



Review

Molecular Dynamics and Evolution of Centromeres in the Genus *Equus*

Francesca M. Piras [†], Eleonora Cappelletti [†], Marco Santagostino , Solomon G. Nergadze ,
Elena Giulotto ^{*} and Elena Raimondi

Department of Biology and Biotechnology “L. Spallanzani”, University of Pavia, 27100 Pavia, Italy; mfrancesca.piras@unipv.it (F.M.P.); eleonora.cappelletti@unipv.it (E.C.); marco.santagostino@unipv.it (M.S.); solomon.nergadze@unipv.it (S.G.N.); elena.raimondi@unipv.it (E.R.)

^{*} Correspondence: elena.giulotto@unipv.it; Tel.: +39-0382985541

[†] These authors contributed equally to this work.

Abstract: The centromere is the chromosomal locus essential for proper chromosome segregation. While the centromeric function is well conserved and epigenetically specified, centromeric DNA sequences are typically composed of satellite DNA and represent the most rapidly evolving sequences in eukaryotic genomes. The presence of satellite sequences at centromeres hampered the comprehensive molecular analysis of these enigmatic loci. The discovery of functional centromeres completely devoid of satellite repetitions and fixed in some animal and plant species represented a turning point in centromere biology, definitively proving the epigenetic nature of the centromere. The first satellite-free centromere, fixed in a vertebrate species, was discovered in the horse. Later, an extraordinary number of satellite-free neocentromeres had been discovered in other species of the genus *Equus*, which remains the only mammalian genus with numerous satellite-free centromeres described thus far. These neocentromeres arose recently during evolution and are caught in a stage of incomplete maturation. Their presence made the equids a unique model for investigating, at molecular level, the minimal requirements for centromere seeding and evolution. This model system provided new insights on how centromeres are established and transmitted to the progeny and on the role of satellite DNA in different aspects of centromere biology.

Keywords: satellite DNA; neocentromeres; centromere repositioning; centromere sliding; karyotype evolution; CENP-A; genus *Equus*



Citation: Piras, F.M.; Cappelletti, E.; Santagostino, M.; Nergadze, S.G.; Giulotto, E.; Raimondi, E. Molecular Dynamics and Evolution of Centromeres in the Genus *Equus*. *Int. J. Mol. Sci.* **2022**, *23*, 4183. <https://doi.org/10.3390/ijms23084183>

Academic Editor: Jiří Fajkus

Received: 8 March 2022

Accepted: 6 April 2022

Published: 10 April 2022

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



Copyright: © 2022 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

The centromere is the chromosomal locus driving chromosome segregation via kinetochore formation and spindle attachment. As a consequence, a functional centromere is required for the stable inheritance of each chromosome during mitosis and meiosis.

The simplest centromere found in nature is the “point centromere” of the budding yeast *Saccharomyces cerevisiae*, where a single modified nucleosome is associated to an about 120 bp long DNA sequence that is both necessary and sufficient to grant full centromere function [1]. An opposite situation is observed in some animal and plant species where “holocentromeres” are spread along the length of the entire chromosomes, the best studied example being the one of *Caenorhabditis elegans* [2]. However, most centromeres of higher eukaryotes, including humans, are of the regional type [1]. These centromeres occupy a defined region along the chromosome, typically encompassing megabases of highly repetitive DNA, called satellite DNA [3]. This type of sequences occupies the functional centromeric domains and the surrounding pericentromeric regions.

S. cerevisiae is the only eukaryotic organism in which the centromeric function is entirely determined by the DNA sequence [4,5]. In all other eukaryotes, although the centromeric function is well conserved, rapid evolution among and within species has

been observed for centromeric DNA sequences and size, making centromeres rapidly evolving genomic regions that can shape karyotypes and drive speciation. The conflict between the need of function conservation and the surprising variety of DNA sequences and strategies involved in centromere formation is known as the “centromere paradox” [6]. A possible interpretation of this paradox is that epigenetic determinants are contributing to the maintenance and inheritance of the centromeric function [5,7,8]. A large body of evidence supports this hypothesis. To this regard, an important observation is the occasional stable transmission of dicentric chromosomes following inactivation of one centromere [9]. Therefore, epigenetic mechanisms define whether a DNA sequence can acquire a centromeric function. Another piece of evidence supporting the epigenetic basis for the centromeric function comes from human clinical neo-centromeres, which confer the ability to segregate correctly to acentric fragments thanks to the acquisition of centromeric function by a hitherto non-centromeric interstitial sequence [10,11]. The discovery of naturally occurring satellite-free centromeres fixed in some animal and plant species indicates that satellite DNA is not strictly required for the centromeric function. These centromeres coexist, in the same karyotype, with canonical satellite-based centromeres and are perfectly functional. The first satellite-free centromere, fixed in a vertebrate species, was discovered in one horse chromosome by our group [12]. We demonstrated that this chromosome is totally devoid of satellite repeats both within the CENP-A binding domain and in the surrounding pericentromeric regions. Thereafter, other examples of satellite-free centromeres have been described in other vertebrate species by our laboratory [13,14] and other groups [15–17] as well as in plants [18–22]. The major epigenetic mark of regional centromeres is CENP-A, a centromere-specific histone H3 variant which interacts with a number of proteins constitutionally or transiently associated to centromeric DNA to establish the kinetochore [23]. Overexpression of CENP-A in *Drosophila* causes its misincorporation in non-centromeric regions and promotes the formation of functional ectopic kinetochores, thus inducing chromosome segregation errors [24,25]; a similar situation was observed in budding yeast [26]. These experiments strongly suggest that CENP-A is both necessary and sufficient for the formation of the kinetochore and that the presence of this histone variant drives the formation of a highly specialized chromatin structure that directs the formation of a functional centromere. Similar experiments of CENP-A overexpression and mistargeting to non-centromeric regions in human chromosomes failed to induce complete kinetochore assembly, suggesting that this process is more complex and additional functions are required [27,28]. Nevertheless, mislocalisation of CENP-A alters the structure and function of endogenous kinetochores and leads to chromosomal instability, suggesting a key role of this protein in specifying centromere function [29].

2. Centromeric and Pericentromeric Satellite DNA

Satellite DNA (SatDNA) is composed by extended tandemly repeated monomers typically located at centromeres but, in some species, also at terminal and intrachromosomal sites [30–32]. Satellite DNA arrays show wide variation in copy number and sequence composition among species, representing the most rapidly evolving sequences in eukaryotic genomes and contributing to shape their karyotypes [31]. SatDNA tends to be homogenized following the principles of concerted evolution. Homogenization of these repeats may occur by unequal crossing over, gene conversion, rolling circle replication or transposition [33,34]. Unequal exchanges and mutational variation in individual satellite monomers may cause the formation of higher-order repeats (HORs) consisting of large units composed by multiple divergent monomers [35,36]. Different families of satellite DNA can be located at the centromere functional core and in the flanking pericentromeric regions [37].

The presence of satellite sequences in the majority of eukaryotic centromeres suggests that they may play a role in kinetochore assembly and/or stability, as inferred since their discovery [38–40]. Even if their sequences are highly divergent, centromeric satellites from different species tend to be AT-rich [41] and are composed of monomers of similar length

(often around 150–180 bp), leading to hypothesise that the repeat unit itself might reflect uniformity in nucleosome phasing and that these constraints, rather than the sequence itself, are needed for heterochromatin propagation [6,42–44]. However, it is difficult to reconcile the existence of centromeres completely devoid of satellite DNA with a pivotal function of this genome component.

The function of satellite DNA remains poorly understood. Pericentromeric and centromeric satellite DNA families are transcribed in a number of animal and plant species (examples are reported in [45–53]; reviewed in [54,55]). Although pericentromeric satellites are embedded in heterochromatin, several findings indicate that pericentric heterochromatin is unexpectedly permissive to transcription. The resulting pericentric transcripts were proposed to contribute to the maintenance of the heterochromatic condensed state (reviewed in [30,56]). Transcription has also been reported for centromeric satellites, and centromeric transcripts were proposed to be involved in CENP-A loading and kinetochore assembly [53,57,58]. However, it is still a matter of debate whether transcription is required for proper centromeric function. Indeed, transcription is not a universal feature of centromeres, with satellite-free centromeres arising either in transcribed regions [11,59–63] or in gene desert regions without evidence of transcription [11,12,59,60,64].

Although dispensable for centromere specification, it was suggested that satellite DNA might contribute to the stability of centromeres, as predicted by the “centromere drive” theory [65,66]. According to this model, during the asymmetric female meiosis, homologous chromosomes may compete for their inclusion in the egg via their “centromere strength”. Since centromeres with more extended satellite DNA arrays may host more CENP-A nucleosomes, it was suggested that centromere strength is related to the length of tandemly repeated sequences [65–69]. Recent cell biological studies indicate that centromeres compete by destabilizing microtubule interactions [70,71]. At the same time, an excessive accumulation of repeated arrays by “selfish” centromeres would lead to expansion of the CENP-A binding domain, which seems to be counteracted by the co-evolution between centromeric proteins and centromeric DNA [65]. In this scenario, satellites at the borders of the centromeric core, during expansion, become more and more degenerated, preventing centromeric protein binding. These degenerate repeats occupy pericentromeric regions, lose their transmission advantage and evolve neutrally [66].

The extreme abundance of pericentromeric compared to centromeric satellite DNA seems in contradiction with the common view that pericentromeric satellites are simply selfish parasitic DNA sequences [72], or fossils of centromere evolution [3]. The maintenance of such extended arrays of tandemly repeated DNA would be too high a burden for the cell [73]. To solve this controversy, pericentromeric satellites were recently proposed to have a structural role in the tridimensional nuclear organisation, driving the formation of chromocenters [73].

The typical presence of satellite repeats at centromeres hampered a comprehensive molecular analysis of these intriguing loci. The discovery of functional centromeres completely devoid of satellite repeats represented a milestone in centromere biology [7,10–13,15–17,74].

3. Evolutionarily New Centromeres

The term “centromerization” was coined by Choo to define the process of centromere formation in a chromosomal region [7]. Centromerization normally concerns the propagation of an existing centromere during replication. Rarely, this phenomenon occurs in regions which are normally non-centromeric. The ectopic centromere that appears occasionally in a hitherto non-centromeric chromosomal region is called neocentromere [7,8,74]. Two different types of neocentromeres have been identified: human clinical neocentromeres and evolutionarily new centromeres (ENC).

While the majority of human neocentromeres are seeded in acentric fragments, conferring on them the ability to segregate, a few human neocentromeres have been described in otherwise normal karyotypes [10,75]. In these cases, the old centromere appears unchanged,

but functionally inactive; conversely, a non-centromeric sequence acquires the ability to recruit a functional kinetochore.

While clinical human neocentromeres are sporadic cases that are not fixed in the population, evolutionarily new centromeres are fixed in some species and represent an aspect of karyotype evolution. Interestingly, several lines of evidence suggest that these two types of neocentromeres are faces of the same coin [76], since the chromosome regions where some human neocentromeres were seeded were orthologous to evolutionary neocentromeres in primates [64,77–79]. These observations indicate that there are chromosome domains which have an inherent potentiality to form centromeres, suggesting that a similar mechanism may originate both human neocentromeres and ENC's [79]. In other words, some chromosome domains may harbor the capability to form centromeres depending on still unknown sequence features or on a peculiar chromatin environment which may favor centromere seeding.

Evolutionarily new centromeres are generated during evolution by repositioning events. Centromere repositioning consists of the movement of the centromere along the chromosome length in the absence of any structural rearrangement. In other words, while the ancestral centromere is inactivated, a centromere appears *de novo* at another chromosomal position. The first unequivocal proof of the existence of this phenomenon was given by Montefalcone and colleagues in 1999 [80]. The group traced the phylogenetic history of chromosome IX in primates by comparative FISH with a panel of BAC clones. They demonstrated that, in some species, the centromere moved and was repositioned along the chromosome while the surrounding markers maintained the original position. Thereafter, a number of ENC's were found in primates and in other mammalian orders (reviewed in [76,81]).

The ENC's initially described in primates are composed of the classical extended arrays of satellite DNA. The first example of a satellite-free evolutionary neocentromere was described in the horse by our laboratory [12]. Later on, our group discovered that the species of the genus *Equus* (horses, asses and zebras) are characterized by an extraordinary number of satellite-free neocentromeres [13,14]. Beyond equids, the only satellite-free centromere identified so far in a mammalian species was discovered in orangutan chromosome 12 [17]. Differently from the equids, this satellite-free centromere and the ancestral satellite-based centromere coexist in the same population as a polymorphism [17]. Thus, the presence of extraordinary numbers of satellite-free centromeres has made the equids a unique model for dissecting several aspects of centromere biology [82–84].

4. Unique Model System: The Genus *Equus*

The genus *Equus* is the only extant genus of the Equidae family, which belongs to the Perissodactyla (odd-toed ungulates) order together with Tapiridae (tapirs) and Rhinocerotidae (rhinoceroses) families. Currently living equids comprise horses (*Equus caballus* and *Equus przewalskii*), African asses (*Equus africanus asinus* and *Equus africanus somaliensis*), Asiatic asses (*Equus hemionus* and *Equus kiang*), and zebras (*Equus grevyi*, *Equus burchelli*, and *Equus zebra hartmannae*).

While the karyotypes of Tapiridae and Rhinocerotidae remained quite stable during evolution and resemble the putative perissodactyl ancestral karyotype, characterized by high chromosomal number and a prevalence of acrocentric chromosomes [85], *Equus* karyotypes underwent a rapid evolution after the divergence from the common ancestor, dated 4.0–4.5 million years ago (Mya) [85,86]. The most recent radiation events, found among asses and zebras, occurred less than 1 Mya and many species and subspecies emerged in this very short evolutionary time [85,87,88].

The evolutionary plasticity of the equid genomes is also proved by the high mobility of insertion DNA elements, such as retrotransposons [89], interstitial telomeres [90,91] and numts (nuclear sequences of mitochondrial origin) [92]. The insertion of these sequences is recognized as a driving force in evolution [91–94], and the extraordinary frequency of

insertion polymorphism in the genus *Equus* [89,90,92] supports the hypothesis that the genome of equid species is in an ongoing evolutionary state.

The numerous speciation events that occurred during *Equus*' evolution were accompanied by extensive karyotype reshaping due to both chromosome rearrangements and centromere repositioning [14,88,95–97]. These phenomena were responsible for a reduction in chromosome number and for the passage from an ancestral karyotype, with the majority of chromosomes being acrocentric, to karyotypes mainly composed by meta- and submeta-centric chromosomes. Several of these (sub)metacentrics harbour a repositioned centromere and, in the majority of them, this centromere is satellite-free [14,96]. The chromosomal distribution of satellite tandem repeats, analysed by FISH in *Equus caballus*, *E. asinus*, *E. grevyi* and *E. burchelli* [14], revealed that several centromeres are devoid of satellite DNA (Figure 1) and that blocks of satellite DNA sequences are present at several non-centromeric chromosome end (Figure 1B–D). The presence of these large satellite DNA arrays at non-centromeric terminal positions on a subset of (sub)metacentrics with a repositioned satellite-less centromere, combined with the knowledge that ancestral chromosomes were known to be acrocentric, strongly suggests that they are relics of ancestral centromeres.

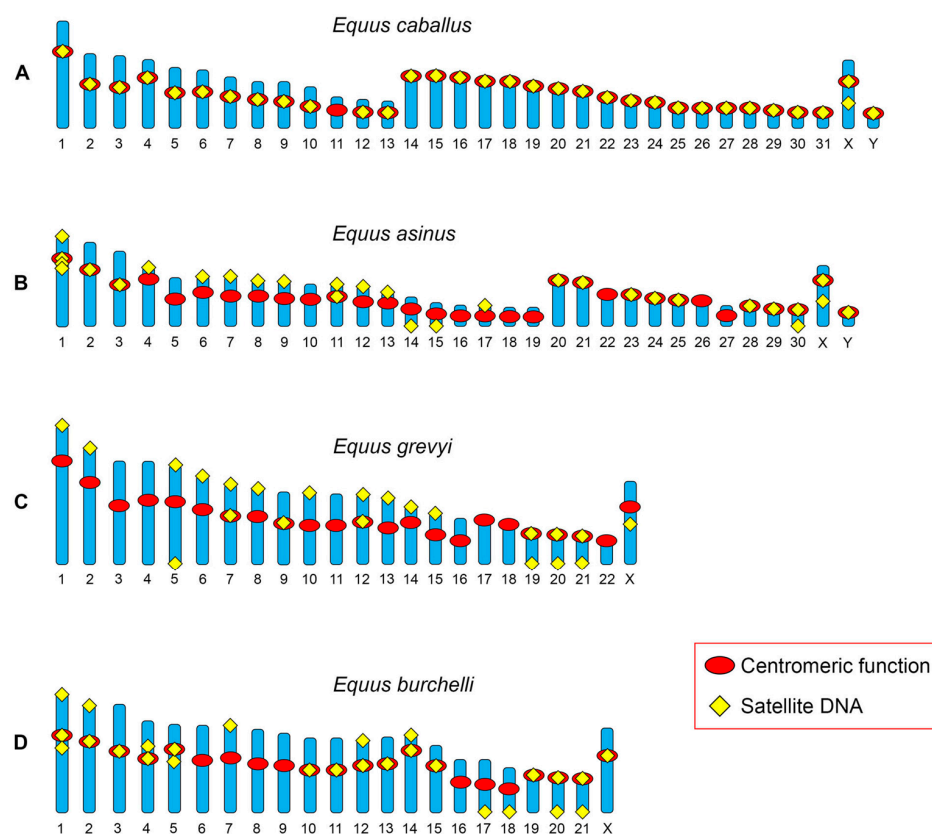


Figure 1. Schematic representation of the chromosomal distribution of satellite DNA obtained by FISH experiments in *E. caballus*, *E. asinus*, *E. grevyi* and *E. burchelli*. (A) In *E. caballus*, all centromeres, with the exception of that of chromosome 11, are associated with satellite DNA sequences. Satellite DNA is also present at an interstitial position on the long arm of chromosome X. (B) In *E. asinus*, 18 centromeres are not associated with satellite DNA sequences at the FISH resolution level. Satellite DNA sequences are present at 13 non-centromeric chromosome ends. Satellite DNA is present at an interstitial position on the long arm of chromosome X. (C) In *E. grevyi*, 17 centromeres are void of satellite DNA and satellite DNA sequences are present at 15 non-centromeric termini and at an interstitial position of the long arm of the X chromosome. (D) In *E. burchelli*, 7 centromeres do not display satellite DNA sequences and 9 non-centromeric chromosome ends are associated with satellite DNA. Satellite DNA loci are present in the pericentromeric region of chromosomes 1, 4 and 5. Modified from [14].

These observations led us to propose a model explaining the birth and evolution of new centromeres [14]. According to this model, satellite-free centromeres are the primary result of a centromere repositioning event which moved the centromeric function from the original position to a new one lacking satellite DNA. On the other hand, the satellite sequences of the old inactivated centromere remain at their original position and are then progressively lost during evolution. Meanwhile, “immature” satellite-free centromeres progressively reach their maturity by acquiring satellite DNA. Thanks to their rapid evolution, the equids offered the possibility to recognize all the steps of this process (Figure 1), while only “mature” neocentromeres, associated with satellite DNA, were initially described in primates [76].

The extraordinary number of satellite-free centromeres in equids is a direct consequence of their rapid and recent evolution [14,88,96,97]. The number of satellite-free centromeres is surprisingly high in asses and zebras (Figure 1), whose species radiated in very recent evolutionary times. Moreover, only in these species were satellite DNA sequences at non-centromeric ends identified as fossil remains of the ancient inactivated centromeres, testifying that these repositioning events were extremely recent. On the other hand, in the horse, whose karyotype shows a prevalence of acrocentric chromosomes and is considered more similar to the ancestral perissodactyl karyotype, only one satellite-free neocentromere is present, suggesting that this species is more “mature” with respect to centromere evolution.

Satellite-free natural centromeres fixed in the *Equus* species [12–14] offer the opportunity to investigate the minimal requirements for centromere seeding, evolution and complete maturation at the molecular level. In particular, using this model system, several key open questions can be addressed: How are the centromeric domains organized at the molecular level? How are centromeric domains transmitted through generations? Which mechanisms drive centromere birth and maturation during evolution? Which is the contribution of satellite DNA to segregation fidelity and inhibition of recombination during meiosis?

4.1. Birth, Evolution and Transmission of Satellite-Free Centromeres

The discovery that the centromere of horse chromosome 11 (ECA11) is completely devoid of satellite DNA [12] was the result of ChIP-on-chip experiments carried out hybridizing horse genomic DNA, purified from chromatin immunoprecipitated with an anti CENP-A antibody, to an array spanning the region containing the primary constriction. The satellite-free centromeric locus of ECA11 does not contain segmental duplications nor protein coding genes. This region is highly conserved in mammals, but the only species in which a centromere is present therein is the horse. These findings are in agreement with the hypothesis that the centromeric function is unrelated to DNA sequence and that the ECA11 centromere was formed very recently during the evolution of the horse lineage. This centromere, in spite of being functional and stable in all horses, did not acquire all the marks typical of mammalian centromeres. It represents the first example of an evolutionary neocentromere caught in a stage of incomplete maturation.

In these early experiments, two peaks corresponding to two CENP-A binding domains were identified in the individual analysed. The same domains were identified when chromatin was immunoprecipitated with an anti CENP-C antibody. When the same experiment was then carried out in five unrelated individuals, it was demonstrated that each horse exhibited a distinct arrangement of CENP-A binding domains [98]. Two CENP-A binding regions were present in three of the five individuals and, in the other two individuals, only one broad peak was observed, which was the result of partial overlapping of two peaks. While single peaks occupy about 100 kb, the entire ECA11 region involved occupies about 500 kb. Single-nucleotide polymorphism (SNP) analysis gave the definitive proof that, when two peaks were observed in the same individual, they were not located on the same chromosome, but each homolog contained a single CENP-A binding domain. These different positional alleles for CENP-A binding were defined as “epialleles” and the

phenomenon was called “centromere sliding”. Positional variation in centromeric domains was also reported within the satellite-based centromere of human chromosome 17, where the centromere can assemble on different alpha satellite arrays [99]. Further evidence of the existence of epialleles was derived by single molecule analysis by immuno-FISH on chromatin fibres. Taken together, these results demonstrated that the position of the CENP-A binding domain can slide within a relatively wide region and allowed us to conclude that centromeric domains are characterized by positional instability which may be physically limited by epigenetic boundaries.

Centromeric domains were then analysed in the donkey by means of ChIP-seq [13], demonstrating that, in this species, more than half of the centromeres (16 out of 31) are devoid of satellite DNA. The 16 satellite-free donkey centromeric domains derived from centromere repositioning events that occurred in this lineage since they are orthologous to horse non-centromeric sequences. Considering that horse and donkey lineages separated recently (about 4–4.5 Mya), we can speculate that there was not enough evolutionary time for satellite DNA accumulation and centromere maturation.

The donkey satellite-free centromeres spanned 54–345 kb and, as already described for the horse satellite-free centromere, contained one or two CENP-A binding domains. The sequence of these domains was de novo assembled and analysed, demonstrating that all of them lay in LINE- and AT-rich regions. AT richness is a typical feature of centromeres in a number of organisms [4,11,100] and might promote the adoption of a non-B DNA configuration which is typically found at centromeres [36,101]. It was recently shown that AT-rich exogenous DNA is also capable of functioning as a centromere in the model organism *Schizosaccharomyces pombe* [102]. It was proposed that AT-rich sequences, which display low nucleosome occupancy, may favour the insertion of LINE-1 elements, which are frequently enriched in natural human centromeres [34] and in clinical neocentromeres [11,79,103]. A peculiar sequence organisation was described in *Drosophila*, where centromeric domains form on islands enriched in LINE retroelements which are embedded within extended arrays of pericentromeric satellite repeats [104]. It is not clear whether AT and LINE richness contributes to the acquisition of centromeric function. Although these sequences are enriched in the donkey satellite-free centromeres, their abundance is not observed at the evolutionary satellite-free neocentromere of horse chromosome 11 [12,13]. The observation that LINE/LTR-rich domains cluster within the nucleus suggests that their abundance at several centromeric domains may be related to function [105]. Therefore, the sequence composition of the satellite-free donkey centromeres may be involved in the organisation into subnuclear domains that promote the functional activation of centromeric chromatin.

The sequences of the donkey satellite-free centromeres were compared with their horse non-centromeric counterparts demonstrating, in some instances, the presence of rearrangements (deletions, amplifications, insertions and inversions). These rearrangements could have occurred before or after neocentromere formation. As it was suggested that CENP-A can be recruited to DNA breaks [106], chromosome rearrangements may have promoted centromere formation. However, this is not a rule, since, in several instances, no relevant sequence change was found in the donkey with respect to the horse. Similarly, Tolomeo and colleagues demonstrated that the sequence underlying the satellite-free polymorphic ENC that they described in the orangutan is identical to its non-centromeric counterpart [17].

Five donkey neocentromeres contain novel tandem repetitions of chromosome-specific sequences that are a single copy in the horse genome. These amplified genomic sequences are unrelated to one another, with amplified units ranging in size from a few to several tens of kilobases. The repeat copy number was variable in the two individuals analysed, suggesting the existence of polymorphism in the population. It is tempting to hypothesize that DNA amplification may represent an intermediate stage toward satellite DNA formation during evolution [13]. In light of these observations, a model for the maturation of a centromere during evolution was proposed, including different routes, some of which involve sequence amplification (Figure 2). Following this new model, after centromere inactivation, associated satellite sequences are maintained at the site of the ancestral centromere while a

neocentromere arises in a new locus, completely devoid of satellite sequences (Figure 2B). Subsequently, satellite sequences are gradually lost at the non-centromeric site (Figure 2C), and, at the functional satellite-free centromere, amplification may occur (Figure 2E,F) as an intermediate step toward complete maturation of the neocentromere. The amplifications observed at some donkey centromeres may be considered as an early seed of chromosome-specific centromeric satellites, suggesting that amplification-like mechanisms can trigger the formation of tandemly repeated DNA sequences within the centromere core. Finally, these neocentromeres will acquire the typical complexity of the mammalian centromeres and will be embedded in tandemly repeated DNA (Figure 2D). It cannot be excluded that sequence amplification may occur before neocentromere formation (Figure 2G) and may favour centromerization, but no data supporting this alternative route are available. It is important to underline that, as indicated in Figure 2, horse or donkey chromosomes corresponding to each maturation step were observed, thus representing evolutionary snapshots which validate the proposed model.

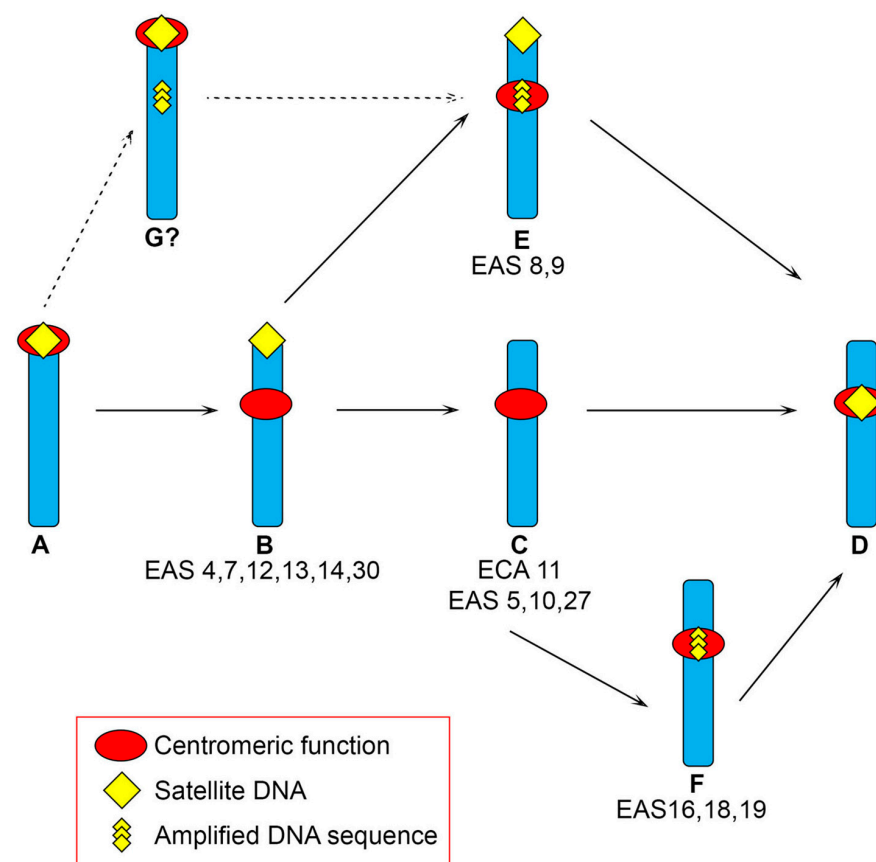


Figure 2. Model for the maturation of a centromere during evolution. Different routes are delineated leading to a mature satellite-based repositioned centromere (D) from an ancestral satellite-based centromere (A) through satellite-free intermediates (B,C,E,F). Route A–D: a neocentromere arises in a satellite-free region and satellite DNA remains at the ancestral position; the ancestral satellite DNA is lost from the non-centromeric terminus and, finally, the neocentromere acquires satellite repeats, giving rise to a “mature” neocentromere. This route follows the previously proposed model [14]. Routes A, B, E, D or A, B, C, F, D: at an already functional satellite-free neocentromere, amplification occurs as an intermediate step toward complete maturation of the neocentromere. Neocentromere maturation and loss of satellite DNA from the old centromere site are independent events that can occur at different stages. The 16 donkey chromosomes carrying a satellite-free neocentromere and the horse chromosome 11 exemplify satellite-free intermediates (B,C,E,F) and are listed below each step. It cannot be excluded that sequence amplification precedes neocentromere formation (G), but no chromosome corresponding to this step was found so far. Modified from [13,14].

The possibility of interspecific breeding in *Equus* species offered the opportunity to follow the transmission of epialleles for CENP-A binding through generations and, thus, to understand how centromeric domains are transmitted to the progeny [13]. Since the hybrid individuals (mules and hinnies) contain two haploid genomes, one from the donkey and one from the horse, the transmission of individual epialleles across generations could be followed. The analysis of CENP-A binding domains in hybrid families revealed that they are inherited as Mendelian traits. However, their position can slide in one generation. Interestingly, the centromeric region in the offspring is always at least partially overlapping the domain of the parent, suggesting that CENP-A nucleosomes are partially maintained at their position. Therefore, centromeres do not jump to a completely new location but, rather, in the course of several generations, slight movements may accumulate, giving rise to nonoverlapping epialleles. Conversely, the position of the centromere is stable during mitotic propagation of cultured cells that were previously immortalized [107], suggesting that the sliding that was observed in the hybrids could have occurred during germ-line differentiation, meiotic division, fertilization or early developmental stages.

4.2. Satellite DNA and Centromere Stability

The presence of well-characterized satellite-free centromeres in a context of canonical satellite-based centromeres in the equid model system was exploited to understand the contribution of satellite DNA to different aspects of centromere biology.

Beyond the extraordinary number of evolutionary neocentromeres, the high plasticity of the equid genomes is also testified by the architectural organisation of centromeric and pericentromeric satellite DNA families. Three main satellite DNA families were described, namely 37cen, 2PI and 137sat. These satellite families differ in the length of their repeat unit, with 37cen consisting of a 221 bp repeat [14,108], 2PI composed of a 23 bp repeat [14] and 137sat made by 137 bp repetitions [109].

While in asses and zebras, the majority of centromeres are satellite-free and many satellite loci are not centromeric, the horse displays a unique satellite-free centromere coexisting with the typical satellite-based mammalian centromeres. High-resolution FISH on combed DNA fibres demonstrated that at least some horse satellite-based centromeres may display a mosaic arrangement of the different satellite DNA families where short arrays of the 2PI and EC137 satellites are closely intermingled and immersed within very large stretches of the 37cen sequence [109]. This organisation suggests that recombination events among centromeric and pericentromeric satellite DNA can occur in the horse genome. The arrays of 37cen embed the centromeric core of horse satellite-based centromeres. Indeed, immunoprecipitation experiments with an anti-CENP-A antibody proved that this satellite family is the only one bound by CENP-A and thus bears the centromeric function [52]. Although centromeric satellites are typically AT-rich [36], 37cen is GC-rich indicating that GC richness is compatible with the centromeric function.

The organisation of horse satellite-based centromeres, with a centromeric core of 37cen satellite and a flanking pericentromere characterized by different sequence composition, is in agreement with the typical organisation of mammalian centromeres. Indeed, it is known that new satellite sequences arise and expand in the inner centromeric core, progressively moving the older satellite families towards the pericentromere, forming layers of different age [110]. Pericentromeric satellites progressively degenerate, losing the ability to be bound by centromeric proteins and avoiding a harmful expansion of the functional centromere [65,66].

As mentioned above, centromeric satellite DNA is transcribed in several species. It has been hypothesized that transcriptional competence may allow centromeric chromatin to assume a partially open conformation that may be important for CENP-A loading. Centromere transcripts have been proposed to favour the formation of a flexible scaffold needed to assembly or stabilize the kinetochore [30,111]. Moreover, they may act in trans on all or on a subset of chromosomes independently of the primary DNA sequence [30,111]. In agreement with this notion, the horse 37cen centromeric satellite was proved, by RNA-seq

approach, to be transcriptionally active [52]. However, the possible role of this centromeric transcript in the epigenetic establishment of centromeric chromatin is still under debate.

Satellite DNA is believed to stabilize centromeres, and a crucial aspect of centromere biology concerns its contribution to chromosome segregation fidelity. Data on the mitotic stability of satellite-free centromeres were already obtained by the analysis of pathologic satellite-free human centromeres. These neocentromeres are often present in the individual in mosaic form, indicating mitotic instability. However, most human pathologic neocentromeres give rise to partial trisomy or tetrasomy and, therefore, the selective disadvantage of cell lines carrying partial aneuploidy rather than an intrinsic mitotic instability of the neocentromere itself may be responsible for the mosaicism observed [11].

The coexistence of canonical satellite-based centromeres with the satellite-free centromere of chromosome 11 was exploited to test whether satellite DNA may influence chromosome segregation in a non-pathologic setting. The mitotic stability of this neocentromere was investigated and compared with that of horse chromosome 13, which is similar in size but has a canonical satellite-based centromere [112]. To this purpose, two independent molecular-cytogenetic approaches, the micronucleus assay and interphase aneuploidy analysis, were used. The mitotic stability of the two chromosomes was the same demonstrating that, in the horse system, centromeric satellite DNA is dispensable for chromosome segregation fidelity. Although the widespread presence of repeated DNA at natural centromeres suggests that there is a positive selection for this arrangement, the above data demonstrate that the mitotic stability of a chromosome is not universally influenced by the presence of highly repeated DNA sequences at its centromere.

4.3. Meiotic Behaviour of Centromeric Domains

An important aspect of centromere biology is the so-called “centromere effect”, that is the suppression of meiotic recombination exerted by the centromere [113]. This phenomenon has been described in all eukaryotes, including humans and other mammals, but it was a matter of debate whether the presence of repetitive DNA contributes to it [114].

The horse model appeared again as the right system to answer this question [115]. Using cytogenetic mapping of recombination foci with respect to centromeres in horse pachytene spreads, it was discovered that the horse neocentromere exerts a crossover suppression as well as canonical centromeres. Thus, this inhibitory effect depends on the centromeric function itself rather than on the presence of highly repetitive DNA [115]. This discovery reinforces the hypothesis that this effect is related to the epigenetic environment of the centromere and is in agreement with the notion that the occurrence of crossovers near centromeres is selectively disadvantageous because it may cause premature sister chromatid separation leading to non-disjunction at the second meiotic division [114,115]. It would be interesting to verify whether centromeres that were inactivated during evolution would have acquired the recombination potential typical of interstitial loci.

During the analysis of the distribution of recombination foci in horse meiosis, a peculiar phenomenon was observed: while the majority of meiotic bivalent chromosomes was labelled with a single immunofluorescence centromeric signal, double-spotted or extended centromeric signals were also detected on a subset of autosomal bivalents with a satellite-based centromere [115]. It is well known that the centromeres of the homologous chromosomes are paired in the chromosome meiotic bivalent [116,117], but the centromeric domains of the paired homologous chromosomes appeared shifted with respect to each other. While the same phenomenon was detected, but remained unexplained, in a few mammalian species affecting very few chromosomes [118,119], the horse showed a surprisingly high frequency of these “double-spotted” centromeres, and their number was variable from cell to cell within the same individual [115]. Their appearance was proposed to be the result of two components: positional variation of the CENP-A binding domains between the two homologous chromosomes and misalignment between the centromeric and pericentromeric satellite DNA arrays during pairing [115]. According to this model, the positional variation of CENP-A binding domains, which was observed at the molecular

level at satellite-free centromeres, seems to also affect satellite-based centromeres. Thus, the centromeric domains in the paired homologous chromosomes seem to occupy different positions along the 37cen array. In addition, the polymorphism in the number of centromeric tandem repeats combined with the high sequence identity among monomers may result in a staggered pairing of the homologous satellite arrays. According to this interpretation, misalignment is variable from meiosis to meiosis and may increase the physical distance between centromeric domains, which may already be in different positions on the two homologs, making centromere sliding visible at the cytogenetic resolution [115]. Clearly, this could not happen at satellite-free centromeres where no repeats are present, and thus the centromere sliding phenomenon could not be “magnified” to be visible at cytogenetic resolution.

5. Conclusions

One of the key enigmas of chromosome biology is how centromeric chromatin is established, transmitted to the progeny and modified during evolution. This is one of the most challenging open questions in biology due to its implications in human cancer and chromosome abnormalities as well as in karyotype evolution and speciation.

Centromeres are typically associated with satellite DNA sequences which hinder the molecular analysis of centromeric chromatin. The discovery of a natural centromere completely devoid of satellite DNA and fixed in a species, the horse, represented a milestone in centromere biology. The subsequent identification of satellite-free centromeres in other equids as well as in other vertebrate and plant species has revealed that this type of centromeric DNA configuration is not so rare in natural karyotypes. However, the genus *Equus* remained unprecedented in terms of uncoupling between satellite DNA and centromeric function. This is the only mammalian genus in which an extraordinary number of satellite-free centromeres were identified so far. Thus, the genus *Equus* represents a powerful model system in centromere biology. The main results obtained using this system demonstrated that: (1) CENP-A binding domains are positionally unstable, generating epialleles which are inherited as mendelian traits; (2) amplified DNA sequences are possible intermediates towards centromere maturation during evolution; (3) the majority of satellite-free centromeres are enriched in AT and LINE sequences; (4) satellite DNA is dispensable for chromosome segregation fidelity and (5) the centromere effect is not related to the presence of satellite DNA but to the centromeric function itself.

Author Contributions: Writing—original draft preparation, F.M.P., E.C., E.G. and E.R.; writing—review and editing, F.M.P., E.C., M.S., S.G.N., E.G. and E.R. All authors have read and agreed to the published version of the manuscript.

Funding: Italian Ministry of Education, University and Research (MIUR) (Dipartimenti di Eccellenza Program (2018–2022)—Dept. of Biology and Biotechnology “L. Spallanzani”, University of Pavia) PRIN Grant No. 2015RA7XZS_002 and from the Consiglio Nazionale delle Ricerche (CNR-Progetto Bandiera Epigenomica).

Data Availability Statement: Not applicable.

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

References

1. Pluta, A.F.; Mackay, A.M.; Ainsztein, A.M.; Goldberg, I.G.; Earnshaw, W.C. The centromere: Hub of chromosomal activities. *Science* **1995**, *270*, 1591–1594. [[CrossRef](#)] [[PubMed](#)]
2. Melters, D.P.; Paliulis, L.V.; Korf, I.F.; Chan, S.W. Holocentric chromosomes: Convergent evolution, meiotic adaptations, and genomic analysis. *Chromosome Res.* **2012**, *20*, 579–593. [[CrossRef](#)] [[PubMed](#)]
3. Malik, H.S.; Henikoff, S. Major evolutionary transitions in centromere complexity. *Cell* **2009**, *138*, 1067–1082. [[CrossRef](#)]
4. Clarke, L.; Carbon, J. The structure and function of yeast centromeres. *Annu. Rev. Genet.* **1985**, *19*, 29–55. [[CrossRef](#)] [[PubMed](#)]
5. Cleveland, D.W.; Mao, Y.; Sullivan, K.F. Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling. *Cell* **2003**, *112*, 407–421. [[CrossRef](#)]

6. Henikoff, S.; Ahmad, K.; Malik, H.S. The centromere paradox: Stable inheritance with rapidly evolving DNA. *Science* **2001**, *293*, 1098–1102. [[CrossRef](#)]
7. Choo, K.H. Centromerization. *Trends Cell Biol.* **2000**, *10*, 182–188. [[CrossRef](#)]
8. Kalitsis, P.; Choo, K.H. The evolutionary life cycle of the resilient centromere. *Chromosoma* **2012**, *121*, 327–340. [[CrossRef](#)]
9. McNulty, S.M.; Sullivan, B.A. Centromere Silencing Mechanisms. *Prog. Mol. Subcell. Biol.* **2017**, *56*, 233–255. [[CrossRef](#)]
10. 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.
11. Marshall, O.J.; Chueh, A.C.; Wong, L.H.; Choo, K.H. Neocentromeres: New insights into centromere structure, disease development, and karyotype evolution. *Am. J. Hum. Genet.* **2008**, *82*, 261–282. [[CrossRef](#)] [[PubMed](#)]
12. Wade, C.M.; Giulotto, E.; Sigurdsson, S.; Zoli, M.; Gnerre, S.; Imsland, F.; Lear, T.L.; Adelson, D.L.; Bailey, E.; Bellone, R.R.; et al. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* **2009**, *326*, 865–867. [[CrossRef](#)] [[PubMed](#)]
13. Nergadze, S.G.; Piras, F.M.; Gamba, R.; Corbo, M.; Cerutti, F.; McCarter, J.G.W.; Cappelletti, E.; Gozzo, F.; Harman, R.M.; Antczak, D.F.; et al. Birth, evolution, and transmission of satellite-free mammalian centromeric domains. *Genome Res.* **2018**, *28*, 789–799. [[CrossRef](#)] [[PubMed](#)]
14. Piras, F.M.; Nergadze, S.G.; Magnani, E.; Bertoni, L.; Attolini, C.; Khoraiuli, L.; Raimondi, E.; Giulotto, E. Uncoupling of satellite DNA and centromeric function in the genus *Equus*. *PLoS Genet.* **2010**, *6*, e1000845. [[CrossRef](#)]
15. Shang, W.H.; Hori, T.; Toyoda, A.; Kato, J.; Popendorf, K.; Sakakibara, Y.; Fujiyama, A.; Fukagawa, T. Chickens possess centromeres with both extended tandem repeats and short non-tandem-repetitive sequences. *Genome Res.* **2010**, *20*, 1219–1228. [[CrossRef](#)]
16. Locke, D.P.; Hillier, L.W.; Warren, W.C.; Worley, K.C.; Nazareth, L.V.; Muzny, D.M.; Yang, S.P.; Wang, Z.; Chinwalla, A.T.; Minx, P.; et al. Comparative and demographic analysis of orang-utan genomes. *Nature* **2011**, *469*, 529–533. [[CrossRef](#)]
17. Tolomeo, D.; Capozzi, O.; Stanyon, R.R.; Archidiacono, N.; D’Addabbo, P.; Catacchio, C.R.; Purgato, S.; Perini, G.; Schempp, W.; Huddleston, J.; et al. Epigenetic origin of evolutionary novel centromeres. *Sci. Rep.* **2017**, *7*, 41980. [[CrossRef](#)]
18. Zhang, H.; Koblížková, A.; Wang, K.; Gong, Z.; Oliveira, L.; Torres, G.A.; Wu, Y.; Zhang, W.; Novák, P.; Buell, C.R.; et al. Boom-Bust Turnovers of Megabase-Sized Centromeric DNA in *Solanum* Species: Rapid Evolution of DNA Sequences Associated with Centromeres. *Plant Cell* **2014**, *26*, 1436–1447. [[CrossRef](#)]
19. Gong, Z.; Wu, Y.; Koblížková, A.; Torres, G.A.; Wang, K.; Iovene, M.; Neumann, P.; Zhang, W.; Novák, P.; Buell, C.R.; et al. Repeatless and repeat-based centromeres in potato: Implications for centromere evolution. *Plant Cell* **2012**, *24*, 3559–3574. [[CrossRef](#)]
20. 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](#)]
21. Fu, S.; Lv, Z.; Gao, Z.; Wu, H.; Pang, J.; Zhang, B.; Dong, Q.; Guo, X.; Wang, X.J.; Birchler, J.A.; et al. De novo centromere formation on a chromosome fragment in maize. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 6033–6036. [[CrossRef](#)] [[PubMed](#)]
22. Liu, Y.; Su, H.; Pang, J.; Gao, Z.; Wang, X.J.; Birchler, J.A.; Han, F. Sequential de novo centromere formation and inactivation on a chromosomal fragment in maize. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1263–E1271. [[CrossRef](#)] [[PubMed](#)]
23. Allshire, R.C.; Karpen, G.H. Epigenetic regulation of centromeric chromatin: Old dogs, new tricks? *Nat. Rev. Genet.* **2008**, *9*, 923–937. [[CrossRef](#)] [[PubMed](#)]
24. Heun, P.; Erhardt, S.; Blower, M.D.; Weiss, S.; Skora, A.D.; Karpen, G.H. Mislocalization of the *Drosophila* centromere-specific histone CID promotes formation of functional ectopic kinetochores. *Dev. Cell* **2006**, *10*, 303–315. [[CrossRef](#)]
25. Moreno-Moreno, O.; Torras-Llort, M.; Azorín, F. Proteolysis restricts localization of CID, the centromere-specific histone H3 variant of *Drosophila*, to centromeres. *Nucleic Acids Res.* **2006**, *34*, 6247–6255. [[CrossRef](#)]
26. Collins, K.A.; Camahort, R.; Seidel, C.; Gerton, J.L.; Biggins, S. The overexpression of a *Saccharomyces cerevisiae* centromeric histone H3 variant mutant protein leads to a defect in kinetochore biorientation. *Genetics* **2007**, *175*, 513–525. [[CrossRef](#)]
27. Van Hooser, A.A.; Ouspenski, I.I.; Gregson, H.C.; Starr, D.A.; Yen, T.J.; Goldberg, M.L.; Yokomori, K.; Earnshaw, W.C.; Sullivan, K.F.; Brinkley, B.R. Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. *J. Cell Sci.* **2001**, *114*, 3529–3542. [[CrossRef](#)]
28. Gascoigne, K.E.; Takeuchi, K.; Suzuki, A.; Hori, T.; Fukagawa, T.; Cheeseman, I.M. Induced ectopic kinetochore assembly bypasses the requirement for CENP-A nucleosomes. *Cell* **2011**, *145*, 410–422. [[CrossRef](#)]
29. Shrestha, R.L.; Ahn, G.S.; Staples, M.I.; Sathyan, K.M.; Karpova, T.S.; Foltz, D.R.; Basrai, M.A. Mislocalization of centromeric histone H3 variant CENP-A contributes to chromosomal instability (CIN) in human cells. *Oncotarget* **2017**, *8*, 46781–46800. [[CrossRef](#)]
30. Biscotti, M.A.; Olmo, E.; Heslop-Harrison, J.S. Repetitive DNA in eukaryotic genomes. *Chromosome Res.* **2015**, *23*, 415–420. [[CrossRef](#)]
31. Garrido-Ramos, M.A. Satellite DNA: An Evolving Topic. *Genes* **2017**, *8*, 230. [[CrossRef](#)] [[PubMed](#)]
32. Hartley, G.; O’Neill, R.J. Centromere Repeats: Hidden Gems of the Genome. *Genes* **2019**, *10*, 223. [[CrossRef](#)] [[PubMed](#)]
33. Dover, G.A.; Tautz, D. Conservation and divergence in multigene families: Alternatives to selection and drift. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1986**, *312*, 275–289. [[CrossRef](#)]
34. Plohl, M.; Meštrović, N.; Mravinac, B. Centromere identity from the DNA point of view. *Chromosoma* **2014**, *123*, 313–325. [[CrossRef](#)] [[PubMed](#)]

35. Cacheux, L.; Ponger, L.; Gerbault-Seureau, M.; Loll, F.; Gey, D.; Richard, F.A.; Escudé, C. The Targeted Sequencing of Alpha Satellite DNA in *Cercopithecus pogonias* Provides New Insight Into the Diversity and Dynamics of Centromeric Repeats in Old World Monkeys. *Genome Biol. Evol.* **2018**, *10*, 1837–1851. [[CrossRef](#)] [[PubMed](#)]
36. Talbert, P.B.; Henikoff, S. What makes a centromere? *Exp. Cell Res.* **2020**, *389*, 111895. [[CrossRef](#)] [[PubMed](#)]
37. 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](#)]
38. Jones, K.W. Chromosomal and nuclear location of mouse satellite DNA in individual cells. *Nature* **1970**, *225*, 912–915. [[CrossRef](#)]
39. Pardue, M.L.; Gall, J.G. Chromosomal localization of mouse satellite DNA. *Science* **1970**, *168*, 1356–1358. [[CrossRef](#)]
40. Yunis, J.J.; Yasmineh, W.G. Heterochromatin, satellite DNA, and cell function. Structural DNA of eucaryotes may support and protect genes and aid in speciation. *Science* **1971**, *174*, 1200–1209. [[CrossRef](#)]
41. 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](#)] [[PubMed](#)]
42. Martínez-Balbás, A.; Rodríguez-Campos, A.; García-Ramírez, M.; Sainz, J.; Carrera, P.; Aymamí, J.; Azorín, F. Satellite DNAs contain sequences that induced curvature. *Biochemistry* **1990**, *29*, 2342–2348. [[CrossRef](#)] [[PubMed](#)]
43. Fitzgerald, D.J.; Dryden, G.L.; Bronson, E.C.; Williams, J.S.; Anderson, J.N. Conserved patterns of bending in satellite and nucleosome positioning DNA. *J. Biol. Chem.* **1994**, *269*, 21303–21314. [[CrossRef](#)]
44. Heslop-Harrison, J.S.; Schwarzscher, T. Nucleosomes and centromeric DNA packaging. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19974–19975. [[CrossRef](#)] [[PubMed](#)]
45. Varadaraj, K.; Skinner, D.M. Cytoplasmic localization of transcripts of a complex G+C-rich crab satellite DNA. *Chromosoma* **1994**, *103*, 423–431. [[CrossRef](#)] [[PubMed](#)]
46. Rudert, F.; Bronner, S.; Garnier, J.M.; Dollé, P. Transcripts from opposite strands of gamma satellite DNA are differentially expressed during mouse development. *Mamm. Genome* **1995**, *6*, 76–83. [[CrossRef](#)] [[PubMed](#)]
47. Lorite, P.; Renault, S.; Rouleux-Bonnin, F.; Bigot, S.; Periquet, G.; Palomeque, T. Genomic organization and transcription of satellite DNA in the ant *Aphaenogaster subterranea* (Hymenoptera, Formicidae). *Genome* **2002**, *45*, 609–616. [[CrossRef](#)]
48. Topp, C.N.; Zhong, C.X.; Dawe, R.K. Centromere-encoded RNAs are integral components of the maize kinetochore. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15986–15991. [[CrossRef](#)]
49. Pathak, D.; Srivastava, J.; Premi, S.; Tiwari, M.; Garg, L.C.; Kumar, S.; Ali, S. Chromosomal localization, copy number assessment, and transcriptional status of BamHI repeat fractions in water buffalo *Bubalus bubalis*. *DNA Cell Biol.* **2006**, *25*, 206–214. [[CrossRef](#)]
50. Lee, H.R.; Neumann, P.; Macas, J.; Jiang, J. Transcription and evolutionary dynamics of the centromeric satellite repeat CentO in rice. *Mol. Biol. Evol.* **2006**, *23*, 2505–2520. [[CrossRef](#)]
51. Ideue, T.; Cho, Y.; Nishimura, K.; Tani, T. Involvement of satellite I noncoding RNA in regulation of chromosome segregation. *Genes Cells* **2014**, *19*, 528–538. [[CrossRef](#)] [[PubMed](#)]
52. Cerutti, F.; Gamba, R.; Mazzagatti, A.; Piras, F.M.; Cappelletti, E.; Belloni, E.; Nergadze, S.G.; Raimondi, E.; Giulotto, E. The major horse satellite DNA family is associated with centromere competence. *Mol. Cytogenet.* **2016**, *9*, 35. [[CrossRef](#)] [[PubMed](#)]
53. Bobkov, G.O.M.; 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](#)]
54. Rošić, S.; Erhardt, S. No longer a nuisance: Long non-coding RNAs join CENP-A in epigenetic centromere regulation. *Cell. Mol. Life Sci.* **2016**, *73*, 1387–1398. [[CrossRef](#)] [[PubMed](#)]
55. Perea-Resa, C.; Blower, M.D. Centromere Biology: Transcription Goes on Stage. *Mol. Cell. Biol.* **2018**, *38*, e00263-18. [[CrossRef](#)]
56. Thakur, J.; Packiaraj, J.; Henikoff, S. Sequence, Chromatin and Evolution of Satellite DNA. *Int. J. Mol. Sci.* **2021**, *22*, 4309. [[CrossRef](#)]
57. 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](#)]
58. 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.e6. [[CrossRef](#)]
59. Alonso, A.; Hasson, D.; Cheung, F.; Warburton, P.E. A paucity of heterochromatin at functional human neocentromeres. *Epigenet. Chromatin* **2010**, *3*, 6. [[CrossRef](#)]
60. Burrack, L.S.; Berman, J. Neocentromeres and epigenetically inherited features of centromeres. *Chromosome Res.* **2012**, *20*, 607–619. [[CrossRef](#)]
61. Krasikova, A.; Fukagawa, T.; Zlotina, A. High-resolution mapping and transcriptional activity analysis of chicken centromere sequences on giant lampbrush chromosomes. *Chromosome Res.* **2012**, *20*, 995–1008. [[CrossRef](#)] [[PubMed](#)]
62. Murillo-Pineda, M.; Valente, L.P.; Dumont, M.; Mata, J.F.; Fachinetti, D.; Jansen, L.E.T. Induction of spontaneous human neocentromere formation and long-term maturation. *J. Cell Biol.* **2021**, *220*, e202007210. [[CrossRef](#)] [[PubMed](#)]
63. Saffery, R.; Sumer, H.; Hassan, S.; Wong, L.H.; Craig, J.M.; Todokoro, K.; Anderson, M.; Stafford, A.; Choo, K.H. Transcription within a functional human centromere. *Mol. Cell* **2003**, *12*, 509–516. [[CrossRef](#)]

64. Cardone, M.F.; Alonso, A.; Paziienza, M.; Ventura, M.; Montemurro, G.; Carbone, L.; de Jong, P.J.; Stanyon, R.; D'Addabbo, P.; Archidiacono, N.; et al. Independent centromere formation in a capricious, gene-free domain of chromosome 13q21 in Old World monkeys and pigs. *Genome Biol.* **2006**, *7*, R91. [[CrossRef](#)]
65. Kursel, L.E.; Malik, H.S. The cellular mechanisms and consequences of centromere drive. *Curr. Opin. Cell Biol.* **2018**, *52*, 58–65. [[CrossRef](#)]
66. Malik, H.S.; Bayes, J.J. Genetic conflicts during meiosis and the evolutionary origins of centromere complexity. *Biochem. Soc. Trans.* **2006**, *34*, 569–573. [[CrossRef](#)]
67. Lampson, M.A.; Black, B.E. Cellular and Molecular Mechanisms of Centromere Drive. *Cold Spring Harb. Symp. Quant. Biol.* **2017**, *82*, 249–257. [[CrossRef](#)]
68. Fishman, L.; Saunders, A. Centromere-associated female meiotic drive entails male fitness costs in monkeyflowers. *Science* **2008**, *322*, 1559–1562. [[CrossRef](#)]
69. Iwata-Otsubo, A.; Dawicki-McKenna, J.M.; Aker, 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.e8. [[CrossRef](#)]
70. Aker, T.; Chmátal, L.; Trimm, E.; Yang, K.; Aonbangkhen, C.; Chenoweth, D.M.; Janke, C.; Schultz, R.M.; Lampson, M.A. Spindle asymmetry drives non-Mendelian chromosome segregation. *Science* **2017**, *358*, 668–672. [[CrossRef](#)]
71. Aker, T.; Trimm, E.; Lampson, M.A. Molecular Strategies of Meiotic Cheating by Selfish Centromeres. *Cell* **2019**, *178*, 1132–1144.e10. [[CrossRef](#)] [[PubMed](#)]
72. Orgel, L.E.; Crick, F.H. Selfish DNA: The ultimate parasite. *Nature* **1980**, *284*, 604–607. [[CrossRef](#)] [[PubMed](#)]
73. Jagannathan, M.; Yamashita, Y.M. Function of Junk: Pericentromeric Satellite DNA in Chromosome Maintenance. *Cold Spring Harb. Symp. Quant. Biol.* **2017**, *82*, 319–327. [[CrossRef](#)] [[PubMed](#)]
74. Amor, D.J.; Choo, K.H. Neocentromeres: Role in human disease, evolution, and centromere study. *Am. J. Hum. Genet.* **2002**, *71*, 695–714. [[CrossRef](#)] [[PubMed](#)]
75. Warburton, P.E.; Cooke, C.A.; Bourassa, S.; Vafa, O.; Sullivan, B.A.; Stetten, G.; Gimelli, G.; Warburton, D.; Tyler-Smith, C.; Sullivan, K.F.; et al. Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. *Curr. Biol.* **1997**, *7*, 901–904. [[CrossRef](#)]
76. Rocchi, M.; Archidiacono, N.; Schempp, W.; Capozzi, O.; Stanyon, R. Centromere repositioning in mammals. *Heredity* **2012**, *108*, 59–67. [[CrossRef](#)]
77. Ventura, M.; Antonacci, F.; Cardone, M.F.; Stanyon, R.; D'Addabbo, P.; Cellamare, A.; Sprague, L.J.; Eichler, E.E.; Archidiacono, N.; Rocchi, M. Evolutionary formation of new centromeres in macaque. *Science* **2007**, *316*, 243–246. [[CrossRef](#)]
78. Ventura, M.; Weigl, S.; Carbone, L.; Cardone, M.F.; Misceo, D.; Teti, M.; D'Addabbo, P.; Wandall, A.; Björck, E.; de Jong, P.J.; et al. Recurrent sites for new centromere seeding. *Genome Res.* **2004**, *14*, 1696–1703. [[CrossRef](#)]
79. Capozzi, O.; Purgato, S.; Verdun di Cantogno, L.; Grosso, E.; Ciccone, R.; Zuffardi, O.; Della Valle, G.; Rocchi, M. Evolutionary and clinical neocentromeres: Two faces of the same coin? *Chromosoma* **2008**, *117*, 339–344. [[CrossRef](#)]
80. Montefalcone, G.; Tempesta, S.; Rocchi, M.; Archidiacono, N. Centromere repositioning. *Genome Res.* **1999**, *9*, 1184–1188. [[CrossRef](#)]
81. Schubert, I. What is behind “centromere repositioning”? *Chromosoma* **2018**, *127*, 229–234. [[CrossRef](#)] [[PubMed](#)]
82. Peng, S.; Petersen, J.L.; Bellone, R.R.; Kalbfleisch, T.; Kingsley, N.B.; Barber, A.M.; Cappelletti, E.; Giulotto, E.; Finno, C.J. Decoding the Equine Genome: Lessons from ENCODE. *Genes* **2021**, *12*, 1707. [[CrossRef](#)] [[PubMed](#)]
83. Giulotto, E.; Raimondi, E.; Sullivan, K.F. The Unique DNA Sequences Underlying Equine Centromeres. *Prog. Mol. Subcell. Biol.* **2017**, *56*, 337–354. [[CrossRef](#)] [[PubMed](#)]
84. Ransom, J.I.; Kaczensky, P. *Wild Equids: Ecology, Management, and Conservation*; JHU Press: Baltimore, MD, USA, 2016; 229p.
85. Trifonov, V.A.; Stanyon, R.; Nesterenko, A.I.; Fu, B.; Perelman, P.L.; O'Brien, P.C.; Stone, G.; Rubtsova, N.V.; Houck, M.L.; Robinson, T.J.; et al. Multidirectional cross-species painting illuminates the history of karyotypic evolution in Perissodactyla. *Chromosome Res.* **2008**, *16*, 89–107. [[CrossRef](#)] [[PubMed](#)]
86. Orlando, L.; Ginolhac, A.; Zhang, G.; Froese, D.; Albrechtsen, A.; Stiller, M.; Schubert, M.; Cappelletti, E.; Petersen, B.; Moltke, I.; et al. Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **2013**, *499*, 74–78. [[CrossRef](#)] [[PubMed](#)]
87. Jónsson, H.; Schubert, M.; Seguin-Orlando, A.; Ginolhac, A.; Petersen, L.; Fumagalli, M.; Albrechtsen, A.; Petersen, B.; Korneliusen, T.S.; Vilstrup, J.T.; et al. Speciation with gene flow in equids despite extensive chromosomal plasticity. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 18655–18660. [[CrossRef](#)] [[PubMed](#)]
88. Trifonov, V.A.; Musilova, P.; Kulemsina, A.I. Chromosome evolution in Perissodactyla. *Cytogenet. Genome Res.* **2012**, *137*, 208–217. [[CrossRef](#)]
89. Santagostino, M.; Khoraiuli, L.; Gamba, R.; Bonuglia, M.; Klipstein, O.; Piras, F.M.; Vella, F.; Russo, A.; Badiale, C.; Mazzagatti, A.; et al. Genome-wide evolutionary and functional analysis of the Equine Repetitive Element 1: An insertion in the myostatin promoter affects gene expression. *BMC Genet.* **2015**, *16*, 126. [[CrossRef](#)]
90. Santagostino, M.; Piras, F.M.; Cappelletti, E.; Del Giudice, S.; Semino, O.; Nergadze, S.G.; Giulotto, E. Insertion of Telomeric Repeats in the Human and Horse Genomes: An Evolutionary Perspective. *Int. J. Mol. Sci.* **2020**, *21*, 2838. [[CrossRef](#)]

91. Sola, L.; Nergadze, S.G.; Cappelletti, E.; Piras, F.M.; Giulotto, E.; Santagostino, M. Telomeric-Like Repeats Flanked by Sequences Retrotranscribed from the Telomerase RNA Inserted at DNA Double-Strand Break Sites during Vertebrate Genome Evolution. *Int. J. Mol. Sci.* **2021**, *22*, 11048. [[CrossRef](#)]
92. Nergadze, S.G.; Lupotto, M.; Pellanda, P.; Santagostino, M.; Vitelli, V.; Giulotto, E. Mitochondrial DNA insertions in the nuclear horse genome. *Anim. Genet.* **2010**, *41* (Suppl. S2), 176–185. [[CrossRef](#)] [[PubMed](#)]
93. Nergadze, S.G.; Santagostino, M.A.; Salzano, A.; Mondello, C.; Giulotto, E. Contribution of telomerase RNA retrotranscription to DNA double-strand break repair during mammalian genome evolution. *Genome Biol.* **2007**, *8*, R260. [[CrossRef](#)] [[PubMed](#)]
94. Nergadze, S.G.; Rocchi, M.; Azzalin, C.M.; Mondello, C.; Giulotto, E. Insertion of telomeric repeats at intrachromosomal break sites during primate evolution. *Genome Res.* **2004**, *14*, 1704–1710. [[CrossRef](#)] [[PubMed](#)]
95. Musilova, P.; Kubickova, S.; Vahala, J.; Rubes, J. Subchromosomal karyotype evolution in Equidae. *Chromosome Res.* **2013**, *21*, 175–187. [[CrossRef](#)] [[PubMed](#)]
96. Carbone, L.; Nergadze, S.G.; Magnani, E.; Misceo, D.; Francesca Cardone, M.; Roberto, R.; Bertoni, L.; Attolini, C.; Francesca Piras, M.; de Jong, P.; et al. Evolutionary movement of centromeres in horse, donkey, and zebra. *Genomics* **2006**, *87*, 777–782. [[CrossRef](#)] [[PubMed](#)]
97. Piras, F.M.; Nergadze, S.G.; Poletto, V.; Cerutti, F.; Ryder, O.A.; Leeb, T.; Raimondi, E.; Giulotto, E. Phylogeny of horse chromosome 5q in the genus *Equus* and centromere repositioning. *Cytogenet. Genome Res.* **2009**, *126*, 165–172. [[CrossRef](#)] [[PubMed](#)]
98. Purgato, S.; Belloni, E.; Piras, F.M.; Zoli, M.; Badiale, C.; Cerutti, F.; Mazzagatti, A.; Perini, G.; Della Valle, G.; Nergadze, S.G.; et al. Centromere sliding on a mammalian chromosome. *Chromosoma* **2015**, *124*, 277–287. [[CrossRef](#)]
99. Maloney, K.A.; Sullivan, L.L.; Matheny, J.E.; Strome, E.D.; Merrett, S.L.; Ferris, A.; Sullivan, B.A. Functional epialleles at an endogenous human centromere. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 13704–13709. [[CrossRef](#)]
100. Chueh, A.C.; Northrop, E.L.; Brettingham-Moore, K.H.; Choo, K.H.; Wong, L.H. LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. *PLoS Genet.* **2009**, *5*, e1000354. [[CrossRef](#)]
101. Kasinathan, S.; Henikoff, S. Non-B-Form DNA Is Enriched at Centromeres. *Mol. Biol. Evol.* **2018**, *35*, 949–962. [[CrossRef](#)]
102. Barbosa, A.C.; Xu, Z.; Karari, K.; Williams, W.; Hauf, S.; Brown, W.R.A. Mutation and selection explain why many eukaryotic centromeric DNA sequences are often A + T rich. *Nucleic Acids Res.* **2022**, *50*, 579–596. [[CrossRef](#)] [[PubMed](#)]
103. Chueh, A.C.; Wong, L.H.; Wong, N.; Choo, K.H. Variable and hierarchical size distribution of L1-retroelement-enriched CENP-A clusters within a functional human neocentromere. *Hum. Mol. Genet.* **2005**, *14*, 85–93. [[CrossRef](#)] [[PubMed](#)]
104. 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](#)]
105. Van de Werken, H.J.G.; Haan, J.C.; Feodorova, Y.; Bijos, D.; Weuts, A.; Theunis, K.; Holwerda, S.J.B.; Meuleman, W.; Pagie, L.; Thanisch, K.; et al. Small chromosomal regions position themselves autonomously according to their chromatin class. *Genome Res.* **2017**, *27*, 922–933. [[CrossRef](#)]
106. Zeitlin, S.G.; Baker, N.M.; Chapados, B.R.; Soutoglou, E.; Wang, J.Y.; Berns, M.W.; Cleveland, D.W. Double-strand DNA breaks recruit the centromeric histone CENP-A. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 15762–15767. [[CrossRef](#)]
107. Vidale, P.; Magnani, E.; Nergadze, S.G.; Santagostino, M.; Cristofari, G.; Smirnova, A.; Mondello, C.; Giulotto, E. The catalytic and the RNA subunits of human telomerase are required to immortalize equid primary fibroblasts. *Chromosoma* **2012**, *121*, 475–488. [[CrossRef](#)]
108. Anglana, M.; Bertoni, L.; Giulotto, E. Cloning of a polymorphic sequence from the nontranscribed spacer of horse rDNA. *Mamm. Genome* **1996**, *7*, 539–541. [[CrossRef](#)]
109. Nergadze, S.G.; Belloni, E.; Piras, F.M.; Khoraiuli, L.; Mazzagatti, A.; Vella, F.; Bensi, M.; Vitelli, V.; Giulotto, E.; Raimondi, E. Discovery and comparative analysis of a novel satellite, EC137, in horses and other equids. *Cytogenet. Genome Res.* **2014**, *144*, 114–123. [[CrossRef](#)]
110. Shepelev, V.A.; Alexandrov, A.A.; Yurov, Y.B.; Alexandrov, I.A. The evolutionary origin of man can be traced in the layers of defunct ancestral alpha satellites flanking the active centromeres of human chromosomes. *PLoS Genet.* **2009**, *5*, e1000641. [[CrossRef](#)]
111. Rošić, 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](#)]
112. Roberti, A.; Bensi, M.; Mazzagatti, A.; Piras, F.M.; Nergadze, S.G.; Giulotto, E.; Raimondi, E. Satellite DNA at the Centromere is Dispensable for Segregation Fidelity. *Genes* **2019**, *10*, 469. [[CrossRef](#)] [[PubMed](#)]
113. Choo, K.H. Why is the centromere so cold? *Genome Res.* **1998**, *8*, 81–82. [[CrossRef](#)] [[PubMed](#)]
114. Talbert, P.B.; Henikoff, S. Centromeres convert but don't cross. *PLoS Biol.* **2010**, *8*, e1000326. [[CrossRef](#)] [[PubMed](#)]
115. Cappelletti, E.; Piras, F.M.; Badiale, C.; Bambi, M.; Santagostino, M.; Vara, C.; Masterson, T.A.; Sullivan, K.F.; Nergadze, S.G.; Ruiz-Herrera, A.; et al. CENP-A binding domains and recombination patterns in horse spermatocytes. *Sci. Rep.* **2019**, *9*, 15800. [[CrossRef](#)] [[PubMed](#)]
116. Kurdzo, E.L.; Dawson, D.S. Centromere pairing—Tethering partner chromosomes in meiosis I. *FEBS J.* **2015**, *282*, 2458–2470. [[CrossRef](#)] [[PubMed](#)]
117. Da Ines, O.; White, C.I. Centromere Associations in Meiotic Chromosome Pairing. *Annu. Rev. Genet.* **2015**, *49*, 95–114. [[CrossRef](#)]

118. Borodin, P.M.; Karamysheva, T.V.; Belonogova, N.M.; Torgasheva, A.A.; Rubtsov, N.B.; Searle, J.B. Recombination map of the common shrew, *Sorex araneus* (Eulipotyphla, Mammalia). *Genetics* **2008**, *178*, 621–632. [[CrossRef](#)]
119. Bikchurina, T.I.; Tishakova, K.V.; Kizilova, E.A.; Romanenko, S.A.; Serdyukova, N.A.; Torgasheva, A.A.; Borodin, P.M. Chromosome Synapsis and Recombination in Male-Sterile and Female-Fertile Interspecies Hybrids of the Dwarf Hamsters (*Phodopus*, Cricetidae). *Genes* **2018**, *9*, 227. [[CrossRef](#)]