The paradox of functional heterochromatin

Patrizio Dimitri,[1,2]* Nicoletta Corradini,[1,2] Fabrizio Rossi,[1,2] and Fiammetta Vernì[1]

Summary

Although heterochromatin has been studied for 80 years, its genetic function and molecular organization have remained elusive. In almost all organisms, heterochromatin has been regarded as genetically inactive chromosome regions. However, from genetic and genomic studies in *Drosophila melanogaster* and other organisms including humans, it is now clear that transcriptionally active domains are present within constitutive heterochromatin. These domains contain essential coding genes whose expression during development ensures the formation of the proper biochemical and morphological phenotypes, together with several gene models defined by genome annotation whose functions still need to be determined. *BioEssays* 27:29–41, 2005.

Introduction

The term heterochromatin was coined by Heitz in 1928[1] to describe chromosomal regions that appear deeply stained at prophase and remain compact throughout the mitotic cell cycle, as opposed to euchromatin, which is condensed at metaphase but diffuse during interphase. Heterochromatin may be either facultative and constitutive.[2] Facultative heterochromatin corresponds to silenced euchromatin (chromosome regions, entire chromosomes or even whole genomes), whilst constitutive heterochromatin, which occurs primarily in large blocks near centromeres and telomeres, consists mostly of repetitive DNA sequences, and maintains its characteristics on both homologous chromosomes.

Constitutive heterochromatin represents a significant fraction of eukaryotic genomes and shares similar genetic and molecular properties in nearly all animal and plant species (Table 1).[3] These properties have led to the traditional view of heterochromatin as a ''genomic wasteland'' or a repository of ''junk'' DNA, with little or no functional significance. However, this idea is becoming obsolete, as constitutive heterochromatin appears to be a genomic compartment carrying important functions with an unconventional structural and functional basis. In fact, studies in *D. melanogaster* have shown that constitutive heterochromatin is important for centromeric activity and chromosomal pairing and contains genes essential for viability and fertility.[4–11] In addition, a large number of gene models has recently been identified in constitutive heterochromatin of yeast, *A. thaliana, D. melanogaster*, rice and humans.[12–19] Here we bring together published and unpublished work to provide a picture of genes located in constitutive heterochromatin and their function. We feel that drawing attention to recent experimental evidence will help to change the current view that constitutive heterochromatin is genetically inert. This has important implications for genome organization and gene expression.

[1]Dipartimento di Genetica e Biologia Molecolare, Università ''La Sapienza''.
[2]Laboratorio di Genomica funzionale e Proteomica di sistemi modello, Università di Roma ''La Sapienza'', Via dei Sardi, 70-00185 Roma, Italy.
Funding agency: Ministero dell'Università e della Ricerca and Consiglio Nazionale delle Ricerche.
*Correspondence to: Patrizio Dimitri, Laboratorio di Genomica funzionale e Proteomica di sistemi modello, Università di Roma ''La Sapienza'', Via dei Sardi, 70-00185 Roma, Italy.
E-mail: patrizio.dimitri@uniroma1.it
DOI 10.1002/bies.20158
Published online in Wiley InterScience (www.interscience.wiley.com).

Facultative heterochromatin: when euchromatin becomes silent

A well-known example of facultative heterochromatin is the inactive X chromosome in somatic cells of female mammals.[20,21] The inactive X becomes heteropycnotic, suggesting that the chromatin in the silent regions is relatively condensed. The essential steps leading to inactivation can be summarized as follows: (1) initiation of the heterochromatinization process in early developmental stages, starting from a specific locus called the inactivation center (XIC), (2) spreading of heterochromatinization along the entire chromosome and (3) once established, the heterochromatic state is maintained through subsequent somatic cell divisions. Heterochromatinization is achieved by changing the chromatin of the X chromosomes from a transcriptionally active to an inactive state. This involves a cascade of chromatin modifications that inhibit the establishment of transcription complexes. These modifications are: (1) methylation of histone H3 lysine-9 and histone H3 lysine-27, (2) hypoacetylation of histones H2A, H3 and H4, (3) decrease of histone H3 lysine-4 methylation, and (4) changes in the time of DNA replication. Such features of inactive chromatin from the X chromosome seem to be shared by inactive chromatin elsewhere in mammalian genomes.

Table 1. The distinctive properties shared by constitutive heterochromatin in animal and plant species[2]

1) strongly reduced level of meiotic recombination
2) low gene density
3) repression of the activity of euchromatic genes when moved nearby, a phenomenon termed position effect variegation (PEV)
4) late replication during S phase
5) enrichment in highly and middle repetitive DNAs
6) transcriptional inertness
7) evolutionary lability, in that closely related species may differ greatly in the amount, location and type of repetitive sequences

Together, these properties have reinforced the idea that constitutive heterochromatin was a dispensable biologial material.

Constitutive heterochromatin and its general properties

Constitutive heterochromatin forms about 5% of the genome in *Arabidopsis thaliana*, 20% in humans, 30% in *Drosophila melanogaster* and up to 85% in certain nematodes,[4,15,22] and shares distinctive properties in nearly all animal and plant species[3] (see Table 1). These properties have often been interpreted as consistent with the view that constitutive heterochromatin is a repository of ''junk'' DNA, with little or no functional significance. Observations in different organisms have supported this idea. First, aneuploidy for some heterochromatic regions in *Drosophila melanogaster* has little effect, as opposed to the dramatic effects of aneuploidy of euchromatin.[23] In other words, all duplications and heterozygous deficiencies and even some homozygous deficiencies are without detectable effect on the viability, fertility and morphology of flies.[24] Second, almost all heterochromatin can be specifically eliminated from somatic nuclei during early development in nematodes and other organisms.[25] Third, in certain polytene tissues of *Drosophila* and other Diptera much of the DNA of constitutive heterochromatin (mainly highly repetitive DNAs) fails to undergo DNA replication during polytenization, and is therefore dramatically underrepresented compared with euchromatin.

These observations need to be revised in the light of additional data. First, only deletions of a specific part of the X-chromosome heterochromatin are homozygous viable, and all known deficiencies of the heterochromatin of major autosomes are homozygous lethal or have negative effects on fitness.[26] Second, in the last two decades several essential genes and gene models have been identified in *Drosophila melanogaster*, yeast, *Arabidopsis*, rice and humans.[12–19] Heterochromatic genes of *Drosophila* tend to be very large: for example, the *Y* chromosome fertility factors have been estimated to comprise up to 4 Mb of heterochromatic DNA.[4] Third, some heterochromatic genes are cryptic and can escape conventional genetic analysis.[27] Thus the density of genetic functions in constitutive heterochromatin might not be as low as previously claimed. Finally, although constitutive heterochromatin is eliminated in the somatic cells of certain organisms, it is retained in germline cells. Similarly, underreplication of heterochromatin in *Drosophila* and other Diptera only occurs in polytene tissues, while diploid cells and germline cells retain a normal content of heterochromatic DNA. It thus appears that, during evolution, different organisms have developed different strategies to get rid of large heterochromatic segments that may be dispensable in certain tissues and/or during specific phases of development. In contrast, the stability of the heterochromatin content in cells of the germline of an organism in which it is eliminated or underreplicated somatically raises the possibility that its maintenance is under selective pressure. In conclusion, the ubiquity and persistence of heterochromatin in a wide variety of eukaryotic genomes suggest that this distinctive genomic component has an adaptive value.

Repetitive DNA sequences in constitutive heterochromatin

Highly repetitive satellite DNA is characteristic of heterochromatin;[3] it is made up of short units that are tandemly repeated up to millions of times, forming large blocks scattered through the heterochromatic regions of all chromosomes. Although highly repetitive DNA lacks coding sequences, several hypotheses have been proposed to explain its ubiquity.[3] According to a recent idea,[28] blocks of satellite DNA do not form a specific structure, but instead may be passive spacers needed to define a centromere.

In addition to satellite DNA, there is growing evidence for an accumulation of transposable element-related sequences (TEs) in constitutive heterochromatin of evolutionarily distant organisms.[29] In fact, 52% of the 20.7 Mb WSG3 of *D. melanogaster* heterochromatin sequence has been estimated to be similar to known TEs, while TE-related sequences only account for about 4–5% of the Release 3 euchromatic sequence.[12,30] The accumulation of TEs represents one of the most intriguing aspects of the structure and organization of heterochromatin.[29,31–33] However, the mechanism of colonization of heterochromatin by TEs remains unclear. The original idea that heterochromatin is merely a ''rubbish bin'' or ''graveyard'' of transposable elements has been modified in light of experimental evidence suggesting that TE-related sequences in heterochromatin may contribute to many of the structural and functional properties of heterochromatin (see discussions below).

Essential heterochromatic genes of *Drosophila melanogaster*

D. melanogaster is the model organism in which the greatest progress in the study of heterochromatin function has been made, due to the ability to combine genetic, cytological and

genomic approaches. Essential genes for viability (lethal mutable genes) and fertility (Y chromosome fertility factors), which reside in constitutive heterochromatin, were initially discovered in *D. melanogaster*. In particular, the *light* gene was the first heterochromatic protein-coding sequence to be molecularly characterized.[34,35] Here we provide a detailed description of a group of essential genes located in the heterochromatin of chromosomes 2 and 3. They were initially identified by recessive lethal mutations genetically linked to heterochromatin.[26,36,37] Complementation analysis using rearrangements with cytologically determined breakpoints in heterochromatin of mitotic chromosomes finally demonstrated that these genes are in pericentric heterochromatin and allowed their mapping.[38,39] Although the release of the *D. melanogaster* heterochromatin sequence by the Berkeley Drosophila Genome Project (BDGP) has greatly facilitated studies on the molecular organization and function of heterochromatic genes,[12,40] most of the known essential genes in the heterochromatin of chromosomes 2 and 3 are still unknown molecularly.

Heterochromatic genes of chromosomes 2 and 3

Genetic analyses have defined at least 32 essential genes mapping to the mitotic heterochromatin of chromosome *2* (Table 2; Fig. 1A,B). However, only a few have been defined molecularly. The *light* gene product is involved in cellular protein trafficking,[44] while *concertina* encodes a maternal α-like subunit of a G protein essential for gastrulation.[45] The *rolled* gene was shown to be required for imaginal disc development and suggested to be involved in cell proliferation.[26,38] The *rolled* product is a mitogen-activated protein (MAP) kinase which is required in the signal transduction pathway of the *sevenless* gene[46] and may also mediate the spindle integrity checkpoint.[47] *Nipped-A* and *Nipped-B* may be involved in chromosome condensation and transcriptional regulation.[41,42] Two other genes in the heterochromatin of the right arm of chromosome *2* (*2R*h), *l(2)41Aa* and *l(2)41Ad* (Table 2; Figs. 1B, 2), which have not yet been characterized molecularly, may be required for chromosome condensation[48] and for proper leg and wing morphogenesis,[26,38] respectively.

Table 2. The known vital genes resident in the heterochromatin of chromosomes 2 and 3

Gene	Mitotic map	Size	Function	Scaffold
l(2)40Fa	h35		unknow	—
l(2)40Fc	h35		unknow	—
light	h35	16kb	cellular-protein traffcking	AABU01002768
concertina	h35	10.5 kb	gastrulation	AABU01002768
l(2)40Fd	h35		unknown	—
l(2)40Fe	h35		unkown	—
l(2)40Ff	h35		unknown	—
l(2)40Fg	h35		unknown	—
l(2)41Ab	h39–40		unknown	—
l(2)41Aa	h41		chromosome condensation	—
rolled	h41	60kb	signal transduction pathway	AABU01001947
l(2)41Ad	h43–h44		legs and wing morphogenesis	—
l(2)41Ae	h46		unknown	—
l(2)41Af	h46		ribosomal protein RpL38	AABU01002769
Nipped-B	h46	39kb	transcriptional regulation, chromosome condensation; mutations in *NIPBL* human homologue are responsible for the Cornelia de Lange syndrome	AABU01002769
l(2)41Ah-Nipped-A	h46	38kb?	transcriptional regulation; significant identity with mammalian TRRAP and yeast Tral proteins	AE03787
l(3)80Fa	h47–49		unknown	—
l(3)80Fb	h47–49		unknown	—
l(3)80Fc	h47–49		unknown	—
l(3)80Fd	h50		unknown	—
l(3)80Fe	h50		unknown	—
l(3)80Ff	h50		unknown	—
l(3)80Fg	h50		unknown	—
l(3)80Fh	h50–51		member of the *trithorax* group	—
l(3)80Fi	h50–51		growth and development	
l(3)80Fj	h50–51		member of the *trithorax* group	—
Parp	h54–55	95kb	polyADP	AABU01002763
l(3)81Fa	h58		unknown	—
l(3)81Fb	h58		unknown	

All the genes have been identified by conventional genetic analysis before the release of the *Drosophila melanogaster* genome sequence.

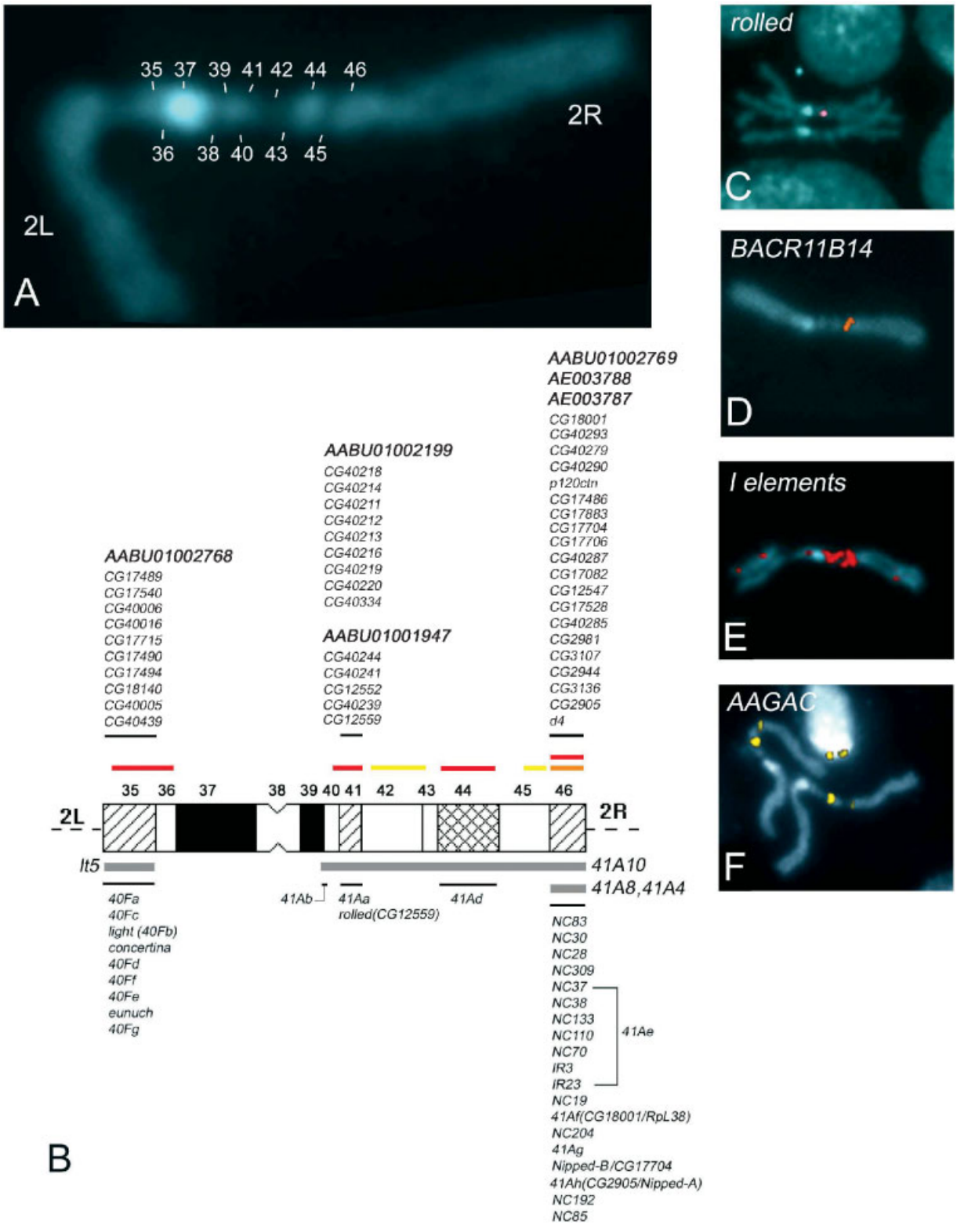
35 37 39 41 42 44 46
36 38 40 43 45
2R
2L
A
rolled
C
BACR11B14
D
I elements
E
AAGAC
F
AABU01002769
AE003788
AE003787
CG18001
CG40293
CG40279
CG40290
p120ctn
CG17486
CG17883
CG17704
CG17706
CG40287
CG17082
CG12547
CG17528
CG40285
CG2981
CG3107
CG2944
CG3136
CG2905
d4
AABU01002199
CG40218
CG40214
CG40211
CG40212
CG40213
CG40216
CG40219
CG40220
CG40334
AABU01001947
CG40244
CG40241
CG12552
CG40239
CG12559
AABU01002768
CG17489
CG17540
CG40006
CG40016
CG17715
CG17490
CG17494
CG18140
CG40005
CG40439
35 36 37 38 39 40 41 42 43 44 45 46
2L
2R
lt5
41A10
41A8,41A4
40Fa
40Fc
light (40Fb)
concertina
40Fd
40Ff
40Fe
eunuch
40Fg
41Ab
41Aa
rolled(CG12559)
41Ad
NC83
NC30
NC28
NC309
NC37
NC38
NC133
NC110
NC70
IR3
IR23
41Ae
NC19
41Af(CG18001/RpL38)
NC204
41Ag
Nipped-B/CG17704
41Ah(CG2905/Nipped-A)
NC192
NC85
B

Figure 1.

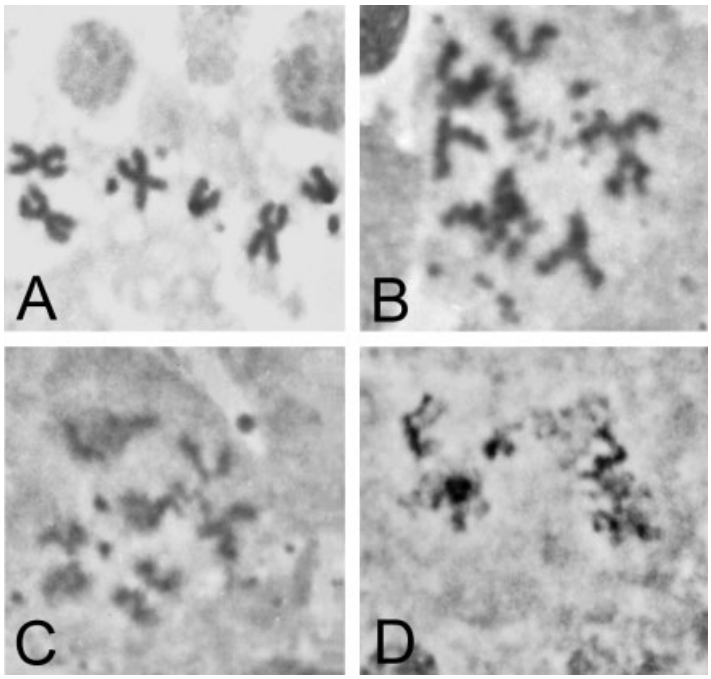

Figure 2. Chromosome condensation defects observed in *l(2)41Aa* mutants. Metaphases were prepared from larval neuroblast cells of third instar larvae emyzygous for *EMS-31* and *IR18*, two lethal alleles of *l(2)41Aa*. **A:** Colchicine-treated control metaphase of the Oregon-R strain stained with orcein; **B–D:** Phenotype of colchicine treated metaphases of *l(2) 41Aa* mutants stained with orcein.

In a recent genetic and bioinformatic analysis of *D. melanogaster* heterochromatin, Myster et al.[42] identified 18 complementation groups in *2R*h. They are all uncovered by two deficiencies, *Df(2R)MS41A10* and *Df(2R)MS41A8*,[38] and are therefore likely to be located in h46, the most-distal portion of *2R*h (Fig. 1B). Three of them correspond to the vital genes *l(2)41Af*, *Nipped-B* and *Nipped A-l(2)41h* previously mapped to h46. Interestingly, S. Marygold and S. Leavers (personal communication) found that *l(2)41Af* is allelic to *M(2)41A*, and corresponds to the gene model CG18001 which encodes the RpL38 ribosomal protein. Lethal mutations in seven other genes identified by Myster et al.[42] all fail to complement a *l(2)41Ae* lethal allele used as tester. This suggests that the *l(2)41Ae* locus, which exhibited complex complementation behaviour,[26] indeed contains different vital genes.

Thirteen vital genes have been mapped to the heterochromatin of chromosome *3* (Table 2; Fig. 3).[36,39,49] Three of the genes from the left arm, *l(3)80Fh, l(3)80Fi* and *l(3)80Fj*, have been cloned and found to correspond to single-copy

Figure 1. Mapping of essential genes, gene models and repetitive sequences to the heterochromatin of chromosome 2. **A:** Banding pattern of mitotic heterochromatin of chromosome *2* after DAPI which is a general indicator of AT-richness and differentiates heterochromatin into several regions with different degrees of fluorescence. The numbers indicate the different heterochromatic regions; *2L* = left arm of chromosome 2; *2R* = right arm of chromosome 2. **B:** Cytogenetic map showing the location of both essential genes (below) and gene models (above) within the heterochromatin of chromosome *2*; dark blocks are brightly fluorescent, cross-hatched blocks are moderately fluorescent; hatched blocks are weakly fluorescent; open blocks are nonfluorescent. The 32 vital genes mapped to the constitutive heterochromatin of chromosome *2*[26,38,41,42] are non-randomly distributed, in that most of them are clustered within regions h35 and h46, which represent the most-distal portions of mitotic heterochromatin.[38] Forty-four additional gene models (above) were also localized to specific heterochromatic regions of chromosome 2.[12,13] They were mapped mainly by FISH with probes of large genomic regions cloned in bacterial artificial chromosomes (BACs), or through their molecular association with known mapped genes. The location of AABU01002199 with its nine gene models in region h41 relies on FISH mapping of a single P-element insertion (LP1) which disrupt *CG40218* (N. Corradini, F. Vernì and P. Dimitri, unpublished). Scaffold designation is shown at the top of each gene model list. Most of the genes, known and predicted, are located in weakly DAPI-fluorescent chromosomal regions, which harbor clusters of transposable element-homologous sequences and are devoid of highly repetitive satellite DNAs.[43] **C:** FISH mapping of the *rolled* gene to region h41; **D:** FISH mapping of BACR11B14 to region h46; note that this BAC carries six genes (*CG2981*, *CG3107*, *CG2944*, *CG3136*, *CG2905/Nipped-A* and *d4*); **E:** FISH mapping of I-element DNA to multiple locations in *2L*h and *2R*h. **F:** FISH mapping of AAGAC satellite DNA to regions h42–43 and h45. The cytological locations of the different DNA sequences mapped by FISH are indicated in the diagramatic map by coloured lines (see panel B); grey lines show the extent of deficiencies.

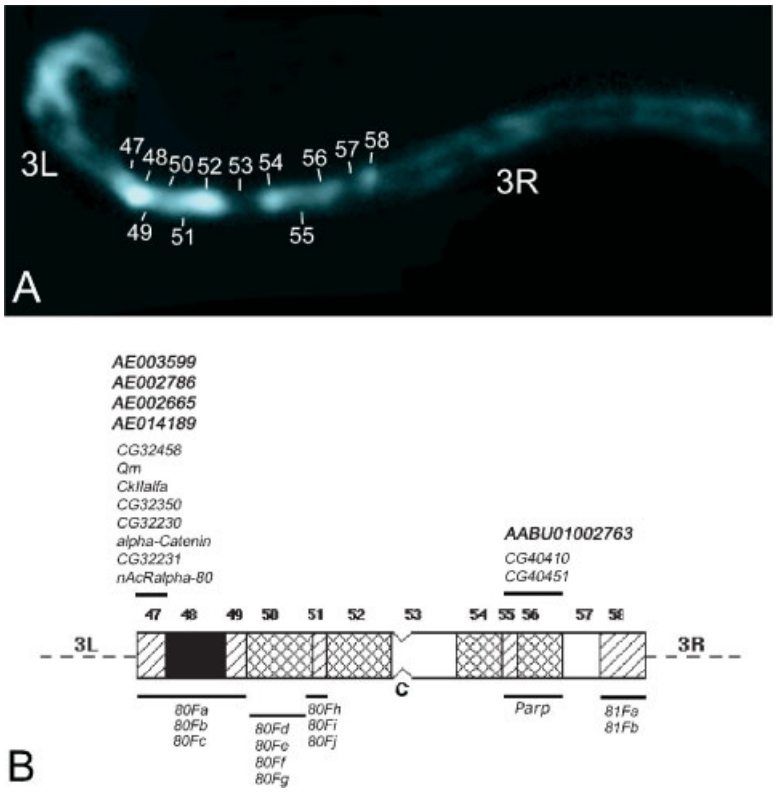

Figure 3. Mapping of essential genes and gene models in the heterochromatin of chromosome *3*. **A:** Banding pattern of the heterochromatin of chromosome *3* after DAPI staining of prometaphase chromosomes; the numbers indicate the different heterochromatic regions; *3L* = left arm of chromosome *3; 3R* = right arm of chromosome *3*. **B:** Cytogenetic map of the heterochromatin of chromosome *3* showing the location of the essential genes (below) and of some of the gene models (above); dark blocks are brightly fluorescent, cross-hatched blocks are moderately fluorescent; hatched blocks are weakly fluorescent; open blocks are nonfluorescent. Scaffold designation is shown at the top of each gene model list. Thus far, a group of 13 vital genes (below) has been associated with the heterochromatin of chromosome *3*.[49] At least seven genes of the left arm (from *l(3)80Fd* to *l(3)80Fj)* map to segment h49–h51 and the most distal genes (from *l(3)80Fa* to *l(3)80Fc)* are located within regions h47–h49. Two genes on the right arm, *l(3)81Fa* and *l(3)81Fb*, map to the distal h58 segment. A group of eight gene models (above) map to h57.[13] PARP, together with two gene models, maps to h55–56.

sequences. Interestingly, *l(3)80Fh* and *l(3)80Fj* appear to be members of the *trithorax* group (*trxG*), while *l(3)80Fi* may have key functions in growth and development.[49] In addition, *Parp*, which encodes a polyADP, has been identified recently and mapped to region h54–55 of the mitotic heterochromatin map.[50] Other single-copy genes such as α*-Cat*, *rp21*, *SCP*, *DSK*, *QIII*, *ziti* and *Dbp80* map to the pericentromeric heterochromatin of chromosome *3*, but it is still not known if these genes are allelic to known vital genes.[36]

Gene models in *Drosophila melanogaster* heterochromatin

The number of active genes in constitutive heterochromatin of *D. melanogaster* may be higher than that defined by genetic analysis, as the annotation of the heterochromatin sequence predicted about 450 computed genes (CG) or gene models[12,51] in all chromosomes. However, the precise location of a large proportion of gene models is still unknown, since they are present in unmapped scaffolds. Using fluorescence in situ hybridization (FISH), we have found that at least 20 gene models map to the h46 region of *2R*h (Fig. 1B), and eight map to the h47 regions of *3L*h[13] (Fig. 3B). Thus at least 28 gene models are located in these regions. It is worth noting that, thus far, of the 19 vital genes mapped to h46 only *41Af, Nipped-B* and *Nipped-A* were found to correspond to gene models present in the assembled sequence of this region.[12,13,42] About another 60 gene models are localized in the euchromatin–heterochromatin transition regions of *2R* and *3L*.[13] Furthermore, Hoskins et al.[12] found that approximately 80 additional gene models are associated with the heterochromatin of chromosomes 2 and 3,[12] some of which have also been mapped to specific regions of mitotic heterochromatin (Fig. 1B). However, some predictions of genes may represent portions of the same gene, or may be spurious. This may have led to an overestimate of the actual number of genes in heterochromatin.

Heterochromatic genes in other species

Are homologues of *D. melanogaster* heterochromatic genes also found in heterochromatin in other organisms? The *light* and *rolled* genes are located in constitutive heterochromatin of closely related species from the melanogaster subgroup[52] (P. Dimitri, unpublished data), but the *Drosophila virilis light* gene is in euchromatin.[52] During evolution transposable elements may promote transfer of genes into constitutive heterochromatin by stimulating chromosome rearrangements or by transferring genes during retrotransposition, similar to the *L1* element in humans.[53] Once in heterochromatin,

recurrent insertions of a plethora of transposable elements could help to generate the present structural organization of heterochromatic genes in *Drosophila*. At present, it is unclear whether the expression pattern of the euchromatic *light* gene in *D. virilis* is different from that of its heterochromatic orthologue in *D. melanogaster*. Homologues of protein-coding genes resident in constitutive heterochromatin of *D. melanogaster*, such as *rolled*, *PARP*, *Nipped-A*, *Nipped-B*, *RpL38* and *p120catenin*, have been found in several organisms, including yeast, mouse and humans, but they are all located in euchromatin. For example, *NIPBL*, the human homolog of the *Drosophila melanogaster Nipped-B* gene, maps to euchromatin of chromosome 5p13 and is widely expressed in fetal and adult tissues. Mutations in *NIPBL* are responsible for Cornelia de Lange syndrome (CdLS), a multiple malformation disorder.[54,55] The genomic region of *NIPBL* spans 188 kb and comprises 47 exons. The protein sequence of human *NIPBL* shares 92% identity with mouse, 88% with rat, and 37% with the *Drosophila melanogaster Nipped-B* gene product. Although the function and regulation of known heterochromatic genes and their homologues in other eukaryotes is far from being elucidated, the data suggest that a heterochromatic location may not be crucial for the proper function of a given gene.

The presence of essential genes in heterochromatin appears to be a conserved trait in the evolution of eukaryotic genomes. Heterochromatic genes have recently been identified in *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Arabidopsis thaliana*, rice and humans. In *S. cerevisiae* expressed genes map within 1 kb of essential centromere sequences while, in *S. pombe*, an unusual clustering of tRNA genes is located within an 80 kb centromere-specific DNA sequence. The centromere-associated B′ and B sequences CEN1 and CEN2 of *S. pombe* contain a cluster of tRNA genes, 22 in the CEN2 region and at least six in the CEN1 region, and some of these tRNAs are expressed.[14]

In *A. thaliana*, at least 47 predicted genes have been identified that are expressed in the centromeric regions. As in *Drosophila*, the single-copy genes of *A. thaliana* are organized in islands of unique sequences that contain transposable element-related DNAs.[15] Proteins encoded by these genes are members of functional families such as membrane proteins, nitrate reductase or polyubiquitin[56–60] (for a list see http://preuss.bsd.uchicago.edu/tables.html). Transposable element-associated gene products such as transposase or RNA helicase are also present in centromeric heterochromatin of *A. thaliana*, which may reflect the abundance of mobile elements in these regions.

Mapping and sequencing of the human genome indicates that pericentromeric heterochromatin is characterized by several blocks of duplicated sequence, probably generated by transposition.[16,17] Fragments of genes, complete genes and repeats are duplicated in pericentromeric regions. Generally the pericentromeric duplications are non-functional pseudogenes, but some mRNAs or ESTs from pericentromeric sequences have been identified. Genes coding for growth factors, immunoglobins κ, λ and D, plasminogen and others have been found in these paralogous sequences (for a list see Ref. 18). Moreover, many pericentromeric paralogous sequences are transcribed in germline, fetal or cancerous tissue,[17,18] suggesting a role for these genes in biological processes.

Finally, sequencing of the rice genome detected fourteen predicted and at least four active genes interspersed in a region of approximately 750 kb comprising the centromere of chromosome 8, which is embedded in heterochromatin.[19]

Heterochromatic genes versus euchromatic genes

In general, in yeast, *Arabidopsis, Drosophila*, rice and humans, both known genes and gene models located in heterochromatin do not seem to have molecular functions that distinguish them from euchromatic genes.[12,34,40,41,45,46,49,50,61] In addition, as for genes found in euchromatin, vital heterochromatic genes of *D. melanogaster* such as *rolled, light Nipped-B* can be widely expressed during development, while others, like the Y-chromosome fertility factors, have tissue- and sex-specific limited expression. This also seems to be true in other organisms. However, heterochromatic genes might have different regulatory requirements from those of euchromatic genes. In fact, these genes are repressed when moved to euchromatin by chromosomal rearrangements, indicating that they are dependent on a heterochromatic location for correct expression[62,63] (see below). As mentioned in the introduction, heterochromatic DNA is replicated later in S-phase than the bulk of euchromatic sequences. A similar dichotomy may also occur at the transcriptional level: heterochromatic and euchromatic genes may differ in their timing of expression during the cell cycle. Cell-cycle-dependent changes in chromosomal localization of heterochromatin-binding proteins have indeed been observed in *Drosophila*.[64]

The molecular organization of heterochromatic and euchromatic genes is clearly different. For example, the known genes located in *D. melanogaster* heterochromatin are on average much longer than euchromatic genes, due to the occurrence of long introns enriched in transposable element (TE)-related DNA sequences.[34,61,65–67] At least 50% of intronic DNA from the *rolled* (60 kb long), *Nipped-B* (39 kb long) and *light* (15 kb long) genes is composed of degenerate retroelements and DNA transposons.[67] The example of the Y-chromosome fertility factors of *D. melanogaster* is even more dramatic, since they contain up to 4 Mb of DNA[4] and carry TE-rich mega-introns.[61,66] Many of the gene models that we mapped to h46 and h47 also have long introns (for example *d4* on 2R and *alpha-catenin* on 3L,[12] but the average density of TEs appears to be lower than that found in introns of known heterochromatic genes.

It has been suggested that TEs not only contribute to the build-up of heterochromatic introns, but may also confer evolutionary plasticity on the structure of genes located in constitutive heterochromatin[67] or may become a functional part of these genes.[29] For example, regulatory sequences of heterochromatic genes may have originated from the long terminal repeats (LTR) of retrotransposons, which contain strong transcriptional enhancers. Evidence for the adaptive significance of a TE-derived sequence in a heterochromatic gene was recently reported for the LTR of *Quasimodo* located in intron I of *chitinase-3*[68] in *D. melanogaster*. Additional candidates for comparable roles may be the LTR portion of *17.6* located in intron IV of *rolled*, and a 120 bp sequence located in the promoter region of *light* which shares 90% identity with the terminal inverted repeat (TIR) of the *S*-element.[67] Further genomic and molecular analyses are required to assess the possible involvement of TE sequences in the function and expression of heterochromatic genes.

The paradox of functional heterochromatin

An important goal is to solve the paradox of "functional heterochromatin". In other words, how can protein-coding genes normally resident in constitutive heterochromatin work properly in an environment that has been thought to be incompatible with gene expression? In fact, active heterochromatic genes may show a nucleosome array characteristic of euchromatic genes, while being flanked by repetitive sequences packaged in a heterochromatic fashion, with long-range nucleosomal ordering.[69] Nevertheless, genes resident in heterochromatin do not seem to be merely euchromatic sequences embedded in a repetitive environment, since their correct expression is dependent on location in heterochromatin.[62,63] We still need to understand, however, how domains of gene expression are organized in heterochromatin and what accounts for the difference between heterochromatic and euchromatic domains. What factors determine the correct expression of single-copy genes resident in heterochromatin? It has been suggested that chromosomal proteins known to be required for the establishment of the heterochromatic state, such as HP1 (see below), are also involved in the control of gene expression in heterochromatin.[5] The normal expression of heterochromatic genes may require a critical concentration of heterochromatin proteins or protein complexes that are present in a limited amount in the nucleus.

Heterochromatin proteins and gene expression

Several conserved nonhistone proteins that play important roles in chromatin structure and gene regulation have been identified as components of heterochromatin in *Drosophila* and other eukaryotes. Here we will focus on HP1, SU(VAR)3-9, SU(VAR)3-7, *ORC*, HP2 and HOAP.

Heterochromatin protein 1 (HP1)

This protein was first identified in *D. melanogaster* by mutations in the gene *Su(var)2-5* that suppress position-effect variegation (PEV), and was found to bind predominantly to the heterochromatic chromocenter of polytene chromosome.[70–72] HP1 has two structural motifs, the chromodomain thought to be involved in chromatin binding[73] and the chromoshadow domain involved in protein–protein interactions.[74] HP1 homologues exist in various other organisms, including fission yeast, mice and humans, where they may have conserved functions. For example, the swi6 homolog of HP1 in *Schizosaccharomyces pombe* localizes to heterochromatin, including the centromeres, telomeres and the donor mating-type loci, and is involved in silencing at these loci. The chromodomain from a mammalian HP1-like protein, M31, can functionally replace that of swi6, showing that chromodomain function is conserved from yeasts to humans.[75,76] In addition to its involvement in heterochromatin formation, HP1 may have a role in telomere capping.[77,78] HP1 also associates with some euchromatic regions in *D. melanogaster*,[71,77,78] suggesting that it is also required for the repression of specific euchromatic genes.[6] This is supported by studies showing that HP1 represses the activity of three genes in polytene chromosome region 31, where it is strongly accumulated,[79] while, in mammals, it is involved in silencing euchromatic genes by interacting with the retinoblastoma (Rb) protein.[80] A detailed cytological analysis of HP1 in *D. melanogaster* populations and closely related species revealed that, in addition to its localization in heterochromatin and telomeres, HP1 is present at about 200 euchromatic sites.[81] Some of these sites (including chromosomal region 31) are conserved in all populations and in several *Drosophila* species. The HP1 euchromatic sites do not appear to be present in regions of intercalary heterochromatin or to be enriched in transposable element-like sequences.

HP1 and gene expression

The above results have reinforced the suggestion that HP1 is involved in genetic silencing. However, the recent finding that HP1 is associated with both ecdysone-induced and heat-shock-induced puffs[82] strongly suggested that HP1 may also act as a gene activator. As previously mentioned, it has been proposed that HP1 or other heterochromatin proteins may be required for the expression of genes resident in constitutive heterochromatin. Experimental evidence consistent with this has been reported from *D. melanogaster*. Firstly, genetic experiments suggested that different Su(var) gene products can interact to guarantee the proper expression of the *light* gene in its normal heterochromatic location.[83] Secondly, the amount of mRNA of *light* and *rolled* heterochromatic genes was found to be reduced about 2.5-fold in HP1 mutant larvae.[84]

HP1 and Su(var) 3-9

It is unclear how HP1 can repress or activate gene expression. The HP1 homologue in mammals associates with chromatin by binding histone H3 tails that are methylated on lysine 9. The specific methyltransferase activity is provided by the product of *Suv39H1*. The binding of the HP1 homologue to histone H3 and to SUV39H1 protein is mediated by the chromodomain and the shadow domain, respectively. Such a triangle of interactions has also been demonstrated in *D. melanogaster*. HP1 and H3 Lys 9met co-localize at the centric heterochromatin and at chromosomal region 31. The methyltransferase activity in *D. melanogaster* is encoded by the *Su(var) 3-9* gene (the homolog of *Suv39H1*), which also is involved in silencing of reporter genes in heterochromatin.[5] As in mammals, the *Su(var) 3-9* product co-localizes with HP1.[85,86] However, it is not yet clear how the methyltransferase is directed to the chromatin.

To obtain more insight into the roles of HP1 and SU(VAR)3-9, large-scale mapping of their target genes has recently been done in *Drosophila* embryonic Kc cells.[87] The results revealed that HP1 and SU(VAR)3-9 bind together to genes and transposable elements in constitutive heterochromatin. Independently of HP1, SU(VAR)3-9 also binds to a distinct set of non-pericentric genes. In addition, HP1, without SU(VAR)3-9, binds to many genes on chromosome 4. Pericentric target genes bound by both HP1 and SU(VAR)3-9 have expression levels similar to those of non-target genes in Kc cells. In contrast, genes bound by SU(VAR)3-9 alone are more repressed. HP1 and SU(VAR)3-9 target genes in pericentric heterochromatin are predominantly embryo-specific, while non-pericentric target genes of SU(VAR)3-9 are male-specific. These findings suggest that, depending on chromosome location, HP1 and SU(VAR)3-9 proteins could act on specific sets of developmentally regulated genes by forming different complexes. On the one hand, HP1 may associate with the chromatin of chromosome 4 independently of SU(VAR)3-9, perhaps by the action of a different H3 K9 methyltransferase. On the other hand, HP1 may fail to bind to SU(VAR)3-9-specific target genes because the histone tails are unmethylated or associated with a different chromatin protein. Surprisingly, Greil et al.[87] were unable to detect any changes in mRNA expression of *light, rolled* or other putative HP1 and *Su(var) 3-9* target genes after induction of dsRNAi of the *Su(var) 205* and *Su(var) 3-9* genes in *Drosophila* Kc cells. These results clearly conflict with previous in vivo data.[83,84] Similarly, dsRNAi-mediated inactivation of HP1 and other telomeric proteins in *Drosophila* S2 or Kc cells does not result in the telomeric fusions usually found in mutant larvae (M. Gatti and G. Cenci, personal communication). Although dsRNAi in tissue culture cells is an extremely powerful method to specifically abolish the expression of several genes, there may be genetic and regulatory differences between cultured cells and the living organism that might explain such discrepancies. For example, in cultured cells some genes might be redundant or a given function might also be provided by protein products different from those required in vivo. Additional explanations have also been proposed.[87] Thus, to test a possible role of heterochromatin complexes on heterochromatic gene expression, further experimental approaches are required: for example, in vivo dsRNAi of heterochromatin protein genes with the Gal4-UAS system.[88]

HP1 and Origin Recognition Complex

It has been proposed that the correct localization of HP1 in *Drosophila* heterochromatin may involve the Origin Recognition Complex (ORC), which is required to initiate eukaryotic DNA replication and to promote transcriptional silencing at the mating-type loci in *S. cerevisiae*.[89] In fact, an underphosphorylated HP1 isoform in the maternally loaded cytoplasm of early *D. melanogaster* embryos is found in a high molecular-weight multiprotein complex that contains ORC subunits.[90] One of the better characterized subunits of the ORC is DmORC2,[91] which associates with heterochromatin in interphase and mitosis. *Orc2* mutants are dominant suppressors of variegation and recessive lethals that exhibit chromosome condensation defects but lack telomere fusions. Compatible with this cytological phenotype, ORC2 and HP1 co-localize in heterochromatin but not at telomeRes It has been proposed that ORC is required to recruit underphosphorylated HP1 isoforms to heterochromatin. It has been suggested that these isoforms function in the nucleation phase of chromatin assembly.[90]

HP1 and Suvar 3-7

Other genes, most of which are suppressors of variegation, seem to play a role in heterochromatin formation. *Suvar 3-7* codes for a protein that contains seven spaced zinc fingers.[92,93] This protein associates with pericentromeric heterochromatin at interphase, and on diploid chromosomes from embryonic nuclei and polytene chromosomes.[94] SU-VAR3-7 and HP1 proteins colocalize not only in heterochromatin, but also at a limited set of sites in euchromatin and at telomeres.[94] They interact genetically and co-immunoprecipitate from nuclear extracts. The chromoshadow domain of HP1 is involved in this interaction. It has been demonstrated that an increase in the amount of these two proteins enhances the silencing effect. Moreover, and in contrast with the effect on euchromatic genes, a decrease in the amounts of both proteins enhances variegation of the *light* gene.[94] Taken together, these results suggest that SU VAR 3-7 and HP1 cooperate in the genomic silencing associated with heterochromatin.[95]

HP1, HP2 and HOAP

HP1 also cooperates with HP2 in heterochromatin-mediated silencing.[96] HP2 has two isoforms, the larger of which

contains two AT-hook domains that bind AT-rich regions. Some alleles of HP2 are PEV suppressors.[96] On polytene chromosomes, this protein co-localizes with HP1 in pericentric heterochromatin, on chromosome 4 , in region 31 and at some but not all euchromatic sites. HP2 and HP1 interact in the two-hybrid screen and co-immunoprecipitate. The HP1–HP2 interaction is mediated by part of the hinge region and by the entire shadow domain. A further chromatin protein called HOAP (HP1/Origin Recognition Complex-Associated Protein), which is similar to HMG proteins, has been recently isolated from an ORC2- and HP1-containing complex.[90,97] HOAP associates with telomeres in mitotic, polytene and interphase chromosomes in a sequence-independent manner.[98] Mutants in the HOAP-coding gene exhibit extensive telomere–telomere fusions, suggesting a role in telomere capping. Although HP1 and HOAP co-localize at telomeres, the presence of HP1 does not seem to be required for the correct localization of HOAP. Perhaps other telomere-binding factors are required to drive HOAP to telomeric sequences. It has recently been demonstrated that peptides that interfere with HP1/HOAP interactions in co-precipitation experiments displace HP1 from the heterochromatic chromocenter of salivary gland polytene chromosomes.[99] Since HOAP has a predominantly telomeric localization,[98] this effect at the chromocenter was unexpected.

Heterochromatin formation mediated by RNA interference in evolutionarily distant organisms

RNA interference (RNAi) is an evolutionarily conserved gene-silencing pathway in eukaryotes. There is growing evidence that RNAi mediated by repetitive sequences is required to establish silencing in heterochromatic domains in evolutionarily distant organisms. Aberrant accumulation of double-stranded RNA from centromeric transposons seems to induce the loss of histone H3 Lys9 methylation in yeast.[100,101] Moreover, in *S. pombe*, retrotransposon long terminal repeats (LTR) can recruit heterochromatin complexes and induce RNAi-dependent chromatin silencing.[102] Centromeric transposable element repeats capable of generating dsRNA may have a role in mediating the formation of heterochromatin in *S. pombe*.[100,103] Similarly, in mouse, small heterochromatic RNAs generated from major satellite transcripts may guide recruitment of the Suv39h HMTases to direct H3-K9 trimethylation and, in turn, DNA methylation in pericentric heterochromatin.[104] Loss of Dicer function, one of the components of the RNAi machinery, in a chicken–human hybrid DT40 cell line that contains human chromosome 21, results in cell death, premature sister-chromatid separation, aberrant accumulation of alpha-satellite transcripts, and abnormalities in the localization of heterochromatin proteins.[105] A functional RNAi machinery for the establishment of the correct pattern of histone methylation and concomitant gene silencing seems to be required in *Drosophila* as well. It has been recently found that mutations in the genes *piwi, aubergine* and *spindle-E* (*homeless*), which code for RNAi components, result in a reduction of H3 Lys9 methylation and delocalization of HP1 and HP2.[106] Thus, the action of genes involved in RNAi could drive the methyltransferase activity of SU(VAR)3-9 to the chromatin and subsequently permit the binding of HP1 to methylated histone H3 tails.[106]

Conclusions

A significant challenge for current research on eukaryotic genomes is to understand heterochromatin sequence, structure and function in model organisms and in humans. Studies on the molecular organization and function of this genomic component have been greatly facilitated by the annotation of the heterochromatin sequence in *D. melanogaster* and other organisms.[12,51] However, we still do not have a complete understanding of the relationship between heterochromatin sequence and function. Most of the essential heterochromatic genes defined by genetic analysis in *D. melanogaster* are still unknown at the molecular level. In addition, detailed genetic and molecular analysis of the functions of gene models in *Drosophila* and other organisms is made difficult by the lack of mutant alleles. There is thus a need for additional studies where both classical genetic (insertional mutagenesis) and reverse genetic approaches (targeted gene replacement and double-stranded-RNA interference) will be required to explore the function of new genes and to validate the functions of those already inferred.

It is now clear that heterochromatin proteins have structural roles in heterochromatin formation, telomere capping and centromere condensation, and can control gene expression. Depending on which chromosomal region they bind, or which partner they have, these proteins may act to repress or activate genes. Thus it is difficult to draw a clear boundary between heterochromatin and euchromatin. However, it must be recalled that genes resident in constitutive heterochromatin are correctly expressed only in their native environment and cannot be simply considered as euchromatin-like functional sequences. At the same time, euchromatic genes that become silent are not simply assimilated to constitutive heterochromatin, but are more similar to X chromosome genes that undergo facultative heterochromatinization in mammalian females. It would be more correct to consider euchromatin and constitutive heterochromatin as two different chromatin compartments, in both of which gene expression can occur, possibly depending on the formation of differential multiprotein complexes. Such multiprotein complexes may share components, but they are not necessarily identical. It is tempting to speculate that TE DNAs, which are abundant in heterochromatic genes, or TE transcripts, may have a key role in the establishment of functional domains of gene expression within constitutive heterochromatin, perhaps by recruiting specific heterochromatin protein complexes.

Acknowledgments

We are grateful to Richard D'Ari for helpful comments on the manuscript. We also wish to thank Steven Marygold for allowing us to quote from his unpublished work.

References

1. Heitz E. 1928. Das Heterochromatin der Moose. Jb Wiss Bot 69:762–818.
2. Brown SW. 1966. Heterochromatin. Science 151:417–425.
3. John B. 1988. The biology of heterochromatin. In: V. R S., editor. Heterochromatin: Molecular and Structural aspects. Cambridge University Press. p. 1–128.
4. Gatti M, Pimpinelli S. 1992. Functional elements in *Drosophila melanogaster* heterochromatin. Annu Rev Genet 26:239–275.
5. Weiler KS, Wakimoto BT. 1995. Heterochromatin and gene expression in *Drosophila*. Ann Rev Genet 29:577–605.
6. Elgin SCR. 1996. Heterochromatin and gene regulation in *Drosophila*. Current Opinion in Genet Dev 6:193–200.
7. Dernburg AF, Sedat JW, Hawley RS. 1996. Direct evidence of a role for heterochromatin in meiotic chromosome segregation. Cell 86:135–146.
8. Karpen GH, Le MG, Le H. 1996. Centric heterochromatin and the efficiency of achiasmate disjunction in Drosophila female meiosis. Science 273:118–122.
9. Eissenberg JC, Hilliker AJ. 2000. Versatility of convinction: heterochromatin as both repressor and an activator of transcription. Genetica 109:19–24.
10. Henikoff S, Ahmad K, Malik HS. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293:1098–1102.
11. Zhimulev IF. 1998. Polytene chromosomes, heterochromatin, and position effect variegation. Adv Genet 37:1–566.
12. Hoskins RA, Smith CD, Carlson JW, Carvalho AB, Halpern A, et al. 2002. Heterochromatic sequences in a *Drosophila* whole-genome shotgun assembly. Genome Biol 3:research0085.1–0085.16.
13. Corradini N, Rossi F, Vernì F, Dimitri P. 2003. FISH analysis of *Drosophila* heterochromatin using BACs and P-elements. Chromosoma 112:26–37.
14. Kuln RM, Clarke L, Carbon J. 1991. Clustered tRNA genes in *Schizosaccharomyces Pombe* centromeric DNA sequence repeats. PNAS 88:1306–1310.
15. Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815.
16. Horvath JE, Schwartz S, Eichler EE. 2000. The mosaic structure of human pericentromeric DNA: a strategy for characteryzing complex regions of the human genome. Genome Res 10:839–852.
17. 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.
18. Horvath JE, Bailey JA, Locke DP, Eichler EE. 2001. Lessons from the human genome: transitions between euchromatin and heterochromatin. Hum Mol Gen 10:2215–2223.
19. Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, et al. 2004. Sequencing of a rice centromere uncovers active genes. Nat Genet 36:138–145.
20. 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.
21. Plath K, Fang J, Mlynarczyk-Evans SK, Cao R, Worringer KA, et al. 2003. Role of histone H3 lysine 27 methylation in X inactivation. Science 300:131–135.
22. Moritz KB, Roth GE. 1976. Complexity of germline and somatic DNA in Ascaris. Nature 259:55–57.
23. Lindsley DL, Sandler L, Baker BS, Carpenter AT, Denell RE, et al. 1972. Segmental aneuploidy and the genetic gross structure of the *Drosophila*. Genome Genetics 71:157–184.
24. Sandler L, Hecht F. 1973. Annotation: genetic effects of aneuploidy. Am J Hum Genet 25:332–339.
25. Pimpinelli S, Goday C. 1989. Unusual kinetochores and chromatin diminution in *Parascaris*. Trends Genet 5:310–315.
26. 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.
27. Pimpinelli S, Sullivan W, Prout M, Sandler L. 1985. On biological functions mapping to the heterochromatin of *Drosophila melanogaster*. Genetics 109:701–724.
28. Csink AK, Henikoff S. 1998. Something from nothing: the evolution and utility of satellite repeats. Trends Genet 14:200–204.
29. Dimitri P, Junakovic N. 1999. Revising the selfish DNA hypothesis: new evidence on accumulation of transposable elements in heterochromatin. Trends Genet 15:123–124.
30. Kaminker JS, Bergman CM, Kronmiller B, Carlson J, Svirskas R, et al. 2002. The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective. Genome Biol 3: RESEARCH0084.
31. Blumenstiel JP, Hartl DL, Lozovsky ER. 2002. Patterns of insertion and deletion in contrasting chromatin domains. Mol Biol Evol 19:2211–2225.
32. Rizzon C, Marais G, Gouy M, Biemont C. 2002. Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster*. Genome Genome Res 12:400–407.
33. 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.
34. Devlin RH, Bingham B, Wakimoto BT. 1990. The organization and expression of the light gene, a heterochromatic gene of *Drosophila melanogaster*. Genetics 125:129–140.
35. Devlin RH, Holm DG, Morin KR, Honda BM. 1990. Identifying single-copy DNA sequence associated with the expression of a heterochromatic gene, the *light* locus of *Drosophila melanogaster*. Genome 33:405–415.
36. 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.
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.
38. Dimitri P. 1991. Cytogenetic analysis of the second chromosome heterochromatin of *Drosophila melanogaster*. Genetics 127:553–564.
39. Koryakov DE, Zhimulev IF, Dimitri P. 2002. Cytogenetic analysis of the third chromosome heterochromatin of *Drosophila melanogaster*. Genetics 160:509–517.
40. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, et al. 2000. The genome sequence of *Drosophila melanogaster*. Science 287: 2185–2195.
41. 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.
42. Myster SH, Wang F, Cavallo R, Christian W, Bhotika S, 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.
43. Dimitri P. 1997. Constitutive heterochromatin and transposable elements in *Drosophila melanogaster*. Genetica 100:85–93.
44. Warner TS, Sinclair DA, Fitzpatrick M, Singh M, Devlin RH, et al. 1998. The light gene of *Drosophila melanogaster* encodes a homologue of VPS41, a yeast gene involved in cellular protein traffic. Genome 41: 236–243.
45. Parks S, Wieschaus E. 1991. The *Drosophila* gastrulation gene concertina encodes a Ga-like protein. Cell 64:447–458.
46. Biggs HW, Zavitz HK, Dikinson B, Van Der Straten A, Brunner D, et al. 1994. The *Drosophila* rolled locus encodes a MAP kinase required in the sevenless signal transduction pathway. EMBO J 13:1628–1635.
47. Inoue YH, Glover DM. 1998. Involvement of the rolled/MAP kinase gene in *Drosophila* mitosis: interaction between genes for the MAP kinase cascade and abnormal spindle. Mol General Genetics 258:334–341.

48. Cenci G, Belloni G, Dimitri P. 2003. *l(2)41Aa*, a heterochromatic gene of *Drosophila melanogaster*, is required for mitotic and meiotic chromosome condensation. Genet Res 81:15–24.
49. Schulze S, Sinclair DA, Silva E, Fitzpatrick KA, Singh M, et al. 2001. Essential genes in proximal 3L heterochromatin of *Drosophila melanogaster*. Mol Gen Genet 264:782–789.
50. 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.
51. Misra S, Crosby MA, Mungall CJ, Matthews BB, Campbell KS, et al. 2002. Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. Genome Biology 3:research0083.1–0083.22.
52. Yasuhara JC, De Crease C, Slade D, Wakimoto BT. 2002. A comparative study of heterochromatic and euchromatic *light* genes in *Drosophila* species. 43rd Annual Drosophila research conference, San Diego CA.
53. Pickeral OK, Makalowski W, Boguski MS, Boeke JD. 2000. Frequent Human Genomic DNA Transduction Driven by LINE-1 Retrotransposition. Genome Res 10:411–415.
54. Krantz ID, McCallum J, De Scipio C, Kaur M, Gillis LA, 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.
55. 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.
56. Copenhaver GP, Nickel K, Kuromori T, Benito M-I, Kaul S, et al. 1999. Genetic definition and sequencing analysis of the *Arabidopsis* centromeres. Science 286:2468–2474.
57. Lin X, Kaul S, Rounsley S, Shea TP, Benito M-I, et al. 1999. Sequence analysis of the chromosome 2 of the plant *Arabidopsis thaliana*. Nature 402:761–768.
58. The European Union Arabidopsis Genome Sequencing Consortium & The Cold Spring Harbor, Washington University in St Louis and PE Biosystems Arabidopsis Sequencing Consortium. 1999. Sequence and analysis of the chromosome 4 of the plant *Arabidopsis thaliana*. Nature 402:769–777.
59. European Union Chromosome 3 Arabidopsis Sequencing Consortium, The Institute for Genomic Research & Kazusa DNA Research Institute. 2000. Sequence and analysis of the chromosome 3 of the plant *Arabidopsis thaliana*. Nature 408:820–822.
60. Theologis A, Ecker JR, Palm CJ, Federspiel NA, Kaul S, et al. 2000. Sequence and analysis of the chromosome 1 of the plant *Arabidopsis thaliana*. Nature 408:816–820.
61. 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.
62. 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.
63. 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.
64. 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.
65. Sinclair DA, Schulze S, Silva E, Fitzpatrick KA, Honda BM. 2000. Essential genes in autosomal heterochromatin of *Drosophila melanogaster*. Genetica 109:9–18.
66. Carvalho AB, Dobo BA, Vibranovski MD, Clark AG. 2001. Identification of five new genes on the Y chromosome of *Drosophila melanogaster*. Proc Natl Acad Sci USA 98:13225–13230.
67. Dimitri P, Junakovic N, Arcà B. 2003. Colonization of heterochromatic genes by transposable elements in *Drosophila*. Mol Biol Evol 20:503–512.
68. McCollum AM, Ganko EW, Barrass PA, Rodriguez JM, McDonald JF. 2002. Evidence for the adaptive significance of an LTR retrotransposon sequence in a *Drosophila* heterochromatic gene. BMC Evol Biol 2:5.
69. 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.
70. James TC, Elgin SCR. 1986. Identification of a nonhistone chromosomal protein associated with heterochromatin in *Drosophila melanogaster* and its gene. Mol Cell Biol 6:3862–3872.
71. James TC, Eissenberg JC, Craig C, Dictrich V, Hobson A, et al. 1989. Distribution pattern of HP1, a heterochromatin-associated nonhistone chromosomal protein of *Drosophila*. Eur J Cell Biol 50:170–180.
72. Eissenberg JC, James TC, Foster-Hartnett DM, Hartnett T, Ngan V, et al. 1990. Mutation in heterochromatin-specific chromosomal protein is associated with suppression of position-effect variegation in *Drosophila melanogaster*. Proc Natl Acad Sci USA 87:9923–9927.
73. Paro R, Hogness DS. 1991. The polycomb protein shares a homologous domain with a heterochromatin-associated protein of *Drosophila melanogaster*. Proc Natl Acad Sci USA 88:263–267.
74. Aasland R, Stewart AF. 1995. The chromo shadow domain, a second chromo domain in heterochromatin-binding protein 1, HP1. Nucleic Acids Res 23:3168–3173.
75. Jones DO, Cowell IG, Sing PB. 2000. Mammalian chromo domain proteins: their role in genome organisation and expression. Bioessays 22:124–137.
76. Wang G, Ma A, Chow CM, Horsley D, Brown NR, et al. 2000. Conservation of heterochromatin protein 1 function. Mol Cell Biol 20:6970–6983.
77. Fanti L, Dorer DR, Berloco M, Henikoff S, Pimpinelli S. 1998a. Heterochromatin protein 1 binds transgene arrays. Chromosoma 107: 286–292.
78. Fanti L, Giovinazzo G, Berloco M, Pimpinelli S. 1998b. The heterochromatin protein 1 prevents telomere fusions in *Drosophila*. Mol Cell 2:527–538.
79. Hwang KK, Eissenberg JC, Worman HJ. 2001. Transcriptional repression of euchromatic genes by *Drosophila* heterochromatin 1 and histone modifiers. Proc Natl Acad Sci USA 98:11423–11427.
80. Nielsen SJ, Schneider R, Bauer UM, Bannister AJ, Morrison A, et al. 2001. Rb targets histone H3 methylation and HP1 to promoters. Nature 412:561–565.
81. Fanti L, Berloco M, Piacentini L, Pimpinelli S. 2003. Chromosomal distribution of heterochromatin 1 (HP1) in *Drosophila*: a cytological map of euchromatic HP1 binding sites. Genetica 117:135–147.
82. Piacentini L, Fanti L, Berloco M, Perrini B, Pimpinelli S. 2003. Heterochromatin protein 1 (HP1) is associated with induced gene expression in *Drosophila* euchromatin. J Cell Biol 161:707–714.
83. Clegg NJ, Honda BM, Whitehead IP, Grigliatti TA, Wakimoto B, 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.
84. 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.
85. Aagaard L, Laible G, Selenko P, Schmid M, Dorn R, et al. 1999. Functional mammalian homologues of the *Drosophila* PEV-modifier *Su(var)3-9* encode centromere-associated proteins which complex with the heterochromatin component M31. EMBO J 18:1923–1938.
86. Schotta G, Ebert A, Krauss V, Fischer A, Hoffman JE, 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.
87. Greil F, van der Kraa I, Delrow J, Smothers JF, de Wit E, 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.
88. Brand AH, Perrimon N. 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118:401–415.
89. Chien CT, Buck S, Sternglanz R, Shore D. 1993. Targeting of SIR1 protein establishes transcriptional silencing at HM loci and telomeres in yeast. Cell 75:531–541.
90. Shareef MM, Badugu R, Kellum R. 2003. HP1/ORC complex and heterochromatin assembly. Genetica 117:127–134.

91. Pak DTS, Pflumm M, Chesnokov I, Huang DW, Kellum R, et al. 1997. The association of *Drosophila* origin recognition complex with Heterochromatin and HP1. Cell 91:311–323.
92. Reuter G, Giarre M, Farah J, Spierer A, Spierer P. 1990. Dependence of position-effect variegation in *Drosophila* on dose of a gene encoding an unusual zinc-finger protein. Nature 344:219–223.
93. Cléard F, Matsarskaia M, Spierer P. 1995. The modifier of position-effect variegation *Suvar(3)7* of *Drosophila*: there are two alternative transcripts and seven scattered zinc fingers, each preceded by a tryptophan box. Nucleic Acids Res 23:796–802.
94. Delattre M, Spierer A, Tonka C-H, Spierer P. 2000. The genomic silencing of position effect variegation in *Drosophila melanogaster*: interaction between the heterochromatin-associated proteins Su(var)3-7 and HP1. J Cell Sci 113:4253–4261.
95. Cléard F, Delattre M, Spierer P. 1997. *SU(VAR)3-7, a Drosophila* heterochromatin-associated protein and companion of HP1 in the genomic silencing of position-effect variegation. EMBO J 16:5280–5288.
96. Shaffer CD, Stephens GE, Thompson BA, Funches L, Bernat JA, et al. 2002. Heterochromatin protein 2 (HP2), a partner of HP1 in *Drosophila* heterochromatin. Genetics 99:14332–14337.
97. Shareef MM, King C, Damaj M, Badagu R, Huang DW, et al. 2001. *Drosophila* heterochromatin protein 1 (HP1)/origin recognition complex (ORC) protein is associated with HP1 and ORC and functions in heterochromatin-induced silencing. Mol Biol Cell 12:1671–1685.
98. Cenci G, Siriaco G, Raffa GD, Kellum R, Gatti M. 2003. The *Drosophila* HOAP protein is required for telomere capping. Nat Cell Biol 5:82–84.
99. Badugu R, Shareef MM, Kellum R. 2003. Novel *Drosophila* Heterochromatin protein 1 (HP1)/origin recognition complex-associated protein (HOAP) repeat motif in HP1/HOAP interactions and chromocenter associations. J Biol Chem 278:34491–34498.
100. Volpe TA, Kidner C, Hall IM, Teng G, Grewal SIS, et al. 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297:1833–1837.
101. Hall IM, Shankaranarayana GD, Noma K, Ayoub N, Cohen A, et al. 2002. Establishment and maintenance of a heterochromatin domain Science 297:2232–2237.
102. Schramke V, Allshire R. 2003. Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin based gene silencing. Science 301:1069–1074.
103. Reinhart BJ, Bartel DP. 2002. Small RNAs correspond to centromere heterochromatic repeats. Science 297:1831.
104. Lehnertz B, Ueda Y, Derijck AHA, Braunschweig U, Perez-Burgos L, 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.
105. Fukagawa T, Nogami M, Yoshikawa M, Ikeno M, Okazaki T, et al. 2004. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol 6:784–791.
106. Pal-Bhadra M, Leibovitch BA, Gandhi SG, Rao M, Bhadra U, et al. 2004. Heterochromatic silencing and HP1 localization in *Drosophila* are dependent on the RNAi machinery. Science 303:669–672.