Decoding the Role of Satellite DNA in Genome Architecture and Plasticity—An Evolutionary and Clinical Affair

Sandra Louzada [1,2], **Mariana Lopes** [1,2], **Daniela Ferreira** [1,2], **Filomena Adega** [1,2], **Ana Escudeiro** [1,2], **Margarida Gama-Carvalho** [2] **and Raquel Chaves** [1,2,*]

[1] Laboratory of Cytogenomics and Animal Genomics (CAG), Department of Genetics and Biotechnology (DGB), University of Trás-os-Montes and Alto Douro (UTAD), 5000-801 Vila Real, Portugal; slouzada@utad.pt (S.L.); lopesfmariana@gmail.com (M.L.); daniela_p_ferreira@hotmail.com (D.F.); filadega@utad.pt (F.A.); anac.escudeiro@gmail.com (A.E.)

[2] Biosystems and Integrative Sciences Institute (BioISI), Faculty of Sciences, University of Lisboa, 1749-016 Lisbon, Portugal; mhcarvalho@fc.ul.pt

[*] Correspondence: rchaves@utad.pt

Abstract: Repetitive DNA is a major organizational component of eukaryotic genomes, being intrinsically related with their architecture and evolution. Tandemly repeated satellite DNAs (satDNAs) can be found clustered in specific heterochromatin-rich chromosomal regions, building vital structures like functional centromeres and also dispersed within euchromatin. Interestingly, despite their association to critical chromosomal structures, satDNAs are widely variable among species due to their high turnover rates. This dynamic behavior has been associated with genome plasticity and chromosome rearrangements, leading to the reshaping of genomes. Here we present the current knowledge regarding satDNAs in the light of new genomic technologies, and the challenges in the study of these sequences. Furthermore, we discuss how these sequences, together with other repeats, influence genome architecture, impacting its evolution and association with disease.

Keywords: satellite DNA; genome architecture; chromosome restructuring; Robertsonian translocations; satellite DNA transcription

1. Introduction

The linear organization of DNA sequences in the genome and how these sequences are packed into chromosomes define their architecture and influence its evolution. Repetitive DNA represents a major organizational component of eukaryotic genomes and includes sequences dispersed throughout the genome like transposable elements (TEs) and tandemly repeated sequences, such as satellite DNA (satDNA) [1,2]. Together with TEs, satDNAs contribute significantly to the differences in genome size between species, accounting for more than 50% of some species total DNA [3]. SatDNAs can be found in varied locations in the chromosomes, such as pericentromeric, subtelomeric and interstitial regions, forming blocks of constitutive heterochromatin (CH) [2–7] that are part of vital structures like centromeres and telomeres [2]. However, satDNA location is not restricted to CH with some satDNAs being found also dispersed throughout euchromatic regions in different species [5,8]. Multiple lines of evidence show that satDNAs have key roles in centromere function, heterochromatin formation and maintenance and chromosome pairing [9–12]. Interestingly, despite their association to critical chromosomal structures, satDNA families can display an astounding sequence variation even among closely related species. This results from their highly dynamic behavior, leading to rapid changes in sequence composition and array size within short evolutionary periods, which can lead to speciation

(reviewed in [13]). Moreover, these sequences have been consistently correlated with fragile sites and evolutionary breakpoint regions in diverse species [14–19] and are intrinsically involved in frequent chromosomal rearrangements like Robertsonian translocations [20,21]. SatDNA dynamics has been shown to promote genome plasticity and to have an active involvement in the modulation of genomic architecture by promoting rearrangements.

Nevertheless, some satDNAs seem to have been preserved or "frozen" across different taxa during long evolutionary periods [22–24] with some of them being transcribed into satellite non-coding RNAs (satncRNAs). Indeed, transcripts of satDNAs have been reported in different species, highlighting a possible role for satncRNAs in the regulation of gene expression, cancer outcomes and aging [25–27]. This suggests that functional constraints may be causing the preservation of these sequences over the time [24,28]. Accordingly, some species centromeric satDNAs have been found to share a 17 bp motif known as the centromere protein B (CENP-B) box, representing the binding site for centromere protein B (CENP-B) [29,30]. It has been demonstrated that the CENP-B box is required for de novo centromere chromatin assembly and CENP-B protein is involved in centromere functions [31]. In this case, the conservation of a sequence motif across diverse mammalian species satDNAs [32] seems to be related to a specific function.

Over the years, different techniques have been used to address satDNA sequences. The advances in sequencing technology and computational approaches have revolutionized the study of these regions, known as the "black holes" of the genome. The increasing number of studies assessing the genomic abundance and sequence variation of satDNAs in different species has led to the coining of new terms to describe the whole collection of repeats (repeatome) and satDNAs (satellitome) in a species genome [33,34], and contributed to improve our knowledge regarding the evolution and function of these sequences [35].

In this review, we contextualize satDNA sequences in the genomes/chromosomes of different species in the light of recent data provided by new technologies and bioinformatic tools and the challenges of studying these DNA sequences and their associated non-coding RNAs. We also discuss the contribution of repetitive sequences to the organization of genomes and their participation in the restructuring of species karyotypes during evolution, focusing on their involvement in rearrangements with evolutionary and clinical significance: Robertsonian translocations. Finally, we address the structural role of satDNA transcripts in the genome.

2. SatDNA Features and Organization in the Genome and Chromosomes: Emerging Technologies and Changing Concepts

The concept of satDNA suffered considerable changes through time. Early experiments historically coined the term "satellite DNA" referring to tandemly arranged sequences that formed satellite bands separate from the rest of the genomic DNA during density gradient centrifugation [36]. Given that no function was initially attributed to these sequences, they were considered as genomic "junk", representing parasites proliferating independently in the genomes [37]. Today, satDNAs are viewed as important genomic functional components. In order to understand participation of these sequences in genome architecture and evolution, we need to briefly address their organizational features, localization and mode of evolution.

SatDNA is typically organized as long arrays of head-to-tail linked repeats and usually present in the genomes in several million copies [1]. The length of the repeating unit (monomer) can range from a few base pairs up to more than 1 kb, forming arrays that may reach 100 Mb in length (reviewed by [38]), and that can form higher-order repeat (HOR) units (e.g., [39–41]). Human chromosome centromeres are populated by α satDNA (*αSAT*) organized in HORs that are structurally distinct and confer chromosome specificity [39,42]. Complex HORs have been found in non-human mammals such as insects, mouse, swine, bovids, horse, dog and elephant (reviewed in [43]), and more recently in Callitrichini monkeys [44] and Teleostei fish [45]. SatDNA arrays are mainly found clustered in heterochromatin, although studies also report the presence of short satDNA arrays dispersed along

euchromatic regions [2–7]. These sequences can be found in varied locations in the chromosomes, such as pericentromeric, subtelomeric and interstitial regions [2,46–48], as well as being part of vital structures like centromeres and telomeres [2].

Usually more than one family of satDNAs can be found in the same genome, thus forming a library, which can be shared among closely related species. The satDNAs within the library may differ in monomer sequence, size, abundance, distribution and location (reviewed in [12]). Expansions and contractions of satDNA arrays can dramatically change the landscape of repetitive sequences, leading to significant differences of satDNA copy number among related species [49,50]. That is the case of the *Drosophila* genus, which contains very dissimilar satDNAs, varying from 0.5% in some species genomes to as high as 50% in others [51,52]. Such striking differences in satDNA abundance in *Drosophila* sp. were proposed to result predominantly from lineage-specific gains accumulated over the past 40 MY of evolution [53], ultimately causing species reproductive barriers [54,55].

The mechanisms proposed to be responsible for the amplification/deletion of repetitive DNA, consequently leading to their rapid evolutionary turnover, are unequal crossing over, replication slippage and rolling circle amplification [56]. SatDNA sequence divergence among species is quite variable, as some repeats are species-specific, while others are widely conserved, being shared across distantly related species [22,24,57]. SatDNAs have a unique mode of evolution, known as concerted evolution, a two-level process in which mutations are homogenized throughout monomers of a repetitive family and concomitantly fixed within a group of reproductively linked organisms [58,59].

The study and characterization of satDNA has lagged behind when compared with other genomic sequences. Throughout time, different methodological approaches have generated insights into the structure, organization, function and evolution of these sequence elements, although this characterization has been significantly hampered by their highly repetitive nature. The advent of high-throughput sequencing technologies and associated bioinformatics tools opened the door to whole genome sequencing projects, and as the technology became more robust and cheaper, the number of sequenced species increased exponentially. In 2018, the Earth BioGenome project was launched, aiming to increase the number of sequenced eukaryotic genomes from 2534 species (of which only 25 comply with the standard for contig and scaffold N50 established by the Genome 10K organization) to characterize the genomes of the 1.5 million known species within a 10 year time frame [60]. Of note, satDNA, as well as other repetitive sequences, have been systematically omitted from the genome projects, due to difficulties in sequence alignment and assembly, given that the read length of current sequencing technologies is unable to span the longer repeats and tandem arrays [61,62]. Nevertheless, high-throughput sequencing contributed significantly to increase our knowledge regarding satDNA sequences [63]. Next generation sequencing (NGS; e.g., Illumina), allied to newly developed bioinformatics tools capable of identifying satDNA sequences in unassembled data (e.g., RepeatExplorer) [64–66], helped uncover the extent of satDNAs present in the genome of different species, revealing unpredicted levels of satDNA diversity (e.g., [34,67–71]). For instances, 62 satDNA families were identified in the genome of the migratory locust, leading to the coining of the term 'satellitome' to refer to the whole collection of satDNA families found in a single genome [34], a part of the 'repeatome', a term proposed previously [33] to refer to the collection of all repetitive sequences in a genome (TEs, satDNAs, etc.). This number has been surpassed by a recent study where 164 satDNA families have been identified in Teleostei fish, being this the biggest satellitome characterized for a given species so far [70]. The availability of a methodology capable of assessing satDNA array abundance and diversity led to an explosion of comparative studies across a wide range of clades, including mammals, insects and plants (e.g., [44,45,69,71–73]) providing insights into these sequences.

The development of sequencing technologies that generate long-range data has allowed the community to overcome some of the limitations imposed by NGS and is fueling the study of repeats. Single-molecule real-time sequencing and nanopore sequencing technologies (commercialized by PacBio and Oxford Nanopore Technologies (ONT), respectively) can generate longer reads capable of

spanning repetitive regions, thus enabling their assembly into contigs (reviewed in [62]). For instances, ONT nanopore sequencers have been shown to generate unprecedented ultra-long reads that can reach mega-base lengths, leading to significant improvements in the human genome assembly [74–77], with some of the repetitive-containing gaps being closed [78,79]. By using long-read methods we are gaining access to important repeated-rich structures, like centromeres, revealing further insights into their sequence content and structure [80]. For instances, *Drosophila* centromeric satDNAs were recently shown to be intermingled with TEs [81]. Other recent studies report the improvement of human Y chromosome centromere assembly [78] and the reconstruction of a 2.8 megabase centromeric satDNA array, with the potential to achieve for the first-time telomere-to-telomere sequencing of the X chromosome [79].

Several studies demonstrate that the combination of different high-throughput sequencing methods (e.g., Illumina, ONT and PacBio) with other techniques, such as optical mapping, cytogenetics and molecular techniques, is beneficial and sometimes essential to determine satDNA features. The use of PacBio long-read sequencing together with optical mapping proved to be helpful in the assembly of satDNA arrays with large monomers and provided insights regarding recombination rates in the Eurasian crow [82]. Positional data derived from fluorescent in situ hybridization (FISH) remains vital to determine the physical location of satDNAs, since such information cannot be achieved for genomes that have not yet been properly assembled (e.g., [34,44,71,81,83]), and sequences mapping by FISH on extended DNA fibers can provide significant assistance to the process of genome assembly, aiding in contig ordering (e.g., [84,85]). Improved techniques based on FISH, helped shedding light into repetitive-rich chromosome regions with centromeric function (e.g., [86]). Other methods have also shown to provide a valid and expedite analysis of repetitive sequences profile, such as PCR-based approaches, that have been used to determine satDNA copy number differences between healthy and cancer cells/tissues [87]. In particular, the use of droplet digital PCR (ddPCR) combined with other methodologies has contributed to the validation and quantification of rare retrotransposon insertion events in different tissues including tumors [88] and the detection and accurate quantification of human *SATII* ncRNA in cancer patients [89]. The integration of genomic, cytogenetic and cell biology data helps to establish a connection between sequence information, its localization in the chromosomes and their interaction with other components of the genome, defining the field of chromosomics [90]. We believe that this approach is essential to fully understand the organization of repetitive sequences.

Other aspects of satDNA biology are also becoming accessible through the use of recent methodologies, such as the characterization of their expression and chromatin state, namely by using RNA sequencing (RNA-seq) and chromatin immunoprecipitation approaches followed by DNA sequencing (CHIP-seq) [91,92]. In particular, for CHIP-seq experiments several studies report the use of a specific antibody against DNA binding centromere-specific histone H3 (CENH3), which is an ortholog for human CENP-A. This methodology has proven to be useful for clarifying the satDNA content in the centromere, improving some organisms reference sequence and uncovering satDNA variability (e.g., [93,94]).

The data generated is now being used to determine satDNA sequences organization in the genome [95], explore predicted evolutionary patterns and hypothesis (e.g., [35,68,96,97]), as well as to shed light into the function of these sequences [81,98]. We are now closer than ever to fully access the sequence information hidden within repetitive-rich chromosome structures like centromeres and telomeres. However, we still need to further develop and adapt currently available approaches to achieve a combination of genomic, cytogenetic and molecular techniques to optimally address these regions, which we propose could be referred to as centrOMICs and telOMICs (Figure 1). SatDNAs represent one of the most intriguing and also interesting components of the genome and their full characterization will help us to better understand genome organization, architecture and evolution.
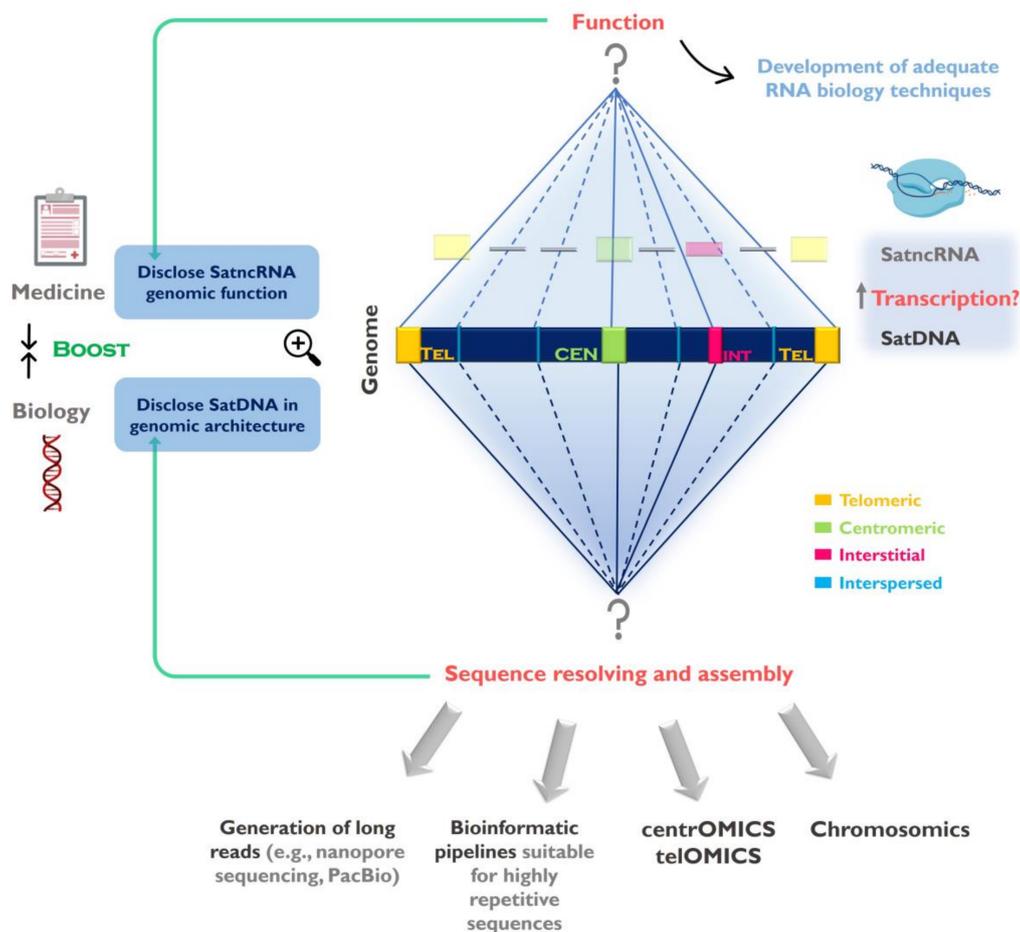
Figure 1. Challenges in the study of satellite DNA (satDNA) sequences and the importance to fully understand the repetitive genomic fraction. SatDNAs can be found clustered at the centromeres, telomeres and forming interstitial heterochromatin (CH) blocks, as well as scattered (interspersed) throughout the chromosomes. The full characterization of satDNAs needs to be addressed in two levels: 1-Disclose satDNAs linear sequence and improve their representation in genome assemblies. Despite currently used sequencing strategies (e.g., next generation sequencing (NGS)) contributed for satDNA studies, the full characterization of these sequences will only be achieved by using sequencing technologies capable of long reads, bioinformatics pipelines suitable for highly repetitive sequences, together with other techniques (e.g., FISH, optical mapping). These strategies need to be directed to specific chromosome structures such as centromeres (centrOMICs) and telomeres (telOMICs), which harbor large amounts of satDNA. Important also is the integration of genomic data with sequence localization in the chromosomes, and their interaction with other components of the genome (chromosomics); 2- Clarify satDNAs function(s) in the genome by studying the satellite non-coding RNAs (satncRNA) and their interaction with other components and structures in the genome. In this field there is the need to develop adequate biology techniques to address repetitive sequences transcription study. The disclosure of satDNA sequences will help to better understand its genomic architecture ant its role in genome restructuring in evolution and disease.

3. Modulating Genome Architecture with SatDNAs

The architecture of genomes confers identity to species. From a generalist point of view, the genomic architectural configuration is the product of a series of sequential molecular events that occurred during the evolutionary process. The impact of these events on genome organization is reflected by chromosome size, number and morphology. Eukaryotic genomes, and particularly, karyotypes, can be viewed as a set of homologous chromosomes, each harboring a combination of syntenic

blocks—conserved blocks that can be differently assembled between species [99]. The events with capability for shaping genomes are based on structural and quantitative chromosomal alterations (e.g., [100]) of variable dimensions, from small to large regions that may completely change the morphology and number of species chromosomes and karyotypes. Amongst these, chromosome fusions (i.e., Robertsonian translocations), fissions (reviewed in [99]) and inversions [101], are perhaps the ones with a stronger impact on the architectural appearance of genomes during species evolution.

Chromosome structural variation may originate from illegitimate non-homologous recombination between different chromosome fields, such as centromeres, chromosome arms and telomeres during meiosis, requiring double strand breaks in at least two chromosomes or chromosome regions [102–105]. The resulting rearranged chromosomes are transmitted either as potentially harmful alterations, or as new variants associated with a selective advantage that will eventually conduct to speciation [99,106,107].

Even before the routine use of advanced molecular technologies, cytogeneticists could realize that the regions where chromosomes break and rearrange (the so-called chromosomal breakpoints) were enriched in constitutive heterochromatin, evidenced by C-bands [57]. Molecular technologies demonstrated that evolutionary breakpoint regions are composed of repeats [107–111]. The involvement of repetitive sequences, including TEs (e.g., [112–115]), segmental duplications (e.g., [108,110,116]) and tandem repeats (e.g., [14–17]), in genome restructuring and evolution is now widely recognized.

The evolutionary rate of tandemly repeated satDNA was shown to be higher than in other genomic sequences, presenting significant changes in short evolutionary times. It is thought that the mechanisms leading to the rapid turnover of these sequences promote chromosome rearrangements and consequently contribute to re-shaping of the genomes. Unequal crossing-over events seem to be responsible for the rapid evolution and divergence found among satDNA families, specifically at the levels of monomer length, nucleotide sequence, complexity and copy number [1,14,49,117,118]. DNA polymerase slippage during DNA replication and recombination in meiosis caused by faulty alignment of repetitive elements further contributes to the instability of these repeat rich regions and to chromosome rearrangements (e.g., [107,119,120]).

SatDNAs can display complex structural organization resulting from the formation of secondary DNA structures, including hairpins, triplexes [121] and even tetraplexes (G-quadruplexes) [122,123]. The formation of such structures can cause problems during genome duplication in the S-phase by slowing down or even stalling the replication fork, resulting in double-strand breaks [124,125]. This damage is then targeted for repair by means of homologous recombination-based mechanisms, which may lead to chromosome and genome architecture alterations due to the selection of identical sequences in non-homologous regions as the template for repair [1,19,126].

Several studies document the presence of TEs intermingled with centromeric satDNA [127,128], in some cases forming complex structures [24,129–131]. TEs are highly represented in some vertebrate species, making up to 60% or more of their genomes. They are characterized by their mobility within genomes using either a direct cut-and-paste mechanism to alter their position (transposons) or requiring an RNA intermediate (retrotransposons) [127,132]. This intrinsic feature makes them active elements of the genome and has been associated with genomic instability. TEs may cause double strand breaks, not only during the transposition process itself, but also by TE–TE ectopic recombination, which may lead to chromosomal rearrangements and consequently to alterations in the genome architecture [133–136]. The integration of TEs in the genome may also result in the disruption of a functional DNA sequence (reviewed in [128]), which can have adverse consequences. Together with segmental duplications, these elements share a high degree of similarity between different intra- and inter-chromosomal regions, making them the perfect templates for non-allelic homologous recombination [137–139]. TEs dynamics has shown to be linked with satDNA origin and evolution. Evidences suggest that some mobile elements may lead to the generation of new repetitive sequences that can be amplified into long arrays of satDNAs [140]. Moreover, it has been suggested that the autonomous LINE-1 retrotransposons could enable amplification and intragenomic movements of satDNA sequences throughout the genome [141]. It thus seems plausible to think that

TEs, especially retrotransposons may, in fact be an adjuvant for satDNA evolution and consequently lead to the creation of genomic innovations.

The dynamic nature of repetitive elements is clearly a basilar reason for genomic plasticity (e.g., [14,102,142,143]) and it is in fact a way of having a low impact on the euchromatic genome [14,97]. Today, an increasing body of evidence strongly validates the involvement of satDNA in the modulation of genomic architectures of a large number of taxa, as in the case of bovids [21,104,144,145], rodents [17,111,143,146], suiformes [57] or genets [147]. This largely extends beyond mammalian evolution, as it can also be observed in insects (e.g., [54,148]), reptiles (e.g., [149]), plants [150] and many other lineages. SatDNAs and TEs can thus be considered the 'engine' triggering genome evolution [14,107], with the regions harboring these sequences functioning as 'hotspots' or fragile sites for structural chromosome rearrangements, leading to species-specific genome architectures [14,17,57,107,139,143,151] and contributing to the generation of key variations responsible for the success of vertebrates [152].

3.1. Repetitive Sequences, Chromosome Instability and Disease

Alterations of genomic architecture can also be pathogenic and have a detrimental effect in organisms, either if occurring at the germinal lineage or somatically. This is the case of many diseases caused or boosted by genomic instability that impacts on nuclear architecture, such as cancer, neurodegenerative disorders and other genetic diseases [153–155]. In fact, alterations in genome architecture can interfere both with chromosomal territories and with topological positioning of chromosomes and genes in the nucleus. Due to the constraints in the regulation of genes and gene networks and to differences in somatic mutation frequencies between genome regions located at the nuclear periphery or core (higher in the periphery) [156], structural variations of critical genome regions can in fact threaten normal cell function. Again, in these situations, repeats seem to be the main actors at play [125,128,153,157]. The repetitive fraction of the eukaryotic genome and in particular, of the mammalian genome, is usually methylated and repressed by a highly condensed chromatin state, which seems to be essential to maintain genome integrity (reviewed in [158,159]). When perturbation of the epigenetic landscape of specific genomic regions occurs (e.g., [160]), repeats that are usually silenced can become active and unconstrained, which may lead to mobilization (in the case of TEs) and an open chromatin state that allows the occurrence of double strand breaks at fragile or hotspot regions. This results in chromosome rearrangements with impact on the three-dimensional genome architecture and gene expression regulation [100], which may lead to disease onset and progression.

3.1.1. Remodeling Genome Architecture Through Robertsonian Translocations from a SatDNA Perspective

The most frequent rearrangements occurring in genomes are Robertsonian translocations (rob). These rearrangements are commonly found in two different genomic scenarios: as an evolutionary rearrangement involved in mammalian karyotypic evolution; and as a chromosomal abnormality with clinical/polymorphic meaning [161,162]. The occurrence of Robertsonian translocations involves a break near or at the centromeric region, followed by the fusion of the entire long (q) arms of two acrocentric chromosomes, forming a dicentric or monocentric chromosome. The associated breakpoints, as well as the subsequent mechanistic steps, have been shown to involve reorganization of satDNA sequences at the centromere level [14,20]. The illegitimate recombination between homologous sequences, such as satDNA on non-homologous chromosomes, has been suggested as a possible path for the occurrence of Robertsonian translocations in mice and humans [163,164]. In fact, the high frequency of rob chromosomes linked to genome remodeling events can be caused not only by the homology of the satDNA sequences shared by the acrocentric chromosomes involved in each translocation, but also by the nicking activity of the centromere protein B (CENP-B), originating the double-strand breaks that precede the fusion events [165]. Robertsonian translocations are complex rearrangements that require, in addition to the double-strand breaks, mechanisms of repair, the silencing of possible additional

centromeric sequences and the adjustment of the amount of CH/satDNA over time, in order to maintain chromosome viability [20,162]. This assigns a primordial task to satDNA in the control, success and viability of Robertsonian translocation events [14,162].

One of the well-known examples of the dual character of these rearrangements is the rob (1;29), which assumes a special relevance as it is the most widespread chromosome rearrangement occurring in domestic cattle with clinical significance [166–169]. In parallel, the rob (1;29) is also a constitutional chromosome rearrangement fixed in several wild bovid species, such as most of the Tragelaphini [170].

The analyses of the sequences at the breakpoint regions preceding a translocation are of great importance in understanding the translocation event [21,131]. These sequences are essentially centromeric satDNAs, whose detailed physical and organizational analysis contributed much to better comprehend the chromosomal mechanism behind the rob (1;29) translocation [20,145]. In 2000, Chaves and colleagues suggested that this chromosomal abnormality might not be a single event [144] and in 2003, using centromeric satDNA sequences, the same group proposed, for the first time, a two-step mechanism for this rearrangement [20]. This translocation mechanism involved, besides the centric fusion of the two acrocentric chromosomes, the loss and reorganization of specific satDNA families that were retained in the translocated chromosome [20]. Later, Di Meo and colleagues [145], using both satDNA and BAC probes, validated the pericentric inversion previously proposed [20]. This event would probably be necessary for satDNA reorganization at the centromeric level, highlighting the active role of satDNA sequences in the translocation mechanism and reinforcing their functional relevance in genome reorganization [14,20].

In humans, the Robertsonian translocations are also the most common structural chromosome abnormality [171,172], with rob (13;14) and rob(14;21) being the most frequent examples [163]. During several decades, aspects such as the high frequency of de novo robs in the human population, their origin during oogenesis, and the non-random participation of the acrocentric chromosomes, have supported the hypothesis that there must be a specific mechanism leading to the formation of these robs [163,164,173]. However, and despite the high frequency of these rearrangements and their clinical implications, there is still insufficient information on the molecular mechanism and exact genomic location of the breakpoints [174]. The rob translocation event has been deeply connected with satDNA sequence homology and consequent recombination [174] giving rise to two alternative explanations: (i) the presence of a homologous inversely-oriented segment on chromosome 14 shared with chromosomes 13 and 21 [163,175]; (ii) the human satellite DNA *SATIII* ability to form uncommon DNA structures that could facilitate the illegitimate recombination [176,177]. However, these hypotheses need further research to be validated. Indeed, in the study of Robertsonian translocations, finding the breakpoint location is a problematic task due to the low resolution of the physical maps at the centromere and short arms of the acrocentric chromosomes [174]. Highly repetitive satDNA undoubtedly represents a major gap in the current human genome assemblies, significantly contributing to the lack of high-resolution sequencing studies in the field of centromere genomics [74,178].

4. Transcribing SatDNAs: Targeting Genomic Functions

The previous sections highlight the role of satDNA sequences in genome architecture and in specific chromosomal rearrangements. However, the participation of these sequences in shaping genome architecture goes beyond their DNA molecule. Currently the transcription of satDNA is a widely accepted feature across species. Different functions have been assigned to satellite non-coding RNAs (satncRNAs) in several cellular contexts, such as cell proliferation, stress response, development or cancer [27] (Figure 2). In fact, satDNA transcripts seem to participate in the most primordial concept of genomic function, being related to centromere structure, chromosome pairing/segregation and kinetochore assembly [2,27,179]. Moreover, recently satDNA transcripts have shown to be associated with male fertility in *Drosophila* sp. [180]. Unfortunately, the function of most satncRNAs remains

unknown or unclear due to the inefficient methodologies currently available to analyze molecules of such repetitive nature.
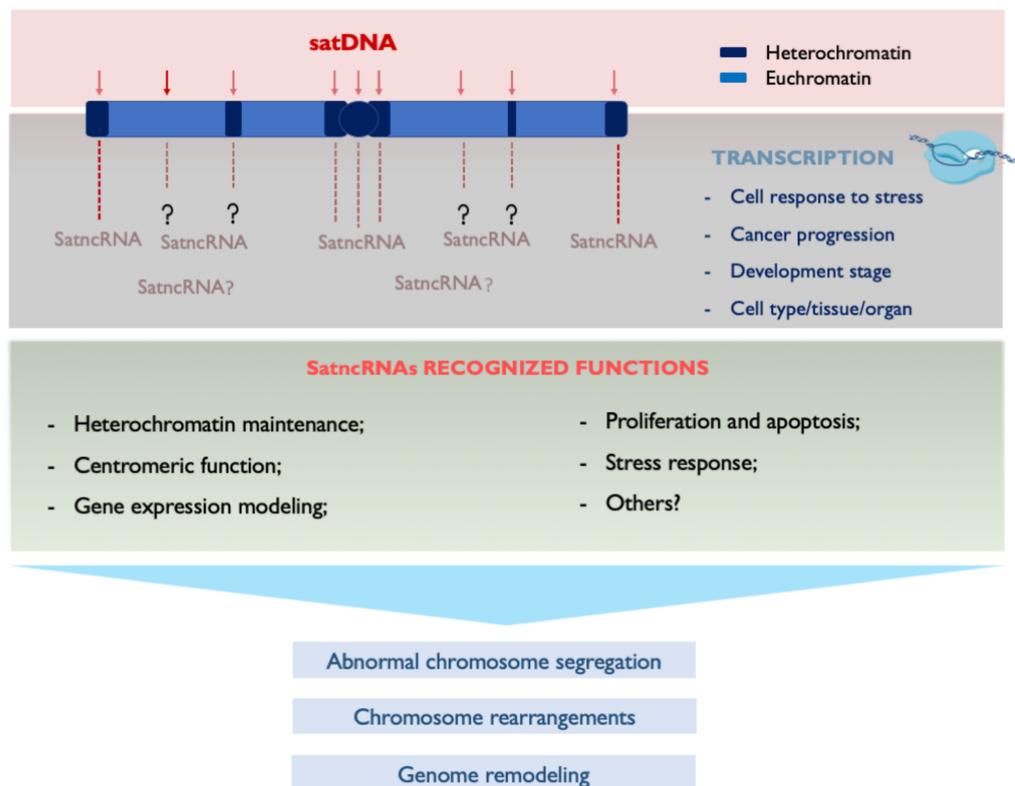
Figure 2. Summary of current knowledge regarding satellite non-coding RNAs (satncRNAs) and how they can contribute to genome remodeling. Even though satDNAs present in the heterochromatin and euchromatin can be transcribed, the most studied satncRNAs are the ones originated from pericentromeric and centromeric satDNAs families. For some satncRNAs reported, chromosome location of the origin satDNA cannot be determined. SatDNA transcription has been shown to be associated to cells response to stress, cancer progression, particular developmental stage and some are differentially expressed in specific cell types, tissues and organs. General recognized functions attributed to satncRNAs are listed. The aberrant expression of satncRNAs may result in abnormal chromosome segregation, and chromosome rearrangements that re-shape the genome and can lead to cancer progression or be fixed during species evolution. Further effort is needed to identify and better characterize satncRNA and their involvement in cellular functions and disease.

Concerning centromeric satDNAs, the human α satellite transcripts (αSAT) have been shown to be crucial for cell cycle progression, as depletion of αSAT resulted in defective centromeric protein A (CENP-A) loading and cell cycle arrest [181]. αSAT ncRNAs also seem to regulate spindle microtubule attachment and sister chromatid disjunction through association with AURORA B proteins [182]. These molecules were further shown to be associated with the SUV39H1 histone methyltransferase, thereby suggesting a regulatory function in heterochromatin maintenance [183–185].

Contrary to what is believed for most condensed genomic regions, centromeric sequences remain transcriptionally active during mitosis [186,187], essentially promoting kinetochore stabilization and centromere cohesion [188,189]. These functions have been similarly attributed to transcriptionally active centromeric satDNAs from other species [190–192], in spite of the observed sequence differences. This suggests that satncRNAs are involved in critical functions, which appear to be associated with their intrinsic molecular characteristics and most probably also with the genomic location of their satDNA sequence.

Pericentromeric satDNA transcripts have been related with pericentric chromatin formation [193–195], acting as molecular scaffolds for the accumulation of HP1 [194]. The presence of human *SATIII* ncRNA can be closely associated with cell response to stress. Particularly, heat shock can trigger *SATIII* transcription by the action of HSF1 (Heat Shock Factor 1) [196,197], giving rise to nuclear stress bodies (nSBs) close to *SATIII* DNA regions [198]. The splicing of relevant genes for stress response may be influenced by *SATIII* ncRNAs, which have been proposed to sequester RNA processing factors and downregulate global transcription [199], providing protection against stress-induced cell death [200]. However, *SATIII* transcription is not thermal stress-exclusive, as a basal level of expression is detectable even in the absence of cellular stress [201]. This same satDNA family can exhibit different genomic locations and its transcripts can be involved in multiple functions, making their study even more difficult. *SatDNA III* from *Drosophila melanogaster* is located at the centromere and pericentromere of the X chromosome and at the pericentromere of chromosomes 2 and 3 [202]. Its transcripts have been shown to play different roles in chromatin silencing/heterochromatinization, centromeric function and upregulation of X-linked genes [191,202–204].

Another interesting case is the *FA-SAT*, the major satDNA sequence of *Felis catus* (cat) genome, located at the (sub)telomeres and (peri)centromeres of chromosomes [153] and also in an interspersed fashion [24]. This satDNA is highly conserved in its primary sequence among Bilateria species (e.g., human, *Drosophila*, oyster, cattle, among others), a rare event observed in satDNA sequences [24]. In these species (non-*Felis* species), an interspersed distribution of this satDNA was proposed, with the exception of the other carnivore analyzed, *Genetta genetta*, in which it also presents a centromeric location. Of note, *FA-SAT* is transcribed in all these species [24], and an important conserved function (in cat and human) was ascribed to this ncRNA as a PKM2 interactor involved in the cross-talk between proliferation and apoptosis [205]. In fact, the absence of this satncRNA in both species results in cell death [205]. These transcripts possibly originate from the transcription of *FA-SAT* interspersed DNA (at current knowledge, the common location among these species). A putative connection between *FA-SAT* ncRNAs with cancer was also recently hypothesized [205].

With the progressive acceptance of satDNA transcriptional activity, the aberrant expression of satncRNAs has been increasingly associated with cancer progression (reviewed in [27]). In fact, the observed overexpression of satncRNAs in stress conditions may be comparable to cancer, since loss of sister chromatids cohesion, incorrect chromosome segregation or aneuploidy are common features of both states [206]. Overexpression of satncRNAs in cancer has been reported alongside decondensation and hypomethylation of pericentromeric DNA [154,207]. The transcription of satDNAs, the change in nuclear architecture and the altered sequestration of transcription factors may all be related to gene expression deregulation induced by the hypomethylation of *SATII* and *SATIII* satDNA sequences [208,209]. SatncRNAs may thus be involved in more general disease contexts associated to chromatin decondensation, DNA breaks and subsequent genomic rearrangements [209] (Figure 2). However, the value of satncRNAs as cancer biomarkers is still an unexplored field [27].

Due to their repetitive nature, high copy number and multiple genomic locations (different chromosomes and/or genomic regions), the study of satncRNAs remains a difficult challenge, namely regarding the original genomic location of their DNA sequence and the determination of their primary sequence(s) (Figure 2). Indeed, the most common next-generation sequencing platforms presents significant limitations in the analysis of satncRNA sequences in RNA-seq libraries, namely regarding their inability to assemble large repetitive transcripts from very short reads. This could be overcome in the future with the application of ultra-long read sequencing technology. The fact that most of the methods currently available that are mainly directed towards the analysis of gene coding sequences set a requirement for essential improvements or adjustments in order to support the efficient study of satncRNAs (Figure 1).

Although the importance of satncRNA in normal cell function and disease states is becoming increasingly accepted by the scientific community, in the wake of recent studies on these transcripts, much remains to be understood about their functions in different contexts. This will only be overcome

through the development of improved methods for the study of repetitive sequences, as well as the commitment of the scientific community to this field of research.

5. Concluding Remarks

In this review we outlined the critical importance of satDNA sequences in driving karyotype evolution and genomic architecture, as well as their involvement in various basic genomic functions. However, to this day, significant technological limitations hinder the progress of this important field in biology and medicine, in particular in the study of diseases involving this genomic fraction. It is imperative to boost the study of satDNA sequences and their transcripts by adapting and developing sequencing technologies and bioinformatics pipelines capable of assembling chromosomes from telomere-to-telomere as well as focused approaches that follow the concept of chromosomics. Significant effort is needed from the entire scientific community to value these important genomic elements, which have been so neglected over time. Only then can we begin to fully understand the largest fraction of our genome.

References

1. Charlesworth, B.; Sniegowski, P.; Stephan, L.W. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* **1994**, *371*, 215–220. [CrossRef] [PubMed]
2. Plohl, M.; Mestrovic, N.; Mravinac, B. Satellite DNA evolution. *Genome Dyn.* **2012**, *7*, 126–152. [CrossRef]
3. López-Flores, I.; Garrido-Ramos, M.A. The Repetitive DNA Content of Eukaryotic Genomes. *Genome Dyn.* **2012**, *7*, 1–28. [CrossRef] [PubMed]
4. Brajković, J.; Feliciello, I.; Bruvo-Mađarić, B.; Ugarković, D. Satellite DNA-like elements associated with genes within euchromatin of the beetle Tribolium castaneum. *G3 Genes Genomes Genet.* **2012**, *2*, 931–941. [CrossRef]
5. Kuhn, G.C.; Küttler, H.; Moreira-Filho, O.; Heslop-Harrison, J.S. The 1.688 repetitive DNA of *Drosophila*: Concerted evolution at different genomic scales and association with genes. *Mol. Biol. Evol.* **2012**, *29*, 7–11. [CrossRef] [PubMed]
6. Larracuente, 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*, 233. [CrossRef] [PubMed]
7. 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]
8. Feliciello, I.; Akrap, I.; Ugarkovic, D. Satellite DNA modulates gene expression in the beetle *Tribolium castaneum* after heat stress. *PLoS Genet.* **2015**, *11*, e1005466. [CrossRef]
9. Volpe, T.A.; Kidner, C.; Hall, I.M.; Teng, G.; Grewal, S.I.; Martienssen, R.A. Regulation of heterochromatic silencing and histone H_3 lysine-9 methylation by RNAi. *Science* **2002**, *297*, 1833–1837. [CrossRef]
10. Martienssen, R.A. Maintenance of heterochromatin by RNA interference of tandem repeats. *Nat. Genet.* **2003**, *35*, 213–214. [CrossRef]
11. Allshire, R.C.; Madhani, H.D. Ten Principles of Heterochromatin Formation and Function. *Nat. Rev. Mol. Cell Biol.* **2017**, *19*, 229–244. [CrossRef] [PubMed]
12. Garrido-Ramos, M.A. Satellite DNA: An Evolving Topic. *Genes* **2017**, *8*, 230. [CrossRef] [PubMed]
13. Ugarković, Đ. Evolution of α-Satellite DNA. In *Encyclopedia of Life Sciences*; John Wiley & Sons, Ltd.: Chichester, UK, 2013.

14. Adega, F.; Guedes-Pinto, H.; Chaves, R. Satellite DNA in the karyotype evolution of domestic animals—Clinical considerations. *Cytogenet. Genome Res.* **2009**, *126*, 12–20. [CrossRef] [PubMed]

15. Ruiz-Herrera, A.; Castresana, J.; Robinson, T.J. Is mammalian chromosomal evolution driven by regions of genome fragility? *Genome Biol.* **2006**, *7*. [CrossRef] [PubMed]

16. Farré, M.; Bosch, M.; López-Giráldez, F.; Ponsà, M.; Ruiz-Herrera, A. Assessing the role of tandem repeats in shaping the genomic architecture of great apes. *PLoS ONE* **2011**, *6*, e27239. [CrossRef] [PubMed]

17. Vieira-da-Silva, A.; Louzada, S.; Adega, F.; Chaves, R. A high-resolution comparative chromosome map of Cricetus cricetus and Peromyscus eremicus reveal the involvement of constitutive heterochromatin in breakpoint regions. *Cytogenet. Genome Res.* **2015**, *145*, 59–67. [CrossRef] [PubMed]

18. De La Fuente, R.; Baumann, C.; Viveiros, M.M. ATRX contributes to epigenetic asymmetry and silencing of major satellite transcripts in the maternal genome of the mouse embryo. *Development* **2015**, *142*, 1806–1817. [CrossRef]

19. Giunta, S.; Funabiki, H. Integrity of the human centromere DNA repeats is protected by CENP-A, CENP-C, and CENP-T. *Proc. Natl. Acad. Sci. USA* **2017**, *11*, 1928–1933. [CrossRef]

20. Chaves, R.; Adega, F.; Heslop-Harrison, J.S.; Guedes-Pinto, H.; Wienberg, J. Complex satellite DNA reshuffling in the polymorphic t (1; 29) Robertsonian translocation and evolutionarily derived chromosomes in cattle. *Chromosome Res.* **2003**, *11*, 641–648. [CrossRef]

21. Escudeiro, A.; Ferreira, D.; Mendes-da-Silva, A.; Heslop-Harrison, J.S.; Adega, F.; Chaves, R. Bovine satellite DNAs—A history of the evolution of complexity and impact in the *Bovidae* Family. *Eur. Zool. J.* **2019**, *86*, 20–37. [CrossRef]

22. Mravinac, B.; Plohl, M.; Mestrović, N.; Ugarković, D. Sequence of PRAT Satellite DNA 'Frozen' in Some Coleopteran Species. *J. Mol. Evol.* **2002**, *54*, 774–783. [CrossRef] [PubMed]

23. Plohl, M.; Petrović, V.; Luchetti, A.; Ricci, A.; Satović, E.; Passamonti, M.; Mantovani, B. Long-term conservation vs high sequence divergence: The case of an extraordinarily old satellite DNA in bivalve mollusks. *Heredity* **2010**, *104*, 543–551. [CrossRef] [PubMed]

24. Chaves, R.; Ferreira, D.; Mendes-da-Silva, A.; Meles, S.; Adega, F. FA-SAT Is an Old Satellite DNA Frozen in Several Bilateria Genomes. *Genome Biol. Evol.* **2017**, *9*, 3073–3087. [CrossRef] [PubMed]

25. Pezer, Z.; Brajković, J.; Feliciello, I.; Ugarkovć, D. Satellite DNA-Mediated Effects on Genome Regulation. *Genome Dyn.* **2012**, *7*, 153–169. [CrossRef] [PubMed]

26. Grenfell, A.W.; Heald, R.; Strzelecka, M. Mitotic noncoding RNA processing promotes kinetochore and spindle assembly in Xenopus. *J. Cell Biol.* **2016**, *214*, 133–141. [CrossRef] [PubMed]

27. Ferreira, D.; Meles, S.; Escudeiro, A.; Mendes-da-Silva, A.; Adega, F.; Chaves, R. Satellite non-coding RNAs: The emerging players in cells, cellular pathways and cancer. *Chromosome Res.* **2015**, *23*, 479–493. [CrossRef] [PubMed]

28. Ferreira, D.; Escudeiro, A.; Adega, F.; Anjo, S.I.; Manadas, B.; Chaves, R. FA-SAT ncRNA interacts with PKM2 protein: Depletion of this complex induces a switch from cell proliferation to apoptosis. *Cell Mol. Life Sci.* **2019**. [CrossRef]

29. Masumoto, H.; Masukata, H.; Muro, Y.; Nozaki, N.; Okazaki, T. Ahumancentromereantigen(CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J. Cell Biol.* **1989**, *109*, 1963–1973. [CrossRef]

30. Muro, Y.; Masumoto, H.; Yoda, K.; Nozaki, N.; Ohashi, M.; Okazaki, T. Centromereprotein Bassembles human centromeric α-satellite DNA at the 17-bp sequence, CENP-B box. *J. Cell Biol.* **1992**, *116*, 585–596. [CrossRef]

31. 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]

32. Kipling, D.; Warburton, P.E. Centromeres, CENP-B and Tigger too. *Trends Genet.* **1997**, *13*, 141–145. [CrossRef]

33. Kim, Y.B.; Oh, J.H.; McIver, L.J.; Rashkovetsky, E.; Michalak, K.; Garner, H.R.; Kang, L.; Nevo, E.; Korol, A.B.; Michalak, P. Divergence of *Drosophila* melanogaster repeatomes in response to a sharp microclimate contrast in Evolution Canyon, Israel. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10630–10635. [CrossRef] [PubMed]

34. 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] [PubMed]

35. Belyayev, A.; Josefiová, J.; Jandová, M.; Kalendar, R.; Krak, K.; Mandák, B. Natural History of a Satellite DNA Family: From the Ancestral Genome Component to Species-Specific Sequences, Concerted and Non-Concerted Evolution. *Int. J. Mol. Sci.* **2019**, *20*, 1201. [CrossRef]

36. Kit, S.J. Equilibrium Sedimentation in Density Gradients of DNA Preparations from Animal Tissues. *J. Mol. Biol.* **1961**, *3*, 711–716. [CrossRef]

37. Orgel, L.E.; Crick, F.H.C. Selfish DNA: The Ultimate Parasite. *Nature* **1980**, *284*, 604–607. [CrossRef]

38. Plohl, M.; Luchetti, A.; Mestrovic, N.; Mantovani, B. Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric (hetero) chromatin. *Gene* **2008**, *409*, 72–82. [CrossRef]

39. Willard, H.F. Chromosome-specific organization of human α satellite DNA. *Am. J. Hum. Genet.* **1985**, *37*, 524–532.

40. Alexandrov, I.A.; Medvedev, L.; Mashkova, T.D.; Kisselev, L.L.; Romanova, L.Y.; Yurov, Y.B. Definition of a new α satellite suprachromosomal family characterized by monomeric organization. *Nucleic Acids Res.* **1993**, *21*, 2209–2215. [CrossRef]

41. McNulty, S.M.; Sullivan, B.A. α satellite DNA biology: Finding function in the recesses of the genome. *Chromosome Res.* **2018**, *26*, 115–138. [CrossRef]

42. Waye, J.S.; Willard, H.F. Chromosome-specific α satellite DNA: Nucleotide sequence analysis of the 2.0 kilobasepair repeat from the human X chromosome. *Nucleic Acids Res.* **1985**, *13*, 2731–2743. [CrossRef] [PubMed]

43. Vlahovic, I.; Gluncic, M.; Rosandic, M.; Ugarkovic, Đ.; Paar, V. Regular Higher Order Repeat Structures in Beetle Tribolium castaneum Genome. *Genome Biol Evol.* **2017**, *9*, 2668–2680. [CrossRef] [PubMed]

44. Araújo, N.P.; de Lima, L.G.; Dias, G.B.; Kuhn, G.C.S.; de Melo, A.L.; Yonenaga-Yassuda, Y.; Stanyon, R.; Svartman, M. Identification and characterization of a subtelomeric satellite DNA in *Callitrichini monkeys*. *DNA Res.* **2017**, *24*, 377–385. [CrossRef] [PubMed]

45. Utsunomia, R.; Ruiz-Ruano, F.J.; Silva, D.M.Z.A.; Serrano, É.A.; Rosa, I.F.; Scudeler, P.E.S.; Hashimoto, D.T.; Oliveira, C.; Camacho, J.P.M.; Foresti, F. A Glimpse into the Satellite DNA Library in *Characidae* Fish (*Teleostei, Characiformes*). *Front. Genet.* **2017**, *8*, 103. [CrossRef]

46. Adega, F.; Chaves, R.; Guedes-Pintos, H. Chromosomal evolution and phylogenetic analysis in *Tayassu pecari* and *Pecari tajacu* (*Tayassuidae*): Tales from constitutive heterochromatin. *J. Genet.* **2007**, *86*, 19–26. [CrossRef]

47. Henikoff, S.; Dalal, Y. Centromeric chromatin: What makes it unique? *Curr. Opin. Genet. Dev.* **2005**, *15*, 177–184. [CrossRef]

48. Paço, A.; Adega, F.; Guedes-Pinto, H.; Chaves, R. Hidden heterochromatin Characterization in the *Rodentia* species *Cricetus cricetus*, *Peromyscus eremicus* (*Cricetidae*) and *Praomys tullbergi* (*Muridae*). *Genet. Mol. Biol.* **2009**, *32*, 58–68. [CrossRef]

49. Ugarković, Đ.; Plohl, M. Variation in satellite DNA profiles—Causes and effects. *EMBO J.* **2002**, *21*, 5955–5959. [CrossRef]

50. Louzada, S.; Vieira-da-Silva, A.; Mendes-da-Silva, A.; Kubickova, S.; Rubes, J.; Adega, F.; Chaves, R. A Novel Satellite DNA Sequence in the Peromyscus Genome (PMSat): Evolution via Copy Number Fluctuation. *Mol. Phylogenet. Evol.* **2015**, *92*, 193–203. [CrossRef]

51. Gall, J.G.; Cohen, E.G.; Polan, M.L. Repetitive DNA sequences in *Drosophila*. *Chromosoma* **1971**, *33*, 319–344. [CrossRef]

52. Lohe, A.R.; Brutlag, D.L. Identical satellite DNA sequences in sibling species of *Drosophila*. *J. Mol. Biol.* **1987**, *194*, 161–170. [CrossRef]

53. Wei, K.H.-C.; Lower, S.E.; Caldas, I.V.; Sless, T.J.S.; Barbash, D.A.; Clark, A.G. Variable rates of simple satellite gains across the *Drosophila* phylogeny. *Mol. Biol. Evol.* **2018**, *35*, 925–941. [CrossRef] [PubMed]

54. 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]

55. Ferree, P.M.; Prasad, S. How can satellite DNA divergence cause reproductive isolation? Let us count the chromosomal ways. *Genet Res Int* **2012**, 430136. [CrossRef]

56. Walsh, J.B. Persistence of tandem arrays: Implications for satellite and simple-sequence DNAs. *Genetics* **1987**, *115*, 553–567.

57. Adega, F.; Chaves, R.; Guedes-Pinto, H. Suiformes orthologous satellite DNAs as a hallmark of collared and white-lipped peccaries (*Tayassuidae*) evolutionary rearrangements. *Micron* **2008**, *39*, 1281–1287. [CrossRef]

58.　Dover, G.A. Molecular drive in multigene families: How biological novelties arise, spread and are assimilated. *Trends Genet.* **1986**, *2*, 159–165. [CrossRef]

59.　Elder, J.F., Jr.; Turner, B.J. Concerted evolution of repetitive DNA sequences in eukaryotes. *Q. Rev. Biol.* **1995**, *70*, 297–320. [CrossRef]

60.　Lewin, H.A.; Robinson, G.E.; Kress, W.J.; Baker, W.J.; Coddington, J.; Crandall, K.A.; Durbin, R.; Edwards, S.V.; Forest, F.; Gilbert, M.T.P.; et al. Earth BioGenome Project: Sequencing life for the future of life. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4325–4333. [CrossRef]

61.　Treangen, T.J.; Salzberg, S.L. Repetitive DNA and next-generation sequencing: Computational challenges and solutions. *Nat. Rev. Genet.* **2011**, *13*, 36–46. [CrossRef]

62.　Van Dijk, E.L.; Jaszczyszyn, Y.; Naquin, D.; Thermes, C. The Third Revolution in Sequencing Technology. *Trends Genet.* **2018**, *34*, 666–681. [CrossRef] [PubMed]

63.　Sedlazeck, F.J.; Lee, H.; Darby, C.A.; Schartz, M.C. Piercing the dark matter: Bioinformatics of long-range sequencing and mapping. *Nat. Rev. Genet.* **2018**, *19*, 329–346. [CrossRef] [PubMed]

64.　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]

65.　Lower, S.S.; McGurk, M.P.; Clark, A.G.; Barbash, D.A. Satellite DNA Evolution: Old Ideas, New Approaches. *Curr. Opin. Genet. Dev.* **2018**, *49*, 70–78. [CrossRef] [PubMed]

66.　Salmela, L.; Rivals, E. LoRDEC: Accurate and efficient long read error correction. *Bioinformatics* **2014**, *30*, 3506–3514. [CrossRef]

67.　Cook, D.E.; Zdraljevic, S.; Tanny, R.E.; Seo, B.; Riccardi, D.D.; Noble, L.M.; Rockman, M.V.; Alkema, M.J.; Braendle, C.; Kammenga, J.E.; et al. The Genetic Basis of Natural Variation in Caenorhabditis elegans Telomere Length. *Genetics* **2016**, *204*, 371–383. [CrossRef]

68.　Wei, K.H.; 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]

69.　Silva, D.M.Z.A.; Utsunomia, R.; Ruiz-Ruano, F.J.; Daniel, S.N.; Porto-Foresti, F.; Hashimoto, D.T.; Oliveira, C.; Camacho, J.P.M.; Foresti, F. High-throughput analysis unveils a highly shared satellite DNA library among three species of fish genus Astyanax. *Sci. Rep.* **2017**, *7*, 12726. [CrossRef]

70.　Utsunomia, R.; Silva, D.M.Z.A.; Ruiz-Ruano, F.J.; Goes, C.A.G.; Melo, S.; Ramos, L.P.; Oliveira, C.; Porto-Foresti, F.; Foresti, F.; Hashimoto, D.T. Satellitome landscape analysis of *Megaleporinus macrocephalus* (*Teleostei, Anostomidae*) reveals intense accumulation of satellite sequences on the heteromorphic sex chromosome. *Sci. Rep.* **2019**, *9*, 5856. [CrossRef]

71.　Palacios-Gimenez, O.M.; Dias, G.B.; De Lima, L.G.; Ramos, É.; Martins, C.; Cabral-de-Mello, D.C. High-throughput analysis of the satellitome revealed enormous diversity of satellite DNAs in the neo-Y chromosome of the cricket *Eneoptera surinamensis*. *Sci. Rep.* **2017**, *7*, 6422. [CrossRef]

72.　Usai, G.; Mascagni, F.; Natali, L.; Giordani, T.; Cavallini, A. Comparative Genome-Wide Analysis of Repetitive DNA in the *Genus Populus* L. *Tree Genet. Genomes* **2017**, *13*, 96. [CrossRef]

73.　Vozdova, M.; Kubickova, S.; Cernohorska, H.; Fröhlich, J.; Rubes, J. Satellite DNA Sequences in Canidae and Their Chromosome Distribution in Dog and Red Fox. *Cytogenet. Genome Res.* **2017**, *150*, 118–127. [CrossRef] [PubMed]

74.　Miga, K.H. Completing the human genome: The progress and challenge of satellite DNA assembly. *Chromosome Res.* **2015**, *23*, 421–426. [CrossRef] [PubMed]

75.　Jain, M.; Olsen, H.E.; Paten, B.; Akeson, M. The Oxford Nanopore MinION: Delivery of nanopore sequencing to the genomics community. *Genome Biol.* **2016**, *17*, 239. [CrossRef]

76.　Kuderna, L.F.K.; Lizano, E.; Julià, E.; Gomez-Garrido, J.; Serres-Armero, A.; Kuhlwilm, M.; Alandes, R.A.; Alvarez-Estape, M.; Juan, D.; Simon, H.; et al. Selective Single Molecule Sequencing and Assembly of a Human Y Chromosome of African Origin. *Nat. Commun.* **2019**, *10*, 4. [CrossRef]

77.　Miga, K.H. Centromeric Satellite DNAs: Hidden Sequence Variation in the Human Population. *Genes* **2019**, *10*, 352. [CrossRef]

78.　Jain, M.; Koren, S.; Miga, K.H.; Quick, J.; Rand, A.C.; Sasani, T.A.; Tyson, J.R.; Beggs, A.D.; Dilthey, A.T.; Fiddes, I.T.; et al. Nanopore Sequencing and Assembly of a Human Genome with Ultra-Long Reads. *Nat. Biotechnol.* **2018**, *36*, 338–345. [CrossRef]

79. 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. *bioRxiv* **2019**, *735928*. [CrossRef]

80. Muller, H.; Gil, J., Jr.; Drinnenberg, I.A. The impact of centromeres on spatial genome architecture. *Trends Genet.* **2019**, *35*, 565–578. [CrossRef]

81. 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]

82. Weissensteiner, M.H.; Pang, A.W.C.; Bunikis, I.; Höijer, I.; Vinnere-Petterson, O.; Suh, A.; Wolf, J.B.W. Combination of Short-Read, Long-Read, and Optical Mapping Assemblies Reveals Large-Scale Tandem Repeat Arrays with Population Genetic Implications. *Genome Res.* **2017**, *27*, 697–708. [CrossRef] [PubMed]

83. Ruiz-Ruano, F.J.; Cuadrado, Á.; Montiel, E.E.; Camacho, J.P.; López-León, M.D. Next Generation Sequencing and FISH Reveal Uneven and Nonrandom Microsatellite Distribution in Two Grasshopper Genomes. *Chromosoma* **2015**, *124*, 221–234. [CrossRef] [PubMed]

84. Skinner, B.M.; Sargent, C.A.; Churcher, C.; Hunt, T.; Herrero, J.; Loveland, J.E.; Dunn, M.; Louzada, S.; Fu, B.; Chow, W.; et al. The Pig X and Y Chromosomes: Structure, Sequence, and Evolution. *Genome Res.* **2016**, *26*, 130–139. [CrossRef] [PubMed]

85. Louzada, S.; Komatsu, J.; Yang, F. Fluorescence in situ hybridization onto DNA fibres generated using molecular combing. In *Fluorescence In Situ Hybridization (FISH)*; Liehr, T., Ed.; Springer: Berlin/Heidelberg, Germany, 2017; pp. 275–293.

86. Giunta, S. Centromere Chromosome Orientation Fluorescent In Situ Hybridization (Cen-CO-FISH) Detects Sister Chromatid Exchange at the Centromere in Human Cells. *Bio-protocol* **2018**, *8*. [CrossRef]

87. Saha, A.; Mourad, M.; Kaplan, M.H.; Chefetz, I.; Malek, S.N.; Buckanovich, R.; Markovitz, D.M.; Contreras-Galindo, R. The Genomic Landscape of Centromeres in Cancers. *bioRxiv* **2019**, *505800*. [CrossRef]

88. White, T.B.; McCoy, A.M.; Streva, V.A.; Fenrich, J.; Deininger, P.L. A droplet digital PCR detection method for rare L1 insertions in tumors. *Mobile DNA* **2014**, *5*, 30. [CrossRef]

89. Kishikawa, T.; Otsuka, M.; Yoshikawa, T.; Ohno, M.; Yamamoto, K.; Yamamoto, N.; Yamamoto, N.; Kotani, A.; Koike, K. Quantitation of circulating satellite RNAs in pancreatic cancer patients. *JCI Insight* **2016**, *1*, e86646. [CrossRef]

90. Deakin, J.E.; Potter, S.; O'Neill, R.; Ruiz-Herrera, A.; Cioffi, M.B.; Eldridge, M.D.B.; Fukui, K.; Marshall Graves, J.A.; Griffin, D.; Grutzner, F.; et al. Chromosomics: Bridging the Gap between Genomes and Chromosomes. *Genes* **2019**, *10*, 627. [CrossRef]

91. Barski, A.; Cuddapah, S.; Cui, K.; Roh, T.Y.; Schones, D.E.; Wang, Z.; Wei, G.; Chepelev, I.; Zhao, K. High-resolution profiling of histone methylations in the human genome. *Cell* **2007**, *129*, 823–837. [CrossRef]

92. Palacios-Gimenez, O.M.; Bardella, V.B.; Lemos, B.; Cabral-de-Mello, D.C. Satellite DNAs are conserved and differentially transcribed among *Gryllus* cricket species. *DNA Res.* **2018**, *25*, 137–147. [CrossRef]

93. Kowar, T.; Zakrzewski, F.; Macas, J.; Kobližková, A.; Viehoever, P.; Weisshaar, B.; Schmidt, T. Repeat Composition of CenH$_3$-chromatin and H$_3$K$_9$me$_2$-marked heterochromatin in Sugar Beet (β vulgaris). *BMC Plant Biol.* **2016**, *16*, 120. [CrossRef] [PubMed]

94. Su, H.; Liu, Y.; Liu, C.; Shi, Q.; Huang, Y.; Han, F. Centromere satellite repeats have undergone rapid changes in polyploid wheat subgenomes. *Plant Cell* **2019**, *31*, 2035–2051. [CrossRef] [PubMed]

95. Khost, D.E.; Eickbush, D.G.; Larracuente, A.M. Single-Molecule Sequencing Resolves the Detailed Structure of Complex Satellite DNA Loci in *Drosophila*. *Genome Res.* **2017**, *27*, 709–721. [CrossRef] [PubMed]

96. De Lima, L.G.; Svartman, M.; Kuhn, C.S. Dissecting the Satellite DNA Landscape in Three Cactophilic *Drosophila* Sequenced Genomes. *G3 Genes Genomes Genet.* **2017**, *7*, 2831–2843. [CrossRef]

97. Flynn, J.M.; Long, M.; Wing, R.A.; Clark, A.G. Evolutionary dynamics of abundant 7 bp satellites in the genome of *Drosophila virilis*. *bioRxiv* **2019**, *693077*. [CrossRef]

98. Hartley, G.; O'neill, R.J. Centromere Repeats: Hidden Gems of the Genome. *Genes* **2019**, *10*, 223. [CrossRef]

99. Ferguson-Smith, M.A.; Trifonov, V. Mammalian karyotype evolution. *Nat. Rev. Genet.* **2007**, *8*, 950–962. [CrossRef]

100. Spielmann, M.; Lupiáñez, D.G.; Mundlos, S. Structural variation in the 3D genome. *Nat. Rev. Genet.* **2018**, *19*, 453–467. [CrossRef]

101. Dobzhansky, T. *Genetics of the Evolutionary Process*; Columbia University Press: New York, NY, USA, 1970.

102. Wichman, H.; Payne, C.; Ryder, O.; Hamilton, M.; Maltbie, M.; Baker, R. Genomic distribution of heterochromatic sequences in equids: Implications to rapid chromosomal evolution. *J. Hered.* **1991**, *82*, 369–377. [CrossRef]

103. Rossi, M.; Redi, C.; Viale, G.; Massarini, A.; Capanna, E. Chromosomal distribution of the major satellite DNA of South American rodents of the genus Ctenomys. *Cytogenet. Cell Genet.* **1995**, *69*, 179–184. [CrossRef]

104. Chaves, R.; Guedes-Pinto, H.; Heslop-Harrison, J.S. Phylogenetic relationships and the primitive X chromosome inferred from chromosomal and satellite DNA analysis in Bovidae. *Proc. Biol. Sci.* **2005**, *272*, 2009–2016. [CrossRef] [PubMed]

105. Adega, F.; Chaves, R.; Guedes-Pinto, H.; Heslop-Harrison, J. Physical organization of the 1.709 satellite IV DNA family in Bovini and Tragelaphini tribes of the Bovidae: Sequence and chromosomal evolution. *Cytogenet. Genome Res.* **2006**, *114*, 140–146. [CrossRef] [PubMed]

106. Maynard Smith, J.; Burian, R.; Kaufman, S.; Alberch, P.; Campbell, J.; Goodwin, B. Developmental constraints and evolution. *Q. Rev. Biol.* **1985**, *60*, 265–287. [CrossRef]

107. Farré, M.; Robinson, T.J.; Ruiz-Herrera, A. An Integrative Breakage Model of genome architecture, reshuffling and evolution: The Integrative Breakage Model of genome evolution, a novel multidisciplinary hypothesis for the study of genome plasticity. *Bioessays* **2015**, *37*, 479–488. [CrossRef] [PubMed]

108. Kehrer-Sawatzki, H.; Cooper, D.N. Molecular mechanisms of chromosomal rearrangement during primate evolution. *Chromosome Res.* **2008**, *16*, 41–56. [CrossRef]

109. Carbone, L.; Harris, R.A.; Vessere, G.M.; Mootnick, A.R.; Humphray, S.; Rogers, J.; Kim, S.K.; Wall, J.D.; Martin, D.; Jurka, J.; et al. Evolutionary breakpoints in the gibbon suggest association between cytosine methylation and karyotype evolution. *PLoS Genet.* **2009**, *5*, e1000538. [CrossRef]

110. Zhao, H.; Bourque, G. Recovering genome rearrangements in the mammalian phylogeny. *Genome Res.* **2009**, *19*, 934–942. [CrossRef]

111. Paço, A.; Chaves, R.; Vieira-da-Silva, A.; Adega, F. The involvement of repetitive sequences in the remodelling of karyotypes: The *Phodopus* genomes (*Rodentia*, *Cricetidae*). *Micron* **2013**, *46*, 27–34. [CrossRef]

112. Bourque, G. Transposable elements in gene regulation and in the evolution of vertebrate genomes. *Curr. Opin. Genet. Dev.* **2009**, *19*, 607–612. [CrossRef]

113. Longo, M.S.; Carone, D.M.; Green, E.D.; O'Neill, M.J.; O'Neill, R.J.; NISC Comparative Sequencing Program. Distinct retroelement classes define evolutionary breakpoints demarcating sites of evolutionary novelty. *BMC Genomics* **2009**, *10*, 334. [CrossRef]

114. Garcia-Perez, J.L.; Widmann, T.J.; Adams, I.R. The impact of transposable elements on mammalian development. *Development* **2016**, *143*, 4101–4114. [CrossRef] [PubMed]

115. Platt, R.N.; Vandewege, M.W.; Ray, D.A. Mammalian transposable elements and their impacts on genome evolution. *Chromosome Res.* **2018**, *26*, 25–43. [CrossRef] [PubMed]

116. Bailey, J.A.; Eichler, E.E. Primate segmental duplications: Crucibles of evolution, diversity and disease. *Nat. Rev. Genet.* **2006**, *7*, 552–564. [CrossRef] [PubMed]

117. Louzada, S.; Paço, A.; Kubickova, S.; Adega, F.; Guedes-Pinto, H.; Rubes, J.; Chaves, R. Different evolutionary trails in the related genomes *Cricetus cricetus* and *Peromyscus eremicus* (*Rodentia*, *Cricetidae*) uncovered by orthologous satellite DNA repositioning. *Micron* **2008**, *39*, 1149–1155. [CrossRef] [PubMed]

118. Mashkova, T.; Oparina, N.; Alexandrov, I.; Zinovieva, O.; Marusina, A.; Yurov, Y.; Lacroix, M.H.; Kisselev, L. Unequal cross-over is involved in human α satellite DNA rearrangements on a border of the satellite domain. *FEBS Lett.* **1998**, *441*, 451–457. [CrossRef]

119. Kelkar, Y.D.; Tyekucheva, S.; Chiaromonte, F.; Makova, K.D. The genome-wide determinants of human and chimpanzee microsatellite evolution. *Genome Res.* **2008**, *18*, 30–38. [CrossRef]

120. Sankoff, D. The where and wherefore of evolutionary breakpoints. *J. Biol.* **2009**, *8*, 66. [CrossRef]

121. Catasti, P.; Chen, X.; Mariappan, S.V.; Bradbury, E.M.; Gupta, G. DNA repeats in the human genome. *Genetica* **1999**, *106*, 15–36. [CrossRef]

122. Yang, D.; Okamoto, K. Structural insights into G-quadruplexes: Towards new anticancer drugs. *Future Med. Chem.* **2010**, *2*, 619–646. [CrossRef]

123. Zhang, Z.; Dai, J.; Veliath, E.; Jones, R.A.; Yang, D.Z. Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K+ solution: Insights into the interconversion of human telomeric G-quadruplex structures. *Nucleic Acids Res.* **2010**, *38*, 1009–1021. [CrossRef]

124. Zhao, J.; Bacolla, A.; Wang, G.; Vasquez, K.M. Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* **2010**, *67*, 43–62. [CrossRef] [PubMed]

125. Black, E.M.; Giunta, S. Repetitive Fragile Sites: Centromere Satellite DNA as a Source of Genome Instability in Human Diseases. *Genes* **2018**, *9*, 615. [CrossRef] [PubMed]

126. Khodaverdian, V.Y.; Hanscom, T.; Yu, A.M.; Yu, T.L.; Mak, V.; Brown, A.J.; Roberts, S.A.; McVey, M. Secondary structure forming sequences drive SD-MMEJ repair of DNA double-strand breaks. *Nucleic Acids Res.* **2017**, *45*, 12848–12861. [CrossRef] [PubMed]

127. De Sotero-Caio, C.G.; Cabral-de-Mello, D.C.; Calixto, M.D.S.; Valente, G.T.; Martins, C.; Loreto, V.; de Souza, M.J.; Santos, N. Centromeric enrichment of LINE-1 retrotransposons and its significance for the chromosome evolution of *Phyllostomid bats*. *Chromosome Res.* **2017**, *25*, 313–325. [CrossRef]

128. Klein, S.J.; O'Neill, R.J. Transposable elements: Genome innovation, chromosome diversity, and centromere conflict. *Chromosome Res.* **2018**, *26*, 5–23. [CrossRef]

129. Mayorov, V.I.; Adkison, L.R.; Vorobyeva, N.V.; Khrapov, E.A.; Kholodhov, N.G.; Rogozin, I.B.; Nesterova, T.B.; Protopopov, A.I.; Sablina, O.V.; Graphodatsky, A.S.; et al. Organization and chromosomal localization of a B1-like containing repeat of Microtus subarvalis. *Mamm. Genome* **1996**, *7*, 593. [CrossRef]

130. Marchal, J.A.; Acosta, M.J.; Bullejos, M.; Puerma, E.; Díaz de la Guardia, R.; Sánchez, A. Distribution of L1-retroposons on the giant sex chromosomes of *Microtus cabrerae* (*Arvicolidae*, *Rodentia*): Functional and evolutionary implications. *Chromosome Res.* **2006**, *14*, 177–186. [CrossRef]

131. Escudeiro, A.; Adega, F.; Robinson, T.J.; Heslop-Harrison, J.S.; Chaves, R. Conservation, divergence and functions of centromeric satellite DNA families in the *Bovidae*. *Genome Biol. Evol.* **2019**, *11*, 1152–1165. [CrossRef]

132. Wessler, S.R. Transposable elements and the evolution of eukaryotic genomes. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17600–17601. [CrossRef]

133. Lim, J.K.; Simmons, J.M. Gross chromosome rearrangements mediated by transposable elements in *Drosophila melanogaster*. *Bioessays* **1994**, *16*, 269–275. [CrossRef]

134. Gray, Y.H.M. It takes two transposons to tango: Transposable-element-mediated chromosomal rearrangements. *Trends Genet.* **2000**, *16*, 461–468. [CrossRef]

135. Hedges, D.J.; Deininger, P.L. Inviting instability: Transposable elements, double-strand breaks, and the maintenance of genome integrity. *Mutat. Res.* **2007**, *616*, 46–59. [CrossRef] [PubMed]

136. Carbone, L.; Harris, R.A.; Gnerre, S.; Veeramah, K.R.; Lorente-Galdos, B.; Huddleston, J.; Meyer, T.J.; Herrero, J.; Roos, C.; Aken, B.; et al. Gibbon genome and the fast karyotype evolution of small apes. *Nature* **2014**, *513*, 195. [CrossRef] [PubMed]

137. Bailey, J.A.; Baertsch, R.; Kent, W.J.; Haussler, D.; Eichler, E.E. Hotspots of mammalian chromosomal evolution. *Genome Biol.* **2004**, *5*, R23. [CrossRef] [PubMed]

138. Cordaux, R.; Batzer, M.A. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* **2009**, *10*, 691–703. [CrossRef]

139. Warren, I.; Naville, M.; Chalopin, D.; Levin, P.; Berger, C.; Galiana, D.; Volff, J.N. Evolutionary impact of transposable elements on genomic diversity and lineage-specific innovation in vertebrates. *Chromosome Res.* **2015**, *23*, 505–531. [CrossRef]

140. Meštrović, N.; Mravinac, B.; Pavlek, M.; Vojvoda-Zeljko, T.; Šatović, E.; Plohl, M. Structural and functional liaisons between transposable elements and satellite DNAs. *Chrom. Res.* **2015**, *23*, 583–596. [CrossRef]

141. Paço, A.; Adega, F.; Chaves, R. Line-1 Retrotransposons: From "parasite" sequences to Functional Elements. *J. Appl. Genet.* **2015**, *56*, 133–145. [CrossRef]

142. Slamovits, C.H.; Rossi, M.S. Satellite DNA: Agent of chromosomal evolution in mammals. *A review. J. Neotrop. Mammal.* **2002**, *9*, 297–308.

143. Chaves, R.; Louzada, S.; Meles, S.; Wienberg, J.; Adega, F. *Praomys tullbergi* (*Muridae*, *Rodentia*) genome architecture decoded by comparative chromosome painting with *Mus* and *Rattus*. *Chromosome Res.* **2012**, *20*, 673–683, Epub 31 July 2012. [CrossRef]

144. Chaves, R.; Heslop-Harrison, J.S.; Guedes-Pinto, H. Centromeric heterochromatin in the cattle rob (1;29) translocation: α-satellite I sequences, In-Situ MspI digestion patterns, chromomycin staining and C-bands. *Chromosome Res.* **2000**, *8*, 621–626. [CrossRef] [PubMed]

145. Di Meo, G.P.; Perucatti, A.; Chaves, R.; Adega, F.; De Lorenzi, L.; Molteni, L.; De Giovanni, A.; Incarnato, D.; Guedes-Pinto, H.; Eggen, A.; et al. Cattle rob(1;29) originating from complex chromosome rearrangements as revealed by both banding and FISH-mapping techniques. *Chromosome Res.* **2006**, *14*, 649–655. [CrossRef] [PubMed]

146. Smalec, B.M.; Heider, T.N.; Flynn, B.L.; O'Neil, R.J. A centromere satellite concomitant with extensive karyotypic diversity across the *Peromyscus* genus defies predictions of molecular drive. *Chromosome Res.* **2019**, *27*, 237–252. [CrossRef] [PubMed]

147. Adega, F.; Matoso, S.R.; Kjöllerström, H.; Vercammen, P.; Raudsepp, T.; Collares-Pereira, M.J.; Fernandes, C.; Oom, M.; Chaves, R. Comparative chromosome painting in genets (*Carnivora*, *Viverridae*, *Genetta*), the only known feliforms with a highly rearranged karyotype. *Cytogenet. Genome Res.* **2018**, *156*, 35–44. [CrossRef] [PubMed]

148. Caceres, M.; Ranz, J.M.; Barbadilla, A.; Long, M.; Ruiz, A. Generation of a widespread *Drosophila* inversion by a transposable element. *Science* **1999**, *285*, 415–418. [CrossRef] [PubMed]

149. Rojo, V.; Martínez-Lage, A.; Giovannotti, M.; González-Tizón, A.M.; Cerioni, P.N.; Barucchi, V.C.; Galán, P.; Olmo, E.; Naveira, H. Evolutionary dynamics of two satellite DNA families in rock lizards of the genus *Iberolacerta* (*Squamata*, *Lacertidae*): Different histories but common traits. *Chromosome Res.* **2015**, *23*, 441–461. [CrossRef] [PubMed]

150. Dela Paz, J.S.; Stronghill, P.E.; Douglas, S.J.; Saravia, S.; Hasenkampf, C.A.; Riggs, C.D. Chromosome fragile sites in *Arabidopsis* harbor matrix attachment regions that may be associated with ancestral chromosome rearrangement events. *PLoS Genet.* **2012**, *8*, e1003136. [CrossRef]

151. Chaves, R.; Santos, S.; Guedes-Pinto, H. Comparative analysis (*Hippotragini* versus *Caprini*, *Bovidae*) of X-chromosome's constitutive heterochromatin by in situ restriction endonuclease digestion: X-chromosome constitutive heterochromatin evolution. *Genetica* **2004**, *121*, 315–325. [CrossRef]

152. Biscotti, M.A.; Olmo, E.; Heslop-Harrison, J.S. Repetitive DNA in Eukaryotic Genomes. *Chromosome Res.* **2015**, *23*, 415–420. [CrossRef]

153. Santos, S.; Chaves, R.; Guedes-Pinto, H. Chromosomal localization of the major satellite DNA family (FA-SAT) in the domestic cat. *Cytogenet. Genome Res.* **2004**, *107*, 119–122. [CrossRef]

154. 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] [PubMed]

155. Zhu, Q.; Pao, G.; Huynh, A.; Suh, H.; Tonnu, N.; Nederlof, P.M.; Gage, F.H.; Verma, I.M. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **2011**, *477*, 179–184. [CrossRef] [PubMed]

156. Smith, K.S.; Liu, L.L.; Ganesan, S.; Michor, F.; De, S. Nuclear topology modulates the mutational landscapes of cancer genomes. *Nat. Struct. Mol. Biol.* **2017**, *24*, 1000–1006. [CrossRef] [PubMed]

157. Guo, C.; Jeong, H.H.; Hsieh, Y.C.; Klein, H.U.; Bennett, D.A.; De Jager, P.L.; Liu, Z.; Shulman, J.M. Tau Activates Transposable Elements in Alzheimer's Disease. *Cell Rep.* **2018**, *23*, 2874–2880. [CrossRef]

158. Francastel, C.; Magdinier, F. DNA methylation in satellite repeats disorders. *Essays Biochem.* **2019**. [CrossRef]

159. Payer, L.M.; Burns, K.H. Transposable elements in human genetic disease. *Nat. Rev. Genet.* **2019**, *20*, 760–772. [CrossRef]

160. Fournier, A.; Mcleer-florin, A.; Lefebvre, C.; Duley, S.; Debernardi, A.; Rousseaux, S.; De Fraipont, F.; Figeac, M.; Kerckaert, J.; De Vos, J.; et al. 1q12 chromosome translocations form aberrant heterochromatic foci associated with changes in nuclear architecture and gene expression in B cell lymphoma. *EMBO Mol. Med.* **2010**, *2*, 159–171. [CrossRef]

161. Ducos, A.; Revay, T.; Kovacs, A.; Hidas, A.; Pinton, A.; Bonnet-Garnier, A.; Molteni, L.; Slota, E.; Switonski, M.; Arruga, M.V.; et al. Cytogenetic screening of livestock populations in Europe: An overview. *Cytogenet. Genome Res.* **2008**, *120*, 26–41. [CrossRef]

162. Iannuzzi, L.; King, W.; Di Berardino, D. Chromosome evolution in domestic bovids as revealed by chromosome banding and FISH-mapping techniques. *Cytogenet. Genome Res.* **2009**, *126*, 49–62. [CrossRef]

163. Therman, E.; Susman, B.; Denniston, C. The nonrandom participation of human acrocentric chromosomes in Robertsonian translocations. *Ann. Hum. Genet.* **1989**, *53*, 49–65. [CrossRef]

164. Page, S.L.; Shin, J.C.; Han, J.Y.; Choo, K.A.; Shaffer, L.G. Breakpoint diversity illustrates distinct mechanisms for Robertsonian translocation formation. *Hum. Mol. Genet.* **1996**, *5*, 1279–1288. [CrossRef] [PubMed]

165. Garagna, S.; Marziliano, N.; Zuccotti, M.; Searle, J.B.; Capanna, E.; Redi, C.A. Pericentromeric organization at the fusion point of mouse Robertsonian translocation chromosomes. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 171–175. [CrossRef] [PubMed]

166. Gustavsson, I. Cytogenetics, distribution and phenotypic effects of a translocation in *Swedish cattle*. *Hereditas* **1969**, *63*, 68–169. [CrossRef] [PubMed]

167. Dyrendahl, I.; Guatavsson, I. Sexual functions, semen characteristics and fertility of bulls carrying the 1/29 chromosome translocation. *Hereditas* **1979**, *90*, 281–289. [CrossRef]

168. Rangel-Figueiredo, T.; Iannuzzi, L. A cattle breed close to 58 diploid number due to high frequency of rob (1; 29). *Hereditas* **1991**, *115*, 73–78. [CrossRef]

169. Rangel-Figueiredo, T.; Iannuzzi, L. Frequency and distribution of rob (1; 29) in three *Portuguese cattle* breeds. *Hereditas* **1993**, *119*, 233–237. [CrossRef]

170. Rubes, J.; Kubickova, S.; Pagacova, E.; Cernohorska, H.; Di Berardino, D.; Antoninova, M.; Vahala, J.; Robinson, T.J. Phylogenomic study of spiral-horned antelope by cross-species chromosome painting. *Chromosome Res.* **2008**, *16*, 935–947. [CrossRef]

171. Gravholt, C.H.; Friedrich, U.; Caprani, M.; Jørgensen, A.L. Breakpoints in Robertsonian translocations are localized to satellite III DNA by fluorescence in situ hybridization. *Genomics* **1992**, *14*, 924–930. [CrossRef]

172. Wang, B.; Nie, B.; Tang, D.; Li, R.; Liu, X.; Song, J.; Wang, W.; Liu, Z. Analysis of meiotic segregation patterns and interchromosomal effects in sperm from 13 robertsonian translocations. *Balk. J. Med. Genet.* **2017**, *20*, 43–50. [CrossRef]

173. Kim, S.R.; Shaffer, L.G. Robertsonian translocations: Mechanisms of formation, aneuploidy, and uniparental disomy and diagnostic considerations. *Genet. Test.* **2002**, *6*, 163–168. [CrossRef]

174. Jarmuz-Szymczak, M.; Janiszewska, J.; Szyfter, K.; Shaffer, L.G. Narrowing the localization of the region breakpoint in most frequent Robertsonian translocations. *Chromosome Res.* **2014**, *22*, 517–532. [CrossRef] [PubMed]

175. Choo, K.H.; Vissel, B.; Brown, R.; Filby, R.G.; Earle, E. Homologous α satellite sequences on human acrocentric chromosomes with selectivity for chromosomes 13, 14 and 21: Implications for recombination between nonhomologues and Robertsonian translocations. *Nucleic Acids Res.* **1988**, *16*, 1273–1284. [CrossRef] [PubMed]

176. Akgun, E.; Zahn, J.; Baumes, S.; Brown, G.; Liang, F.; Romanienko, P.J.; Lewis, S.; Jasin, M. Palindrome resolution and recombination in the mammalian germ line. *Mol. Cell. Biol.* **1997**, *17*, 5559–5570. [CrossRef] [PubMed]

177. 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] [PubMed]

178. Li, H. Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* **2014**, *30*, 2843–2851. [CrossRef] [PubMed]

179. Biscotti, M.A.; Canapa, A.; Forconi, M.; Olmo, E.; Barucca, M. Transcription of tandemly repetitive DNA: Functional roles. *Chromosome Res.* **2015**, *23*, 463–477. [CrossRef] [PubMed]

180. Mills, W.K.; Lee, Y.C.G.; Kochendoerfer, A.M.; Dunleavy, E.M.; Karpen, G.H. RNA from a simple-tandem repeat is required for sperm maturation and male fertility in *Drosophila melanogaster*. *Elife* **2019**, *8*, 48940. [CrossRef]

181. McNulty, S.M.; Sullivan, L.L.; Sullivan, B.A. Human Centromeres Produce Chromosome-Specific and Array-Specific α Satellite Transcripts that Are Complexed with CENP-A and CENP-C. *Dev. Cell* **2017**, *42*, 226–240. [CrossRef]

182. Liu, H.; Qu, Q.; Warrington, R.; Rice, A.; Cheng, N.; Yu, H. Mitotic transcription installs Sgo1 at centromeres to coordinate chromosome segregation. *Mol. Cell* **2015**, *59*, 426–436. [CrossRef]

183. 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. Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA: DNA hybrid formation. *Elife* **2017**, *6*, e25293. [CrossRef]

184. Johnson, W.L.; Yewdell, W.T.; Bell, J.C.S.; McNulty, M.; Duda, Z.; O'Neill, R.J.; Sullivan, B.A.; Straight, A.F. RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *Elife* **2017**, *6*, e25299. [CrossRef] [PubMed]

185. Shirai, A.; Kawaguchi, T.; Shimojo, H.; Muramatsu, D.; Ishida-Yonetani, M.; Nishimura, Y.; Kimura, H.; Nakayama, J.I.; Shinkai, Y. Impact of nucleic acid and methylated H3K9 binding activities of Suv39h1 on its heterochromatin assembly. *Elife* **2017**, *6*, e25317. [CrossRef] [PubMed]

186. Chan, F.L.; Marshall, O.J.; Saffery, R.; Kim, B.W.; Earle, E.; Choo, K.H.; 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]

187. Chan, F.L.; Wong, L.H. Transcription in the maintenance of centromere chromatin identity. *Nucleic Acids Res.* **2012**, *40*, 11178–11188. [CrossRef] [PubMed]

188. Grenfell, A.W.; Strzelecka, M.; Heald, R. Transcription brings the complex(ity) to the centromere. *CellCycle* **2017**, *16*, 235–236. [CrossRef] [PubMed]

189. Smurova, K.; De Wulf, P. Centromere and pericentromere transcription: Roles and regulation . . . in sickness and in health. *Front. Genet.* **2018**, *9*, 674. [CrossRef] [PubMed]

190. Bouzinba-Segard, H.; Guais, A.; Francastel, C. Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8709–8714. [CrossRef]

191. Rošič, S.; Köhler, F.; Erhardt, S. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J. Cell Biol.* **2014**, *207*, 335–349. [CrossRef]

192. Chan, D.Y.L.; Moralli, D.; Khoja, S.; Monaco, Z.L. Noncoding Centromeric RNA Expression Impairs Chromosome Stability in Human and Murine Stem Cells. *Dis. Markers* **2017**, 7506976. [CrossRef]

193. Lu, J.; Gilbert, D.M. Proliferation-dependent and cell cycle regulated transcription of mouse pericentric heterochromatin. *J. Cell Biol.* **2007**, *179*, 411–421. [CrossRef]

194. Maison, C.; Bailly, D.; Roche, D.; Montes de Oca, R.; Probst, A.V.; Vassias, I.; Dingli, F.; Lombard, B.; Loew, D.; Quivy, J.P.; et al. Sumoylation promotes de novo targeting of HP1alpha to pericentric heterochromatin. *Nat. Genet.* **2011**, *43*, 220–227. [CrossRef] [PubMed]

195. Vourc'h, C.; Biamonti, G. Transcription of Satellite DNAs in Mammals. *Prog. Mol. Subcell. Biol.* **2011**, *51*, 95–118. [CrossRef] [PubMed]

196. Jolly, C.; Metz, A.; Govin, J.; Vigneron, M.; Turner, B.M.; Khochbin, S.; Vourc'h, C. Stress-induced transcription of satellite III repeats. *J. Cell Biol.* **2004**, *164*, 25–33. [CrossRef] [PubMed]

197. Rizzi, N.; Denegri, M.; Chiodi, I.; Corioni, M.; Valgardsdottir, R.; Cobianchi, F.; Riva, S.; Biamonti, G. Transcriptional activation of a constitutive heterochromatic domain of the human genome in response to heat shock. *Mol. Biol. Cell* **2004**, *15*, 543–551. [CrossRef] [PubMed]

198. Biamonti, G. Nuclear stress bodies: A heterochromatin affair? *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 493–498. [CrossRef]

199. Biamonti, G.; Vourc'h, C. Nuclear stress bodies. *Cold Spring Harb Perspect. Biol.* **2010**, *2*. [CrossRef]

200. Goenka, A.; Sengupta, S.; Pandey, R.; Parihar, R.; Mohanta, G.C.; Mukerji, M.; Ganesh, S. Human satellite-III non-coding RNAs modulate heat-shock-induced transcriptional repression. *J. Cell Sci.* **2016**, *129*, 3541–3552. [CrossRef]

201. Valgardsdottir, R.; Chiodi, I.; Giordano, M.; Rossi, A.; Bazzini, S.; Ghigna, C.; Riva, S.; Biamonti, G. Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Res.* **2008**, *36*, 423–434. [CrossRef]

202. Kuhn, G.C. Satellite DNA transcripts have diverse biological roles in *Drosophila'*. *Heredity* **2015**, *115*, 1–2. [CrossRef]

203. Usakin, L.; Abad, J.; Vagin, V.V.; de Pablos, B.; Villasante, A.; Gvozdev, V.A. Transcription of the 1.688 satellite DNA family is under the control of RNA interference machinery in Drosophila melanogaster ovaries. *Genetics* **2007**, *176*, 1343–1349. [CrossRef]

204. Menon, D.U.; Coarfa, C.; Xiao, W.; Gunaratne, P.H.; Meller, V.H. siRNAs from an X-linked satellite repeat promote X-chromosome recognition in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 16460–16465. [CrossRef] [PubMed]

205. Ferreira, D.; Escudeiro, A.; Adega, F.; Chaves, R. DNA Methylation Patterns of a Satellite Non-coding Sequence—FA-SAT in Cancer Cells: Its Expression Cannot Be Explained Solely by DNA Methylation. *Front. Genet.* **2019**, *10*, 101. [CrossRef] [PubMed]

206. Sana, J.; Faltejskova, P.; Svoboda, M.; Slaby, O. Novel classes of non-coding RNAs and cancer. *J. Transl. Med.* **2012**, *10*, 103. [CrossRef] [PubMed]

207. Eymery, A.; Horard, B.; El Atifi-Borel, M.; Fourel, G.; Berger, F.; Vitte, A.L.; Van den Broeck, A.; Brambilla, E.; Fournier, A.; Callanan, M.; et al. A transcriptomic analysis of human centromeric and pericentric sequences in normal and tumor cells. *Nucleic Acids Res.* **2009**, *37*, 6340–6354. [CrossRef] [PubMed]

208. Ehrlich, M.; Sanchez, C.; Shao, C.; Nishiyama, R.; Kehrl, J.; Kuick, R.; Kubota, T.; Hanash, S.M. ICF, an immunodeficiency syndrome: DNA methyltransferase 3B involvement, chromosome anomalies, and gene dysregulation. *Autoimmunity* **2008**, *41*, 253–271. [CrossRef]

209. Saksouk, N.; Simboeck, E.; Dejardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenet. Chromatin* **2015**, *8*, 3. [CrossRef]