Available online at www.sciencedirect.com
ScienceDirect
Gene 409 (2008) 72–82
GENE
www.elsevier.com/locate/gene

Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin

Miroslav Plohl [a,*], Andrea Luchetti [b], Nevenka Meštrović [a], Barbara Mantovani [b]

[a] Department of Molecular Genetics, Ruđer Bošković Institute, Bijenička 54, HR-10002 Zagreb, Croatia
[b] Dipartimento di Biologia E. S., Università degli Studi di Bologna, Via Selmi 3, I-40126 Bologna, Italy

Received 6 August 2007; received in revised form 8 November 2007; accepted 20 November 2007
Available online 4 December 2007

Abstract

Satellite DNAs (tandemly repeated, non-coding DNA sequences) stretch over almost all native centromeres and surrounding pericentromeric heterochromatin. Once considered as inert by-products of genome dynamics in heterochromatic regions, recent studies showed that satellite DNA evolution is interplay of stochastic events and selective pressure. This points to a functional significance of satellite sequences, which in (peri) centromeres may play some fundamental functional roles. First, specific interactions with DNA-binding proteins are proposed to complement sequence-independent epigenetic processes. The second role is achieved through RNAi mechanism, in which transcripts of satellite sequences initialize heterochromatin formation. In addition, satellite DNAs in (peri)centromeric regions affect chromosomal dynamics and genome plasticity. Paradoxically, while centromeric function is conserved through eukaryotes, the profile of satellite DNAs in this region is almost always species-specific. We argue that tandem repeats may be advantageous forms of DNA sequences in (peri)centromeres due to concerted evolution, which maintains high intra-array and intrapopulation sequence homogeneity of satellite arrays, while allowing rapid changes in nucleotide sequence and/or composition of satellite repeats. This feature may be crucial for long-term stability of DNA-protein interactions in centromeric regions.

Keywords: Centromere; Concerted evolution; Library model; Pericentromeric heterochromatin; Repetitive sequences

1. Introduction

Repetitive DNA sequences build a significant portion of every eukaryotic genome, and among them, a large fraction comprises sequences repeated in tandem, commonly known as satellite DNAs (Charlesworth et al., 1994; Elder and Turner, 1995; Schmidt and Heslop-Harrison, 1998). The term "satellite DNA" is historical, because this kind of sequences was initially isolated from satellite bands in experiments with gradient centrifugation, due to the difference in A+T content from the rest of genomic DNA (Szybalski, 1968). Since no protein coding function could be primarily associated with satellite DNAs, early hypothesis considered them as useless genomic elements accumulated as junk (Ohno, 1972), or alternatively, as sequences that represent genomic parasites proliferating for their own sake (Orgel and Crick, 1980). An opposite view suggested the involvement of satellite DNAs in a series of functions ranging from chromosome organization and pairing to cell metabolism and speciation (John and Miklos, 1979). More recent studies supported these functionalist assumptions concerning the association of satellite DNAs with complex features of eukaryotic chromosomes (for example, Csink and Henikoff, 1998; Henikoff et al., 2001; Sullivan et al., 2001). Accordingly, satellite sequences are the main constituent of centromeric and pericentromeric heterochromatin, two epigenetically determined regions responsible for correct pairing and disjunction of eukaryotic chromosomes in cell divisions (see for example Arney and Fisher, 2004; Hall et al., 2004; Bloom, 2007 for reviews). The extreme diversity of satellite DNAs, in nucleotide sequence, complexity, genomic abundance, as well as the existence of a large number of unrelated satellite DNA families,

Abbreviations: bp, base pair(s); ChIP, chromatin immunoprecipitation; HOR, higher-order repeat; kb, kilo base pair(s); Mb, mega base pair(s); Myr(s), Million year(s).

* Corresponding author. Tel.: +385 1 4561 083; fax: +385 1 4561 177.
E-mail address: plohl@irb.hr (M. Plohl).

0378-1119/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2007.11.013

blurs the putative functions these sequences might have in a genome. Despite progressive accumulation of data, direct evidences about function(s) of satellite sequences and mechanisms of their action in particular processes are still mostly lacking, and great part of our conclusions rely more on intuitive deduction than on direct experimental evidences.

Some conceptual properties of satellite DNA composition and evolution were already reviewed elsewhere (Charlesworth et al., 1994; Elder and Turner, 1995; Schmidt and Heslop-Harrison, 1998; Ugarković and Plohl, 2002). Here, in the light of recently published data, we present some general features of satellite DNA sequences and discuss possible significance of satellite repeats and their evolution in centromeric and pericentromeric genomic regions.

2. Basic features of satellite repeats

Satellite DNAs can be defined as highly reiterated non-coding DNA sequences, organized as long arrays of head-to-tail linked repeats located in the constitutive heterochromatin, the eukaryotic chromosomal regions that remain condensed throughout the cell cycle (Heitz, 1928). Basic repeating units, satellite DNA monomers, are often A+T rich and range in length from only few bp up to more than 1 kb, building up to 100 Mb long arrays. The preferential monomer length of 150–180 bp and 300–360 bp detected in many satellites in both plants and animals is often considered to mirror requirements of DNA length wrapped around one or two nucleosomes (Schmidt and Heslop-Harrison, 1998; Henikoff et al., 2001). Satellite DNA contribution to total genomic content varies significantly among species, exceeding sometimes 50% of total DNA (Elder and Turner, 1995; Schmidt and Heslop-Harrison, 1998), and consequently they are involved in the enormous variation of genome size in eukaryotes (Doolittle and Sapienza, 1980; Cavalier-Smith, 1985; Gregory et al., 2007). Megabase-long arrays of repeats in heterochromatin mainly differentiate satellite DNAs from two other categories of tandemly repeated non-coding DNA elements, i.e. mini- and microsatellites. These elements are characterized by short repeating units (2–6 bp for microsatellites and 15–60 bp for minisatellites) repeated in moderate abundance, and located predominantly in euchromatic chromosomal regions (Charlesworth et al., 1994). Further, microsatellite arrays can be often also detected in heterochromatin (Gindullis et al., 2001).

Current strategy for satellite DNA detection and characterization is genomic DNA digestion with restriction endonucleases, followed by sequence analysis of randomly cloned monomers or short multimers. This approach is still ongoing, even after the burst of large-scale genome sequencing projects: low variability of satellite monomers and paucity of intervening sequences inserted into megabase-sized arrays impose serious limitations on current sequencing and mapping techniques to assemble tandemly repeated motifs into large contigs. Since satellite DNAs are dominant components in the heterochromatin, it remains underrepresented in outputs of genome projects, and the information concerning general organization and functional significance of its DNA sequences is still only limited (Eichler et al., 2004; Nagaki et al., 2004; Rudd and Willard, 2004; Hoskins et al., 2007).

3. Concerted evolution

Even if satellite DNA monomers are present in many thousands copies per genome, sequence divergence between monomers of the same family is often very low, usually up to 15% (see for example, King and Cummings, 1997). However, the divergence can be much higher in some cases, such as in monomeric alpha-satellite, in which repeating units diverge for up to 30% (Rudd and Willard, 2004). High repeat homogeneity is achieved by non-independent evolution of monomers. It is a consequence of molecular drive, a two-level process in which mutations are homogenized throughout members of a repetitive family, and concomitantly fixed within a group of reproductively linked organisms (Dover, 1982, 1986). The consequence is concerted evolution of monomers constituting a satellite DNA family.

Sequence homogenization is due to diverse molecular mechanisms of nonreciprocal transfer, such as unequal crossover, gene conversion, rolling circle replication and reinsertion, and transposon-mediated exchange (Stephan, 1986; Dover, 2002; Glinka et al., 2006). While it is not clear which of the above reported mechanisms is preferentially involved in sequence homogenization, it is generally acknowledged that these mechanisms act more efficiently within localized subsets of satellite repeats, while efficiency drops progressively when changes are homogenized between arrays on the same chromosome, homologous and heterologous chromosomes (Fig. 1; Dover, 1986). Because of differences in rates of local and global sequence homogenization, adjacent monomers show a higher degree of sequence similarity than those retrieved at random, and can be often grouped into subsets or subfamilies, defined by diagnostic mutations (Willard and Waye, 1987; Durfy and Willard, 1989; Schindelhauer and Schwarz, 2002; Hall et al., 2005; Roizes, 2006). Distinctive groups of monomer variants are usually chromosome-specific. As predicted by theoretical models (Smith, 1976; Stephan, 1989), monomers at array ends are more divergent than those located centrally due to the low efficiency of homogenization mechanisms (predominantly unequal crossover) in bordering regions of the satellite array (Mashkova et al., 1998; McAllister and Werren, 1999; Bassi et al., 2000; Schueler et al., 2005).

Adjacent monomer variants can be sometimes homogenized together and form a new, composite higher-order repeat (HOR) unit in which former monomers became subrepeats or subunits (Willard and Waye, 1987; Warburton and Willard, 1990). Since a HOR is a homogenization unit, HORs generally show high level of sequence identity, while substantial sequence divergence is accumulated among constituent subunits. For example, in the human alpha-satellite, HORs are highly homogeneous and are typically 97–100% identical, while subunits within them are on average ~70% identical (Willard and Waye, 1987; Roizes, 2006 and references therein). The increase in repeat unit length and complexity by merging shorter repeat motifs into a HOR seems to be a common trend in at least some satellite

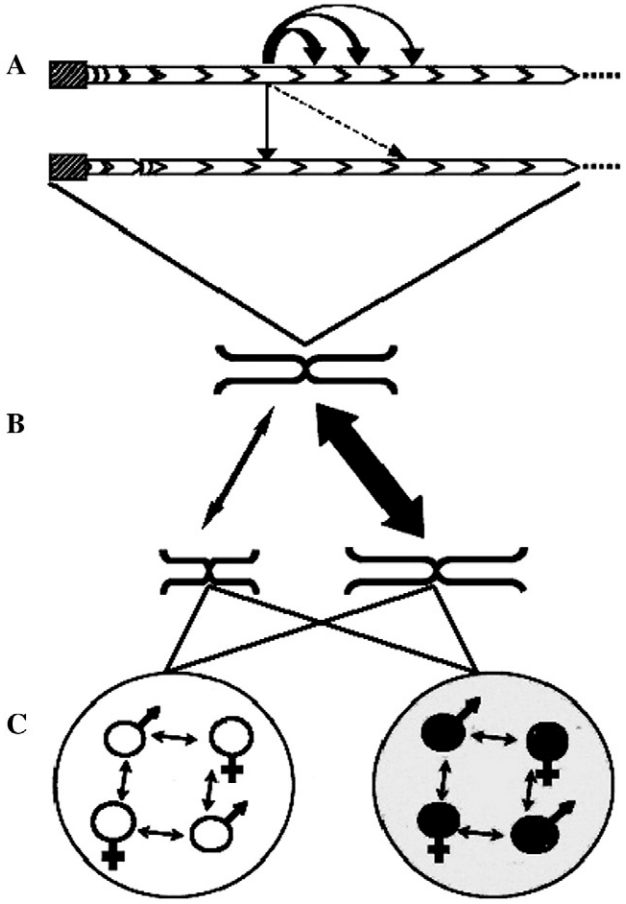

Fig. 1. Molecular drive and concerted evolution. A) Homogenization of mutated variant throughout the members of a repetitive family within an array and between sister chromatids. Arrow appearance is correlated with the homogenization efficiency. B) Efficiency of homogenization between homologous and non-homologous chromosomes. C) Variant spread among individuals (fixation) depends on bisexuality and population factors. Reproductive isolation leads to fixation of different repeat variants in genomes of different evolutionary units (white and black circles).

DNAs and/or organisms (for example, Willard and Waye, 1987; Modi et al., 2004; Mravinac et al., 2005). This process can be predicted theoretically if low unequal cross-over to mutation ratios are assumed as specific for centromeric regions (Stephan and Cho, 1994).

While homogenization depends on mechanisms of genomic turnover, fixation results from random chromosomal assortment in sexual reproduction through meiosis and amphimixis, depending thus on population factors. The final outcome of concerted evolution is higher repeat homogeneity within lineages (strains, populations, subspecies, species, etc.) than between them (Dover, 1982; Rudd et al 2006; Ellingsen et al., 2007). This model has received considerable support from other repetitive DNA elements, such as rRNA genes (Dadejová et al., 2007) and a repeated protein coding gene (Carmon et al., 2007). Effects of fixation are excluded in satellite DNAs from parthenogenetic organisms, due to the lack of bisexual reproduction. As expected, satellite sequences show higher intra- than intergroup homogeneity in bisexual taxa, while unisexuals exhibit similar variability disregarding the level of taxonomic position and parthenogenetic mechanism (Mantovani, 1998; Luchetti et al., 2003; Fig. 2). Exceptionally, specific biological traits can lead to the non-concerted patterns of satellite DNA evolution (Lorite et al., 2004; Luchetti et al., 2006). In eusocial termites, for example, satellite monomers accumulate mutations which have no possibility to be spread (or eliminated) owing to the limited number of reproducers, thus leading to a high level of variability uniformly distributed among taxa (Luchetti et al., 2006).

Due to the nature of homogenization and fixation, molecular drive is assumed to be an entirely stochastic process (Smith, 1976; Stephan, 1986, Charlesworth et al., 1994), during which mutations rapidly accumulate in a gradual manner (Bachmann and Sperlich, 1993). An opposite effect, i.e. an extreme conservation of nucleotide sequence, can be also predicted by the model, if "non-desirable" mutations are preferentially eliminated instead of being spread throughout a satellite family (Ohta and Dover, 1984; Dover and Flavell, 1984; Strachan et al., 1985). The nucleotide sequence of some satellite families indeed remained "frozen" for long periods, even for tens of Myrs (for example, Arnason et al., 1992; Heikkinen et al., 1995; Vershinin et al., 1996; Mravinac et al., 2002; Cesari et al., 2003; Robles et al., 2004; Meštrović et al., 2006a). Although the basis for favoring one sequence variant over another is usually not known (Dover, 1987), it might mirror constraints imposed on satellite sequences by some functional interactions. In that case, the evolution of at least some satellites seems to be driven by an interplay of selective constraints and stochastic events (Hall et al., 2003, 2005; Meštrović et al., 2006b).

4. Satellite DNAs in and around centromeres

4.1. Diversity of DNA sequences at the centromere

Centromere is a multidomain locus necessary for poleward chromosomal segregation in mitosis and meiosis. Functional centromeres are usually embedded into large blocks of pericentromeric heterochromatin, but chromatin structure in centromeres is distinct from that in heterochromatin and in euchromatin (Sullivan and Karpen, 2004). Principal DNA components underlying the majority of centromeres in plants and animals are satellite repeats, as corroborated by chromatin immunoprecipitation (ChIP) data (e.g. Nagaki et al., 2003; Zhong et al., 2002; Lee et al., 2005). While centromere structure and function is conserved through complex eukaryotes, DNA sequences in that region are paradoxically variable (Henikoff et al., 2001). For example, centromere-specific immunoprecipitation revealed divergent satellite DNAs in centromeres of several rice species (Lee et al., 2005). The extreme diversity of rapidly evolving satellite repeats, or even their absence in neocentromeres (e.g., Wong and Choo, 2001) agree with the idea about epigenetic determinants of centromere function and inheritance; with satellite sequences being considered to be neither necessary nor sufficient for centromere dynamics (Karpen, 1994; Henikoff et al., 2001). Nevertheless, the abundance of satellite DNAs in the majority of native centromeres indicates that satellite DNA amplification together with retrotransposon

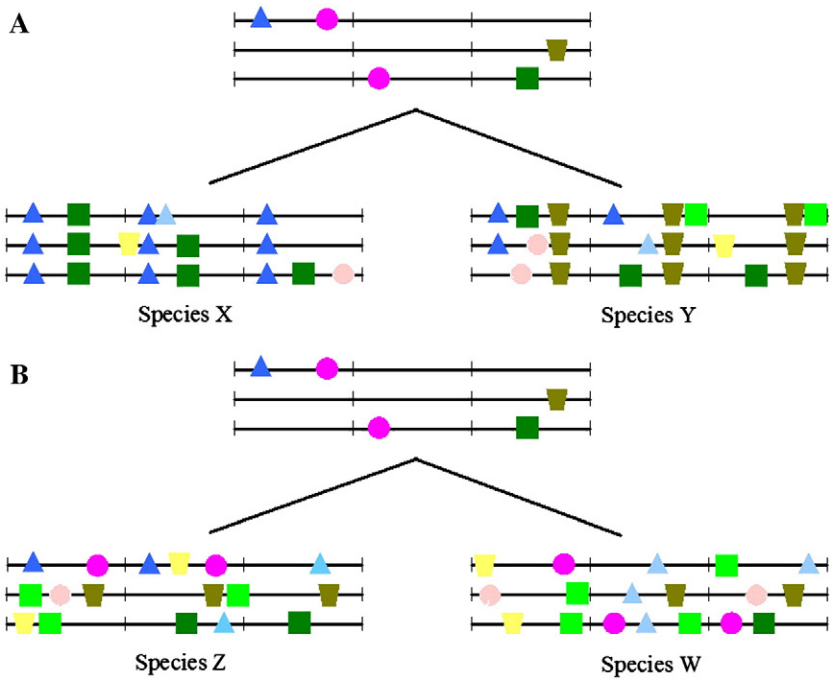

Fig. 2. Effect of reproductive strategies on concerted evolution. Segmented line represents monomer variants within one individual, symbols indicate mutations, groups of lines species. A) In sexual species mutations spread both throughout repeats (homogenization) and organisms (fixation), or they disappear. Transition stages of concerted evolution are reflected by incompletely homogenized mutations (see also Fig. 4 in Strachan et al., 1985). New mutations (light colors) appear at the same time and absolute identity of repeats is never achieved. Reproductively isolated units (populations, species etc.) accumulate group-specific mutations. B) According to the theory of concerted evolution, the lack of panmissy and chromosome reshuffling leads to the failure of variant fixation in the unisexual species.

accumulation are important for centromere expansion and stabilization in the evolution from gene-containing non-satellite DNA neocentromere towards typical eukaryotic centromere (Nagaki et al., 2004; Ma and Jackson, 2006).

Satellite arrays in the centromere are usually much longer than it is necessary for the centromeric function. ChIP experiments using antibodies against centromere-specific histones revealed that 30–50% of chromosome-specific higher-order alpha-satellite repeats forms centromeres in human chromosomes 7, 17, and X, while the rest of the homogeneous arrays constitute the flanking heterochromatin (Lam et al., 2006). Similar studies showed that only about 15% of 180-bp satellite resides in *Arabidopsis* centromeres (Nagaki et al., 2003).

4.2. Organizational patterns of centromeric-pericentromeric satellite DNA arrays and intervening sequences

Satellite families in (peri)centromeric regions vary significantly in copy number, nucleotide sequence, organizational patterns and number and nature of inserted non-satellite DNA sequences. Domains formed by single satellites are usually several hundreds kb, or even Mb long, such as in humans (Shiels et al., 1997; Mahtani and Willard, 1998; Schueler et al., 2001), *Drosophila* (Sun et al., 2003), or *Arabidopsis* (Heslop-Harrison et al., 2003). Among individuals, array length of a single satellite can be highly polymorphic. For example, array length in alpha-satellite from human X varies almost 3 times (Mahtani and Willard, 1990).

In *Arabidopsis*, despite a small genome, a 180-bp centromeric satellite is present on all five chromosome pairs, forming up to 4 Mb long tracts, occasionally interrupted by retrotransposons and other types of repetitive DNA elements (Heslop-Harrison et al., 1999, 2003; Kumekawa et al., 2000, 2001). About 300 kb long arrays of centromerically located mouse minor satellite are surrounded with about 10 times more abundant major satellite, both of them being interrupted with short stretches of two GC-rich families (Kuznetsova et al., 2006 and references therein). Interestingly, all four mouse satellite DNAs were detected in threads of locally decondensed pericentromeric heterochromatin that physically links mitotic chromosomes (Kuznetsova et al., 2007). Organizational pattern in which two satellite families form an irregular patchwork of interspersed short arrays (<70 kb) of each satellite has been observed in the beetle *Tribolium madens*. The same pattern probably extends over functional centromeres of all chromosomes (Durajlija-Žinić et al., 2000).

The best studied long-range organization is with no doubt that of alpha-satellite DNA in the pericentromeric and centromeric region of human chromosomes (Rudd and Willard, 2004). Briefly, two major types of phylogenetically distinct alpha-satellite DNA exist. Multimers of ~171 bp monomers form centrally located, chromosome-specific HORs, flanked with domains of irregularly interspersed variants, called monomeric alpha-satellite. Homogeneous array of X-chromosome-specific HOR, DXZ1, spans several megabases, and includes the region of kinetochore formation (Mahtani and Willard, 1998, Schueler et al., 2001; Lam et al., 2006). Such pattern of two distinct organizational forms of alpha-satellite is characteristic for centromeric heterochromatin in all human and higher primate chromosomes (Rudd and Willard, 2004; Rudd et al., 2006). It

has been proposed that highly homogeneous alphoid arrays may facilitate proper centromeric function (Basu et al., 2005; Roizes, 2006).

The contribution of satellite repeats can be only partial in some centromeric regions. A well-known example is given by the centromeric domain of rice chromosome 8, with only 68.4 kb of the satellite DNA family CentO, split into three short arrays (Cheng et al., 2002; Wu et al., 2004; Nagaki et al., 2004). The whole domain into which these arrays are embedded is at least 2 Mb long and built mostly of transposable elements. Besides small amounts of satellite repeats, it also contains unique DNA sequences, including transcriptionally active genes (see also below). ChIP experiments using the rice CENH3 histone-like protein enabled location of the kinetochore within 750 kb long region, and this functional region includes all arrays of CentO satellite (Wu et al., 2004; Nagaki et al., 2004). Based on these results, it has been hypothesized that low abundance of satellite repeats in the centromere represents an early stage in the centromere evolution, characterized by progressive accumulation of satellite repeats in mature centromeres. Short satellite arrays in the rice chromosome 8 enabled also, for the first time, DNA sequence assembly and continuous reading through the whole centromeric heterochromatin of one native chromosome (Wu et al., 2004; Nagaki et al., 2004).

One of the most intriguing finding about satellite array organization is given by the occurrence of transcriptionally active genes within domains of satellite DNAs. Surprisingly, these genes require the heterochromatic environment for their normal activity; if translocated to euchromatin their expression is impaired (Howe et al., 1995). Recent assembly and mapping of non-satellite components in *Drosophila melanogaster* heterochromatin revealed a minimum of 230 to 254 protein coding genes, conserved in related species (Smith et al., 2007). Long arrays of satellite repeats can represent introns of these genes, such as in the case of the gene for dynein (Kurek et al., 2000). Transcriptionally active genes have been found in the centromeric domain of rice chromosome 8, also near satellite repeats in the region of kinetochore formation (Nagaki et al., 2004; Wu et al., 2004; Yan and Jiang, 2007). However, besides transposable elements, no unique sequences and gene candidates could be detected in *Arabidopsis* (Hosouchi et al., 2002), or human centromeres (Schueler et al., 2001).

5. The evolution of satellite families and the centromeric function

As discussed above, satellite DNAs in (peri)centromeric heterochromatin as well as in genome in general, represent rapidly evolving components. Consequently, even among the most closely related species, they differ in nucleotide sequence, copy number, and/or composition of satellite families (reviewed in Schmidt and Heslop-Harrison, 1998; Ugarković and Plohl, 2002). Rapid evolution of satellite DNA sequences is possible owing to the accumulation of nucleotide divergences, usually with a high rate and in a gradual manner (Bachmann and Sperlich, 1993). Gradual accumulation of mutations follows phylogeny at different hierarchical ranks. At the species level, centromeric satellite DNAs were informative in phylogenetic studies of the *Drosophila obscura* group (Bachmann and Sperlich, 1993), or in the study of the fish family Sparidae (Garrido-Ramos et al., 1999). Determination of ecotype-specific variants in the *Arabidopsis thaliana* 180-bp satellite indicated accumulation of divergences within the last ~5 Myr (Hall et al., 2003; Ito et al., 2007). Even within a genome, distinct forms of satellite DNAs can accumulate mutations with different rates, adding to the diversity of sequence patterns in (peri)centromeric areas. Alpha-satellite repeats, for example, occur as monomeric and higher-order units; these two distinct forms accumulate mutations with different evolutionary rates. Interestingly, centromerically located higher-order units diverge more rapidly than pericentromerically located monomeric repeats (Rudd et al., 2006).

Accumulation of mutations in satellite families is not the only way to alter specific profiles of satellite repeats in short evolutionary periods. Since more than one satellite family exists in a genome, expansions and contractions of satellite arrays can efficiently change a landscape of DNA sequences in heterochromatin by replacing one dominant (major) satellite repeat with another one less represented (reviewed in Ugarković and Plohl, 2002). In this, unequal crossover is proposed to be the major mechanism responsible for dramatic fluctuations in the copy number of satellite DNAs (Smith, 1976). The occurrence of species-specific profiles as a consequence of copy number changes in a set of satellite DNAs shared by related genomes was originally explained through the library model (Fry and Salser, 1977), and experimentally verified in the study of satellite DNAs shared by species of the insect genus *Palorus* (Meštrović et al., 1998; Fig. 3). Copy number changes may be, but are not necessarily, accompanied by rapid evolution of nucleotide sequences, and can explain species-specificity of satellite profiles even when satellite sequences remain “frozen” during long evolutionary periods (Meštrović et al., 1998; Mravinac et al., 2002; Bruvo et al., 2003, see also Section 3). Not only distinct satellite DNAs, but also monomer variants from a single family can be distributed in genomes in the form of a library (Cesari et al., 2003). In the constitution of a library, besides stochastic events, selection might represent a limiting factor for persistence of particular satellite sequences, as indicated by the study of inter-satellite variability in a set of related repeats differentially amplified in a group of taxa (Meštrović et al., 2006b). In addition to nucleotide changes and expansions-contractions of satellite arrays, large-scale changes, such as segmental duplications, play an important role in the rapid evolution of DNA sequences in and around centromeric regions (for example, Cardone et al., 2004; Hall et al., 2004; Ventura et al., 2007).

Satellite repeats may be the preferred form of DNA sequences in functional centromeres and their flanking regions just because of their unique characteristic to maintain sequence homogeneity over long stretches of DNA, and simultaneously to change rapidly in evolution, as explained in the above paragraphs. Concerted evolution and abundance of satellite repeats may stabilize interactions with DNA-binding proteins and eliminate effects of possible unwanted mutations, and in the same time the whole array can rapidly adopt new sequence variant which can better fit the mentioned interactions. In a recent model proposed by Dawe

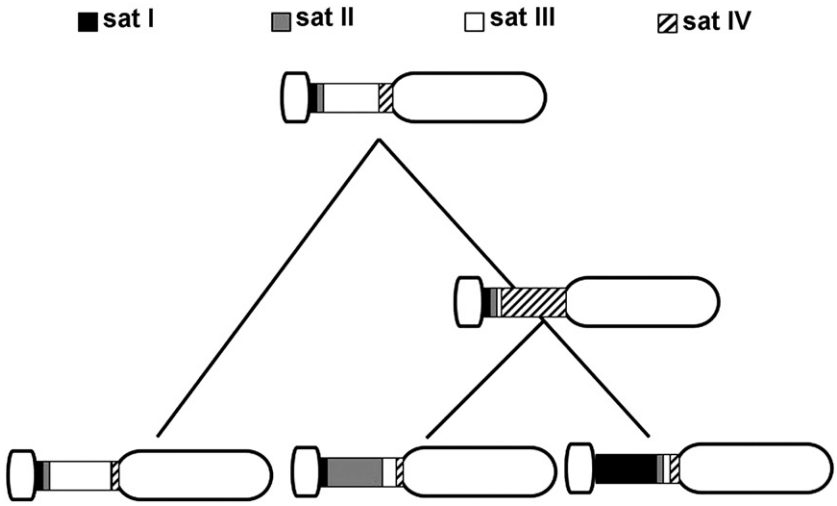

Fig. 3. The library model. In a genome, several satellite DNA families are coexisting on chromosomes, one family being often at high copy number as the major satellite. Different families can be preferentially amplified in the derived chromosomes, therefore changing the relative contribution of each family to the (peri) centromeric chromatin and leading to species-specific profiles of satellite repeats.

and Henikoff (2006), centromeres are defined in a dual way, through an interplay of epigenetic factors and through interactions between "selfish" repetitive DNA sequences and protein components in a centromere, mediated by meiotic drive. Rapid evolution of satellite DNA sequences without impairment of centromeric function can be explained assuming that DNA-binding centromeric proteins such as CENH3 and CENP-C coevolve with satellite DNAs (Fig. 4A). This coevolution may be driven either by changes in satellite DNAs (Malik and Henikoff, 2001; Talbert et al., 2004), or by satellite repeats competition to better fit the chromatin environment (Dawe and Henikoff, 2006). In other words, both DNA and protein evolution drive each other in a centromere, thus providing a stable, but flexible system, able to work on genetic and epigenetic platforms and, if necessary, to rescue chromosomal function by forming new centromeres on non-specialized locations (Dawe and Henikoff, 2006). In agreement with the library model, it can be additionally proposed that one satellite family can replace another one if their sequences are of similar functional value, as discussed recently by Meštrović et al. (2006b). In an ideal case, rapid replacements of equivalent DNA sequences would be possible without alterations in binding affinities (Fig. 4B). However, it can be reasonably expected that the evolution of DNA and DNA-binding centromeric proteins is an integrative result of all these processes in a particular organism. Whatever the scenario could be, a proposed direct consequence is that divergences in satellite sequences and corresponding proteins accumulated between individuals can cause incompatibilities in hybrids and eventually lead to reproductive isolation acting thus as a trigger in speciation process (Meštrović et al., 1998; Henikoff et al., 2001; Hall et al., 2005).

6. Functional potential of satellite DNAs

6.1. Informational capacity of centromeric satellites

To participate in functional interactions in heterochromatin and/or in the functional centromere, satellite repeats should contain sequence motifs recognized by protein components. The best known is CENP-B box, the 17-bp long sequence motif found in its functional form in a subset of higher-order alpha-satellite monomers. The motif binds the CEN-B protein, suggested to facilitate kinetochore formation (Ikeno et al., 1994; Masumoto et al., 2004; Schueler et al., 2005). Motifs resembling to CENP-B box were observed in diverse satellite families from various species, but their true functional significance is not

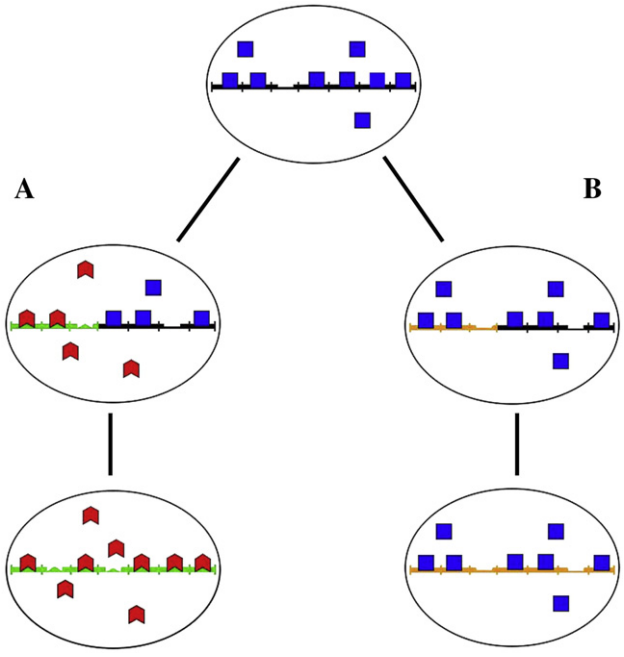

Fig. 4. Evolution of satellite DNAs in centromeric regions. A) Centromeric DNA-binding proteins and underlying satellite repeats drive each other's evolution. A mutated satellite monomer variant (green) with modified binding site is able to bind the mutated centromeric protein (red); they will replace the old protein/satellite pair. B) Rapid changes in satellite DNA profiles without affecting DNA-binding proteins is possible by copy number changes in a set of satellite families with equivalent binding affinities.

known (Canapa et al., 2000; Lorite et al., 2004). Besides CENP-B box, other binding sites might also exist in satellite DNAs to recognize, for example, CENP-C or other centromeric proteins (Henikoff and Dalal, 2005). Uneven distribution of polymorphism observed in alpha-satellite and in *A. thaliana* 180-bp satellite might help to identify unknown DNA sequence motifs that evolve under constraints due to functional interactions (Hall et al., 2003). Alternatively, segments with reduced polymorphism may act as sites that promote recombination among monomers (Hall et al., 2005). Sequence segments with differential substitution rates were also observed in satellite DNAs from arthropods (Mravinac et al., 2004; Luchetti et al., 2004), nematodes (Meštrović et al., 2005), and molluscs (Petrović and Plohl, 2005).

Besides sequence motifs, other features of satellite repeats can be informative in putative functional interactions. Similarity in monomer length of many centromeric satellites (often around 170 bp), and rare length-affecting mutations, led to the assumption that the repeat unit itself might reflect uniformity in nucleosome phasing and heterochromatin propagation (Henikoff et al., 2001). Many, but not all, (peri)centromeric satellite DNAs can adopt tertiary structure due to the sequence-induced bent helix axis (Martinez-Balbas et al., 1990; Fitzgerald et al., 1994). For example, structural features and repeat unit length are similar in evolutionary unrelated satellites from a group of tenebrionid beetles, indicating constraints above the nucleotidic sequence level (Plohl et al., 1998). Inverted sequence elements able to form stable cruciform structures are suggested to act as signals involved in DNA-protein interactions (Bigot et al., 1990). Long inverted subunits might be evolutionary favored in some satellite monomers due to increased stability of secondary structure (Mravinac et al., 2005). Direct and inverse repeated sequence elements in satellite monomers could be related to transposition as a mechanism in the process of concerted evolution (Dover, 2002). Different classes of transposable elements were recognized as sources of (peri)centromeric satellites in various organisms, for example in whales (Kapitonov et al., 1998), *Drosophila* (Heikkinen et al., 1995; Miller et al., 2000), and potato (Tek et al., 2005).

6.2. Satellite DNAs and transcription

Transcriptional activity was not expected for repetitive DNA sequences residing in the transcriptionally suppressive heterochromatin environment. However, satellite transcripts were until now detected in many animal and plant taxa indicating that satellite transcription might be a general phenomenon (for example, Varadaraj and Skinner, 1994; Lorite et al., 2002; Rudert et al., 1995; Pathak et al., 2006; Lee et al., 2006). Satellite DNAs never show any prominent open reading frame, and accordingly, transcript translation has never been demonstrated. But, there are other possible functional roles; for example, the strand or tissue or stage specificity observed in some cases suggests the involvement of satellite transcripts in regulatory functions (Varadaraj and Skinner 1994; Lorite et al., 2002; Pathak et al., 2006). Some satellite DNAs from insects, nematodes and amphibians produce hammerhead structure with a possible ribozymic activity (Rojas et al., 2000 and references therein). Transcripts of centromeric satellite in maize were shown to remain tightly bound within centromeric chromatin and contribute to initiation and stabilization of kinetochore chromatin structure (Topp et al., 2004).

Another proposed role of satellite DNA transcripts attracted particular attention during recent years. It has been observed that satellite DNA transcripts are involved in the initiation of histone H3 methylation, a necessary prerequisite in heterochromatin formation and maintenance (Volpe et al., 2002; Martienssen, 2003). Transcripts from centromeric satellites are processed to produce small interfering RNAs (siRNA) that mobilize a number of proteins and specifically target their coding sequence. This sequence is then packed into the transcription-inhibiting heterochromatin structure (reviewed in Grewal and Elgin, 2007). This mechanism requires low-levels of transcription and may be universal, since siRNAs processed from centromeric satellite repeats were identified in several eukaryotic species (Lee et al., 2006 and references therein). Fukagawa et al. (2004) demonstrated that Dicer-related RNA interference machinery is involved in the formation of the heterochromatin structure in higher vertebrate cells. However, the relationship between transcription of centromeric satellite repeats and centromeric silencing/centromere function is unclear, possibly being much more complex than those reported in *Schizosaccharomyces pombe* (Martienssen et al., 2005; Grewal and Elgin, 2007).

In conclusion, it is evident that transcripts from centromeric satellites may play in quite different scenarios. If a lot of work is still required to understand the roles of these satellite transcripts, possibly many more efforts are needed to highlight the relationships between satellite DNA and proteins or molecules responsible for transcriptional activities of satellite DNAs (Kropotov et al., 2006; Blattes et al., 2006).

6.3. Satellite DNAs and chromosome dynamics

Among different, although indirect, potential roles of satellite DNAs, chromosomal repatterning should be mentioned. However, there is no evidence that karyotipic repatterning leads to speciation *per se*; it is better suggested that it may just enhance reproductive isolation (Coghlan et al., 2005). For example, a link between satellite DNAs and chromosomal instability was studied in the genus *Ctenomys*, one of the most specious and karyotypically diverse mammalian taxon. The high karyotypic variability was associated with amplifications and deletions of the major *Ctenomys* satellite DNA and with the number of species (Slamovits et al., 2001; Hartmann and Scherthan, 2004; Ellingsen et al., 2007). Opposite to the possible diversification role played by satellite sequences in *Ctenomys*, the organization of tandem repeats at centromeres and telomeres of mouse telocentric chromosomes seems to reflect a mechanism of frequent recombinational exchanges between non-homologous chromosomes which promotes the maintenance of the telocentric karyotype (Kalitsis et al., 2006). Finally, the amplification of satellite DNA sequences in the cat chromosomes was related to mitotic instability, which could explain

the exhibition of complex patterns of chromosome aberrations detected in the analyzed fibrosarcoma (Santos et al., 2006).

Satellite DNAs appear also involved in genome restructuring during development in different organisms. The process of chromatin diminution is known to occur during development in different organisms, such as in the nematodes *Parascaris univalens* and *Ascaris suum*, in copepods and in a hagfish. The quantity of lost DNA ranges up to 94%, and is mainly composed of satellite sequences (Stanley et al., 1984; Drouin, 2006, and references therein). The hypothesis that RNAi-related mechanisms are involved in chromatin diminution has been put forward (Drouin, 2006). Chromosomal breaking regions suggest that satellite DNAs represent "recombinational hot-spots" of a genome reorganization. The elimination process may ensure the maintenance of a somatic genome and at the same time allow extremely rapid and profound evolutionarily changes in the organization of the germ line (Bachmann-Waldmann et al., 2004).

7. A final comment

At this point of our knowledge, it is evident that satellite DNAs in (peri)centromeric regions must play at least some fundamentally important functional roles. Satellite repeats in both plants and animals are likely to be involved in specific interactions, thus complementing epigenetic processes. Another role of satellite DNAs in heterochromatin formation and maintenance is initialized by satellite transcripts processed in the form of small interfering RNAs. While the first role depends on information encoded in DNA sequence or in some other features of satellite repeats, the second one is thought to be sequence-independent.

Profiles of satellite repeats in centromeric and pericentromeric regions change rapidly in evolution as a consequence of efficient spread and homogenization of new mutations within a repetitive family, and/or because of amplifications and contractions of diverse repetitive families that replace each other as dominant satellite sequence. Arrays of tandemly repeated monomers may be evolutionary favored form of DNA sequence in and around centromeres because of their unique characteristic to maintain sequence homogeneity over long DNA segments, and, if necessary, during long time-periods. In the same time, satellite DNAs have enormous potential to change extremely rapidly in nucleotide sequence and/or in copy number. These features can be under constraints in order to maintain the best fit with the DNA-binding components.

It is clear that the diversity of satellite repeats and their organizational patterns would require far more detailed studies in various experimental systems in order to reach more general conclusions about functional organization of satellite repeats in a centromere and in the surrounding heterochromatin. This ultimate goal is contrasted by the consistent technical difficulties in reconstructing contigs from the regions of repetitive nature.

Acknowledgements

This work was supported by the Research Fund of the Republic of Croatia, project No. 098-0982913-2756 to M. Plohl, and by RFO funds from the University of Bologna to B. Mantovani.

References

Arnason, U., Gretarsdottir, S., Widegren, B., 1992. Mysticete (baleen whale) relationships based upon the sequence of the common cetacean DNA satellite. Mol. Biol. Evol. 9, 1018–1028.

Arney, K.L., Fisher, A.G., 2004. Epigenetic aspects of differentiation. J. Cell Sci. 117, 4355–4363.

Bachmann, L., Sperlich, D., 1993. Gradual evolution of a specific satellite DNA family in *Drosophila ambigua*, *D. tristis*, and *D. obscura*. Mol. Biol. Evol. 10, 647–659.

Bachmann-Waldmann, C., Jentsch, S., Tobler, H., Muller, F., 2004. Chromatin diminution leads to rapid evolutionary changes in the organization of the germ line genomes of the parasitic nematodes *A. suum* and *P. univalens*. Mol. Biochem. Parasitol. Mar. 134, 53–64.

Bassi, C., et al., 2000. Molecular structure and evolution of DNA sequences located at the alpha satellite boundary of chromosome 20. Gene 256, 43–50.

Basu, J., Stromberg, G., Compitello, G., Willard, H.F., Van Bokkelen, G., 2005. Rapid creation of BAC-based human artificial chromosome vectors by transposition with synthetic alpha-satellite arrays. Nucleic Acids Res. 33, 587–596.

Bigot, Y., Hamelin, M.H., Periquet, G., 1990. Heterochromatin condensation and evolution of unique satellite-DNA families in two parasitic wasp species: *Diadromus pulchellus* and *Eupelmus vuilleti* (Hymenoptera). Mol. Biol. Evol. 7, 351–364.

Blattes, R., et al., 2006. Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. EMBO J. 25, 2397–2408.

Bloom, K., 2007. Centromere dynamics. Curr. Opin. Genet. Dev. 17, 151–156.

Bruvo, B., Pons, J., Ugarković, Đ., Juan, C., Petitpierre, E., Plohl, M., 2003. Evolution of low-copy number and major satellite DNA sequences coexisting in two *Pimelia* species-groups (Coleoptera). Gene 312, 85–94.

Canapa, A., Barucca, M., Cerioni, P.N., Olmo, E., 2000. A satellite DNA containing CENP-B box-like motifs is present in the antarctic scallop *Adamussium colbecki*. Gene 247, 175–180.

Cardone, M.F., et al., 2004. Evolution of beta satellite DNA sequences: evidence for duplication-mediated repeat amplification and spreading. Mol. Biol. Evol. 21, 1792–1799.

Carmon, A., Wilkin, M., Hassan, J., Baron, M., MacIntyre, R., 2007. Concerted evolution within the *Drosophila* dumpy gene. Genetics 176, 309–325.

Cavalier-Smith, T., 1985. Selfish DNA and the origin of introns. Nature 315, 283–284.

Cesari, M., Luchetti, A., Passamonti, M., Scali, V., Mantovani, B., 2003. PCR amplification of the *Bag*320 satellite family reveals the ancestral library and past gene conversion events in *Bacillus rossius* (Insecta Phasmatodea). Gene 312, 289–295.

Charlesworth, B., Sniegowski, P.E., Stephan, W., 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 37, 215–220.

Cheng, Z., et al., 2002. Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell 14, 1691–1704.

Coghlan, A., Eichler, E.E., Oliver, S.G., Paterson, A.H., Stein, L., 2005. Chromosome evolution in eukaryotes: a multi-kingdom perspective. Trends Genet. 21, 673–682.

Csink, A.K., Henikoff, S., 1998. Something from nothing: the evolution and utility of satellite repeats. Trends Genet. 14, 200–204.

Dadejová, M., et al., 2007. Transcription activity of rRNA genes correlates with a tendency towards intergenomic homogenization in *Nicotiana* allotetraploids. New Phytol. 174, 658–668.

Dawe, R.K., Henikoff, S., 2006. Centromeres put epigenetics in the driver's seat. Trends Biochem. Sci. 31, 662–669.

Doolittle, W.F., Sapienza, C., 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature 284, 601–603.

Dover, G.A., 1982. Molecular drive: a cohesive mode of species evolution. Nature 299, 111–117.

Dover, G.A., 1986. Molecular drive in multigene families: how biological novelties arise, spread and are assimilated. Trends Genet. 2, 159–165.

Dover, G.A., 1987. DNA turnover and the molecular clock. J. Mol. Evol. 26, 47–58.
Dover, G.A., 2002. Molecular drive. Trends Genet. 18, 587–589.
Dover, G.A., Flavell, R.B., 1984. Molecular coevolution: DNA divergence and the maintenance of function. Cell 38, 623–624.
Drouin, G., 2006. Chromatin diminution in the copepod Mesocyclops edax: diminution of tandemly repeated DNA families from somatic cells. Genome 49, 657–665.
Durajlija-Žinić, S., Ugarković, Đ., Cornudella, L., Plohl, M., 2000. A novel interspersed type of organization of satellite DNAs in *Tribolium madens* heterochromatin. Chromosom. Res. 8, 201–212.
Durfy, S.J., Willard, H.F., 1989. Patterns of intra- and interarray sequence variation in alpha-satellite from the human X chromosome: evidence for short-range homogenization of tandemly repeated DNA sequences. Genomics 5, 810–821.
Eichler, E.E., Clark, R.A., She, X., 2004. An assessment of the sequence gaps: unfinished business in a finished human genome. Nat. Rev., Genet. 5, 345–354.
Elder Jr, J.F., Turner, B.J., 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. Q. Rev. Biol. 70, 297–320.
Ellingsen, A., Slamovits, C.H., Rossi, M.S., 2007. Sequence evolution of the major satellite DNA of the genus *Ctenomys* (Octodontidae, Rodentia). Gene 392, 283–290.
Fitzgerald, D.J., Dryden, G.L., Bronson, E.C., Williams, J.S., Anderson, J.N., 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-alpha satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. Cell 12, 1069–1084.
Fukagawa, T., et al., 2004. Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat. Cell Biol. 6, 784–791.
Garrido-Ramos, M.A., de la Herran, R., Jamilena, R., Lozano, R., Ruiz-Rejon, C., Ruiz-Rejon, M., 1999. Evolution of centromeric satellite DNA and its use in phylogenetic studies of the Sparidae family (Pisces, Perciformes). Mol. Phylogenet. Evol. 12, 200–204.
Gindullis, F., Desel, C., Galasso, I., Schmidt, T., 2001. The large-scale organization of the centromeric region in beta species. Genome Res. 11, 253–265.
Glinka, S., De Lorenzo, D., Stephan, W., 2006. Evidence of gene conversion associated with a selective sweep in *Drosophila melanogaster*. Mol. Biol. Evol. 23, 1869–1878.
Gregory, T.R., et al., 2007. Eukaryotic genome size databases. Nucleic Acids Res. 35, D332–D338.
Grewal, S.I., Elgin, S.C., 2007. Transcription and RNA interference in the formation of heterochromatin. Nature 447, 399–406.
Hall, S.E., Kettler, G., Preuss, D., 2003. Centromere Satellites From *Arabidopsis* populations: maintenance of conserved and variable domains. Genome Res. 13, 195–205.
Hall, A.E., Keith, K.C., Hall, S.E., Copenhaver, G.P., Preuss, D., 2004. The rapidly evolving field of plant centromeres. Curr. Opin. Plant Biol. 7, 108–114.
Hall, S.E., Luo, S., Hall, A.E., Preuss, D., 2005. Differential rates of local and global homogenization in centromere satellites from *Arabidopsis* relatives. Genetics 170, 1913–1927.
Hartmann, N., Scherthan, H., 2004. Characterization of ancestral chromosome fusion points in the Indian muntjac deer. Chromosoma 112, 213–220.
Heikkinen, E., Launonen, V., Muller, E., Bachmann, L., 1995. The pvB370 BamHI satellite DNA family of the *Drosophila virilis* group and its evolutionary relation to mobile dispersed genetic pDv elements. J. Mol. Evol. 41, 604–614.
Heitz, E., 1928. Das Heterochromatin der Moose. I. Jahrb. Wiss. Bot. 69, 762–818.
Henikoff, S., Dalal, Y., 2005. Centromeric chromatin: what makes it unique? Curr. Opin. Genet. Dev. 15, 177–184.
Henikoff, S., Ahmad, K., Malik, H.S., 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293, 1098–1102.
Heslop-Harrison, J.S., Murata, M., Ogura, Y., Schwarzacher, T., Motoyoshi, F., 1999. Polymorphisms and genomic organization of repetitive DNA from centromeric regions of *Arabidopsis* chromosomes. Plant Cell 11, 31–42.
Heslop-Harrison, J.S., Brandes, A., Schwarzacher, T., 2003. Tandemly repeated DNA sequences and centromeric chromosomal regions of *Arabidopsis* species. Chromosom. Res. 11, 241–253.
Hoskins, R.A., et al., 2007. Sequence finishing and mapping of *Drosophila melanogaster* heterochromatin. Science 316, 1625–1628.
Hosouchi, T., Kumekawa, N., Tsuruoka, H., Kotani, H., 2002. Physical map-based sizes of the centromeric regions of *Arabidopsis thaliana* chromosomes 1, 2, and 3. DNA Res. 9, 117–121.
Howe, M., Dimitri, P., Berloco, M., Wakimoto, B.T., 1995. Cis-effects of heterochromatin on heterochromatic and euchromatic gene activity in *Drosophila melanogaster*. Genetics 140, 1033–1045.
Ikeno, M., Masumoto, H., Okazaki, T., 1994. Distribution of CENP-B boxes reflected in CREST centromere antigenic sites on long-range alpha-satellite DNA arrays of human chromosome 21. Hum. Mol. Genet. 3, 1245–1257.
Ito, H., Miura, A., Takashima, K., Kakutani, T., 2007. Ecotype-specific and chromosome-specific expansion of variant centromeric satellites in *Arabidopsis thaliana*. Mol. Genet. Genomics 277, 23–30.
John, B., Miklos, G.L.G., 1979. Functional aspects of satellite DNA and heterochromatin. Int. Rev. Cyt. 58, 1–114.
Kalitsis, P., Griffiths, B., Choo, K.H., 2006. Mouse telocentric sequences reveal a high rate of homogenization and possible role in Robertsonian translocation. Proc. Natl. Acad. Sci. USA 103, 8786–8791.
Kapitonov, V.V., Holmquist, G.P., Jurka, J., 1998. L1 repeat is a basic unit of heterochromatin satellites in cetaceans. Mol. Biol. Evol. 15, 611–612.
Karpen, G., 1994. Position-effect variegation and the new biology of heterochromatin. Curr. Biol. 4, 281–291.
King, L.M., Cummings, M.P., 1997. Satellite DNA repeat sequence variation is low in three species of burying beetles in the genus *Nicrophorus* (Coleoptera: Silphidae). Mol. Biol. Evol. 14, 1088–1095.
Kropotov, A., et al., 2006. Constitutive expression of the human peroxiredoxin V gene contributes to protection of the genome from oxidative DNA lesions and to suppression of transcription of noncoding DNA. FEBS J. 273, 2607–2617.
Kumekawa, N., Hosouchi, T., Tsuruoka, H., Kotani, H., 2000. The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 5. DNA Res. 7, 315–321.
Kumekawa, N., Hosouchi, T., Tsuruoka, H., Kotani, H., 2001. The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 4. DNA Res. 8, 285–290.
Kurek, R., Reugels, A.M., Lammermann, U., Bunemann, H., 2000. Molecular aspects of intron evolution in dynein encoding mega–genes on the heterochromatic Y chromosome of *Drosophila* sp. Genetica 109, 113–123.
Kuznetsova, I.S., Podgornaya, O., Ferguson-Smith, M.A., 2006. High-resolution organization of mouse centromeric and pericentromeric DNA. Cytogenet. Genome Res. 112, 248–255.
Kuznetsova, I.S., et al., 2007. Evidence for the existence of satellite DNA-containing connection between metaphase chromosomes. J. Cell Biochem. 101, 1046–1061.
Lam, A.L., Boivin, C.D., Bonney, C.F., Rudd, M.K., Sullivan, B.A., 2006. Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. Proc. Natl. Acad. Sci. USA 103, 4186–4191.
Lee, H.R., et al., 2005. Chromatin immunoprecipitation cloning reveals rapid evolutionary patterns of centromeric DNA in *Oryza* species. Proc. Natl. Acad. Sci. USA 102, 11793–11798.
Lee, H.R., Neumann, P., Macas, J., Jiang, J., 2006. Transcription and evolutionary dynamics of the centromeric satellite repeat CentO in rice. Mol. Biol. Evol. 23, 2505–2520.
Lorite, P., Renault, S., Rouleux-Bonnin, F., Bigot, S., Periquet, G., Palomeque, T., 2002. Genomic organization and transcription of satellite DNA in the ant *Aphaenogaster subterranea* (Hymenoptera, Formicidae). Genome 45, 609–616.
Lorite, P., Carrillo, J.A., Tinaut, A., Palomeque, T., 2004. Evolutionary dynamics of satellite DNA in species of the genus *Formica* (Hymenoptera, Formicidae). Gene 332, 159–168.
Luchetti, A., et al., 2003. Unisexuality and molecular drive: Bag320 sequence diversity in *Bacillus* Taxa (Insecta Phasmatodea). J. Mol. Evol. 56, 587–596.
Luchetti, A., Marino, A., Scanabissi, F., Mantovani, B., 2004. Genomic dynamics of a low copy number satellite DNA family in *Leptestheria dahalacensis* (Crustacea, Branchiopoda, Conchostraca). Gene 342, 313–320.
Luchetti, A., Marini, M., Mantovani, B., 2006. Non-concerted evolution of the RET76 satellite DNA family in *Reticulitermes* taxa (Insecta, Isoptera). Genetica 128, 123–132.

Ma, J., Jackson, S.A., 2006. Retrotransposon accumulation and satellite amplification mediated by segmental duplication facilitate centromere expansion in rice. Genome Res. 16, 251–259.

Mahtani, M.M., Willard, H.F., 1990. Pulsed-field gel analysis of alpha-satellite DNA at the human X chromosome centromere: high-frequency polymorphisms and array size estimate. Genomics 7, 607–613.

Mahtani, M.M., Willard, H.F., 1998. Physical and genetic mapping of the human X chromosome centromere: repression of recombination. Genome Res. 8, 100–110.

Malik, H.S., Henikoff, S., 2001. Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*. Genetics 157, 1293–1298.

Mantovani, B., 1998. Satellite sequence turnover in parthenogenetic systems: the apomictic triploid hybrid *Bacillus lynceorum* (Insecta, Phasmatodea). Mol. Biol. Evol. 15, 1288–1297.

Martienssen, R.A., 2003. Maintenance of heterochromatin by RNA interference of tandem repeats. Nat. Genet. 35, 213–214.

Martienssen, R.A., Zaratiegui, M., Goto, D.B., 2005. RNA interference and heterochromatin in the fission yeast *Schizosaccharomyces pombe*. Trends Genet 21, 450–456.

Martinez-Balbas, A., et al., 1990. Satellite DNAs contain sequences that induced curvature. Biochemistry 29, 2342–2348.

Mashkova, T., et al., 1998. Unequal crossing-over is involved in human alpha satellite DNA rearrangements on border of the satellite domain. FEBS Lett. 441, 451–457.

Masumoto, H., Nakano, M., Ohzeki, J., 2004. The role of CENP-B and alpha-satellite DNA: de novo assembly and epigenetic maintenance of human centromeres. Chromosom. Res. 12, 543–556.

McAllister, B.F., Werren, J.H., 1999. Evolution of tandemly repeated sequences: what happens at the end of an array? J. Mol. Evol. 48, 469–481.

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.

Meštrović, N., Randig, O., Abad, P., Plohl, M., Castagnone-Sereno, P., 2005. Conserved and variable domains in satellite DNAs of mitotic parthenogenetic root–knot nematode species. Gene 362, 44–50.

Meštrović, N., Castagnone-Sereno, P., Plohl, M., 2006a. High conservation of the differentially amplified MPA2 satellite DNA family in parthenogenetic root–knot nematodes. Gene 376, 260–267.

Meštrović, N., Castagnone-Sereno, P., Plohl, M., 2006b. Interplay of selective pressure and stochastic events directs evolution of the MEL172 satellite DNA library in root–knot nematodes. Mol. Biol. Evol. 23, 2316–2325.

Miller, W.J., Nagel, A., Bachmann, J., Bachmann, L., 2000. Evolutionary dynamics of the SGM transposon family in the *Drosophila obscura* species group. Mol. Biol. Evol. 17, 1597–1609.

Modi, W.S., Ivanov, S., Gallagher, D.S., 2004. Concerted evolution and higher-order repeat structure of the 1.709 (satellite IV) family in bovids. J. Mol. Evol. 58, 460–465.

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., Ugarković, Đ., Franjević, D., Plohl, M., 2005. Long inversely oriented subunits form a complex monomer of *Tribolium brevicornis* satellite DNA. J. Mol. Evol. 60, 513–525.

Nagaki, K., Talbert, P.B., Zhong, C.X., Dawe, R.K., Henikoff, S., Jiang, J., 2003. Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of *Arabidopsis thaliana* centromeres. Genetics 163, 1221–1225.

Nagaki, K., et al., 2004. Sequencing of a rice centromere uncovers active genes. Nat. Genet. 36, 138–145.

Ohno, S., 1972. So much "junk" DNA in our genome. Brookhaven Symp. Biol. 23, 366–370.

Ohta, T., Dover, G.A., 1984. The cohesive population genetics of molecular drive. Genetics 108, 501–521.

Orgel, L.E., Crick, F.H., 1980. Selfish DNA: the ultimate parasite. Nature 284, 604–607.

Pathak, D., et al., 2006. Chromosomal localization, copy number assessment, and transcriptional status of BamHI repeat fractions in water buffalo *Bubalus bubalis*. DNA Cell Biol. 25, 206–214.

Petrović, V., Plohl, M., 2005. Sequence divergence and conservation in organizationally distinct subfamilies of *Donax trunculus* satellite DNA. Gene 362, 37–43.

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, 234–239.

Robles, F., de la Herran, R., Ludwig, A., Ruiz Rejon, C., Ruiz Rejon, M., Garrido-Ramos, M.A., 2004. Evolution of ancient satellite DNAs in sturgeon genomes. Gene 338, 133–142.

Roizes, G., 2006. Human centromeric alphoid domains are periodically homogenized so that they vary substantially between homologues. Mechanism and implications for centromere functioning. Nucleic Acids Res. 34, 1912–1924.

Rojas, A.A., et al., 2000. Hammerhead-mediated processing of satellite pDo500 family transcripts from *Dolichopoda* cave crickets. Nucleic Acids Res. 28, 4037–4043.

Rudd, M.K., Willard, H.F., 2004. Analysis of the centromeric regions of the human genome assembly. Trends Genet. 20, 529–533.

Rudd, M.K., Wray, G.A., Willard, H.F., 2006. The evolutionary dynamics of alpha-satellite. Genome Res. 16, 88–96.

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.

Santos, S., Chaves, R., Adega, F., Bastos, E., Guedes-Pinto, H., 2006. Amplification of the major satellite DNA family (FA-SAT) in a cat fibrosarcoma might be related to chromosomal instability. J. Heredity 97, 114–118.

Schindelhauer, D., Schwarz, T., 2002. Evidence for a fast, intrachromosomal conversion mechanism from mapping of nucleotide variants within a homogeneous alpha-satellite DNA array. Genome Res. 12, 1815–1826.

Schmidt, T., Heslop-Harrison, J.S., 1998. Genomes, genes and junk: the large-scale organization of plant chromosomes. Trends Plant Sci. 3, 195–199.

Schueler, M.G., Higgins, A.W., Rudd, M.K., Gustashaw, K., Willard, H.F., 2001. Genomic and genetic definition of a functional human centromere. Science 294, 109–115.

Schueler, M.G., et al., 2005. Progressive proximal expansion of the primate X chromosome centromere. Proc. Natl. Acad. Sci. USA 102, 10563–10568.

Shiels, C., Coutelle, C., Huxley, C., 1997. Contiguous arrays of satellites 1, 3, and beta form a 1.5-Mb domain on chromosome 22p. Genomics 44, 35–44.

Slamovits, C.H., Cook, J.A., Lessa, E.P., Rossi, M.S., 2001. Recurrent amplifications and deletions of satellite DNA accompanied chromosomal diversification in South American Tuco-tucos (Genus *Ctenomys*, Rodentia: Octodontidae): a phylogenetic approach. Mol. Biol. Evol. 18, 1708–1719.

Smith, G.P., 1976. Evolution of repeated DNA sequences by unequal crossover. Science 191, 528–535.

Smith, C.D., Shu, S., Mungall, C.J., Karpen, G.H., 2007. The Release 5.1 annotation of *Drosophila melanogaster* heterochromatin. Science 316, 1586–1591.

Stanley, H.P., Kasinsky, H.E., Bols, N.C., 1984. Meiotic chromatin diminution in a vertebrate, the holocephalan fish *Hydrolagus collie* (Chondrichthyes, Holocephali). Tissue Cell 16 (2), 203–215.

Stephan, W., 1986. Recombination and the evolution of satellite DNA. Genet. Res. 47 (3), 167–174.

Stephan, W., 1989. Tandem-repetitive noncoding DNA: forms and forces. Mol. Biol. Evol. 6, 198–212.

Stephan, W., Cho, S., 1994. Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. Genetics 136 (1), 333–341.

Strachan, T., Webb, D., Dover, G.A., 1985. Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. EMBO J. 4, 1701–1708.

Sullivan, B.A., Karpen, G.H., 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nat. Struct. Mol. Biol. 11, 1076–1083.

Sullivan, B.A., Blower, M.D., Karpen, G.H., 2001. Determining centromere identity: cyclical stories and forking paths. Nat. Rev., Genet. 2, 584–596.

Sun, X., Le, H.D., Wahlstrom, J.M., Karpen, G.H., 2003. Sequence analysis of a functional *Drosophila* centromere. Genome Res. 13 (2), 182–194.

Szybalski, W., 1968. Use of cesium sulfate for equilibrium density gradient centrifugation. Methods Enzymol. 12B, 330–360.

Talbert, P.B., Bryson, T.D., Henikoff, S., 2004. Adaptive evolution of centromere proteins in plants and animals. J. Biol. 3, 18–69.

Tek, A.L., Song, J., Macas, J., Jiang, J., 2005. Sobo, a recently amplified satellite repeat of potato, and its implications for the origin of tandemly repeated sequences. Genetics 170, 1231–1238.

Topp, C.N., Zhong, C.X., Dawe, R.K., 2004. Centromere-encoded RNAs are integral components of the maize kinetochore. Proc. Natl. Acad. Sci. U. S. A. 101, 15986–15991.

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

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

Ventura, M., et al., 2007. Evolutionary formation of new centromeres in macaque. Science 316, 243–246.

Vershinin, A.V., Alkhimova, E.G., Heslop-Harrison, J.S., 1996. Molecular diversification of tandemly organized DNA sequences and heterochromatic chromosome regions in some Triticeae species. Chromosom. Res. 4, 517–525.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I.S., 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science 297, 1833–1837.

Warburton, P.E., Willard, H.F., 1990. Genomic analysis of sequence variation in tandemly repeated DNA. Evidence for localized homogeneous sequence domains within arrays of alpha-satellite DNA. J. Mol. Biol. 216, 3–16.

Willard, H.F., Waye, J.S., 1987. Chromosome-specific subsets of human alpha satellite DNA: analysis of sequence divergence within and between chromosomal subsets and evidence for an ancestral pentameric repeat. J Mol. Evol. 25, 207–214.

Wong, L.H., Choo, K.H., 2001. Centromere on the move. Genome Res. 11, 513–516.

Wu, J., et al., 2004. Composition and structure of the centromeric region of rice chromosome 8. Plant Cell 16, 967–976.

Yan, H., Jiang, J., 2007. Rice as a model for centromere and heterochromatin research. Chromosom. Res. 15, 77–84.

Zhong, C.X., et al., 2002. Centromeric retroelements and satellites interact with maize kinetochore protein CENH3. Plant Cell 14, 2825–2836.