



Review

Sequence, Chromatin and Evolution of Satellite DNA

Jitendra Thakur ^{1,*}, Jenika Packiaraj ¹ and Steven Henikoff ^{2,3}

¹ Department of Biology, Emory University, Atlanta, GA 30322, USA; jenika.s.packiaraj@emory.edu

² Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA; steveh@fredhutch.org

³ Fred Hutchinson Cancer Research Center, Howard Hughes Medical Institute, Seattle, WA 98109, USA

* Correspondence: jthakur@emory.edu

Abstract: Satellite DNA consists of abundant tandem repeats that play important roles in cellular processes, including chromosome segregation, genome organization and chromosome end protection. Most satellite DNA repeat units are either of nucleosomal length or 5–10 bp long and occupy centromeric, pericentromeric or telomeric regions. Due to high repetitiveness, satellite DNA sequences have largely been absent from genome assemblies. Although few conserved satellite-specific sequence motifs have been identified, DNA curvature, dyad symmetries and inverted repeats are features of various satellite DNAs in several organisms. Satellite DNA sequences are either embedded in highly compact gene-poor heterochromatin or specialized chromatin that is distinct from euchromatin. Nevertheless, some satellite DNAs are transcribed into non-coding RNAs that may play important roles in satellite DNA function. Intriguingly, satellite DNAs are among the most rapidly evolving genomic elements, such that a large fraction is species-specific in most organisms. Here we describe the different classes of satellite DNA sequences, their satellite-specific chromatin features, and how these features may contribute to satellite DNA biology and evolution. We also discuss how the evolution of functional satellite DNA classes may contribute to speciation in plants and animals.

Keywords: repetitive DNA; heterochromatin; centromeres; telomeres; H3K9me3; CENP-A; non-B-form DNA



Citation: Thakur, J.; Packiaraj, J.; Henikoff, S. Sequence, Chromatin and Evolution of Satellite DNA. *Int. J. Mol. Sci.* **2021**, *22*, 4309. <https://doi.org/10.3390/ijms22094309>

Academic Editors: Miroslav Plohl, Eva Šatović and Raquel Chaves

Received: 1 April 2021
Accepted: 17 April 2021
Published: 21 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Eukaryotic DNA sequences are present either as single copies in the genome, or as repeats, which can be either interspersed throughout the genome or locally repeated in tandem. Interspersed repeats mostly comprise transposable elements and their relics, which are scattered throughout the genome, whereas tandem repeats are found mostly at or around centromeres and telomeres. These distinct classes of repetitive DNA were characterized almost two decades before the introduction of Sanger and Maxam–Gilbert DNA sequencing based on analysis of reannealing kinetics [1]. When DNA is denatured and allowed to reanneal, single-copy DNA anneals slowly because the DNA strand complementary to any given strand is present at low concentration, whereas the concentration of the complementary strand doubles when it is present in two copies, and so on. Unlike bacterial genomic DNA, which anneals as a single component, DNA from multicellular eukaryotes, such as mammals, anneals as three components: The slowly annealing single-copy DNA typically accounts for as much as half of the total, with the remainder being middle-repetitive dispersed repeats and highly repetitive tandem repeats. The highly repetitive fraction was also identified by buoyant density gradient centrifugation. A high-concentration aqueous solution of cesium chloride (CsCl) centrifuged at high speed will form a gradient, and where DNA bands in a CsCl gradient at equilibrium is determined by its base composition. The term “satellite DNA” was coined to describe DNA with a base composition different enough from most other DNA in the genome that it forms a satellite

band above (AT-rich DNA) or below (GC-rich DNA) the main band containing the bulk of the genome [1]. Subsequently, restriction enzyme digestion and electrophoresis of the DNA fragments generated on agarose gels followed by Southern blot hybridization using DNA probes were used to clone and characterize repetitive DNA [2–5]. With the advancement of genomics with next-generation sequencing technologies and bioinformatics tools for repeat analysis (e.g., Tandem Repeat Finder [6], RepeatExplorer [7], TAREAN [8] etc.), the cataloging of the repetitive regions has gained momentum, which together with genetic and cell biological approaches has revealed that many satellite families are involved in distinct cellular functions.

2. Satellite DNA Sequences

Satellite DNAs are classified into microsatellites, minisatellites, satellites and macrosatellites based on the monomeric repeat length. Microsatellites, most relevant to medicine, also called simple sequence repeat (SSR) or short tandem repeat (STR), are small (2–6 bp in size) tandem repeats. Microsatellite copy numbers are characteristically variable within a population. The human genome contains ~3% microsatellites [9,10]. Telomeric regions are occupied by microsatellites that can be repeated multiple times, eventually forming the bulk of the telomeric region up to 15 kb on human chromosomes [11]. Due to their polymorphic nature microsatellites are used in gene discovery by linkage analysis for identification purposes in paternity testing or forensic DNA testing [12–14]. Minisatellite monomeric repeat units are larger (~15 bp) but fewer in number than microsatellites. Minisatellite arrays exhibit a mean length of 0.5–30 kb and are found in euchromatic regions of the genome of vertebrates, fungi and plants. Several thousand chromosomal loci are enriched with minisatellites in the human genome [15,16]. Minisatellites are highly variable in array size and classified as variable number of tandem repeats (VNTRs). In the human genome, VNTRs have been used for DNA fingerprinting to type individuals [17]. Each organism has a unique arrangement pattern, the only exception being multiple individuals from a single zygote (e.g., identical twins) [18]. Satellites consist of tandemly repeated arrays that are present at centromeres, pericentromeric regions and subtelomeric regions. Most satellite arrays are characterized by either simple repeats or complex repeats that involve increase in repeat unit length and complexity by merging shorter repeat motifs [19–24]. Macrosatellite monomer repeat units are much larger and range up to a few kilobases in length [25]. Examples of human macrosatellites are D4Z4, a 3.3 kb macrosatellite located at the subtelomeric regions of chromosomes 4q35 and 10q26, DXZ4, a 3 kb CpG-rich macrosatellite is present in 12–100 tandem copies on chromosome Xq23 and NBL2, a 1.4 kb macrosatellite repeat mainly found on the short arm of acrocentric chromosomes 13, 14, 15 and 21 [25,26]. Satellite DNA exhibits a high degree of polymorphism due to variations in the number of repeat units caused by mutations involving several mechanisms [27]. This hypervariability among satellite DNA repeats of related and unrelated organisms makes them excellent markers for mapping, characterizing the genomes, genotype-phenotype correlation, marker-assisted selection of the crop plants, molecular ecology and diversity-related studies [14–16,28].

2.1. Functional Satellite Classes

2.1.1. Centromeric Satellites

Centromeres are the sites of spindle binding and are located within the primary constrictions, the “pinched-waist” of chromosomes. Spindles bound to centromeres pull replicated chromatids/homologous chromosomes apart to mediate faithful chromosome segregation during mitosis and meiosis. Centromere function in chromosome segregation is conserved across eukaryotes, but centromeric satellites are the fastest evolving genomic elements and closely related species exhibit dramatic variations in their centromeric DNA sequences [29,30]. Centromeric sequences in the budding yeast *Saccharomyces cerevisiae* are non-repetitive and are bound by centromeric proteins in a sequence-dependent manner [31]. However, centromeres of most other eukaryotes are occupied by arrays of satellite repeats

that span up to several megabases [2,32,33]. The best studied centromeric satellites are those found in primates, mice, plants and insects (Table 1).

Table 1. Centromeric satellite repeat units of mammals, plants and insects.

Organism	Repeat Unit	Length	References
Human	α -satellite	171-bp	[4]
<i>M. musculus</i>	Minor satellite	120-bp	[34,35]
	MS3	150-bp	
	MS4	300-bp	
<i>Z. mays</i>	CentC	156-bp	[36–38]
	CRM	~7000-bp	
<i>O. sativa</i>	RCS2 or CentO	155-bp	[39]
	CRR	4400- and 9200-bp	
<i>L. nivea</i>	LCS1	178-bp	[40]
<i>A. thaliana</i>	pAL1	178-bp	[41,42]
<i>R. sativus</i>	pRA5/pRB	~177-bp	[43]
<i>D. melanogaster</i>	Prodsat 10-mer AATAACATAG	10-bp	[44,45]
	AATAT	5-bp	
	AATAG	5-bp	
	AAGAT	5-bp	
	G2/Jockey-3	Up to a few kb	
<i>D. simulans</i>	10-mer AATAGAATTG	10-bp	[44,45]
	Simcent1	154-bp	
	Simcent2	88-bp	
	G2/Jockey-3	Up to a few kb	

Primate Centromeric Satellites

Centromeres of primates consist of α -satellites that may constitute as much as 10% of the total number of repeats in primate genomes [46]. α -satellites are the most abundant satellites in the human genome representing half of the total human satellite DNA [10,47]. Individual monomers in the human α -satellite pool are up to 60% divergent and have been classified by Alexandrov into five suprachromosomal groups or families, based on the organization of repeat monomers [19,48]. Suprachromosomal families 1-2 (SF1-2) are tandem arrays of two alternating α -satellite units. Two α -satellite units of a dimer can differ up to ~30%. These dimeric units were independently identified as the most abundant functional centromere associated α -satellites based on their binding to centromere marker proteins in cell lines from individuals belonging to five different populations [49]. Dimeric arrays are present on multiple chromosomes and are highly enriched for centromeric proteins, therefore forming the majority of “functional” α -satellites [49]. One of the two α -satellite units of a dimer contain the CENP-B box, the binding site for Centromere Protein B (CENP-B) [50,51]. The SF3 family contains a higher-order repeat (HOR) organization, which results from the simultaneous amplification and homogenization of a block of more than two α -satellite monomers [19,48,52]. As a result, individual α -satellite repeat units within a HOR are typically more divergent (~70% identical) than adjacent HOR blocks (up to 97–100% sequence identity) [53,54]. Unlike SF1-2 dimers, SF3 α -satellite HORs are mostly chromosome-specific. Centromeric arrays of chromosomes 1, 11, 17 and X contain SF3 HORs in which the repeating unit is a block of five monomers (pentameric) [48]. The HOR with highest number of

repeating units are present on the centromere of chromosome Y (34-mer HOR) [55]. A recent study has shown that even on HORs, the most prominent basic pattern is the dimeric unit with the presence of a CENP-B box at every alternative monomer [56]. The remaining α satellites are present as monomers and do not bind to enough centromeric proteins to form functional centromeres. The HOR organization had been recently found to exist in hominoids as well as the Old and new world monkeys [57–61].

On rare occasions, when a human native satellite centromere is inactivated, centromere function is acquired by an ectopic non-satellite locus called neocentromere [21,62,63]. More than 100 cases of human neocentromeres have been studied and, in most cases, they were found to occur in regions that are gene deserts [64,65]. Neocentromeres can also be experimentally induced by deleting native centromeres [64,65]. In some experimentally tractable organisms, neocentromeres form primarily at locations adjacent to the native centromeres or in the subtelomeric regions [66–69].

Mouse Centromeric Satellites

Mouse chromosomes are telocentric, with centromeric satellite arrays located at or near one end of the chromosome. Despite sharing ~70% sequence identity in their genes, mice and humans carry completely different satellite repeats. *Mus musculus* centromeres consist of tandem arrays of 120-bp minor satellite (MiSat), which contains the CENP-B box and comprise 1–2% of the genome [33,34]. MiSats are more homogeneous within a *Mus* species as compared to α -satellites of primate species [24,33,34,70]. Additional centromeric satellites are the ~150-bp MS3 and the ~300-bp MS4, both of which cytologically co-localize with MiSats at *M. musculus* centromeres [35,71]. The association of MS3 and MS4 with functional centromeric markers remains to be elucidated, but both of these satellites contain CENP-B-like boxes indicating that they might act as functional centromeric satellites [35]. *M. spretus*, which separated from *M. musculus* ~1–2 MYA, has retained a small fraction of MiSats, which are distributed throughout the entire centromeric and pericentric domain, unlike in *M. musculus* where MiSats are exclusively within the centromeric regions [70,72,73]. *M. caroli*, separated from *M. musculus* ~6 MYA, lacks minor satellites and has evolved different families of satellites with monomer lengths of 60 and 79 bp, which localize to the telocentric ends [72,73]. The *M. caroli*-specific 79-bp satellite monomer includes a distant variant of the CENP-B box, which contains all nine bases necessary for its interaction with Cenp-B, and is capable of binding to *M. musculus* Cenp-B in an interspecific cell line [50].

Plant Centromere Satellites

Similar to primate and mouse centromeric satellites, the centromeric repeat units in most flowering plants are AT-rich in composition and approximately nucleosomal in length (Table 1). Centromeric regions of several plant species are colonized by retroelements that belong to the Ty3/gypsy endogenous retrovirus family [30,74–80]. The integrases of plant centromeric retroelements contain a chromodomain, which is thought to ensure correct integration of the retrotransposon to the centromeric region and are collectively referred to as putative centromeric targeting domains [74,75,81]. Maize (*Zea mays*) satellite repeat, CentC, is a 156-bp monomeric repeat unit and is highly interspersed with centromeric retrotransposons of maize (CRM) [36]. Interspersed CentC/CRM arrays range from ~300–3000 kbp [82]. *Z. mays* contain a nonessential selfish B chromosome, which is highly heterochromatic and maintained in populations through a system of non-Mendelian inheritance. The B chromosome contains a neocentromere, the 370-kb Pmel fragment, which is important for maize B chromosome transmission and shares high sequence similarity to repeats in knobs of maize standard chromosomes [37,83,84]. Rice (*Oryza sativa*) centromeres are dominated by a 155-bp rice-specific CentO satellite repeat unit and a centromere-specific retrotransposon, CRR [39,85]. CentO is present in ~6000 copies in the rice genome that share 85–90% sequence identity [39,85]. The holocentric plant *Luzula nivea* and nine other *Luzula* species contain a 178 bp repeat unit, LCS1, which is similar to rice CentO and is potentially centromeric in origin [40,86,87]. *Arabidopsis thaliana* centromeres

consist of tandem arrays of a 178-bp pAL1 repeat unit spanning ~1–3 Mb and comprising ~3% of the genome [32,41]. Other members of the *Arabidopsis* genus contain pAL1-like tandem repeats that are ~60–80% identical [88].

Drosophila Centromeric Satellites

Among nearly 1 million known insect species, satellite DNA has been studied for Diptera, Hymenoptera, Bombyx and Tribolium [30], and functional centromeric satellites have been functionally characterized in *D. melanogaster* and its sister species, *D. simulans* (Table 1). *Drosophila* centromeric satellites differ from satellites that dominate centromeres of most other complex eukaryotes in that repeat units are not of nucleosomal length, but rather are very short, mostly 5–10 bp. In early studies, a region sufficient for centromere function centromeres in *D. melanogaster* was mapped on an X-derived minichromosome, Dp1187, to a 420-kb locus containing the AAGAG and AATAT satellites interspersed with “islands” of complex sequences [89,90]. *D. melanogaster* centromeric satellite repeats are mostly tandem arrays of short 5- to 12-bp sequences, often following the pattern RRNRN, where R is a purine and N is any nucleotide [91]. Two recent studies have characterized *D. melanogaster* and *D. simulans* functional centromeres based on their binding to centromeric protein marker CENP-A [44,45]. Our group found that *D. melanogaster* centromeres are largely occupied by the Prodsat 10-mer AATAACATAG, AATAT, AATAG, and AAGAT repeats and have diverged in their sequences and numbers in *D. simulans*. *D. simulans* centromeric regions are enriched in a simple 10-mer AATAGAATTG repeat and two complex sequence families simcent1 and simcent2 [45]. Both the simcent1 and simcent2 families, contain a local tandem duplication of a core sequence of ~76 bp (simcent1) or 44 bp (simcent2) within the longer complex sequence [45]. Highly abundant Prodsat comprises roughly $5 \pm 2\%$ of the *D. melanogaster* genome but is absent in other species. Chang et al. found that all *D. melanogaster* centromeres are occupied by a non-long terminal repeat (non-LTR) retroelement in the Jockey family, G2/Jockey-3, which is also present at the centromeres of *D. simulans* [44]. It is possible that G2/Jockey-3 has evolved to target centromeres, regions that lack meiotic recombination and so may serve as “safe havens” for parasitic elements. This finding is reminiscent of the recent colonization of maize centromeres by the CR2 retrotransposon [92]. It will be interesting to determine whether the presence of G2/Jockey-3 has acquired any functional role at *Drosophila* centromeres.

2.1.2. Pericentromeric Satellites

Pericentric regions flank centromeres and are bound by cohesin, which prevents premature separation of sister chromatids during chromosome segregation. Pericentric satellites are the most abundant satellites in their genomes and in mouse constitute ~10% of the total. In interphase nuclei, pericentromeric satellites from multiple chromosomes are clustered into chromocenters in several eukaryotes including mouse, flowering plants and *Drosophila* [93–96]. In *D. melanogaster*, defective chromocenter formation results in micronuclei formation due to budding from the interphase nucleus, DNA damage and cell death [96].

Mammalian Pericentric Satellites

Unlike human centromeric regions, which are occupied by a single satellite family (α -satellite), pericentric regions are occupied by several different satellite families (Table 2). Pericentric regions of human chromosomes 3, 4, 9, 13, 14, 15, 21, and 22 contain satellites I, II, and III, which are composed of a mixture of simple short repeated sequences satellites 1 (17 bp, 25 bp), 2 (10–80 bp) and 3 (5 bp, 10 bp), respectively [22,97–103]. A subset of GC-rich β -satellites (Sau3A DNA family) with 68-bp monomer repeat units are also present at the pericentric regions of multiple chromosomes (chromosome 9; acrocentrics—chromosomes 13, 14, 15, 21, and 22) and the secondary constriction of chromosome 1 [23,104,105]. Pericentromeric regions of human chromosomes 8, X, and Y contain tandem arrays of 220-bp GC-rich γ -satellite usually forming 10- to 200-kb clusters flanked by α -satellite DNA [104,106]. Two acrocent-

tric chromosomes can fuse to form a metacentric chromosome through Robertsonian (Rb) translocation in humans [107–112]. Human Rb translocation breakpoints predominantly span the pericentric regions of acrocentric chromosomes 13, 14, and 21 and involve pericentric satellites [108,109,113,114]. The majority of the Rb translocations retain satellite III DNA [109]. Several properties of satellite DNA, such as extensive sequence identity between chromosomes and large uninterrupted arrays, make them ideal for non-homologous recombination and Rb translocations. Similarly, mouse telocentric ends that span pericentric regions are hotspots for Rb translocations in wild populations of *M. musculus* located in Western Europe and North Africa [109,115]. Mouse pericentric regions are occupied by major satellites (MaSats), the TRPC-21A family repeats and TLC (TeLoCentric) repeats [24,34]. MaSat is the most abundant mouse satellite comprising ~10% of the genome and is present at pericentric regions of all chromosomes. The MaSat repeat unit is 234 bp long, is composed of four different 58–60 bp simple repeats, and exhibits monomer variability of up to 30% [24,34,116]. While the MaSat satellite is confined to the pericentric regions in *M. musculus*, it is also present in the chromosomal arms in *M. spretus* and *M. caroli* [20]. The TRPC-21A family repeat, which has a basic 21-bp repeat unit, resembles human pericentric satellites (Sat I, II and III) and is more GC-rich than MiSat and MaSat [24]. The TLC satellite has a 146-bp repeat unit and is located between MiSat and telomeres [24]. The subgenus *Mus* includes 14 species separated into three geographic clades: Southeast Asian, Indian and Palearctic (which includes *M. musculus*). MaSat and MiSat are only found in the Indian and Palearctic clades, whereas the TLC satellite is present in all three clades but absent in some species (*M. famulus*, *M. cypriacus*, and *M. m. musculus*) within each clade [73]. This suggests that TLC must have appeared in the ancestor of the *Mus* subgenus and was then lost independently in some species within each clade. MiSat and MaSat might have arisen and amplified after the divergence of the Southeast Asian clade. This may also suggest that TLC satellites gave rise to MiSat and MaSat sequences before the divergence of the Indian and Palearctic groups, as TLC sequences have higher sequence similarity with MiSat and MaSat than they have between each other [73].

Plant Pericentric Satellites

The majority of flowering plant species in which pericentric satellites have been characterized carry species-specific multiple pericentric satellites similar to those of insects [117–119,129]. *Arabidopsis* pericentric regions are occupied by multiple satellite families including 180-bp repeat family pAL1/pAtMr, Athila retrotransposons, 500-bp and 160-bp tandem repeat families [117–119]. The 500-bp repeats co-localize with centromeres, while the 160-bp repeats localize adjacent to the centromeres. Interestingly, the 500-bp repeat unit is related to pAL1 and proposed to have originated by duplication of one-half of an ancestral pAL1 repeat followed by insertion of a foreign element [117]. *Arabidopsis* pericentric regions are also occupied by four dispersed repeat units, which average 1–2 kb in length. These dispersed repeats are present with copy numbers of 20–300 and are estimated to span up to 1.5 Mb on each side of the *Arabidopsis* centromeres [117]. Pericentric satellites of potato *Solanum bulbocastanum* carry a unique 5.9-kb tandem repeat 2D8 repeat, which is highly homologous to the intergenic spacer (IGS) of the 18S.25S ribosomal RNA genes. Each 2D8 unit is arranged as a dimer of ~3 kb monomers showing some divergence from one another and differing in AT-content. Sequence, structure and copy number of the 2D8 repeat is highly variable throughout the *Solanum* genus, suggesting that it is evolutionarily dynamic [120].

Table 2. Pericentromeric satellite repeat units of mammals, plants and insects.

Organism	Repeat Unit	Monomer Length	Reference
Human	Satellite I	17-bp, 25-bp	[23,62,102,105]
	Satellite II	10–80-bp	
	Satellite III	5-bp, 10-bp	
	α Satellites	68-bp	
	α Satellites	220-bp	
<i>M. musculus</i>	Major satellites	234-bp	[24,34]
	TRPC-21A	21-bp	
	TLC	146-bp	
<i>A.thaliana</i>	pAL1/pAtMr	180-bp	[42,117–119]
	Athila retrotransposons	Up to 14,100-bp	
	500 bp repeat	500-bp	
	160 bp repeat	160-bp	
	Dispersed repeats-163A, 164A, 278A and 106B	1000–2000-bp	
<i>S. bulbocastanum</i>	2D8	5900-bp	[120]
<i>D. melanogaster</i>	1.688 gm/cm ³ buoyant density family	353-bp, 356-bp, 260-bp and 359-bp	[121–123]
	Rsp	120-bp + 120-bp	
	Prod	10-bp	
	AAGAT	5-bp	
	AATAG	5-bp	
	AAGAG	5-bp	
<i>P. subdepressus</i>	T2A5T octanucleotide containing 72-bp-long repeat	8-bp, 22-bp	[124]
<i>C. americana</i>	189 bp repeat	189-bp	[125]
<i>C. carnifex</i>	Complex satellite repeat	211-bp, 477-bp, 633-bp	[126]
<i>T. castaneum</i>	TCAST1	377-bp and 362-bp	[127,128]
	Tcast2a	359-bp	
	Tcast2b	179-bp	
	Tcast2c	300-bp	

Insect Pericentric Satellites

In insects, pericentric satellites have been well-studied in many species (Table 2). Similar to centromeric regions, pericentric regions of *D. melanogaster* contain several different satellite repeats—the 1.688 gm/cm³ buoyant density family of related satellite repeats (353-bp and 356-bp repeats on chromosome 3, a 260-bp repeat on chromosome 2 and a 359-bp repeat on chromosome X), Prod satellite [(AATAACATAG)_n], Responder (Rsp) satellite (a dimer of two 120-bp monomers that share 84% identity), AAGAT, AATAG and AAGAG repeats [121–123]. The 359-bp satellite-containing array is located immediately adjacent to the rDNA locus on the X chromosome and is suggested to play a role in regulating the expression of the rDNA genes [130]. Rsp satellite located on chromosome 2 is

highly variable, ranging from ~10 to over 3000 repeats per array among individuals and can be deleterious in specific genetic backgrounds [131]. Although most Rsp satellites are located in the pericentric regions, a small number of Rsp satellites are also present in euchromatic loci throughout the genome [132]. When the homologous chromosome 2 carries a Segregation Distorter (SD) allele, Rsp satellite causes spermatids bearing it to degenerate causing loss of half of sperms, which results in high transmission frequencies of the SD-carrying chromosome [133]. Retention of Rsp satellite in the *Drosophila* genome despite its harmful effects has been suggested due to its requirement for the organismal fitness as deletion of Rsp satellite array is associated with a reduction in the fitness [134].

The pericentric satellite repeats of *Chrysolina* (leaf beetles) species are highly species-specific (Table 2). *C. carnifex* pericentromeric species-specific complex satellite monomers are organized into three types of repeats: monomers (211-bp) and higher-order repeats in the form of dimers (477-bp) or trimers (633-bp) [126]. *Palorus subdepressus* (tenebrionid beetles) pericentric regions of all 20 chromosomes contain a highly abundant 72-bp satellite that comprises 20% of the genome [124]. The 72-bp *P. subdepressus* pericentric repeat is composed of two copies of T2A5T octanucleotide alternating with 22-bp units of an inverted repeat. Interestingly, phylogenetic clustering of *P. subdepressus* pericentric satellite sequence variants separates them into two clades organized in alternating dimeric 144-bp repeat units [124].

The red flour beetle *Tribolium castaneum* satellite repeats TCAST1 and TCAST2 comprise >35% of the genome and are present at both centromeric as well as pericentromeric regions of all 20 chromosomes [127,135,136]. Interestingly, evolutionarily young variants of TCAST2 have been detected in the natural populations of *T. castaneum*. These populations carry three distinct TCAST2 subfamilies, which differ in monomer size, genome organization, and subfamily specific mutations [128]. Subfamilies Tcast2a (359-bp repeat) and Tcast2b (179-bp repeat) are tandemly arranged within pericentromeric regions, whereas Tcast2c (300-bp repeat) is preferentially dispersed within euchromatic regions. TCAST2 subfamilies show either overrepresentation or almost complete absence of a particular subfamily in some strains [128]. Population-specific *T. castaneum* satellites are proposed to have arisen from homologous recombination, probably stimulated by environmental stress and partial organization of TCAST2 satellite DNA in the form of single repeats dispersed within euchromatin [128,136].

2.1.3. Telomeric and Subtelomeric Satellites

DNA polymerases replicate DNA from 5'-to-3' so cannot by themselves fully replicate the very ends of chromosomes, which can cause telomere shortening at every replication round. In the large majority of eukaryotes, this end-replication problem is prevented by telomerase, which adds de novo telomeric repeats to the G-rich 3'-overhang [137]. A loss of telomere maintenance function leads to telomere shortening, which results in human diseases such as bone marrow failure, cancer and premature aging. In humans, telomerase is active only in germline and in stem cells but not in somatic cells, limiting the number of cell divisions to reduce cancer proliferation and the risk of accumulating mutations that could lead to genomic instability and tumorigenesis. Most telomeric repeats fall into the category of microsatellites, whereas subtelomeric repeats are categorized as satellites. Telomeres consist of short tandem repeats of 5'-TTAGGG-3' in most vertebrates, 5'-TTAGG-3' in most insects and 5'-TTTAGGG-3' in plants [11,138–140]. Telomeres in humans consist of thousands of repeats of 5'-TTAGGG-3' ending in a G-rich 3'-overhang 30–300 nucleotides in length [141]. Telomeric motifs are also present within chromosomes, where they are called interstitial telomeric sequences. In contrast to telomeres from most groups of organisms, *Drosophila* telomeres are maintained without encoding telomerase, but rather have domesticated retrotransposons for this purpose [142]. *Drosophila* telomeres are composed of three subdomains—a cap that distinguishes telomeres from a DNA double-strand break, retrotransposon arrays that maintain telomere length by targeted transposition to telomeres, and proximal telomere-associated sequences (TAS) consisting of a mosaic of complex repeats [143]

Telomeric retrotransposon arrays consist of up to 12 kb of tandem arrays of telomere-specific LINE-like elements: HeT-A, TART and TAHRE [144–146]. Telomeres of lower Diptera, *Anopheles*, *Rhynchosciara*, and *Chironomus* are different from both microsatellite telomeres of most eukaryotes and retrotransposon telomeres of *Drosophila*. Lower Diptera telomeres carry long tandem repeats (satellites) and telomere elongation and maintenance has been suggested to take place by unequal crossover during homologous recombination [147–150].

Subtelomeric sequences vary among eukaryotes and between different chromosomes within a species. Comparative analysis of fully sequenced subtelomeres of human 4p, 16p, 22q, Xq and Yq, has revealed the presence of proximal and distal subtelomeric domains that are separated by a stretch of degenerate TTAGGG repeats [151–153]. While the proximal TAS subtelomeric domain is gene-poor and shows much longer uninterrupted homology to a few chromosome ends, the telomere-distal subtelomeric domain is found at only a few chromosome ends and contains nonessential genes [151,154–156]. Additionally, human subtelomeric domains also show high copy number variation [157].

2.2. Roles of Satellite DNA Sequences

The exact role of most satellite DNA sequences remains unclear in the processes in which they are involved because of their rapid evolution and technical intractability. However, a few protein-binding sequence motifs have been identified in satellite DNA. Another DNA sequence feature is the potential to fold into non-B-form DNA secondary structures, and some satellites are enriched for predicted DNA secondary structures.

2.2.1. DNA Sequence Motifs

Thus far, two protein-binding satellite sequence motifs, CENP-B-box and pJ α direct repeat, have been described for α -satellites in primates. The mammalian CENP-B box is a 17-bp sequence motif (5'-CTTCGTTGGAAACGGGA-3') that is bound by the conserved CENP-B protein both in vitro and in vivo [158–160]. CENP-B protein was evidently domesticated from a pogo-like transposase, which was proposed to have been a source of ancestral genetic variation among mammalian satellites [51,161,162]. Another sequence motif, a 9-bp alphoid-derived direct repeat (GTGAAAAG), is found at the junction of α -satellite and pericentric satellites in chromosomes 13, 14, and 21 and binds to a putative α -satellite DNA binding protein, pJ α [163,164]. Whether pJ α has a role in centromere function is unknown.

2.2.2. DNA Secondary Structures

Predicted DNA secondary structures are highly enriched on centromeric satellites of the mouse and primates (Figure 1) [165,166]. Human centromeric α -satellites and the C-rich strand of *D. melanogaster* dodeca satellite are predicted to form dimeric i-motif structures generated by the association of two parallel DNA duplexes combined in an antiparallel manner in vitro [166,167]. Inverted and palindromic repeats and especially dyad and cruciform structures can potentially act as nucleosome-positioning signals as an alternative to the intrinsic curvature of the DNA and such structures have been predicted in satellites of several insects, the parasitic wasps (*Diadromus* and *Eupelmus*), *Tribolium* sp., *Trichogramma brassicae* and *Chironomus pallidivittatus* [168–172]. The predicted DNA curvature resulting from A-T tracts is a strong feature of the pericentric satellites in several insect species (Figure 1) [125,168,173,174]. The potential role of predicted DNA secondary structures and curvature within satellite DNA is not yet established, but such regions may serve as the binding surface for DNA-bending proteins such as CENP-B [165,175].

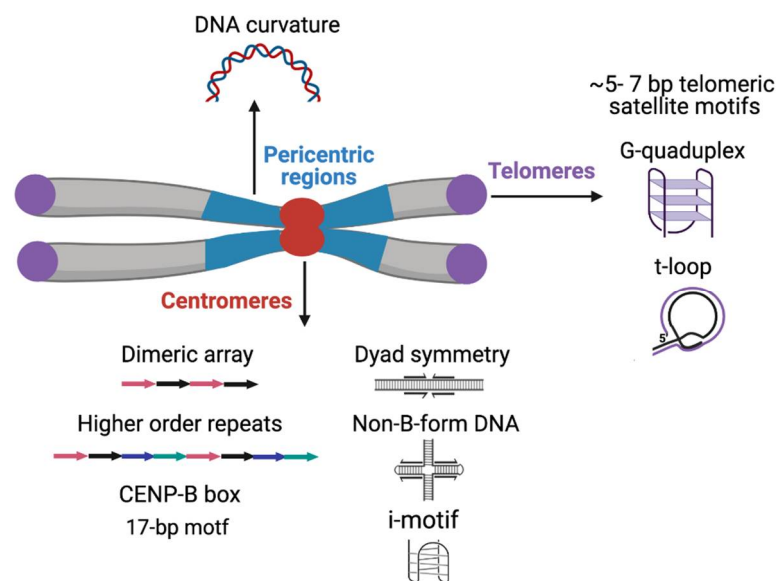


Figure 1. Sequence motifs, basic organizational units and predicted non-B-form secondary DNA structures on various functional classes of satellite DNA.

The G-rich strand at the 3' ends of the chromosome at telomeres is longer than the C strand, forming a 3' G-overhang that is bound by the hexameric shelterin complex [176]. Shelterin regulates telomerase to prevent telomere loss and inhibits DNA repair to avoid telomere ends being mistaken for damaged DNA and activating the DNA damage response [176,177]. The G-overhang allows the formation of a tertiary structure termed the t-loop, where the G-overhang folds back and inserts into the upstream duplex telomeric DNA, displacing the G-rich strand, thereby hiding the end of the chromosome (Figure 1) [178–180]. According to the t-loop end protection model, the shelterin component TRF2 promotes end protection by binding and stabilizing t-loops [181].

Telomeric short tandem repeats of the G-rich strand can form into secondary structures *in vitro*, such as a four-stranded G-quadruplex structure derived by intramolecular folding of the G-rich single-stranded extension (Figure 1). G-quadruplex structures have been identified *in vivo* in ciliates but the direct evidence of G-quadruplex structures in mammalian cells is still lacking [182,183]. Indirect support for the existence of G-quadruplexes at human telomeres has come from experiments with G-quadruplex stabilizing ligands that result in the depletion of telomeric proteins, including shelterin complex components [184–187]. Telomeric G-quadruplexes have been speculated to play a role in processes that include telomere associations, such as the sister chromatid cohesion during DNA replication or telomere clustering into a “bouquet” during meiosis [188–190].

3. Satellite Chromatin

Satellite DNA biology has remained unclear due to the technical intractability of long rapidly evolving tandem repeat arrays, so that the focus of functional studies has largely shifted to satellite chromatin. Based on structure and function, chromatin has been cytologically classified into less compact euchromatin and more compact heterochromatin. Satellite regions are primarily heterochromatic in nature but can be transcribed into non-coding RNA (ncRNA). The nucleosomes of satellite heterochromatin are associated with histone variants, DNA methylation, and histone modifications that prevent transcription. The structure of centromeric chromatin, which marks centromeres, is distinct from both euchromatin and heterochromatin. Heterochromatin and centromeric chromatin are bound by distinct protein components that contribute to the establishment of underlying chromatin states (Figure 2) [191–193].

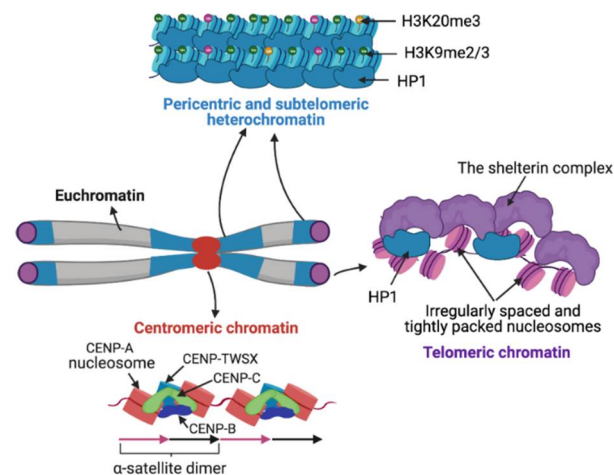


Figure 2. Satellite chromatin organization. In contrast to euchromatin, pericentric and subtelomeric regions are compact, marked by repressive modifications such as H3K9me3 and H3K20me3 and bound by Heterochromatin-associated Protein 1 (HP1). At centromeres, Centromere Protein A (CENP-A) nucleosomes form tight complexes with DNA binding centromeric proteins CENP-B, CENP-C and the histone-fold containing protein complex CENP-TWSX. Telomeric ends contain irregularly spaced and tightly packed nucleosomes separated by 10–20 bp DNA linkers. Telomeric chromatin is tightly associated with the shelterin complex and HP1.

3.1. Centromeric Chromatin

Although centromeric DNA sequences are not evolutionarily conserved, most eukaryotic lineages assemble nucleosomes in which histone H3 is replaced by its (cenH3) variant, Centromeric protein A (CENP-A) [194–196]. CENP-A nucleosomes serve as the chromatin foundation for the kinetochore, an assemblage of multi-protein complexes that make direct connections with growing and shrinking spindle microtubules to mediate chromosome segregation [197,198]. CENP-A is loaded onto centromeres by CENP-A specific chaperone Holiday Junction Recognition Protein (HJURP); however, it is not known how HJURP recognizes centromeric satellite sequences. HJURP was originally described as a protein that binds to cruciform structures representing recombination intermediates, and we have proposed that HJURP recognizes centromeric satellites through their predicted cruciform DNA structures that are enriched on satellite centromeres and neocentromeres [165].

Only a subset of α -satellite arrays on a given chromosome assemble CENP-A chromatin and are functional [49,199]. The majority of CENP-A nucleosomes are assembled on abundant SF1 dimeric α -satellite arrays, where they are precisely and strongly positioned [49]. As these dimeric arrays contain the densest CENP-B boxes, they are expected to bind to high amounts of CENP-B. Indeed, SF1 dimers are associated with high levels of CENP-B that are correlated with high levels of CENP-A assembled on them, suggesting that CENP-B binding enhances CENP-A recruitment [49,196,200]. However, the presence of a CENP-B box is not necessary for the formation of CENP-A chromatin as Y centromeres and neocentromeres that lack CENP-B boxes and CENP-B protein efficiently assemble centromere chromatin [201]. SF2 HOR α -satellite arrays assemble intermediate levels of CENP-A [49]. Centromeres containing these HORs are fully functional, suggesting that CENP-A amounts present on them are sufficient for proper centromere function [199]. Indeed, a single human centromere is shown to contain only \sim 400 CENP-A molecules [202]. Monomeric α -satellites that are present at the centromere edges proximal to pericentric regions contain very low levels of CENP-A, possibly arising from the residual spreading from the core centromeres [49,200,203].

CENP-A nucleosomes are tightly associated with other DNA-bound centromere-specific proteins CENP-B, CENP-C and the CENP-TWSX histone-fold complex to form a coherent assemblage (Figure 2) [203–205], resulting in protection of more than nucleosomal length DNA [200]. CENP-A chromatin is efficiently recovered under high salt conditions, although a small fraction

of CENP-A is found in transiently incorporating intermediate particles that are sensitive to high salt [200,206]. Centromeric satellites are transcribed into ncRNAs that are necessary for centromere function chromosomes segregation [207–209]. Centromeric transcripts are involved in the structural stabilization of the centromeric chromatin [206,208,210]. How centromeric chromatin might affect the transcription of underlying satellites is not yet understood.

The patterns of centromeric DNA methylation are variable across different species and tissues. In mice, DNA methylation levels vary depending on tissue type with somatic cells, sperm and oocytes containing high, intermediate and low DNA methylation levels, respectively [211]. Human and *Arabidopsis* centromeric satellites are significantly methylated, but rice and maize centromeric satellites are hypomethylated [212–215]. A recent study showed that the Japanese killifish medaka (*Oryzias latipes*) contains a specific pattern of hypomethylated and hypermethylated domains in centromeric repeats [216]. Human CENP-B preferentially binds to the unmethylated CENP-B box DNA rather than the methylated form. DNA methylation of the CENP-B box reduces human CENP-B binding and the DNMT inhibitor, 5-aza-2'-deoxycytidine, leads to the redistribution of CENP-B [217]. However, CENP-C, another component of the CENP-A chromatin complex, interacts with the de novo DNA methyltransferase DNMT3B, and both CENP-C and DNMT3B are involved in repressing centromeric transcription [218]. Further work is needed to determine whether or not repression of transcription by DNA methylation in these cases helps to maintain functional centromeres.

3.2. Pericentric Heterochromatin

When protein-coding genes are juxtaposed to pericentric heterochromatin due to genomic rearrangements, they undergo silencing in some somatic cells but not in others, a phenomenon called position-effect variegation (PEV) [219–221]. Human pericentric satellites are silenced in normal cells but are aberrantly expressed in many cancer cells, which coincides with the accumulation of methyl-CpG binding protein 2 at pericentromeres [222,223]. Pericentromeric satellites assemble highly condensed constitutive heterochromatin, which once established during development, remains permanently condensed during interphase. In contrast, facultative heterochromatin is assembled on developmentally regulated loci, which can switch between compacted and decompact chromatin states during gene repression and activation. Pericentric heterochromatin serves to maintain sister chromatid cohesion around centromeres to prevent premature chromatid separation during chromosome segregation [224]. Pericentric heterochromatin is characterized by methylation of Lysine-9 of histone H3 (H3K9me) and trimethylation of Lysine-20 of histone H4 (H4K20me3) catalyzed by Suv29H1 and SUV4-20H1&2 histone methyltransferases, respectively (Figure 2) [225,226]. Pericentric chromocenters can be seen as large round structures upon H3K9me staining [116,227]. The presence of H3K9me3 on pericentric heterochromatin provides binding sites for heterochromatin protein 1 (HP1), which then recruits SUV4-20H which trimethylates H4K20 [225,228,229]. Pericentric satellites are uniformly DNA hypermethylated across different species and tissues, making pericentric chromatin environment repressive [211].

Despite its highly condensed structure, pericentric heterochromatin is transcribed into non-coding RNAs. Mouse MaSats are transcribed from both strands into ncRNAs of heterogeneous length [230]. During embryogenesis, pericentric satellites undergo a transient peak of strand-specific transcription precisely at the time of chromocenter formation during the 2-cell stage. Interference with MaSat transcripts using locked nucleic acid (LNA)-DNA gapmers results in developmental arrest before completion of chromocenter formation [231]. Mouse pericentric satellites are also expressed during spermatogenesis to produce MaSat repeat transcripts that are involved in the recruitment of SUV39H2, which trimethylates H3K9 on the pericentric heterochromatin required for meiotic chromosome segregation [232]. Additionally, transcription from MaSats increases significantly during neuronal differentiation, implying that satellite transcripts might play a role in differentiating post-mitotic neurons [233]. MaSat transcripts remain associated with chromatin

and have a secondary structure that favors RNA:DNA hybrid formation [234]. MaSat transcripts participate in the recruitment to scaffold attachment factor B (SAFB) to pericentric regions [235]. H3K9me3 chromatin signals on pericentric satellite decrease in the absence of structural RNA [236–238]. Together these results suggest that pericentric heterochromatin is permissive to transcription, and the resulting transcripts may act as a structural scaffold to maintain the native condensed state of heterochromatin.

Apart from their function in sister chromatin cohesion, pericentromeric regions are also proposed to be involved in cellular differentiation or reprogramming. Mouse MaSats and human γ -satellites on chromosome 8 contain multiple recognition sites for the zinc finger protein, Ikaros [239]. Ikaros binds directly to both the promoters of many lymphoid-specific genes and pericentromeric satellite DNA in actively dividing mouse hematopoietic cells [239–242]. These results have led to the model that Ikaros facilitates repression of many lymphoid-specific genes by targeting them to pericentric heterochromatin [239–242]. Interestingly human pericentric γ -satellites that carry CTCF and Ikaros binding sites allow transcriptionally permissive chromatin in the nearby transgene to protect it from heterochromatic silencing [243].

3.3. Telomeric and Subtelomeric Heterochromatin

Similar to pericentric regions, reporter genes introduced in proximity to telomeres undergo silencing, suggesting that the telomeric chromatin environment is repressive [244,245]. Telomeric satellites assemble an unusual chromatin structure characterized by irregularly spaced and tightly packed nucleosomes separated by 10–20 bp DNA linkers (Figure 2) [246–248]. In most but not all eukaryotes, telomeric repeats (5–8 bp) are unfavorable for nucleosome formation, being out of phase with the 10-bp DNA double-helical periodicity that accommodates regular arginine minor groove contacts around the histone octamer surface [249,250]. Telomeric and subtelomeric regions are characterized by repressive heterochromatic marks H3K9me3, H3K79me2, H4K20me3 and HP1 binding (Figure 2) [251,252]. The telomere capping complex called the shelterin complex is tightly bound to telomeric repeat DNA and contributes to protecting of telomeric ends by compacting telomeric chromatin [253]. In *Drosophila* HP1, encoded by Su(var)205, is part of the telomere capping complex, called terminin, which includes rapidly-evolving telomere-specific proteins that prevent telomere fusion [254]. HP1 is recruited to telomeres independently of the presence of H3K9me3 and spreads into neighboring retroelements arrays where HP1 chromodomain binds H3K9me3 [255,256]. The loss of heterochromatic marks at telomeres and subtelomeric regions are linked to elongated telomeres [251,257]. Additionally, Retinoblastoma (RB) family proteins are also involved in promoting H4K20me3 and DNA methylation levels at both telomeres and pericentromeres by promoting the recruitment of Suv4-20h HMTase and HP1 to these loci [258]. Lack of the Rb family proteins results in decreased H4K20 tri-methylation, which coincides with chromosome segregation defects and telomere elongation [258,259]. Telomeric sequences lack CG dinucleotides, the known substrates for mammalian DNMTs, and therefore telomeres lack DNA methylation. However, subtelomeric sequences are heavily methylated and are proposed to contribute to telomeric position-effect silencing and regulation of homologous recombination and telomere length [245,260–262]. A decrease in subtelomeric DNA methylation leads to extreme telomere elongation [261,262]. Although deficiency of DNMTs does not alter the levels of histone methylation at telomeres, it causes a dramatic increase in telomere recombination and telomere length [262].

Telomeric and subtelomeric sequences in vertebrates are transcribed by RNA polymerase II into UUAGGG-repeat-containing RNAs (TERRAs or TelRNA), which act as structural and regulatory components of telomeres [263,264]. TERRAs are heterogeneous in length (100–9000 bp) and localize to telomeric chromatin in cis, a feature reported earlier for the XIST ncRNA, which mediates mammalian X-chromosome inactivation [264,265]. The TERRA levels increase when HMTases Suv39h and Suv4-20h are depleted, suggesting that the heterochromatin silences telomeric repeats [263,266]. Telomeric RNAs have been proposed to control telomere structure as well as telomere elongation by telom-

erase [264,267]. Interestingly, G-rich human TERRA form G-quadruplexes that might inhibit telomerase activity, similar to telomeric DNA [268–271].

4. Satellite Evolution

Satellite DNA sequences undergo high evolutionary turnover involving both stochastic events and selective pressures [30,52]. The high rate of satellite evolution produces species-specific satellite families that can be used for species identification and inferring phylogeny [19,191,261,272–275]. Several molecular mechanisms underlie satellite sequence complexity, homogeneity, array size variability, and rapid evolution (Figure 3).

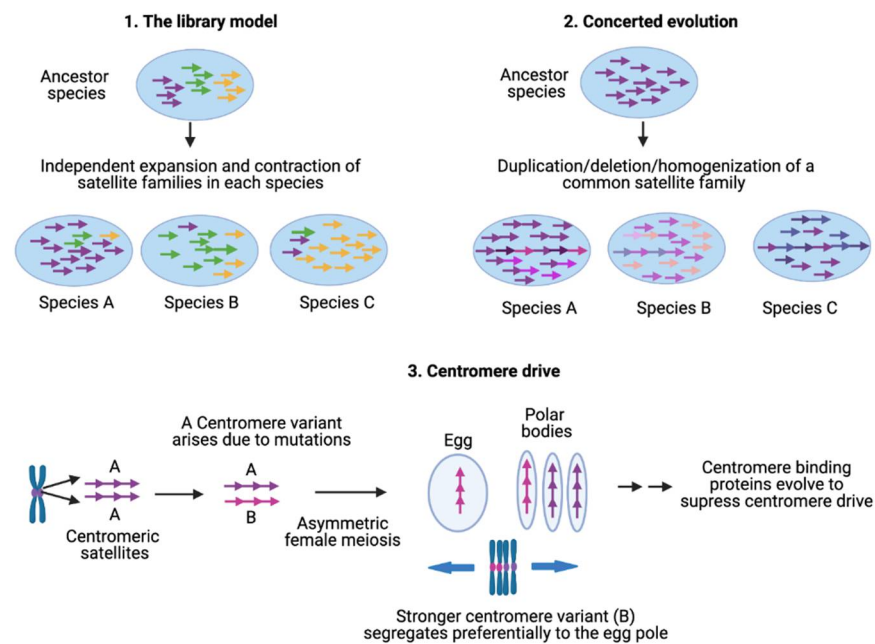


Figure 3. Satellite evolution. The library model can account for the divergent evolution of species-specific satellite families through independent expansion and contraction of satellite families in diverging species. Concerted evolution occurs via recombination processes that result in expansions, contractions and homogenizations of tandemly repetitive sequences. Centromere drive is a selfish process in which stronger centromeres preferentially orient towards the egg pole during asymmetric female meiosis. Drive is suppressed when mutation of a centromere-binding protein restores centromere balance, resulting in an arms race between centromeric satellite DNA and centromere proteins.

4.1. Library Model for Satellite DNA Evolution

According to the library hypothesis of satellite DNA evolution, closely related species share a set of conserved satellite DNA families that originated in the ancestor, but each of which is differentially amplified in each species [276,277]. Amplification of a given satellite DNA family in one species will make it appear as a low-copy satellite in sister species. A simple amplification event will allow higher monomers in the related species to remain significantly conserved. Interspecific sequence conservation and absence of species-specific mutations was first demonstrated in four *Palorus* congeneric species separated by long evolutionary periods of up to 60 million years [276]. Each of these *Palorus* species contains a single pericentromeric AT-rich satellite DNA on all chromosomes, accounting for 20–40% of the genome. All four sequences are present in each of these *Palorus* species and show high conservation of sequences, repeat length and organization. In a given *Palorus* species, one of the four satellites is amplified while the three others are present as low-copy-number repeats accounting for ~0.05% of the genome. The low-copy-number satellites are dispersed between the large arrays of the major satellite over the whole heterochromatic block. The library model has since been confirmed in plants, mammals, nematodes and

insects [173,278–280]. Satellite generation and amplification are proposed to result from the activity of transposable elements and replication of extrachromosomal tandem repeat circles by rolling-circle replication and reinsertion into the genome by unequal crossing over [281–285].

4.2. Concerted Evolution of Satellites

Given that satellite DNAs are among the fastest evolving DNA, different species acquire species-specific satellite families. The amount of satellite DNA as a fraction of the total genome size also varies drastically among closely related species. For example, *Drosophila simulans* has 5% of the genome composed of satellite DNA while only the 0.5% of the *D. erecta* genome is satellite DNA [286]. However, certain satellite families are shared by several species in a genus e.g., α -satellites in primates. In the absence of selective forces, once a given sequence has been amplified into multiple copies in the genome, each copy can incorporate mutations independently and become diverged from each other, leading to a high divergence within a species. However, members of a satellite family show a low divergence rate because satellite repeats undergo concerted evolution in which monomers of a satellite family are homogenized within a species. The outcome of this homogenization is either identical tandem repeats as in mouse telomeres or dimers or higher-order repeats as in human centromeric regions. A few groups of human chromosomes are undergoing concerted evolution as they share large arrays of satellites (1, 5 & 19, 13 & 21, 14 & 22) [287,288].

Several models have been proposed to explain the evolution of satellite DNAs. Crow and Kimura proposed a population model assuming that stabilizing selection would operate to maintain copy numbers [289]. According to their model, the distribution of repeat numbers per genome should reach an equilibrium point stabilized by unequal crossing-over and stabilizing selection, while the selection for optimal array size could be the major player of the evolutionary outcome [289]. A frequently cited model for satellite organization and evolution was developed by Smith using computer-generated simulations [290,291]. Smith argued that DNA sequences that are not constrained by selection have the natural tendency to become repetitive through unequal crossover processes. Smith simulated the effects of repeated out-of-register crossover between sister chromatids over an initially nonrepetitive sequence and showed that the satellite periodicity was the inevitable outcome. The Smith model was further elaborated by Perelson and Bell by including the effects of a variable mutation rate relative to the fixation time of a repeat variant [292]. They found that repeating units become heterogeneous if the mutation rate exceeds the rate of spread of a variant through an array. However, if the fixation rate exceeds the mutation rate, the resulting repeat arrays are homogenous due to the higher likelihood of stochastic removal of mutations before amplification [292]. The Smith model was further elaborated by Stephan (1989) and Stephan and Cho (1994), who showed that for a given sequence length, the out-of-register recombination to generate long repeat arrays is allowed if the mitotic recombination rate between sister chromatids is higher than the base pair mutation rate [293,294].

Dover and colleagues argued that concerted evolution occurs via molecular drive, an evolutionary process emerging from the activities of a number of ubiquitous mechanisms of DNA turnover, such as gene conversion, unequal crossing over, replication slippage, rolling circle replication, and multiple TE insertions [295–297]. A specific molecular drive mechanism based on known properties of centromeric chromatin was recently proposed by Rice, who argued that collapse of the replication fork when it encounters CENP-A nucleosome-associated complexes and/or tightly bound CENP-B results in DNA breaks that are repaired by out-of-register reinitiations [298]. Such break-induced recombination (BIR) had been described for expansions and contractions of rDNA arrays, induced by fork collisions with RNA Polymerases [299]. At centromeres, BIR may also lead to ratcheting up of copy number as CENP-A complexes are assembled onto newly replicated DNA, increasing the probability of further collisions, breaks and reinitiations. It is also possible that both

BIR and unequal reciprocal recombination contribute to satellite evolution, for example, expansions by BIR and deletions by unequal exchange.

4.3. Centromere Drive

Whereas molecular drive might account for satellite DNA evolution under conditions of relaxed selection, centromeres are under intense selection, in that loss of a single centromere or of a centromere protein results in death of the cell. The centromere drive model was introduced to explain the rapid evolution of both centromeric DNA sequences and essential centromere proteins [29,300]. During female meiosis, only one of the four gametic products will be included in the egg, whereas the other three products will be lost in the polar body. This process of meiotic drive was originally envisioned to explain selfish non-Mendelian segregation of heterochromatic maize knobs [301], and was applied to centromeres by proposing competition for the egg pole by centromere-directed reorientation of the meiosis I tetrad. For example, variants of centromeric satellite arrays that enhance the spindle binding during chromosome segregation in meiosis have been observed to selfishly drive their own transmission in this fashion [302–304]. It is also possible that specific satellite sequence variants that increase the efficiency of proper spindle attachment increase their frequency in the population. For example, human centromeric dimeric arrays might be expanding due to their more efficient recruitment of centromeric protein shown by centromeric chromatin profiling [49,196,200,203]. Analogous to centromeric satellites, subtelomeric satellites of different organisms show no sequence similarity and have been proposed to be subjected to telomere drive in which rapid sequence changes in the subtelomeric DNA can trigger adaptation of telomeric proteins to restore telomere homeostasis [305–307].

4.4. Repeat Evolution and Speciation

Bateson–Dobzhansky–Muller incompatibilities between rapidly evolving gene pairs contribute to the hybrid failure (sterility/inviability) and speciation [308]. For example, *D. simulans* Lethal hybrid rescue (Lhr) and *D. melanogaster* Hybrid male rescue (Hmr) have functionally diverged and interact with each other to cause lethality in F1 hybrid males [309]. LHR is adaptively evolving and localizes to pericentric regions, suggesting that rapidly evolving heterochromatic satellite DNA sequences may be driving the evolution of the incompatibility genes [309,310]. Hmr and Lhr encode proteins that form a heterochromatic complex with HP1 and repress transcripts from satellite DNAs and many families of transposable elements [310]. Hmr and Lhr mutations also cause massive overexpression of telomeric TEs and significant telomere lengthening. Hmr and Lhr regulate three major types of heterochromatic sequences responsible for the significant differences between *D. melanogaster* and *D. simulans* and thus have high potential to cause genetic conflicts with host fitness. Adaptive divergence of heterochromatin proteins in response to satellite DNA divergence might be an important underlying force driving the evolution of hybrid incompatibility genes. Direct evidence for the involvement of satellite DNA in speciation came from study where *D. melanogaster* locus Zhr, which maps to the pericentric 359-bp satellite array on the X chromosomes, caused lethality in F1 daughters from crosses between *D. simulans* females and *D. melanogaster* males [311]. Hybrid females show mitotic defects due to lagging X chromatids during early embryogenesis. A deletion of the entire 359-bp satellite array leads to the normal segregation of X chromosomes, while translocation of the 359-bp satellite array to the Y chromosome results in Y chromosome segregation defects. The 359-bp satellite array is absent in *D. simulans*, suggesting that the absence or divergence of factors for required for species-specific heterochromatin formation in the *D. simulans* maternal cytoplasm causing F1 female lethality.

5. Conclusions and Future Perspective

Satellite DNA repeats are among the most diverse and versatile genomic sequence classes contributing to several key biological processes, and satellite deregulation is seen in

various cancers and other diseases. Satellites are hypertranscribed in lung, kidney, ovary, colon, and prostate cancers, possibly due to loss of silencing and hypomethylation of satellite DNA being associated with Wilms tumors, mammary adenocarcinomas, ovarian epithelial carcinomas and immunodeficiency-centromeric instability-facial anomalies (ICF) syndrome [312–316]. However, satellites remain less studied than non-repetitive regions because their highly repetitive nature makes their genetic manipulation challenging and rapid satellite evolution makes identification of conserved sequence motifs difficult. Furthermore, linear assembly maps for satellite regions and the knowledge of the full satellite complement of insect, mammalian and plant species remains incomplete. However, with the improvement of long-read sequencing technology, it has recently become possible to obtain end-to-end maps of the human X chromosome that span centromeric and telomeric regions [317]. Additionally, using the long-read sequencing, the centromeric satellites have been assembled for chromosome 8 and utilized for evolutionary analysis of α -satellite arrays [318]. The rapid improvement in long-read sequencing technology should soon reveal the precise arrangement of satellite blocks, which will provide evidence of the rearrangements that have occurred during their emergence and evolution. Transcription of satellite chromatin, whether heterochromatin or specialized telomeric chromatin, is poorly understood relative to transcribed euchromatin, but advances in understanding gene transcriptional regulation should be applicable to satellites. With improving chromatin extraction and genomic profiling technologies, it is now possible to determine the conformation of satellite chromatin complexes at base-pair resolution [200,203]. We also look forward to better understanding of satellite recognition and determining the degree to which DNA secondary structures are involved in satellite biology at both centromeres and telomeres.

Funding: J.T. is supported by the Department of Biology, Emory University start-up funds. S.H. is an Investigator of the Howard Hughes Medical Institute.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kit, S. Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. *J. Mol. Biol.* **1961**, *3*, 711–716. [[CrossRef](#)]
2. Rudd, M.K.; Willard, H.F. Analysis of the centromeric regions of the human genome assembly. *Trends Genet.* **2004**, *20*, 529–533. [[CrossRef](#)] [[PubMed](#)]
3. Willard, H.F. Chromosome-specific organization of human alpha satellite DNA. *Am. J. Hum. Genet.* **1985**, *37*, 524–532.
4. Schueler, M.G.; Higgins, A.W.; Rudd, M.K.; Gustashaw, K.; Willard, H.F. Genomic and genetic definition of a functional human centromere. *Science* **2001**, *294*, 109–115. [[CrossRef](#)] [[PubMed](#)]
5. Rudd, M.; Schueler, M.; Willard, H. Sequence Organization and Functional Annotation of Human Centromeres. *Cold Spring Harb. Symp. Quant. Biol.* **2003**, *68*, 141–150. [[CrossRef](#)] [[PubMed](#)]
6. Benson, G. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **1999**, *27*, 573–580. [[CrossRef](#)] [[PubMed](#)]
7. Novák, P.; Neumann, P.; Pech, J.; Steinhaisl, J.; Macas, J. RepeatExplorer: A Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* **2013**, *29*, 792–793. [[CrossRef](#)] [[PubMed](#)]
8. Novák, P.; Ávila Robledillo, L.; Koblížková, A.; Vrbová, I.; Neumann, P.; Macas, J. TAREAN: A computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Res.* **2017**, *45*, e111. [[CrossRef](#)]
9. Subramanian, S.; Mishra, R.K.; Singh, L. Genome-wide analysis of microsatellite repeats in humans: Their abundance and density in specific genomic regions. *Genome Biol.* **2003**, *4*, R13. [[CrossRef](#)]
10. International Human Genome Sequencing Consortium Initial sequencing and analysis of the human genome. *Nat. Cell Biol.* **2001**, *409*, 860–921. [[CrossRef](#)]
11. Gomes, N.M.; Shay, J.W.; Wright, W.E. Telomere biology in Metazoa. *FEBS Lett.* **2010**, *584*, 3741–3751. [[CrossRef](#)] [[PubMed](#)]
12. Kashi, Y.; King, D.; Soller, M. Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet.* **1997**, *13*, 74–78. [[CrossRef](#)]
13. Hearne, C.M.; Ghosh, S.; Todd, J.A. Microsatellites for linkage analysis of genetic traits. *Trends Genet.* **1992**, *8*, 288–294. [[CrossRef](#)]
14. Zietkiewicz, E.; Rafalski, A.; Labuda, D. Genome Fingerprinting by Simple Sequence Repeat (SSR)-Anchored Polymerase Chain Reaction Amplification. *Genomics* **1994**, *20*, 176–183. [[CrossRef](#)] [[PubMed](#)]

15. Ramel, C. Mini- and microsatellites. *Environ. Health Perspect.* **1997**, *105*, 781–789. [[CrossRef](#)] [[PubMed](#)]
16. Näslund, K.; Saetre, P.; Von Salomé, J.; Bergström, T.F.; Jareborg, N.; Jazin, E. Genome-wide prediction of human VNTRs. *Genomics* **2005**, *85*, 24–35. [[CrossRef](#)] [[PubMed](#)]
17. Jeffreys, A.J.; Wilson, V.; Thein, S.L. Individual-specific ‘fingerprints’ of human DNA. *Nat. Cell Biol.* **1985**, *316*, 76–79. [[CrossRef](#)]
18. Azuma, C.; Kamiura, S.; Nobunaga, T.; Negoro, T.; Saji, F.; Tanizawa, O. Zygosity determination of multiple pregnancy by deoxyribonucleic acid fingerprints. *Am. J. Obstet. Gynecol.* **1989**, *160*, 734–736. [[CrossRef](#)]
19. Alexandrov, A.; Kazakov, A.; Tumeneva, I.; Shepelev, V.; Yurov, Y. Alpha-satellite DNA of primates: Old and new families. *Chromosoma* **2001**, *110*, 253–266. [[CrossRef](#)]
20. Wong, A.K.C.; Biddle, F.G.; Rattner, J.B. The chromosomal distribution of the major and minor satellite is not conserved in the genus *Mus*. *Chromosoma* **1990**, *99*, 190–195. [[CrossRef](#)]
21. Voullaire, L.E.; Slater, H.R.; Petrovic, V.; Choo, K.H. A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: Activation of a latent centromere? *Am. J. Hum. Genet.* **1993**, *52*, 1153–1163. [[PubMed](#)]
22. Choo, K.; Earle, E.; McQuillan, C. A homologous subfamily of satellite III DNA on human chromosomes 14 and 22. *Nucleic Acids Res.* **1990**, *18*, 5641–5648. [[CrossRef](#)]
23. Waye, J.S.; Willard, H.F. Human beta satellite DNA: Genomic organization and sequence definition of a class of highly repetitive tandem DNA. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 6250–6254. [[CrossRef](#)] [[PubMed](#)]
24. Komissarov, A.S.; Gavrilo, E.V.; Demin, S.J.; Ishov, A.M.; Podgornaya, O.I. Tandemly repeated DNA families in the mouse genome. *BMC Genom.* **2011**, *12*, 531. [[CrossRef](#)] [[PubMed](#)]
25. Dumbovic, G.; Forcales, S.-V.; Perucho, M. Emerging roles of macrosatellite repeats in genome organization and disease development. *Epigenetics* **2017**, *12*, 515–526. [[CrossRef](#)]
26. Chadwick, B.P. Macrosatellite epigenetics: The two faces of DXZ4 and D4Z4. *Chromosoma* **2009**, *118*, 675–681. [[CrossRef](#)]
27. Tautz, D.; Renz, M. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Res.* **1984**, *12*, 4127–4138. [[CrossRef](#)]
28. Kayser, M.; Schneider, P.M. DNA-based prediction of human externally visible characteristics in forensics: Motivations, scientific challenges, and ethical considerations. *Forensic Sci. Int. Genet.* **2009**, *3*, 154–161. [[CrossRef](#)]
29. Henikoff, S.; Ahmad, K.; Malik, H.S. The Centromere Paradox: Stable Inheritance with Rapidly Evolving DNA. *Science* **2001**, *293*, 1098–1102. [[CrossRef](#)]
30. Melters, D.P.; Bradnam, K.R.; Young, H.A.; Telis, N.; May, M.R.; Ruby, J.G.; Sebra, R.; Peluso, P.; Eid, J.; Rank, D.; et al. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol.* **2013**, *14*, R10. [[CrossRef](#)]
31. Fitzgerald-Hayes, M.; Clarke, L.; Carbon, J. Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. *Cell* **1982**, *29*, 235–244. [[CrossRef](#)]
32. Round, E.K.; Flowers, S.K.; Richards, E.J. *Arabidopsis thaliana* Centromere Regions: Genetic Map Positions and Repetitive DNA Structure. *Genome Res.* **1997**, *7*, 1045–1053. [[CrossRef](#)] [[PubMed](#)]
33. Kipling, D.; Ackford, H.E.; Taylor, B.A.; Cooke, H.J. Mouse minor satellite DNA genetically maps to the centromere and is physically linked to the proximal telomere. *Genomics* **1991**, *11*, 235–241. [[CrossRef](#)]
34. Komissarov, A.S.; Kuznetsova, I.S.; Podgornaia, O.I. Mouse centromeric tandem repeats In Silico and In Situ. *Genetika* **2010**, *46*, 1217–1221. [[CrossRef](#)] [[PubMed](#)]
35. Kuznetsova, I.S.; Prusov, A.N.; Erukashvily, N.I.; Podgornaya, O.I. New types of mouse centromeric satellite DNAs. *Chromosom. Res.* **2005**, *13*, 9–25. [[CrossRef](#)] [[PubMed](#)]
36. Ananiev, E.V.; Phillips, R.L.; Rines, H.W. Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13073–13078. [[CrossRef](#)]
37. Ananiev, E.V.; Phillips, R.L.; Rines, H.W. Complex structure of knobs and centromeric regions in maize chromosomes. *TSitologia Genet.* **2000**, *34*, 11–15.
38. Zhong, C.X.; Marshall, J.B.; Topp, C.; Mroczek, R.; Kato, A.; Nagaki, K.; Birchler, J.A.; Jiang, J.; Dawe, R.K. Centromeric Retroelements and Satellites Interact with Maize Kinetochores Protein CENH3. *Plant Cell* **2002**, *14*, 2825–2836. [[CrossRef](#)]
39. Cheng, Z.; Dong, F.; Langdon, T.; Ouyang, S.; Buell, C.R.; Gu, M.; Blattner, F.R.; Jiang, J. Functional Rice Centromeres Are Marked by a Satellite Repeat and a Centromere-Specific Retrotransposon. *Plant Cell* **2002**, *14*, 1691–1704. [[CrossRef](#)]
40. Haizel, T.; Lim, Y.; Leitch, A.R.; Moore, G. Molecular analysis of holocentric centromeres of *Luzula* species. *Cytogenet. Genome Res.* **2005**, *109*, 134–143. [[CrossRef](#)]
41. Murata, M.; Ogura, Y.; Motoyoshi, F. Centromeric repetitive sequences in *Arabidopsis thaliana*. *Jpn. J. Genet.* **1994**, *69*, 361–370. [[CrossRef](#)]
42. Simoens, C.; Gielen, J.; Van Montagu, M.; Inzé, D. Characterization of highly repetitive sequences of *Arabidopsis thaliana*. *Nucleic Acids Res.* **1988**, *16*, 6753–6766. [[CrossRef](#)]
43. Grellet, F.; Delcasso, D.; Panabières, F.; Delseny, M. Organization and evolution of a higher plant alphoid-like satellite DNA sequence. *J. Mol. Biol.* **1986**, *187*, 495–507. [[CrossRef](#)]
44. Chang, C.-H.; Chavan, A.; Palladino, J.; Wei, X.; Martins, N.M.C.; Santinello, B.; Chen, C.-C.; Erceg, J.; Beliveau, B.J.; Wu, C.-T.; et al. Islands of retroelements are major components of *Drosophila* centromeres. *PLoS Biol.* **2019**, *17*, e3000241. [[CrossRef](#)] [[PubMed](#)]

45. Talbert, P.B.; Kasinathan, S.; Henikoff, S. Simple and Complex Centromeric Satellites in *Drosophila Sibling Species*. *Genetics* **2018**, *208*, 977–990. [[CrossRef](#)]
46. Aldrup-MacDonald, M.E.; Sullivan, B.A. The Past, Present, and Future of Human Centromere Genomics. *Genes* **2014**, *5*, 33–50. [[CrossRef](#)] [[PubMed](#)]
47. Levy, S.; Sutton, G.; Ng, P.C.; Feuk, L.; Halpern, A.L.; Walenz, B.P.; Axelrod, N.; Huang, J.; Kirkness, E.F.; Denisov, G.; et al. The Diploid Genome Sequence of an Individual Human. *PLoS Biol.* **2007**, *5*, e254. [[CrossRef](#)] [[PubMed](#)]
48. Alexandrov, I.A.; Medvedev, L.; Mashkova, T.; Kisselev, L.; Romanova, L.; Yurov, Y.; Aiexandrov, I. Definition of a new alpha satellite suprachromosomal family characterized by monomeric organization. *Nucleic Acids Res.* **1993**, *21*, 2209–2215. [[CrossRef](#)] [[PubMed](#)]
49. Henikoff, J.G.; Thakur, J.; Kasinathan, S.; Henikoff, S. A unique chromatin complex occupies young α -satellite arrays of human centromeres. *Sci. Adv.* **2015**, *1*, e1400234. [[CrossRef](#)]
50. Kipling, D.; Mitchell, A.R.; Masumoto, H.; E Wilson, H.; Nicol, L.; Cooke, H.J. CENP-B binds a novel centromeric sequence in the Asian mouse *Mus caroli*. *Mol. Cell. Biol.* **1995**, *15*, 4009–4020. [[CrossRef](#)]
51. Kipling, D.; Warburton, P.E. Centromeres, CENP-B and Tigger too. *Trends Genet.* **1997**, *13*, 141–145. [[CrossRef](#)]
52. Plohl, M.; Meštrović, N.; Mravinac, B. Satellite DNA Evolution. *Genome Dyn.* **2012**, *7*, 126–152. [[CrossRef](#)]
53. Wayne, J.S.; Durfy, S.J.; Pinkel, D.; Kenwrick, S.; Patterson, M.; Davies, K.E.; Willard, H.F. Chromosome-specific alpha satellite DNA from human chromosome 1: Hierarchical structure and genomic organization of a polymorphic domain spanning several hundred kilobase pairs of centromeric DNA. *Genomics* **1987**, *1*, 43–51. [[CrossRef](#)]
54. Rudd, M.K.; Wray, G.A.; Willard, H.F. The evolutionary dynamics of α -satellite. *Genome Res.* **2005**, *16*, 88–96. [[CrossRef](#)] [[PubMed](#)]
55. Tyler-Smith, C.; Brown, W.R. Structure of the major block of alphoid satellite DNA on the human Y chromosome. *J. Mol. Biol.* **1987**, *195*, 457–470. [[CrossRef](#)]
56. Rice, W.R. A Game of Thrones at Human Centromeres I. Multifarious structure necessitates a new molecular/evolutionary model. *bioRxiv* **2020**. [[CrossRef](#)]
57. Pike, L.M.; Carlisle, A.; Newell, C.; Hong, S.-B.; Musich, P.R. Sequence and evolution of rhesus monkey alphoid DNA. *J. Mol. Evol.* **1986**, *23*, 127–137. [[CrossRef](#)]
58. Alves, G.; Seuánez, H.N.; Fanning, T. Alpha satellite DNA in neotropical primates (Platyrrhini). *Chromosoma* **1994**, *103*, 262–267. [[CrossRef](#)] [[PubMed](#)]
59. Sujjwattanarat, P.; Thapana, W.; Srikulnath, K.; Hirai, Y.; Hirai, H.; Koga, A. Higher-order repeat structure in alpha satellite DNA occurs in New World monkeys and is not confined to hominoids. *Sci. Rep.* **2015**, *5*, 10315. [[CrossRef](#)]
60. Terada, S.; Hirai, Y.; Hirai, H.; Koga, A. Higher-order repeat structure in alpha satellite DNA is an attribute of hominoids rather than hominids. *J. Hum. Genet.* **2013**, *58*, 752–754. [[CrossRef](#)] [[PubMed](#)]
61. Koga, A.; Hirai, Y.; Terada, S.; Jahan, I.; Baicharoen, S.; Arsaithamkul, V.; Hirai, H. Evolutionary Origin of Higher-Order Repeat Structure in Alpha-Satellite DNA of Primate Centromeres. *DNA Res.* **2014**, *21*, 407–415. [[CrossRef](#)] [[PubMed](#)]
62. Vance, G.H.; Curtis, C.A.; Heerema, N.A.; Schwartz, S.; Palmer, C.G. An apparently acentric marker chromosome originating from 9p with a functional centromere without detectable alpha and beta satellite sequences. *Am. J. Med. Genet.* **1997**, *71*, 436–442. [[CrossRef](#)]
63. Depinet, T. Characterization of neo-centromeres in marker chromosomes lacking detectable alpha-satellite DNA. *Hum. Mol. Genet.* **1997**, *6*, 1195–1204. [[CrossRef](#)]
64. Marshall, O.J.; Choo, K.H.A. Neocentromeres Come of Age. *PLoS Genet.* **2009**, *5*, e1000370. [[CrossRef](#)]
65. Marshall, O.J.; Chueh, A.C.; Wong, L.H.; Choo, K.A. Neocentromeres: New Insights into Centromere Structure, Disease Development, and Karyotype Evolution. *Am. J. Hum. Genet.* **2008**, *82*, 261–282. [[CrossRef](#)]
66. Maggert, K.A.; Karpen, G.H. The activation of a neocentromere in *Drosophila* requires proximity to an endogenous centromere. *Genetics* **2001**, *158*, 1615–1628.
67. Thakur, J.; Sanyal, K. Efficient neocentromere formation is suppressed by gene conversion to maintain centromere function at native physical chromosomal loci in *Candida albicans*. *Genome Res.* **2013**, *23*, 638–652. [[CrossRef](#)]
68. Williams, B.C.; Murphy, T.D.; Goldberg, M.L.; Karpen, G.H. Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nat. Genet.* **1998**, *18*, 30–38. [[CrossRef](#)]
69. Shang, W.-H.; Hori, T.; Martins, N.M.; Toyoda, A.; Misu, S.; Monma, N.; Hiratani, I.; Maeshima, K.; Ikeo, K.; Fujiyama, A.; et al. Chromosome Engineering Allows the Efficient Isolation of Vertebrate Neocentromeres. *Dev. Cell* **2013**, *24*, 635–648. [[CrossRef](#)] [[PubMed](#)]
70. Garagna, S.; Redi, C.; Capanna, E.; Andayani, N.; Alfano, R.; Doi, P.; Viale, G. Genome distribution, chromosomal allocation, and organization of the major and minor satellite DNAs in 11 species and subspecies of the genus *Mus*. *Cytogenet. Genome Res.* **1993**, *64*, 247–255. [[CrossRef](#)]
71. Kuznetsova, I.; Podgornaya, O.; Ferguson-Smith, M. High-resolution organization of mouse centromeric and pericentromeric DNA. *Cytogenet. Genome Res.* **2006**, *112*, 248–255. [[CrossRef](#)]
72. Boursot, P.; Auffray, J.; Brittondavidian, J.; Bonhomme, F. The evolution of house mice. *Annu. Rev. Ecol. Syst.* **1993**, *24*, 119–152. [[CrossRef](#)]

73. Cazaux, B.; Catalan, J.; Justy, F.; Escudé, C.; Desmarais, E.; Britton-Davidian, J. Evolution of the structure and composition of house mouse satellite DNA sequences in the subgenus *Mus* (Rodentia: Muridea): A cytogenomic approach. *Chromosoma* **2013**, *122*, 209–220. [[CrossRef](#)]
74. Neumann, P.; Navrátilová, A.; Koblížková, A.; Kejnovský, E.; Hříbová, E.; Hobza, R.; Widmer, A.; Doležel, J.; Macas, J. Plant centromeric retrotransposons: A structural and cytogenetic perspective. *Mob. DNA* **2011**, *2*, 4. [[CrossRef](#)]
75. Kordiš, D. A genomic perspective on the chromodomain-containing retrotransposons: Chromoviruses. *Gene* **2005**, *347*, 161–173. [[CrossRef](#)]
76. Nasuda, S.; Hudakova, S.; Schubert, I.; Houben, A.; Endo, T.R. Stable barley chromosomes without centromeric repeats. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 9842–9847. [[CrossRef](#)] [[PubMed](#)]
77. Bao, W.; Zhang, W.; Yang, Q.; Zhang, Y.; Han, B.; Gu, M.; Xue, Y.; Cheng, Z. Diversity of centromeric repeats in two closely related wild rice species, *Oryza officinalis* and *Oryza rhizomatis*. *Mol. Genet. Genom.* **2006**, *275*, 421–430. [[CrossRef](#)] [[PubMed](#)]
78. Zhang, T.; Talbert, P.B.; Zhang, W.; Wu, Y.; Yang, Z.; Henikoff, J.G.; Henikoff, S.; Jiang, J. The CentO satellite confers translational and rotational phasing on cenH3 nucleosomes in rice centromeres. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4875–E4883. [[CrossRef](#)] [[PubMed](#)]
79. Yang, X.; Zhao, H.; Zhang, T.; Zeng, Z.; Zhang, P.; Zhu, B.; Han, Y.; Braz, G.T.; Casler, M.D.; Schmutz, J.; et al. Amplification and adaptation of centromeric repeats in polyploid switchgrass species. *New Phytol.* **2018**, *218*, 1645–1657. [[CrossRef](#)] [[PubMed](#)]
80. Houben, A.; Schroeder-Reiter, E.; Nagaki, K.; Nasuda, S.; Wanner, G.; Murata, M.; Endo, T.R. CENH3 interacts with the centromeric retrotransposon cereba and GC-rich satellites and locates to centromeric substructures in barley. *Chromosoma* **2007**, *116*, 275–283. [[CrossRef](#)] [[PubMed](#)]
81. Gao, X.; Hou, Y.; Ebina, H.; Levin, H.L.; Voytas, D.F. Chromodomains direct integration of retrotransposons to heterochromatin. *Genome Res.* **2008**, *18*, 359–369. [[CrossRef](#)] [[PubMed](#)]
82. Jin, W.; Melo, J.R.; Nagaki, K.; Talbert, P.B.; Henikoff, S.; Dawe, R.K.; Jiang, J. Maize Centromeres: Organization and Functional Adaptation in the Genetic Background of Oat. *Plant Cell* **2004**, *16*, 571–581. [[CrossRef](#)] [[PubMed](#)]
83. Kaszás, E.; Birchler, J.A. Meiotic transmission rates correlate with physical features of rearranged centromeres in maize. *Genetics* **1998**, *150*, 1683–1692.
84. Kaszás, E.; Birchler, J.A. Misdivision analysis of centromere structure in maize. *EMBO J.* **1996**, *15*, 5246–5255. [[CrossRef](#)] [[PubMed](#)]
85. Dong, F.; Miller, J.T.; Jackson, S.A.; Wang, G.-L.; Ronald, P.C.; Jiang, J. Rice (*Oryza sativa*) centromeric regions consist of complex DNA. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 8135–8140. [[CrossRef](#)]
86. Jiang, J.; Birchler, J.A.; Parrott, W.A.; Dawe, R.K. A molecular view of plant centromeres. *Trends Plant Sci.* **2003**, *8*, 570–575. [[CrossRef](#)]
87. Birchler, J.A.; Gao, Z.; Han, F. Plant Centromeres. In *Advanced Structural Safety Studies*; Metzler, J.B., Ed.; Springer: New York, NY, USA, 2011; Volume 4, pp. 133–142.
88. Lermontova, I.; Sandmann, M.; Demidov, D. Centromeres and kinetochores of *Brassicaceae*. *Chromosom. Res.* **2014**, *22*, 135–152. [[CrossRef](#)]
89. Sun, X.; Le, H.D.; Wahlstrom, J.M.; Karpen, G.H. Sequence Analysis of a Functional *Drosophila* Centromere. *Genome Res.* **2003**, *13*, 182–194. [[CrossRef](#)]
90. Sun, X.; Wahlstrom, J.; Karpen, G. Molecular Structure of a Functional *Drosophila* Centromere. *Cell* **1997**, *91*, 1007–1019. [[CrossRef](#)]
91. Lohe, A.R.; Brutlag, D.L. Multiplicity of satellite DNA sequences in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 696–700. [[CrossRef](#)]
92. Schneider, K.L.; Xie, Z.; Wolfgruber, T.K.; Presting, G.G. Inbreeding drives maize centromere evolution. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E987–E996. [[CrossRef](#)]
93. Jones, K.W. Chromosomal and Nuclear Location of Mouse Satellite DNA in Individual Cells. *Nat. Cell Biol.* **1970**, *225*, 912–915. [[CrossRef](#)] [[PubMed](#)]
94. Pardue, M.L.; Gall, J.G. Chromosomal Localization of Mouse Satellite DNA. *Science* **1970**, *168*, 1356–1358. [[CrossRef](#)]
95. Fransz, P.; De Jong, J.H.; Lysak, M.; Castiglione, M.R.; Schubert, I. Interphase chromosomes in Arabidopsis are organized as well defined chromocenters from which euchromatin loops emanate. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14584–14589. [[CrossRef](#)] [[PubMed](#)]
96. Jagannathan, M.; Cummings, R.; Yamashita, Y.M. A conserved function for pericentromeric satellite DNA. *eLife* **2018**, *7*, e34122. [[CrossRef](#)] [[PubMed](#)]
97. Tagarro, I.; Fernández-Peralta, A.M.; González-Aguilera, J.J. Chromosomal localization of human satellites 2 and 3 by a FISH method using oligonucleotides as probes. *Qual. Life Res.* **1994**, *93*, 383–388. [[CrossRef](#)]
98. Vissel, B.; Nagy, A.; Choo, K. A satellite III sequence shared by human chromosomes 13,14, and 21 that is contiguous with α satellite DNA. *Cytogenet. Genome Res.* **1992**, *61*, 81–86. [[CrossRef](#)]
99. Meyne, J.; Goodwin, E.H.; Moyzis, R.K. Chromosome localization and orientation of the simple sequence repeat of human satellite I DNA. *Chromosoma* **1994**, *103*, 99–103. [[CrossRef](#)] [[PubMed](#)]
100. Moyzis, R.K.; Albright, K.L.; Bartholdi, M.F.; Cram, L.S.; Deaven, L.L.; Hildebrand, C.E.; Joste, N.E.; Longmire, J.L.; Meyne, J.; Schwarzscher-Robinson, T. Human chromosome-specific repetitive DNA sequences: Novel markers for genetic analysis. *Chromosoma* **1987**, *95*, 375–386. [[CrossRef](#)]

101. Beauchamp, R.S.; Mitchell, A.R.; Buckland, R.A.; Bostock, C.J. Specific arrangements of human satellite III DNA sequences in human chromosomes. *Chromosoma* **1979**, *71*, 153–166. [[CrossRef](#)]
102. Prosser, J.; Frommer, M.; Paul, C.; Vincent, P. Sequence relationships of three human satellite DNAs. *J. Mol. Biol.* **1986**, *187*, 145–155. [[CrossRef](#)]
103. Tagarro, I.; Wiegant, J.; Raap, A.K. Assignment of human satellite 1 DNA as revealed by fluorescent In Situ hybridization with oligonucleotides. *Qual. Life Res.* **1994**, *93*, 125–128. [[CrossRef](#)] [[PubMed](#)]
104. Lee, C.; Wevrick, R.; Fisher, R.B.; Ferguson-Smith, M.A.; Lin, C.C. Human centromeric DNAs. *Qual. Life Res.* **1997**, *100*, 291–304. [[CrossRef](#)] [[PubMed](#)]
105. Greig, G.M.; Willard, H.F. β satellite DNA: Characterization and localization of two subfamilies from the distal and proximal short arms of the human acrocentric chromosomes. *Genomics* **1992**, *12*, 573–580. [[CrossRef](#)]
106. Schueler, M.G.; Dunn, J.M.; Bird, C.P.; Ross, M.T.; Viggiano, L.; Rocchi, M.; Willard, H.F.; Green, E.D. Progressive proximal expansion of the primate X chromosome centromere. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10563–10568. [[CrossRef](#)]
107. Welborn, J. Acquired Robertsonian translocations are not rare events in acute leukemia and lymphoma. *Cancer Genet. Cytogenet.* **2004**, *151*, 14–35. [[CrossRef](#)]
108. Page, S.L.; Shin, J.-C.; Han, J.-Y.; Choo, K.H.A.; Shaffer, L.G. Breakpoint diversity illustrates distinct mechanisms for Robertsonian translocation formation. *Hum. Mol. Genet.* **1996**, *5*, 1279–1288. [[CrossRef](#)]
109. Sullivan, B.A.; Jenkins, L.S.; Karson, E.M.; Leana-Cox, J.; Schwartz, S. Evidence for structural heterogeneity from molecular cytogenetic analysis of dicentric Robertsonian translocations. *Am. J. Hum. Genet.* **1996**, *59*, 167–175.
110. Bandyopadhyay, R.; Heller, A.; Knox-DuBois, C.; McCaskill, C.; Berend, S.A.; Page, S.L.; Shaffer, L.G. Parental Origin and Timing of De Novo Robertsonian Translocation Formation. *Am. J. Hum. Genet.* **2002**, *71*, 1456–1462. [[CrossRef](#)]
111. Capanna, E.; Gropp, A.; Winking, H.; Noack, G.; Civitelli, M.-V. Robertsonian metacentrics in the mouse. *Chromosoma* **1976**, *58*, 341–353. [[CrossRef](#)]
112. Nachman, M.W.; Searle, J.B. Why is the house mouse karyotype so variable? *Trends Ecol. Evol.* **1995**, *10*, 397–402. [[CrossRef](#)]
113. Han, J.Y.; Choo, K.H.; Shaffer, L.G. Molecular cytogenetic characterization of 17 rob(13q14q) Robertsonian translocations by FISH, narrowing the region containing the breakpoints. *Am. J. Hum. Genet.* **1994**, *55*, 960–967. [[PubMed](#)]
114. Choo, K.H. Role of acrocentric cen-pter satellite DNA in Robertsonian translocation and chromosomal non-disjunction. *Mol. Biol. Med.* **1990**, *7*, 437–449.
115. Garagna, S.; Page, J.; Fernandez-Donoso, R.; Zuccotti, M.; Searle, J.B. The Robertsonian phenomenon in the house mouse: Mutation, meiosis and speciation. *Chromosoma* **2014**, *123*, 529–544. [[CrossRef](#)]
116. Guenatri, M.; Bailly, D.; Maison, C.; Almouzni, G. Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J. Cell Biol.* **2004**, *166*, 493–505. [[CrossRef](#)]
117. Thompson, H.L.; Schmidt, R.; Dean, C. Identification and distribution of seven classes of middle-repetitive DNA in the *Arabidopsis thaliana* genome. *Nucleic Acids Res.* **1996**, *24*, 3017–3022. [[CrossRef](#)]
118. Brandes, A.; Thompson, H.; Dean, C.; Heslop-Harrison, J.S. Multiple repetitive DNA sequences in the paracentromeric regions of *Arabidopsis thaliana* L. *Chromosom. Res.* **1997**, *5*, 238–246. [[CrossRef](#)]
119. Bauwens, S.; Van Oostveldt, P.; Engler, G.; Van Montagu, M. Distribution of the rDNA and three classes of highly repetitive DNA in the chromatin of interphase nuclei of *Arabidopsis thaliana*. *Chromosoma* **1991**, *101*, 41–48. [[CrossRef](#)]
120. Stupar, R.M.; Song, J.; Tek, A.L.; Cheng, Z.; Dong, F.; Jiang, J. Highly condensed potato pericentromeric heterochromatin contains rDNA-related tandem repeats. *Genetics* **2002**, *162*, 1435–1444.
121. Kuhn, G.C.S.; Küttler, H.; Moreira-Filho, O.; Heslop-Harrison, J.S.; Heslop-Harrison, J. The 1.688 Repetitive DNA of *Drosophila*: Concerted Evolution at Different Genomic Scales and Association with Genes. *Mol. Biol. Evol.* **2011**, *29*, 7–11. [[CrossRef](#)]
122. Losada, A.; Villasante, A. Autosomal location of a new subtype of 1.688 satellite DNA of *Drosophila melanogaster*. *Chromosom. Res.* **1996**, *4*, 372–383. [[CrossRef](#)]
123. Houtchens, K.; Lyttle, T.W. Responder (Rsp) Alleles in the Segregation Distorter (SD) System of Meiotic Drive in *Drosophila* may Represent a Complex Family of Satellite Repeat Sequences. *Genetics* **2003**, *117*, 291–302. [[CrossRef](#)]
124. Plohl, M.; Mestrovic, N.; Bruvo, B.; Ugarkovic, D. Similarity of Structural Features and Evolution of Satellite DNAs from *Palorus subdepressus* (Coleoptera) and Related Species. *J. Mol. Evol.* **1998**, *46*, 234–239. [[CrossRef](#)] [[PubMed](#)]
125. Lorite, P.; Palomeque, T.; Garner, I.; Petitpierre, E. Characterization and chromosome location of satellite DNA in the leaf beetle *Chrysolina americana* (Coleoptera, Chrysomelidae). *Genetics* **2000**, *110*, 143–150. [[CrossRef](#)]
126. Palomeque, T.; Muñoz-López, M.; Carrillo, J.A.; Lorite, P. Characterization and evolutionary dynamics of a complex family of satellite DNA in the leaf beetle *Chrysolina carnifex* (Coleoptera, Chrysomelidae). *Chromosom. Res.* **2005**, *13*, 795–807. [[CrossRef](#)] [[PubMed](#)]
127. Feliciello, I.; Chinali, G.; Ugarkovic, D. Structure and population dynamics of the major satellite DNA in the red flour beetle *Tribolium castaneum*. *Genetics* **2011**, *139*, 999–1008. [[CrossRef](#)] [[PubMed](#)]
128. Feliciello, I.; Akrap, I.; Brajkovic, J.; Zlatac, I.; Ugarkovic, D. Satellite DNA as a Driver of Population Divergence in the Red Flour Beetle *Tribolium castaneum*. *Genome Biol. Evol.* **2015**, *7*, 228–239. [[CrossRef](#)]
129. Shaw, D.D. Centromeres: Moving chromosomes through space and time. *Trends Ecol. Evol.* **1994**, *9*, 170–175. [[CrossRef](#)]
130. Blattes, R.; Monod, C.; Susbielle, G.; Cuvier, O.; Wu, J.-H.; Hsieh, T.-S.; Laemmli, U.K.; Käs, E. Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. *EMBO J.* **2006**, *25*, 2397–2408. [[CrossRef](#)]

131. Wu, M.-L. Association between a satellite DNA sequence and the responder of segregation distorter in *D. melanogaster*. *Cell* **1988**, *54*, 179–189. [[CrossRef](#)]
132. Larracuent, A.M. The organization and evolution of the Responder satellite in species of the *Drosophila melanogaster* group: Dynamic evolution of a target of meiotic drive. *BMC Evol. Biol.* **2014**, *14*, 1–12. [[CrossRef](#)]
133. Sandler, L.; Novitski, E. Meiotic Drive as an Evolutionary Force. *Am. Nat.* **1957**, *91*, 105–110. [[CrossRef](#)]
134. Wu, C.-I.; True, J.R.; Johnson, N. Fitness reduction associated with the deletion of a satellite DNA array. *Nat. Cell Biol.* **1989**, *341*, 248–251. [[CrossRef](#)]
135. Pavlek, M.; Gelfand, Y.; Plohl, M.; Meštrović, N. Genome-wide analysis of tandem repeats in *Tribolium castaneum* genome reveals abundant and highly dynamic tandem repeat families with satellite DNA features in euchromatic chromosomal arms. *DNA Res.* **2015**, *22*, 387–401. [[CrossRef](#)]
136. Ugarkovic, D.; Podnar, M.; Plohl, M. Satellite DNA of the red flour beetle *Tribolium castaneum*—Comparative study of satellites from the genus *Tribolium*. *Mol. Biol. Evol.* **1996**, *13*, 1059–1066. [[CrossRef](#)]
137. Blackburn, E.H. Telomeres and telomerase: Their mechanisms of action and the effects of altering their functions. *FEBS Lett.* **2004**, *579*, 859–862. [[CrossRef](#)]
138. Bebikhov, D.V. Repeating sequences, organizing the telomeric region of chromosomes from the eukaryotic genome. *Genetika* **1993**, *29*, 373–387.
139. Vitková, M.; Král, J.; Traut, W.; Zrzavý, J.; Marec, F. The evolutionary origin of insect telomeric repeats, (TTAGG) N. *Chromosom. Res.* **2005**, *13*, 145–156. [[CrossRef](#)]
140. Riha, K.; Shippen, R.E. Telomere structure, function and maintenance in *Arabidopsis*. *Chromosom. Res.* **2003**, *11*, 263–275. [[CrossRef](#)] [[PubMed](#)]
141. Makarov, V.L.; Hirose, Y.; Langmore, J.P. Long G Tails at Both Ends of Human Chromosomes Suggest a C Strand Degradation Mechanism for Telomere Shortening. *Cell* **1997**, *88*, 657–666. [[CrossRef](#)]
142. Casacuberta, E. *Drosophila*: Retrotransposons Making up Telomeres. *Viruses* **2017**, *9*, 192. [[CrossRef](#)] [[PubMed](#)]
143. Mason, J.M.; Ransom, J.; Konev, A.Y. A Deficiency Screen for Dominant Suppressors of Telomeric Silencing in *Drosophila*. *Genetics* **2004**, *168*, 1353–1370. [[CrossRef](#)]
144. Sheen, F.M.; Levis, R.W. Transposition of the LINE-like retrotransposon TART to *Drosophila* chromosome termini. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 12510–12514. [[CrossRef](#)]
145. Mason, J.M.; Biessmann, H. The unusual telomeres of *Drosophila*. *Trends Genet.* **1995**, *11*, 58–62. [[CrossRef](#)]
146. Levis, R.W.; Ganesan, R.; Houtchens, K.; Tolar, L.A.; Sheen, F.-M. Transposons in place of telomeric repeats at a *Drosophila* telomere. *Cell* **1993**, *75*, 1083–1093. [[CrossRef](#)]
147. Roth, C.W.; Kobeski, F.; Walter, M.F.; Biessmann, H. Chromosome end elongation by recombination in the mosquito *Anopheles gambiae*. *Mol. Cell. Biol.* **1997**, *17*, 5176–5183. [[CrossRef](#)] [[PubMed](#)]
148. Biessmann, H.; Kobeski, F.; Walter, M.F.; Kasravi, A.; Roth, C.W. DNA organization and length polymorphism at the 2L telomeric region of *Anopheles gambiae*. *Insect Mol. Biol.* **1998**, *7*, 83–93. [[CrossRef](#)]
149. Compton, A.; Liang, J.; Chen, C.; Lukyanchikova, V.; Qi, Y.; Potters, M.; Settlege, R.; Miller, D.; Deschamps, S.; Mao, C.; et al. The Beginning of the End: A Chromosomal Assembly of the New World Malaria Mosquito Ends with a Novel Telomere. *Genes Genomes Genet.* **2020**, *10*, 3811–3819. [[CrossRef](#)]
150. Madalena, C.R.G.; Amabis, J.M.; Gorab, E. Unusually short tandem repeats appear to reach chromosome ends of *Rhynchosciara americana* (Diptera: Sciariidae). *Chromosoma* **2010**, *119*, 613–623. [[CrossRef](#)]
151. Flint, J.; Thomas, K.; Micklem, G.; Raynham, H.; Clark, K.; Doggett, N.A.; Andrew, A.; Higgs, D.R. The relationship between chromosome structure and function at a human telomeric region. *Nat. Genet.* **1997**, *15*, 252–257. [[CrossRef](#)]
152. Flint, J.; Bates, G.P.; Clark, K.; Dorman, A.; Willingham, D.; Roe, B.A.; Micklem, G.; Higgs, D.R.; Louis, E.J. Sequence comparison of human and yeast telomeres identifies structurally distinct subtelomeric domains. *Hum. Mol. Genet.* **1997**, *6*, 1305–1313. [[CrossRef](#)]
153. Chute, I.; Le, Y.; Ashley, T.; Dobson, M.J. The Telomere-Associated DNA from Human Chromosome 20p Contains a Pseudotelomere Structure and Shares Sequences with the Subtelomeric Regions of 4q and 18p. *Genomics* **1997**, *46*, 51–60. [[CrossRef](#)]
154. Riethman, H.; Ambrosini, A.; Paul, S. Human subtelomere structure and variation. *Chromosom. Res.* **2005**, *13*, 505–515. [[CrossRef](#)]
155. Mefford, H.C.; Trask, B.J. The complex structure and dynamic evolution of human subtelomeres. *Nat. Rev. Genet.* **2002**, *3*, 91–102. [[CrossRef](#)]
156. Ambrosini, A.; Paul, S.; Hu, S.; Riethman, H. Human subtelomeric duplicon structure and organization. *Genome Biol.* **2007**, *8*, R151. [[CrossRef](#)]
157. Riethman, H. Human subtelomeric copy number variations. *Cytogenet. Genome Res.* **2008**, *123*, 244–252. [[CrossRef](#)]
158. Ohzeki, J.-I.; Nakano, M.; Okada, T.; Masumoto, H. CENP-B box is required for De Novo centromere chromatin assembly on human alphoid DNA. *J. Cell Biol.* **2002**, *159*, 765–775. [[CrossRef](#)]
159. Fachinetti, D.; Han, J.S.; McMahon, M.A.; Ly, P.; Abdullah, A.; Wong, A.J.; Cleveland, D.W. DNA Sequence-Specific Binding of CENP-B Enhances the Fidelity of Human Centromere Function. *Dev. Cell* **2015**, *33*, 314–327. [[CrossRef](#)]
160. Masumoto, H.; Masukata, H.; Muro, Y.; Nozaki, N.; Okazaki, T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J. Cell Biol.* **1989**, *109*, 1963–1973. [[CrossRef](#)]

161. Mateo, L.; González, J. Pogo-like Transposases Have Been Repeatedly Domesticated into CENP-B-Related Proteins. *Genome Biol. Evol.* **2014**, *6*, 2008–2016. [[CrossRef](#)]
162. Casola, C.; Hucks, D.; Feschotte, C. Convergent Domestication of pogo-like Transposases into Centromere-Binding Proteins in Fission Yeast and Mammals. *Mol. Biol. Evol.* **2007**, *25*, 29–41. [[CrossRef](#)]
163. Rosandić, M.; Paar, V.; Basar, I.; Glunčić, M.; Pavin, N.; Pilaš, I. CENP-B box and pJ α sequence distribution in human alpha satellite higher-order repeats (HOR). *Chromosom. Res.* **2006**, *14*, 735–753. [[CrossRef](#)]
164. Paar, V.; Basar, I.; Rosandić, M.; Glunčić, M. Consensus Higher Order Repeats and Frequency of String Distributions in Human Genome. *Curr. Genom.* **2007**, *8*, 93–111. [[CrossRef](#)]
165. Kasinathan, S.; Henikoff, S. Non-B-Form DNA Is Enriched at Centromeres. *Mol. Biol. Evol.* **2018**, *35*, 949–962. [[CrossRef](#)]
166. Garavís, M.; Escaja, N.; Gabelica, V.; Villasante, A.; González, C. Centromeric Alpha-Satellite DNA Adopts Dimeric i-Motif Structures Capped by AT Hoogsteen Base Pairs. *Chem. A Eur. J.* **2015**, *21*, 9816–9824. [[CrossRef](#)]
167. Garavís, M.; Méndez-Lago, M.; Gabelica, V.; Whitehead, S.L.; González, C.; Villasante, A. The structure of an endogenous *Drosophila centromere* reveals the prevalence of tandemly repeated sequences able to form i-motifs. *Sci. Rep.* **2015**, *5*, 13307. [[CrossRef](#)]
168. Barceló, F.; Gutierrez, F.; Barjau, I.; Portugal, J. A Theoretical Perusal of the Satellite DNA Curvature in Tenebrionid Beetles. *J. Biomol. Struct. Dyn.* **1998**, *16*, 41–50. [[CrossRef](#)]
169. Rojas-Rousse, D.; Bigot, Y.; Periquet, G. DNA insertions as a component of the evolution of unique satellite DNA families in two genera of parasitoid wasps: *Diadromus* and *Eupelmus* (Hymenoptera). *Mol. Biol. Evol.* **1993**, *10*, 383–396. [[CrossRef](#)]
170. Landais, I.; Chavigny, P.; Castagnone, C.; Pizzol, J.; Abad, P.; Vanlerberghe-Masutti, F. Characterization of a highly conserved satellite DNA from the parasitoid wasp *Trichogramma brassicae*. *Gene* **2000**, *255*, 65–73. [[CrossRef](#)]
171. Mravinac, B.; Ugarković, Đ.; Franjević, D.; Plohl, M. Long Inversely Oriented Subunits Form a Complex Monomer of *Tribolium brevicornis* Satellite DNA. *J. Mol. Evol.* **2005**, *60*, 513–525. [[CrossRef](#)]
172. Rosén, M.; Castillejo-López, C.; Edström, J.-E. Telomere terminating with centromere-specific repeats is closely associated with a transposon derived gene in *Chironomus pallidivittatus*. *Chromosoma* **2002**, *110*, 532–541. [[CrossRef](#)] [[PubMed](#)]
173. Pons, J.; Bruvo, B.; Petitpierre, E.; Plohl, M.; Ugarković, D.; Juan, C. Complex structural features of satellite DNA sequences in the genus *Pimelia* (Coleoptera: Tenebrionidae): Random differential amplification from a common ‘satellite DNA library’. *Heredity* **2004**, *92*, 418–427. [[CrossRef](#)] [[PubMed](#)]
174. Barcelo, F.; Pons, J.; Petitpierre, E.; Barjau, I.; Portugal, J. Polymorphic Curvature of Satellite DNA in Three Subspecies of the Beetle *Pimelia sparsa*. *J. Biol. Inorg. Chem.* **1997**, *244*, 318–324. [[CrossRef](#)]
175. Lobov, I.B.; Tsutsui, K.; Mitchell, A.R.; Podgornaya, O.I. Specificity of SAF-A and lamin B binding in vitro correlates with the satellite DNA bending state. *J. Cell. Biochem.* **2001**, *83*, 218–229. [[CrossRef](#)]
176. de Lange, T. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes Dev.* **2005**, *19*, 2100–2110. [[CrossRef](#)] [[PubMed](#)]
177. Denchi, E.L. Give me a break: How telomeres suppress the DNA damage response. *DNA Repair* **2009**, *8*, 1118–1126. [[CrossRef](#)]
178. Griffith, J.D.; Comeau, L.; Rosenfield, S.; Stansel, R.M.; Bianchi, A.; Moss, H.; de Lange, T. Mammalian Telomeres End in a Large Duplex Loop. *Cell* **1999**, *97*, 503–514. [[CrossRef](#)]
179. Ruis, P.; Boulton, S.J. The end protection problem—An unexpected twist in the tail. *Genes Dev.* **2021**, *35*, 1–21. [[CrossRef](#)]
180. Muñoz-Jordán, J.L.; Cross, G.A.; De Lange, T.; Griffith, J.D. t-loops at trypanosome telomeres. *EMBO J.* **2001**, *20*, 579–588. [[CrossRef](#)]
181. Tomáška, L.; Cesare, A.J.; Alturki, T.M.; Griffith, J.D. Twenty years of t-loops: A case study for the importance of collaboration in molecular biology. *DNA Repair* **2020**, *94*, 102901. [[CrossRef](#)]
182. Paeschke, K.; Simonsson, T.; Postberg, J.; Rhodes, D.; Lipps, H.J. Telomere end-binding proteins control the formation of G-quadruplex DNA structures In Vivo. *Nat. Struct. Mol. Biol.* **2005**, *12*, 847–854. [[CrossRef](#)]
183. Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H.J.; Plückthun, A. In Vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with *Styloynchia lemnae* macronuclei. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8572–8577. [[CrossRef](#)] [[PubMed](#)]
184. Pagano, B.; Amato, J.; Iaccarino, N.; Cingolani, C.; Zizza, P.; Biroccio, A.; Novellino, E.; Randazzo, A. Looking for Efficient G-Quadruplex Ligands: Evidence for Selective Stabilizing Properties and Telomere Damage by Drug-Like Molecules. *ChemMedChem* **2015**, *10*, 640–649. [[CrossRef](#)]
185. Granotier, C. Preferential binding of a G-quadruplex ligand to human chromosome ends. *Nucleic Acids Res.* **2005**, *33*, 4182–4190. [[CrossRef](#)] [[PubMed](#)]
186. Tahara, H.; Shin-Ya, K.; Seimiya, H.; Yamada, H.; Tsuruo, T.; Ide, T. G-Quadruplex stabilization by telomestatin induces TRF2 protein dissociation from telomeres and anaphase bridge formation accompanied by loss of the 3' telomeric overhang in cancer cells. *Oncogene* **2006**, *25*, 1955–1966. [[CrossRef](#)] [[PubMed](#)]
187. Rodriguez, R.; Müller, S.; Yeoman, J.A.; Trentesaux, C.; Riou, J.-F.; Balasubramanian, S. A Novel Small Molecule That Alters Shelterin Integrity and Triggers a DNA-Damage Response at Telomeres. *J. Am. Chem. Soc.* **2008**, *130*, 15758–15759. [[CrossRef](#)]
188. Canudas, S.; Smith, S. Differential regulation of telomere and centromere cohesion by the Scc3 homologues SA1 and SA2, respectively, in human cells. *J. Cell Biol.* **2009**, *187*, 165–173. [[CrossRef](#)] [[PubMed](#)]
189. Saint-André, C.D.L.R. Alternative ends: Telomeres and meiosis. *Biochimie* **2008**, *90*, 181–189. [[CrossRef](#)]

190. Sen, D.; Gilbert, W. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nat. Cell Biol.* **1988**, *334*, 364–366. [[CrossRef](#)]
191. Csink, A.K.; Henikoff, S. Something from nothing: The evolution and utility of satellite repeats. *Trends Genet.* **1998**, *14*, 200–204. [[CrossRef](#)]
192. Blower, M.D.; Sullivan, B.A.; Karpen, G.H. Conserved Organization of Centromeric Chromatin in Flies and Humans. *Dev. Cell* **2002**, *2*, 319–330. [[CrossRef](#)]
193. Sullivan, B.A.; Karpen, G.H. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1076–1083. [[CrossRef](#)]
194. Palmer, D.K.; Day, M.H.W.; Wener, M.H.; Andrews, B.S.; Margolis, R.L. A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *J. Cell Biol.* **1987**, *104*, 805–815. [[CrossRef](#)] [[PubMed](#)]
195. Westhorpe, F.G.; Straight, A.F. The Centromere: Epigenetic Control of Chromosome Segregation during Mitosis. *Cold Spring Harb. Perspect. Biol.* **2014**, *7*, a015818. [[CrossRef](#)] [[PubMed](#)]
196. Henikoff, S.; Thakur, J.; Kasinathan, S.; Talbert, P.B. Remarkable Evolutionary Plasticity of Centromeric Chromatin. *Cold Spring Harb. Symp. Quant. Biol.* **2017**, *82*, 71–82. [[CrossRef](#)]
197. Westermann, S.; Cheeseman, I.M.; Anderson, S.; Yates, J.R.; Drubin, D.G.; Barnes, G. Architecture of the budding yeast kinetochore reveals a conserved molecular core. *J. Cell Biol.* **2003**, *163*, 215–222. [[CrossRef](#)]
198. Welburn, J.P.; Cheeseman, I.M. Toward a Molecular Structure of the Eukaryotic Kinetochore. *Dev. Cell* **2008**, *15*, 645–655. [[CrossRef](#)]
199. Hayden, K.E.; Strome, E.D.; Merrett, S.L.; Lee, H.-R.; Rudd, M.K.; Willard, H.F. Sequences Associated with Centromere Competency in the Human Genome. *Mol. Cell Biol.* **2012**, *33*, 763–772. [[CrossRef](#)]
200. Thakur, J.; Henikoff, S. Unexpected conformational variations of the human centromeric chromatin complex. *Genes Dev.* **2018**, *32*, 20–25. [[CrossRef](#)]
201. Hasson, D.; Panchenko, T.; Salimian, K.J.; Salman, M.U.; Sekulic, N.; Alonso, A.; Warburton, P.E.; Black, B.E. The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat. Struct. Mol. Biol.* **2013**, *20*, 687–695. [[CrossRef](#)]
202. Bodor, D.L.; Mata, J.F.; Sergeev, M.; David, A.F.; Salimian, K.J.; Panchenko, T.; Cleveland, D.W.; Black, B.E.; Shah, J.V.; Jansen, L.E. The quantitative architecture of centromeric chromatin. *eLife* **2014**, *3*, e02137. [[CrossRef](#)] [[PubMed](#)]
203. Thakur, J.; Henikoff, S. CENPT bridges adjacent CENPA nucleosomes on young human α -satellite dimers. *Genome Res.* **2016**, *26*, 1178–1187. [[CrossRef](#)]
204. Kato, H.; Jiang, J.; Zhou, B.-R.; Rozendaal, M.; Feng, H.; Ghirlando, R.; Xiao, T.S.; Straight, A.F.; Bai, Y. A Conserved Mechanism for Centromeric Nucleosome Recognition by Centromere Protein CENP-C. *Science* **2013**, *340*, 1110–1113. [[CrossRef](#)]
205. Ando, S.; Yang, H.; Nozaki, N.; Okazaki, T.; Yoda, K. CENP-A, -B, and -C Chromatin Complex That Contains the I-Type α -Satellite Array Constitutes the Prekinetochore in HeLa Cells. *Mol. Cell Biol.* **2002**, *22*, 2229–2241. [[CrossRef](#)] [[PubMed](#)]
206. Bobkov, G.O.; Gilbert, N.; Heun, P. Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation. *J. Cell Biol.* **2018**, *217*, 1957–1972. [[CrossRef](#)] [[PubMed](#)]
207. Chan, F.L.; Marshall, O.J.; Saffery, R.; Kim, B.W.; Earle, E.; Choo, K.H.A.; Wong, L.H. Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1979–1984. [[CrossRef](#)] [[PubMed](#)]
208. Wong, L.H.; Brettingham-Moore, K.H.; Chan, L.; Quach, J.M.; Anderson, M.A.; Northrop, E.L.; Hannan, R.; Saffery, R.; Shaw, M.L.; Williams, E.; et al. Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res.* **2007**, *17*, 1146–1160. [[CrossRef](#)]
209. Du, Y.; Topp, C.N.; Dawe, R.K. DNA Binding of Centromere Protein C (CENPC) Is Stabilized by Single-Stranded RNA. *PLoS Genet.* **2010**, *6*, e1000835. [[CrossRef](#)]
210. McNulty, S.M.; Sullivan, L.L.; Sullivan, B.A. Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are Complexed with CENP-A and CENP-C. *Dev. Cell* **2017**, *42*, 226–240. [[CrossRef](#)]
211. Yamagata, K.; Yamazaki, T.; Miki, H.; Ogonuki, N.; Inoue, K.; Ogura, A.; Baba, T. Centromeric DNA hypomethylation as an epigenetic signature discriminates between germ and somatic cell lineages. *Dev. Biol.* **2007**, *312*, 419–426. [[CrossRef](#)]
212. Luo, S.; Preuss, D. Strand-biased DNA methylation associated with centromeric regions in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 11133–11138. [[CrossRef](#)]
213. Li, Y.; Miyanari, Y.; Shirane, K.; Nitta, H.; Kubota, T.; Ohashi, H.; Okamoto, A.; Sasaki, H. Sequence-specific microscopic visualization of DNA methylation status at satellite repeats in individual cell nuclei and chromosomes. *Nucleic Acids Res.* **2013**, *41*, e186. [[CrossRef](#)]
214. Koo, D.-H.; Han, F.; Birchler, J.A.; Jiang, J. Distinct DNA methylation patterns associated with active and inactive centromeres of the maize B chromosome. *Genome Res.* **2011**, *21*, 908–914. [[CrossRef](#)]
215. Yan, H.; Kikuchi, S.; Neumann, P.; Zhang, W.; Wu, Y.; Chen, F.; Jiang, J. Genome-wide mapping of cytosine methylation revealed dynamic DNA methylation patterns associated with genes and centromeres in rice. *Plant J.* **2010**, *63*, 353–365. [[CrossRef](#)]
216. Ichikawa, K.; Tomioka, S.; Suzuki, Y.; Nakamura, R.; Doi, K.; Yoshimura, J.; Kumagai, M.; Inoue, Y.; Uchida, Y.; Irie, N.; et al. Centromere evolution and CpG methylation during vertebrate speciation. *Nat. Commun.* **2017**, *8*, 1833. [[CrossRef](#)]
217. Mitchell, A.R.; Jeppesen, P.; Nicol, L.; Morrison, H.; Kipling, D. Epigenetic control of mammalian centromere protein binding: Does DNA methylation have a role? *J. Cell Sci.* **1996**, *109*, 2199–2206. [[PubMed](#)]

218. Gopalakrishnan, S.; Sullivan, B.A.; Trazzi, S.; Della Valle, G.; Robertson, K.D. DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions. *Hum. Mol. Genet.* **2009**, *18*, 3178–3193. [[CrossRef](#)]
219. Tartof, K.D.; Bishop, C.; Jones, M.; Hobbs, C.A.; Locke, J. Towards an understanding of position effect variegation. *Dev. Genet.* **1989**, *10*, 162–176. [[CrossRef](#)] [[PubMed](#)]
220. Girton, J.R.; Johansen, K.M. Chapter 1 Chromatin Structure and the Regulation of Gene Expression: The Lessons of PEV in *Drosophila*. *Adv. Genet.* **2008**, *61*, 1–43. [[CrossRef](#)] [[PubMed](#)]
221. Sass, G.L.; Henikoff, S. Comparative analysis of position-effect variegation mutations in *Drosophila melanogaster* delineates the targets of modifiers. *Genetics* **1998**, *148*, 733–741.
222. Hall, L.L.; Byron, M.; Carone, D.M.; Whitfield, T.W.; Pouliot, G.P.; Fischer, A.; Jones, P.; Lawrence, J.B. Demethylated HSATII DNA and HSATII RNA Foci Sequester PRC1 and MeCP2 into Cancer-Specific Nuclear Bodies. *Cell Rep.* **2017**, *18*, 2943–2956. [[CrossRef](#)] [[PubMed](#)]
223. Landers, C.C.; Rabeler, C.A.; Ferrari, E.K.; D’Alessandro, L.R.; Kang, D.D.; Malisa, J.; Bashir, S.M.; Carone, D.M. Ectopic expression of pericentric HSATII RNA results in nuclear RNA accumulation, MeCP2 recruitment, and cell division defects. *Chromosoma* **2021**, *130*, 75–90. [[CrossRef](#)] [[PubMed](#)]
224. Bernard, P.; Maure, J.-F.; Partridge, J.F.; Genier, S.; Javerzat, J.-P.; Allshire, R.C. Requirement of Heterochromatin for Cohesion at Centromeres. *Science* **2001**, *294*, 2539–2542. [[CrossRef](#)] [[PubMed](#)]
225. Schotta, G.; Sengupta, R.; Kubicek, S.; Malin, S.; Kauer, M.; Callén, E.; Celeste, A.; Pagani, M.; Opravil, S.; De La Rosa-Velazquez, I.A.; et al. A chromatin-wide transition to H4K20 monomethylation impairs genome integrity and programmed DNA rearrangements in the mouse. *Genes Dev.* **2008**, *22*, 2048–2061. [[CrossRef](#)] [[PubMed](#)]
226. Schotta, G.; Lachner, M.; Sarma, K.; Ebert, A.; Sengupta, R.; Reuter, G.; Reinberg, D.; Jenuwein, T. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* **2004**, *18*, 1251–1262. [[CrossRef](#)] [[PubMed](#)]
227. Almouzni, G.; Probst, A.V. Heterochromatin maintenance and establishment: Lessons from the mouse pericentromere. *Nucleus* **2011**, *2*, 332–338. [[CrossRef](#)] [[PubMed](#)]
228. Hahn, M.; Dambacher, S.; Dulev, S.; Kuznetsova, A.Y.; Eck, S.; Wörz, S.; Sadic, D.; Schulte, M.; Mallm, J.-P.; Maiser, A.; et al. Suv4-20h2 mediates chromatin compaction and is important for cohesin recruitment to heterochromatin. *Genes Dev.* **2013**, *27*, 859–872. [[CrossRef](#)]
229. Jørgensen, S.; Schotta, G.; Sørensen, C.S. Histone H4 Lysine 20 methylation: Key player in epigenetic regulation of genomic integrity. *Nucleic Acids Res.* **2013**, *41*, 2797–2806. [[CrossRef](#)]
230. Lehnertz, B.; Ueda, Y.; Derijck, A.A.; Braunschweig, U.; Perez-Burgos, L.; Kubicek, S.; Chen, T.; Li, E.; Jenuwein, T.; Peters, A.H. Suv39h-Mediated Histone H3 Lysine 9 Methylation Directs DNA Methylation to Major Satellite Repeats at Pericentric Heterochromatin. *Curr. Biol.* **2003**, *13*, 1192–1200. [[CrossRef](#)]
231. Probst, A.; Okamoto, I.; Casanova, M.; El Marjou, F.; Le Baccon, P.; Almouzni, G. A Strand-Specific Burst in Transcription of Pericentric Satellites Is Required for Chromocenter Formation and Early Mouse Development. *Dev. Cell* **2010**, *19*, 625–638. [[CrossRef](#)] [[PubMed](#)]
232. Yadav, R.P.; Mäkelä, J.-A.; Hyssälä, H.; Cisneros-Montalvo, S.; Kotaja, N. DICER regulates the expression of major satellite repeat transcripts and meiotic chromosome segregation during spermatogenesis. *Nucleic Acids Res.* **2020**, *48*, 7135–7153. [[CrossRef](#)]
233. Kishi, Y.; Kondo, S.; Gotoh, Y. Transcriptional Activation of Mouse Major Satellite Regions during Neuronal Differentiation. *Cell Struct. Funct.* **2012**, *37*, 101–110. [[CrossRef](#)] [[PubMed](#)]
234. Camacho, O.V.; Galan, C.; Swist-Rosowska, K.; Ching, R.; Gamalinda, M.; Karabiber, F.; De La Rosa-Velazquez, I.; Engist, B.; Koschorz, B.; Shukeir, N.; et al. Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *eLife* **2017**, *6*, e25293. [[CrossRef](#)] [[PubMed](#)]
235. Huo, X.; Ji, L.; Zhang, Y.; Lv, P.; Cao, X.; Wang, Q.; Yan, Z.; Dong, S.; Du, D.; Zhang, F.; et al. The Nuclear Matrix Protein SAFB Cooperates with Major Satellite RNAs to Stabilize Heterochromatin Architecture Partially through Phase Separation. *Mol. Cell* **2020**, *77*, 368–383. [[CrossRef](#)] [[PubMed](#)]
236. Maison, C.; Bailly, D.; Peters, A.H.; Quivy, J.-P.; Roche, D.; Taddei, A.; Lachner, M.; Jenuwein, T.; Almouzni, G. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.* **2002**, *30*, 329–334. [[CrossRef](#)]
237. Thakur, J.; Henikoff, S. Architectural RNA in chromatin organization. *Biochem. Soc. Trans.* **2020**, *48*, 1967–1978. [[CrossRef](#)] [[PubMed](#)]
238. Thakur, J.; Fang, H.; Llagas, T.; Disteche, C.M.; Henikoff, S. Architectural RNA is required for heterochromatin organization. *bioRxiv* **2019**, 784835. [[CrossRef](#)]
239. Brown, K.E.; Guest, S.S.; Smale, S.T.; Hahn, K.; Merckenschlager, M.; Fisher, A.G. Association of Transcriptionally Silent Genes with Ikaros Complexes at Centromeric Heterochromatin. *Cell* **1997**, *91*, 845–854. [[CrossRef](#)]
240. Hahn, K.; Cobb, B.S.; Mccarty, A.S.; Brown, K.E.; Klug, C.A.; Lee, R.; Akashi, K.; Weissman, I.L.; Fisher, A.G.; Smale, S.T. Helios, a T cell-restricted Ikaros family member that quantitatively associates with Ikaros at centromeric heterochromatin. *Genes Dev.* **1998**, *12*, 782–796. [[CrossRef](#)] [[PubMed](#)]
241. Cobb, B.S.; Morales-Alcelay, S.; Kleiger, G.; Brown, K.E.; Fisher, A.G.; Smale, S.T. Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes Dev.* **2000**, *14*, 2146–2160. [[CrossRef](#)]

242. Gurel, Z.; Ronni, T.; Ho, S.; Kuchar, J.; Payne, K.J.; Turk, C.W.; Dovat, S. Recruitment of Ikaros to Pericentromeric Heterochromatin Is Regulated by Phosphorylation. *J. Biol. Chem.* **2008**, *283*, 8291–8300. [[CrossRef](#)]
243. Kim, J.-H.; Ebersole, T.; Kouprina, N.; Noskov, V.N.; Ohzeki, J.-I.; Masumoto, H.; Mravinac, B.; Sullivan, B.A.; Pavlicek, A.; Dovat, S.; et al. Human gamma-satellite DNA maintains open chromatin structure and protects a transgene from epigenetic silencing. *Genome Res.* **2009**, *19*, 533–544. [[CrossRef](#)]
244. Koering, C.E.; Pollice, A.; Zibella, M.P.; Bauwens, S.; Puisieux, A.; Brunori, M.; Brun, C.; Martins, L.; Sabatier, L.; Pulitzer, J.F.; et al. Human telomeric position effect is determined by chromosomal context and telomeric chromatin integrity. *EMBO Rep.* **2002**, *3*, 1055–1061. [[CrossRef](#)]
245. Baur, J.A.; Zou, Y.; Shay, J.W.; Wright, W.E. Telomere Position Effect in Human Cells. *Science* **2001**, *292*, 2075–2077. [[CrossRef](#)] [[PubMed](#)]
246. Pisano, S.; Pascucci, E.; Cacchione, S.; de Santis, P.; Savino, M. AFM imaging and theoretical modeling studies of sequence-dependent nucleosome positioning. *Biophys. Chem.* **2006**, *124*, 81–89. [[CrossRef](#)] [[PubMed](#)]
247. Makarov, V.L.; Lejnine, S.; Bedoyan, J.; Langmore, J.P. Nucleosomal organization of telomere-specific chromatin in rat. *Cell* **1993**, *73*, 775–787. [[CrossRef](#)]
248. Mechelli, R.; Anselmi, C.; Cacchione, S.; de Santis, P.; Savino, M. Organization of telomeric nucleosomes: Atomic force microscopy imaging and theoretical modeling. *FEBS Lett.* **2004**, *566*, 131–135. [[CrossRef](#)]
249. Filesi, I.; Cacchione, S.; de Santis, P.; Rossetti, L.; Savino, M. The main role of the sequence-dependent DNA elasticity in determining the free energy of nucleosome formation on telomeric DNAs. *Biophys. Chem.* **2000**, *83*, 223–237. [[CrossRef](#)]
250. Ichikawa, Y.; Morohashi, N.; Nishimura, Y.; Kurumizaka, H.; Shimizu, M. Telomeric repeats act as nucleosome-disfavouring sequences In Vivo. *Nucleic Acids Res.* **2013**, *42*, 1541–1552. [[CrossRef](#)]
251. Blasco, M.A. The epigenetic regulation of mammalian telomeres. *Nat. Rev. Genet.* **2007**, *8*, 299–309. [[CrossRef](#)]
252. Jones, B.; Su, H.; Bhat, A.; Lei, H.; Bajko, J.; Hevi, S.; Baltus, G.A.; Kadam, S.; Zhai, H.; Valdez, R.; et al. The Histone H3K79 Methyltransferase Dot1L Is Essential for Mammalian Development and Heterochromatin Structure. *PLoS Genet.* **2008**, *4*, e1000190. [[CrossRef](#)]
253. Bandaria, J.N.; Qin, P.; Berk, V.; Chu, S.; Yildiz, A. Shelterin Protects Chromosome Ends by Compacting Telomeric Chromatin. *Cell* **2016**, *164*, 735–746. [[CrossRef](#)] [[PubMed](#)]
254. Raffa, G.D.; Ciapponi, L.; Cenci, G.; Gatti, M. Terminin: A protein complex that mediates epigenetic maintenance of *Drosophila* telomeres. *Nucl.* **2011**, *2*, 383–391. [[CrossRef](#)]
255. Frydrychova, R.C.; Mason, J.M.; Archer, T.K. HP1 Is Distributed within Distinct Chromatin Domains at *Drosophila* Telomeres. *Genetics* **2008**, *180*, 121–131. [[CrossRef](#)]
256. Raffa, G.D.; Siriaco, G.; Cugusi, S.; Ciapponi, L.; Cenci, G.; Wojcik, E.; Gatti, M. The *Drosophila* modigliani (moi) gene encodes a HOAP-interacting protein required for telomere protection. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 2271–2276. [[CrossRef](#)] [[PubMed](#)]
257. Benetti, R.; Gonzalo, S.; Jaco, I.; Schotta, G.; Klatt, P.; Jenuwein, T.; Blasco, M.A. Suv4-20h deficiency results in telomere elongation and derepression of telomere recombination. *J. Cell Biol.* **2007**, *178*, 925–936. [[CrossRef](#)] [[PubMed](#)]
258. Gonzalo, S.; García-Cao, M.; Fraga, M.F.; Schotta, G.; Peters, A.H.F.M.; Cotter, S.E.; Eguía, R.; Dean, D.C.; Esteller, M.; Jenuwein, T.; et al. Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nat. Cell Biol.* **2005**, *7*, 420–428. [[CrossRef](#)] [[PubMed](#)]
259. Gonzalo, S.; Blasco, M.A. Role of Rb Family in the Epigenetic Definition of Chromatin. *Cell Cycle* **2005**, *4*, 752–755. [[CrossRef](#)] [[PubMed](#)]
260. Pedram, M.; Sprung, C.N.; Gao, Q.; Lo, A.W.I.; Reynolds, G.E.; Murnane, J.P. Telomere Position Effect and Silencing of Transgenes near Telomeres in the Mouse. *Mol. Cell. Biol.* **2006**, *26*, 1865–1878. [[CrossRef](#)]
261. Garrido-Ramos, M.A. Satellite DNA: An Evolving Topic. *Genes* **2017**, *8*, 230. [[CrossRef](#)]
262. Gonzalo, S.; Jaco, I.; Fraga, M.F.; Chen, T.; Li, E.; Esteller, M.; Blasco, M.A. DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nat. Cell Biol.* **2006**, *8*, 416–424. [[CrossRef](#)] [[PubMed](#)]
263. Schoeftner, S.; Blasco, M.A. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat. Cell Biol.* **2007**, *10*, 228–236. [[CrossRef](#)] [[PubMed](#)]
264. Azzalin, C.M.; Reichenbach, P.; Khoraiuli, L.; Giulotto, E.; Lingner, J. Telomeric Repeat Containing RNA and RNA Surveillance Factors at Mammalian Chromosome Ends. *Science* **2007**, *318*, 798–801. [[CrossRef](#)] [[PubMed](#)]
265. Schoeftner, S.; Blasco, M.A. Chromatin regulation and non-coding RNAs at mammalian telomeres. *Semin. Cell Dev. Biol.* **2010**, *21*, 186–193. [[CrossRef](#)]
266. Horard, B.; Gilson, E. Telomeric RNA enters the game. *Nat. Cell Biol.* **2008**, *10*, 113–115. [[CrossRef](#)]
267. Savitsky, M.; Kwon, D.; Georgiev, P.; Kalmykova, A.; Gvozdev, V. Telomere elongation is under the control of the RNAi-based mechanism in the *Drosophila* germline. *Genes Dev.* **2006**, *20*, 345–354. [[CrossRef](#)]
268. Xu, Y.; Kimura, T.; Komiyama, M. Human telomere RNA and DNA form an intermolecular G-quadruplex. *Nucleic Acids Symp. Ser.* **2008**, *52*, 169–170. [[CrossRef](#)] [[PubMed](#)]
269. Zahler, A.M.; Williamson, J.R.; Cech, T.R.; Prescott, D.M. Inhibition of telomerase by G-quartet DNA structures. *Nat. Cell Biol.* **1991**, *350*, 718–720. [[CrossRef](#)]
270. Xu, Y.; Kaminaga, K.; Komiyama, M. Human telomeric RNA in G-quadruplex structure. *Nucleic Acids Symp. Ser.* **2008**, *52*, 175–176. [[CrossRef](#)] [[PubMed](#)]

271. Phan, A.T. Human telomeric G-quadruplex: Structures of DNA and RNA sequences. *FEBS J.* **2010**, *277*, 1107–1117. [[CrossRef](#)]
272. Meštrović, N.; Mravinac, B.; Juan, C.; Ugarković, Đ.; Plohl, M. Comparative study of satellite sequences and phylogeny of five species from the genus *Palorus* (Insecta, Coleoptera). *Genome* **2000**, *43*, 776–785. [[CrossRef](#)] [[PubMed](#)]
273. Willard, H.F. Evolution of alpha satellite. *Curr. Opin. Genet. Dev.* **1991**, *1*, 509–514. [[CrossRef](#)]
274. Biscotti, M.A.; Olmo, E.; Heslop-Harrison, J.S. (Pat) Repetitive DNA in eukaryotic genomes. *Chromosom. Res.* **2015**, *23*, 415–420. [[CrossRef](#)] [[PubMed](#)]
275. Palomeque, T.; Lorite, P. Satellite DNA in insects: A review. *Heredity* **2008**, *100*, 564–573. [[CrossRef](#)] [[PubMed](#)]
276. Mestrovic, N.; Plohl, M.; Mravinac, B.; Ugarkovic, D. Evolution of satellite DNAs from the genus *Palorus*—experimental evidence for the “library” hypothesis. *Mol. Biol. Evol.* **1998**, *15*, 1062–1068. [[CrossRef](#)] [[PubMed](#)]
277. Fry, K.; Salser, W. Nucleotide sequences of HS- α satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* **1977**, *12*, 1069–1084. [[CrossRef](#)]
278. Ruiz-Ruano, F.J.; López-León, M.D.; Cabrero, J.; Camacho, J.P.M. High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Sci. Rep.* **2016**, *6*, 28333. [[CrossRef](#)]
279. Cesari, M.; Luchetti, A.; Passamonti, M.; Scali, V.; Mantovani, B. Polymerase chain reaction amplification of the Bag320 satellite family reveals the ancestral library and past gene conversion events in *Bacillus rossius* (Insecta Phasmatodea). *Gene* **2003**, *312*, 289–295. [[CrossRef](#)]
280. del Bosque, M.Q.; Navajas-Pérez, R.; Panero, J.; Fernández-González, A.J.; Garrido-Ramos, M. A satellite DNA evolutionary analysis in the North American endemic dioecious plant *Rumex hastatulus* (Polygonaceae). *Genome* **2011**, *54*, 253–260. [[CrossRef](#)]
281. del Bosque, M.E.Q.; López-Flores, I.; Suárez-Santiago, V.N.; Garrido-Ramos, M.A. Satellite-DNA diversification and the evolution of major lineages in *Cardueae* (Carduoideae Asteraceae). *J. Plant Res.* **2014**, *127*, 575–583. [[CrossRef](#)]
282. Feliciello, I.; Picariello, O.; Chinali, G. The first characterisation of the overall variability of repetitive units in a species reveals unexpected features of satellite DNA. *Gene* **2005**, *349*, 153–164. [[CrossRef](#)]
283. Meštrović, N.; Mravinac, B.; Pavlek, M.; Vojvoda-Zeljko, T.; Šatović, E.; Plohl, M. Structural and functional liaisons between transposable elements and satellite DNAs. *Chromosom. Res.* **2015**, *23*, 583–596. [[CrossRef](#)]
284. Belyayev, A.; Josefiová, J.; Jandová, M.; Mahelka, V.; Krak, K.; Mandák, B. Transposons and satellite DNA: On the origin of the major satellite DNA family in the *Chenopodium* genome. *Mob. DNA* **2020**, *11*, 1–10. [[CrossRef](#)]
285. del Bosque, M.E.Q.; López-Flores, I.; Suárez-Santiago, V.N.; Garrido-Ramos, M.A. Differential spreading of *Hinf*I satellite DNA variants during radiation in *Centaureinae*. *Ann. Bot.* **2013**, *112*, 1793–1802. [[CrossRef](#)] [[PubMed](#)]
286. Wei, K.H.-C.; Grenier, J.K.; Barbash, D.A.; Clark, A.G. Correlated variation and population differentiation in satellite DNA abundance among lines of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 18793–18798. [[CrossRef](#)] [[PubMed](#)]
287. Durfy, S.J.; Willard, H.F. Concerted evolution of primate alpha satellite DNA: Evidence for an ancestral sequence shared by gorilla and human X chromosome alpha satellite. *J. Mol. Biol.* **1990**, *216*, 555–566. [[CrossRef](#)]
288. Waye, J.S.; Willard, H.F. Concerted evolution of alpha satellite DNA: Evidence for species specificity and a general lack of sequence conservation among alphoid sequences of higher primates. *Chromosoma* **1989**, *98*, 273–279. [[CrossRef](#)]
289. Kimura, M.; Crow, J.F. The number of alleles that can be maintained in a finite population. *Genetics* **1964**, *49*, 725–738. [[CrossRef](#)]
290. Smith, G. Evolution of repeated DNA sequences by unequal crossover. *Science* **1976**, *191*, 528–535. [[CrossRef](#)]
291. Smith, G.P. Unequal Crossover and the Evolution of Multigene Families. *Cold Spring Harb. Symp. Quant. Biol.* **1974**, *38*, 507–513. [[CrossRef](#)] [[PubMed](#)]
292. Perelson, A.S.; Bell, G.I. Mathematical models for the evolution of multigene families by unequal crossing over. *Nat. Cell Biol.* **1977**, *265*, 304–310. [[CrossRef](#)] [[PubMed](#)]
293. Stephan, W.; Cho, S. Possible role of natural selection in the formation of tandem-repetitive noncoding DNA. *Genet.* **1994**, *136*, 333–341. [[CrossRef](#)]
294. Stephan, W. Tandem-repetitive noncoding DNA: Forms and forces. *Mol. Biol. Evol.* **1989**, *6*, 198–212. [[CrossRef](#)] [[PubMed](#)]
295. Dover, G.A. Molecular drive: A cohesive mode of species evolution. *Nature* **1982**, *299*, 111–117. [[CrossRef](#)] [[PubMed](#)]
296. Dover, G. Molecular drive. *Trends Genet.* **2002**, *18*, 587–589. [[CrossRef](#)]
297. Liao, D.; Pavelitz, T.; Kidd, J.R.; Kidd, K.K.; Weiner, A.M. Concerted evolution of the tandemly repeated genes encoding human U2 snRNA (the RNU2 locus) involves rapid intrachromosomal homogenization and rare interchromosomal gene conversion. *EMBO J.* **1997**, *16*, 588–598. [[CrossRef](#)]
298. Rice, W.R. Game of Thrones at Human Centromeres II. A new molecular/evolutionary model. *bioRxiv* **2019**. [[CrossRef](#)]
299. Kobayashi, T. Recombination Regulation by Transcription-Induced Cohesin Dissociation in rDNA Repeats. *Science* **2005**, *309*, 1581–1584. [[CrossRef](#)]
300. Malik, H.S.; Henikoff, S. Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*. *Genetics* **2001**, *157*, 1293–1298.
301. Rhoades, M.M. Preferential Segregation in Maize. *Genetics* **1942**, *27*, 395–407. [[CrossRef](#)]
302. Iwata-Otsubo, A.; Dawicki-McKenna, J.M.; Akeru, T.; Falk, S.J.; Chmátal, L.; Yang, K.; Sullivan, B.A.; Schultz, R.M.; Lampson, M.A.; Black, B.E. Expanded Satellite Repeats Amplify a Discrete CENP-A Nucleosome Assembly Site on Chromosomes that Drive in Female Meiosis. *Curr. Biol.* **2017**, *27*, 2365–2373. [[CrossRef](#)] [[PubMed](#)]
303. Chmátal, L.; Gabriel, S.I.; Mitsainas, G.P.; Martínez-Vargas, J.; Ventura, J.; Searle, J.B.; Schultz, R.M.; Lampson, M.A. Centromere Strength Provides the Cell Biological Basis for Meiotic Drive and Karyotype Evolution in Mice. *Curr. Biol.* **2014**, *24*, 2295–2300. [[CrossRef](#)]

304. Fishman, L.; Saunders, A. Centromere-Associated Female Meiotic Drive Entails Male Fitness Costs in Monkeyflowers. *Science* **2008**, *322*, 1559–1562. [[CrossRef](#)]
305. Saint-Leandre, B.; Christopher, C.; Levine, M.T. Adaptive evolution of an essential telomere protein restricts telomeric retrotransposons. *eLife* **2020**, *9*, e60987. [[CrossRef](#)]
306. Saint-Leandre, B.; Levine, M.T. The Telomere Paradox: Stable Genome Preservation with Rapidly Evolving Proteins. *Trends Genet.* **2020**, *36*, 232–242. [[CrossRef](#)]
307. Pontremoli, C.; Forni, D.; Cagliani, R.; Pozzoli, U.; Clerici, M.; Sironi, M. Evolutionary rates of mammalian telomere-stability genes correlate with karyotype features and female germline expression. *Nucleic Acids Res.* **2018**, *46*, 7153–7168. [[CrossRef](#)]
308. Cutter, A.D. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. *Trends Ecol. Evol.* **2012**, *27*, 209–218. [[CrossRef](#)]
309. Brideau, N.J.; Flores, H.A.; Wang, J.; Maheshwari, S.; Wang, X.; Barbash, D.A. Two Dobzhansky-Muller Genes Interact to Cause Hybrid Lethality in *Drosophila*. *Science* **2006**, *314*, 1292–1295. [[CrossRef](#)] [[PubMed](#)]
310. Satyaki, P.R.V.; Cuykendall, T.N.; Wei, K.H.-C.; Brideau, N.J.; Kwak, H.; Aruna, S.; Ferree, P.M.; Ji, S.; Barbash, D.A. The Hmr and Lhr Hybrid Incompatibility Genes Suppress a Broad Range of Heterochromatic Repeats. *PLoS Genet.* **2014**, *10*, e1004240. [[CrossRef](#)] [[PubMed](#)]
311. Ferree, P.M.; Barbash, D.A. Species-Specific Heterochromatin Prevents Mitotic Chromosome Segregation to Cause Hybrid Lethality in *Drosophila*. *PLoS Biol.* **2009**, *7*, e1000234. [[CrossRef](#)] [[PubMed](#)]
312. Ting, D.T.; Lipson, D.; Paul, S.; Brannigan, B.W.; Akhavanfard, S.; Coffman, E.J.; Contino, G.; Deshpande, V.; Iafrate, A.J.; Letovsky, S.; et al. Aberrant Overexpression of Satellite Repeats in Pancreatic and Other Epithelial Cancers. *Science* **2011**, *331*, 593–596. [[CrossRef](#)]
313. Bersani, F.; Lee, E.; Kharchenko, P.V.; Xu, A.W.; Liu, M.; Xega, K.; MacKenzie, O.C.; Brannigan, B.W.; Wittner, B.S.; Jung, H.; et al. Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 15148–15153. [[CrossRef](#)]
314. Ferreira, D.; Meles, S.; Escudeiro, A.; Mendes-da-Silva, A.; Adegas, F.; Chaves, R. Satellite non-coding RNAs: The emerging players in cells, cellular pathways and cancer. *Chromosom. Res.* **2015**, *23*, 479–493. [[CrossRef](#)] [[PubMed](#)]
315. Jackson, K.; Yu, M.C.; Arakawa, K.; Fiala, E.; Youn, B.; Fiegl, H.; Müller-Holzner, E.; Widschwendter, M.; Ehrlich, M. DNA hypomethylation is prevalent even in low-grade breast cancers. *Cancer Biol. Ther.* **2004**, *3*, 1225–1231. [[CrossRef](#)] [[PubMed](#)]
316. Vukic, M.; Daxinger, L. DNA methylation in disease: Immunodeficiency, Centromeric instability, Facial anomalies syndrome. *Essays Biochem.* **2019**, *63*, 773–783. [[CrossRef](#)] [[PubMed](#)]
317. Miga, K.H.; Koren, S.; Rhie, A.; Vollger, M.R.; Gershman, A.; Bzikadze, A.; Brooks, S.; Howe, E.; Porubsky, D.; Logsdon, G.A.; et al. Telomere-to-telomere assembly of a complete human X chromosome. *Nature* **2020**, *585*, 79–84. [[CrossRef](#)]
318. Logsdon, G.A.; Vollger, M.R.; Hsieh, P.; Mao, Y.; Liskovych, M.A.; Koren, S.; Nurk, S.; Mercuri, L.; Dishuck, P.C.; Rhie, A.; et al. The structure, function and evolution of a complete human chromosome 8. *Nat. Cell Biol.* **2021**, 1–7. [[CrossRef](#)]