans are involved in genetic diseases. For example, NIPBL, the human homologue of Nipped-B, maps to euchromatin of chromosome 5p13 and is widely expressed in fetal and adult tissues. Mutations in NIPBL are responsible for the Cornelia de Lange syndrome, a multiple malformation disorder (Krantz *et al*, 2004; Tonkin *et al*, 2004). The predicted gene CG17528 encodes a putative microtubule-binding protein, which contains different, highly conserved functional domains: two tandemly repeated doublecortin (DCX) domains, which are characteristic of some microtubule-binding proteins, and a carboxy-terminal serine/threonine kinase domain. The human orthologues of CG17528, DCX, DCKL1 and DCKL2, might be implicated in lissencephaly, a genetic disorder characterized by severe mental retardation. Another interesting example is that of CG40218. The CG40218 protein belongs to the evolutionarily conserved family of bucentaur (BCNT)-like proteins found in several animals and plants (*A. thaliana*, *O. sativa*, *Neurospora*, *S. cerevisiae*, *Caenorhabditis elegans*, mosquitoes, flies, mice and humans). Little is known about the functions of the BCNT-like family. For example, craniofacial development protein 1 (CFDP1), the human orthologue of CG40218, encodes a 299 aa protein phosphorylated by casein kinase II, and whose function is still unknown. Intriguingly, this gene maps to chromosome 16 in 16q22.2–q22.3, in proximity to several loci associated with inherited craniofacial diseases such as Fanconi anaemia type A (Diekwisch *et al*, 1999).

Conclusions

Although some interesting features are beginning to emerge from structural and functional studies on genes located in heterochromatin, much still remains to be learned. In particular, we need to (1) link the sequences of genes in heterochromatin to its functions; (2) identify the factors that determine correct expression of those genes and (3) understand their evolutionary dynamics.

Acknowledgements

This work was supported by grants from Ministero dell'Università e della Ricerca 'Progetti di ricerca di interesse nazionale' and Istituto Pasteur-Fondazione Cenci Bolognetti.

References

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG *et al* (2000). The genome sequence of *Drosophila melanogaster*. *Science* **287**: 2185–2195.

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 786–791.

Bernstein E, Allis CD (2005). RNA meets chromatin. *Genes Dev* **19**: 1635–1655.

Biggs HW, Zavitz HK, Dikinson B, Van Der Straten A, Brunner D, Hafen E *et al* (1994). The *Drosophila rolled* locus encodes a MAP kinase required in the sevenless signal transduction pathway. *EMBO J* **13**: 1628–1635.

Brandt J, Schrauth S, Veith AM, Froschauer A, Haneke T, Schultheis C *et al* (2005). Transposable elements as a source of genetic innovation: expression and evolution of a family of retrotransposon-derived neogenes in mammals. *Gene* **345**: 101–111.

Brown SW (1966). Heterochromatin. *Science* **151**: 417–425.

Brun ME, Ruault M, Ventura M, Roizes G, De Sario A (2003). Juxtacentromeric region of human chromosome 21: a boundary between centromeric heterochromatin and euchromatic chromosome arms. *Gene* **312**: 41–50.

Carvalho AB, Dobo BA, Vibranovski MD, Clark AG (2001). Identification of five new genes on the Y chromosome f *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **98**: 13225–13230.

Castillo-Davis CI, Mekhedov SL, Hartl DL, Koonin EV, Kondrashov FA (2002). Selection for short introns in highly expressed genes. *Nat Genet* **31**: 415–418.

Clegg NJ, Honda BM, Whitehead IP, Grigliatti TA, Wakimoto B, Brock HW *et al* (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.

Corradini N, Rossi F, Verní F, Dimitri P (2003). FISH analysis of *Drosophila* heterochromatin using BACs and P-elements. *Chromosoma* **112**: 26–37.

Coulthard AB, Eberl DF, Sharp CB, Hilliker AJ (2003). Genetic analysis of the second chromosome centromeric heterochromatin of *Drosophila melanogaster*. *Genome* **46**: 343–352.

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

Dernburg AF, Sedat JW, Hawley RS (1996). Direct evidence of a role for heterochromatin in meiotic chromosome segregation. *Cell* **86**: 135–146.

Devlin RH, Bingham B, Wakimoto BT (1990a). The organization and expression of the *light* gene, a heterochromatic gene of *Drosophila melanogaster*. *Genetics* **125**: 129–140.

Devlin RH, Holm DG, Morin KR, Honda BM (1990b). Identifying single-copy DNA sequence associated with the expression of a heterochromatic gene, the *light* locus of *Drosophila melanogaster*. *Genome* **33**: 405–415.

Diekwisch TG, Marches F, Williams A, Luan X (1999). Cloning, gene expression, and characterization of CP27, a novel gene in mouse embryogenesis. *Gene* **235**: 19–30.

Dimitri P (1991). Cytogenetic analysis of the second chromosome heterochromatin of *Drosophila melanogaster*. *Genetics* **127**: 553–564.

Dimitri P (1997). Constitutive heterochromatin and transposable elements in *Drosophila melanogaster*. *Genetica* **100**: 85–93.

Dimitri P, Corradini N, Rossi F, Verní F (2005a). The paradox of functional heterochromatin. *BioEssays* **27**: 29–41.
Dimitri P, Junakovic N (1999). Revising the selfish DNA hypothesis: new evidence on accumulation of transposable elements in heterochromatin. *Trends Genet* **15**: 123–124.
Dimitri P, Junakovic N, Arcà B (2003). Colonization of heterochromatic genes by transposable elements in *Drosophila*. *Mol Biol Evol* **20**: 503–512.
Dimitri P, Verní F, Mei E, Rossi F, Corradini N (2005b). Transposable elements as artisans of the heterochromatic genome. *Cytogenet Genome Res* **110**: 165–172.
Eberl D, Duyf BJ, Hilliker AH (1993). The role of heterochromatin in the expression of a heterochromatic gene, the *rolled* gene of *Drosophila melanogaster*. *Genetics* **134**: 277–292.
Eissenberg JC, Hilliker AJ (2000). Versatility of conviction: heterochromatin as both repressor and an activator of transcription. *Genetica* **109**: 19–24.
Elgin SCR (1996). Heterochromatin and gene regulation in *Drosophila*. *Curr Opin GenetDev* **6**: 193–200.
FlyBase (2006). http://www.flybase.net
Fitzpatrick KA, Sinclair DA, Schulze SR, Syrzycka M, Honda BM (2005). A genetic and molecular profile of third chromosome centric heterochromatin in *Drosophila melanogaster*. *Genome* **48**: 571–584.
Fukagawa T, Nogami M, Yoshikawa M, Ikeno M, Okazaki T, Takami Y *et al* (2004). Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat Cell Biol* **6**: 784–791.
Gatti M, Pimpinelli S (1992). Functional elements in *Drosophila melanogaster* heterochromatin. *Annu Rev Genet* **26**: 239–275.
Giordano E, Rendina R, Peluso I, furia M (2002). RNAi triggered by symmetrically transcribed transgenes in *Drosophila melanogaster*. *Genetics* **160**: 637–648.
Greil F, van der Kraan I, Delrow J, Smothers JF, de Wit E, Bussemaker HJ *et al* (2003). Distinct HP1 and *Su(var)3–9* complexes bind to sets of developmentally coexpressed genes depending on chromosomal location. *Genes Dev* **17**: 2825–2838.
Hall IM, Shankaranarayana GD, Noma K, Ayoub N, Cohen A, Grewal SI (2002). Establishment and maintenance of a heterochromatin domain. *Science* **297**: 2232–2237.
Heitz E (1928). Das heterochromatin der moose. *Jb Wiss Bot* **69**: 762–818.
Henikoff S, Ahmad K, Malik HS (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**: 1098–1102.
Hilliker AJ (1976). Genetic analysis of the centromeric heterochromatin of chromosome 2 of *Drosophila melanogaster*: deficiency mapping of EMS-induced lethal complementation groups. *Genetics* **83**: 765–782.
Horvath JE, Schwartz S, Eichler EE (2000). The mosaic structure of human pericentromeric DNA: a strategy for characterizing complex regions of the human genome. *Genome Res* **10**: 839–852.
Hoskins RA, Smith CD, Carlson JW, Carvalho AB, Halpern A, kaminker JS *et al* (2002). Heterochromatic sequences in a *Drosophila* whole-genome shotgun assembly. *Genome Biol* **3**: research0085.1–0085.16.
John B (1988). The biology of heterochromatin. In: Verma RS (ed) *Heterochromatin: Molecular and Structural Aspects*. Cambridge University Press. pp 1–128.
Jordan IK, Rogozin IB, Glazko GV, Koonin EV (2003). Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet* **19**: 68–72.
Kapitonov VV, Jurka J (2005). RAG1 core and V(D)J recombination signal sequences were derived from *Transib* transposons. *PLoS Biol* **3**: e181 [E-pub].
Karpen GH, Le MG, Le H (1996). Centric heterochromatin and the efficiency of achiasmate disjunction in *Drosophila* female meiosis. *Science* **273**: 118–122.
Kidwell MG, Lish D (2000). Transposable elements and host genome evolution. *Trends Ecol Evol* **15**: 95–99.
Koryakov DE, Zhimulev IF, Dimitri P (2002). Cytogenetic analysis of the third chromosome heterochromatin of *Drosophila melanogaster*. *Genetics* **160**: 509–517.
Krantz ID, McCallum J, De Scipio C, Kaur M, Gillis LA, Yaeger D *et al* (2004). Cornelia de Lange syndrome is caused by mutations in *NIPBL*, the human homolog of *Drosophila melanogaster Nipped-B*. *Nat Genet* **36**: 631–635.
Kuln RM, Clarke L, Carbon J (1991). Clustered tRNA genes in *Schizosaccharomyces pombe* centromeric DNA sequence repeats. *Proc Natl Acad Sci USA* **88**: 1306–1310.
Kurek RA, Reugels M, Lammermann U, Buenemann H (2000). Molecular aspects of intron evolution in dynein encoding mega-genes on the heterochromatic Y chromosome of *Drosophila* sp. *Genetica* **109**: 113–123.
Lehnertz B, Ueda Y, Derijck AHA, Braunschweig U, Perez-Burgos L, Kubicek S *et al* (2003). *Suv39h*-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13**: 1192–1200.
Lohe AR, Hilliker AJ, Roberts PA (1993). Mapping simple repeated DNA sequences in heterochromatin of *Drosophila melanogaster*. *Genetics* **134**: 1149–1174.
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.
Marchant GE, Holm DG (1988). Genetic analysis of the heterochromatin of chromosome 3 in *Drosophila melanogaster*. II. Vital loci identified through EMS mutagenesis. *Genetics* **120**: 519–532.
Marygold SJ, Coelho CM, Leevers SJ (2005). Genetic analysis of *RpL38* and *RpL5*, two minute genes located in the centric heterochromatin of chromosome 2 of *Drosophila melanogaster*. *Genetics* **169**: 683–695.
Miller WJ, McDonald JF, Nouaud D, Anxolabehere D (1999). Molecular domestication: more than a sporadic episode in evolution. *Genetica* **107**: 197–207.
Moritz KB, Roth GE (1976). Complexity of germline and somatic DNA in *Ascaris*. *Nature* **259**: 55–57.
Myster SH, Wang F, Cavallo R, Christian W, Bhotika S, Anderson CT *et al* (2004). Genetic and bioinformatic analysis of 41C and the 2R heterochromatin of *Drosophila melanogaster:* a window on the heterochromatin–euchromatin junction. *Genetics* **166**: 807–822.
Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, Jones KM *et al* (2004). Sequencing of a rice centromere uncovers active genes. *Nat Genet* **36**: 138–145.
Pal-Bhadra M, Leibovitch BA, Gandhi SG, Rao M, Bhadra U, Birchler JA *et al* (2004). Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. *Science* **303**: 669–672.
Parks S, Wieschaus E (1991). The *Drosophila* gastrulation gene *concertina* encodes a Ga-like protein. *Cell* **64**: 447–458.
Pimpinelli S, Berloco M, Fanti L, Dimitri P, Bonaccorsi S, Marchetti E *et al* (1995). Transposable elements are stable structural components of *Drosophila melanogaster* heterochromatin. *Proc Natl Acad Sci USA* **92**: 3804–3808.
Platero JS, Csink AK, Quintanilla A, Henikoff S (1998). Changes in chromosomal localization of heterochromatin-binding proteins during the cell cycle in *Drosophila*. *J Cell Biol* **140**: 1297–1306.
Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, Nusinov DA *et al* (2003). Role of histone H3 lysine 27 methylation in X inactivation. *Science* **300**: 131–135.
Plath K, Mlynarczyk-Evans S, Nusinov DA, Panning B (2002). Xist RNA and the mechanism of X chromosome inactivation. *Annu Rev Genet* **36**: 233–278.
Reinhart BJ, Bartel DP (2002). Small RNAs correspond to centromere heterochromatic repeats. *Science* **297**: 1831.

Reiter LT, Potocki L, Chien S, Gribskov M, Bier E (2001). A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res* **11**: 1114–1125.

Rollins RA, Morcillo P, Dorsett D (1999). *Nipped-B*, a *Drosophila* homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. *Genetics* **152**: 577–593.

Schulze SR, McAllister BF, Sinclair DA, Fitzpatrick KA, Marchetti M, Pimpinelli S *et al* (2006). Heterochromatic genes in *Drosophila*: a comparative analysis of two genes. *Genetics* **173**: 1433–1445.

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.

Smith CD, Yandell M, Edgar RC, Kennedy C, Carlson J, Shengqiang S *et al* (2005). The Drosophila Heterochromatin Genome Project (DHGP): genes and repeat annotation. Seventh International Conference on Drosophila Heterochromatin, Gubbio, Italy.

Somma P, Fasulo B, Cenci G, Cundari E, Gatti M (2002). Molecular dissection of cytokinesis by RNA interference in *Drosophila* cultured cells. *Mol Biol Cell* **13**: 2448–2460.

Sun F, Haynes K, Simpson CL, Lee SD, Collins L, Wuller J *et al* (2004). *Cis*-acting determinants of heterochromatin formation on *Drosophila melanogaster* chromosome four. *Mol Cell Biol* **24**: 8210–8220.

Tonkin ET, Wang TJ, Lisgo S, Bamshad MJ, Strachan T (2004). *NIPBL*, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly *Nipped-B*, is mutated in Cornelia de Lange syndrome. *Nat Genet* **36**: 636–641.

Tulin A, Stewart D, Spradling AC (2002). The *Drosophila* heterochromatic gene encoding poly (ADP-ribose) polymerase (PARP) is required to modulate chromatin structure during development. *Genes Dev* **16**: 2108–2119.

Verdel A, Jia S, Gerber S, Sugiyama T, Gygi S, Grewal SI *et al* (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**: 672–676.

Volpe TA, Kidner C, Hall IM, Teng G, Grewal SI, Mrtienssen RA (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **297**: 1833–1837.

Von Stenberg RM, Novick GE, Gao GP, Herrera RJ (1992). Genome canalization: the coevolution of transposable and interspersed repetitive elements with single copy DNA. *Genetica* **86**: 106–137.

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

Weiler KS, Wakimoto BT (1995). Heterochromatin and gene expression in *Drosophila*. *Annu Rev Genet* **29**: 577–605.

Williams SM, Robbins LG (1992). Molecular genetic analysis of *Drosophila* rRNA arrays. *Trends Genet* **8**: 335–340.

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

Yasuhara JC, Marchetti M, Fanti L, Pimpinelli S, Wakimoto BT (2003). A strategy for mappimg the heterochromatin of chromosome 2 of *Drosophila melanogaster*. *Genetica* **117**: 217–226.

Yasuhara JC, Wakimoto BT (2006). Oxymoron no more: the expanding world of heterochromatic genes. *Trends Genet* **22**: 330–338.