Chapter 8
Transcription of Satellite DNAs in Insects

Željka Pezer, Josip Brajković, Isidoro Feliciello, and Đurđica Ugarković

Abstract The very complex life cycle and extreme diversity of insect life forms require a carefully regulated network of biological processes to switch on and off the right genes at the right time. Chromatin condensation is an important regulatory mechanism of gene silencing as well as gene activation for the hundreds of functional protein genes harbored in heterochromatic regions of different insect species. Being the major heterochromatin constituents, satellite DNAs (satDNAs) serve important roles in heterochromatin regulation in insects in general. Their expression occurs in all developmental stages, being the highest during embryogenesis. satDNA transcripts range from small RNAs, corresponding in size to siRNAs, and piwiRNAs, to large, a few kb long RNAs. The long transcripts are preferentially nonpolyadenylated and remain in the nucleus. The actively regulated expression of satDNAs by *cis* or *trans* elements as well as by environmental stress, rather than constitutive transcription, speaks in favor of their involvement in differentiation, development, and environmental response.

8.1 Satellite DNAs in Insects

Satellite DNAs (satDNAs) are the major DNA component of eukaryotic heterochromatin. These noncoding sequences constitute a considerable part of the genomic DNA in many insect species, which can reach over half of the genomic content

Ž. Pezer, J. Brajković, and Đ. Ugarković (✉)
Ruđer Bošković Institute, Bijenička 54, 10001 Zagreb, P.O. Box 1016, Croatia
e-mail: ugarkov@irb.hr

I. Feliciello
Ruđer Bošković Institute, Bijenička 54, 10001 Zagreb, P.O. Box 1016, Croatia
and
Dipartimento di Medicina Clinica e Sperimentale, Università degli Studi di Napoli Federico II, via Pansini 5, 80131 Napoli, Italy

Đ. Ugarković (ed.), *Long Non-Coding RNAs*, Progress in Molecular and Subcellular Biology 51, DOI 10.1007/978-3-642-16502-3_8,

(Davis and Wyatt 1989). Composed of tandemly reiterated arrays, usually millions of base pairs long, satDNA is located mainly in the pericentromeric and telomeric regions of chromosomes and are the major building elements of functional centromeres in many eukaryotes including insects (Ugarković 2009a).

The complexity of satDNA sequences varies between insect species. Some are very simple, composed of short repeats oriented in a head-to-tail fashion, such as ten *Drosophila melanogaster* satellites that are between 5 and 10 bp long (Lohe and Roberts 1988). The nucleotide sequences of ten simple repeats in *D. melanogaster* conforms to a formula $(AAN)_m(AN)_n$ where N is any nucleotide. In addition to simple sequence monomers, *D. melanogaster* and other related *Drosophila* species share also a complex satellite of 359 bp, known as 1.688 satellite according to CsCl buoyant density. All *D. melanogaster* satellites are AT rich and, together, they make approximately 20% of the genome. Satellite monomers in insect can also exceed 1,000 bp, as revealed in some beetle species (Pons 2004). However, most of insect satDNAs fall into two size classes: one ranging from 140 to 190 bp and the other in the range of 300–400 bp (reviewed in Palomeque and Lorite (2008)). The total length of the satellite tandem arrays varies from less than 100 bp to over 100 Mb between species. Some satellites are restricted to particular chromosomes such as *Drosophila* 359 bp satellite that is located on X chromosome, while most of the insect satellites are spread on all chromosomes of particular species. Satellites present on the same chromosome can be organized into separate arrays or can exhibit interspersed type of organization characteristic for beetle species *Tribolium madens* (Durajlija Žinić et al. 2000). In many insect species, satDNAs encompass both centromeric and pericentromeric regions. However, satellite sequences in centromeric and pericentromeric regions are almost indistinguishable, even in the best studied *Drosophila* centromere (Sun et al. 1997). Moreover, due to the technical difficulties associated with the sequencing and assembling of highly repetitive regions, the detailed structure and organization of heterochromatic and centromeric regions for many insect species is not known.

satDNAs sometimes possess an additional level of organization, called higher order repeats (HORs), which are best described for α satellite of primates and are characteristic for human centromeres (Schueler et al. 2001). For example, the predominant form of α satellite on human chromosome 17 is a 2.7 kbp HOR unit consisting of 16 α satellite monomers (Waye and Willard 1986). In insects, however, HORs are not found very often and are usually in the form of complex dimers and trimers (Palomeque and Lorite 2008).

It has been demonstrated that related insect species share a set of common satDNAs, differentially amplified between species, so that changes in the number of copies produce a species-specific profile of satDNAs (Meštrović et al. 1998; Bruvo-Mađarić et al. 2007). According to the "library" hypothesis (Fry and Salser 1977), some of the satDNAs originally contained in the common ancestor could be amplified during speciation. In each of the descendants, usually one satDNA would exist as a major satellite sequence while the others would remain as low-copy-number satellites. It is proposed that satellite sequences residing within a library exhibit certain structural characteristics that enable them to confer a centromeric

role and therefore could represent a source for the evolution of new centromere (Ugarković 2008, 2009b).

Satellite sequences are known to evolve fast, changing their nucleotide sequence and copy number by following a pattern known as concerted evolution. Changes among repeats create new variants, which are spread and homogenized within the genome by a variety of DNA turnover mechanisms, such as unequal crossing over, gene conversion, replication slippage, and rolling circle replication (Charlesworth et al. 1994). Variants are subsequently fixed within the reproductive group of organisms in a stochastic process known as molecular drive (Dover 2002). Taken together, recombinational mechanisms and molecular drive seem to be the major cause of high turnover of satDNA sequences, resulting in the significant sequence divergence of satDNAs and changes in copy number, even between closely related species (Ugarković and Plohl 2002).

8.2 Functional Elements Within satDNAs

Due to the homogenizing effects of concerted evolution, satDNAs usually display low internal sequence variability. However, comparison of monomer sequences of a given satDNA reveals that some monomer regions are more conserved while others show higher mutation rates (Borstnik et al. 1994; Romanova et al. 1996; Mravinac et al. 2004). Such a nonuniform rate of evolution along the sequence indicates the presence of selective pressure on satDNAs. It can be proposed that selection affects pericentromeric satellite repeats whose transcripts play a role in heterochromatin establishment through RNA interference mechanism (Volpe et al. 2002). In fission yeast, *Schizosaccharomyces pombe*, analysis of siRNAs involved in heterochromatin formation showed that they derive preferentially from the most conserved regions of repeats (Djupedal et al. 2009). This indicates that conservation is more probably due to functional constraints than to frequent events of homologous recombination causing sequence identity. Therefore, conserved regions found in other satDNAs could be functional in the sense that they represent a preferential source of siRNAs that recruit protein complexes responsible to heterochromatin formation.

Due to possible functional constraints on satDNAs, it is not surprising that some characteristics of satDNAs are shared between many eukaryotic organisms (Ugarković 2005). Probably the most common feature of satDNA is its intrinsic curvature. Satellite repeats are generally AT rich, and the periodical distribution of AT tracts causes DNA to bend into a super-helical tertiary structure (Fitzgerald et al. 1994). This sequence-dependent property is thought to be responsible for the tight packing of DNA and proteins in heterochromatin (Ugarković et al. 1992; Fitzgerald et al. 1994). Conserved CENP-B box-like motifs have been identified within satDNA of mammals and insects (Kipling and Warburton 1997; Lorite et al. 2002a; Mravinac et al. 2004). The CENP-B box is a 17 bp motif in human α-satDNA and a binding site for centromere protein B (CENP-B) (Masumoto et al. 1989) whose homologs have been found in many eukaryotes. Not every repeat of

α satellite contains a functional CENP-B box, but they appear at regular intervals in human centromeres and seem to be essential for centromeric chromatin assembly (Ohzeki et al. 2002).

Given their relatively simple sequence and the lack of any significant open-reading frame, previously reported transcription of satDNA has been ascribed to read-through from upstream genes and transposable elements (Diaz et al. 1981; Wu et al. 1986; Gaubatz and Cutler 1990). However, promoter elements and transcription start sites as well as binding motifs for transcription factors have been mapped within some satellites. Putative internal promoters have been reported in the wasp *Diadromus pulchellus* (Renault et al. 1999) where motifs cognate to RNA Pol II and III are present within the satellite monomer sequence. In schistosome satDNA, which encodes an active ribozyme, a functional RNA Pol III promoter is present (Ferbeyre et al. 1998). The sequence of highly conserved satellite 2 found in distant families of salamanders shares structural and functional properties with the typical vertebrate small nuclear RNA (snRNA) promoter (Coats et al. 1994).

The *Drosophila* GAGA transcription factor that binds GA/CT-rich elements in promoters of many *Drosophila* genes and activates transcription by opening chromatin structure was found associated with heterochromatin throughout the cell cycle. It is proposed that GAGA factor directly interacts with a GA/CT-rich subset of satDNA repeats and modifies heterochromatin structure (Raff et al. 1994). Human satellite III has a binding motif for the heat-shock transcription factor 1, which drives its RNA Pol II-dependent transcription in stress conditions (Metz et al. 2004; Chap. 5 in this book). γ-satDNA, the abundant pericentromeric sequence of all murine chromosomes, contains conserved binding sites for ubiquitous transcription factor Yin Yang 1 (YY1) (Shestakova et al. 2004). YY1 belongs to the Polycomb group of proteins involved in gene regulation during development. It has been found to be associated with γ-satDNA in proliferating cells, but the association strongly diminishes during transition to the quiescent state (G_0). It has been proposed that the interaction of YY1 with γ-satDNA could lead to the targeting of proteins required for heterochromatization or to the silencing of euchromatic genes by bringing them in close proximity to pericentromeric heterochromatin.

Even though some time ago satDNAs were considered useless evolutionary remnants, the functional significance of their sequences is becoming ever more clear. The existence of conserved motifs and structural properties as well as emerging evidence on their widespread transcriptional activity prompt us to reexamine these highly abundant eukaryotic sequences.

8.3 Preserved satDNA Repeats in Insects

Beetles of the family Tenebrionidae (Coleoptera) represent convenient models for studying satDNAs. Located in the (peri)centromeric heterochromatin of virtually all chromosomes, satDNAs comprise up to half of the genome size in most of the

coleopteran species analyzed so far (Ugarković and Plohl 2002). They are characterized by low intraspecific variability of the basic repeating units. However, their sequences differ dramatically between species, with some exceptions such as satDNAs from the genus *Pimelia* (Pons et al. 1997).

Members of the genus *Palorus* share a "library" of satDNAs – a collection of common satellite sequences, differentially amplified in each species. Besides the species-specific major satellite, each species contains several low-copy-number satellites, which account for approximately 0.05% of the genome and are dispersed over the whole heterochromatic block, interrupting large arrays of the major satellite. Shared satellites are extremely conserved with respect to sequence, monomer length, and tandem repeat organization (Meštrović et al. 1998). Satellites from some *Palorus* species are widely distributed even beyond the level of the genus, such as PRAT and PSUB, major satellites that comprise 40 and 20% of the respective genome size of *P. ratzeburgii* and *P. subdepressus* (Ugarković et al. 1992; Plohl et al. 1998). These two satellites share no sequence homology but have similar repeat lengths of 142 bp (PRAT) and 144 bp (PSUB). In both sequences, a motif homologous to the human CENP-B box has been identified (Mravinac et al. 2005). PRAT and PSUB are found in a low number of copies in species belonging to the related genera *Tribolium*, *Tenebrio, Pimelia*, and *Latheticus*, subfamily Pimelinae, and family Chrysomelidae (Mravinac et al. 2002, 2005). Despite the fact that some of these species diverged from the *Palorus* group up to 60 Myr ago, PRAT and PSUB sequences remained virtually unchanged, showing no species diagnostic mutations and retaining the nonrandom pattern of variability along the sequence. It is assumed that the extreme preservation of these two satellites over a long period is related to their functional significance (Mravinac et al. 2005; Ugarković 2005).

Ancient satDNAs up to 80 Myr old have also been reported in some species of fish (De la Herrán et al. 2001) and whales (Arnason et al. 1984), and their remarkable preservation is thought to be related to the low mutation rates generally observed in aquatic environments (De la Herrán et al. 2001). However, functional constraints have been implicated in the preservation of salamander satellite 2 sequence – its promoter activity and the self-cleavage ability of its transcripts have remained conserved for 200 Myr (Green et al. 1993).

8.4 Transcription of Insects satDNAs

Insects represent one of the most diverse groups of animals, accounting for more than half of all known living organisms. Despite this, they are poorly represented as model organisms for the study of satDNA function. With the present study, satDNA transcription has been described in species from only four orders: Hymenoptera, Orthoptera, Diptera, and Coleoptera (Table 8.1).

Table 8.1 Transcription of satDNAs in insects

Species	satDNA		satDNA transcription			
	satDNA/repeat unit	Amount in the genome	Size of transcripts	Strand specific	Other features	Reference
Orthoptera						
Dolichopoda baccettii	pDo500 420-510-bp	5%	Monomeric and multimeric	N/A	Ribozyme, self-cleavage activity	Rojas et al. (2000)
Hymenoptera						
Symphyta. Diprionidae						
Diprion pini	Ps 311-bp	1.8–4%	Heterogeneous	Differential expression	Preferential transcription in females	Rouleux-Bonnin et al. (1996)
	Pv 280-bp	0.1–1.3%	Heterogeneous	N/A		Rouleux-Bonnin et al. (1996)
	Ec 148-bp	1.6–7%	Heterogeneous	N/A		Rouleux-Bonnin et al. (1996)
Apocrita. Apidae						
Bombus terrestris	BT 422-bp	2.4–3.2%	Discrete sizes (2.1 kb, 1.55 kb, 1.05 kb, 0.6 kb)	Strand specific in embryos Both strands in images		Rouleux-Bonnin et al. (2004)
Apocrita. Eupelmidae						
Eupelmus orientalis	107-bp	7%	Heterogeneous	N/A	Differential expression between sexes	Renault et al. (1999)
Eupelmus vuilleti	107-bp	25%	Heterogeneous	N/A	Sex-specific	Renault et al. (1999)
Apocrita. Formicidae						
Aphaenogaster subterranea	APSU 162-bp	4–16%	Long	Both strands		Lorite et al. (2002a, b)
Apocrita. Ichneumonidae						
Diadromus collaris	512-bp	5%	Discrete sizes	N/A	Differential expression between sexes	Renault et al. (1999)
Diadromus pulchellus	320-bp	15%	Discrete sizes (1.9 kb, 1.5 kb, 1.2 kb, 0.62 kb, 0.32 kb)	Both strands	Nonpolyadenylated, differential expression between stages Putative promoter elements	Renault et al. (1999)

Diptera						
Drosophila hydei	YSI 600-bp, YLI 693-bp, YLII 77-bp, YLIII 4-bp	3–9% of Y chromosomal DNA	Heterogeneous	Strand specific	Nuclei of primary spermatocytes	Trapitz et al. (1988)
Drosophila melanogaster	1.688 g/cm^3 (four subfamilies: 254, 353, 356 and 359-bp)	4%	siRNA (19–28 nt)	Both strands	In ovaries and testes	Aravin et al. (2003), Usakin et al. (2007)
	AAGAC repeats from satDNA 1.686 g/cm^3	2.4%	Heterogeneous	N/A	On Y chromosome loops of primary spermatocytes	Bonaccorsi et al. (1990)
Coleoptera						
Family Tenebrionidae						
Tribolium castaneum	TCAST 360-bp	17%	Long: ~0.5–5 kb; small: 21–28 nt	Threefold difference	Differential expression between stages partially polyadenylated	Pezer et al. unpublished
Palorus genalis	PGEN 429-bp	30%	Heterogeneous (~0.5–5 kb)	N/A		Pezer and Ugarković unpublished
Palorus ratzeburgii	PRAT 142-bp	40%	Heterogeneous (~0.5–5 kb)	Tenfold difference	Located in nuclei and cytoplasm Partially polyadenylated Putative promoter elements	Pezer and Ugarković (2008)
Palorus subdepressus	PSUB 144-bp	20%	Heterogeneous (~0.5–5 kb)	Both strands	Located in nuclei and cytoplasm Partially polyadenylated Putative promoter elements	Pezer and Ugarković (2009)

8.4.1 Hymenoptera

Transcription of satDNAs seems to be a general phenomenon in Hymenopteran suborder Apocrita, which includes ants, wasps, and bees (Renault et al. 1999; Lorite et al. 2002b). Satellite expression rates vary significantly between sexes, and female sex-specific transcripts are detected. Also, the amount of satellite transcripts differs among the queen, the worker, and the male. The presence of specific transcription factors might influence different rates of satellite expression in males and females, which suggests that transcription is regulated in *trans* rather than by elements inside the satellite sequence itself. The wasp *Diprion pini* has a higher satDNA expression in females, despite lower satDNA sequence content compared to the male genome, indicating that transcription is not constitutive (Rouleux-Bonnin et al. 1996). Gender-specific satDNA transcription in hymenopteran species could be related to sexual differentiation at the chromatin level and it is proposed that long multimeric transcripts probably have a structural role (Rouleux-Bonnin et al. 2004).

In addition to gender-specific expression, satDNAs seem to be differentially transcribed during Hymenopteran insect development, displaying higher expression in earlier stages. In the wasp *Diadromus pulchellus,* single-stranded, nonpolyadenylated transcripts heterogeneous in size but with discrete bands of 1.9 kb and 0.6 kb were detected. satDNA seems to be differentially expressed during developments since the transcripts were found to be more abundant in embryos and larvae than adults (Renault et al. 1999). In addition, putative promoters and transcription initiation site were mapped within the monomer sequence. In the bumble bee *Bombus terrestris*, multimeric transcripts arise from one strand preferentially in embryos opposed to both strands in imagos (Rouleux-Bonnin et al. 2004). It seems therefore that satDNA transcription interferes with development as well as sexual and caste-differentiation in this insect order.

Stage-specific transcription associated with differentiation has also been observed in systems other than insects. The most abundant mouse γ satDNA is differentially expressed in cells of the developing central nervous system as well as in adult liver and testis (Rudert et al. 1995; Chap. 5 in this book). In chicken and zebrafish, transcription of alphoid repeat sequences displays a specific temporal and spatial expression pattern during embryogenesis (Li and Kirby 2003).

8.4.2 Diptera

In Diptera, transcription of satDNA has been reported in primary spermatocytes of *D. melanogaster* and *D. hydei*. (Trapitz et al. 1988; Bonaccorsi et al. 1990). Transcripts of simple satellite sequence AAGAC that are highly heterogeneous in size, ranging from less than 1 kb to 10s of kb, have been found on Y chromosome loops. They do not appear to migrate to cytoplasm and are degraded during the first meiotic prophase. It has been proposed that the transcripts probably act as fertility

factors by providing a structural framework for accumulation of Y-encoded proteins involved in sperm differentiation.

In addition to the transcripts of simple sequence satellite AAGAC, transcripts of 1.688 *D. melanogaster* satellite with complex repeat unit of 359 bp were found in germinal tissues. Transcription proceeds from both DNA strands and is under the control of RNA interference machinery (Usakin et al. 2007). It is revealed that RNAi is necessary to maintain the silenced state of centromeric and pericentromeric 1.688 repeats located specifically on X chromosome. The heterochromatic locus on X chromosome that contains 1.688 satellite is responsible for hybrid female lethality in crosses between *D. simulans* females and *D. melanogaster* males (Ferree and Barbash 2009). It induces mitotic failure in early embryos due to the inability of 1.688 (359 bp) satellite block to form a proper heterochromatin state. Both *Drosophila* species share common satellites that differ in amount and location between the species, but 1.688 (359 bp) satellite is 50 times more abundant in *D. melanogaster* relative to *D. simulans*. It is proposed that hybrid female lethality occurs owing to the absence of the 1.688 satellite-derived small RNAs in the maternal cytoplasm that are required for heterochromatin establishment on 1.688 satellite array.

Transcription of complex telomeric repeats that are characteristic for chromosomal ends in Diptera was demonstrated in *Chironomus thummi* (Martínez-Guitarte et al. 2008). Transcripts are heterogeneous in length and correspond to multimers of the repeat. Moreover, transcription of telomeric repeats is not constitutive and is activated under conditions of environmental stress, such as heat shock.

8.4.3 Orthoptera

satDNAs were studied in the cave cricket genus *Dolichopoda (*Rhaphidophoridae). Three different satellite families were characterized in this species, among which pDo500 satellite is present in all species of the genus. The 500-bp satDNA family is actively expressed in the form of long multimeric transcripts, although the monomeric transcripts are also detected (Rojas et al. 2000). The transcripts act as ribozymes, as they have the ability to adopt hammerhead-like secondary structures and self-cleave in vitro. It is also possible that hammerhead sequences from the pDo500 satellite can *trans*-cleave host transcripts in the cells of *Dolichopoda*. The physiological role of these ribozymes is unknown, but it can be proposed that they may affect certain regulatory mechanisms in the cell. High sequence conservation of their corresponding satellites and active transcription suggests that they are under selective pressure (Rojas et al. 2000).

The hammerhead ribozyme structures associated with transcribed satDNA sequences have also been found in salamanders (Epstein and Gall 1987) and schistostomes (Ferbeyre et al. 1998). All hammerhead ribozymes detected in animal satDNA so far have been shown to self-cleave in *cis* long multimeric satellite transcripts into monomers.

8.4.4 Coleoptera

Within order Coleoptera, expression of satDNAs was investigated in species of genera *Palorus* and *Tribolium* that belong to family Tenebrionidae. Species of the two genera are characterized by the presence of large blocks of pericentromeric heterochromatin on all chromosomes and molecular analyses suggested that these blocks are composed almost exclusively of satDNAs that comprise up to 40% of the whole genome and encompass the regions of functional centromere (Ugarković et al. 1996; Durajlija Žinić et al. 2000).

In beetles *Palorus ratzeburgii*, *Palorus subdepressus*, and *Tribolium castaneum*, the major satDNAs called PRAT, PSUB, and TCAST, respectively, are continuously expressed during larval, pupal, and imago stages. The transcripts are of heterogeneous size, ranging from 0.5 kb to more than 5 kb, and originate from both strands of satDNA, albeit with a difference in expression between the two strands (Pezer and Ugarković 2009). Most of the transcripts are detected in the nucleus and are not polyadenylated. Although transcription from both DNA strands could potentially activate the RNA interference (RNAi) pathway, no processing of long PRAT and PSUB transcripts into small interfering RNAs (siRNA) was detected. However, small RNAs cognate to the major satellite TCAST (Ugarković et al. 1996) have been detected in the red flour beetle *T. castaneum*, (unpublished results). Small RNAs are more abundant in embryos than in later developmental stages, ranging in size between 21 and 26 nt with a predominant size of 24 nt. According to their size, these RNAs could be assigned to small interfering RNAs (siRNAs 21–23 nt) and piwiRNAs (piRNAs, 24–26 nt). piRNAs are characterized as the long class of siRNAs that bind to the Piwi clade of Argonaute proteins (Hamilton et al. 2002; Aravin et al. 2003). It is proposed that both types of small RNAs function as guide molecules during heterochromatin formation. The piwiRNAs, also known as repeat-associated RNAs ranging 23–26 nt in size, are most abundant in testes and early embryos, which may be related to dramatic changes in heterochromatin structure that occur in these stages. In addition to siRNAs and piRNAs, some components of the RNAi machinery have been identified in the sequenced genome of *T. castaneum*, such as Dicer and Argonaute protein families but not the RNA-dependent RNA polymerase (RdRP) gene (Tomoyasu et al. 2008). RdRP transcribes single-stranded RNA from an RNA template and is important for the production of siRNA as well as the amplification of the RNAi effect in fungi, protists, nematodes, and plants. However, it seems to be lacking in insects and vertebrates.

Multiple transcription initiation and termination sites as well as putative RNA Pol II promoter elements were mapped within PRAT and PSUB sequences. Overlapping promoter-like sequences on both DNA strands and the close position of transcription initiation sites suggest bidirectional activity of putative promoters. Presence of 5′ cap structure on portion of PRAT transcripts and susceptibility of transcription to inhibition of Pol II further confirm role of RNA Pol II in transcription of satDNAs. In addition, motifs similar to A and B boxes, associated with RNA

Pol III transcription, are located in PSUB and PRAT satellites. Thus, Pol III or some other polymerase(s) might be responsible for the production of the main nonpolyadenylated fraction of transcripts. Involvement of different RNA polymerases in the production of siRNAs was demonstrated in plants. Noncoding transcripts generated by RNA Pol II in *Arabidopsis* act as a scaffold for the recruitment of two other polymerases, Pol IV and Pol V, which seem to be important for the production of siRNAs (Zheng et al. 2009). The interplay of all three polymerases is required for siRNA-mediated transcriptional gene silencing in *Arabidopsis*.

The heterogeneous transcript size of satDNAs in beetles could be explained by the multiple transcription initiation and termination sites. Read-through transcription from the nearby gene promoters and transposable elements cannot be excluded either, although there is a strong indication that satDNAs are transcribed as autonomous transcription units from own promoters that reside within the satellite sequences (Pezer and Ugarković 2008, 2009). However, a smaller portion of PRAT and PSUB satDNA transcripts is polyadenylated and is found in the cytoplasm. Cytoplasmic localization and the presence of a polyA tail have been reported before for the satellite transcripts of various species. Polyadenylated transcripts of the G + C-rich satDNA of the Bermuda land crab are present in the cytoplasm of different tissues (Varadaraj and Skinner 1994). Satellite 2 is an abundant tandemly repeated sequence distributed in clusters throughout the genome of the newt *Notophthalamus viridescens* and is transcribed on lampbrush chromosomes. However, stable, strand-specific transcripts homologous to satellite 2 are present in the cytoplasm in a variety of different tissues (Epstein et al. 1986). Satellite III DNA is transcribed in response to stress in human cells, generating heterogeneous-sized RNAs that contain a polyA tail but remain in the nucleus (Valgardsdottir et al. 2005; see Chap. 5 in this book). In addition, many eukaryotic long ncRNAs that have regulatory roles are always polyadenylated (Amaral and Mattick 2008). For instance, the polyA tail is part of the mature *Xist* RNA, which mediates X chromosome inactivation in dosage compensation (Lucchesi et al. 2005; see Chap. 3 in this book)

In conclusion, expression of satDNAs in beetles is developmentally regulated and proceeds in the form of long, stable, nonpolyadenylated transcripts that remain mostly in the nucleus where they probably play a structural role in the organization of pericentromeric heterochromatin. A small portion of transcripts is exported to the cytoplasm where they perform an unknown role. In addition, long transcripts are processed into small RNAs, 21–26 nt long, that are proposed to function as guide molecules during heterochromatin formation.

8.5 Heterochromatin Formation in *Drosophila*: Role of Heterochromatic Transcripts

D. melanogaster heterochromatin is prominent in pericentromeric regions and is mostly comprised of satDNA and transposon elements (TE). As in fission yeast *Schizosaccharomyce pombe*, it is associated with histone H3 methylation on lysine 9

(H3K9) by the histone methylase Su(var)3-9 that enables recruitment of heterochromatin protein HP1 necessary to maintain and spread heterochromatic state (Ebert et al. 2006). It has been speculated for a long time whether an endogenous siRNA pathway, similar to those in *S. pombe*, is involved in the formation of heterochromatin in *Drosophila*. Small RNA molecules related to several types of repetitive DNA have been isolated from *D. melanogaster* (Aravin et al. 2003). These repeat-associated RNAs, 23–26 nt in size, are most abundant in testes and early embryos, which may be related to the regulation of transposon activity and the dramatic changes in heterochromatin structure that occur in these stages.

Examination and analysis of small RNA libraries obtained from different developmental stages of fly revealed the presence of TE-derived small RNAs in all stages: in early embryos, most of them correspond to 25 nt long piwiRNAs. They are formed in gonads from long transcripts of TEs and induce silencing of TEs through a feedback regulatory mechanism involving the Piwi subfamily of Argonaute proteins (Brennecke et al. 2007). In other developmental phases, 25 nt piRNAs are partially replaced by a population of 21 nt long RNAs that also derive from long TE transcripts. Due to the limitation of method of high throughput deep sequencing that is restricted to nontandemly repeated DNA, small RNAs that derive from satDNA were not systematically examined. However, siRNA deriving from 1.688 satellite have a size range between 19 and 28 nt and were detected in early embryos as well as in larvae (Aravin et al. 2003). It has been shown that a nuclear pool of TE-derived 21 nt long siRNAs is involved in heterochromatin formation in somatic cells of *Drosophila* and that components of the RNAi pathway participate in heterochromatin process (Fagegaltier et al. 2009). This implicates similarity between mechanisms of heterochromatin formation in *S. pombe* and *Drosophila* and points to the role of pericentromeric transcripts, either satDNA or transposon-derived, in heterochromatin formation. The possible mechanism by which repeat-derived siRNAs could promote heterochromatin formation in *Drosophila* is by tethering complementary nascent transcript of satDNAs and transposons and guiding chromatin modifiers, such as histone methylase Su(var)3-9, that induce H3K9 methylation. Identification of proteins that tether siRNAs to chromatin in *Drosophila* and other animals needs, however, to be elucidated.

In *Drosophila*, distinct heterochromatic loci are the source of primary piRNAs, which target a large number of transposons that are active in the germline and induce degradation of their transcripts. The mechanism of silencing is not well explained but includes Piwi proteins loaded with piRNAs that target and cleave RNA molecules. It is also proposed that piRNAs could promote chromatin modifications and, recently, a role for Rhino protein, one of the HP1-like proteins, in piRNA generation has been established (Klattenhoff et al. 2009). Rhino protein in germ cells of *Drosophila* replaces HP1 and seems to promote the expression of piRNAs and transposon silencing. It is suggested that piRNAs associated with Piwi protein target Rhino to transposon clusters and promote the production of additional piRNAs from the cluster. In this way, a link between the two major transposon defense pathways involving heterochromatin and RNA silencing mechanism exists (Klattenhoff et al. 2009). This could also represent another mechanism of

heterochromatin formation specific for germline. It can be also proposed that numerous satDNA-derived piRNAs present in germ cells could also contribute to heterochromatin establishment and maintenance by similar mechanism that might include Piwi protein and HP1 germline analog Rhino.

8.6 Possible Regulatory Role of satDNAs and Their Transcripts

Although satellite repeats show remarkable restriction in their distribution along chromosomes to pericentromeric and subtelomeric heterochromatin, there are several exceptions involving minor amounts of satellite sequences present in euchromatin. Such examples of limited localization of satellite sequences in euchromatin involve a simple and a complex satellite. Eight tandem repeats of *D. melanogaster* satellite AATAC are found in front of the s38 chorion gene on X chromosome (Spradling et al. 1987). The 359-bp repeats of the 1.688 satellite, located predominantly in pericentromeric heterochromatin of X chromosome, are also found in other positions of the same chromosome (Tartof et al. 1984). In beetle *T. castaneum* 360 bp repeats of abundant centromeric and pericentromeric satellite TCAST are found dispersed in the vicinity of genes on all chromosomes (unpublished results). The discovery of short satellite segments interspersed among the genes in euchromatic portion of genomes suggest possible regulatory role of these sequences, since they are often source of regulatory elements such as promoters and/or transcription factors binding sites (Ugarković 2005, Fig. 8.1). Recently, a regulatory role of 32 bp satellite repeats located in the intron of the major histocompatibility complex gene (MHIIβ) of fish *Salvelinus fontinalis*, on MHIIβ gene expression was demonstrated (Croisetiere et al. 2010). The level of gene expression depends on temperature being higher at lower temperatures as well as on the length of satellite repeats:

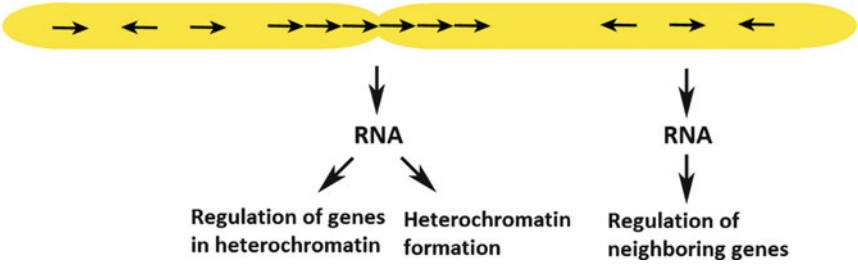

Fig. 8.1 Role of satDNAs and corresponding transcripts in the regulation of genes located in heterochromatin and euchromatin. Transcripts of tandemly repeated satellite repeats, located in (peri)centromeric regions, play a role in heterochromatin formation as well as in the regulation of the genes located in heterochromatin. Transcripts of satDNA repeats dispersed within the euchromatin could play a role in the regulation of the neighboring genes. Transcription of satellite repeats is temperature sensitive, and the role of transcripts in the environmental stress response is proposed

a longer satellite array induces reduced expression. Although the mechanism of *cis*-acting satellite gene regulation is not clear, there is evidence that temperature-sensitive satellites could play an important role in the gene regulation of the adaptive immune response.

Influence of satDNAs and their transcripts on gene regulation could not refer only to genes located in euchromatin but also on heterochromatic genes (Fig. 8.1). It is known that important developmental genes are located in heterochromatin, as revealed for *D. melanogaster,* (Pimpinelli et al. 1985) and that the proximity of heterochromatin is an important regulatory requirement for their function (Dimitri et al. 2009). Heterochromatin is also involved in gene silencing, and this process is developmentally programmed in *Drosophila* and mammals (Lu et al. 1998). Heterochromatin formation in *D. melanogaster* is influenced by transcripts of satDNA elements and transposons present in heterochromatin (see Sect. 8.5 in this Chapter). On the other hand, insect development is very sensitive to changes in the environment, particularly temperature. With a lowering temperature, the length of the development period is prolonged, and at a critical temperature, development ceases altogether. It has been shown that in beetle *T. castaneum,* expression of satDNA is temperature sensitive during embryogenesis, being significantly decreased at low temperatures where development is stopped (unpublished results). It can be proposed that decrease of satDNA expression affects heterochromatin formation during embryogenesis and in this way influences activity of heterochromatin-localized developmental genes. Temperature-sensitive expression of heterochromatic satDNAs also indicates their involvement in the signaling mechanism responsible for insect development, differentiation, and stress response (Fig. 8.1).

8.7 Conclusion

satDNAs are major heterochromatin constituents in many insect species and are found to be transcribed during all developmental stages. Transcripts are heterogeneous in size ranging from long multimers to small interfering RNAs. Their role in heterochromatin establishment and regulation is proposed although the detailed molecular mechanism and proteins involved are not elucidated yet. The satDNA transcription is not constitutive but associated with development and differentiation and is actively regulated by environmental factors such as temperature. It is proposed that satDNAs play a role in regulation of genes positioned within heterochromatin as well as those located in the vicinity of satellite elements in euchromatin. Mechanism of gene regulation is not explained but could be related to the presence of active regulatory elements within satDNAs such as promoters and transcription factor binding sites as well as corresponding satellite transcripts. Further studies are needed in order to explain the complex role of satDNAs and their transcripts in the signaling mechanism responsible for insect development, differentiation, and stress response.

Acknowledgments This work was supported by EU FP6 Marie Curie Transfer of Knowledge Grant MTKD-CT-2006-042248 and grant 00982604 from the Croatian Ministry of Science. Isidoro Feliciello is Marie Curie Fellow at Ruder Boskovic Institute.

References

Amaral PP, Mattick JS (2008) Noncoding RNA in development. Mamm Genome 19:454–492

Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, Snyder B et al (2003) The small RNA profile during *Drosophila melanogaster* development. Dev Cell 5:337–350

Arnason U, Höglund M, Widegren B (1984) Conservation of highly repetitive DNA in cetaceans. Chromosoma 89:238–242

Bonaccorsi S, Gatti M, Pisano C, Lohe A (1990) Transcription of a satellite DNA on two Y chromosome loops of *Drosophila melanogaster*. Chromosoma 99(4):260–266

Borstnik B, Pumpernik D, Lukman D, Ugarković Đ, Plohl M (1994) Tandemly repeated pentanucleotides in DNA sequences of eukaryotes. Nucleic Acids Res 22(16):3412–3417

Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon GJ (2007) Discrete small RNA generating loci as master regulators of transposon activity in Drosophila. Cell 128:1089–1103

Bruvo-Mađarić B, Plohl M, Ugarković Đ (2007) Wide distribution of related satellite DNA families within the genus Pimelia (Tenebrionidae). Genetica 130:35–42

Charlesworth B, Sniegowski P, Stephan W (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215–220

Coats SR, Zhang Y, Epstein LM (1994) Transcription of satellite 2 DNA from the newt is driven by a snRNA type of promoter. Nucleic Acids Res 22:4697–4704

Croisetiere S, Bernatchez L, Belhumeur P (2010) Temperature and length-dependent modulation of the MH class IIβ gene expression in brook charr (*Salvelinus fontinalis*) by a *cis*-acting minisatellite. Mol Immun 47:1817–1829

Davis CA, Wyatt GR (1989) Distribution and sequence homogeneity of an abundant satellite DNA in the beetle, *Tenebrio molitor*. Nucleic Acids Res 17(14):5579–5586

De la Herrán R, Fontana F, Lanfredi M, Congiu L, Leis M, Rossi R, Ruiz Rejón C, Ruiz Rejón M, Garrido-Ramos MA (2001) Slow rates of evolution and sequence homogenization in an ancient satellite DNA family of sturgeons. Mol Biol Evol 18:432–436

Diaz MO, Barsacchi-Pilone G, Mahon KA, Gall JG (1981) Transcripts from both DNA strands of a satellite DNA occur on lampbrush chromosome loops of the newt *Notophthalmus*. Cell 24:649–659

Dimitri P, Caizzi R, Giordano E, Carmela Accardo M, Lattanzi G, Biamonti G (2009) Constitutive heterochromatin: a surprising variety of expressed sequences. Chromosoma 118:419–435

Djupedal I, Kos-Braun IC, Mosher RA, Söderholm N, Simmer F, Hardcastle TJ, Fender A, Heidrich N, Kagansky A, Bayne E, Wagner EG, Baulcombe DC, Allshire RC, Ekwall K (2009) Analysis of small RNA in fission yeast; centromeric siRNAs are potentially generated through a structured RNA. EMBO J 28(24):3832–3844

Dover G (2002) Molecular drive. Trends Genet 18:587–589

Durajlija Žinić S, Ugarković Đ, Cornudella L, Plohl M (2000) A novel interspersed type of organization of satellite DNAs in *Tribolium madens* heterochromatin. Chromosome Res 8:201–212

Ebert A, Lein S, Schotta G, Reuter G (2006) Histone modification and the control of heterochromatin gene silencing in *Drosophila*. Chromosome Res 14:377–392

Epstein LM, Mahon KA, Gall JG (1986) Transcription of a satellite DNA in the newt. J Cell Biol 103:1137–1144

Epstein LM, Gall JG (1987) Self-cleaving transcripts of a satellite DNA in a newt. Cell 48:535–543

Fagegaltier D, Bougé AL, Berry B, Poisot E, Sismeiro O, Coppée JY, Théodore L, Voinnet O, Antoniewski C (2009) The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. Proc Natl Acad Sci USA 106:21258–21263

Ferbeyre G, Smith JM, Cedergren R (1998) Schistosome satellite DNA encodes active hammerhead-ribozymes. Mol Cell Biol 18:3880–3888

Ferree PM, Barbash DA (2009) Species-specific heterochromatin prevents mitotic chromosome segregation to cause hybrid lethality in *Drosophila*. PLoS Biol 7:e1000234

Fitzgerald DJ, Dryden GL, Bronson EC, Williams JS, Anderson JN (1994) Conserved pattern of bending in satellite and nucleosome positioning DNA. J Biol Chem 269:21303–21314

Fry K, Salser W (1977) Nucleotide sequences of HS-α satellite DNA from kangaroo rat Dipodomys ordii and characterisation of similar sequences in other rodents. Cell 12:1069–1084

Gaubatz JW, Cutler RG (1990) Mouse satellite DNA is transcribed in senescent cardiac muscle. J Biol Chem 265:17753–17758

Green B, Pabon-Pena LM, Graham TA, Peach SE, Coats SR, Epstein LM (1993) Conserved sequence and functional domains in satellite 2 from three families of salamanders. Mol Biol Evol 10:732–750

Hamilton A, Voinnet O, Chappell L, Baulcombe D (2002) Two classes of short interfering RNA in RNA silencing. EMBO J 21:4671

Kipling D, Warburton PE (1997) Centromeres, CENP-B and *Tigger* too. Trends Genet 13:141–145

Klattenhoff C, Xi H, Li C, Lee S, Xu J, Khurana JS, Zhang F, Schultz N, Koppetsch BS, Nowosielska A, Seitz H, Zamore PD, Weng Z, Theurkauf WE (2009) The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters. Cell 138(6):1137–1149

Li YX, Kirby ML (2003) Coordinated and conserved expression of alphoid repeat and alphoid repeat-tagged coding sequences. Dev Dyn 228:72–81

Lohe A, Roberts P (1988) Evolution of satellite DNA sequences in *Drosophila*. In: Verma RS (ed) Heterochromatin: molecular and structural aspects. Cambridge University Press, Cambridge, pp 148–186

Lorite P, Carrillo JA, Tinaut A, Palomeque T (2002a) Comparative study of satellite DNA in ants of the *Messor* genus. Gene 297(1–2):113–122

Lorite P, Renault S, Rouleux-Bonnin F, Bigot S, Periquet G, Palomeque T (2002b) Genomic organization and transcription of satellite DNA in the ant *Aphaenogaster subterranea* (Hymenoptera, Formicidae). Genome 45:609–616

Lu BY, Ma J, Eissenberg JC (1998) Developmental regulation of heterochromatin-mediated silencing in *Drosophila*. Development 125:2223–2234

Lucchesi JC, Kelly WG, Panning B (2005) Chromatin remodeling in dosage compensation. Annu Rev Genet 39:615–651

Martínez-Guitarte JL, Díez JL, Morcillo G (2008) Transcription and activation under environmental stress of the complex telomeric repeats of *Chironomus thummi*. Chromosome Res 16 (8):1085–1096

Masumoto H, Masukata H, Muro Y, Nozaki N, Okazaki T (1989) A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. J Cell Biol 109:1963–1973

Meštrović N, Plohl M, Mravinac B, Ugarković Đ (1998) Evolution of satellite DNAs from the genus *Palorus* – experimental evidence for the 'library' hypothesis. Mol Biol Evol 15:1062–1068

Metz A, Soret J, Vourc'h C, Tazi J, Jolly C (2004) A key role for stress-induced satellite III transcripts in the relocalization of splicing factors into nuclear stress granules. J Cell Sci 117: 4551–4558

Mravinac B, Plohl M, Meštrović N, Ugarković Đ (2002) Sequence of PRAT satellite DNA "frozen" in some coleopteran species. J Mol Evol 54:774–783

Mravinac B, Plohl M, Ugarković Đ (2004) Conserved patterns in the evolution of *Tribolium* satellite DNAs. Gene 332:169–177

Mravinac B, Plohl M, Ugarković Đ (2005) Preservation and high sequence conservation of satellite DNAs suggest functional constraints. J Mol Evol 61:542–550

Ohzeki J, Nakano M, Okada T, Masumoto H (2002) CENP-B box is required for *de novo* centromere chromatin assembly on human alphoid DNA. J Cell Biol 159:765–775

Palomeque T, Lorite P (2008) Satellite DNA in insects: a review. Heredity 100:564–573

Pezer Z, Ugarković Đ (2008) RNA Pol II promotes transcription of centromeric satellite DNA in beetles. PLoS One 3:e1594

Pezer Z, Ugarković Đ (2009) Transcription of pericentromeric heterochromatin in beetles – satellite DNAs as active regulatory elements. Cytogenet Genome Res 124:268–276

Pimpinelli S, Sullivan W, Prout M, Sandler L (1985) On biological functions mapping to the heterochromatin of *Drosophila melanogaster*. Genetics 109:701–724

Plohl M, Meštrović N, Bruvo B, Ugarković Đ (1998) Similarity of structural features and evolution of satellite DNAs from Palorus subdepressus (Coleoptera) and related species. J Mol Evol 46:234–249

Pons J (2004) Cloning and characterization of a transponsable-like repeat in the heterochromatin of the darkling beetle *Misolampus goudoti*. Genome 47:769–774

Pons J, Bruvo B, Juan C, Petitpierre E, Plohl M, Ugarković Đ (1997) Conservation of satellite DNA in species of the genus *Pimelia* (Tenebrionidae, Coleoptera). Gene 205:183–190

Raff JW, Kellum R, Alberts B (1994) The Drosophila GAGA transcription factor is associated with specific regions of heterochromatin throughout the cell cycle. EMBO J 13:5977–5983

Renault S, Rouleux-Bonnin F, Periquet G, Bigot Y (1999) Satellite DNA transcription in *Diadromus pulchellus* (Hymenoptera). Insect Biochem Mol Biol 29:103–111

Rojas AA, Vázquez-Tello A, Ferbeyre G, Venanzetti F, Bachmann L, Paquin B et al (2000) Hammerhead-mediated processing of satellite pDo500 family transcripts from *Dolichopoda* cave crickets. Nucleic Acids Res 28:4037–4043

Romanova LY, Deriagin GV, Mashkova TD, Tumeneva IG, Mushegian AR, Kisselev LL, Alexandrov IA (1996) Evidence for selection in evolution of alpha satellite DNA: the central role of CENP-B/pJ alpha binding region. J Mol Biol 261(3):334–340

Rouleux-Bonnin F, Renault S, Bigot Y, Periquet G (1996) Transcription of four satellite DNA subfamilies in Diprion pini (Hymenoptera, Symphyta, Diprionidae). Eur J Biochem 238: 752–759

Rouleux-Bonnin F, Bigot S, Bigot Y (2004) Structural and transcriptional features of *Bombus terrestris* satellite DNA and their potential involvement in the differentiation process. Genome 47:877–888

Rudert F, Bronner S, Garnier J-M, Dollé P (1995) Transcripts from opposite strands of gamma satellite DNA are differentially expressed during mouse development. Mamm Genome 6:76–83

Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF (2001) Genomic and genetic definition of a functional human centromere. Science 294:109–115

Shestakova EA, Mansuroglu Z, Mokrani H, Ghinea N, Bonnefoy E (2004) Transcription factor YY1 associates with pericentromeric *y*-satellite DNA in cycling but not in quiescent (G_0) cells. Nucleic Acids Res 32:4390–4399

Spradling AC, de Cicco DV, Wakimoto BT, Levine JF, Katfayan LJ, Cooley L (1987) Amplification of the X-linked *Drosophila* chorion gene cluster requires a region upstream from the s38 chorion gene. EMBO J 6:1045–1053

Sun X, Wahlstrom J, Karpen GH (1997) Molecular structure of a functional Drosophila centromere. Cell 91:1007–1019

Tartof KD, Hobbs C, Jones M (1984) A structural basis for variegating position effects. Cell 37:869–878

Tomoyasu Y, Miller SC, Tomita S, Schoppmeier M, Grossman D, Bucher G (2008) Exploring systemic RNA interference in insects: a genome-wide survey for RNAi genes in *Tribolium*. Genome Biol 9:R10

Trapitz P, Wlaschek M, Bunemann H (1988) Structure and function of Y chromosomal DNA II. Analysis of lampbrush loop associated transcripts in nuclei of primary spermatocytes of

Drosophila hydei by in situ hybridization using asymmetric RNA probes of four different families of repetitive DNA. Chromosoma 96:159–170

Ugarković Đ (2005) Functional elements residing within satellite DNAs. EMBO Rep 6: 1035–1039

Ugarković Đ (2008) Satellite DNA libraries and centromere evolution. Open Evol J 2:1–6

Ugarković Đ (2009a) Centromere: structure and evolution. Springer, Berlin

Ugarković Đ (2009b) Centromere-competent DNA: structure and evolution. Prog Mol Subcell Biol 48:53–76

Ugarković Đ, Plohl M, Lucijanić-Justić V, Borštnik B (1992) Detection of satellite DNA in *Palorus ratzeburgii*: analysis of curvature profiles and comparison with *Tenebrio molitor* satellite DNA. Biochimie 74:1075–1082

Ugarković Đ, Podnar M, Plohl M (1996) Satellite DNA of the red flour beetle *Tribolium castaneum* – comparative study of satellites from the genus *Tribolium*. Mol Biol Evol 13: 1059–1066

Ugarković Đ, Plohl M (2002) Variation in satellite DNA profiles – causes and effects. EMBO J 21(22):5955–5959

Usakin L, Abad J, Vagin VV, de Pablos B, Villasante A et al (2007) Transcription of the 1.688 Satellite DNA family is under the control of RNA interference machinery in *Drosophila melanogaster* ovaries. Genetics 176:1343–1349

Valgardsdottir R, Chiodi I, Giordano M, Cobianchi F, Riva S, Biamonti G (2005) Structural and functional characterization of noncoding repetitive RNAs transcribed in stressed human cells. Mol Biol Cell 16:2597–2604

Varadaraj K, Skinner DM (1994) Cytoplasmic localization of transcripts of a complex G + C-rich crab satellite DNA. Chromosoma 103:423–431

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

Waye JS, Willard HF (1986) Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome. Moll Cell Biol 6:3156–3165

Wu ZG, Murphy C, Gall JG (1986) A transcribed satellite DNA from the bullfrog *Rana catesbeiana*. Chromosoma 93(4):291–297

Zheng B, Wang Z, Li S, Yu B, Liu JY, Chen X (2009) Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. Genes Dev 23:2850–2860