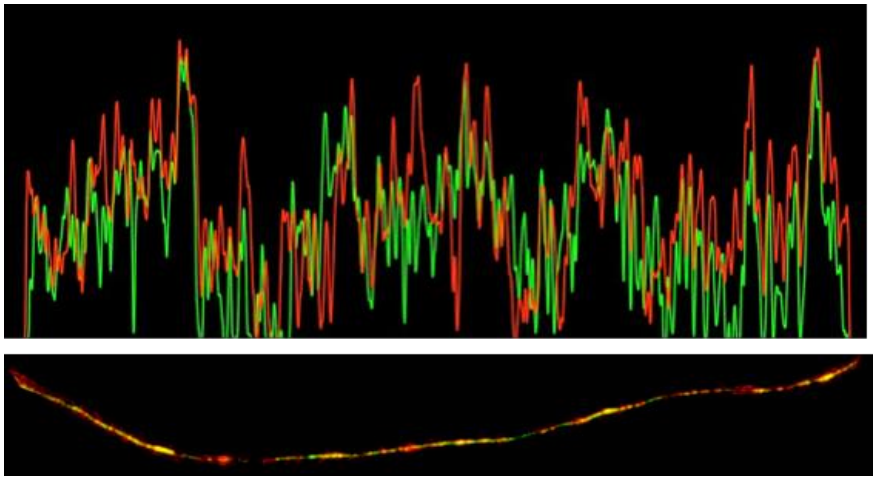
UNIVERSITÀ
DI PAVIA

Dipartimento di Biologia e Biotecnologie “Lazzaro Spallanzani”

Molecular cytogenetic investigation of mammalian centromere biology: the *Equus* species as a model system



Annalisa Manuela Roberti

Dottorato di Ricerca in
Genetica, Biologia Molecolare e Cellulare
Ciclo XXXII – A.A. 2016-2019

Dipartimento di Biologia e Biotecnologie “Lazzaro Spallanzani”

Molecular cytogenetic investigation of mammalian centromere biology: the *Equus* species as a model system

Annalisa Manuela Roberti

Supervised by Prof. Elena Raimondi

Dottorato di Ricerca in
Genetica, Biologia Molecolare e Cellulare
Ciclo XXXII – A.A. 2016-2019.

CONTENTS

ABSTRACT

The centromere is the chromosomal *locus* which drives proper genome inheritance during eukaryotic cell division. Most vertebrate centromeres are composed of extended arrays of tandemly repeated DNA sequences named satellite DNA, suggesting that satellite DNA may contribute to the establish centromere function. However, while centromere function is highly conserved during evolution, centromeric DNA sequences are extremely variable across species: this contradiction in known as "the centromere paradox". Indeed, it has been unequivocally demonstrated that centromere identity is determined by epigenetic marks, the most important of which is the CENP-A protein, a centromere specific variant of the H3 histone.

This PhD project arises from a collaboration between the laboratory of Molecular Cytogenetics directed by Professor Elena Raimondi and the laboratory of Molecular and Cellular Biology directed by Professor Elena Giulotto, aimed at studying the molecular organization of mammalian centromeres using the equid species as a model system. Species belonging to the genus *Equus* are an unprecedented model for the analysis of mammalian centromere birth, evolution and complete maturation. Indeed, an exceedingly high number of centromere repositioning events occurred during their evolution leading to the formation of evolutionarily new centromeres void of satellite DNA (1 out of 32 in the horse, 16 out of 31 in the donkey, 9 out of 23 in the Grevyi's zebra, 11 out of 22 in the Burchell's zebra, 10 out of 28 in the Hartmann's zebra and 14 out of 25 in the African donkey).

Taking advantage of the equid model system, we firstly investigated one of the most intriguing opened questions in mammalian centromere biology: which is the role played by satellite DNA at the centromere? Indeed, the existence in equid karyotypes of canonical satellite-based centromeres together with evolutionarily new satellite-free centromeres, offers the opportunity to directly compare their behaviour during cell division. Literature data mainly concern pathological and experimental human neocentromeres, but, to our knowledge, no systematic analysis has ever been carried out on the stability of satellite-free natural centromeres.

Here I analysed the mitotic stability of horse chromosome 11 (ECA11), whose centromere is satellite-free, and compared it with that of chromosome 13 (ECA13) which is similar in size and has a centromere containing long stretches of the canonical horse centromeric satellite DNA. The results, obtained by two different chromosome segregation assays, under normal and mitotic stress conditions, demonstrated that the segregation accuracy of these two chromosomes is similar, thus suggesting that satellite DNA is dispensable for transmission fidelity.

As mentioned above, centromere function is determined by epigenetic factors. In the second part of my PhD project, I studied the epigenetic landscape of horse and donkey centromeres by means of high resolution molecular cytogenetic approaches. Contrary to molecular and next generation sequencing approaches, molecular cytogenetics allows to directly compare the architectural organization of satellite-based and satellite-free centromeres. Moreover, information about DNA protein and protein-protein interactions can be obtained. My results demonstrated that satellite-free centromeres, as well as the satellite-based ones, are immersed in a heterochromatic environment, even if they contain much smaller amounts of constitutive heterochromatin. This implicates that, independently from sequence composition, heterochromatin specific post-translational histone modifications are needed to drive centromere function. Satellite-less centromeres also contain facultative heterochromatin marks thus indicating that centrochromatin has a partially opened conformation that is needed for CENP-A loading. Markers of transcriptional competence are also present at horse and donkey satellite-free and satellite-based centromeres. However, we observed a high variability in the amount of these modifications among centromeres an among different cells, independently from the presence or absence of satellite DNA. It can be hypothesized that the centromeric chromatin environment varies in different phases of the cell cycle. It is also tempting to speculate that, in each cell, a threshold amount of centromere transcripts is needed for proper centromere function, independently from which centromeric DNA cluster is actively transcribed. In this hypothesis, it must

be assumed that, in analogy with what demonstrated in man, centromere transcripts can act in *trans* also in equids. Finally, all horse and donkey centromeres are associated to the H4K20me1 histone modification which perfectly co-localizes with the CENP-A protein. This result is in agreement with literature data that demonstrate that this histone marker is needed at for the maturation of CENP-A nucleosomes and for kinetochore assembly.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Elena Raimondi for supporting me scientifically and for being a “guide”. I will always be grateful to her.

I would like to thank also Prof. Elena Giulotto and her laboratory for precious collaboration.

ABBREVIATIONS

BAC	Bacterial Artificial Chromosome
CATD	CENP-A Targeting Domain
CCAN	Constitutive-Centromere-Associated-Network
CENP	CENtromere Protein
ChIP-seq	Chromatin ImmunoPrecipitation - Sequencing
CREST	Calcinosis, Raynaud's syndrome, Esophageal dysmotility, Sclerodactyly and Telangiectasia
CBMN	Cytokinesis-Blocked MicroNucleus
DAPI	4',6-diamidino-2-phenylindole
EAS	*Equus asinus*
ECA	*Equus caballus*
ENCs	Evolutionarily New Centromeres
FISH	Fluorescence *In Situ* Hybridization
HAC	Human Artificial Chromosome
HJURP	Holliday Junction Recognition Protein
INCENP	INner CENtromere Protein
non-LTR	Non-Long Terminal Repeat
LINE	Long Interspersed Nuclear Elements
OWM	Old World Monkeys
PAK	*Perissodactyla* Ancestral Karyotype
SINE	Short Interspersed Elements
SNBs	Stress-induced Nuclear Bodies
TE	Transposable Elements

1. REVIEW OF THE LITERATURE

1.1 THE CENTROMERE

The centromere is an essential chromosomal *locus* which ensures accurate segregation of replicated chromosomes to daughter cells during cell division (Plohl *et al*., 2014).

The centromere plays its role by recruiting the kinetochore, a proteinaceous macromolecular structure that mediates the attachment of the centromere chore to the microtubules of the mitotic and meiotic spindle (Fig. 1). Defects in centromere or kinetochore function can lead to the loss of genomic information, resulting in developmental defects and disease.

The centromere was first observed by light microscopy as the primary constriction of condensed mitotic chromosomes (Flemming, 1882).

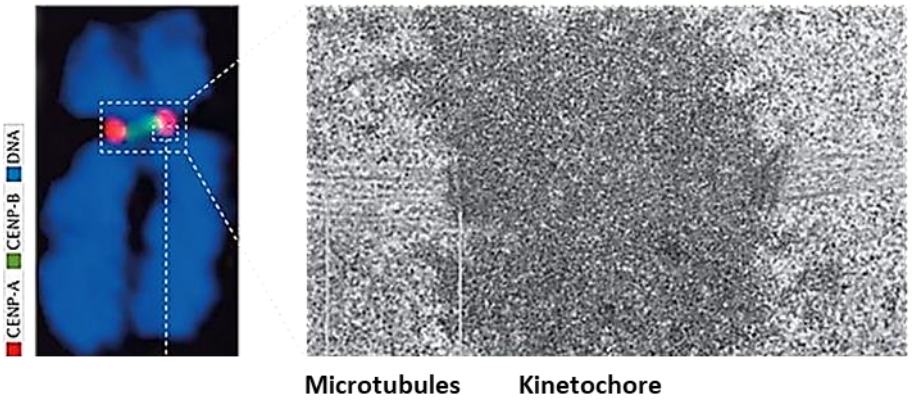


Figure 1 – Images of the centromere at increasing resolution. Left: immunofluorescence image of a metaphase chromosome where CENP-A is red labelled, and CENP-B is green labelled. Right: electron micrograph of the centromeric region of a metaphase chromosome showing centromeric chromatin (dark cloud), kinetochore and microtubules (adapted from McKinley and Cheeseman, 2016).

Eukaryotic centromeres are classified in three categories: point, regional and diffused (Pluta *et al*., 1995). Point centromeres are assembled on a specific short DNA sequence, which is sufficient for the centromeric function, while regional centromeres occupy extended regions containing highly repeated DNA sequences that are variable among species. On the contrary, diffused centromeres, are characteristic of "holocentric chromosomes", which have a kinetochore plate extending along the entire chromosome length. This type of centromeres occurs in a number of arthropods and plants as well as in the model organism *C. elegans*. (McKinley and Cheeseman, 2016).

Point centromeres are present only in the budding yeasts *Saccharomyces cerevisiae.* Point centromeres encompass about 125 bp and consist of three Conserved DNA Elements (CDEs): CDEI, CDEII, and CDEIII (Fig. 2-C). The centromeric sequence in *S. cerevisiae* is necessary and sufficient for kinetochore recruitment and for chromosome segregation (Hyman and Sorger, 1995). CDEI (8 bp) and CDEIII (25 bp) are conserved in all *S. cerevisiae* centromeres and point mutations in these two elements can abolish the function. CDE III binds a four subunits protein complex, Cbf3, which interacts with the Cbf1 protein which, in turn, binds the CDEI sequence (Hemmerich *et al*., 2000; Gardner *et al*., 2001). The resulting complex forms a scaffold for kinetochore assembly. CDEII is not conserved in terms of sequence, but in length (between 78 and 86 bp) and AT enrichment (90%). CDEII, interacts with the histone H3 centromere specific variant Cse4 (named CENP-A in vertebrates) which is the epigenetic marker of functional centromeres (Hyman and Sorger, 1995; Stoler *et al*., 1995; Cleveland *et al*., 2003).

A well-characterized regional centromere is the one of *Schizosaccharomyces pombe* (Figure 2-C), whose centromeric domain extends for 40-100 kb. The centromeres of *S. pombe* contain an AT rich 4-7 Kb long core sequence, flanked by inverted repeats called K', K'', L and B'. Progressive deletions have been introduced experimentally within these sequences to understand what is the minimum DNA sequence necessary for centromere function. Such experiments showed that at least part of the repeated sequence K'' is essential for a correct segregation. Instead, the role

of K'+ L sequences at mitosis was demonstrated to be very minimum, as deletions of one or both copies of K' and L have little or no effect on mitotic centromere function (Hahnenberger *et al*., 1991).

The centromeres of *Candida albicans* are regional as well and contain an about 3 kb long unique sequence (Figure 2-D), which is enriched in AT (65%) and resembles the sequence of the central core of *S. pombe* centromeres, however, *C. albicans* centromeres do not contain repeated sequences flanking the central core (Sanyal *et al*., 2004). On the other hand, the presence of repeated sequences at the centromere is a property of *Drosophyla melanogaster* centromeres (Figure 2-B), which are characterized by the presence of relatively short blocks of satellite DNA (AATAT and AAGAG) which, taken together, constitute a 220 kb sequence named "Bora Bora". For full centromere function Bora Bora, which nucleates kinetochore formation, must be flanked on either the 5' or 3' side by a 200 kb satellite DNA sequence which regulates sister chromatid pairing (Pluta *et al.,* 1995). Furthermore, transposable elements are interspersed within the centromeric DNA and represent 10% of it (Sun *et al*., 2003; Chang *et al*., 2019).

The centromeres of *Mus musculus* are a typical example of mammalian centromeres (Fig. 2-A). They are characterized by the repetition of long stretches of satellite DNA, distinguished in major satellite (made of 234 bp monomers) and minor satellite (120 bp monomers) and localized respectively in the pericentromeric and centromeric region, respectively. The major satellite is associated to the HP1 protein, responsible for maintaining the heterochromatic state, while the centromeric protein CENP-A is associated to the minor satellite (Guenatri *et al*., 2004).

A more complex centromeric organization is present in humans (Figure 2-A). Human centromeres contain extended satellite DNA arrays composed of 171 bp monomers arranged in a tandem head-to-tail fashion.

This satellite DNA is known as alpha satellite and arrays can cover mega base pairs. Individual monomers share 50%–70% sequence identity. A defined number of monomers can give rise to higher order repeats (HOR) that are themselves repeated so that at a given centromeric *locus*, the alpha

satellite array can span from 250 kb to 5,000 kb. Some monomers within each HOR contain a conserved 17 bp sequence, CENP-B box, which binds the protein CENP-B, that is thought to facilitate kinetochore formation even if its role is a matter of debate (Aldrup-Macdonald and Sullivan, 2014; Plohl *et al*., 2014). In the regions flanking the higher order arrays, monomers are randomly arranged in the pericentromeric region, and their sequence becomes progressively divergent from the consensus as the distance from the functional centromere core increases (Aldrup-Macdonald and Sullivan, 2014).

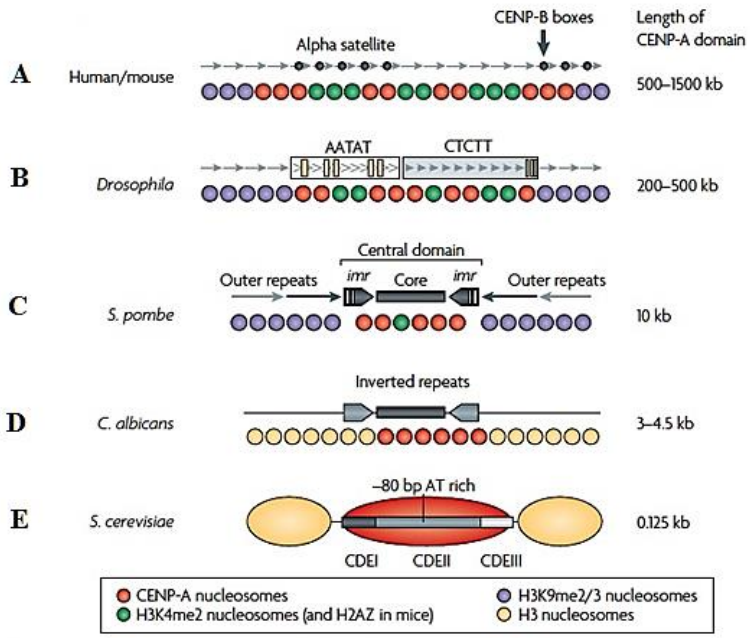


Figure 2 – Centromere structure and organization. Schematic representation of the DNA sequence arrangement at the centromere of different model organisms (Allshire and Karpen, 2008).

A number of mammalian and plant species have centromeres containing stretches of satellite DNA similar in their organization to the

human alpha satellite. For example, in *Arabidopsis thaliana* centromere repeats are made up of 178 bp long units, in Oryza sativa the repeated units are 155 bp long. However, even if the large majority of higher eukaryotic centromeres contain long stretches of satellite DNA, the size of the repeated stretches and the sequence composition of the repeated units is extremely variable (Kursel and Malik, 2016).

In some organisms the centromere is diffused along the entire length of the chromosome (holocentric) and therefore a primary constriction cannot be identified. Holocentricity has evolved multiple times in plants and insects (Kursel and Malik, 2016). The driving forces behind the transition from monocentric to holocentric chromosomes remain poorly understood (Pluta *et al.,* 1995).

At the sequence level, a common feature of the centromeres of the large majority of studied species is the presence of highly repeated sequences, however, these sequences are neither necessary nor sufficient to assembly a functional kinetochore (Choo, 2000; Plohl *et al*., 2014).

The discovery of active centromeres which are completely void of the canonical highly repeated DNA sequences, the so called “neocentromeres” is the proof that repeated sequences are not necessary for centromere function (Fig 3-C) (Voullaire *et al*., 1993; Marshall *et al*., 2008). On the other hand, dicentric chromosomes, with two identical and separate blocks of centromeric highly repeated DNA exist (Therman *et al*., 1986; Sullivan and Willard, 1998; Choo, 2000). These rearranged chromosomes are actually pseudo-dicentrics and can segregate correctly since only one centromeric satellite DNA block is able to recruit a functional kinetochore, while the other satellite-block is epigenetically inactive (Fig. 3-B). Pseudo-dicentrics demonstrate that satellite DNA at the centromere is not sufficient to establish centromere function (Fig. 3-B).

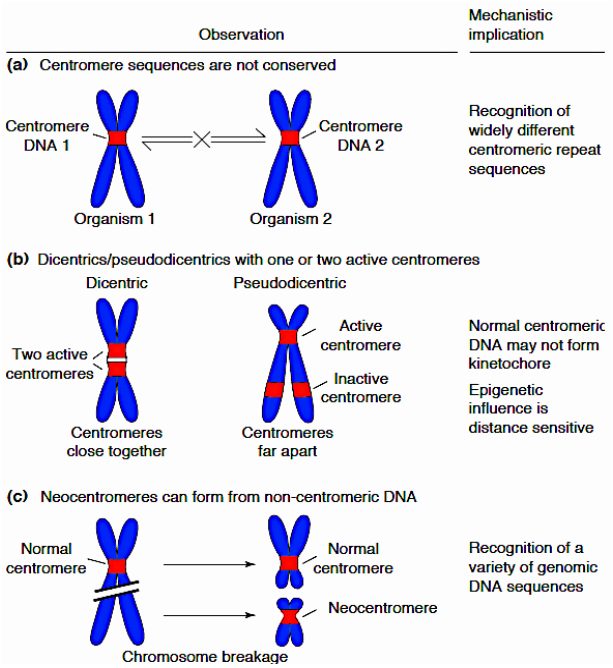


Figure 3 – Observations implicating an epigenetic control model for centromerization. (A) Absence of a conserved centromeric DNA sequence in different organisms and within an organism; (B) In dicentric chromosomes, the presence of two functionally active centromeres when the centromeres are close together, but not when they are far apart, suggests that centromerization is epigenetically controlled and the possibility that the control mechanism can spread over a distance beyond the core centromere region; (C) Neocentromere formation at many different non-centromeric sites suggests a versatility of centromerization that is independent of both the primary nucleotide sequence and the highly repetitive nature of the normal centromeric DNA (Choo, 2000).

All together, these evidences turn the centromere into a structure defined more epigenetically, than genetically.

Among the proteins which are constitutively or transiently associated to the centromeres, the most important epigenetic determinant of centromere function is CENP-A (also named CEN-H3), a centromere-specific

variant of the H3 histone (McKinley and Cheeseman, 2016). CEN-H3 is present at all active centromeres, independently from the underlying DNA sequence, and its function is critical for kinetochore assembly (Barra and Fachinetti, 2018).

Although histone H3 is evolutionarily conserved, the centromeric histone H3 variant in related species of mammals, flies, and plants diverge in the DNA binding region (Yoda *et al*., 2000; Henikoff and Malik, 2002; Malik *et al*., 2002;). Indeed, CENP-A interacts only with centromeric DNA sequences that are highly variable among species. In other words, the observed rapid evolution of centromeric DNA sequences is possible only assuming coevolution of CENP-A and other DNA-binding proteins (Dawe and Henikoff, 2006).

The phenomenon of centromere function conservation in spite of DNA sequence evolutionary divergence, is called “centromere paradox” (Henikoff *et al*., 2001). This paradox can be explained assuming that the higher order chromatin structure of the centromere drives its position and function (Grady *et al*., 1992; Koch, 2000). This hypothesis reconciles data showing the irrelevance of the primary DNA sequence for centromere function with data that demonstrate the conservation of centromeric functional domains (Lamb and Birchler, 2003).

1.2 CENTROMERIC AND PERICENTROMERIC REPETITIVE SEQUENCES

The “regional” centromeres, the most common type of eukaryotic centromeres, are characterised by the presence of two types of repetitive sequences: satellite DNA (satDNA) and transposable elements (TEs) (Pluta *et al*., 1995; Plohl *et al*., 2008). Satellite DNA is composed of tandemly repeated DNA sequences. In a large number of species, the length of the repeat units forming the centromeric satellite ranges from 170 to 340 bp, that is to say a DNA length compatible with that needed to wrap one or two nucleosomes (Henikoff *et al*., 2001). Some sequence motifs within satDNAs are conserved in evolutionarily distant species, suggesting that the evolution of such elements is influenced by selective constraints (Mravinac *et al*., 2005; Petraccioli *et al*., 2015). A well described example is the CENP-B box, a 17 bp motif binding the centromeric protein CENP-B. This sequence box was first described in human centromeric alpha-satellite DNA (Masumoto *et al*., 1989) and then found in other mammals (Earnshaw *et al*., 1987; Sullivan and Glass, 1991; Schueler *et al*., 2001), in insects (Lorite *et al*., 2002; Lorite *et al*., 2017), molluscs (Canapa *et al*., 2000) and nematodes (Mestrovic *et al*., 2013).

The second type of highly repeated sequences found in the centromeric and pericentromeric chromatin are transposable elements. Based on the mechanism by which TEs move in the genome, they are classified as transposons and retrotrasposons. Transposons can transpose via a “cut and paste” mechanism, while retroelements move via a “copy and paste” mechanism through an RNA intermediate which is reverse transcribed and then inserted in the genome. The so called “non-LTR elements” are the retroelements most represented at the centromere. They comprise two subtypes of retroelements: the short interspersed nuclear elements (SINE) and the long interspersed nuclear elements (LINE) (Biscotti *et al*., 2015; Scarpato *et al*., 2015). Retrotransposons characterize the centromeres of a wide range of angiosperms (Neumann *et al.,* 2011). In rice

and maize, all centromeres have different retrotransposon families (Nagaki *et al*., 2004b; Bao *et al*., 2006;), and retrotransposons are the main component of banana and some wheat centromeres (Cizkova *et al*., 2013; Li *et al*., 2013).

A common feature of centromeres containing satellite DNA is that the chromatin surrounding the functional centromere core is enriched in TEs. For example, human centromere cores contain stretches of alpha satellite DNA, while pericentromeres are composed of unrelated and heterogeneous satDNA families (e.g., gamma-satellite and SatIII) and LINE elements (Schueler *et al*., 2001). Phylogenetic analyses and TEs annotation demonstrated that when present in the centromere core, TEs have been recently inserted, while older retrotransposons typically lie outside the functional centromere (Wolfgruber *et al*., 2009) and show a distribution comparable to that observed in human pericentromeric regions (Schueler *et al.,* 2001).

1.2.1 CENTROMERIC AND PERICENTROMERIC SATELLITE DNA

Satellite DNA consists in repetitive DNA sequences arranged as very long arrays clustered in specific chromosomal regions. The total amount of satellite DNA in the genome of one eukaryotic species can highly vary. In plants, for example, satDNA can represent from 0.1% up to 36% of the genome (Macas *et al*., 2000; de la Herran *et al*., 2001; Levy *et al*., 2007; Hribova *et al*., 2010; Ambrozova *et al*., 2011; Cizkova *et al*., 2013; Emadzade *et al*., 2014; Garrido-Ramos, 2015; Miga, 2015). Within the animal kingdom, the total satDNA content ranges from less than 0.5% to more than 50%, both in invertebrates and vertebrates (Garrido-Ramos *et al*., 1994; Mestrovic *et al*., 1998; Mravinac *et al*., 2005; Levy *et al*., 2007; Kopecna *et al*., 2014; Miga, 2015; Subirana *et al*., 2015; Subirana and Messeguer, 2013). The term satellite DNA takes its origin from early caesium chloride gradient centrifugation experiments, in which genomic DNA fragments composed of tandemly repeated short sequences were isolated due their different density with respect to bulk DNA, thus forming “satellite” bands in the gradient (Kit, 1961).

Satellite DNA comprises a wide range of sequences that differ in nucleotide sequence, sequence complexity, repeat unit length and abundance, with only 2 features shared: ability to build long arrays of tandem head-to- tail arranged repeats and ability to form highly condensed heterochromatic regions (Plohl *et al*., 2008). Each satDNA family consists in a library of monomer variants that can be shared by related species. Expansions and/or contraction of different variants from this library may result in rapid changes in satDNA distribution and abundance profiles, even among closely related species (Kuhn *et al*., 2008; Plohl *et al*., 2012; Rojo *et al*., 2015). Because of the dynamic changes of satDNA in evolutionary time scale, these sequences can be species or genus specific (Garrido-Ramos, 2015). Moreover, different satDNA families may be also present in single species. For example, there are up to 15 different satellite DNA families in *Pisum sativum* (Macas *et al*., 2007), 62 families in *Locusta migratoria* (Ruiz-Ruano *et al*., 2016) and 9 in human (Levy *et al*., 2007; Miga, 2015). However,

usually only one or a few families are predominant in each species (Macas *et al*., 2007). The most abundant satDNA family within the human genome is the alpha-satellite (Lander *et al*., 2001).

The evolutionary divergence of satDNAs is demonstrated not only by the presence of variable satDNA families in one species and among different species, but also by the striking difference in the complexity of the organizational patterns and contribution of particular sequence types to repeat-based centromeres among species (Plohl *et al*., 2014). For example, while all rice centromeres share the same DNA sequence (Macas *et al*., 2007; Macas *et al*., 2010), the alpha-satellite of human centromeres, is chromosome-specific, with various degrees of polymorphism, making the organization of the human centromeric DNA different not only among individuals in the population, but also among chromosomes in the same individual. Polymorphism may concern the total number of HORs, the number of monomers making up the HORs and the sequence of monomers due to the presence of various SNPs (Aldrup-Macdonald and Sullivan, 2014).

Satellite DNA is also present in the pericentromeric regions where its abundance can be higher than that of centromeric core satellite DNA (Garrido-Ramos, 2017). Cytologically, pericentromeric satellite DNA is organized in chromocenters, which are dense aggregates of pericentromeric satellite DNA from multiple heterologous chromosomes clearly visible in interphase nuclei (Jones, 1970; Pardue and Gall, 1970; Mayer *et al*., 2005). Chromocenters show epigenetic modifications associated with chromatin compaction and transcriptional repression such as DNA methylation and methylation of histone H3 and H4 (Saksouk *et al*., 2015; Nishibuchi and Déjardin, 2017).

As mentioned before, human centromeres host families of repetitive DNA other than alpha satellite, satellites 1, 2, 3, β and γ, that are localized in the pericentric regions. These may be organized in tandem arrays of variable length (from 5 bp up to 220 bp) or they may be composed of interspersed repeats (LINEs and SINEs). Satellite 1, 2, 3 are A-T rich and their amount is highly variable in the pericentromeric regions of most human chromosomes. They are evolutionarily conserved, suggesting a functional role (Grady *et al*.,

1992). Another well characterized pericentromeric satellite-DNA family is the mayor satellite present at all mouse chromosomes, with the exception of chromosome Y (Pardue and Gall, 1970). Major satellite monomers are 234 bp long and are organized in uninterrupted arrays that vary in extension from a minimum of 240 kb to more than 2000 kb (Vissel and Choo, 1989).

1.2.2 FUNCTIONS OF CENTROMERIC SATELLITE DNA

The functional significance of satDNA has been long debated. Because of its association with the heterochromatic fraction of the genome, the supposed absence of transcriptional activity as well as the high sequence divergence across species, in the early seventies, this DNA was defined junk (Ohno, 1972; Palazzo and Gregory, 2014). However, this large genomic DNA fractions is not lost or diminished during evolution, thus indicating that it undergoes selective pressure and therefore should have a functional role. Several evidences show the involvement of centromeric and pericentromeric satellite DNA in stabilizing interactions with DNA binding proteins, in maintaining heterochromatin architecture, in sustaining kinetochore formation, and in controlling chromatid association and disjunction during cell divisions (Plohl *et al.,* 2008; Pezer *et al*., 2012; Bierhoff *et al*., 2014; Bersani *et al*., 2015; Biscotti *et al.,* 2015). Moreover, the ability of satellite DNA to change rapidly during evolution might give rise to reproductive barriers and promote speciation (Plohl *et al.,* 2014; Biscotti *et al.,* 2015).

Independently on sequence conservation, centromeric satDNAs may stabilize CENP-A nucleosomes and promote functional interactions by forming secondary and tertiary structures. Different combinations of nucleotides can build similar higher order structures; therefore, the sequence can be altered as long as the structure itself is not impaired (Lamb and Birchler, 2003; Plohl *et al*., 2008). The functional role of satellite DNA in stabilizing CENP-A nucleosomes could be a driving force in the birth of new

satellite DNA families that can colonize centromeres and could explain why evolutionarily new centromeres (see chapter 1.3.3), void of repetitive DNA sequences, can evolve to mature centromeres containing reiterated DNA families (Zhang *et al*., 2013; Fukagawa and Earnshaw, 2014). Tandem duplication of a new sequence with a selective advantage for CENP-A stabilization or transposition of existing satDNAs could populate an epigenetically defined neocentromere (Gong *et al*., 2012; Zhang *et al*., 2013; Nergadze *et al*., 2018).

While a function for satellite DNA in kinetochore assembly and/or stability has been inferred almost since its discovery (Pardue *et al.,* 1970; Yunis and Yasmineh, 1971), a prevailing wrong idea was that these sequences are not actively transcribed into RNA. The analysis of the linear organization of histones within centromeres using immunofluorescence on chromatin fibres revealed that CENP-A nucleosome domains are intermingled with nucleosomes containing H3K4me2 and H3K36me2 which are epigenetic marks of transcriptionally active chromatin (Blower *et al*., 2002; Sullivan and Karpen, 2004; Bergmann *et al*., 2011). Indeed, transcripts derived from centromeric and pericentromeric repetitive sequences have been identified in several organisms such as yeast (Choi *et al*., 2011; Ohkuni and Kitagawa, 2011; Choi *et al*., 2012), rice (Neumann *et al*., 2007) and maize (Topp *et al*., 2004), insects (Pezer and Ugarkovic, 2008; Rosic *et al*., 2014), beetles (Pezer and Ugarković., 2008), mouse (Ferri *et al*., 2009) and man (Saffery *et al*., 2003; Wong *et al*., 2007; Quenet and Dalal, 2014). It has been demonstrated that kinetochore assembly and centromere function are strictly linked to the transcription of the repetitive DNA sequences which they contain (Ohkuni and Kitagawa, 2011; Bergmann *et al*., 2012a; Bergmann *et al*., 2012b; Chan and Wong, 2012; Quénet and Dalal, 2014). Indeed, defects in centromeric and/or pericentromeric sequences transcription lead to chromosome mis-segregation during cell division as it has been documented in yeast (Ohkuni and Kitagawa, 2011), in human HeLa cells (Chan *et al*., 2012), and in Human Artificial Chromosomes (HACs) (Bergmann *et al*., 2012b). For example, small-interfering RNAs (siRNAs), processed by the ribonuclease Dicer are transcribed from pericentromeric

tandem repeats in *S. pombe* where they are involved in heterochromatin formation and maintenance (Volpe *et al.,* 2002; Motamedi *et al*., 2004; Verdel *et al*., 2004). In hybrid chicken cells containing a single human chromosome, the loss of Dicer protein leads to defects in centromere heterochromatin and chromosome segregation, pointing out the importance of siRNA for heterochromatin assembly (Fukagawa *et al*., 2004). Not only small RNAs, but also long non-coding RNAs have been found at the centromere. At human centromeres, CENP-A recruitment and its subsequent loading require the transcription of a 1.3 kb lncRNA that directly binds CENP-A and its chaperone HJURP (Quénet and Dalal, 2014). Centromeric transcripts have been found also in *D. melanogaster*; in particular, a specific block of satellite DNA (knows as 359-bp satellite) (Lohe *et al*., 1993; Sun *et al*., 2003; Blattes *et al*., 2006) which is present only at the centromere of the X chromosome, produces a transcript which is involved in the correct localization of the centromere constitutive proteins, CENP-A and CENP-C, as well as of outer kinetochore proteins (Rošic *et al.,* 2014).

Long ncRNAs, not processed into siRNA and highly variable in size, are characteristic of the centromeres of rice, wallaby, mouse and beetles (Bouzinba-Segard *et al*., 2006; Carone *et al*., 2009; Pezer and Ugarkovic, 2009; Pezer *et al*., 2012). In rice, long transcripts from centromeric satDNA are processed in RNAs of 40 nt and they act together with 21–24 nt long siRNAs derived from the pericentromeric portion of the same satellite (Lee *et al*., 2006).

In man, centromeric RNAs have been shown to be involved in the response to stress conditions and in tumour progression. For the first time, Denegri and co-workers demonstrated that in interphase nuclei of heat-shocked HeLa cells, the pericentromeric heterochromatin of chromosomes 9, 12 and 15 colocalize with Stress-induced Nuclear Bodies (SNBs) (Stress-induced Nuclear Bodies) which are known to be sites of transcripts accumulation (Denegri *et al*., 2002). Concerning tumour progression, it has been demonstrated that BRCA1-deficient cancer cells are defective in pericentromeric heterochromatin formation and show activation of pericentromeric alpha-satellite DNA transcription (Zhu *et al*., 2011).

Transcription of non-coding RNAs has also been hypothesized to be involved in the emergence of new centromeres. An L1 transposable element was found to be transcribed in a human neocentromere defining the boundaries of the CENP-A domain (Chueh *et al*., 2005; Chueh *et al*., 2009).

The transition of TEs from a silenced to a transcriptionally active state may promote the recruitment of CENP-A nucleosomes, leading to the rescue of acentric chromosome fragments following the inactivation of the native centromere (O'Neill *et al*., 1998). How an ectopic site becomes activated, enabling the recruitment of CENP-A nucleosomes in the absence of chromosome damage, as happens for evolutionary centromere repositioning events, is unknown.

1.3 NEOCENTROMERES

Neocentromeres are fully functional centromeres that arise on non-repetitive DNA sequences (Marshall *et al.,* 2008). Due to the absence of satellite DNA, which is indeed the most represented type of DNA at most vertebrate centromeres, neocentromeres represent the strongest evidence supporting the epigenetic nature of the centromere (Barra and Fachinetti, 2018). Most of the initial information on neocentromeres derived from the analysis of human clinical data derived from cytogenetic screenings (Marshall *et al.,* 2008). Neocentromeres associated with pathological conditions are usually a consequence of structural chromosome rearrangements that remove or disrupt the resident centromere, so that the event of “neocentromerization” acts as a mechanism to rescue acentric fragments that otherwise should be lost (Choo, 2000; Blom *et al*., 2010).

Partial/complete deletion of a centromere has been used as a mechanism to induce new centromere formation in different model organisms such as *D. melanogaster* (Williams *et al*., 1998; Maggert and Karpen, 2001; Olszak *et al*., 2011), *Schizosaccharomyces pombe* (Ishii *et al*.,

2008), *Candida albicans (Ketel et al., 2009)* and several plant species (Nasuda *et al*., 2005; Gong *et al*., 2009; Topp *et al*., 2009).

In addition, in a number of instances, ectopic centromeres formation occurs in evolutionary time scale, leading to the formation of the so called Evolutionarily New Centromeres (ENCs) (Montefalcone *et al*., 1999; Ventura *et al*., 2001; Cardone *et al*., 2006; Piras *et al*., 2009; Wade *et al*., 2009; Rocchi *et al*., 2012; Nergadze *et al*., 2018). These centromeres are centromeres that moved to a new position without any observable chromosomal rearrangement and are one of the driving forces for karyotype evolution in a number of species (Rocchi *et al.,* 2012).

1.3.1 - HUMAN NEOCENTROMERES

Most of the knowledge about the formation and function of neocentromeres derives from the cytogenetic analysis of neocentromeres found in clinical cases (Burrack and Berman, 2012). The first case was identified in 1993 and the marker chromosome, designated mardel (10), originated from a *de novo* rearrangement of chromosome 10 which gave rise to a ring chromosome containing the endogenous centromere and a linear chromosome fragment (Voullaire *et al*., 1993). This acentric maker chromosome was rescued by the spontaneous formation of a new centromere at the cytogenetic band 10q25, an euchromatic region of the chromosome arm that did not undergo any sequence change (du Sart *et al*., 1997; Lo *et al*., 2001).

More than 100 pathological neocentromeres have been described to date showing differences both in the chromosomal position and in the features of the DNA sequence on which they form (Marshall *et al*., 2008; Alonso *et al*., 2010; Klein *et al*., 2012).

Human neocentromeres derive from structural chromosome rearrangements, such as interstitial deletions or inverted duplications, both generating a chromosome fragment lacking a conventional centromere.

Both types of rearrangements, reported in Figure 4, originate from chromosome breaks that, in the first case, generate a centric ring chromosome and an acentric fragment (on which the neocentromere forms) (Fig. 4-A) or an acentric ring (on which the neocentromere forms) and a centric fragment (Fig. 4-B), depending on the site of deletion; in the second case, chromosome breaks lead to the formation of an acentric chromosome fragment (on which the neocentromere forms) with an inverted duplication (Fig. 4-C; fig. 4-D). In the cases in which a neocentromeric marker chromosome is present in an unbalanced karyotype (inv-dup) it is associated with mosaicism presumably due to the counter selection of trisomic or tetrasomic cells (Amor and Choo, 2002). Meanwhile, chromosome fragments caused by interstitial deletions lead to a balanced karyotype at the cytogenetic level; however phenotypic effect may result from a "ring syndrome," due to a mosaicism for the marker or to the interruption of critical genes at the sites of chromosome breakage (Amor and Choo, 2002; Marshall *et al*., 2008). Therefore, human neocentromeres are generally associated to developmental delay and heterogeneous clinical signs (Amor *et al*., 2005; Mascarenhas *et al*., 2008; Burnside *et al*., 2011; Hemmat *et al*., 2012; Ma *et al*., 2015).

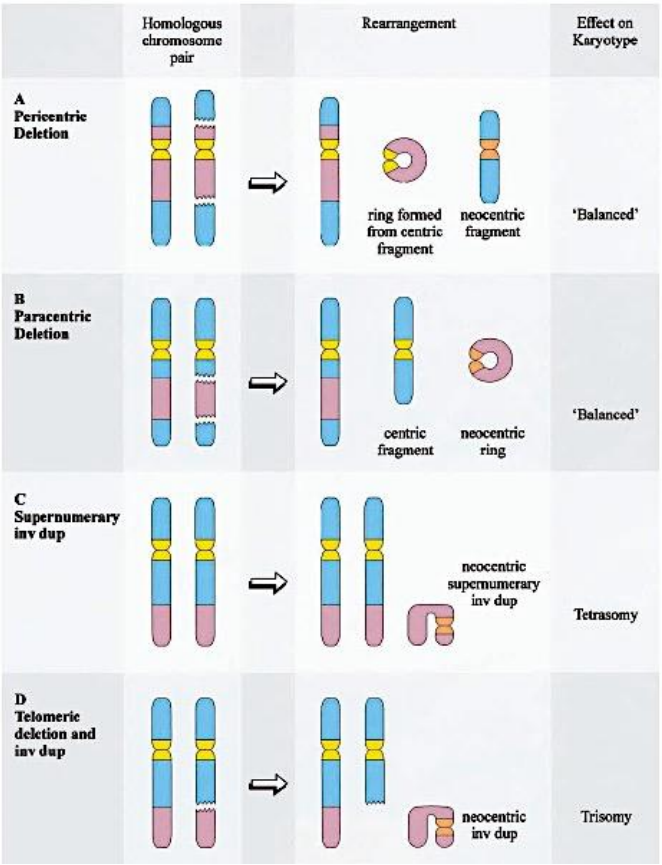


Figure 4 – Chromosome rearrangements commonly associated with clinical neocentromere formation. The most common chromosome rearrangements are interstitial deletions (A and B) and inverted duplications (C and D) (Amor and Choo, 2002).

Neocentromere formation is an event that can accidentally occur in cancer, as a mechanism that cancer cells use to rescue acentric chromosomes mostly containing amplified genes that confer a selective advantage to cancer cells. Examples of tumours in which neocentric chromosomes have been described are liposarcoma, retinoblastoma, non-Hodgkin's lymphoma, acute myeloid leukaemia and lung cancer (Morrissette

et al., 2001; Italiano *et al*., 2006; Mascarenhas *et al*., 2008; de Figueiredo *et al*., 2009; Blom *et al*., 2010).

Some cases of well-differentiated liposarcomas are cytogenetically characterized by the presence of a supernumerary ring or giant chromosome containing the amplification of the 12q14-15 region (Italiano *et al*., 2009), and in some instances, a chromosome 8-derived ring chromosome has been described in acute myeloid leukaemia (de Figueiredo *et al.,* 2009).

Most of the initial information on neocentromeres has stemmed from human clinical data gathered through cytogenetic screening. Figure 5 summarizes 60 reported cases of constitutional human neocentromeres in which chromosomal origin has been investigated (Amor and Choo, 2002). So far, human neocentromeres have been identified on 21 out of the 22 autosomes, as well as on chromosomes X and Y and more than 100 have been characterized (Marshal *et al.,* 2008; Alonso *et al.,* 2010; Klein *et al.,* 2012). Neocentromeres are extremely heterogeneous in chromosome position and DNA associated sequence. Indeed, some neocentromeres arise in heterochromatic regions containing repeated DNA sequences (Hasson *et al.,* 2011), but most of them are localized in single copy DNA sequences (Alonso *et al.,* 2010). However, the distribution of reported neocentromeric sites across the human genome is not completely random, some chromosome regions being hot spots such as 3q, 13q, 15q and Yq (Figure 5). Moreover, distal chromosome regions appear to be more susceptible to neocentromere formation, compared with proximal regions (Marshall *et al*., 2008; Liehr *et al*., 2010).

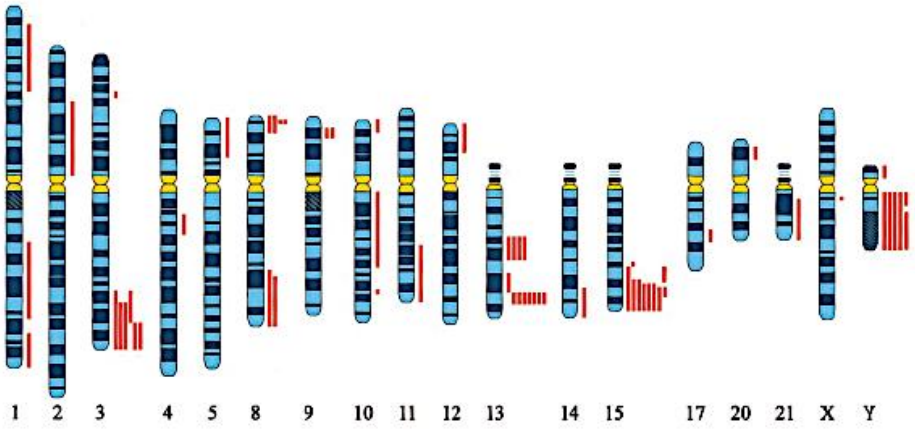


Figure 5 – Sites of formation of constitutional neocentromeres within the human genome. The mapped position of each neocentromere is indicated by red bars on the right of the chromosome ideograms. Longer bars indicate neocentromere sites that have not been precisely localized (Amor and Choo, 2002).

Despite the divergence of the sequence underlying human neocentromeres, some common features can be identified, such as a high AT content and an increase in LINEs content (Chueh *et al*., 2009).

In rare cases, human neocentromeres form on otherwise normal chromosomes, without any rearrangement, presumably following inactivation of the native centromere through unclear mechanisms (Amor *et al*., 2004; Liehr *et al*., 2010). Unlike pathological human neocentromeres, this type of neocentromeres do not show mosaicism and segregate permanently for at least three generations (Rocchi *et al.,* 2012).

1.3.2 ARTIFICIAL NEOCENTROMERES

Although the structure and functional characteristics of neocentromeres are known, comparatively little is known about the mechanism which drives their formation. The first example of neocentromerization induced experimentally was reported in *Drosophila melanogaster*. In an initial study, a neocentromere arised after irradiation of a mini chromosome derived from the X chromosome (Williams *et al*., 1998; Maggert and Karpen, 2001). Significantly, the rearrangements within the X derived mini chromosome relocated the neocentromere very close to the endogenous centromere. In contrast, when the same fragment was released from a normal X chromosome, where it is 40 Mb apart from the native centromere, no neocentromere formation was observed (Williams *et al.,* 1998). A further experiment in Drosophila showed that CENP-A over-production results in ectopic kinetochore assembly at the boundaries between heterochromatin and euchromatin (Olszak *et al*., 2011).

Interestingly, an association may exist between DNA damage and neocentromere formation. In human HeLa cells, in which overexpression of CENP-A was transiently induced, ectopic CENP-A localization was observed at Double Strand Break (DSB) sites (Zeitlin *et al*., 2009). This data, together with the identification of CENP-A chaperone HJURP at sites of DNA damage, suggest that sites of DSB bound by CENP-A might initiate neocentromere formation as a mechanism to prevent the loss of otherwise acentric chromosomes (Kato *et al*., 2007; Zeitlin *et al*., 2009).

In barley (Hordeum vulgare), a neocentromere was formed through specie-specific crossbreeding and marker selection. A spontaneous breakage produced a chromosome 7 derived isochromosome (mirror images of chromosome 7 short arm), containing a neocentromere with no centromeric repeated DNA sequence. The newly formed centromere was localised in close proximity to the chromosome 7 native centromere and was transmitted normally (Nasuda *et al*., 2005).

Neocentromeres have also been induced in fungi by the deletion of the endogenous centromere using a homologous recombination strategy in

Candida albicans (Ketel *et al*., 2009) or by the Cre-Lox system in fission yeast (Ishii *et al*., 2008). In *C. albicans*, a human pathogenic fungus with small regional centromeres, 3-4 kb, (Baum *et al*., 2006; Ketel *et al*., 2009), neocentromeres arised close to the *locus* of the constitutive centromere following deletion of the constitutive centromere which was replaced by a selectable marker gene (Ketel *et al.,* 2009). In the fission yeast *S. pombe,* neocentromere formation was induced by deletion of the native centromere DNA via Cre-Lox excision. Similarly, to the ectopic centromeres induced by over-production of CENP -A in *D. melanogaster*, *S. pombe* neocentromeres were found at the borders between euchromatin and heterochromatin (Ishii *et al.,* 2008). In addition, neocentromere formation in *S. pombe* required heterochromatin marker proteins. Thus, in *S. pombe* as in *D. melanogaster*, the chromatin context strongly influences neocentromere formation.

Despite a lot of information came out from the studies carried out in the model organisms described above, the process of neocentromere formation in vertebrate cells remains poorly understood. Shang and collaborators (Shang *et al*., 2013) established a chromosome-engineering system in chicken DT40 cells that allowed them to efficiently isolate neocentromere-containing chromosomes. The authors used the Z chromosome of DT40 chicken cells; this chromosome is present in single copy and does not contain satellite DNA at the native centromere. In order to generate DT40 chicken cells containing neocentromeres, a 127 kb region, which included a 35 kb CENP-A binding domain, was conditionally removed and cells that maintained the chromosome despite the loss of the centromere, were selected. The authors demonstrated that, though neocentromeres can be formed in any region of the Z chromosome, 76% of those regions are adjacent to the site where the original centromere was previously located (Shang *et al.,* 2013). The preference for non-random neocentromere formation near the endogenous centromere was thought to be due to the presence of CENP-A in the flanking regions (Shang *et al.,* 2013).

These secondary CENP-A sites do not allow the nucleation of a kinetochore in the presence of the constitutive centromere; nevertheless, these flanking regions can become able to recruit a functional kinetochore

when the constitutive centromere is removed or inactivated (Ketel *et al.,* 2009).

The transformation/transduction of putative centromere DNA sequences into cells has been performed in several organisms to define the minimal region needed for *de novo* centromere formation (Kalitsis and Choo, 2012). Human alpha satellite DNA, integrated in mammalian cell lines, was able to recruit centromeric proteins but it did not ensure stable segregation (Haaf *et al*., 1992; Larin *et al*., 1994). Positive results were obtained by Harrington and collaborators in 1997 (Harrington *et al*., 1997), that combined long arrays of alpha satellite DNA with telomeric and genomic DNA, obtaining mini chromosomes stably transmitted in HT1080 cells. However, centromere formation occurred only in HT1080 cells, presumably providing trans acting elements needed for centromere seeding. Moreover, only alpha satellite arrays containing CENP-B boxes were able to form *de novo* centromeres (Ikeno *et al*., 1998; Grimes *et al*., 2002). Furthermore, arrays containing mutated CENP-B boxes did not form *de novo* human artificial chromosomes (Ohzeki *et al*., 2002). Thus, the presence of CENP-B binding sites is required for human centromeres assembly *in vitro*. This is in contrast with the fact the centromere of Y chromosome, while being perfectly functional, lacks CENP-B boxes. These findings point to the existence of un-understood key differences between *de novo* artificial centromeres versus natural ones.

1.3.3 EVOLUTIONARILY NEW CENTROMERES

Evolutionarily New Centromeres (ENCs) derive from the repositioning of a centromere during evolution without any structural chromosome rearrangement, in other words, in evolutionary time, centromere function slides in a new position while the ancestral centromere becomes inactive (Montefalcone *et al*., 1999; Burrack and Berman, 2012). The first case of evolutionary centromere repositioning was described by Montefalcone and colleagues (Montefalcone *et al.,* 1999) while studying the evolution of chromosome 9 in primates. FISH experiments, using BAC probes specific for chromosome 9 in different primate species, revealed that the position of the centromere changed while the order of molecular markers was conserved (Figure 6).

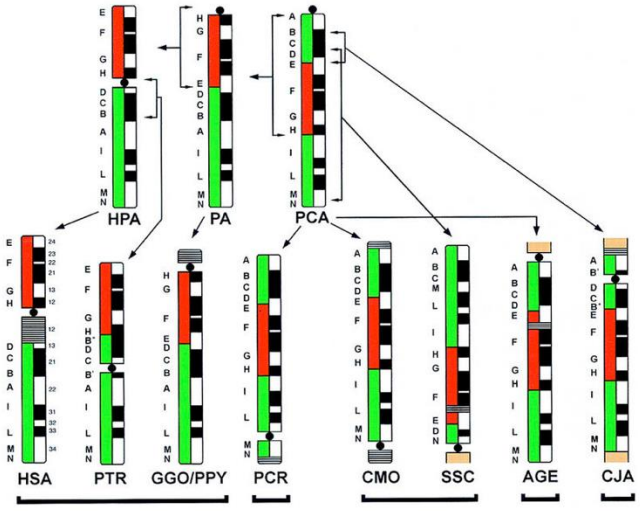


Figure 6 – Diagram representing clear evidences of a process of centromere repositioning in primates. Regions homologous to human 9p (red) and 9q (green) are reported on the left of each ideogram, while the G-banding pattern is indicated on the right. From left to right**:** Pongidae ancestor (PA)**;** HPA (HSA/PTR common ancestor); *Primate Common Ancestor* (*PCA);* Homo sapiens (HSA); common chimpanzee (Pan troglodytes, PTR), gorilla (Gorilla gorilla, GGO), and orangutan (Pongo pygmaeus, PPY); Cercopithecidae (Old World monkey, OWM), silvered leaf-monkey (Presbytis cristata, PCR); Platyrrhinae (New World monkeys, NWM), dusky titi (Callicebus molloch, CMO, Callicebinae), spider monkey (Ateles geoffroy, AGE, Atelinae), common marmoset (Callithrix jacchus, CJA, Callitrichinae), and squirrel monkey (Saimiri sciureus, SSC, Saimirinae). (Montefalcone *et al.,* 1999).

A number of cases of ENCs have been found in primates (Ventura *et al*., 2001; Eder *et al*., 2003; Misceo *et al*., 2005; Cardone *et al*., 2006; Roberto *et al*., 2007; Stanyon *et al*., 2008; Rocchi *et al*., 2012) and in other mammals such as equids (Piras *et al*., 2009; Wade *et al*., 2009; Nergadze *et al*., 2018), *Macropus eugenii* (Ferreri *et al*., 2005), Ryukyu spiny rat (Kobayashi *et al*., 2008).

In recent years, the widespread practice of prenatal diagnosis allowed to identify asymptomatic human neocentromeres; similarly, to ENCs they are derived from centromere repositioning, suggesting that ENCs and human neocentromeres are two sides of the same coin. It's worth noticing that in some cases the *locus* where a new centromere is seeded during evolution or in human pathology is the same and this may imply that some chromosome domains are hot spots for centromerization (Cardone *et al*., 2006; Rocchi *et al*., 2009). Furthermore, it has been discovered that some sites where pathological neocentromeres occur correspond to inactivated ancestral centromeres (Rocchi *et al.,* 2009; Rocchi *et al.,* 2012).

A chromosome domain that seems to promote the formation of human neocentromeres is the region 15q24-26, characterized by a high concentration of segmental duplications, which represent relics of the pericentromeric segmental duplications flanking the ancestral inactive centromere (Ventura *et al*., 2003; Ventura *et al*., 2007). Human chromosomes 15 and 14 were generated by fission of an ancestral chromosome, which appears to be composed of these two chromosomes arranged head-tail (Fig. 7). Following the fission event, a neocentromere arised on both fragments. Indeed, one new centromere emerged in human chromosome 15, corresponding to the telomeric region of the short arm of macaque homologous chromosome 7 (MMU7) (Fig. 7). A second centromere emerged on chromosome 14 and corresponded to the fission point of MMU7.

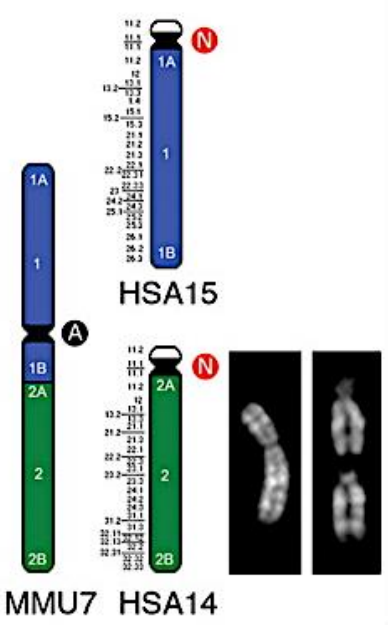


Figure 7 – Evolution of human chromosomes 15 and 14. Chromosomes 15 (HSA15) and 14 (HSA14) in *Hominoidea* were generated by fission of an ancestral MMU7 chromosome, which appears to be composed of these two chromosomes arranged head-tail (Adapted from Ventura *et al.,* 2007).

The relationship between human neocentromeres and ENCs described so far regards chromosomes in which hotspots for the formation of clinical neocentromeres (3q, 8p, 13q, and 15q) have been found (Voullaire *et al*., 1993; Blennow *et al*., 1994; Maraschio *et al*., 1996; Rowe *et al*., 2000; Teshima *et al*., 2000; Warburton *et al*., 2000). However, though no region of chromosome 6 is found to be a hotspot, Capozzi and co-workers (Capozzi *et al*., 2009) reported a family in which a centromere repositioning event occurred on this chromosome causing centromere sliding to the ancestral position (6p22.1) (Eder *et al*., 2003; Raaum *et al*., 2005).

Additional intriguing relationships between human neocentromeres and ENCs have been reported for chromosomes 13 and 3. Chromosome 13 is one

of the most conserved chromosomes among primates (hominoids, Old World monkeys, and New World monkeys), Carnivora (cat), Perissodactyla (horse), and Cetartiodactyla (pig). Two independent centromere repositioning events occurred in OWMs and pigs (*Sus scrofa*) with the centromere repositioned in the same location (13q21). Strikingly, this region is also the place where some pathological neocentromeres were reported (Morrissette *et al*., 2001; Li *et al*., 2002; Warburton *et al*., 2000). Sequence analysis of this chromosome domain revealed the presence of a 3,9 Mb region totally void of genes. It is supposed that the absence of genes may be a critical component for the seeding and maturation of a novel centromere. (Cardone *et al*., 2006).

A further case emphasizing the relationship between pathological and evolutionarily neocentromeres is the human chromosome 3**.** In this chromosome, a repositioned centromere associated with a normal phenotype and a clinical neocentromere were seeded to the 3q26 chromosomal domain. This region corresponds to the position of the centromere in OWM common ancestor (Ventura *et al*., 2004). Therefore, the same domain has been used as seeding point for an ENC and pathological/non-pathological human neocentromeres.

The frequency of centromere repositioning in evolution is surprisingly high, making these events a powerful driving force in karyotype evolution. Ventura and collaborators (Ventura *et al*., 2007), by the comparison of human and macaque karyotypes, identified 14 ENCs; nine of which occurred in the macaque lineage and five in the human lineage. As the last common ancestor of macaques and humans is estimated at about 25 million years ago, ENCs formed about once every three million years, meanwhile in the same range of time only four translocations occurred (about one translocation every twelve million years) (Rocchi *et al*., 2009).

The process of centromere repositioning may be influenced by forces acting to keep the region of the new centromere unaltered, the presence of genes can be one of these forces (Rocchi *et al*., 2012). Some authors examined the expression of genes embedded in neocentromeric regions pointing out that presence of neocentromeres does not affect gene expression per se (Lam *et al*., 2006; Nagaki *et al*., 2004a; Saffery *et al*., 2003;

Yan *et al*., 2006). However, the accumulation of satellite DNA and the potential remodelling of the pericentromeric regions can negatively affect gene structure and, consequently, expression. The absence of genes around the neocentromerization regions can therefore be a condition favouring, or at least not opposing, ENCs fixation. Lomiento and co-workers (Lomiento *et al*., 2008) found that ENCs fixed in primates were preferentially seeded in gene-deserts.

Whatever characteristics favour the neocentromere formation, it is evident that the site of the repositioned neocentromere tends to become more complex in evolutionary time, acquiring most of the characteristics of natural centromeres. Meanwhile, the chromosomal domain containing the inactive ancestral centromere undergoes rapid depletion of specific centromeric sequences. An example is represented by the human chromosomal region 2q21 where the endogenous centromere underwent inactivation following the fusion of the telomeres of two ancestral chromosomes, raising present time human chromosome 2 (IJdo *et al*., 1991). In this chromosome, the ancestral centromere lost almost totally alphoid DNA sequences (Montefalcone *et al*., 1999; Ventura *et al*., 2003).

An excellent model for the study of the dynamics and mechanisms underlying the evolution of the centromeric region is represented by the species belonging to the genus *Equus*. As will be detailed later on in this introduction, the presence of different steps of centromere maturation in equid species allowed to propose a model to explain the birth, evolution and complete maturation of mammalian centromeres (Piras *et al*., 2010; Nergadze *et al*., 2018).

1.4 CENTROMERIC AND PERICENTROMERIC EPIGENETIC ENVIRONMENT

The differences observed among the centromeric sequences in eukaryotes together with the existence of neocentromeres and pseudo-dicentric chromosomes (see paragraph 1.1) demonstrate the lack of a universal sequence motif underlying centromere function, the latter being determined by epigenetic mechanisms. As mentioned above, the main epigenetic mark of the eukaryotic centromere function is the protein CENP-A, a centromere-specific variant of histone H3 (McKinley and Cheeseman, 2016). CENP-A is required for kinetochore assembly and its deletion is associated with mis-localization of most kinetochore proteins. Furthermore, a CENP-A over-expression leads to the formation of the kinetochore in ectopic sites due to its recruitment in normally non-centromeric regions (Allshire and Karpen, 2008). However, single CENP-A nucleosomes can be found at non-centromeric sites throughout the chromosomes in human cells indicating that the presence of CENP-A alone is not sufficient for centromere activity (Bodor *et al*., 2014). Indeed, centromere identity is determined also by the chromatin landscape characterising the centromere core and the surrounding regions (Blower *et al*., 2002; Peters *et al*., 2003; Sullivan and Karpen, 2004; Bergmann *et al*., 2011).

In general, different chromatin domains are characterized by the presence of different epigenetic markers which modify the flexible N-terminal tails of the four core histones (H2A, H2B, H3 and H4). These post-translational modifications include acetylation, methylation, phosphorylation and ubiquitination (Jenuwein and Allis, 2001; Fischle *et al*., 2003). Histone modifications are indicators of transcriptionally active or repressed chromatin. The 'histone code' hypothesis suggests that combinations of specific histone modifications create a complex, functional hierarchy for chromatin regulation (Jenuwein and Allis, 2001).

The centromere core and the pericentromeric regions are distinguished by distinct chromatin signatures that are crucial for their function. A critical balance between repressive and activating post-translational

histone modifications give rise to the so called “centrochromatin” (Blower *et al*., 2002; Peters *et al*., 2003; Sullivan *et al*., 2004; Bergmann *et al*., 2011).

Moreover, the presence at the centromere core and at the pericentromere of histone modifications typical of transcriptionally permissive chromatin suggests that transcription of these domains may play a role in centromere propagation and function (Folco *et al*., 2008; Choi *et al*., 2011; Chan and Wong, 2012; Rosic *et al*., 2014; Quenet and Dalal, 2014).

1.4.1 CENTROMERIC PROTEINS

Centromeric proteins were firstly isolated in early ‘80s from the serum of patients affected by the CREST syndrome (Calcinosis, Raynaud’s phenomenon, Esophageal dysmotility, Sclerodactily, Telangiectasia) (Moroi *et al*., 1980; Earnshaw and Rothfield, 1985). These patients produce anti centromere antibodies that were used to isolate the centromere proteins CENP-A, CENP-B and CENP-C.

Centromeric proteins form a complex structure, called kinetochore, which appears as a trilaminar disc characterized by an electron dense layer in the middle of two compact layers (Jokelainen, 1967; Comings and Okada, 1971; McEwen *et al*., 2007).

The inner kinetochore layer is assembled on the centromeric chromatin formed by nucleosomes containing CENP-A interspersed with nucleosomes containing the classical histone H3. It consists of 16 proteins complex, called CCAN (Constitutive Centromere-Associated Network) (Cheeseman and Desai, 2008; Hori *et al*., 2008a; Hori *et al*., 2008b) (Fig.8 - blue coloured box). This protein network is divided in five functional domains: CENP-C, the CENP-L-N complex (Carroll *et al*., 2010; Hinshaw and Harrison, 2013), the CENP-O-P-Q-U-R complex (Hori *et al*., 2008b; Hornung *et al*., 2014), the CENP-H-I-K complex (Cheeseman and Desai, 2008; McKinley *et al*., 2015) and the CENP-T-W-S-X complex (Nishino *et al*., 2012). Depletion of components within the CENP-H-I-K-M complex causes cell-cycle arrest,

kinetochore-assembly defects and severe chromosome mis-segregation (Fukagawa *et al*., 2001; Nishihashi *et al*., 2002; Goshima *et al*., 2003; Foltz *et al*., 2006; Izuta *et al*., 2006; Okada *et al*., 2006). The CENP-O-P-Q-U-R domain (also known as CENP-50/PBIP) (Foltz *et al*., 2006; Izuta *et al*., 2006; Okada *et al*., 2006) is non-essential for viability, but is required to ensure recovery after spindle damage (Minoshima *et al*., 2005; Okada *et al*., 2006). The CCAN complex localizes at the centromere during the entire cell cycle, mediating the link between kinetochore and centrochromatin and making up the platform on which the outer layer of the kinetochore is assembled (Nagpal *et al*., 2015).

The outer layer of the kinetochore is formed by the KMN complex (KNL1–MIS12–NDC80) (Fig. 7 – magenta coloured box) which is the microtubule binding-site (Screpanti *et al*., 2011; Malvezzi *et al*., 2013). A number of fibrillar projections depart from the outer plate forming the so-called fibrous corona (Yen *et al*., 1991; Cooke *et al*., 1993). The latter is constituted by kinesins, the molecular motors of the microtubules (CENP-E / Cenp-meta and Kif18A / Klp67A), dyneins and the associated proteins LIS1 and the ROD–ZW10–Zwilch (RZZ complex). The cooperation among all these components drives chromosome alignment and segregation (Yen *et al*., 1991; Cooke *et al*., 1997; Yao *et al*., 1997; Gandhi *et al*., 2004; McEwen *et al*., 2007; Stumpff *et al*., 2008; Barisic and Geley, 2011).

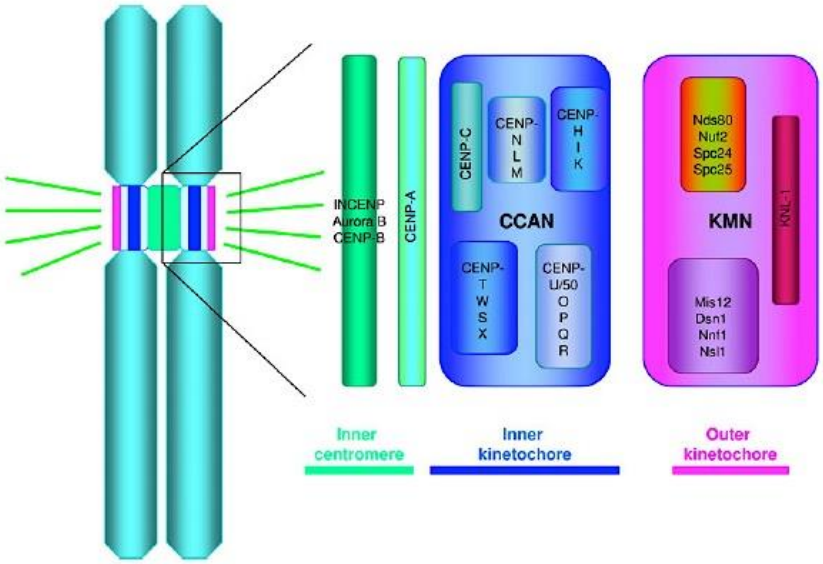


Figure 8 – Schematic model of trilaminar structure of the kinetochore. The area between the sister-chromatids is the inner centromere zone (green), which contains the chromosome passenger protein complex (CPC) formed by borealin, survivin, INCENP, Aurora B, condensin, and CENP-B. The inner kinetochore (blue) consists of five groups of CENPs known as CCAN (CENP-C, CENPT/W/S/X, CENP-N/L/M, CENP-H/I/K, and CENP-O/P/Q/R/U), which assemble directly on centromeric chromatin where CENP-A nucleosomes are deposited (light green). The outer kinetochore (magenta) lies on the inner kinetochore and has three representatives: KNL1, the Mis12 complex (Mis12, Dsn1, Nsl1, and Nnf1) and the Ndc80 complex (Ndc80, Nuf2, Spc24, and Spc25) (Perpelescu and Fukagawa, 2011).

CENP-A, the centromeric histone H3 variant, is the main actor in the formation of a functional kinetochore, as its presence is essential for kinetochore recruitment. CENP-A (Fig. 9) contains a histone fold domain which shares 62% homology with human histone H3 and an amino-terminal tail that differs more significantly from histone H3 and among CENP-A proteins from different species (Goutte-Gattat *et al*., 2013). Within the

histone fold domain, the first loop (L1) and the second alpha-helix (α2) form the CENP-A Targeting Domain (CATD), which is sufficient for centromere localization and confers a unique rigidity to the CENP-A nucleosome that is its defining characteristic (Black *et al*., 2004; Black *et al*., 2007b). When in cells knocked down for CENP-A, the histone fold domain of histone H3 is substituted with the CATD domain, it gains the ability to be recruited at the centromere and to ensure cell viability (Black *et al*., 2007a; Black *et al*., 2007b).

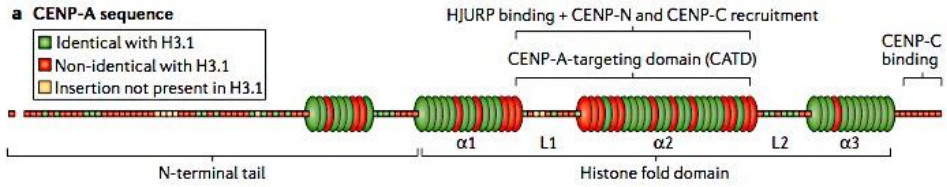


Figure 9 – Model of human CENP-A primary and secondary structure, showing conservation with respect to histone H3. Each segment corresponds to a single amino acid and is coloured according to its conservation with respect to human H3.1. Barrels represent α-helices, and rods represent loops. Within the histone-fold domain, the helices are designated α1 to α3 and the loops are designated L1 and L2. L1 and α2 comprise the CENP-A-targeting domain, which is sufficient to target CENP-A to centromeres (McKinley and Cheeseman, 2016).

CENP-A binds CENP-N and CENP-C through the CATD domain. CENP-C presents six sites of interaction with CENP-A carboxyterminal tail with other histones within the CENPA nucleosome and with the CENPA CATD (Carroll *et al*., 2010; Guse *et al*., 2011; Kato *et al*., 2013; Logsdon *et al*., 2015; Westhorpe *et al*., 2015).

The interaction between CENP-C and CENP-A is essential for CENP-A deposition process. Indeed, the correct deposition of CENP-A, which, unlike other histones, occurs in the G1 phase, is ensured by the interaction of CENP-C with the protein M18BP1 which is a subunit of the MIS18 complex.

This complex localizes at the centromeres transiently during a short period of the cell cycle from telophase and persisting through early G1 phase (Maddox *et al*., 2007) and allows the recruitment of a protein called "HJURP", also involved in the CENP-A loading (Barnhart *et al*., 2011). The interaction between the complex MIS18/HJURP with CENP-C allows CENP-A targeting at pre-existing CENP-A containing nucleosomes (McKinley and Cheeseman, 2016).

While the correct CENP-A loading is essential for a proper kinetochore assembly, a correct kinetochore attachment to the mitotic spindle is essential in order to avoid chromosomal segregation errors. Therefore, defective attachments of kinetochores to the mitotic spindle must be monitored. This monitoring activity is performed by a protein complex, called Chromosomal Passenger Complex (CPC), which is formed by the enzymatic component Aurora B and the three regulatory and targeting components INCENP, Survivin and Borealin (Ruchaud *et al*., 2007). This complex has a characteristic behaviour in mitosis: from prophase, up to metaphase is located on the kinetochore, from anaphase onwards it is redistributed on the spindle microtubules and cell cortex (Carmena *et al*., 2012). These dynamic changes in CPC localisation throughout mitosis ensure the effective and spatially restricted phosphorylation of substrates involved in chromosome condensation, correction of erroneous kinetochore-microtubule attachments, activation of the spindle assembly checkpoint (SAC), and cytokinesis (Sandall *et al*., 2006; Ruchaud *et al*., 2007; Carmena *et al*., 2012).

1.4.2 HISTONE MODIFICATIONS AT THE CENTROMERE

As already mentioned in the introduction of this chapter, centromere function is strictly dependent on the chromatin environment in which the functional centromere core is embedded (Blower *et al*., 2002; Peters *et al*., 2003; Sullivan and Karpen, 2004; Bergmann *et al*., 2011).

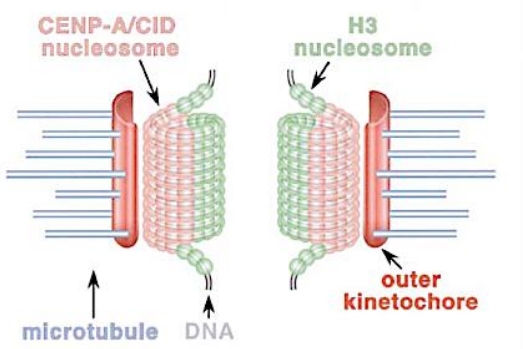
The organization of centromere chromatin is peculiar due to the presence of both euchromatic and heterochromatic marks whose relative amount is critical for the creation of the permissive conditions needed for CENP-A loading and for the maintenance of centromere identity (Bergmann *et al*., 2012b; Quenet and Dalal, 2014). Early studies associated centromeres to heterochromatin (Flemming, 1882; Vig, 1982), moreover, the heterochromatic state of pericentromeric chromatin is demonstrated by the presence of high levels of di- and tri-methylation of histone H3 at lysine 9 (H3K9me2, H3K9me3) (Peters *et al*., 2003; Rice *et al*., 2003; Alonso *et al*., 2010; Shang *et al*., 2013), although, non-repetitive centromeres and neocentromeres frequently lack surrounding heterochromatin (Peters *et al*., 2003; Rice *et al*., 2003).

In particular, H3K9me3 is a typical marker of constitutive heterochromatin. In *D. melanogaster* and in humans, this histone modification is concentrated in the pericentromeric region and does not overlap the centromere core (Sullivan and Karpen, 2004). Also, in *S. pombe* the central centromere core is flanked on both sides by heterochromatin marked by H3K9me3 (Partridge *et al*., 2000).

H3K9me2 is a marker of facultative heterochromatin and, as HeK9me3, is present in the pericentromeric regions. Immuno-fluorescence analysis of chromatin fibres showed that this post-translational histone modification can overlap the borders of the CENP-A containing domain, but never spreads over them (Sullivan and Karpen, 2004). Conversely, chromatin immunoprecipitation analyses indicated that, in rice, this histone modification is present also within the functional centromere core (Nagaki *et al*., 2004a).

Several roles have been attributed to the heterochromatin flanking

the centromere core. It may be required to organize the higher order structure of centromeric chromatin. In this regard, a model has been proposed by Blower and co-workers (Blower *et al*., 2002). This model predicts that pericentromeric heterochromatin may interact with the interspersed H3 domains to produce or maintain a cylinder structure, in which CENP-A containing nucleosome are pushed toward the external face of the cylinder, thus facilitating interactions with the kinetochore; meanwhile H3 nucleosomes are sequestered in the region between the



paired sister kinetochores (Fig. 10) (Blower *et al*., 2002).

Figure 10 – 3D Organization Model of centromere chromatin of *Drosophila* and Humans metaphases chromosomes. Blocks of CENP-A and H3 nucleosomes are arranged so that CENP-A nucleosomes are pushed towards the outer face of the metaphase chromosome while H3 nucleosomes are sequestered in the region between the paired sister kinetochores (Blower *et al*., 2002).

Heterochromatin at the centromere may also be necessary to adopt a conformation that maintains cohesion between sister chromatids (Bernard *et al*., 2001; Nonaka *et al*., 2002). Furthermore, it may define the borders of the centromeric functional domain and prevent CENP-A chromatin spreading in adjacent regions (Maggert and Karpen, 2000; Maggert and Karpen 2001).

Though heterochromatin is present at the constitutive centromeres,

most satellite-free neocentromeres are not associated with it (Shang *et al*., 2013; Alonso *et al*., 2010). Human neocentromeres (Alonso *et al*., 2010), neocentromeres induced in DT40 chicken cells (Shang *et al*., 2013) and in *C. elegans* are assembled in the absence of heterochromatin (Yuen *et al*., 2011; Padeken and Heun, 2013). However, 3D analysis performed by Nishimura and co-workers (Nishimura *et al*., 2019) in DT40 chicken cells, demonstrated that, in interphase nuclei, heterochromatin free neocentromeres interact with far heterochromatic regions by means of large loops. These observations suggest that also centromeres not showing the typical heterochromatic histone modifications, however, need to interact with a heterochromatic environment.

While histone modification typical of constitutive and facultative heterochromatin generally characterize pericentromeric regions, the centromere functional core is characterized by the presence of post-translational histone modifications such as H3K4me2 (histone 3 lysine 4 di-methylation) and H3K36me2 (histone 3 lysine 36 di-methylation) that correlate with transcriptionally competent chromatin (Sullivan and Karpen, 2004; Bergmann *et al*., 2011). Using Human Artificial Chromosomes (HACs) it was demonstrated that centromere core recruits RNA polymerase II (Bergmann *et al*., 2012b; Chan and Wong, 2012) and produces transcripts (Saffery *et al*., 2003; Yan *et al*., 2005; Wong *et al*., 2007). The knock down of the transcriptionally competent pattern or the inhibition of transcription is detrimental for centromere function. Indeed, experiments performed on human cells containing a HAC showed that H3K4me2 depletion affects the levels of centromeric transcripts and centromere-associated RNAP II; the decrease in centromeric transcription causes a progressive loss of CENP-A (Molina *et al*., 2016).

A further histone modification localizing at the centromere core is the mono-methylation of lysine 20 of histone H4 (H4K20me1) which is present in CENP-A containing nucleosomes. This histone modification was specifically detected at both repetitive DNA containing and repetitive DNA free centromeres in chicken DT40 cells, where it is needed for kinetochore assembly (Hori *et al*., 2014). Cell cycle analysis demonstrated that this

modification does not occur before CENP-A recruitment at the centromeres indicating that H4K20me1 may be involved in the "maturation" of CENP-A nucleosome (Hori *et al*., 2014). H4K20me1 modification must occur during G1 only after CENP-A deposition, this would allow Constitutive-Centromere-Associated-Network (CCAN) proteins to recognize H4K20me1 in centrochromatin and thereby kinetochore assembly. The observation that, in chicken, ectopic low levels of CENP-A located at non-centromeric loci are not able *per se* to form a new centromere (Shang *et al*., 2013), can be due to their lack of the H4K20me1 modification (Hori *et al*., 2014). Importantly, this modification is frequently associated with transcription, which has been shown to be important for kinetochore maintenance and assembly (Nakano *et al*., 2008; Bergmann *et al*., 2011; Bergmann *et al*., 2012). Thus, H4K20me1 might be linked to the regulation of centromeric transcription. Alternatively, or in addition, it may have a role in establishing the structure of centromeric chromatin, thereby facilitating kinetochore assembly, for example, by providing a preferred binding site for factors that may promote the CENP-T-W-S-X complex association with chromatin (Nishino *et al*., 2012).

1.4.3 DNA METHYLATION AT THE CENTROMERE

DNA methylation, that is the addition of methyl groups to the DNA molecule, contributes, together with post-translational histone modifications to define the epigenetic environment needed for centromere function (Lehnertz *et al*., 2003; Saksouk *et al*., 2015; Wong *et al*., 2006). Five-methylcytosine, in particular, is widespread throughout the mammalian genome and found on up to 80% of CpG dinucleotides (Vinson and Chatterjee, 2012).

DNA methylation generally acts as transcription repressor (Hsieh, 1997; Lorincz *et al*., 2004), the best described example being the hyper-methylation of DNA in the X chromosome of mammalian female somatic cells (Kaslow and Migeon, 1987). The regulatory function of DNA methylation has also been studied in depth at imprinting control regions (Li *et al*., 1993) and, more recently, at pericentromeric regions (Saksouk *et al*., 2015).

Though the presence of repetitive DNA at the centromere hampers detailed molecular analysis, studies performed in patients affected by the ICF (Immunodeficiency, Centromeric region instability, Facial anomalies) syndrome offered the first insights in the role of DNA methylation in centromere stability (Xu *et al*., 1999). ICF Patients show loss of function of DNMT3b [DNA (cytosine-5-)-methyltransferase]. This condition leads to a substantial decrease in DNA methylation at both centromeric (Miniou *et al*., 1997) and pericentromeric DNA sequences (Hansen *et al*., 1999) which is associated with the presence of large regions of decondensed pericentromeric heterochromatin. DNA methylation has been demonstrated to be associated with the H3K9me3 histone modification, which is marker of constitutive heterochromatin (see paragraph above). H3K9me3 was originally shown to be a prerequisite for DNA methylation in *Neurospora crassa* (Tamaru and Selker, 2001) and in plants (Jackson *et al*., 2002). H3K9me3 and DNA methylation systems also interact at mammalian pericentromeric heterochromatin. Suv39h methyl-transferase knockout mouse cells reveal an altered DNA methylation profile, specifically at

pericentromeric satellite DNA repeated sequences (Lehnertz *et al*., 2003).

Biochemical analyses demonstrated that the Methyl-CpG-Binding Domain protein 1 (Mdb1) interacts with the Suv39h1/2 histone methyltransferases, which modifies histone H3 at lysine 9 (Lunyak *et al*., 2002; Fuks *et al*., 2003). As detailed above, the methylated form of histone H3 is generally associated with transcriptional repression. Accordingly, the Mdb1 and H3K9me are concentrated at pericentromeric regions (Fuks *et al*., 2003). The recognition of methylated sequences at the centromere mediated by Mdb1 drives H3K9me to those regions. The interaction between Mbd1 and H3K9 methyltransferase may play an essential role in the correct maintenance of H3K9 methylation in heterochromatin during DNA replication (Zhang and Reinberg, 2001; Lehnertz *et al*., 2003; Sarraf and Stancheva, 2004).

DNA hypo-methylation in heterochromatic regions has been reported in several cancers: breast, ovary and liver (Bernardino *et al*., 1997; Narayan *et al*., 1998; Qu *et al*., 1999; Saito *et al*., 2001). It has been suggested that DNA hypo-methylation can be associated to genomic and chromosomal instability (Almeida *et al*., 1993; Lengauer *et al*., 1997; Chen *et al*., 1998).

Analysis of the DNA methylation state at satellite DNA free neocentromeres, demonstrated that, within the neocentromeric chromatin, there are specific sites of active transcription and the centromere chromatin boundaries are defined by DNA hypo-methylation. These results suggest also that transcription of centromere DNA may occur through the selective hypo-methylation of pockets of sequences without compromising the overall silent chromatin state and function of the centromere (Wong *et al*., 2006).

Further studies have been conducted in maize, using the supernumerary B chromosome as a model (Koo *et al*., 2011). The functional centromere of this chromosome spans ~700 kb and contains five distinct arrays of the ZmB repeat; each of these arrays is flanked by the CentC satellite intermingled with the CRM retrotransposon (Jin *et al*., 2005; Lamb *et al*., 2005). Despite the presence of repetitive DNA sequences, it was demonstrated that the centromere of the supernumerary B chromosome, shows a peculiar state of DNA methylation, where the ZmB repeats arrays

are hyper-methylated and their flanking repeats (CentC-CRM) are hypo-methylated (Koo *et al*., 2011). However, inactivation of the B centromere resulted in a drastic increase of DNA methylation of its underlying sequences. Probably, the loss of CENP-A may lead the underlying CentC-CRM sequences to be "exposed" to DNA methyltransferases thus being heavily methylated (Koo *et al*, 2011).

1.4.4 CENTROMERE TRANSCRIPTION

Historically, centromeres were considered simply heterochromatin-rich and thus transcriptionally silent. The discovery of post-translational histone modifications associated with transcriptionally active chromatin such as mono-methylation of lysine 4 and di- and tri-methylation of lysine 36 of histone H3 (H3K4me1, H3K4me2, H3K36me2, and H3K36me3 (Sullivan and Karpen, 2004; Gopalakrishnan *et al*., 2009; Bergmann *et al*., 2011; Bergmann *et al*., 2012b) changed the way of looking at the centromere.

Interestingly, transcripts, ranging in sizes from 35 to 5,000 bp and deriving both from the pericentromeric region and from the centromere core, have been identified in a number of organisms including rice, maize, beetle, tammar wallaby, mouse, and man (Denegri *et al*., 2002; Hall *et al*., 2003; Ugarkovic, 2005; Carone *et al*., 2009; Eymery *et al*., 2009; Brown *et al*., 2012; Gent and Dawe, 2012; Koo *et al*., 2016; Rosic and Erhardt, 2016).

In the model organism *S. pombe*, long non-coding RNAs (lncRNAs), transcribed from pericentromeric repetitive sequences, are processed into siRNAs which recruit factors involved in heterochromatin assembly and maintenance (Volpe *et al*., 2002; Motamedi *et al*., 2004; Verdel *et al*., 2004). Indeed, the impairment of the RNA interference (RNAi) pathway in this organism resulted in severe chromosome segregation defects (Hall *et al*. 2003). Later, chromatin immunoprecipitation was used by Choi and co-workers (Choi *et al*., 2011) to demonstrate that RNA polymerase II is present at *S. pombe* centromeres.

The involvement or pericentromeric transcription in heterochromatin assembly was also demonstrated in other organisms. Pathways resembling RNA interference were found in rice, *Arabidopsis*, *Drosophila* and mouse (Lippman and Martienssen, 2004; Neumann *et al*., 2007; Hsieh *et al*., 2011).

In man, the first evidence of pericentromeric heterochromatin transcription came from a work by Denegri and co-workers (Denegri *et al*., 2002) in which the authors demonstrated that in interphase nuclei of heat-shocked HeLa cells, the pericentromeric heterochromatin of chromosomes 9, 12 and 15 colocalizes with Stress-induced Nuclear Bodies (SNBs) which are known to be sites of transcripts accumulation. More recently, a growing number of studies demonstrated the presence of transcripts deriving from the centromere core, indicating that these transcripts play a role in centromere/kinetochore assembly and maintenance (Wong *et al*., 2007; Li *et al*., 2008; Chueh *et al*., 2009; Ferri *et al*., 2009; Gent and Dawe 2012; Quénet and Dalal, 2014).

Centromere transcripts were found in a number of evolutionarily distant species such as man (Wong *et al*., 2007; Chueh *et al*., 2009), *Plasmodium falciparum* (Li *et al*., 2008), mouse (Ferri *et al*., 2009) and tammar wallaby (Carone *et al*., 2009), suggesting that non coding RNA is an ancient component of centrochromatin.

It was demonstrated that centromere transcripts or small CEN RNA (CENtromeric RNA) derivatives form ribonucleoprotein complexes with several proteins which localize constitutively or transiently at the centromere, such as CENP-A, HJURP (CENP-A chaperone), CENP-C and CPC (Chromosomal Passenger Complex). The interaction between CENP-A and CEN RNAs was first observed at a human neocentromere (Chueh *et al*., 2009). LINE-1 elements within the CENP-A-binding region of a neocentromere on 10q25 were transcribed in non-coding RNAs that were integrated within the CENP-A chromatin. Later, Quénet and Dalal (Quénet and Dalal, 2014) demonstrated that both CENP-A and HJURP interact with CEN RNAs and that this association is important for CENP-A loading at centromere. In the same study, it was shown that the inhibition of Pol II

during early G1 halved CENP-A amount at the centromere and nearly zeroed HJURP although the total cellular levels of CENP-A and HJURP were unaffected (Quenet and Dalal, 2014). Moreover, in Hela cells, centromeric α-satellite RNAs are present in the nucleoli and, at mitosis, are associated to the foundational kinetochore protein CENP-C and to INCENP, a component of the chromosomal passenger complex (CPC). Indeed, the treatment with single-stranded RNA-specific RNases results in a significant delocalization of CENP-C and in the complete delocalization of INCENP and survivin, indicating that that the assembly of INCENP and Survivin and of a mitotic pool of CENP-C at the metaphase centromere is dependent on the presence of single-stranded RNA (Wong *et al*., 2007).

The binding of CENP-A and CENP-C to CEN DNA and alpha-satellite RNA promotes kinetochore assembly, including the recruitment of the CPC (INCENP, Survivin, Borealin, and Aurora B), which regulates chromosome-spindle attachment and activates the Spindle Assembly Check-point (SAC), thus preventing chromosome misalignment (Hindriksen *et al*., 2017). Co-immmuno-precipitation experiments demonstrated the direct interaction of Aurora B with CEN RNA (Ferri *et al*., 2009; Ideue *et al*., 2014). The RNA-dependent inner kinetochore localization of CPC is mediated by at least two RNA-binding domains: one that is present in Aurora B and one in Survivin or Borealin (Blower, 2016). Moreover, it was observed that the knockdown of α-satellite RNA in HeLa cells (Ideue *et al*., 2014) or inhibition of transcription with triptolide in Xenopus egg extracts (Blower, 2016) reduced the centromeric localization of Aurora-B and caused improper spindle attachment and chromosome unalignment.

It is well known that in eukaryotes during cell division DNA transcription stops almost completely (Liang *et al*., 2015). On the contrary, centromeres remain transcriptionally active and elongating RNA Polymerase II (Pol II) can be found at the centromere up to metaphase (Chan *et al*., 2012).

Overall, the literature data reported in this paragraph, lead to conclude that a transcriptionally permissive environment characterizes centromeres in eukaryotes and a fine regulation of centromeric

transcription is crucial for the centromere function. Indeed, on one hand the depletion of active chromatin markers (H3K4me2) at the centromere inhibits RNA Pol II transcription causing the loss of CENP-A loading due to the failure in HJURP recruitment (Bergmann *et al*., 2011). On the other hand, a completely open chromatin configuration, permissive for high level transcription is incompatible with centromere function as well (Nakano *et al*., 2008).

1.5 THE GENUS *EQUUS*

Perissodactyla (from ancient greek, perissós means odd and dáktylos means finger), is an order of eutherian mammals characterized by limbs with odd fingers. This order is currently made up of two sub-orders: *Ceratomorpha* and *Hippomorpha*; to the *Ceratomorpha* belong two families: *Tapiridae*, which currently includes four species, and *Rhinocerotidae,* which currently includes five species. The *Equidae* family belongs to the *Hippomorpha* and is represented by a single extant *genus ,* the *genus Equus*, which in turn includes: two horse species (domestic horse, *Equus caballus* and Preswalsky horse, *Equus Przewalski*), five donkey species (onager, *Equus hemionusonager*, African wild ass, *Equus africanus*, Asian wild ass, *Equus hemionus*, domestic ass, *Equus asinus*, and Tibetan emion, *Equus kiang*) and four zebra species (Grevy zebra, *Equus grevyi*, zebra of the plains, *Equus quagga*, mountain zebra, *Equus zebra hartmannae*, and Burchelli zebra, *Equus burchelli*) (Trifonov *et al*., 2008).

Paleontological and molecular studies indicate that the divergence of Perissodactyla occurred in Laurasia about 56 million years ago (Murphy *et al*., 2007). The first radiation events within the *Tapiridae* and *Rhinocerotidae* families dated about 20-30 million years ago, while the radiation within *Equidae* dates back to about 3 million years ago. In particular, the divergence of the horse from the ancestor goes back approximately to 2.4 million years ago, while the zebra and donkey speciation took place about 0.9 million

years ago (Oakenfull and Clegg, 1998).

Despite the recent divergence, the morphological similarity and the possibility of crossing between different species, *Equidae* have very different karyotypes in terms of chromosome number and structure, chromosome number ranging from 32 in *E. zebra* to 66 in *E. przewalskii* (Trifonov *et al*., 2008; Jonsson *et al*., 2014). Studies based on comparative mapping, comparative chromosome painting and genome sequencing allowed the reconstruction of the *Perissodactyla* ancestral karyotype (PAK), that comprised 74-78 chromosomes, mostly acrocentric (Trifonov *et al*., 2008; Trifonov *et al*., 2012). The uncertainly in the chromosome number can be explained by the ancestral polymorphic state of some *Perissodactyla* chromosomes and/or by fusion and fission events (Trifonov *et al*., 2008).

Due to its domestication and its social importance, the domestic horse, is the most studied species among *Perissodactyla*. The horse genome was entirely sequenced and assembled in 2009 (Wade *et al*., 2009). Thanks to the technique of "chromosome painting", it was possible to reconstruct the hypothetical Equid ancestral karyotype which is composed of 70-76 chromosomes and differs from the *Perissodactyla* ancestral karyotype (PAK) only for two fusions (PAK4/8 and PAK9/36).

In turn, the domestic horse karyotype (2n = 64) differs from the equid ancestral karyotype for two centric fissions (PAK1 and PAK18) and five centric fusions (PAK21/1, 18/1, 32/15, 31/17, 18/22 resulting in ECA2, 3, 6, 8, 10, respectively); the karyotype of Przewalski horse acquired an additional fission (PAK 3). A non-Robertsonian fusion (PAK2/34 = ECA4p/31) and a centric fission (PAK3 = ECA5p/q) are common for non-caballine equids (zebras and asses) (Trifonov *et al*., 2012).

In light of these data and comparing the frequency of rearrangements for each phylogenetic branch of the *Perissodactyla* tree (Fig. 11), it can be deduced that, in the Equid karyotype, an extremely high number of rearrangements occurred in a very short space of time, the rate of rearrangements per million years ranging from 2.92 to 22.2 (Lindsay *et al*., 1980; Trifonov *et al*., 2008), that is to say about 80-fold that of ancient ceratomorphs, making equid genome evolution one of the most rapid in non-rodent mammals.

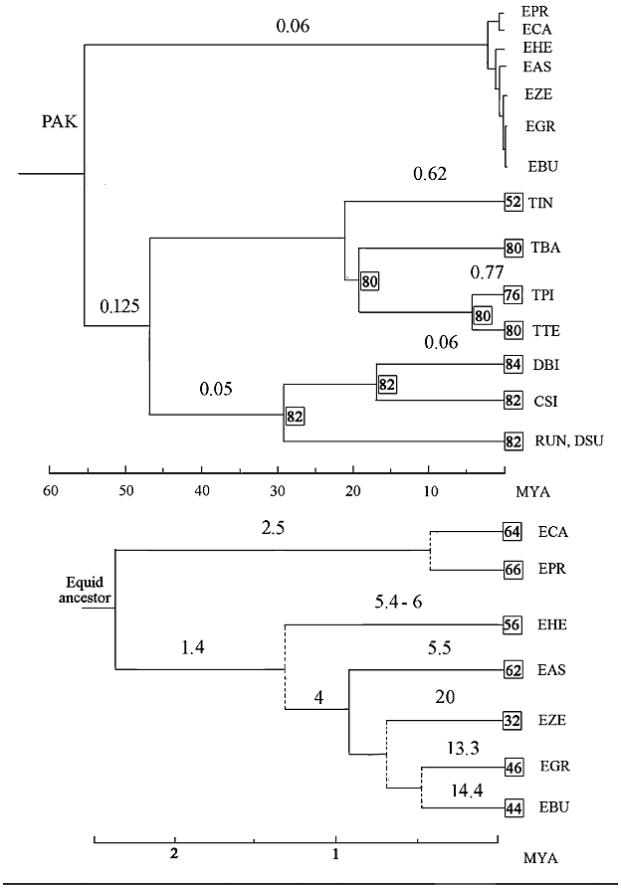


Figure 11 – Rates of karyotype evolution in perissodactyla (A) and in equids (B). Numbers in squares indicate diploid chromosome numbers. Numbers above the branches indicate the average rearrangements rate (rearrangement per million years) (Trifonov *et al*., 2008).

1.5.1 THE CENTROMERES IN THE GENUS *EQUUS*

The species of the genus *Equus* represent an ideal model for the study of centromere organization and behaviour. Indeed, the evolution of the karyotypes of these species has been characterised by a surprisingly high number of centromere repositioning events that lead to the formation of a number of "immature" evolutionarily new centromeres which are void of the canonical satellite DNA sequences. In particular, in the horse 1 chromosome out of 32 is satellite-free, 16 out of 31 in the donkey, 11 out of 22 in the Burchell's zebra , 9 out of 23 in the Grevy's zebra, 10 out of 28 in the Hartmann's zebra and 14 out of 25 in the African donkey (Carbone *et al*., 2006; Wade *et al*., 2009; Piras *et al*., 2010; Raimondi *et al*., 2011; Geigl *et al*., 2016; Nergadze *et al*., 2018). The centromere of horse chromosome 11 (ECA11), was the first discovered natural satellite-free mammalian centromere and was completely sequenced in 2009 (Wade *et al*., 2009). It was hypothesized that this centromere has not acquired yet the classical satellite DNA sequences, and for this reason, it can be considered evolutionarily "immature" (Piras *et al*., 2010). The analysis of the chromosomal distribution of the two major horse satellite DNA families (37cen and 2PI) in four equid species, *E. caballus* (ECA), *E. asinus* (EAS), *E. grevyi* (EGR) and *E. burchelli* (EBU) (Piras *et al*., 2010) showed that some chromosomes lack these sequences at the centromere, but they are present in terminal position (Fig. 7), where they could represent the relics of the ancestral centromere. Notably, in the horse only one chromosome is satellite-free DNA, while in the donkey and in the zebras several chromosomes lack satellite DNA at the cytogenetic resolution level. The absence of satellite DNA FISH signals from some centromeres raised the question about the presence of undiscovered repetitive DNA families at such centromeres. To address this question, FISH analyses on metaphases chromosomes from these four equid species were performed using their total genomic DNA as probe. The results showed a few additional FISH signals in non-centromeric regions (Fig. 12). Thus, strongly suggesting that a

number of centromeres of the domestic donkey, Grevy's zebra and Burchelli's zebra are void of satellite DNA (Piras *et al*., 2010; Nergadze *et al*., 2014)

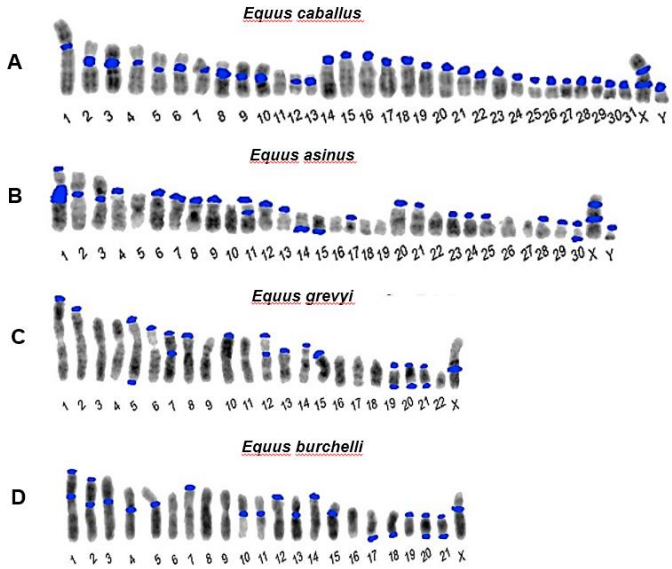


Figure 12 – Schematic representation of the distribution of satellite DNA on metaphase chromosomes from *Equus caballus* (A), *Equus asinus* (B), *Equus grevyi* (C) and *Equus burchelli* (D) (adapted from Piras *et al*., 2010).

A novel class of satellite DNA in horses and other equids, named EC137 (137 bp unit-repeat), was then isolated. FISH experiments on metaphases chromosomes showed that this class of satellite DNA is more represented in *E. caballus*, with respect to other equid species (*E. asinus*, *E. grevyi* and *E. burchelli*). Moreover, high-resolution FISH experiments on combed chromatin fibres of the horse highlighted a peculiar organization of satellite DNA sequences at the centromeric level. Small arrays (2-8 kb) of highly interspersed 2PI and EC137 sequences are immersed within large blocks (hundreds pf kb) of the 37cen sequence (Nergadze *et al*., 2014). This

organization leaded to the hypothesis that the 37cen satellite DNA sequence is present at the functional centromere domain, whereas 2PI and EC137 play an accessory role being presumably involved in the organization of the pericentromeric chromatin domains (Nergadze *et al*., 2014). Equid centromeres offer snapshots of the process of evolutionary maturation of mammalian centromeres. A model resembling these steps (Fig. 13) was proposed by Piras and co-workers (Piras *et al*., 2010). According to this scheme, and as proposed also by other authors (Montefalcone *et al.*, 1999; Ventura *et al.*, 2001; Amor and Choo, 2002) the first step would consist in the sliding of the centromere function to a new position lacking satellite DNA, while the satellite DNA from the old centromere remains in the terminal position (Fig. 13-B). A subsequent step would be the loss of satellite sequences located at the terminal position of the chromosome (Fig. 13-C). Finally, the new centromere could reach the complete maturation by the acquisition of satellite DNA (Fig. 13-D). Based on this model, the equid satellite-free centromeres correspond to steps B and C in Figure 13.

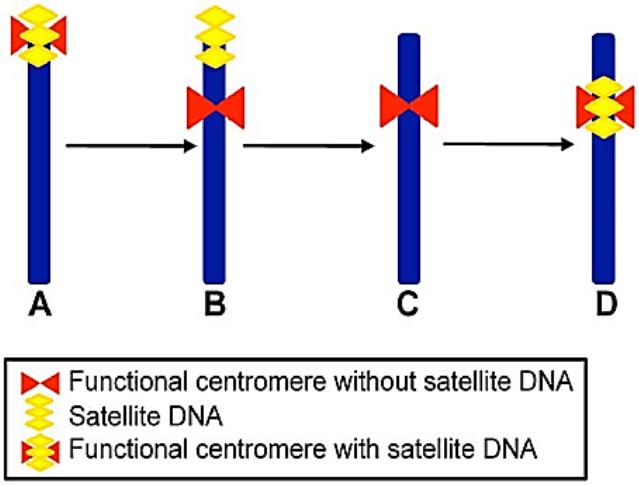


Figure 13 – Schematic representation of a four-step mechanism for neocentromere formation during evolution. (A) Acrocentric ancestral chromosome carrying satellite DNA (yellow) at its terminal centromere (red). (B) Submetacentric chromosome derived from centromere repositioning; this chromosome maintained satellite DNA sequences (yellow) at the terminal position, coinciding with the old centromere site, while the neocentromere (red) is devoid of repetitive sequences. (C) Submetacentric chromosome derived from (B) in which the terminal satellite sequences have been lost. (D) Submetacentric chromosome in its full "maturation" stage carrying satellite DNA (yellow) at the neocentromere (Piras *et al*., 2010).

In a recent paper Nergadze and co-workers (Nergadze *et al*., 2018) demonstrated that among the 16 donkey satellite-free neocentromeres five contain amplified DNA sequences. This discovery prompted the authors to add two alternative routes to the model shown in Fig. 14. The model presented in Fig. 14 takes into account the hypothesis that sequence amplification may be an intermediate step towards centromere complete maturation (routes A, B, E, D or A, B, C, F, D of **Fig. 14).**

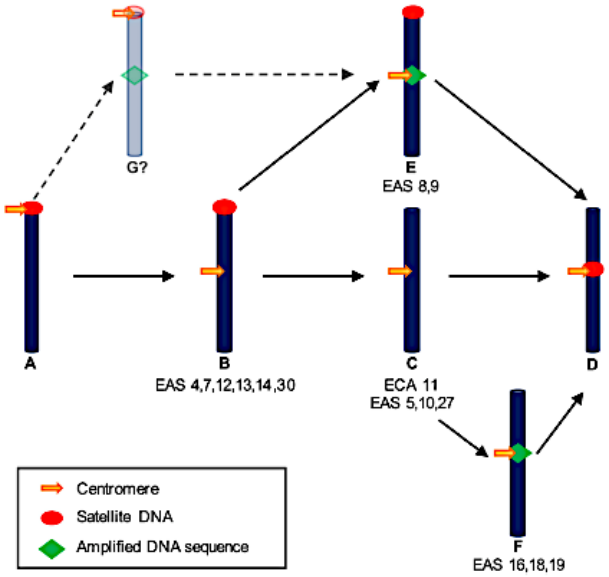


Figure 14 – Model for the maturation of a centromere during evolution. Several ways lead to fully mature satellite based repositioned centromere (D), from an ancestral centromere with satellite repetitions (A) through the intermediates without satellites (B, C, E, F) (Nergadze *et al*., 2018).

The presence of satellite-free centromeres in the Equid species supports the idea that satellite DNA is not essential for centromere function. However, literature data demonstrated that satellite DNA is transcribed and that these transcripts are involved in heterochromatin formation and CENP-A loading at the centromere (Volpe *et al*., 2002; Lippman and Martienssen, 2004; Neumann *et al*., 2007; Bergmann *et al*., 2011; Hsieh *et al*., 2011; Quenet and Dalal, 2014). Therefore, in order to understand if also in the genus *Equus* satellite DNA was transcribed, Cerutti and co-workers (Cerutti *et al*., 2016) performed RNA-seq experiments in the horse and demonstrated that the 37cen sequence, one of the two major satellite-families in the genus *Equus*, is transcribed. Interestingly, literature data

showed that centromeric transcripts may act in *trans* on all or on a subset of chromosomes, independently of the primary DNA sequence (Rošić *et al*., 2014; Biscotti *et al*., 2015; Rošić *et al*., 2016). Therefore, it could be hypothesized that, also in the horse the 37cen transcripts could act both in *cis* and in *trans*, thus stabilizing the satellite-free centromere of ECA 11.

A further characteristic of Equid centromeres is their extreme plasticity. ChIP-seq experiments were performed in the horse to characterize, at the sequence level, the functional core of the satellite-free centromere of chromosome 11 (Purgato *et al*., 2015). Surprisingly, in single individual horses, two CENP-A binding peaks were present on chromosomes 11, each one of about 100kb in length. Single chromosome analysis was performed by immuno-FISH on DNA fibres thus demonstrating that each CENP-A binding peak represents a positional allele. In other words, each homologous chromosome 11 binds CENP-A in a slightly different position and this phenomenon generates functional allelism. The analysis of this allelism in 5 independent individual horses showed that the centromeric CENP-A domain on horse chromosome 11 can slide within a region of about 500 kb (Purgato *et al*., 2015). The possible configurations of the centromere of horse chromosome 11 are called "epialleles" or alleles that do not differ in sequence, but in the position of the epigenetic "CENP-A" marker and therefore, in the position of the centromeric function. For this reason, this condition is called “functional polymorphism”.

Functional allelism at satellite-free centromeres was recently demonstrated also in the donkey (Nergadze *et al*., 2018). Moreover, the analysis of the transmission of CENPA binding domains in horse/donkey hybrids showed that positional movement can occur in one generation and it is on the order of 50 - 80 kb (Nergadze *et al*., 2018). These observations strongly support the hypothesis that the centromeric functional domain can be flexible, though limited by some sort of boundaries, such as specific patterns of chromatin marks (Sullivan and Karpen 2004; Martins *et al*., 2016).

2. AIMS OF THE WORK

The present research was carried out in the laboratory of Molecular Cytogenetics of the Department of Biology and Biotechnology of the University of Pavia, leaded by Professor Elena Raimondi and is the result of a collaboration with the laboratory of Molecular Biology of the same Department, leaded by Professor Elena Giulotto. My Ph.D. research project was aimed at studying centromere biology using the species belonging to the genus *Equus* as a biological model. These species radiated quite recently from the common ancestor, nevertheless, their karyotypes are very divergent due to the occurrence of a surprisingly high number of centromere repositioning events.

My research project was divided in two parts with two different, but complementary, aims:

1) ***Understanding the role played by satellite DNA in guaranteeing chromosome segregation fidelity***

Though the presence of satellite-DNA is a common feature of most eukaryotic centromeres, the existence of completely satellite-free natural centromeres opens the question of the functional role played by satellite DNA at the centromere. Due to their karyotype peculiarities, equid species offer the unique opportunity to directly compare the mitotic behaviour of chromosomes with satellite-free centromeres with that of chromosomes with satellite-based centromeres.

I set up chromosome two independent segregation assays (interphase aneuploidy analysis and micronucleus assay), coupled with FISH with chromosome specific DNA probes to compare the mitotic stability of a horse chromosome with a satellite-free centromere with respect to that of a chromosome with a satellite-based centromere. Moreover, to investigate the response of the two types of centromere to mitotic stress, I performed the same assays under normal conditions and after exposure to the spindle inhibitor nocodazole.

2) *Dissecting the equid centromere function from an epigenetic point of view*

The main epigenetic factor defining centromere function is the CENP-A protein, a centromere specific histone H3 variant (Blower *et al*., 2002). However, additional molecular features contribute to defining centromere activity (Jenuwein *et al*., 2001; Fischle *et al*., 2003), including the composition of pericentromeric chromatin and post-translational modifications of CENP-A itself (Hori *et al*., 2014; Bergmann *et al*., 2011, 2012; Nakano *et al*., 2008). Indeed, a centromere specific ratio between typical euchromatic and typical heterochromatic histone modifications is crucial for centromere identity.

In order to study the epigenetic environment of equid centromeres, I performed double immuno-fluorescence experiments with antibodies directed against the CENP-A marker and antibodies directed against different histone post-translational modifications (H3K4me2, H3K9me2, H3K9me3, H3K27me2, H4K20me1). The experiments were set up on horse and donkey metaphase chromosomes and chromatin fibres.

In addition, to directly compare the epigenetic state of satellite-free and satellite-based centromeres, immuno-FISH experiments on metaphase chromosomes and chromatin fibres were performed. Four chromosome specific DNA probes were used to identify the satellite–free centromeres of horse chromosome 11 and of donkey chromosome 7 and the satellite-based centromeres of horse chromosome 13 and of donkey chromosome 3. Finally, three colour immuno-FISH on metaphase chromosomes and chromatin fibres using a probe specific for the centromeric functional horse satellite DNA family 37cen with anti CENP-A and anti H3K9me3 antibodies allowed me to analyse the organization of constitutive heterochromatin at horse satellite-based centromeres.

3. MATERIALS AND METHODS

3.1 CELL LINES

Primary fibroblast cell line from the horse (HSF-D) and the domestic donkey (EASn) previously isolated and established in the laboratory of Molecular and Cellular Biology (Prof. Elena Giulotto), were used. Fibroblasts were cultured in Dulbecco's modified Eagle's medium (Euroclone), supplemented with 20 % foetal bovine serum (Euroclone), 2 mM glutamine, 2 % non-essential amino acids and 1 % penicillin/streptomycin. Cells were maintained at 37°C in a humidified atmosphere of 5% CO_2.

3.2 METAPHASE SPREAD PREPARATION

Mitotically active cells were collected flushing the medium on the cell monolayer and then were centrifuged at 1200 rpm (Z380 centrifuge, Hermle) for 10 minutes. The pellet was resuspended in 75 mM KCl hypotonic solution then incubated at 37°C for 15 minutes. Cold fixative (acetic acid:methanol - 1:3) was added and centrifuged at 1200 rpm (Z380 centrifuge, Hermle) for 30 minutes. The fixation was repeated twice. Then, slides were prepared by dropping the cell suspension perpendicularly to the slides and then air-dried. They were stored in an appropriate volume of cold fixative, at -20°C.

3.3 DRUG TREATEMENT FOR INTERPHASE NUCLEI ANALYSIS AND CYTOKINESIS-BLOCK

Fifty thousand cells were seeded in 5 cm diameter Petri dishes containing 24x24 mm coverslips. After 24 hours culture period, cells were exposed to nocodazole (Sigma) 100 nM. Since nocodazole is soluble in DiMethyl SulfOxide (DMSO), control cells were treated with the same DMSO concentration used for nocodazole treatment. For the CBMN assay, 24 hours after the addition of nocodazole (or DMSO in control samples) cytochalasin-

B (Cyt-B), 5 μg/ml, was added to the cultures. After 48 hours of nocodazole treatment, the slides were processed for interphase nuclei and micronuclei preparations as follows. Slides were treated with 75 mM KCl hypotonic solution and incubated at 37°C for 15 minutes. Cold (-20°C) fixative (methanol:acetic acid – 3:1) was added for 30 minutes. The fixation was repeated twice.

3.4 METAPHASE SPREAD PREPARATION FOR IMMUNOFLUORESCENCE

Mitotically active cells were collected as before. They were centrifuged at 1400 rpm (Z380 centrifuge, Hermle) for 8 minutes and suspended in 75 mM KCl, 0.8 % Na-citrate, H_2Obd (1:1:1) hypotonic solution for 15 minutes for the horse and 10 minutes for the donkey. The cell density in the hypotonic mixture was adjusted to $14*10^4$ cells/ml. After the hypotonic treatment, the cell suspension was cytocentrifuged onto glass slides at 750 rpm for 4 minutes (Z300 centrifuge, Hermle), then placed vertically and air dried.

3.5 CHROMATIN FIBRE PREPARATION

The cells were treated with trypsin, collected and then centrifuged at 1400 rpm for 8 minutes (Z380 centrifuge, Hermle). The supernatant was removed, and the pellet resuspended in PBS, then centrifuged again at 1400 rpm for 8 minutes (Z380 centrifuge, Hermle). The supernatant was removed, and the pellet was resuspended in an appropriate volume of 75 mM KCl, 0.8% Na-citrate, H_2Obd (1:1:1) hypotonic solution to obtain a final concentration of $7x10^4$ cells/ml for the horse and $6x10^4$ for the donkey. The treatment with the hypotonic solution was carried out at 37 ° C for 15 minutes for the horse and 20 minutes for the donkey. Slides were cytocentrifuged at 1500 rpm for 4 minutes and 1750 rpm for 4 minutes for horse and donkey respectively (Z300 cytocentrifuge, Hermle), then placed vertically and air dried. Slides were treated with a lysis buffer (2.5 mM TRIS

pH 8, 500 mM NaCl, 0.2 M urea, 1% Triton X-100) for 20 minutes. The slides were then removed from the solution with a constant speed of 300 µm/sec using an apparatus equipped with an electric pulley**.** The constant speed allows the unidirectional distension of the fibres.

3.6 DNA PROBES

Bacterial artificial chromosome (BAC) derived from horse CHORI-241 BAC library were used (Leeb *et al.,* 2006). Their cytogenetic position was validated by fluorescent in situ hybridization (FISH) on metaphase chromosomes.

- CHORI241-21D14 (chr11: 27,639,936 – 27,829,952)
- CHORI241-387D1(chr8: 41,985,542 – 42,131,402)
- CHORI241-22C1 (chr13: 7,346,775 - 7,544,907)
- CHORI241-150D4 (chr3: 41,487,711-41,273,927)
- CHORI241-25H7 (chr5: 95,112,604-95,305,354)

Lambda phage 37cen clone previously isolated from a horse genomic library in lambda phage (Anglana *et al.,* 1996) was used.

- 37cen - 221 bp repeat (Accession number: AY029358)

3.6.1 PLASMID DNA PURIFICATION

Satellite DNA containing recombinant plasmids was extracted from 10 ml of bacterial cultures. Bacterial clones were plated with 20 ml of LB medium (agar 1.8 % and ampicillin 100 µg/ml) and incubated at 37°C overnight. An isolated colony was taken from each plate and placed in 5 ml of liquid culture (LB medium and ampicillin 100 µg/ml). The tubes were then incubated under constant stirring at 37°C overnight. Aliquots were centrifuged at 13000 rpm (Mini-spin microcentrifuge, Eppendorf) for 1 minute. The supernatant was removed, and the pellet was resuspended in an appropriate volume of GTE buffer (50 mM glucose, 25 mM Tris-HCl pH 8, 10 mM EDTA pH 8). Then were added 0.2 N NaOH, 1% SDS and the test tubes

were incubated for 5 minutes on ice. The lysis reaction was stopped adding potassium acetate 3 M pH 4.8 and incubating the solution on ice for 5 minutes. The tubes were then centrifuged at 13000 rpm (Mini-spin microcentrifuge, Eppendorf) for 10 minutes and the supernatant was transferred to a new tube. RNase (20 μg/ml) was added to eliminate the RNA and the solution was incubated at 37°C for 20 minutes. Then, a volume of chloroform equivalent to that already contained in the tubes were added and centrifuged at 13000 rpm (Mini-spin microcentrifuge, Eppendorf) for 1 minute. The DNA remains in the upper phase was collected and gently transferred to another tube. The chloroform extraction was repeated twice. An equal volume of isopropanol was added. The tubes were then centrifuged at 13000 rpm (Mini-spin microcentrifuge, Eppendorf) for 10 minutes. Finally, the pellet was washed with 70 % Et-OH and resuspended in an appropriate volume of water.

3.6.2 BAC DNA PURIFICATION

Bacterial clones were plated with 20 ml of LB medium (agar 1.8% and chloramphenicol 12.5 μg/ml) and incubated at 37°C overnight. An isolated colony was taken from each plate and placed in 100 ml of liquid culture (LB medium and chloramphenicol 100 μg/ml). The extraction was carried out with Qiagen Plasmid purification kit®, according to supplier instructions.

3.6.3 LABELLING AND PRECIPITATION OF THE PROBES

Probes were labelled by nick translation with Cy3-dUTP (Perkin Elmer), Alexa488-dUTP (Invitrogen), digoxigenin-11-dUTP or biotin-16-dUTP (Roche). The nick translation reaction, performed at 15°C, had two different times depending on the d-UTP used (directly or indirect labelled). For nick translation with Cy3-dUTP (Perkin Elmer), Alexa488-dUTP (Invitrogen) time was 30 minutes (plasmid) and 60 minutes (BACs), instead for nick translation with digoxigenin-11-dUTP or biotin-16-dUTP (Roche) time was 90 minutes (plasmids) or for 180 minutes (BACs) and then, blocked with 0.5 mM EDTA.

Precipitation of the probes was obtained by adding 7.5 M ammonium acetate and absolute ethanol to the solution. Probes were then resuspended in a hybridization solution (50 % formaldehyde, 10 % dextran sulphate, 1X Denhart solution, 0.1 % SDS, 40 mM Na_2HPO4 pH 6.8 in 2XSSC) to a final concentration of 30 ng/µl when the probe was hybridized on chromatin fibres or 20 ng/µl when the probe was hybridized on metaphase chromosomes.

3.7 FISH – FLUORESCENCE *IN SIT*U HYBRIDIZATION

3.7.1 SLIDE AGING

Slides were aged at 90°C for 1 hour and 30 minutes. Then they were treated at 37°C for 30 minutes in a solution consisting of 0.005% pepsin/0.01 M HCl. Slides were washed for 3 times of 5 minutes at room temperature using PBS, a buffer consisting of PBS, 1M MgCl2, and a buffer of PBS, 1M MgCl2, 4% paraformaldehyde. Then slides were dehydrated for 5 minutes in the ethanol series (70 % EtOH; 90 % EtOH; and 100 % EtOH).

3.7.2 *IN SITU* HYBRIDIZATION ON METAPHASE CHROMOSOMES, INTERPHASE NUCLEI AND MICRONUCLEI

Slides were denaturated at 72°C for 4 minutes. Then the probes previously denaturated for 8 minutes at 80°C, were put on the slides and the slides were incubated at 37°C overnight in a moist chamber.

3.7.3 *IN SITU* HYBRIDIZATION ON EXTENDED CHROMATIN FIBRES

The slides were dehydrated in 3 steps, each of 3 minutes, in the ethanol series (EtOH 75%, 95%, 100%). The slides were then air-dried. Subsequently, the chromatin fibres on the slides were denatured with a

solution of 70% formamide in 2XSSC at 80°C for 4 minutes; the probes were instead denatured at 80°C for 8 minutes. After this step, the slides were immersed in a solution of 2XSSC at 4°C for 2 minutes and then subjected to the ethanol series as described previously. The slides were treated with denatured probe and the hybridization reaction was conducted at 37°C in a moist chamber for 12 hours. After the hybridization, the slides were washed 3 times for 5 minutes in 50% formamide, 2XSSC at 42°C. In case of indirectly probe labelling, slides were treated with blocking solution (3% BSA, 4XSSC, 0.1% Tween20), and incubated at 37°C for 30 minutes. Detection of the probe were then performed (see paragraph 3.7.4) Finally, the slides were mounted with 30µl of a solution of DAKO-DAPI (DAKO: 5 µl/ml).

3.7.4 POST HYBRIDIZATION WASHES AND PROBE DETECTION

Post-hybridization washes were performed in 50 % formamide, 2XSSC at 42°C. Then the slides were washed three times in 4XSSC, 0.1% Tween20 at 42°C temperature for 5 minutes. Slides hybridized with probes in which fluorescent dye molecules are directly bound to the dUTPs (Cy3-dUTP, Alexa488-dUTP) are counterstained with DAPI (4 ', 6'-Diamidino-2-phenylindole hydrochloride) (1 µg/ml) and then mounted with DAKO. Slides hybridized with BIO and DIG-labelled probes were treated with 100 µl of "blocking solution" (4XSSC, 0.1 % Tween20, 3 % BSA) and incubated at 37°C for 30 minutes in a moist chamber. DIG and BIO -labelled probes were detected with rhodamine and FITC respectively, using five successive layers of antibodies as follows: (i) anti-DIG-Rhod (sheep) 1/200 (Roche) (ii) anti-sheep Rhod 1/100 (CHEMICON) (iii) ExtrAvidin-FITC 1/100 (Sigma-Aldrich) (iv) biotinylated anti-avidin 1/200 (Sigma-Aldrich) (v) ExtrAvidin-FITC 1/100 (Sigma-Aldrich). All antibodies were incubated for 30 min at 37°C, then washed for three times of 5 minutes using 4XSSC, 0.1 % Tween20 at 42°C after each step of antibody detection. After the last antibody, the slides were washed two times for 5 minutes using 4XSSC, 0.1% Tween20 at 42°C. An additional wash in 4XSSC at room temperature were performed. Metaphase and chromatin fibres were stained with DAPI (4 ', 6'-Diamidino-2-phenylindole hydrochloride) (1 µg/ml) and then mounted with DAKO.

3.8 IMMUNOFLUORESCENCE

The slides were treated with the primary antibody (anti-CENP-A serum isolated in the lab of Professor Kevin Sullivan); anti-H3K9me3 antibody (abcam8898) and anti-H4K20me1 antibody (abcam 9051). The slides were subsequently incubated at 4°C for 24 hours in moist chamber. Slides were then washed for three times in KB^- (0,01 TRIS-HC l pH 8, 0.15 M NaCl, 0.5% BSA) of 5 minutes each. The slides were subsequently treated with 35µl of secondary antibody (anti-sheep conjugated with rhodamine (Jackson ImmunoResearch); anti-rabbit conjugated with Cy2 (Jackson ImmunoResearch); anti-rabbit conjugated with Alexa Fluor 647 (abcam 150179)). The slides were incubated at 37°C for 1 hour in moist chamber and then washed twice in KB^- for 5 minutes. Slides were treated with 4% paraformaldehyde in KCM for 7 minutes, followed by 2 washes in distilled water of 3 minutes each. If after the immunofluorescence an *In Situ* Fluorescence Hybridization was performed, the slides were immersed in cold methanol-acetic acid (3:1) for 15 minutes and washed in 2XSSC at room temperature for 2 minutes. Otherwise, the slides were stained with DAPI (4', 6'-Diamidino-2-phenylindole hydrochloride) (1 µg/ml) and then mounted with DAKO.

3.9 IMMUNO-FISH

For the slide preparation see previous paragraphs 3.4 and 3.5. For fixation, the slides were treated with a solution of 4% paraformaldehyde in KCM (0,12 M KCl, 0,08 mM NaCl, 0,01 M Tris-HCl pH 8, 0.5 mM EDTA, 0.1% Triton) for 10 minutes. Finally, the slides were treated with KCM for 5 minutes. Further steps are described in paragraphs 3.7.2 - 3.7.4.

3.10 MICROSCOPE ANALYSIS

The slides were analysed with a fluorescence microscope Axioplan (Zeiss) equipped with a cooled charge-coupled device (CCD) camera

(Photometrics). The CCD camera is characterized by a photo-sensor made of a matrix of silicon crystals sensitive to light. The photons contact the sensor and are converted into an electric charge proportional to the intensity of the light. After the exposure to light, the electric charge is transferred to the computer and converted into the binary system. The emission of each pixel is converted into a grey scale value which depends on the intensity and wavelength of the incident light. The tones in the scale vary from 256 to 4096. The images were acquired and pseudo-coloured with the IPLab spectrum software (Digital Pixel Advanced Imaging System, Brighton) and then processed using the software Photoshop®.

4. RESULTS – PART 1

CENTROMERIC SATELLITE DNA IS DISPENSABLE FOR MITOTIC TRANSMISSION FIDELITY

The coexistence of satellite-free and satellite-based centromere in equid species allowed us to study the role of played by satellite DNA in directing chromosome segregation fidelity. We analysed the mitotic stability of horse chromosome 11 (ECA11), whose centromere is satellite-free, and compared it with that of chromosome 13 (ECA13), which has similar size and a centromere containing long stretches of the canonical horse centromeric satellite DNA families (Piras *et al*., 2010; Nergadze *et al*., 2014; Cerutti *et al*., 2016).

Two chromosome stability assays, the interphase aneuploidy analysis and the Cytokinesis-Blocked MicroNucleus (CBMN) assay were performed. In both assays FISH experiments with chromosome specific DNA probes allowed us to unequivocally identify ECA11 and ECA 13.

The interphase aneuploidy analysis, together with the use of specific probes for centromeric sequences, allows the detection of chromosome specific aneuploidies, that is to say the gain or loss of single chromosomes in interphase nuclei (Raimondi *et al*., 1989; Faas *et al*., 2011). Different mechanisms can lead to aneuploidy: (i) nondisjunction, which can involve single or multiple chromosomes; (ii) merotelic attachment, which occurs when one kinetochore is attached to both mitotic spindle poles; and (iii) formation of multipolar/monopolar spindles.

The CBMN assay represents a reliable test for the detection of the gain, loss and breakage of chromosomes (Fenech, 2007). Centric and acentric chromosome fragments, as well as whole chromosomes unable to migrate to one pole during anaphase, can be included in telophase into micronuclei, so called because of their reduced size when compared to that of the cell nucleus. As micronuclei are expressed only in cells that have completed at least one cell division, in the micronucleus assay the use of

inhibitors of cell cycle progression (cytokinesis-block), such as cytochalasin, is recommended. In the CBMN assay only bi-nucleated cells, which underwent a cell division, are scored (Fenech and Morley, 1985; Kirsch-Volders *et al*., 2011).

Both assays were performed under control conditions and after exposure of the cells to mitotic stress induced by the spindle inhibitor nocodazole. The use of this drug was aimed at amplifying the difference, if any, in segregation fidelity between ECA11, with a satellite-free centromere and ECA13, whose centromere is satellite-based, and at identifying possible differences in their sensitivity to conditions perturbing cell division.

4.1 GROWTH CURVE

In order to assess the proper dose of Nocodazole to be used, two different doses, 100 nM and 200 nM, were tested. After a pre-culture period of 24h, cells were exposed to nocodazole for about two cell cycles (48 hours), then harvested and counted. The trypan blue stain was used in order to consider only viable cells.

In **Figure 15** the results of these experiments are reported. The highest dose (200nM) (grey line) induced a reduction in cell growth and a significant cell death after 48-hour treatment. Meanwhile, the lower dose (100 nM) impacted less negatively on the cell growth, causing growth depression after 48 hours with respect to the control (blue line). Therefore, as in this work the chemical is used in order to stress the cell and amplify possible differences in the stability of chromosome 11 with respect to chromosome 13, the best dose for our aims was 100 nM.

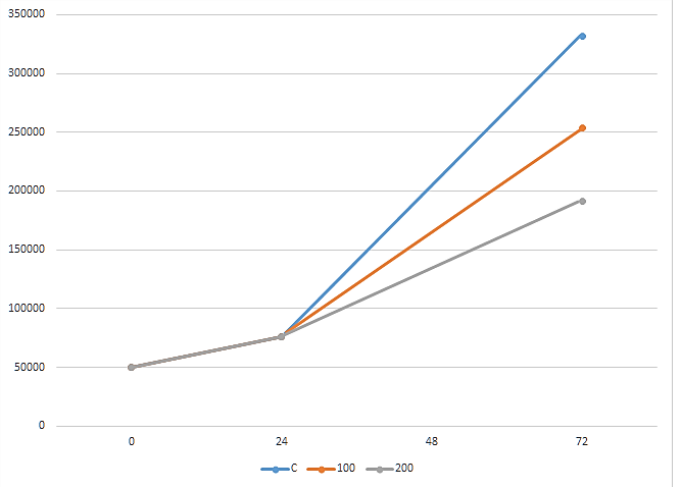


Figure 15 – Growth curve. Blue control, orange nocodazole 100nM, grey nocodazole 200nM.

4.2 INTERPHASE ANEUPLOIDY ANALYSIS

In order to compare the mitotic stability of horse chromosome 11 (ECA11) with that of horse chromosome 13 (ECA13), interphase FISH experiments were carried out with two BAC probes derived from the horse CHORI-241 BAC library (Leeb *et al*., 2006) and specific for ECA11 and ECA13, respectively (CHORI241-21D14, chr11: 27,639,936–27,829,952 and CHORI241-22C1, chr13: 7,346,775–7,544,907). In **Figure 16** two horse metaphase spreads, hybridized with the ECA11 and ECA13 specific probe respectively, are shown.

The position of the FISH signal on ECA11 coincides with the primary constriction (Fig. 16-A). As the centromere of ECA13 is satellite-based, a probe mapping on the short arm close to the centromere was used, as shown in Figure 16-B.

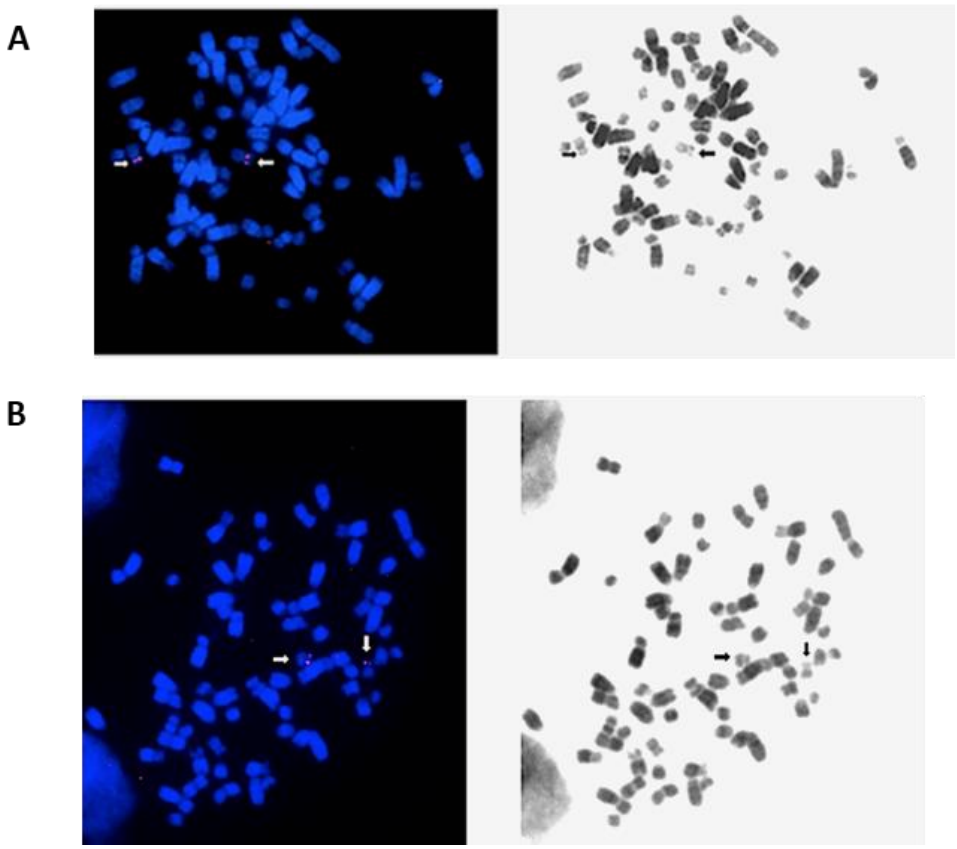


Figure 16 – Chromosomal localization by FISH of the BAC probes used. (A) FISH with the ECA11 specific BAC probe. (B) FISH with the ECA13 specific BAC probe. Chromosome 11 and 13 are marked by the arrows. Probe localization is marked by the red FISH signals in the images on the left. Chromosomes are counterstained with DAPI. Reverse DAPI banding is shown in the images on the right.

Interphase nuclei were analysed both in control condition and after exposure to nocodazole for a period of 48 hours (about two cell cycles).
Nuclei were screened for the presence of aneuploidy, monosomy, trisomy and nullisomy. The frequency of aneuploid cells was evaluated by counting the number of fluorescence spots, corresponding to each probe on

interphase nuclei, numbers differing from two indicating an aneuploid condition. At first, the two probes were labelled with two different fluorochromes (CY3 and CY2) and interphase nuclei were hybridized with both probes on the same slide. In particular, the probe for ECA11 was labelled with a fluorochrome (CY2) which emits in green spectrum, while the probe for ECA13 was labelled with a fluorochrome (CY3) which emits in red spectrum. **Figure 17** shows examples of the results of two-colour FISH experiments.

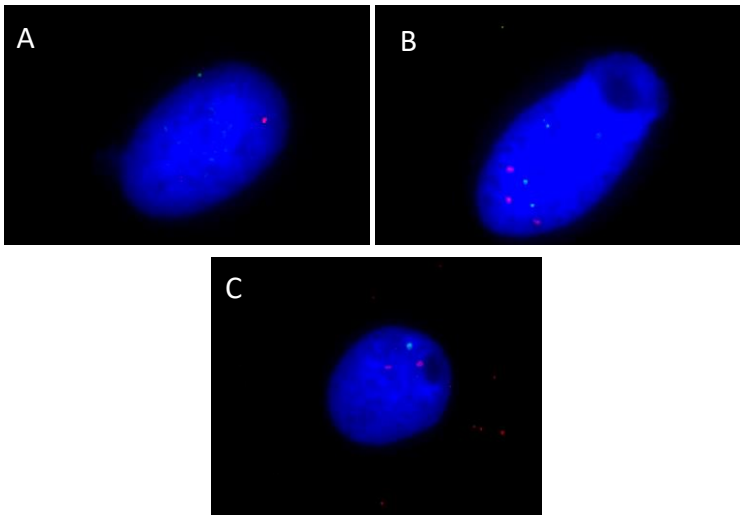


Figure 17. – Two colour FISH on interphase nuclei. A) Nucleus, monosomic for chromosome 11 and chromosome 13. B) Nucleus trisomic for chromosome 11 and chromosome 13. C) Nucleus monosomic for chromosome 11 and disomic for chromosome 13.

However, while performing the analysis, we observed that the two fluorophores displayed a remarkable variability in their stability within the same slide and among different slides, with the red fluorophore (CY3) being more stable than the green one (CY2). The difference in fluorophore stability could be done to the intrinsic batch-to-batch variability of the antibodies. - both the production of the antibody and its labeling with fluorochromes.

Indeed, since the frequency of chromosome specific aneuploidy and micronuclei was expected to be low, even minor differences in the efficiency or stability of different fluorophores would have impaired the results. Therefore, for the next experiments we decided to label both probes with the most stable fluorophore, which was the red one, CY3.

Figure 18 shows examples of the results of FISH experiments in which the probe for ECA11 and the probe for ECA13 were both labelled with the fluorophore CY3. The number of signals was then counted on interphase nuclei hybridized with the chromosome 11 or with the chromosome 13 specific probe.

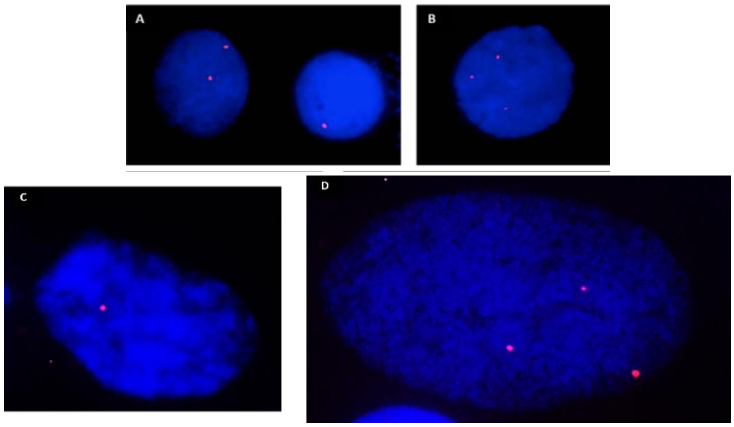


Figure 18 – FISH on interphase nuclei - (A) A nucleus, disomic for chromosome 11 is flanked by a monosomic one. (B) Nucleus trisomic for chromosome 13. (C) Nucleus monosomic for chromosome 13. D) Nucleus trisomic for chromosome 11.

Samples of 4.000 interphase nuclei, from three replicate experiments for chromosome 11 and 13, respectively, were analysed under control conditions. The results of the analysis are reported in **Table 1**. The χ^2 test was employed for the statistical analysis of the difference in the number of nuclei aneuploid for ECA11 or ECA13. The standard error calculated in the three replicates is reported in the table. Table 1 also shows that the percentages of nullisomic, monosomic and trisomic nuclei. Since this assay

does not allow to determine whether nuclei with four hybridization signals are tetrasomic or tetraploid, these nuclei were not included in the analysis.

Table 1. Interphase FISH under control conditions

Chromosome	Total Number of Nuclei	Diploid Nuclei (%)	Aneuploid Nuclei (% ± SE)	Type of Aneuploidy	Number (%)
ECA11	4000	3924 (98,10)	76 (1.90 ± 0.196)	*nullisomy*	5 (0,12)
				monosomy	30 (0,75)
				trisomy	41 (1,02)
ECA13	4000	3921 (98,03)	79 (1.98 ± 0.329)	*nullisomy*	8 (0,20)
				monosomy	32 (0,80)
				trisomy	39 (0,98)

The χ^2 test demonstrated that there was no difference in the number of nuclei aneuploid for ECA11 or for ECA13.

In order to compare the stability of the two chromosomes under stress condition, cells were exposed to 100 nM nocodazole for 48 hours and the number of aneuploid nuclei was counted in samples of about 1.000 nuclei from two replicates for ECA11 and ECA13, respectively, under control and stress conditions **(Table 2)**, the standard error calculated in the two replicates is reported in the table. Nocodazole is a well described antimitotic agent which interferes with the function of spindle and cytoplasmic microtubules by binding to tubulin. The use of this drug was aimed at amplifying the difference, if any, in segregation fidelity between ECA11, whose centromere is satellite-free, and ECA13, with a satellite-based centromere. Moreover, the test was aimed at identifying possible differences in the sensitivity of the two centromeres to conditions perturbing cell division.

Table 2. Interphase FISH under mitotic stress

Chromosome	Total Number of Nuclei	Diploid Nuclei (%)	Aneuploid Nuclei (% ± SE)	Type of Aneuploidy	Number (%)
Control					
ECA11	1107	1084 (97,92)	23 (2.08 ± 0.014)	*nullisomy*	0
				monosomy	11 (0,99)
				trisomy	12 (1,08)
ECA13	1117	1098 (98,30)	19 (1.70 ± 0.057)	*nullisomy*	0
				monosomy	11 (0,98)
				trisomy	8 (0,72)
Nocodazole 100 nM					
ECA11	1045	1000 (95,70)	45 (4.31 ± 0.085)	*nullisomy*	0
				monosomy	13 (1,24)
				trisomy	32 (3,01)
ECA13	1023	976 (95,40)	47 (4.60 ± 0.042)	*nullisomy*	0
				monosomy	12 (1,17)
				trisomy	35 (3,42)

Again, the statistical analysis of the results, performed by the χ^2 test, did not show any difference in the number of nuclei aneuploid for ECA11 or for ECA13 under control conditions and after drug treatment.

4.3 MICRONUCLEUS ASSAY

To test whether the formation of micronuclei is influenced by the presence/absence of centromeric satellite DNA, micronuclei containing ECA11 or ECA13 were identified by FISH with specific probes, the probes used were the same described for the interphase aneuploidy assay.

Also, in this assay, to minimize the experimental error, micronuclei containing chromosome 11 or chromosome 13 were counted in independent experiments, in which both probes were red labelled with Cy3. In a first set of experiments the frequency of spontaneously occurring micronuclei was counted in the same samples of 4.000 nuclei previously

analysed for aneuploidy. Since the observed frequencies of micronuclei were low, the statistical analysis of the results was performed using the Fisher's exact test. The results are reported in **Table 3.** The standard error calculated in the three replicates is reported in the table.

Table 3. Spontaneously occurring micronuclei under control conditions without cytokinesis-block

Chromosome	Total Number of Nuclei	Total Number of Micronuclei (% ± SE)	Micronuclei Containing the Chromosome (% ± SE)	Micronuclei not Containing the Chromosome (%)
ECA11	4000	63 (1.6 ± 0.098)	7 (11.1 ± 1.143)	56 (88,9)
ECA13	4000	55 (1.4 ± 0.127)	6 (10.9 ± 2.390)	49 (89,1)

No statistical difference was observed between the frequency of ECA11 and ECA 13 containing micronuclei.

Thereafter, the analysis was performed under control conditions and after exposure to 100 nM nocodazole. In order to score only cells which have completed mitosis during nocodazole treatment the Cytokinesis-Block MicroNucleus (CBMN) assay was applied. In this assay, the analysis is restricted only to micronuclei in binucleated cells, that is to say cells which underwent one cell duplication after drug treatment. To inhibit cytokinesis, thus allowing the formation of binucleated cells, cytochalasin was added to the cell cultures for the last 24 hours of nocodazole treatment (total treatment with nocodazole 48 hours) (Kirsch-Volders M. *et al*., 2002; Fenech M., 2007). Samples of control and treated cells were Giemsa stained to verify the presence of a high rate of binucleated cells with micronuclei, if present, being part of their cytoplasm **(Fig. 19)**.

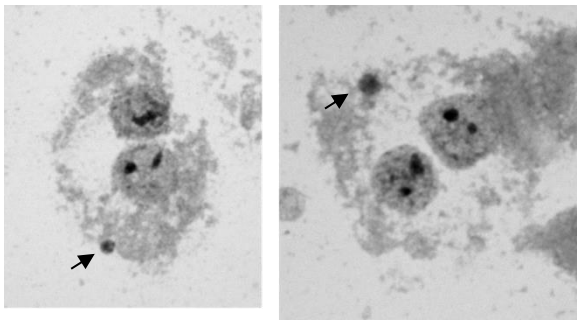


Figure 19. Binucleated cells containing micronuclei after treatment with cytochalasin - Giemsa staining. Micronuclei are indicated by the black arrows.

The results of the CBMN assay are reported in **Table 4,** the standard error, calculated in two replicates, is reported in the table. Nocodazole treatment caused a significant increase in the total number of micronuclei with respect to control both for chromosome 11 and 13, as demonstrated by the χ^2 test (p-value= 3×10^{-5} and 9×10^{-4}, respectively.). However, control as well as treated cells showed a comparable percentage of micronuclei containing each chromosome (Fisher's exact test). Examples of bi-nucleated cells with micronuclei are shown in **Figure 20**.
The BAC signals specific for ECA 11 and ECA 13 centromeres, are red labelled in the figure.

Table 4. CBMN under mitotic stress

Chromosome	Total Number of BN Cells	Total Number of Micronuclei (% ± SE)	Micronuclei Containing the Chromosome (% ± SE)	Micronuclei not Containing the Chromosome (%)
		Control		
ECA11	2000	31 (1.6 ± 0.002)	6 (19.4 ± 0.019)	25 (80,6)
ECA13	2000	32 (1.6 ± 0.000)	6 (18.8 ± 0.000)	26 (81,2)
		Nocodazole 100 nM		
ECA11	2000	75 (3.8 ± 0.004)	11 (14.7 ± 0.0004)	64 (85,3)
ECA13	2000	65 (3.3 ± 0.009)	10 (15.4 ± 0.0180)	55 (84,6)

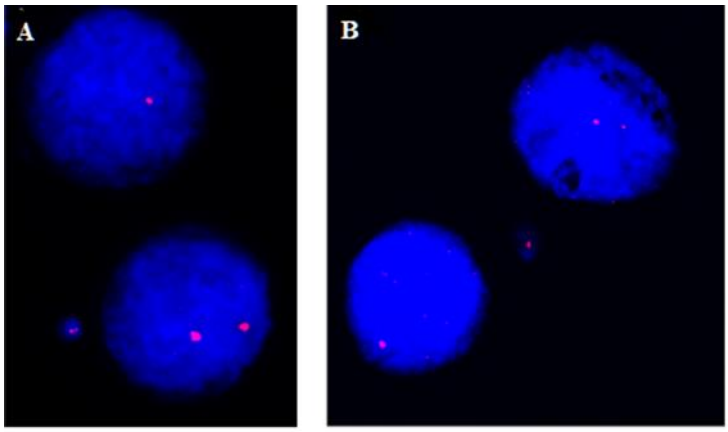


Figure 20 – FISH on micronuclei. A) Bi-nucleated Cyt-B treated cell with a disomic nucleus, a monosomic nucleus, and a chromosome 11 positive micronucleus. B) Bi-nucleated Cyt-B treated cell with a disomic nucleus, a monosomic nucleus, and a chromosome 13 positive micronucleus.

5. DISCUSSION – PART 1

The interphase nuclei aneuploidy assay and the micronucleus assay applied in this work were aimed at investigating the role played by satellite DNA at the centromere in directing proper chromosome segregation during mitosis. Equid species with their karyotypes characterized by the coexistence of satellite-based and satellite-free centromeres, represent an ideal model for the analysis of the differences in the behaviour of these two kinds of centromeres. I applied the assays cited above to detect possible differences in segregation fidelity between a horse chromosome with a satellite-less centromere (ECA11) and a horse chromosome with a satellite-based one (ECA 13).

Interphase analysis, coupled with FISH with chromosome specific probes, demonstrated that the two chromosomes segregate correctly during cell division and no difference exists in their segregation fidelity. Furthermore, the percentages of nullisomic, monosomic and trisomic nuclei for the two chromosomes were similar (Table 1). The assay was performed by analysing independently the segregation of ECA11 and that of ECA13 in one colour FISH experiments. This choice was determined by our previous observation that different fluorophores can give surprisingly different results due to their emission efficiency and stability after exposure to UV light. The expected frequency of single chromosome aneuploidy is very low, therefore, also slight differences in fluorophore stability and emission efficiency could impair the results. For this reason, our assay did not allow to determine whether nuclei containing four signals corresponded to tetrasomic or tetraploid conditions and therefore they were not included in the analysis of aneuploid cells. The percentage of nuclei containing four signals was 0,40% and 0,36% for ECA11 and for ECA13, respectively. The treatment with nocodazole induced, as expected, a statistically significant increase in the number of aneuploid nuclei for both chromosomes (χ^2 test, p-value $3x10^{-3}$ and $1x10^{-4}$), confirming that nocodazole treatment was effective. However, the difference in the number of aneuploidy nuclei

containing each chromosome was comparable under mitotic stress as well as in controls (Table 2).

These results demonstrate that the satellite-free horse centromere, as well as the satellite-based ones, directs correctly chromosome segregation, moreover, the two types of centromeres are equally sensitive to cell division disturbance by nocodazole.

The micronucleus assay gave results comparable to those obtained by the previous test. Also, in this case, no difference in the frequency of micronuclei formation was observed both in control condition, with and without cytokinesis-block, and in stress conditions after exposure to 100 nM nocodazole. The results of the micronucleus assay therefore confirmed that the mitotic behaviour of the two chromosomes was comparable and that the absence of satellite DNA at the centromere does not affect its sensitivity to the spindle inhibitor nocodazole (see Table 3 and 4).

Thus, two independent molecular-cytogenetic approaches demonstrated that, in the horse fibroblast system, the segregation fidelity of a chromosome is not influenced by the presence of highly repetitive DNA sequences at its centromere.

6. CONCLUSION AND PERSPECTIVES – PART 1

Equid species have karyotypes which are characterized by the coexistence of chromosomes with canonical satellite-based centromeres with chromosomes whose centromeres are satellite-free (Wade *et al*., 2009; Piras *et al*., 2010; Raimondi *et al*., 2011; Nergadze *et al*., 2014; Purgato *et al*., 2015; Cerutti *et al*., 2016; Nergadze *et al*., 2018). This exceptional biological model offers the opportunity to directly investigate, in a natural environment, the behaviour of these differently organized centromeres. The two chromosome stability assays performed to understand the role of satellite DNA in driving proper chromosome segregation unequivocally demonstrated that there is no difference in the segregation fidelity of chromosome 11 (ECA 11), with a satellite-free centromere and that of chromosome 13 (ECA 13) with a canonical satellite-based centromere.

Interestingly, sequence analysis of the centromere of horse chromosome 11 showed that no motif resembling a CENP-B box is present (unpublished results), suggesting that this protein at the equid centromeres is not fundamental for centromere function. The role of the CENP-B protein at the centromere is a matter of debate. (Masumoto *et al*., 1989; Miga *et al*., 2014; Fachinetti *et al*., 2015; Drinnenberg *et al*., 2016; Logsdon *et al*., 2019). Recently, Logsdon and co-workers (Logsdon *et al*., 2019) demonstrated that artificial human neocentromeres don't require for their formation neither alphoid-DNA, nor CENP-B binding sequences, being originated only by an initial CENP-A nucleosome seeding.

The fact that the absence of satellite DNA at a horse centromere does not influence chromosome segregation fidelity confirms that factors, other than the DNA sequence, are important for correct centromere function. These factors are mostly epigenetic. Indeed, literature data demonstrate that a fine balance among post-translation histone modifications is needed at the centromere (Blower *et al*., 2002; Peters *et al*., 2003; Sullivan *et al*., 2004; Bergmann *et al*., 2011). Transcription competence is another important characteristic of the centromere core (Denegri *et al*., 2002; Volpe

et al., 2002; Hall *et al*., 2003; Ugarkovic, 2005; Carone *et al*., 2009; Eymery *et al*., 2009; Brown *et al*., 2012; Gent and Dawe, 2012; Koo *et al*., 2016; Rosic and Erhardt, 2016) (as it was discussed in paragraph 1.4.4).

The results of the first part of my Ph.D. work strongly suggests that satellite DNA at the centromere plays roles other than the maintenance of chromosome segregation fidelity. One possible function, based on our previous results on centromere sliding (Purgato *et al*., 2015; Nergadze *et al*., 2018) is that satellite DNA may contribute to constrain the borders of the functional centromere domain within non-coding genomic regions. Moreover, it may be hypothesized that, despite the absence of the canonical centromeric satellite DNA sequences, horse chromosome 11 is stabilized by the chromatin context in which it is located. It is also tempting to speculate that transcripts derived from the satellite DNA of the other canonical horse centromeres may act in trans thus complementing satellite DNA deficiency at ECA11 (Cerutti *et al*., 2016).

In the next future, it will be interesting to study in depth the role of centromeric transcripts in the horse and in other equid species. An intriguing possibility will be to analyse, by means RNA-FISH coupled with immunofluorescence, the association of centromeric transcripts with satellite-less centromeres and to assess whether their localization changes during the cell-cycle.

7. RESULTS – PART2

HISTONE MODIFICATIONS AT EQUID SATELLITE-BASED AND SATELLITE-FREE CENTROMERES

The presence of satellite DNA at most of eukaryotic centromeres has been hypothesized to provide a suitable chromatin environment for centromere maintenance and for the cohesion and separation of sister chromatids (Henikoff and Dalal, 2005). However, the lack of a universal centromere sequence and the existence of centromeres completely void of repetitive DNA (neocentromeres) suggest that the centromere identity is defined epigenetically (Sullivan and Karpen, 2004; Hall *et al*., 2012).

The centromere specific histone H3-like protein CENP-A is a conserved marker of centromere function (Blower *et al*., 2002). Additional molecular features contribute to define centromere activity, including the composition of centromeric and pericentromeric chromatin and post-translational modifications of CENP-A itself. Indeed, a centromere specific ratio between typical euchromatic and typical heterochromatic histone modifications is crucial for centromere identity, creating a 'permissive' chromatin structure needed for CENP-A recruitment (McKinley and Cheeseman, 2016; Mellone and Allshire, 2003; Quenet and Dalal, 2014).

The genus *Equus* is a powerful model system to study the structure, identity and function of mammalian centromeres, since satellite-based and satellite-free centromeres coexist in single karyotypes (Wade *et al*., 2009; Piras *et al*., 2010; Nergadze *et al*., 2018). The second part of this work was focused on the study of the epigenetic environment of horse and donkey satellite-free and satellite-based centromeres. Several post-translational histone modifications were analysed: H3K9me3 (constitutive heterochromatin marker), H3K27me3 and H3K9me2 (facultative heterochromatin markers), H3K4me2 (transcriptionally competent chromatic marker) and H4K20me1 (CENP-A containing nucleosomes marker).

Double immunofluorescence was performed on horse and donkey metaphase chromosomes and chromatin fibres to analyse the arrangement of the histone modifications mentioned above. Moreover, immuno-FISH on horse and donkey metaphases chromosomes and on horse and donkey chromatin fibres was performed in order to deeply analyse the heterochromatin centrochromatin environment as identified by the H3K9me3 histone modification.

7.1 CONSTITUTIVE HETEROCHROMATIC MARKER H3K9me3

The trimethylation of lysine 9 of histone H3 (H3K9me3) is the best-known constitutive heterochromatin marker. It was demonstrated to be present in the pericentromeric regions of Drosophila, mouse and human chromosomes (Peters *et al*., 2003; Rice *et al*., 2003) and absent at the centromere functional core. On human chromosomes, H3K9me3 is concentrated in the pericentromeric regions of chromosomes that contain large blocks of satellite DNA and is also located in regions far away from CENP-A-containing domains (Sullivan and Karpen, 2004).

7.1.1 Double immunofluorescence on metaphase chromosomes

In order to analyse the architectural organization of the H3K9me3 histone modification, I performed double immunofluorescence experiments on horse and donkey metaphase chromosomes with antibodies against CENP-A, to mark centromere function, and antibodies against H3K9me3 to mark constitutive heterochromatin. In **Figure 21**, examples of the results of double immune-fluorescence analysis of horse and donkey metaphase chromosomes is shown. The H3K9me3 signal is pseudo-coloured in green while CENP-A is pseudo-coloured in red.

In the horse (Fig. 21-A), two spots corresponding to CENP-A were clearly observed at all the centromeres. The H3K9me3 fluorescence signal was very bright at all centromeres with the exception of one couple of centromeres (white arrows in Fig. 20-A) which showed a very weak signal. A sample of 20 metaphase spreads was analysed and they all displayed the same picture.

Sequence analysis and the analysis of the distribution of satellite DNA in the horse (Wade *et al*., 2009; Piras *et al*., 2010) demonstrated that the centromere of chromosome 11 is void of satellite DNA, therefore we suppose that the centromeres showing a faint H3K9me3 signal are those of chromosomes 11.

In the donkey (Fig. 21-B), as observed in the horse, double spotted CENP-A signals were present at all centromeres. Conversely, the distribution of H3K9me3 signals was different. Heterochromatic signals were highly variable in intensity among chromosomes. The analysis of 20 metaphases showed that 32 centromeres always displayed a faint (white arrows) or barely perceptible (red arrows) fluorescence signal. Notably, most of the chromosomes which showed a weak centromeric heterochromatic signal exhibited a large H3K9me3 positive region at one telomere.

These results are in agreement with previous data concerning the distribution of satellite DNA families in the donkey (Piras *et al*., 2010) and with the results of sequence analysis (Nergadze *et al.,* 2018). We hypothesize that the donkey centromeres showing a faint H3K9me3 signal are indeed those with satellite less centromeres, some of which conserve residual satellite DNA sequence at one non centromeric terminus (Piras *et al*., 2010).

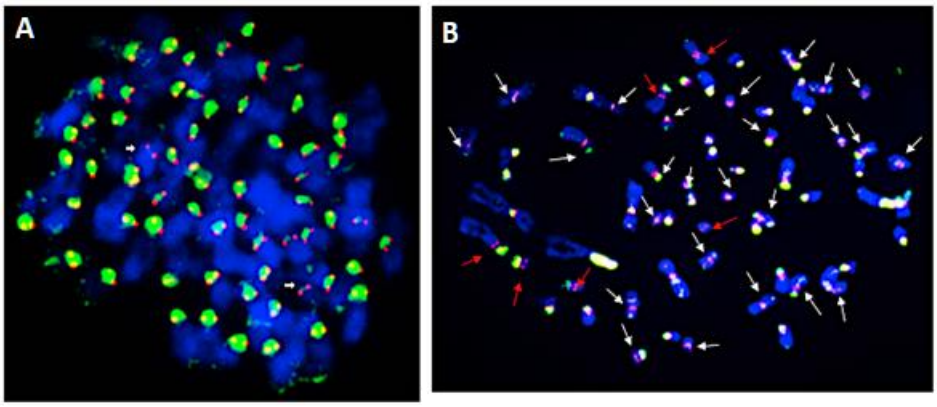


Figure 21 – Two colour immunofluorescence on horse (A) and donkey (B) metaphase chromosomes. CENP-A protein is red labelled while the trimethylated histone 3 (H3K9me3) is green labelled. The arrows indicate centromeres with a faint (white arrows) or barely perceptible (red arrows) heterochromatic signal.

7.1.2 Immuno-FISH on metaphase chromosomes

The results obtained from the double immunofluorescence experiments, raised the question if those centromeres showing a faint heterochromatic signal were the satellite-less ones. In order to answer this question, we compared the intensity of the heterochromatic fluorescence signal observed in satellite-based centromeres with that observed in the satellite-less ones in the horse and in the donkey. To this purpose, we set up immuno-FISH analyses using an antibody directed against the H3K9me3 epigenetic modification and chromosome specific BAC probes for the identification of satellite-less centromere and satellite-based ones.

We first compared the satellite-less centromere of horse chromosome 11 (ECA11) with the satellite-based one of chromosome 13 (ECA13). The BACs used as probes (CH241-21D14 for ECA11 and CH241-22C1 for ECA 13) were previously described (see chapter 3, paragraph 3.6 of this thesis).

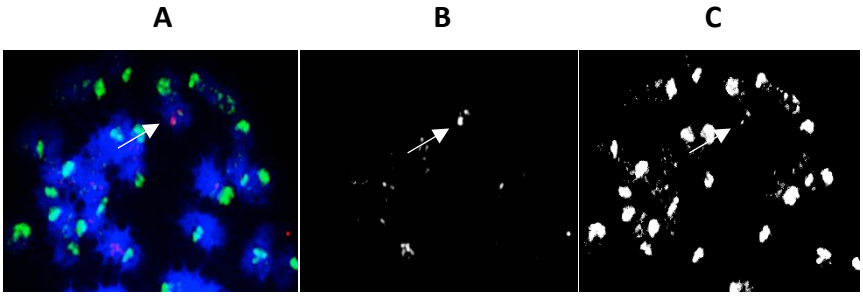


Figure 22 – Immuno-FISH analysis on ECA11. A) Merged image. Metaphase chromosomes are blue coloured. In red the FISH signal identifies ECA11 centromere while the green signal identifies the H3K9me3 histone modification B) Only the FISH signal is shown. C) Only the H3K9me3 immuno-signal is shown.

Figure 22-A shows a merged image of immuno-FISH on horse metaphase chromosomes. It can be observed that the FISH signal and the immuno-signal colocalize on the chromosome. In panel B) only the red channel is shown. An intense FISH signal is clearly visible at ECA11 centromere (indicated by the white arrows in panel B). The white arrows in panel C indicate the H3K9me3 signal which is very weak. The analysis of 20 images gave comparable results

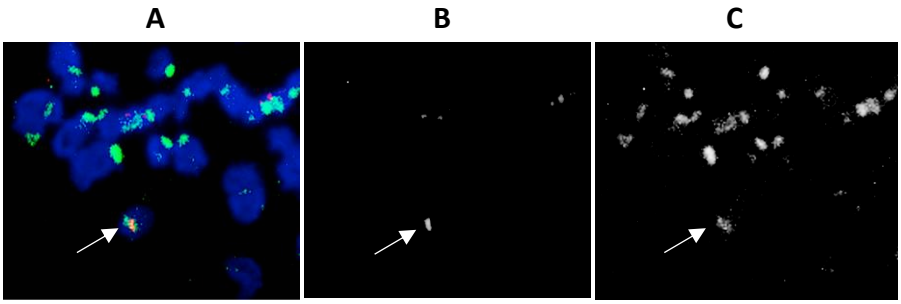


Figure 23 – Immuno-FISH analysis on ECA13. A) Merged image. Metaphase chromosomes are blue coloured. In red the FISH signal identifies ECA13 and the green signal identifies the H3K9me3 histone modification. B) Only the 22C1 FISH signal is shown. C) Only the H3K9me3 immuno-signal is shown.

Figure 23-A shows the results of the immune-FISH experiment performed to analyse the satellite-based centromere of ECA13. As in the previous experiment, 20 metaphases were analysed. The H3K9me3 signal was always very intense if compared with that observed on the centromere of ECA11 **(Fig. 22)**.

I then analysed by immuno-FISH a satellite-based and a satellite-less centromere in the donkey. The satellite–less centromere of EAS7 was compared with the satellite-based centromere of EAS3. The BAC used to identify EAS7 was selected from the horse CHORI BAC library. It was localized on ECA8, which is homologous to EAS7, in a region corresponding to EAS7 centromere (coordinates: ECA8: 41,985,542 - 42,131,402). To identify EAS3 the BAC CH241-150D4 was used. This BAC localizes on ECA3, which is homologous to EAS3, in a region proximal to EAS3 centromere (coordinates ECA3: 41,487,711-41,273,927).

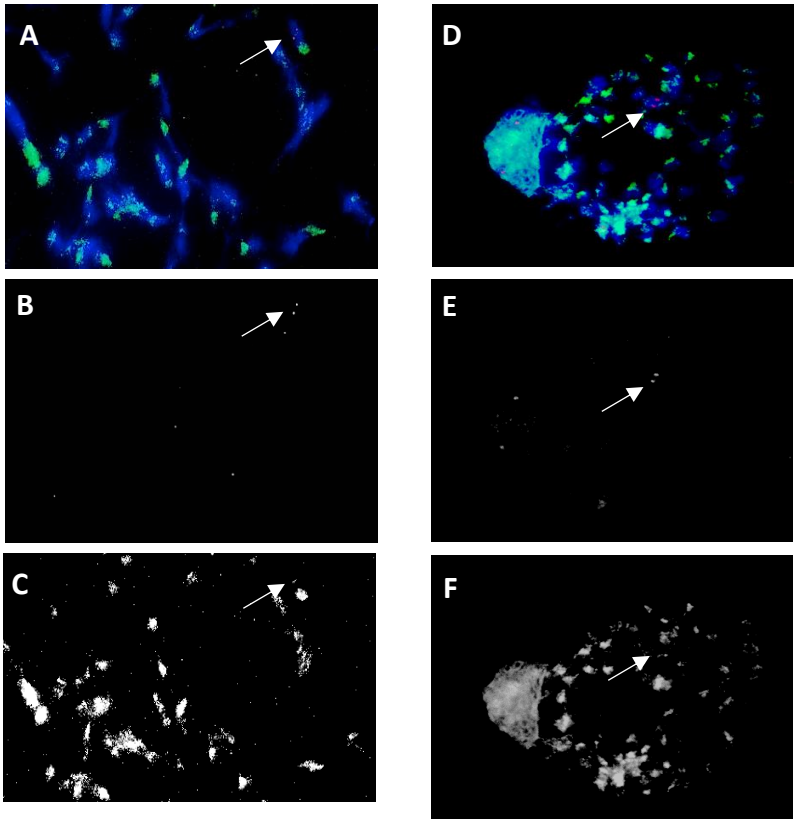


Figure 24 – Immuno-FISH analysis on EAS7. A-D) Merged images. Metaphase chromosomes are blue coloured. In red the FISH signal identifies the centromere of donkey chromosome 7, the green signal identifies the H3K9me3 histone modification B-E) Only the FISH signal is shown C-F) Only the H3K9me3 immuno-signal is shown**.**

The results of immuno-FISH on EAS7 are shown in **Figure 24**. The analysis of 20 metaphases always showed a very faint heterochromatic signal on the centromere of EAS7 (panel C, F in Figure 24). The heterochromatic signal was then analysed at the satellite-based centromere of donkey chromosome 3 **(Fig. 25)**.

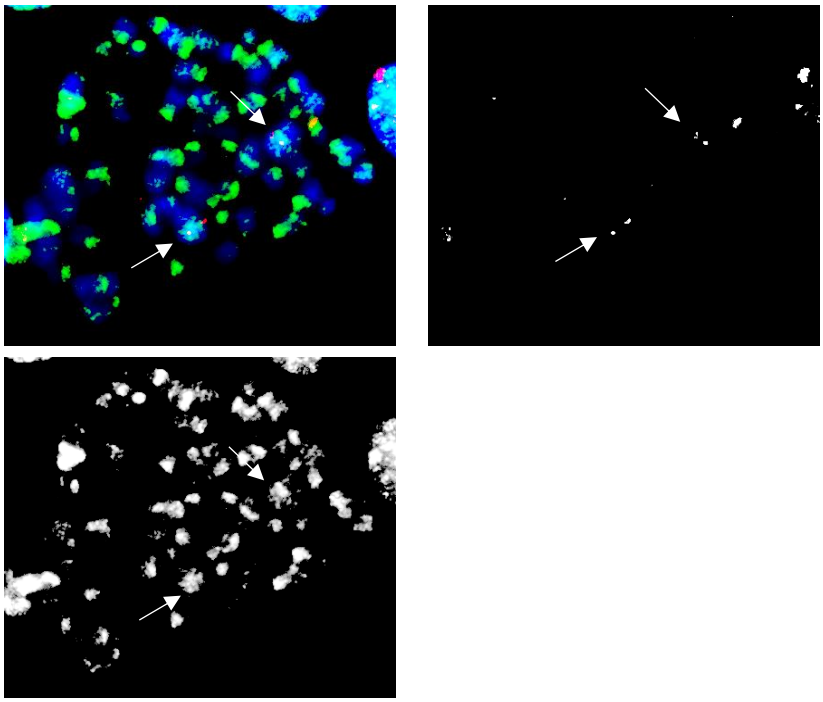


Figure 25 – Immuno-FISH analysis on EAS3. A) Merged image. Metaphase chromosomes are blue coloured. In red the FISH signal identifies EAS3 chromosome and the green signal identifies the H3K9me3 histone modification. B) Only the 150D4 FISH signal is shown. C) Only the H3K9me3 immuno-signal is shown.

The fluorescence signal identifying H3K9me3 was always very strong in 20 metaphase spreads (**Fig. 25**, panel C), if compared with that observed at the centromere of EAS7 (Fig. 24).

To further confirm the data obtained by the double immuno-fluorescence and immuno-FISH experiments described above and as a positive control, I performed three colour immuno-FISH experiments on horse metaphase chromosomes. The purpose of this experiment was to identify centromere function using the anti CENP-A antibody, heterochromatin using the anti H3K9me3 antibody and satellite DNA using a plasmid probe complementary to the main horse satellite DNA family

37cen. The experiment was aimed at having a positive control of the presence of strong heterochromatic signals at all satellite-based centromeres. The results are shown in **Figure 26**. The intensity of the heterochromatic signals (panel D) was very variable among centromeres. Actually, in the previous two-colour immuno-FISH experiment the heterochromatic signals were stronger (see Fig. 23). However, for three colour immuno-FISH, a different, infrared emitting, secondary antibody had to be used to identify the H3K9me3 histone modification. The different signal intensity observed in the two experiment may therefore be due to the different emission efficiency of the secondary antibodies.

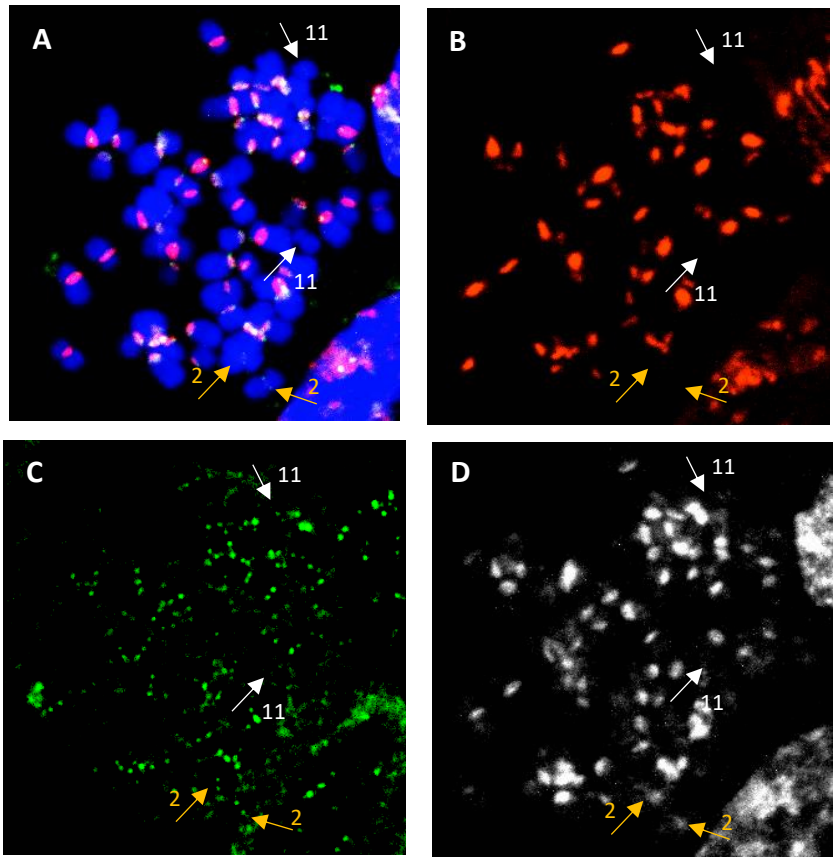


Figure 26 – Three colour immuno-FISH on horse metaphase chromosomes. A) Merged image. Metaphase chromosomes are blue coloured. In red the FISH signal identifies the functional centromeric satellite DNA (37cen), the green signal identifies the CENP-A protein and the grey signal identifies the H3K9me3 histone modification. B) Only the 37cen FISH signal is shown. C) Only the CENP-A immuno-signal is shown D) Only the H3K9me3 immuno-signal is shown.

As expected, four horse centromeres were negative to the 37cen FISH signal (panel B in fig. 26). Indeed, the centromere of ECA11 is completely satellite-free and the centromere of ECA2 is satellite DNA based positive, but contains only the 2PI satellite DNA family, while is 37cen negative (see Piras *et al*., 2010). Two of the FISH negative centromeres,

seemingly those of ECA11 which is satellite DNA free (white arrows in all the panels), showed a not perceptible histone modification signal (panel D). The other two FISH negative centromeres, presumably corresponding to ECA2, which contains the 2PI satellite DNA family (yellow arrows in all the panels), showed a heterochromatin signal comparable to that observed on other centromeres. The analysis of 10 metaphase spreads gave similar results.

7.1.3 Double immunofluorescence and immuno-FISH on chromatin fibres

In order to analyse the architectural organisation of the H3K9me3 post-translational histone modification at the functional centromere core, high resolution double immuno-fluorescence and immuno-FISH experiments were performed on horse and donkey extended chromatin fibres. A quantitative analysis of the intensity of the fluorescence signals on chromatin fibres was performed using the ImageJ and Plot2 software. In each experiment a total number of 15 chromatin fibres was analysed.

For double immunofluorescence, an anti CENP-A and an anti H3K9me3 antibody were used. **Figure 27** shows the results obtained from the analysis of horse chromatin fibres (Fig. 27-A) and those obtained from the analysis of donkey chromatin fibres (Fig. 27-B). It can be observed that in both images, constitutive heterochromatin is dispersed all along the CENP-A positive region, the fluoresce intensity profiles of CENP-A and H3K9me3 showing a comparable distribution.

In the horse, in which only the centromeres of chromosome 11 are satellite-free, all the fibres analysed displayed a fluorescence profile similar to that shown in Figure 26-A. On the contrary, in the donkey, whose karyotype is composed of 16 couples of chromosomes with satellite-free centromeres and 15 couples of chromosomes with satellite-based centromeres, 8 out of 15 fibres, that is about half, showed the profile

displayed in Figure 26-C, which is characterized by a much less intense heterochromatic fluorescence profile. Therefore, we supposed that the two kinds of profiles corresponded to satellite-based (Fig. 26-B) and satellite-less (Fig. 26-C) centromeres, respectively.

To confirm this prediction, we performed immuno-FISH experiments, using an antibody against CENP-A and a probe complementary to the functional centromere sequence of a satellite-free centromere on chromatin fibres from the horse and from the donkey.

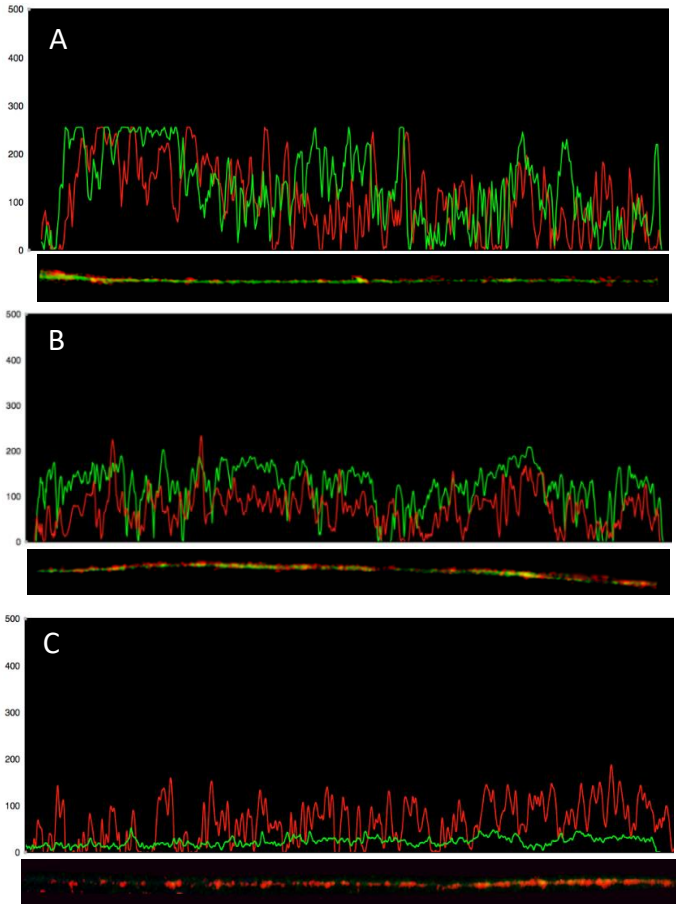


Figure 27 – Two colour immunofluorescence on horse (A) and donkey (B, C) chromatin fibres. The graph over the image of the fibre in A, B and C, reports the quantitative analysis (ImageJ software) of the intensity of the fluorescence signals. The CENP-A protein is red labelled while H3K9me3 is green labelled.

The BAC probes used were the same used for the immuno-FISH experiments on metaphase chromosomes (Paragraph 7.1.2).

Figure 28 shows that both the horse (Fig. 28-A) and the donkey satellite-free centromeres (Fig. 28-B) identified by the FISH signal are integrated in a heterochromatic environment identified by the H3K9me3 histone mark. However, the fluoresce intensity profile clearly demonstrates that H3K9me3 fluorescence intensity at these satellite-less centromeric domains is much lower than that shown in Fig. 26-A and B and comparable to that shown in Fig. 26-C. These results confirm that both donkey and horse satellite-free centromeres, while being embedded in a heterochromatic environment, however, contain a much lower amount of heterochromatin marks with respect to the satellite-based ones.

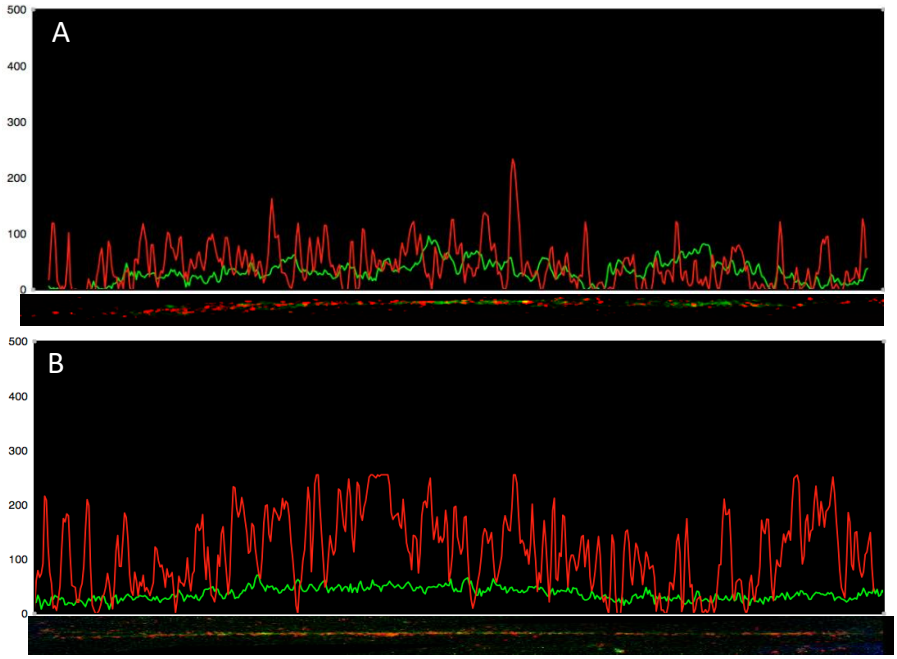


Figure 28 – Immuno-FISH on horse (A) and donkey (B) chromatin fibres. The graph over the image of the fibre in A and B, reports the quantitative analysis (ImageJ software) of the intensity of the fluorescence signals. In A and B, the BAC signal is red labelled while H3K9me3 is green labelled.

In order to have a further positive and negative control of the double immuno-fluorescence and immuno-FISH experiments described above, I analysed the H3K9me3 histone modification in a donkey non centromeric heterochromatic region and in a donkey non centromeric euchromatic region.

I first analysed a genomic heterochromatic region by immuno-FISH using an antibody against the H3K9me3 histone modification and a BAC probe specific for a non centromeric region which was shown to be heterochromatic by ChIP-sequencing. The BAC probe used was CH241-21D14, mapping on horse chr11: 27,639,936 – 27,829,952 which detects an orthologous heterochromatin region on donkey chromosome 13. A sample image of the results obtained is shown in **Figure 29**. As expected, the region identified by the BAC is associated to a very strong H3K9me3 fluorescence profile.

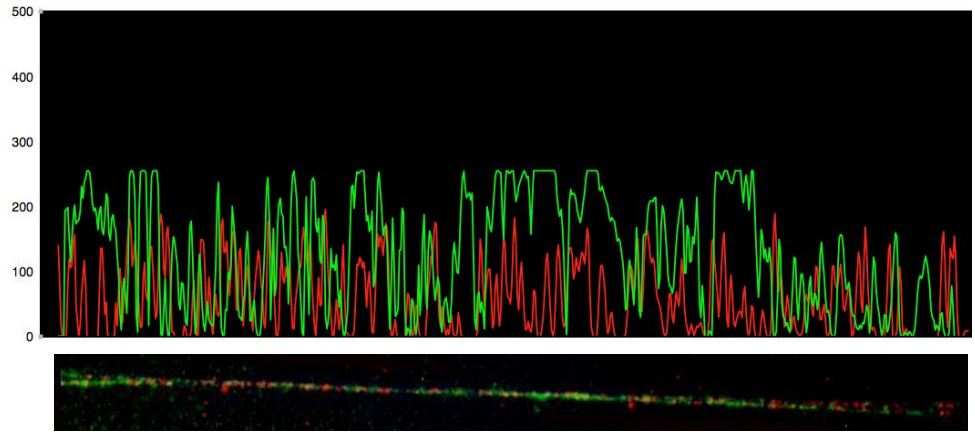


Figure 29 – Immuno-FISH on donkey chromatin fibres. The graph over the image reports the quantitative analysis (ImageJ software) of the intensity of the fluorescence signals. The 21D14 FISH signal is red labelled while H3K9me3 is green labelled.

Then I analysed a donkey euchromatic region identified by the BAC CH241-25H7, mapping on horse chromosome 5 (coordinates 95,112,604-

95,305,354) which detects an orthologous euchromatic region on donkey chromosome 16 (EAS16), as confirmed by ChIP-sequencing. The results shown in **Figure 30**, demonstrate that, as expected, in the euchromatic region identified by the BAC, no fluorescence signal corresponding to the heterochromatic marker H3K9me3 was visible.

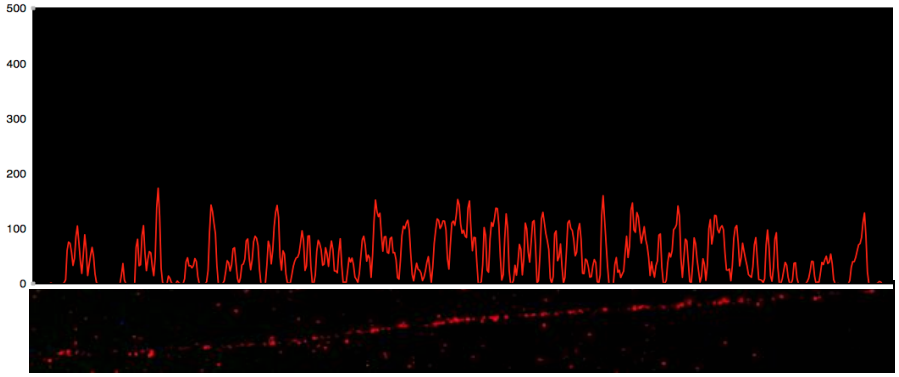


Figure 30 – Immuno-FISH on donkey chromatin fibres. The graph over the image reports the quantitative analysis (ImageJ software) of the intensity of the fluorescence signals. The 25H7 FISH signal is red labelled while H3K9me3 is green labelled.

Finally, to analyse satellite-based functional centromeric regions, I performed three-colour immuno-FISH experiments on horse chromatin fibres. This experiment was aimed at having a positive control of the presence of strong heterochromatic signals at all satellite-based centromeres. The centromere function was marked with an anti CENP-A antibody, heterochromatin by anti H3K9me3 antibody and the centromeric satellite DNA by a plasmid probe complementary to the main horse satellite DNA family 37cen. **Figure 31** demonstrates that at satellite-based centromeres a strong heterochromatic signal (green) is present both at the centromere core identified by CENP-A (grey) and in the pericentromeric region. As expected, the 37cen signal (red) is present at the functional centromere core and extends in the flanking regions.

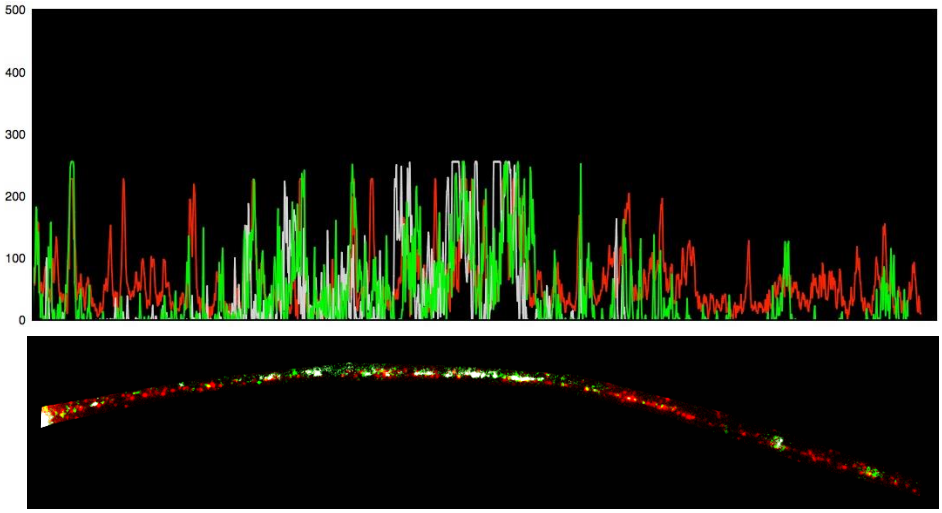


Figure 31 – Three colour immuno-FISH on horse chromatin fibres. In red the FISH signal identifies the functional centromeric satellite DNA (37cen), the grey colour identifies the CENP-A protein and the green colour identifies the H3K9me3 histone modification.

7.2 H3K9me2, H3K27me3 AND H3K4me2 MARKERS

The dimethylation of lysine 9 of histone H3 (H3K9me2) is a marker of facultative heterochromatin; this type of histone modification is present in silenced regions that can be activated when necessary (Mravinac *et al*., 2005). This modification is present in the pericentromeric regions both in man and in Drosophila. (Sullivan and Karpen, 2004; Bailey *et al*., 2016). H3K27me3 can be found both in chromatin regions containing satellite DNA and in regions containing single copy DNA (Aldrup-Macdonald and Sullivan, 2014; Miga, 2015).

The dimethylation of histone H3 at lysine 4 (H3K4me2) is involved in the transition from an inactive to an active chromatin state. This modification has been found at promoters and transcribed genomic regions, and it is related to a not necessarily active euchromatin (Schneider *et al*., 2004; Lam *et al*., 2006). At the centromere this histone modification was

found in the functional core where correlates with transcriptionally competent chromatin (Bergmann *et al.*, 2011; Sullivan and Karpen, 2004).

7.2.1 Double immunofluorescence on metaphase chromosomes and chromatin fibres

The distribution of H3K9me2 was analysed in horse **(Fig. 32-A)** and in donkey **(Fig. 32-B)** metaphase chromosomes by immunofluorescence using antibodies against H3K9me2 and antibodies against CENP-A. No difference in the intensity and distribution of the fluorescence signals was observed in the two species, the two markers co-localising at all the centromeres. These results indicate that satellite-free and satellite-based centromeres both contain heterochromatin prone to be opened, but not necessarily transcribed.

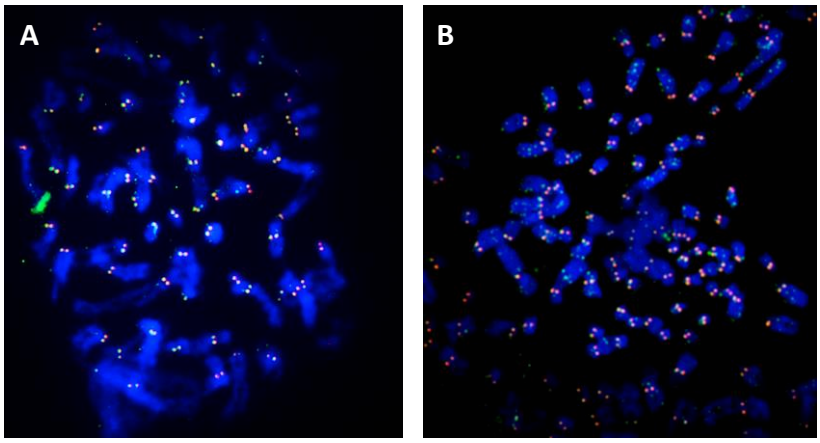


Figure 32 – Two colour immunofluorescence on horse (A) and donkey (B) metaphase chromosomes. Chromosomes are stained with DAPI, blue, the CENP-A protein is red labelled and H3K9me2 is green labelled.

The analysis of the same histone modification oh horse and donkey extended chromatin fibres confirmed the data. Again, no difference was observed between horse and donkey and all the analysed fibres (15) showed the same pattern were CENPA and H3K9me2 signals co-localized. Examples

of a horse chromatin fibre (Fig. 33-A) and of a donkey chromatin fibre (Fig. 33-B) are reported in **Figure 33**.

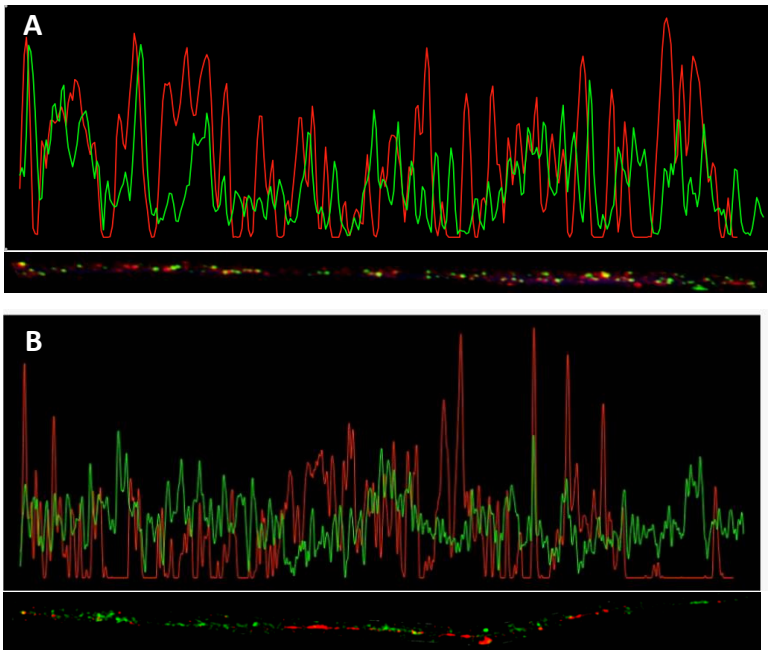


Figure 33 – Two colour immunofluorescence on horse (A) and donkey (B) chromatin fibres. The graph over the images reports the quantitative analysis (ImageJ software) of the intensity of the fluorescence signals. In A and B, the CENP-A protein is red labelled and H3K9me2 is green labelled.

The second facultative heterochromatin marker studied was H3K27me3. Double immuno-fluorescence on horse **(Fig. 34-A)** and donkey **(Fig. 34-B)** metaphase chromosomes did not show any significant difference in the intensity and distribution of fluorescence signals between the two species and all the centromeres were positive to both CENP-A and H3K27me3. Interestingly, the pattern of distribution of the H3K27me3 marker somehow recalled a G banding pattern. This result is not surprising since G banding identifies late replicating chromosome regions, which are typically heterochromatic and gene-poor.

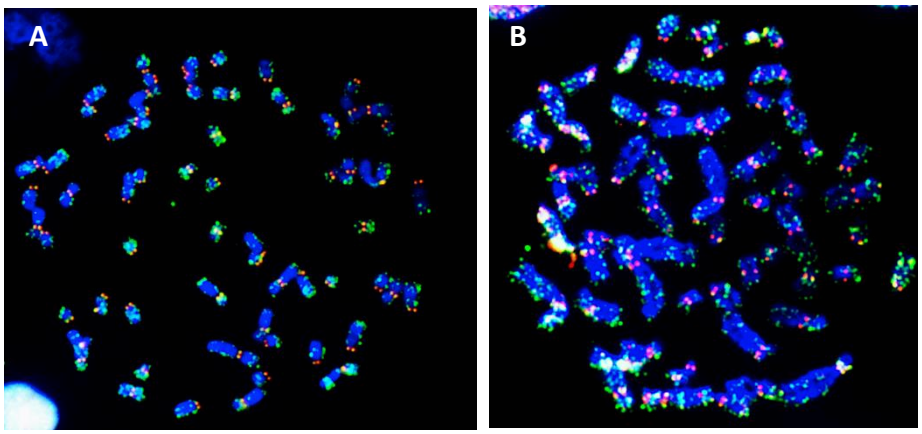


Figure 34 – Two colour immunofluorescence on horse (A) and donkey (B) metaphase chromosomes. Metaphase chromosomes are stained with DAPI, blue, CENP-A is red labelled and H3K27me3 is green labelled.

The analysis of extended chromatin fibres confirmed that, in the horse and in the donkey, the histone modification H3K27me3 was present at the centromere core, identified by the CENP-A protein. In **Figure 35** examples of a horse chromatin fibre (Fig. 34-A) and of a donkey chromatin fibre (Fig. 34-B) are shown.

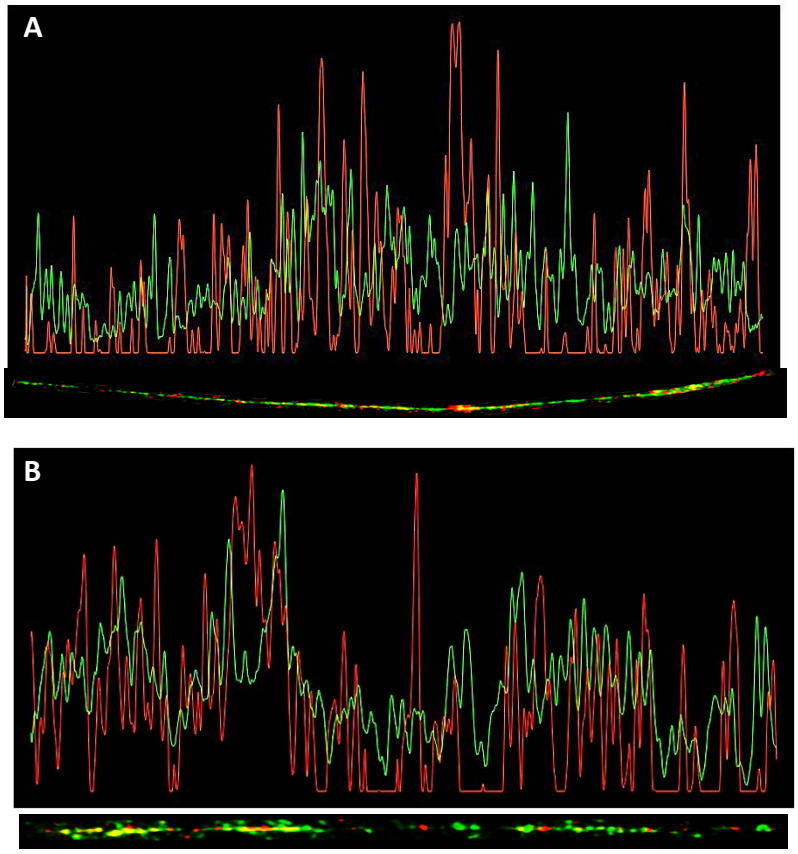


Figure 35 – Two colour immunofluorescence on horse (A) and donkey (B) chromatin fibres. The graph over the images of the fibres in A and B, represents the quantitative analysis (ImageJ software) of the intensity of the fluorescence signals. In A and B, the CENP-A protein is red labelled while H3K27me3 is green labelled.

Thereafter, I analysed the distribution of H3K4me2 on horse **(Fig. 36-A)** and donkey **(Fig. 36-B)** metaphase chromosomes. No significant difference was observed between horse and donkey. The signal was variable among centromeres; however, most of centromeres showed a very faint (absent, in some cases) signal, apart from some centromeres showing a stronger H3K4me2 signal.

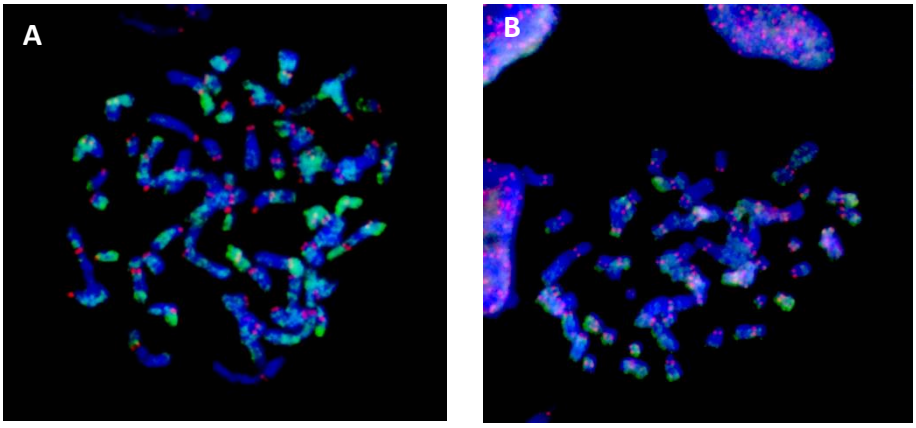


Figure 36 – Two colour immunofluorescence on horse (A) and donkey (B) metaphase chromosomes. Metaphase chromosomes are stained with DAPI, blue, CENP-A is red labelled and H3K4me2 is green labelled.

Two colour-immunofluorescence oh horse and donkey chromatin fibres showed that most of centromeres (8 out of 10 fibres in the horse and 7 out of 10 fibres in the donkey) had a very faint H3K4me2 signal **(Fig. 37)**, in both horse **(Fig. 37-A)** and donkey **(Fig. 37-B)**.

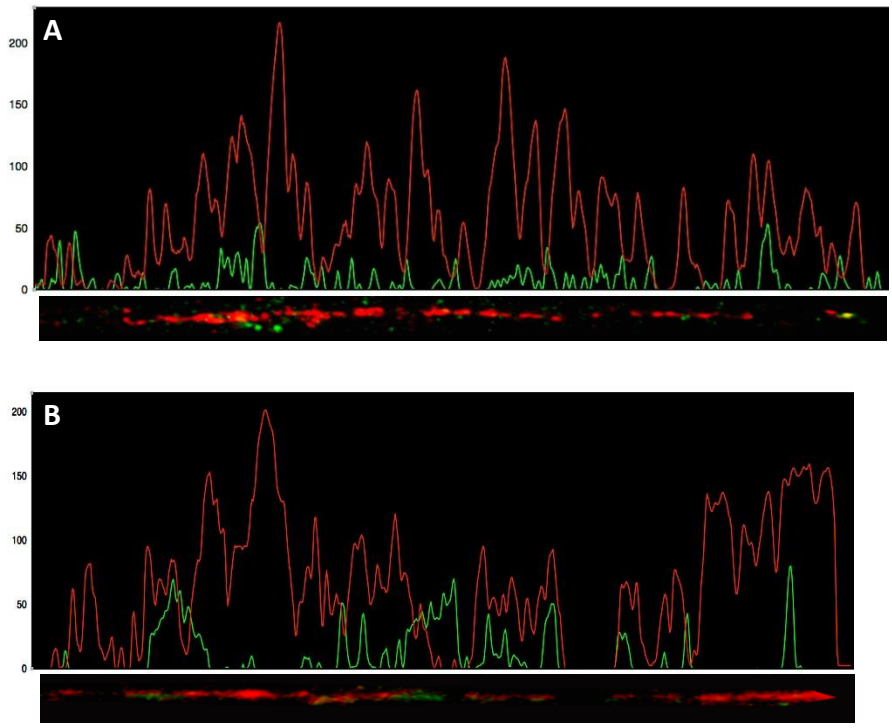


Figure 37 – Two colour immunofluorescence on horse (A) and donkey (B) chromatin fibres. The graph over the image of the fibre in A and B, represents the quantitative analysis (ImageJ software) of the intensity of the fluorescence signals.

In A and B, the CENP-A protein is red labelled while H3K4me2 is green labelled. The minority of fibres (2 out 10 of fibres in the horse and 3 out of 10 fibres in the donkey) showed a stronger H3K4me2 signal. In **Figure 38** examples of horse (Fig. 38-A) and donkey chromatin fibres (Fig. 38-B) with a strong H3K4me2 green fluorescence are reported.

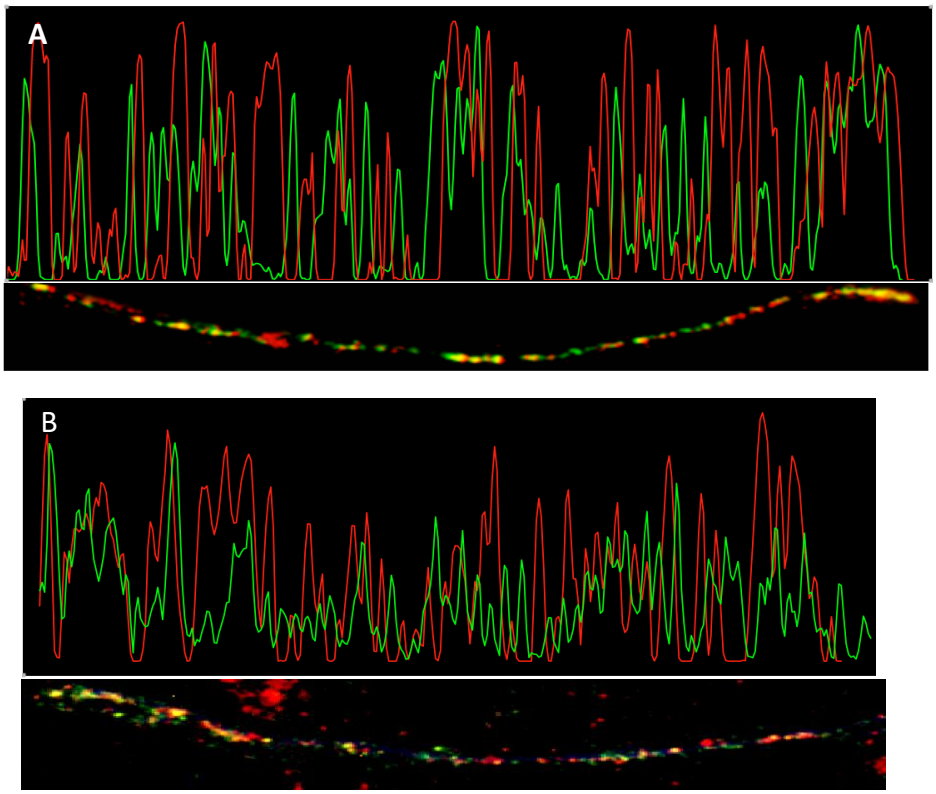


Figure 38 – Two colour immunofluorescence on horse (A) and donkey (B) chromatin fibres. The graph over the image of the fibre in A and B, reports the quantitative analysis (ImageJ software) of the intensity of the fluorescence signals. In A and B, the CENP-A protein is red labelled while H3K4me2 is green labelled.

7.3 H4K20me1 MARKER

The last histone modification analysed was the mono-methylation of lysine 20 of histone H4 (H4K20me1), which has been reported to be associated to CENP-A nucleosomes at the centromere functional core (Hori *et al*., 2014). This modification is supposed to be important for the maturation of CENP-A nucleosomes. In addition, in regions other than the centromere, it is associated with transcription (Beck *et al*., 2012).

7.3.1 Double immunofluorescence on metaphase chromosomes and chromatin fibres

Double-immunofluorescence experiments were performed on horse **(Fig. 39-A)** and donkey **(Fig. 39-B)** metaphase chromosomes using an antibody against H4K20me1 and an antibody against CENP-A to detect the centromere. The data demonstrated that both horse and donkey centromeres are enriched in this histone modification. No difference was noted between satellite-based and satellite-free centromeres, all the centromeric immuno-fluorescence signals having the same intensity.

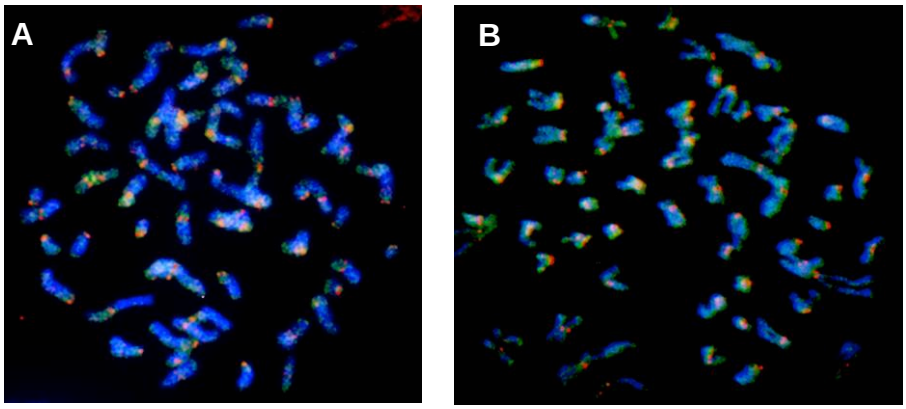


Figure 39 – Two colour immunofluorescence on horse (A) and donkey (B) metaphase chromosomes. Metaphase chromosomes are stained with DAPI, blue, CENP-A is red labelled and H4K20me1 is green labelled.

Double immunofluorescence on chromatin fibres confirmed the data obtained on metaphases. In addition, this experiment showed, as expected for a histone modification associated with CENP-A containing nucleosomes, a perfect co-localisation between CENP-A and H4K20me1 both in horse **(Fig. 40-A)** and in donkey chromatin fibres (**Fig. 40-B**). Moreover, as this marker is associated with transcriptionally competent chromatin, its presence at both satellite-free and satellite-based centromeres prompts to hypothesize that a partially opened, transcriptionally competent chromatin state is needed for centromere function

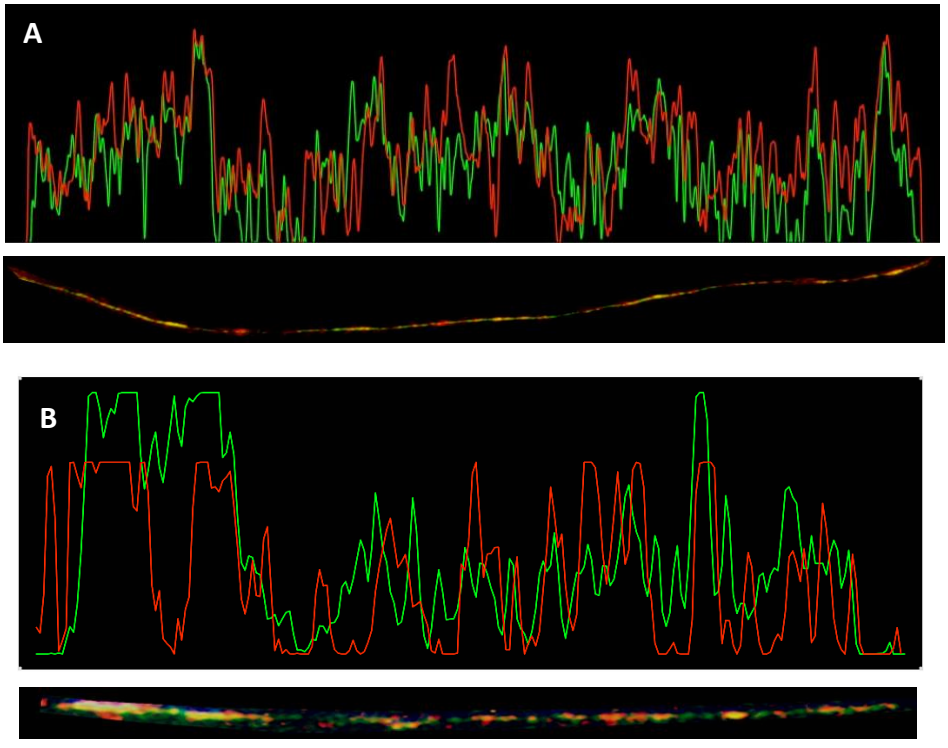


Figure 40 – Two colour immunofluorescence on horse (A) and donkey (B) chromatin fibres. The graph over the image of the fibre in A and B, reports the quantitative analysis (ImageJ software) of the intensity of the fluorescence signals. In A and B, the CENP-A protein is red labelled while H4K20me1 is green labelled.

8. DISCUSSION – PART 2

The lack of a universal centromere sequence and the existence of centromeres completely void of repetitive DNA (neocentromeres) suggest that the centromere identity is defined epigenetically (Sullivan and Karpen, 2004; Hall *et al*., 2012). The genus *Equus* is a powerful model for the study of the centromere because of the presence of a high number of evolutionarily new centromeres. These are satellite-DNA-free novel centromeres originated by centromere repositioning events that occurred during evolution in the absence of structural chromosome rearrangements (Carbone *et al*., 2006; Wade *et al*., 2009; Piras *et al*., 2010; Purgato *et al*., 2015; Nergadze *et al*., 2018). Interestingly, these satellite-free centromeres coexist in the same karyotype with the canonical satellite-based ones. This characteristic makes this an ideal model for the direct comparison of the molecular and epigenetic organization of satellite-free and satellite–based mammalian centromeres. However, the presence of highly repeated DNA sequences at satellite-based centromeres hampers their analysis using molecular and next generation sequencing approaches. Our analysis, based on the direct investigation of chromatin fibres on microscope slides by means of high-resolution cytogenetic approaches, is the method of election to make direct comparisons between satellite containing and satellite void centromeres present in the same karyotype. In addition, semi-quantitative estimates of the relative amount of specific chromatin markers can be performed.

I analysed, by double-immunofluorescence and immuno-FISH on horse and donkey metaphase chromosomes and chromatin fibres, five post-translational histone modifications: H3K9me3, H3K27me3, H3K9me2, H3K4me2 and H4K20me1.

H3K9me3

H3K9me3 is a marker of constitutive heterochromatin (Peters *et al*., 2003; Rice *et al*., 2003) and it was present both at horse and at donkey centromeres. Although all horse and donkey centromeres were immersed into a heterochromatic environment, one couple of horse chromosomes and 16 couples of donkey chromosomes displayed a very weak fluorescence signal. Since we previously demonstrated that the centromere of horse chromosome11 and 16 donkey centromere are satellite-less (Wade *et al*., 2009; Piras *et al*., 2010; Purgato *et al*., 2015; Nergadze *et al*., 2018), this result clearly suggested that the centromeres associated to a faint H3K9me3 fluorescence signal were the satellite-free ones.

The hypothesis that those centromeres showing a faint heterochromatic signal were the satellite-free ones was then confirmed by immuno-FISH experiments both on horse and donkey metaphases, using BAC probes specific for a satellite-free and a satellite-based centromere. The results showed a striking difference in the amount of the H3K9me3 modification, with the satellite-based centromeres showing a stronger heterochromatic signal with respect to the satellite-free ones.

To validate these results, I set up immuno-FISH on horse and donkey chromatin fibres; the results, while confirming the data obtained by immuno-FISH on metaphase chromosomes, added information about the architectural distribution of the H3K9me3 histone modification which was demonstrated to be present all along the CENP-A domain.

Interestingly, double immunofluorescence on metaphase chromosomes showed a peculiar centromere organization (Chapter 7. Paragraph 7.1.1, Fig. 21). In the horse and in the donkey, sister CENP-A spots faced the outside of the primary constriction, while the heterochromatic core, as identified by the presence of the H3K9me3 modification, was confined in the middle, just between the two CENP-A spots. This result perfectly reflects the hypothetical three-dimensional

arrangement of super-coiled centrochromatin. It has been hypothesized that centromeric chromatin supercoils to form a cylindrical structure, thus forcing the alignment of nucleosomes with the same composition: to promote proper kinetochore assembly during cell division, CENP-A nucleosomes must be exposed towards the outer face of the cylinder to be able to interact with spindle fibres (Fig. 41) (Blower *et al*., 2002).

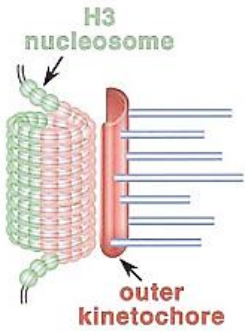


Figure 41 – Three-dimensional arrangement of super coiled centrochromatin. Blocks of CENP-A nucleosomes are pushed to the external face of the centrochromatin, to interact with the kinetochore inner plate, while H3-containing nucleosomes lie between the sister chromatids (Blower *et al*., 2002).

H3K9me2 and H3K27me3

The analysis of two markers of facultative heterochromatin (H3K9me2 and H3K27me3) (Nagaki *et al*., 2004; Sullivan and Karpen, 2004; Mravinac *et al*., 2005; Beiley *et al*., 2016) demonstrated that all centromeres were positive to the immunostaining, indicating that both satellite-less and satellite-based centromeres contain facultative heterochromatin. Contrary to constitutive heterochromatin, this kind of heterochromatin is prone to be decondensed and transcribed at occurrence.

H3K4me2

The results obtained by the analysis of the H3K4me2 histone modification, which is a marker of transcriptional competent chromatin (Schneider *et al*., 2004; Lam *et al*., 2006) on metaphase chromosomes demonstrated that this modification is absent or present in a small amount at horse and donkey centromeres. There was no evidence of any difference in the amount of this modification between satellite-based and satellite-free centromeres.

High resolution analysis of horse and donkey chromatin fibres showed variability in the H3K4me2 signal intensity among different fibres in both species, with about half of the fibres showing a strong histone modification signal and half showing a very low signal intensity. This variability could reflect a change in the centrochromatin epigenetic environment in different phases of the cell cycle. However, since a great variability in the intensity of fluorescence signals was evident also at metaphase, we speculate that the epigenetic state of single centromeres may vary among chromosomes and among cells. Indeed, it is plausible to hypothesize that a steady state amount of centromere transcripts is needed for proper centromere function and that this steady level can be maintained even if the transcriptional activity at single centromeres varies (Quénet *et al*., 2017). Literature data demonstrate that H3K4me2 and transcription are important for CENP-A recruitment and for the maintenance of centromere structure in mammals (Sullivan and Karpen, 2004). Centromere chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin (Bergmann *et al*., 2011). In view of its importance for centromere function, we plan to apply immuno-FISH to deepen the study of this chromatin modification, similarly to what we previously made for the analysis of constitutive heterochromatin.

H4K20me1

Finally, the analysis of the H4K20me1 histone modification on horse and donkey metaphase chromosomes demonstrated the presence of this histone modification at all centromeres in both species, this means that the presence of satellite DNA sequence at the centromere is not necessary for the recruitment of this histone modification. The analysis of chromatin fibres demonstrated that there is a perfect co-localization between the H4K20me1 and the CENP-A fluorescence signals. This result was expected since it has been demonstrated that this histone modification is present at CENP-A containing nucleosomes where it seems to be important for kinetochore assembly (Hori *et al*., 2014; Nishino *et al*., 2013).

9. CONCLUSION AND PERSPECTIVES – PART 2

The results obtained by the cytogenetic approach were compared with the results of ChIP-seq obtained by our collaborators in the Laboratory of Molecular Biology (Elena Giulotto, personal communication). The ChIP-seq approach allows to study, at the DNA sequence level, the organization of satellite–free centromeres, but is not suitable to analyse the organisation of centromeres containing highly repeated DNA sequences. On the contrary, the cytogenetic approach is less sensitive, but allows to characterize both satellite-free and satellite-based centromeres.

Table 1 summarises the results concerning those histone modifications that have been analysed by ChIP-seq and by molecular cytogenetics.

The cytogenetic approach allowed to demonstrate that there is a striking difference in the amount of constitutive heterochromatin, as marked by the H3K9me3 histone modification, between satellite-based centromeres and satellite-free centromeres. This result is relevant since suggests that centromeric single copy DNA, as well as the highly repetitive sequences, is able to condense in heterochromatin. However, the efficiency of this process is lower at satellite-free with respect to satellite-based centromeres. To our knowledge, presumably due to the absence of a suitable model system and to the use of molecular and next generation sequencing approaches, similar data have never been published before.

In addition, the cytogenetic approach showed that most of the centromeres in both horse and donkey metaphase chromosomes were slightly positive or negative for the H3K4me2 signal. However, a great variability was observed among chromosomes and this variation was apparently unrelated to the presence of satellite DNA. ChIP-seq results demonstrated that satellite–free centromeres do not contain this modification. On the other hand, high resolution analysis of chromatin fibres showed that, both in the horse and in the donkey, there are two kinds of organization: about half of the centromeres have a very low amount of this

histone modification, while the others have a much higher amount. The observed variability could be explained by considering that different fibres could represent centromeres in different phases of the cell cycle suggesting that the epigenetic environment at the centromeres can vary throughout cell cycle. However, since this modification is a marker of transcriptional competence, it can be supposed that the transcriptional activity of different centromere varies among cells and in different phases of the cell cycle, but the total amount of centromeric transcripts is maintained stable in the cell. It will be important to perform immuno-FISH experiments in order to evaluate the amount of this modification and its variability at single centromeres.

Finally, cytogenetics demonstrated that the H4K20me1 modification is present at all centromere. Consistently, ChIP-seq demonstrated the presence of this modification at satellite-free centromeres

Table 1 – Summary of the results obtained by Immunofluorescence, Immuno-FISH and ChIP-seq.

Histone modification	Functional significance	ChIP-seq satellite-free	Immuno-fluorescence satellite-free	Immuno-fluorescence satellite-based
H3K9me3	constitutive heterochromatin	+	-/+	++
H3K4me2	transcriptionally competent	-	-/+	-/+
H4K20me1	CENP-A associated transcriptionally competent	+	+	+

Altogether, the results of the second part of my PhD work demonstrated that, in the horse and in the donkey, the heterochromatic environment defined by the H3K9me3 histone mark is maintained both at satellite-free and at satellite-based centromeres, even though, as demonstrated by high resolution immuno-FISH experiments, in a lower amount at the satellite-free centromeres. It can be hypothesized that heterochromatic marks are a prerequisite for the recruitment of CENP-A in

ectopic sites, thus triggering neocentromere seeding. Literature data (Lam *et al.*, 2006; Olszak *et al*., 2011) suggest a functional role of heterochromatin which could define the borders of the centromeric functional domain and prevent CENP-A chromatin spreading in adjacent regions (Maggert and Karpen, 2000; Maggert and Karpen 2001). Previous works (Purgato *et al*., 2015; Nergadze *et al*., 2018), revealed that the satellite-free centromeres in horse and donkey are not fixed, but they can slide inside a region of about 500 kb. This evidence supports the hypothesis that the recruitment of satellite DNA during the process of centromere maturation, may reduce the flexibility of centromere position, stabilizing it. Therefore, the low amount of the constitutive heterochromatic marker H3K9me3 at the satellite-free centromeres, not being enough to limit the CENP-A chromatin spreading, could explain the flexibility in the position of the CENP-A domain.

While a transcriptionally competent heterochromatic environment (H3K4me2) seems to be little represented at horse and donkey metaphase chromosomes, a high variability in the amount of this histone modification characterises different horse and donkey fibres, suggesting that the epigenetic environment at the centromeres can vary during the cell cycle. In addition, all centromeres contain facultative heterochromatin prone to be opened, as confirmed by the presence of the H3K9me2 and H3K27me3 histone modifications. This result is not surprising since a relaxed chromatin environment is known to be essential for CENP-A loading (Bergmann *et al.*, 2012).

Finally, the presence at satellite-free and at satellite-based centromeres of the H4K20me1 marker, as confirmed by both approaches, is in agreement with literature data which demonstrated that this histone modification, ensuring CENP-A containing nucleosome maturation, promotes kinetochore assembly.

Our data combined with ChIP-seq results demonstrate the presence of a relaxed chromatin environment at satellite-free centromeres. We previously demonstrated that satellite DNA is transcribed at horse centromeres (Cerutti *et al*., 2016), but we have no evidence of the existence of transcripts derived from satellite-free centromeric sequences (Elena

Giulotto, personal communication). It is tempting to speculate that the permissive state of satellite-free centrochromatin makes it able to interact with trans-acting RNAs which derive from satellite-based centromeres. In the next future, we are planning to characterize centromere transcripts and to set up RNA-FISH experiments in order to directly determine their ability to interact with satellite-free centromeres.

Finally, it will be worth investigating, by means of cell synchronization, the variability in the amount of transcriptional competent chromatin markers in different phases of the cell cycle.

10. REFERENCES

Aldrup-Macdonald, M.E., and Sullivan, B.A. (2014). The past, present, and future of human centromere genomics. Genes (Basel) *5*, 33-50.

Allshire, R.C., and Karpen, G.H. (2008). Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nat Rev Genet *9*, 923-937.

Almeida, A., Kokalj-Vokac, N., Lefrancois, D., Viegas-Pequignot, E., Jeanpierre, M., Dutrillaux, B., and Malfoy, B. (1993). Hypomethylation of classical satellite DNA and chromosome instability in lymphoblastoid cell lines. Human genetics *91*, 538-546.

Alonso, A., Hasson, D., Cheung, F., and Warburton, P.E. (2010). A paucity of heterochromatin at functional human neocentromeres. Epigenetics Chromatin *3*, 6.

Ambrozova, K., Mandakova, T., Bures, P., Neumann, P., Leitch, I.J., Koblizkova, A., Macas, J., and Lysak, M.A. (2011). Diverse retrotransposon families and an AT-rich satellite DNA revealed in giant genomes of Fritillaria lilies. Ann Bot *107*, 255-268.

Amor, D.J., Bentley, K., Ryan, J., Perry, J., Wong, L., Slater, H., and Choo, K.H. (2004). Human centromere repositioning "in progress". Proc Natl Acad Sci U S A *101*, 6542-6547.

Amor, D.J., and Choo, K.H. (2002). Neocentromeres: role in human disease, evolution, and centromere study. Am J Hum Genet *71*, 695-714.

Amor, D.J., Voullaire, L., Bentley, K., Savarirayan, R., and Choo, K.H. (2005). Mosaic monosomy of a neocentric ring chromosome maps brachyphalangy and growth hormone deficiency to 13q31.1-13q32.3. Am J Med Genet A *133a*, 151-157.

Bailey, A.O., Panchenko, T., Shabanowitz, J., Lehman, S.M., Bai, D.L., Hunt, D.F., Black, B.E., and Foltz, D.R. (2016). Identification of the Post-translational Modifications Present in Centromeric Chromatin. Mol Cell Proteomics *15*, 918-931.

Bao, W., Zhang, W., Yang, Q., Zhang, Y., Han, B., Gu, M., Xue, Y., and Cheng, Z. (2006). Diversity of centromeric repeats in two closely related wild rice species, Oryza officinalis and Oryzarhizomatis. Molecular Genetics and Genomics *275*, 421-430.

Barisic, M., and Geley, S. (2011). Spindly switch controls anaphase: spindly and RZZ functions in chromosome attachment and mitotic checkpoint control. Cell cycle *10*, 449-456.

Barnhart, M.C., Kuich, P.H., Stellfox, M.E., Ward, J.A., Bassett, E.A., Black, B.E., and Foltz, D.R. (2011). HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. J Cell Biol *194*, 229-243.

Barra, V., and Fachinetti, D. (2018). The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. Nat Commun *9*, 4340.

Baum, M., Sanyal, K., Mishra, P.K., Thaler, N., and Carbon, J. (2006). Formation of functional centromeric chromatin is specified epigenetically in Candida albicans. Proc Natl Acad Sci U S A *103*, 14877-14882.

Beck, D.B., Oda, H., Shen, S.S., and Reinberg, D. (2012). PR-Set7 and H4K20me1: at the crossroads of genome integrity, cell cycle, chromosome condensation, and transcription. Genes Dev *26*, 325-337.

Bergmann, J.H., Jakubsche, J.N., Martins, N.M., Kagansky, A., Nakano, M., Kimura, H., Kelly, D.A., Turner, B.M., Masumoto, H., Larionov, V.*, et al.* (2012a). Epigenetic engineering: histone H3K9 acetylation is compatible with kinetochore structure and function. J Cell Sci *125*, 411-421.

Bergmann, J.H., Martins, N.M., Larionov, V., Masumoto, H., and Earnshaw, W.C. (2012b). HACking the centromere chromatin code: insights from human artificial chromosomes. Chromosome Res *20*, 505-519.

Bergmann, J.H., Rodriguez, M.G., Martins, N.M., Kimura, H., Kelly, D.A., Masumoto, H., Larionov, V., Jansen, L.E., and Earnshaw, W.C. (2011). Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. Embo j *30*, 328-340.

Bernard, P., Maure, J.F., Partridge, J.F., Genier, S., Javerzat, J.P., and Allshire, R.C. (2001). Requirement of heterochromatin for cohesion at centromeres. Science *294*, 2539-2542.

Bernardino, J., Roux, C., Almeida, A., Vogt, N., Gibaud, A., Gerbault-Seureau, M., Magdelenat, H., Bourgeois, C.A., Malfoy, B., and Dutrillaux, B. (1997). DNA hypomethylation in breast cancer: an independent parameter of tumor progression? Cancer Genet Cytogenet *97*, 83-89.

Bersani, F., Lee, E., Kharchenko, P.V., Xu, A.W., Liu, M., Xega, K., MacKenzie, O.C., Brannigan, B.W., Wittner, B.S., Jung, H.*, et al.* (2015). Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. Proc Natl Acad Sci U S A *112*, 15148-15153.

Bierhoff, H., Postepska-Igielska, A., and Grummt, I. (2014). Noisy silence: non-coding RNA and heterochromatin formation at repetitive elements. Epigenetics *9*, 53-61.

Biscotti, M.A., Canapa, A., Forconi, M., Olmo, E., and Barucca, M. (2015). Transcription of tandemly repetitive DNA: functional roles. Chromosome Res *23*, 463-477.

Black, B.E., Brock, M.A., Bédard, S., Woods, V.L., and Cleveland, D.W. (2007a). An epigenetic mark generated by the incorporation of CENP-A into centromeric nucleosomes. Proceedings of the National Academy of Sciences *104*, 5008-5013.

Black, B.E., Foltz, D.R., Chakravarthy, S., Luger, K., Woods, V.L., Jr., and Cleveland, D.W. (2004). Structural determinants for generating centromeric chromatin. Nature *430*, 578-582.

Black, B.E., Jansen, L.E., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J.V., and Cleveland, D.W. (2007b). Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. Mol Cell *25*, 309-322.

Blattes, R., Monod, C., Susbielle, G., Cuvier, O., Wu, J.H., Hsieh, T.S., Laemmli, U.K., and Kas, E. (2006). Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. Embo j *25*, 2397-2408.

Blennow, E., Telenius, H., de Vos, D., Larsson, C., Henriksson, P., Johansson, O., Carter, N.P., and Nordenskjold, M. (1994). Tetrasomy 15q: two marker chromosomes with no detectable alpha-satellite DNA. Am J Hum Genet *54*, 877-883.

Blom, E., Heyning, F.H., and Kroes, W.G. (2010). A case of angioimmunoblastic T-cell non-Hodgkin lymphoma with a neocentric inv dup(1). Cancer Genet Cytogenet *202*, 38-42.

Blower, M.D. (2016). Centromeric Transcription Regulates Aurora-B Localization and Activation. Cell Rep *15*, 1624-1633.

Blower, M.D., Sullivan, B.A., and Karpen, G.H. (2002). Conserved organization of centromeric chromatin in flies and humans. Dev Cell *2*, 319-330.

Bodor, D.L., Mata, J.F., Sergeev, M., David, A.F., Salimian, K.J., Panchenko, T., Cleveland, D.W., Black, B.E., Shah, J.V., and Jansen, L.E. (2014). The quantitative architecture of centromeric chromatin. Elife *3*, e02137.

Bouzinba-Segard, H., Guais, A., and Francastel, C. (2006). Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. Proc Natl Acad Sci U S A *103*, 8709-8714.

Brown, J.D., Mitchell, S.E., and O'Neill, R.J. (2012). Making a long story short: noncoding RNAs and chromosome change. Heredity (Edinb) *108*, 42-49.

Burnside, R.D., Ibrahim, J., Flora, C., Schwartz, S., Tepperberg, J.H., Papenhausen, P.R., and Warburton, P.E. (2011). Interstitial deletion of proximal 8q including part of the centromere from unbalanced segregation of a paternal deletion/marker karyotype with neocentromere formation at 8p22. Cytogenet Genome Res *132*, 227-232.

Burrack, L.S., and Berman, J. (2012). Neocentromeres and epigenetically inherited features of centromeres. Chromosome Res *20*, 607-619.

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

Capozzi, O., Purgato, S., D'Addabbo, P., Archidiacono, N., Battaglia, P., Baroncini, A., Capucci, A., Stanyon, R., Della Valle, G., and Rocchi, M. (2009). Evolutionary descent of a human chromosome 6 neocentromere: a jump back to 17 million years ago. Genome Res *19*, 778-784.

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.* (2006). Evolutionary movement of centromeres in horse, donkey, and zebra. Genomics *87*, 777-782.

Cardone, M.F., Alonso, A., Pazienza, M., Ventura, M., Montemurro, G., Carbone, L., de Jong, P.J., Stanyon, R., D'Addabbo, P., Archidiacono, N.*, et al.* (2006). Independent centromere formation in a capricious, gene-free domain of chromosome 13q21 in Old World monkeys and pigs. Genome Biol *7*, R91.

Carmena, M., Wheelock, M., Funabiki, H., and Earnshaw, W.C. (2012). The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. Nat Rev Mol Cell Biol *13*, 789-803.

Carone, D.M., Longo, M.S., Ferreri, G.C., Hall, L., Harris, M., Shook, N., Bulazel, K.V., Carone, B.R., Obergfell, C., O'Neill, M.J.*, et al.* (2009). A new class of retroviral and satellite encoded small RNAs emanates from mammalian centromeres. Chromosoma *118*, 113-125.

Carroll, C.W., Milks, K.J., and Straight, A.F. (2010). Dual recognition of CENP-A nucleosomes is required for centromere assembly. J Cell Biol *189*, 1143-1155.

Cerutti, F., Gamba, R., Mazzagatti, A., Piras, F.M., Cappelletti, E., Belloni, E., Nergadze, S.G., Raimondi, E., and Giulotto, E. (2016). The major horse satellite DNA family is associated with centromere competence. Mol Cytogenet *9*, 35.

Chan, F.L., Marshall, O.J., Saffery, R., Kim, B.W., Earle, E., Choo, K.H., and Wong, L.H. (2012). Active transcription and essential role of RNA polymerase II at the centromere during mitosis. Proc Natl Acad Sci U S A *109*, 1979-1984.

Chan, F.L., and Wong, L.H. (2012). Transcription in the maintenance of centromere chromatin identity. Nucleic Acids Res *40*, 11178-11188.

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.* (2019). Islands of retroelements are major components of Drosophila centromeres. PLoS Biol *17*, e3000241.

Cheeseman, I.M., and Desai, A. (2008). Molecular architecture of the kinetochore-microtubule interface. Nat Rev Mol Cell Biol *9*, 33-46.

Chen, R.Z., Pettersson, U., Beard, C., Jackson-Grusby, L., and Jaenisch, R. (1998). DNA hypomethylation leads to elevated mutation rates. Nature *395*, 89-93.

Choi, E.S., Stralfors, A., Castillo, A.G., Durand-Dubief, M., Ekwall, K., and Allshire, R.C. (2011). Identification of noncoding transcripts from within CENP-A chromatin at fission yeast centromeres. J Biol Chem *286*, 23600-23607.

Choi, E.S., Stralfors, A., Catania, S., Castillo, A.G., Svensson, J.P., Pidoux, A.L., Ekwall, K., and Allshire, R.C. (2012). Factors that promote H3 chromatin integrity during transcription prevent promiscuous deposition of CENP-A(Cnp1) in fission yeast. PLoS Genet *8*, e1002985.

Choo, K.H. (2000). Centromerization. Trends Cell Biol *10*, 182-188.

Chueh, A.C., Northrop, E.L., Brettingham-Moore, K.H., Choo, K.H., and Wong, L.H. (2009). LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. PLoS Genet *5*, e1000354.

Chueh, A.C., Wong, L.H., Wong, N., and Choo, K.H. (2005). Variable and hierarchical size distribution of L1-retroelement-enriched CENP-A clusters within a functional human neocentromere. Hum Mol Genet *14*, 85-93.

Cizkova, J., Hribova, E., Humplikova, L., Christelova, P., Suchankova, P., and Dolezel, J. (2013). Molecular analysis and genomic organization of major DNA satellites in banana (Musa spp.). PLoS One *8*, e54808.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Comings, D.E., and Okada, T.A. (1971). Fine structure of kinetochore in Indian muntjac. Exp Cell Res *67*, 97-110.

Cooke, C.A., Bazett-Jones, D.P., Earnshaw, W.C., and Rattner, J.B. (1993). Mapping DNA within the mammalian kinetochore. J Cell Biol *120*, 1083-1091.

Cooke, C.A., Schaar, B., Yen, T.J., and Earnshaw, W.C. (1997). Localization of CENP-E in the fibrous corona and outer plate of mammalian kinetochores from prometaphase through anaphase. Chromosoma *106*, 446-455.

Dawe, R.K., and Henikoff, S. (2006). Centromeres put epigenetics in the driver's seat. Trends Biochem Sci *31*, 662-669.

de Figueiredo, A.F., Mkrtchyan, H., Liehr, T., Soares Ventura, E.M., de Jesus Marques-Salles, T., Santos, N., Ribeiro, R.C., Abdelhay, E., and Macedo Silva, M.L. (2009). A case of childhood acute myeloid leukemia AML (M5) with a neocentric chromosome neo(1)(qter-->q23 approximately 24::q23 approximately 24-->q43-->neo-->q43-->qter) and tetrasomy of chromosomes 8 and 21. Cancer Genet Cytogenet *193*, 123-126.

de la Herran, R., Robles, F., Cunado, N., Santos, J.L., Ruiz Rejon, M., Garrido-Ramos, M.A., and Ruiz Rejon, C. (2001). A heterochromatic satellite DNA is highly amplified in a single chromosome of Muscari (Hyacinthaceae). Chromosoma *110*, 197-202.

Denegri, M., Moralli, D., Rocchi, M., Biggiogera, M., Raimondi, E., Cobianchi, F., De Carli, L., Riva, S., and Biamonti, G. (2002). Human chromosomes 9, 12, and 15 contain the nucleation sites of stress-induced nuclear bodies. Mol Biol Cell *13*, 2069-2079.

Drinnenberg, I.A.; Henikoff, S.; Malik, H.S. (2016) Evolutionary Turnover of Kinetochore Proteins: A Ship of Theseus? Trends Cell. Biol., *26,* 498–510.

du Sart, D., Cancilla, M.R., Earle, E., Mao, J.I., Saffery, R., Tainton, K.M., Kalitsis, P., Martyn, J., Barry, A.E., and Choo, K.H. (1997). A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA. Nat Genet *16*, 144-153.

Earnshaw, W.C., and Rothfield, N. (1985). Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. Chromosoma *91*, 313-321.

Earnshaw, W.C., Sullivan, K.F., Machlin, P.S., Cooke, C.A., Kaiser, D.A., Pollard, T.D., Rothfield, N.F., and Cleveland, D.W. (1987). Molecular cloning of cDNA for CENP-B, the major human centromere autoantigen. J Cell Biol *104*, 817-829.

Eder, V., Ventura, M., Ianigro, M., Teti, M., Rocchi, M., and Archidiacono, N. (2003). Chromosome 6 phylogeny in primates and centromere repositioning. Mol Biol Evol *20*, 1506-1512.

Emadzade, K., Jang, T.S., Macas, J., Kovarik, A., Novak, P., Parker, J., and Weiss-Schneeweiss, H. (2014). Differential amplification of satellite PaB6 in chromosomally hypervariable Prospero autumnale complex (Hyacinthaceae). Ann Bot *114*, 1597-1608.

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.* (2009). A transcriptomic analysis of human centromeric and pericentric sequences in normal and tumor cells. Nucleic Acids Res *37*, 6340-6354.

Faas, B.H., Cirigliano, V., and Bui, T.H. (2011). Rapid methods for targeted prenatal diagnosis of common chromosome aneuploidies. Semin Fetal Neonatal Med *16*, 81-87.

Fachinetti, D.; Han, J.S.; McMahon, M.A.; Ly, P.; Abdullah, A.; Wong, A.J.; Cleveland, D.W. (2015). DNA Sequence-Specific Binding of CENP-B Enhances the Fidelity of Human Centromere Function. Dev. Cell *33,* 314–327

Fenech, M. (2007). Cytokinesis-block micronucleus cytome assay. Nat Protoc *2*, 1084-1104.

Fenech, M., and Morley, A.A. (1985). Measurement of micronuclei in lymphocytes. Mutat Res *147*, 29-36.

Ferreri, G.C., Liscinsky, D.M., Mack, J.A., Eldridge, M.D., and O'Neill, R.J. (2005). Retention of latent centromeres in the Mammalian genome. J Hered *96*, 217-224.

Ferri, F., Bouzinba-Segard, H., Velasco, G., Hube, F., and Francastel, C. (2009). Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. Nucleic acids research *37*, 5071-5080.

Fischle, W., Wang, Y., and Allis, C.D. (2003). Binary switches and modification cassettes in histone biology and beyond. Nature *425*, 475-479.

Flemming, W. (1882). Zellsubstanz, kern und zelltheilung (Vogel).

Folco, H.D., Pidoux, A.L., Urano, T., and Allshire, R.C. (2008). Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. Science *319*, 94-97.

Foltz, D.R., Jansen, L.E., Black, B.E., Bailey, A.O., Yates, J.R., 3rd, and Cleveland, D.W. (2006). The human CENP-A centromeric nucleosome-associated complex. Nat Cell Biol *8*, 458-469.

Fukagawa, T., and Earnshaw, W.C. (2014). The centromere: chromatin foundation for the kinetochore machinery. Dev Cell *30*, 496-508.

Fukagawa, T., Mikami, Y., Nishihashi, A., Regnier, V., Haraguchi, T., Hiraoka, Y., Sugata, N., Todokoro, K., Brown, W., and Ikemura, T. (2001). CENP-H, a constitutive centromere component, is required for centromere targeting of CENP-C in vertebrate cells. Embo j *20*, 4603-4617.

Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T., and Oshimura, M. (2004). Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nat Cell Biol *6*, 784-791.

Fuks, F., Hurd, P.J., Deplus, R., and Kouzarides, T. (2003). The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Res *31*, 2305-2312.

Gandhi, R., Bonaccorsi, S., Wentworth, D., Doxsey, S., Gatti, M., and Pereira, A. (2004). The Drosophila kinesin-like protein KLP67A is essential for mitotic and male meiotic spindle assembly. Molecular biology of the cell *15*, 121-131.

Gardner, R.D., Poddar, A., Yellman, C., Tavormina, P.A., Monteagudo, M.C., and Burke, D.J. (2001). The spindle checkpoint of the yeast Saccharomyces cerevisiae requires kinetochore function and maps to the CBF3 domain. Genetics *157*, 1493-1502.

Garrido-Ramos, M.A. (2015). Satellite DNA in Plants: More than Just Rubbish. Cytogenet Genome Res *146*, 153-170.

Garrido-Ramos, M.A., Jamilena, M., Lozano, R., Ruiz Rejon, C., and Ruiz Rejon, M. (1994). Cloning and characterization of a fish centromeric satellite DNA. Cytogenet Cell Genet *65*, 233-237.

Geigl EM, B.-D.S., Beja Pereira A, Cothran EG, Hrabar H, Oysunsuen T, Pruvost M. (2016). Genetics and paleogenetics of equids, In Wild equids.

Gent, J.I., and Dawe, R.K. (2012). RNA as a structural and regulatory component of the centromere. Annu Rev Genet *46*, 443-453.

Gong, Z., Wu, Y., Koblizkova, A., Torres, G.A., Wang, K., Iovene, M., Neumann, P., Zhang, W., Novak, P., Buell, C.R.*, et al.* (2012). Repeatless and repeat-based centromeres in potato: implications for centromere evolution. Plant Cell *24*, 3559-3574.

Gong, Z., Yu, H., Huang, J., Yi, C., and Gu, M. (2009). Unstable transmission of rice chromosomes without functional centromeric repeats in asexual propagation. Chromosome Res *17*, 863-872.

Gopalakrishnan, S., Sullivan, B.A., Trazzi, S., Della Valle, G., and Robertson, K.D. (2009). DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions. Hum Mol Genet *18*, 3178-3193.

Goshima, G., Kiyomitsu, T., Yoda, K., and Yanagida, M. (2003). Human centromere chromatin protein hMis12, essential for equal segregation, is independent of CENP-A loading pathway. J Cell Biol *160*, 25-39.

Goutte-Gattat, D., Shuaib, M., Ouararhni, K., Gautier, T., Skoufias, D.A., Hamiche, A., and Dimitrov, S. (2013). Phosphorylation of the CENP-A amino-terminus in mitotic centromeric chromatin is required for kinetochore function. Proc Natl Acad Sci U S A *110*, 8579-8584.

Grady, D.L., Ratliff, R.L., Robinson, D.L., McCanlies, E.C., Meyne, J., and Moyzis, R.K. (1992). Highly conserved repetitive DNA sequences are present at human centromeres. Proc Natl Acad Sci U S A *89*, 1695-1699.

Grimes, B.R., Rhoades, A.A., and Willard, H.F. (2002). Alpha-satellite DNA and vector composition influence rates of human artificial chromosome formation. Mol Ther *5*, 798-805.

Guenatri, M., Bailly, D., Maison, C., and Almouzni, G. (2004). Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. J Cell Biol *166*, 493-505.

Guse, A., Carroll, C.W., Moree, B., Fuller, C.J., and Straight, A.F. (2011). In vitro centromere and kinetochore assembly on defined chromatin templates. Nature *477*, 354-358.

Haaf, T., Warburton, P.E., and Willard, H.F. (1992). Integration of human alpha-satellite DNA into simian chromosomes: centromere protein binding and disruption of normal chromosome segregation. Cell *70*, 681-696.

Hahnenberger, K.M., Carbon, J., and Clarke, L. (1991). Identification of DNA regions required for mitotic and meiotic functions within the centromere of Schizosaccharomyces pombe chromosome I. Mol Cell Biol *11*, 2206-2215.

Hall, I.M., Noma, K., and Grewal, S.I. (2003). RNA interference machinery regulates chromosome dynamics during mitosis and meiosis in fission yeast. Proc Natl Acad Sci U S A *100*, 193-198.

Hall, L.E., Mitchell, S.E., and O'Neill, R.J. (2012). Pericentric and centromeric transcription: a perfect balance required. Chromosome Res *20*, 535-546.

Hansen, R.S., Wijmenga, C., Luo, P., Stanek, A.M., Canfield, T.K., Weemaes, C.M., and Gartler, S.M. (1999). The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. Proc Natl Acad Sci U S A *96*, 14412-14417.

Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., and Willard, H.F. (1997). Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nat Genet *15*, 345-355.

Hemmat, M., Wang, B.T., Warburton, P.E., Yang, X., Boyar, F.Z., El Naggar, M., and Anguiano, A. (2012). Neocentric X-chromosome in a girl with Turner-like syndrome. Mol Cytogenet *5*, 29.

Hemmerich, P., Stoyan, T., Wieland, G., Koch, M., Lechner, J., and Diekmann, S. (2000). Interaction of yeast kinetochore proteins with centromere-protein/transcription factor Cbf1. Proc Natl Acad Sci U S A *97*, 12583-12588.

Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. Science *293*, 1098-1102.

Henikoff, S., and Dalal, Y. (2005). Centromeric chromatin: what makes it unique? Curr Opin Genet Dev *15*, 177-184.

Henikoff, S., and Malik, H.S. (2002). Centromeres: selfish drivers. Nature *417*, 227.

Hindriksen, S., Lens, S.M.A., and Hadders, M.A. (2017). The Ins and Outs of Aurora B Inner Centromere Localization. Front Cell Dev Biol *5*, 112.

Hinshaw, S.M., and Harrison, S.C. (2013). An Iml3-Chl4 heterodimer links the core centromere to factors required for accurate chromosome segregation. Cell Rep *5*, 29-36.

Hori, T., Amano, M., Suzuki, A., Backer, C.B., Welburn, J.P., Dong, Y., McEwen, B.F., Shang, W.H., Suzuki, E., Okawa, K.*, et al.* (2008a). CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. Cell *135*, 1039-1052.

Hori, T., Okada, M., Maenaka, K., and Fukagawa, T. (2008b). CENP-O class proteins form a stable complex and are required for proper kinetochore function. Molecular biology of the cell *19*, 843-854.

Hori, T., Shang, W.H., Toyoda, A., Misu, S., Monma, N., Ikeo, K., Molina, O., Vargiu, G., Fujiyama, A., Kimura, H.*, et al.* (2014). Histone H4 Lys 20 monomethylation of the CENP-A nucleosome is essential for kinetochore assembly. Dev Cell *29*, 740-749.

Hornung, P., Troc, P., Malvezzi, F., Maier, M., Demianova, Z., Zimniak, T., Litos, G., Lampert, F., Schleiffer, A., Brunner, M.*, et al.* (2014). A cooperative mechanism drives budding yeast kinetochore assembly downstream of CENP-A. J Cell Biol *206*, 509-524.

Hribova, E., Neumann, P., Matsumoto, T., Roux, N., Macas, J., and Dolezel, J. (2010). Repetitive part of the banana (Musa acuminata) genome investigated by low-depth 454 sequencing. BMC Plant Biol *10*, 204.

Hsieh, C.L. (1997). Stability of patch methylation and its impact in regions of transcriptional initiation and elongation. Mol Cell Biol *17*, 5897-5904.

Hsieh, C.L., Lin, C.L., Liu, H., Chang, Y.J., Shih, C.J., Zhong, C.Z., Lee, S.C., and Tan, B.C. (2011). WDHD1 modulates the post-transcriptional step of the centromeric silencing pathway. Nucleic Acids Res *39*, 4048-4062.

Hyman, A.A., and Sorger, P.K. (1995). Structure and function of kinetochores in budding yeast. Annu Rev Cell Dev Biol *11*, 471-495.

Ideue, T., Cho, Y., Nishimura, K., and Tani, T. (2014). Involvement of satellite I noncoding RNA in regulation of chromosome segregation. Genes Cells *19*, 528-538.

Ikeno, M., Grimes, B., Okazaki, T., Nakano, M., Saitoh, K., Hoshino, H., McGill, N.I., Cooke, H., and Masumoto, H. (1998). Construction of YAC-based mammalian artificial chromosomes. Nat Biotechnol *16*, 431-439.

IJdo J,W., Baldini, A., Ward, D.C., Reeders, S.T., and Wells, R.A. (1991). Origin of human chromosome 2: an ancestral telomere-telomere fusion. Proc Natl Acad Sci U S A *88*, 9051-9055.

Ishii, K., Ogiyama, Y., Chikashige, Y., Soejima, S., Masuda, F., Kakuma, T., Hiraoka, Y., and Takahashi, K. (2008). Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. Science *321*, 1088-1091.

Italiano, A., Attias, R., Aurias, A., Perot, G., Burel-Vandenbos, F., Otto, J., Venissac, N., and Pedeutour, F. (2006). Molecular cytogenetic

characterization of a metastatic lung sarcomatoid carcinoma: 9p23 neocentromere and 9p23-p24 amplification including JAK2 and JMJD2C. Cancer Genet Cytogenet *167*, 122-130.

Italiano, A., Maire, G., Sirvent, N., Nuin, P.A., Keslair, F., Foa, C., Louis, C., Aurias, A., and Pedeutour, F. (2009). Variability of origin for the neocentromeric sequences in analphoid supernumerary marker chromosomes of well-differentiated liposarcomas. Cancer Lett *273*, 323-330.

Izuta, H., Ikeno, M., Suzuki, N., Tomonaga, T., Nozaki, N., Obuse, C., Kisu, Y., Goshima, N., Nomura, F., Nomura, N.*, et al.* (2006). Comprehensive analysis of the ICEN (Interphase Centromere Complex) components enriched in the CENP-A chromatin of human cells. Genes Cells *11*, 673-684.

Jackson, J.P., Lindroth, A.M., Cao, X., and Jacobsen, S.E. (2002). Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature *416*, 556-560.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Jin, W., Lamb, J.C., Vega, J.M., Dawe, R.K., Birchler, J.A., and Jiang, J. (2005). Molecular and functional dissection of the maize B chromosome centromere. Plant Cell *17*, 1412-1423.

Jokelainen, P.T. (1967). The ultrastructure and spatial organization of the metaphase kinetochore in mitotic rat cells. J Ultrastruct Res *19*, 19-44.

Jones, K.W. (1970). Chromosomal and nuclear location of mouse satellite DNA in individual cells. Nature *225*, 912-915.

Jonsson, H., Schubert, M., Seguin-Orlando, A., Ginolhac, A., Petersen, L., Fumagalli, M., Albrechtsen, A., Petersen, B., Korneliussen, T.S., Vilstrup, J.T.*, et al.* (2014). Speciation with gene flow in equids despite extensive chromosomal plasticity. Proc Natl Acad Sci U S A *111*, 18655-18660.

Kalitsis, P., and Choo, K.H. (2012). The evolutionary life cycle of the resilient centromere. Chromosoma *121*, 327-340.

Kaslow, D.C., and Migeon, B.R. (1987). DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation. Proceedings of the National Academy of Sciences *84*, 6210-6214.

Kato, H., Jiang, J., Zhou, B.R., Rozendaal, M., Feng, H., Ghirlando, R., Xiao, T.S., Straight, A.F., and Bai, Y. (2013). A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. Science *340*, 1110-1113.

Kato, T., Sato, N., Hayama, S., Yamabuki, T., Ito, T., Miyamoto, M., Kondo, S., Nakamura, Y., and Daigo, Y. (2007). Activation of Holliday junction recognizing protein involved in the chromosomal stability and immortality of cancer cells. Cancer Res *67*, 8544-8553.

Ketel, C., Wang, H.S., McClellan, M., Bouchonville, K., Selmecki, A., Lahav, T., Gerami-Nejad, M., and Berman, J. (2009). Neocentromeres form efficiently at multiple possible loci in Candida albicans. PLoS Genet *5*, e1000400.

Kirsch-Volders, M., Plas, G., Elhajouji, A., Lukamowicz, M., Gonzalez, L., Vande Loock, K., and Decordier, I. (2011). The in vitro MN assay in 2011: origin and fate, biological significance, protocols, high throughput methodologies and toxicological relevance. Arch Toxicol *85*, 873-899.

Kit, S. (1961). Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. J Mol Biol *3*, 711-716.

Klein, E., Rocchi, M., Ovens-Raeder, A., Kosyakova, N., Weise, A., Ziegler, M., Meins, M., Morlot, S., Fischer, W., Volleth, M.*, et al.* (2012). Five novel locations of Neocentromeres in human: 18q22.1, Xq27.1 approximately 27.2, Acro p13, Acro p12, and heterochromatin of unknown origin. Cytogenet Genome Res *136*, 163-166.

Kobayashi, T., Yamada, F., Hashimoto, T., Abe, S., Matsuda, Y., and Kuroiwa, A. (2008). Centromere repositioning in the X chromosome of XO/XO mammals, Ryukyu spiny rat. Chromosome Res *16*, 587-593.

Koch, J. (2000). Neocentromeres and alpha satellite: a proposed structural code for functional human centromere DNA. Hum Mol Genet *9*, 149-154.

Koo, D.-H., Han, F., Birchler, J.A., and Jiang, J. (2011). Distinct DNA methylation patterns associated with active and inactive centromeres of the maize B chromosome. Genome research *21*, 908-914.

Koo, D.H., Zhao, H., and Jiang, J. (2016). Chromatin-associated transcripts of tandemly repetitive DNA sequences revealed by RNA-FISH. Chromosome Res *24*, 467-480.

Kopecna, O., Kubickova, S., Cernohorska, H., Cabelova, K., Vahala, J., Martinkova, N., and Rubes, J. (2014). Tribe-specific satellite DNA in non-domestic Bovidae. Chromosome Res *22*, 277-291.

Kuhn, G.C., Sene, F.M., Moreira-Filho, O., Schwarzacher, T., and Heslop-Harrison, J.S. (2008). Sequence analysis, chromosomal distribution and long-range organization show that rapid turnover of new and old pBuM satellite DNA repeats leads to different patterns of variation in seven species of the Drosophila buzzatii cluster. Chromosome Res *16*, 307-324.

Kursel, L.E., and Malik, H.S. (2016). Centromeres. Curr Biol *26*, R487-r490.

Lam, A.L., Boivin, C.D., Bonney, C.F., Rudd, M.K., and Sullivan, B.A. (2006). Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. Proc Natl Acad Sci U S A *103*, 4186-4191.

Lamb, J.C., and Birchler, J.A. (2003). The role of DNA sequence in centromere formation. Genome Biol *4*, 214.

Lamb, J.C., Kato, A., and Birchler, J.A. (2005). Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. Chromosoma *113*, 337-349.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Larin, Z., Fricker, M.D., and Tyler-Smith, C. (1994). De novo formation of several features of a centromere following introduction of a Y alphoid YAC into mammalian cells. Hum Mol Genet *3*, 689-695.

Lee, H.R., Neumann, P., Macas, J., and Jiang, J. (2006). Transcription and evolutionary dynamics of the centromeric satellite repeat CentO in rice. Mol Biol Evol *23*, 2505-2520.

Leeb, T., Vogl, C., Zhu, B., de Jong, P.J., Binns, M.M., Chowdhary, B.P., Scharfe, M., Jarek, M., Nordsiek, G., Schrader, F., *et al.* (2006). A human-horse comparative map based on equine BAC end sequences. Genomics *87*, 772–776.

Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol *13*, 1192-1200.

Lengauer, C., Kinzler, K.W., and Vogelstein, B. (1997). DNA methylation and genetic instability in colorectal cancer cells. Proc Natl Acad Sci U S A *94*, 2545-2550.

Levy, S., Sutton, G., Ng, P.C., Feuk, L., Halpern, A.L., Walenz, B.P., Axelrod, N., Huang, J., Kirkness, E.F., Denisov, G.*, et al.* (2007). The diploid genome sequence of an individual human. PLoS Biol *5*, e254.

Li, B., Choulet, F., Heng, Y., Hao, W., Paux, E., Liu, Z., Yue, W., Jin, W., Feuillet, C., and Zhang, X. (2013). Wheat centromeric retrotransposons: the new ones take a major role in centromeric structure. Plant J *73*, 952-965.

Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362-365.

Li, F., Sonbuchner, L., Kyes, S.A., Epp, C., and Deitsch, K.W. (2008). Nuclear non-coding RNAs are transcribed from the centromeres of Plasmodium falciparum and are associated with centromeric chromatin. J Biol Chem *283*, 5692-5698.

Li, S., Malafiej, P., Levy, B., Mahmood, R., Field, M., Hughes, T., Lockhart, L.H., Wu, Z., Huang, M., Hirschhorn, K*., et al.* (2002). Chromosome 13q neocentromeres: molecular cytogenetic characterization of three additional cases and clinical spectrum. Am J Med Genet *110*, 258-267.

Liang, K., Woodfin, A.R., Slaughter, B.D., Unruh, J.R., Box, A.C., Rickels, R.A., Gao, X., Haug, J.S., Jaspersen, S.L., and Shilatifard, A. (2015). Mitotic Transcriptional Activation: Clearance of Actively Engaged Pol II via Transcriptional Elongation Control in Mitosis. Mol Cell *60*, 435-445.

Liehr, T., Kosyakova, N., Weise, A., Ziegler, M., and Raabe-Meyer, G. (2010). First case of a neocentromere formation in an otherwise normal chromosome 7. Cytogenet Genome Res *128*, 189-191.

Lindsay, E.H., Opdyke, N.D., and Johnson, N.M. (1980). Pliocene dispersal of the horse Equus and late Cenozoic mammalian dispersal events. Nature *287*, 135.

Lippman, Z., and Martienssen, R. (2004). The role of RNA interference in heterochromatic silencing. Nature *431*, 364-370.

Lo, A.W., Craig, J.M., Saffery, R., Kalitsis, P., Irvine, D.V., Earle, E., Magliano, D.J., and Choo, K.H. (2001). A 330 kb CENP-A binding domain and altered replication timing at a human neocentromere. Embo j *20*, 2087-2096.

Logsdon, G.A., Barrey, E.J., Bassett, E.A., DeNizio, J.E., Guo, L.Y., Panchenko, T., Dawicki-McKenna, J.M., Heun, P., and Black, B.E. (2015). Both tails and the centromere targeting domain of CENP-A are required for centromere establishment. J Cell Biol *208*, 521-531.

Logsdon, G.A., Gambogi, C.W., Liskovykh, M.A., Barrey, E.J., Larionov, V., Miga, K.H., Heun, P., Black, B.E. (2019) Human Artificial Chromosomes that Bypass Centromeric DNA. Cell *178,* 624-639

Lohe, A.R., Hilliker, A.J., and Roberts, P.A. (1993). Mapping simple repeated DNA sequences in heterochromatin of Drosophila melanogaster. Genetics *134*, 1149-1174.

Lomiento, M., Jiang, Z., D'Addabbo, P., Eichler, E.E., and Rocchi, M. (2008). Evolutionary-new centromeres preferentially emerge within gene deserts. Genome Biol *9*, R173.

Lorincz, M.C., Dickerson, D.R., Schmitt, M., and Groudine, M. (2004). Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. Nature structural & molecular biology *11*, 1068.

Lorite, P., Carrillo, J.A., Tinaut, A., and Palomeque, T. (2002). Comparative study of satellite DNA in ants of the Messor genus. Gene *297*, 113-122.

Lorite, P., Muñoz-López, M., Carrillo, J.A., Sanllorente, O., Vela, J., Mora, P., Tinaut, A., Torres, M.I., and Palomeque, T. (2017). Concerted evolution, a slow process for ant satellite DNA: study of the satellite DNA in the Aphaenogaster genus (Hymenoptera, Formicidae). Organisms Diversity & Evolution *17*, 595-606.

Lunyak, V.V., Burgess, R., Prefontaine, G.G., Nelson, C., Sze, S.H., Chenoweth, J., Schwartz, P., Pevzner, P.A., Glass, C., Mandel, G.*, et al.* (2002). Corepressor-dependent silencing of chromosomal regions encoding neuronal genes. Science *298*, 1747-1752.

Ma, R., Peng, Y., Zhang, Y., Xia, Y., Tang, G., Chang, J., Guo, R., Gui, B., Huang, Y., Chen, C.*, et al.* (2015). Partial trisomy 2q33.3-q37.3 in a patient with an inverted duplicated neocentric marker chromosome. In Mol Cytogenet (England), p. 10.

Macas, J., Neumann, P., and Navratilova, A. (2007). Repetitive DNA in the pea (Pisum sativum L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and Medicago truncatula. BMC Genomics *8*, 427.

Macas, J., Neumann, P., Novak, P., and Jiang, J. (2010). Global sequence characterization of rice centromeric satellite based on oligomer frequency analysis in large-scale sequencing data. Bioinformatics *26*, 2101-2108.

Macas, J., Pozarkova, D., Navratilova, A., Nouzova, M., and Neumann, P. (2000). Two new families of tandem repeats isolated from genus Vicia using genomic self-priming PCR. Mol Gen Genet *263*, 741-751.

Maddox, P.S., Hyndman, F., Monen, J., Oegema, K., and Desai, A. (2007). Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. J Cell Biol *176*, 757-763.

Maggert, K.A., and Karpen, G.H. (2000). Acquisition and metastability of centromere identity and function: sequence analysis of a human neocentromere. Genome Res *10*, 725-728.

Maggert, K.A., and Karpen, G.H. (2001). The activation of a neocentromere in Drosophila requires proximity to an endogenous centromere. Genetics *158*, 1615-1628.

Malik, H.S., Vermaak, D., and Henikoff, S. (2002). Recurrent evolution of DNA-binding motifs in the Drosophila centromeric histone. Proc Natl Acad Sci U S A *99*, 1449-1454.

Malvezzi, F., Litos, G., Schleiffer, A., Heuck, A., Mechtler, K., Clausen, T., and Westermann, S. (2013). A structural basis for kinetochore recruitment of the Ndc80 complex via two distinct centromere receptors. Embo j *32*, 409-423.

Maraschio, P., Tupler, R., Rossi, E., Barbierato, L., Uccellatore, F., Rocchi, M., Zuffardi, O., and Fraccaro, M. (1996). A novel mechanism for the origin of supernumerary marker chromosomes. Hum Genet *97*, 382-386.

Marshall, O.J., Chueh, A.C., Wong, L.H., and Choo, K.H. (2008). Neocentromeres: new insights into centromere structure, disease development, and karyotype evolution. Am J Hum Genet *82*, 261-282.

Mascarenhas, A., Matoso, E., Saraiva, J., Tonnies, H., Gerlach, A., Juliao, M.J., Melo, J.B., and Carreira, I.M. (2008). First prenatally detected small supernumerary neocentromeric derivative chromosome 13 resulting in a non-mosaic partial tetrasomy 13q. Cytogenet Genome Res *121*, 293-297.

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

Mayer, R., Brero, A., von Hase, J., Schroeder, T., Cremer, T., and Dietzel, S. (2005). Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. BMC Cell Biol *6*, 44.

McEwen, B.F., Dong, Y., and VandenBeldt, K.J. (2007). Using electron microscopy to understand functional mechanisms of chromosome alignment on the mitotic spindle. Methods Cell Biol *79*, 259-293.

McKinley, K.L., and Cheeseman, I.M. (2016). The molecular basis for centromere identity and function. Nat Rev Mol Cell Biol *17*, 16-29.

McKinley, K.L., Sekulic, N., Guo, L.Y., Tsinman, T., Black, B.E., and Cheeseman, I.M. (2015). The CENP-L-N Complex Forms a Critical Node in an Integrated Meshwork of Interactions at the Centromere-Kinetochore Interface. Mol Cell *60*, 886-898.

Mellone, B.G., and Allshire, R.C. (2003). Stretching it: putting the CEN(P-A) in centromere. Curr Opin Genet Dev *13*, 191-198.

Mestrovic, N., Pavlek, M., Car, A., Castagnone-Sereno, P., Abad, P., and Plohl, M. (2013). Conserved DNA Motifs, Including the CENP-B Box-like, Are Possible Promoters of Satellite DNA Array Rearrangements in Nematodes. PLoS One *8*, e67328.

Mestrovic, N., Plohl, M., Mravinac, B., and Ugarkovic, D. (1998). Evolution of satellite DNAs from the genus Palorus--experimental evidence for the "library" hypothesis. Mol Biol Evol *15*, 1062-1068.

Miga, K.H. (2015). Completing the human genome: the progress and challenge of satellite DNA assembly. Chromosome Res *23*, 421-426.

Miniou, P., Jeanpierre, M., Bourc'his, D., Coutinho Barbosa, A.C., Blanquet, V., and Viegas-Pequignot, E. (1997). alpha-satellite DNA methylation in normal individuals and in ICF patients: heterogeneous methylation of constitutive heterochromatin in adult and fetal tissues. Hum Genet *99*, 738-745.

Minoshima, Y., Hori, T., Okada, M., Kimura, H., Haraguchi, T., Hiraoka, Y., Bao, Y.C., Kawashima, T., Kitamura, T., and Fukagawa, T. (2005). The constitutive centromere component CENP-50 is required for recovery from spindle damage. Mol Cell Biol *25*, 10315-10328.

Misceo, D., Cardone, M.F., Carbone, L., D'Addabbo, P., de Jong, P.J., Rocchi, M., and Archidiacono, N. (2005). Evolutionary history of chromosome 20. Mol Biol Evol *22*, 360-366.

Molina, O., Vargiu, G., Abad, M.A., Zhiteneva, A., Jeyaprakash, A.A., Masumoto, H., Kouprina, N., Larionov, V., and Earnshaw, W.C. (2016). Epigenetic engineering reveals a balance between histone modifications and transcription in kinetochore maintenance. Nat Commun *7*, 13334.

Montefalcone, G., Tempesta, S., Rocchi, M., and Archidiacono, N. (1999). Centromere repositioning. Genome Res *9*, 1184-1188.

Moroi, Y., Peebles, C., Fritzler, M.J., Steigerwald, J., and Tan, E.M. (1980). Autoantibody to centromere (kinetochore) in scleroderma sera. Proceedings of the National Academy of Sciences *77*, 1627-1631.

Morrissette, J.D., Celle, L., Owens, N.L., Shields, C.L., Zackai, E.H., and Spinner, N.B. (2001). Boy with bilateral retinoblastoma due to an unusual ring chromosome 13 with activation of a latent centromere. Am J Med Genet *99*, 21-28.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789-802.

Mravinac, B., Plohl, M., and Ugarkovic, D. (2005). Preservation and high sequence conservation of satellite DNAs suggest functional constraints. J Mol Evol *61*, 542-550.

Nagaki, K., Cheng, Z., Ouyang, S., Talbert, P.B., Kim, M., Jones, K.M., Henikoff, S., Buell, C.R., and Jiang, J. (2004a). Sequencing of a rice centromere uncovers active genes. Nat Genet *36*, 138-145.

Nagaki, K., Neumann, P., Zhang, D., Ouyang, S., Buell, C.R., Cheng, Z., and Jiang, J. (2004b). Structure, divergence, and distribution of the CRR centromeric retrotransposon family in rice. Molecular biology and evolution *22*, 845-855.

Nagpal, H., Hori, T., Furukawa, A., Sugase, K., Osakabe, A., Kurumizaka, H., and Fukagawa, T. (2015). Dynamic changes in CCAN organization through CENP-C during cell-cycle progression. Mol Biol Cell *26*, 3768-3776.

Nakano, M., Cardinale, S., Noskov, V.N., Gassmann, R., Vagnarelli, P., Kandels-Lewis, S., Larionov, V., Earnshaw, W.C., and Masumoto, H. (2008). Inactivation of a human kinetochore by specific targeting of chromatin modifiers. Dev Cell *14*, 507-522.

Narayan, A., Ji, W., Zhang, X.Y., Marrogi, A., Graff, J.R., Baylin, S.B., and Ehrlich, M. (1998). Hypomethylation of pericentromeric DNA in breast adenocarcinomas. Int J Cancer *77*, 833-838.

Nasuda, S., Hudakova, S., Schubert, I., Houben, A., and Endo, T.R. (2005). Stable barley chromosomes without centromeric repeats. Proc Natl Acad Sci U S A *102*, 9842-9847.

Nergadze, S.G., Belloni, E., Piras, F.M., Khoriauli, L., Mazzagatti, A., Vella, F., Bensi, M., Vitelli, V., Giulotto, E., and Raimondi, E. (2014). Discovery and comparative analysis of a novel satellite, EC137, in horses and other equids. Cytogenet Genome Res *144*, 114-123.

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.* (2018). Birth, evolution, and transmission of satellite-free mammalian centromeric domains. Genome Res *28*, 789-799.

Neumann, P., Yan, H., and Jiang, J. (2007). The centromeric retrotransposons of rice are transcribed and differentially processed by RNA interference. Genetics *176*, 749-761.

Nishibuchi, G., and Déjardin, J. (2017). The molecular basis of the organization of repetitive DNA-containing constitutive heterochromatin in mammals. Chromosome Research *25*, 77-87.

Nishihashi, A., Haraguchi, T., Hiraoka, Y., Ikemura, T., Regnier, V., Dodson, H., Earnshaw, W.C., and Fukagawa, T. (2002). CENP-I is essential for centromere function in vertebrate cells. Dev Cell *2*, 463-476.

Nishimura, K., Komiya, M., Hori, T., Itoh, T., and Fukagawa, T. (2019). 3D genomic architecture reveals that neocentromeres associate with heterochromatin regions. J Cell Biol *218*, 134-149.

Nishino, T., Takeuchi, K., Gascoigne, K.E., Suzuki, A., Hori, T., Oyama, T., Morikawa, K., Cheeseman, I.M., and Fukagawa, T. (2012). CENP-T-W-S-X forms a unique centromeric chromatin structure with a histone-like fold. Cell *148*, 487-501.

Nonaka, N., Kitajima, T., Yokobayashi, S., Xiao, G., Yamamoto, M., Grewal, S.I.S., and Watanabe, Y. (2002). Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. Nature Cell Biology *4*, 89-93.

O'Neill, R.J., O'Neill, M.J., and Graves, J.A. (1998). Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid. Nature *393*, 68-72.

Oakenfull, E.A., and Clegg, J.B. (1998). Phylogenetic relationships within the genus Equus and the evolution of alpha and theta globin genes. J Mol Evol *47*, 772-783.

Ohkuni, K., and Kitagawa, K. (2011). Endogenous transcription at the centromere facilitates centromere activity in budding yeast. Curr Biol *21*, 1695-1703.

Ohno, S. So much'junk'DNA in our genome.

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

Okada, M., Cheeseman, I.M., Hori, T., Okawa, K., McLeod, I.X., Yates, J.R., 3rd, Desai, A., and Fukagawa, T. (2006). The CENP-H-I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. Nat Cell Biol *8*, 446-457.

Olszak, A.M., van Essen, D., Pereira, A.J., Diehl, S., Manke, T., Maiato, H., Saccani, S., and Heun, P. (2011). Heterochromatin boundaries are hotspots for de novo kinetochore formation. Nat Cell Biol *13*, 799-808.

Padeken, J., and Heun, P. (2013). Centromeres in nuclear architecture. In Cell Cycle (United States), pp. 3455-3456.

Palazzo, A.F., and Gregory, T.R. (2014). The case for junk DNA. PLoS genetics *10*, e1004351.

Pardue, M.L., and Gall, J.G. (1970). Chromosomal localization of mouse satellite DNA. Science *168*, 1356-1358.

Partridge, J.F., Borgstrom, B., and Allshire, R.C. (2000). Distinct protein interaction domains and protein spreading in a complex centromere. Genes Dev *14*, 783-791.

Perpelescu, M., and Fukagawa, T. (2011). The ABCs of CENPs. Chromosoma *120*, 425-446.

Peters, A.H., Kubicek, S., Mechtler, K., O'Sullivan, R.J., Derijck, A.A., Perez-Burgos, L., Kohlmaier, A., Opravil, S., Tachibana, M., Shinkai, Y.*, et al.* (2003). Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. Mol Cell *12*, 1577-1589.

Petraccioli, A., Odierna, G., Capriglione, T., Barucca, M., Forconi, M., Olmo, E., and Biscotti, M.A. (2015). A novel satellite DNA isolated in Pecten jacobaeus shows high sequence similarity among molluscs. Mol Genet Genomics *290*, 1717-1725.

Pezer, Z., Brajkovic, J., Feliciello, I., and Ugarkovc, D. (2012). Satellite DNA-mediated effects on genome regulation. Genome Dyn *7*, 153-169.

Pezer, Z., and Ugarkovic, D. (2008). Role of non-coding RNA and heterochromatin in aneuploidy and cancer. Semin Cancer Biol *18*, 123-130.

Pezer, Z., and Ugarkovic, D. (2009). Transcription of pericentromeric heterochromatin in beetles--satellite DNAs as active regulatory elements. Cytogenet Genome Res *124*, 268-276.

Piras, F.M., Nergadze, S.G., Magnani, E., Bertoni, L., Attolini, C., Khoriauli, L., Raimondi, E., and Giulotto, E. (2010). Uncoupling of satellite DNA and centromeric function in the genus Equus. PLoS Genet *6*, e1000845.

Piras, F.M., Nergadze, S.G., Poletto, V., Cerutti, F., Ryder, O.A., Leeb, T., Raimondi, E., and Giulotto, E. (2009). Phylogeny of horse chromosome 5q in the genus Equus and centromere repositioning. Cytogenet Genome Res *126*, 165-172.

Plohl, M., Luchetti, A., Mestrovic, N., and Mantovani, B. (2008). Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. Gene *409*, 72-82.

Plohl, M., Mestrovic, N., and Mravinac, B. (2012). Satellite DNA evolution. Genome Dyn *7*, 126-152.

Plohl, M., Mestrovic, N., and Mravinac, B. (2014). Centromere identity from the DNA point of view. Chromosoma *123*, 313-325.

Pluta, A.F., Mackay, A.M., Ainsztein, A.M., Goldberg, I.G., and Earnshaw, W.C. (1995). The centromere: hub of chromosomal activities. Science *270*, 1591-1594.

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.* (2015). Centromere sliding on a mammalian chromosome. Chromosoma *124*, 277-287.

Qu, G., Dubeau, L., Narayan, A., Yu, M.C., and Ehrlich, M. (1999). Satellite DNA hypomethylation vs. overall genomic hypomethylation in ovarian epithelial tumors of different malignant potential. Mutat Res *423*, 91-101.

Quenet, D., and Dalal, Y. (2014). A long non-coding RNA is required for targeting centromeric protein A to the human centromere. Elife *3*, e03254.

Quenet, D., Sturgill, D., Olson, M., and Dalal, Y. (2017). CENP-A associated lncRNAs influence chromosome segregation in human cells. BioRxiv, 097956.

Raaum, R.L., Sterner, K.N., Noviello, C.M., Stewart, C.B., and Disotell, T.R. (2005). Catarrhine primate divergence dates estimated from complete mitochondrial genomes: concordance with fossil and nuclear DNA evidence. J Hum Evol *48*, 237-257.

Raimondi, E., Piras, F.M., Nergadze, S.G., Di Meo, G.P., Ruiz-Herrera, A., Ponsa, M., Ianuzzi, L., and Giulotto, E. (2011). Polymorphic organization of constitutive heterochromatin in Equus asinus (2n = 62) chromosome 1. Hereditas *148*, 110-113.

Raimondi, E., Scariolo, S., De Sario, A., and De Carli, L. (1989). Aneuploidy assays on interphase nuclei by means of in situ hybridization with DNA probes. Mutagenesis *4*, 165-169.

Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y., and Allis, C.D. (2003). Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol Cell *12*, 1591-1598.

Roberto, R., Capozzi, O., Wilson, R.K., Mardis, E.R., Lomiento, M., Tuzun, E., Cheng, Z., Mootnick, A.R., Archidiacono, N., Rocchi, M.*, et al.* (2007). Molecular refinement of gibbon genome rearrangements. Genome Res *17*, 249-257.

Rocchi, M., Archidiacono, N., Schempp, W., Capozzi, O., and Stanyon, R. (2012). Centromere repositioning in mammals. Heredity (Edinb) *108*, 59-67.

Rocchi, M., Stanyon, R., and Archidiacono, N. (2009). Evolutionary new centromeres in primates. Prog Mol Subcell Biol *48*, 103-152.

Rojo, V., Martinez-Lage, A., Giovannotti, M., Gonzalez-Tizon, A.M., Nisi Cerioni, P., Caputo Barucchi, V., Galan, P., Olmo, E., and Naveira, H. (2015). Evolutionary dynamics of two satellite DNA families in rock lizards of the genus Iberolacerta (Squamata, Lacertidae): different histories but common traits. Chromosome Res *23*, 441-461.

Rosic, S., and Erhardt, S. (2016). No longer a nuisance: long non-coding RNAs join CENP-A in epigenetic centromere regulation. Cell Mol Life Sci *73*, 1387-1398.

Rosic, S., Kohler, F., and Erhardt, S. (2014). Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. J Cell Biol *207*, 335-349.

Rowe, A.G., Abrams, L., Qu, Y., Chen, E., and Cotter, P.D. (2000). Tetrasomy 15q25-->qter: cytogenetic and molecular characterization of an analphoid supernumerary marker chromosome. Am J Med Genet *93*, 393-398.

Ruchaud, S., Carmena, M., and Earnshaw, W.C. (2007). Chromosomal passengers: conducting cell division. Nat Rev Mol Cell Biol *8*, 798-812.

Ruiz-Ruano, F.J., Lopez-Leon, M.D., Cabrero, J., and Camacho, J.P.M. (2016). High-throughput analysis of the satellitome illuminates satellite DNA evolution. Sci Rep *6*, 28333.

Saffery, R., Sumer, H., Hassan, S., Wong, L.H., Craig, J.M., Todokoro, K., Anderson, M., Stafford, A., and Choo, K.H. (2003). Transcription within a functional human centromere. Mol Cell *12*, 509-516.

Saito, Y., Kanai, Y., Sakamoto, M., Saito, H., Ishii, H., and Hirohashi, S. (2001). Expression of mRNA for DNA methyltransferases and methyl-CpG-binding proteins and DNA methylation status on CpG islands and pericentromeric satellite regions during human hepatocarcinogenesis. Hepatology *33*, 561-568.

Saksouk, N., Simboeck, E., and Dejardin, J. (2015). Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin *8*, 3.

Sandall, S., Severin, F., McLeod, I.X., Yates, J.R., 3rd, Oegema, K., Hyman, A., and Desai, A. (2006). A Bir1-Sli15 complex connects centromeres to microtubules and is required to sense kinetochore tension. Cell *127*, 1179-1191.

Sanyal, K., Baum, M., and Carbon, J. (2004). Centromeric DNA sequences in the pathogenic yeast Candida albicans are all different and unique. Proc Natl Acad Sci U S A *101*, 11374-11379.

Sarraf, S.A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Molecular cell *15*, 595-605.

Scarpato, M., Angelini, C., Cocca, E., Pallotta, M.M., Morescalchi, M.A., and Capriglione, T. (2015). Short interspersed DNA elements and miRNAs: a novel hidden gene regulation layer in zebrafish? Chromosome Res *23*, 533-544.

Schneider, R., Bannister, A.J., Myers, F.A., Thorne, A.W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol *6*, 73-77.

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

Screpanti, E., De Antoni, A., Alushin, G.M., Petrovic, A., Melis, T., Nogales, E., and Musacchio, A. (2011). Direct binding of Cenp-C to the Mis12 complex joins the inner and outer kinetochore. Curr Biol *21*, 391-398.

Shang, W.H., Hori, T., Martins, N.M., Toyoda, A., Misu, S., Monma, N., Hiratani, I., Maeshima, K., Ikeo, K., Fujiyama, A*., et al.* (2013). Chromosome engineering allows the efficient isolation of vertebrate neocentromeres. Dev Cell *24*, 635-648.

Stanyon, R., Rocchi, M., Capozzi, O., Roberto, R., Misceo, D., Ventura, M., Cardone, M.F., Bigoni, F., and Archidiacono, N. (2008). Primate chromosome evolution: ancestral karyotypes, marker order and neocentromeres. Chromosome Res *16*, 17-39.

Stoler, S., Keith, K.C., Curnick, K.E., and Fitzgerald-Hayes, M. (1995). A mutation in CSE4, an essential gene encoding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. Genes Dev *9*, 573-586.

Stumpff, J., Von Dassow, G., Wagenbach, M., Asbury, C., and Wordeman, L. (2008). The kinesin-8 motor Kif18A suppresses kinetochore movements to control mitotic chromosome alignment. Developmental cell *14*, 252-262.

Subirana, J.A., Alba, M.M., and Messeguer, X. (2015). High evolutionary turnover of satellite families in Caenorhabditis. BMC Evol Biol *15*, 218.

Subirana, J.A., and Messeguer, X. (2013). A satellite explosion in the genome of holocentric nematodes. PLoS One *8*, e62221.

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

Sullivan, B.A., and Willard, H.F. (1998). Stable dicentric X chromosomes with two functional centromeres. Nat Genet *20*, 227-228.

Sullivan, K.F., and Glass, C.A. (1991). CENP-B is a highly conserved mammalian centromere protein with homology to the helix-loop-helix family of proteins. Chromosoma *100*, 360-370.

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

Tamaru, H., and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature *414*, 277-283.

Teshima, I., Bawle, E.V., Weksberg, R., Shuman, C., Van Dyke, D.L., and Schwartz, S. (2000). Analphoid 3qter markers. Am J Med Genet *94*, 113-119.

Therman, E., Trunca, C., Kuhn, E.M., and Sarto, G.E. (1986). Dicentric chromosomes and the inactivation of the centromere. Hum Genet *72*, 191-195.

Topp, C.N., Okagaki, R.J., Melo, J.R., Kynast, R.G., Phillips, R.L., and Dawe, R.K. (2009). Identification of a maize neocentromere in an oat-maize addition line. Cytogenet Genome Res *124*, 228-238.

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

Trifonov, V.A., Musilova, P., and Kulemsina, A.I. (2012). Chromosome evolution in Perissodactyla. Cytogenet Genome Res *137*, 208-217.

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.* (2008). Multidirectional cross-species painting illuminates the history of karyotypic evolution in Perissodactyla. Chromosome Res *16*, 89-107.

Ugarkovic, D. (2005). Functional elements residing within satellite DNAs. EMBO Rep *6*, 1035-1039.

Ventura, M., Antonacci, F., Cardone, M.F., Stanyon, R., D'Addabbo, P., Cellamare, A., Sprague, L.J., Eichler, E.E., Archidiacono, N., and Rocchi, M. (2007). Evolutionary formation of new centromeres in macaque. Science *316*, 243-246.

Ventura, M., Archidiacono, N., and Rocchi, M. (2001). Centromere emergence in evolution. Genome Res *11*, 595-599

Ventura, M., Mudge, J.M., Palumbo, V., Burn, S., Blennow, E., Pierluigi, M., Giorda, R., Zuffardi, O., Archidiacono, N., Jackson, M.S*., et al.* (2003). Neocentromeres in 15q24-26 map to duplicons which flanked an ancestral centromere in 15q25. Genome Res *13*, 2059-2068.

Ventura, M., Weigl, S., Carbone, L., Cardone, M.F., Misceo, D., Teti, M., D'Addabbo, P., Wandall, A., Bjorck, E., de Jong, P.J*., et al.* (2004). Recurrent sites for new centromere seeding. Genome Res *14*, 1696-1703.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672-676.

Vig, B.K. (1982). Sequence of centromere separation: role of centromeric heterochromatin. Genetics *102*, 795-806.

Vinson, C., and Chatterjee, R. (2012). CG methylation. Epigenomics *4*, 655-663.

Vissel, B., and Choo, K.H. (1989). Mouse major (gamma) satellite DNA is highly conserved and organized into extremely long tandem arrays: implications for recombination between nonhomologous chromosomes. Genomics *5*, 407-414.

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

Voullaire, L.E., Slater, H.R., Petrovic, V., and Choo, K.H. (1993). A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? Am J Hum Genet *52*, 1153-1163.

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.* (2009). Genome sequence, comparative analysis, and population genetics of the domestic horse. Science *326*, 865-867.

Warburton, P.E., Dolled, M., Mahmood, R., Alonso, A., Li, S., Naritomi, K., Tohma, T., Nagai, T., Hasegawa, T., Ohashi, H.*, et al.* (2000). Molecular cytogenetic analysis of eight inversion duplications of human chromosome 13q that each contain a neocentromere. Am J Hum Genet *66*, 1794-1806.

Westhorpe, F.G., Fuller, C.J., and Straight, A.F. (2015). A cell-free CENP-A assembly system defines the chromatin requirements for centromere maintenance. J Cell Biol *209*, 789-801.

Williams, B.C., Murphy, T.D., Goldberg, M.L., and Karpen, G.H. (1998). Neocentromere activity of structurally acentric mini-chromosomes in Drosophila. Nat Genet *18*, 30-37.

Wolfgruber, T.K., Sharma, A., Schneider, K.L., Albert, P.S., Koo, D.H., Shi, J., Gao, Z., Han, F., Lee, H., Xu, R.*, et al.* (2009). Maize centromere structure and evolution: sequence analysis of centromeres 2 and 5 reveals dynamic Loci shaped primarily by retrotransposons. PLoS Genet *5*, e1000743.

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.* (2007). Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. Genome Res *17*, 1146-1160.

Wong, N.C., Wong, L.H., Quach, J.M., Canham, P., Craig, J.M., Song, J.Z., Clark, S.J., and Choo, K.H. (2006). Permissive transcriptional activity at the centromere through pockets of DNA hypomethylation. PLoS Genet *2*, e17.

Xu, G.L., Bestor, T.H., Bourc'his, D., Hsieh, C.L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J.J., and Viegas-Pequignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature *402*, 187-191.

Yan, H., Ito, H., Nobuta, K., Ouyang, S., Jin, W., Tian, S., Lu, C., Venu, R.C., Wang, G.L., Green, P.J.*, et al.* (2006). Genomic and genetic characterization of rice Cen3 reveals extensive transcription and evolutionary implications of a complex centromere. Plant Cell *18*, 2123-2133.

Yan, H., Jin, W., Nagaki, K., Tian, S., Ouyang, S., Buell, C.R., Talbert, P.B., Henikoff, S., and Jiang, J. (2005). Transcription and histone modifications in the recombination-free region spanning a rice centromere. Plant Cell *17*, 3227-3238.

Yao, X., Anderson, K.L., and Cleveland, D.W. (1997). The microtubule-dependent motor centromere-associated protein E (CENP-E) is an integral component of kinetochore corona fibers that link centromeres to spindle microtubules. J Cell Biol *139*, 435-447.

Yen, T.J., Compton, D.A., Wise, D., Zinkowski, R.P., Brinkley, B.R., Earnshaw, W.C., and Cleveland, D.W. (1991). CENP-E, a novel human centromere-associated protein required for progression from metaphase to anaphase. Embo j *10*, 1245-1254.

Yoda, K., Ando, S., Morishita, S., Houmura, K., Hashimoto, K., Takeyasu, K., and Okazaki, T. (2000). Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. Proc Natl Acad Sci U S A *97*, 7266-7271.

Yuen, K.W., Nabeshima, K., Oegema, K., and Desai, A. (2011). Rapid de novo centromere formation occurs independently of heterochromatin protein 1 in C. elegans embryos. Curr Biol *21*, 1800-1807.

Yunis, J.J., and Yasmineh, W.G. (1971). Heterochromatin, satellite DNA, and cell function. Structural DNA of eucaryotes may support and protect genes and aid in speciation. Science *174*, 1200-1209.

Zeitlin, S.G., Baker, N.M., Chapados, B.R., Soutoglou, E., Wang, J.Y., Berns, M.W., and Cleveland, D.W. (2009). Double-strand DNA breaks recruit the centromeric histone CENP-A. Proc Natl Acad Sci U S A *106*, 15762-15767.

Zhang, T., Talbert, P.B., Zhang, W., Wu, Y., Yang, Z., Henikoff, J.G., Henikoff, S., and Jiang, J. (2013). The CentO satellite confers translational and rotational phasing on cenH3 nucleosomes in rice centromeres. Proc Natl Acad Sci U S A *110*, E4875-4883.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes & development *15*, 2343-2360.

Zhu, Q., Pao, G.M., Huynh, A.M., Suh, H., Tonnu, N., Nederlof, P.M., Gage, F.H., and Verma, I.M. (2011). BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. Nature *477*, 179-184.

LIST OF ORIGINAL MANUSCRIPTS

PEER REVIEW PAPER

- **Annalisa Roberti**, Mirella Bensi, Alice Mazzagatti, Francesca M. Piras, Solomon G. Nergadze, Elena Giulotto* and Elena Raimondi* "Satellite DNA at the Centromere is Dispensable for segregation fidelity". Gene. 2019; 10: 469.

MEETING ABSTRACTS AND ORAL COMMUNICATIONS

- **Annalisa Roberti**, Alessia Ferro, Francesca M. Piras, Elena Giulotto and Elena Raimondi "Chromatin Fibre Analysis for the Epigenetic Characterization of Equid centromeres". Convegno AGI/SIMAG – Associazione Genetica Italiana/Società Italiana di Mutagenesi Ambientale e Genomica. Cortona, 26-28 September 2019.
- Marco Corbo, **Annalisa Roberti**, Francesca M. Piras, Eleonora Cappelletti, Mirella Bensi, Solomon G. Nergadze, Elena Raimondi, Elena Giulotto "The epigenetic landscape of mammalian centromeres: a cytogenetic approach". 2ND Joint Annual Symposium. Pavia, 20-21-22 June 2018
- **Annalisa Roberti**, Marco Corbo, Mirella Bensi, Francesca Piras, Elena Giulotto and Elena Raimondi "Epigenetic Marks at Satellite-Free and Satellite-Based Centromeres". FISV conference - Italian Federation of Life Sciences. Roma, 18-21 September 2018.
- **Annalisa Roberti**, Marco Corbo, Solomon G. Nergadze, Mirella Bensi, Francesca M. Piras, Elena Giulotto and Elena Raimondi "Satellite-Free and Satellite-Based Centromeres: Epigenetic Marks". Havemayer Conference. 12th International Horse Genome Workshop. Pavia, 12-15 September 2018
- **Annalisa Roberti**, Solomon G. Nergadze, Mirella Bensi, Riccardo Gamba, Marco Corbo, Francesca M. Piras, Eleonora Cappelletti, Elena Giulotto and Elena Raimondi. "Epigenetic modifications at satellite-less evolutionarily new centromeres". Convegno AGI - Associazione Genetica Italiana. Cortona, 7-9 September 2017.
- Mazzagatti A, Langella A, **Roberti A**, Bensi M, Piras F. M., Cappelletti E., GambaR, Giulotto E and Raimondi E "The epigenetic landscape of equid centromeres: a cytogenetic approach". FISV conference - Italian Federation of Life Sciences. Roma, September 2016.

genes

MDPI

Brief Report

Satellite DNA at the Centromere Is Dispensable for Segregation Fidelity

Annalisa Roberti, Mirella Bensi, Alice Mazzagatti, Francesca M. Piras, Solomon G. Nergadze, Elena Giulotto * and Elena Raimondi *

Department of Biology and Biotechnology "L. Spallanzani", University of Pavia, Via Ferrata 1, 27100 Pavia, Italy; annalisamanuel.roberti01@universitadipavia.it (A.R.); mirella.bensi@unipv.it (M.B.); a.mazzagatti@qmul.ac.uk (A.M.); mfrancesca.piras@unipv.it (F.M.P.); solomon.nergadze@unipv.it (S.G.N.)

* Correspondence: elena.giulotto@unipv.it (E.G.); elena.raimondi@unipv.it (E.R.)

Received: 7 June 2019; Accepted: 19 June 2019; Published: 20 June 2019

Abstract: The typical vertebrate centromeres contain long stretches of highly repeated DNA sequences (satellite DNA). We previously demonstrated that the karyotypes of the species belonging to the genus *Equus* are characterized by the presence of satellite-free and satellite-based centromeres and represent a unique biological model for the study of centromere organization and behavior. Using horse primary fibroblasts cultured in vitro, we compared the segregation fidelity of chromosome 11, whose centromere is satellite-free, with that of chromosome 13, which has similar size and a centromere containing long stretches of satellite DNA. The mitotic stability of the two chromosomes was compared under normal conditions and under mitotic stress induced by the spindle inhibitor, nocodazole. Two independent molecular-cytogenetic approaches were used—the interphase aneuploidy analysis and the cytokinesis-block micronucleus assay. Both assays were coupled to fluorescence in situ hybridization with chromosome specific probes in order to identify chromosome 11 and chromosome 13, respectively. In addition, we tested if the lack of centromeric satellite DNA affected chromatid cohesion under normal and stress conditions. We demonstrated that, in our system, the segregation fidelity of a chromosome is not influenced by the presence of long stretches of tandem repeats at its centromere. To our knowledge, the present study is the first analysis of the mitotic behavior of a natural satellite-free centromere.

Keywords: centromere; horse; satellite DNA; satellite-free centromere; segregation fidelity; aneuploidy; micronucleus assay; Fluorescence In Situ Hybridization

1. Introduction

The centromere is a *locus* essential for chromosome segregation and genome integrity; this function is carried out by supporting kinetochore assembly and attachment to spindle microtubules. Most vertebrate centromeres are composed of long stretches of highly repeated DNA sequences named satellite DNA [1,2]. Nonetheless, kinetochore assembly is not dependent on the primary DNA sequence, and the centromeres are epigenetically determined, the modified H3 histone CENP-A (CENtromere Protein A) being the marker of centromere function. The existence of completely satellite-free pathological [3–5] and natural [6–10] centromeres raises the question of whether satellite DNA at centromeres plays any functional role. Centromeric repetitive DNA is typically devoid of active genes, thus it may aid the formation of a heterochromatic environment, which would favor the stability of the chromosome during mitosis and meiosis [1,2,5]. In several species, centromeric satellite DNA is transcribed, and it has been suggested that transcription of the centromeric regions may be important for chromatin opening and CENP-A loading; these transcripts are believed to provide a flexible scaffold that allows assembly or stabilization of the kinetochore proteins and may act in trans

on all or on a subset of chromosomes independently of the primary DNA sequence [11–14]. Moreover, most mammalian centromeric satellite DNA sequences contain binding sites for the constitutive centromeric protein CENP-B (CENtromere Protein B), whose role is still a matter of debate [15–18].

A crucial issue in centromere biology concerns the contribution of satellite DNA to chromosome segregation fidelity. To our knowledge, the mitotic stability of satellite-free centromeres was not carefully investigated. Data from the analysis of pathologic satellite-free centromeres indicate that these marker chromosomes are often present in the individual in mosaic form. Since most human neocentromeres give rise to partial trisomy or tetrasomy, it is plausible to suppose that the selective disadvantage of partial aneuploidy is responsible for this mosaicism [5].

Human artificial chromosomes containing fully functional centromeres have been constructed by different approaches [19–22], and it was demonstrated that to be propagated in culture, they require alpha-satellite arrays, including binding sites for the CENP-B protein (CENP-B box) [23,24]. Later on, human artificial chromosomes with a conditional centromere were used to manipulate the epigenetic state of chromatin and to elucidate the requirements for proper centromere function [25–27]. These studies together with high resolution immunofluorescence on chromatin fibers [28] revealed that a centromere-specific balance between typical euchromatic and typical heterochromatic post-translational histone modifications is essential for kinetochore activity, with highly repetitive DNA stretches having a central role.

We previously described a biological model system represented by species belonging to the genus *Equus* that is especially suitable for the dissection of centromere function [6,14,29–34]; in these species, the centromere function and the position of satellite DNA turned out to be often uncoupled [30]. Moreover, centromere repositioning (that is, centromere movement along the chromosome without rearrangement [35]) was unexpectedly frequent in *Equus* species [36] and generated satellite-free centromeres. In particular, one horse (ECA, *Equus Caballus*) chromosome (ECA11) and 16 donkey (EAS, *Equus Asinus*) chromosomes (EAS4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 16, 18, 19, 27, 30, X) were demonstrated to be completely satellite-free at the sequence level [6,34].

Strikingly, the satellite-free centromere of horse chromosome 11 and the 16 satellite-free donkey centromeres showed a sliding behavior. In other words, we demonstrated that the functional centromeric domain as defined by CENP-A binding at these centromeres can move within a region spanning 500-600 kb, thus generating functional alleles or "epialleles" [33,34].

In the present paper, we analyzed the mitotic stability of horse chromosome 11 (ECA11), whose centromere is completely satellite-free, and compared it with that of chromosome 13 (ECA13), which has similar size and a centromere containing long stretches of the canonical horse centromeric satellite DNA families [30,32]. The comparison was performed under normal conditions and after exposure of the cells to mitotic stress induced by the spindle inhibitor nocodazole. Two chromosome stability assays, interphase aneuploidy analysis and the cytokinesis-block micronucleus assay, were combined with fluorescence in situ hybridization (FISH) with chromosome specific probes.

2. Materials and Methods

2.1. Cell Line and Metaphase Spread Preparation

A horse skin fibroblast cell line was previously established [30]. The fibroblasts were cultured in Dulbecco's modified Eagle's medium (CELBIO) supplemented with 20% fetal calf serum (Bio-west), 2 mM glutamine, 2% non-essential amino acids, and 1x penicillin/streptomycin. Cells were maintained at 37 °C in a humidified atmosphere of 5% CO2.

Mitotically active cells were collected by flushing the medium on the cell monolayer. Metaphase spreads were prepared following the standard air-drying procedure.

2.2. Drug Treatment and Cytokinesis-Blok

Fifty thousand cells were seeded in 5 cm diameter Petri dishes containing 24 × 24 mm coverslips. After a 24 h culture period, cells were exposed to nocodazole (Sigma, St. Louis, MO, USA) 100 nM. Since nocodazole is soluble in DiMethyl SulfOxide (DMSO), control cells were treated with the same DMSO concentration used for nocodazole treatment. For the cytokinesis-block micronucleus (CBMN) assay, 24 h after the addition of nocodazole (or DMSO in control samples) cytochalasin-B (Cyt-B), 5 µg/mL was added to the cultures. After 48 h of nocodazole treatment, the slides were processed for interphase nuclei and micronuclei preparations as follows. Slides were treated with 75 mM KCl hypotonic solution and incubated at 37 °C for 15 min. Cold (−20 °C) fixative (methanol:acetic acid—3:1) was added for 30 min. The fixation was repeated twice.

2.3. BAC Extraction and FISH

Two bacterial artificial chromosomes (BACs) derived from horse CHORI-241 BAC library were used (CHORI241-21D14, chr11: 27,639,936-27,829,952; CHORI241-22C1, chr13: 7,346,775-7,544,907). The BACs were extracted from 100 mL of liquid culture (LB medium and chloramphenicol 12.5 µg/mL). The extraction was carried out with Qiagen Plasmid purification kit®, according to the supplier instructions.

The BACs were labeled by nick translation with Cy3-dUTP (Cyanine 3-dUTP, Perkin Elmer) and hybridized to metaphase spreads, interphase nuclei, or micronuclei preparations. For each slide, 250 ng of probe were used. Slides were denatured at 72 °C in 70% formamide 2 × SSC and immediately incubated overnight at 37 °C with the denatured probe in 50% formamide 2 × SSC. Post-hybridization washes were performed in 50% formamide 2 × SSC at 42 °C. Slides were counterstained with DAPI (4′, 6′-Diamidino-2-phenylindole hydrochloride, 1 µg/mL) and mounted in DAKO mounting medium.

A fluorescence microscope (Zeiss Axioplan) equipped with a cooled, grey scale charge coupled device (CCD) camera (Photometrics) was used for image capturing. Digital grey scale images for Cy3 and DAPI fluorescence signals were acquired separately, pseudo-colored, and merged using the IpLab software.

3. Results and Discussion

3.1. Interphase Aneuploidy Analysis

Aneuploidy is the presence of extra or missing chromosomes in a cell. Different mechanisms can lead to aneuploidy: (i) nondisjunction, which can involve single or multiple chromosomes; (ii) merotelic attachment, which occurs when one kinetochore is attached to both mitotic spindle poles; and (iii) formation of multipolar/monopolar spindles. FISH in interphase is a rapid and reliable molecular cytogenetic approach for the targeted detection of aneuploidy [37]. The number of FISH signals per cell is counted using chromosome specific probes. This analysis is currently used for human prenatal diagnosis and for the rapid screening of cancer cell populations.

In the present study, we used this test to compare the segregation fidelity of horse chromosome 11, whose centromere is satellite-free, with that of chromosome 13, which has a similar size and the centromeric function associated to satellite DNA [30]. Interphase FISH experiments were carried out with two BAC probes derived from the horse CHORI-241 BAC library [38] and specific for ECA11 and ECA13, respectively (CHORI241-21D14, chr11: 27,639,936–27,829,952 and CHORI241-22C1, chr13: 7,346,775–7,544,907). The localization of the ECA11 centromeric BAC is shown in Figure 1a. Since the centromere of chromosome 13 contains long stretches of satellite DNA and no single copy sequence co-localizes with the centromere, for the identification of this chromosome, we used a BAC clone mapping on the short arm close to the centromere (Figure 1b).

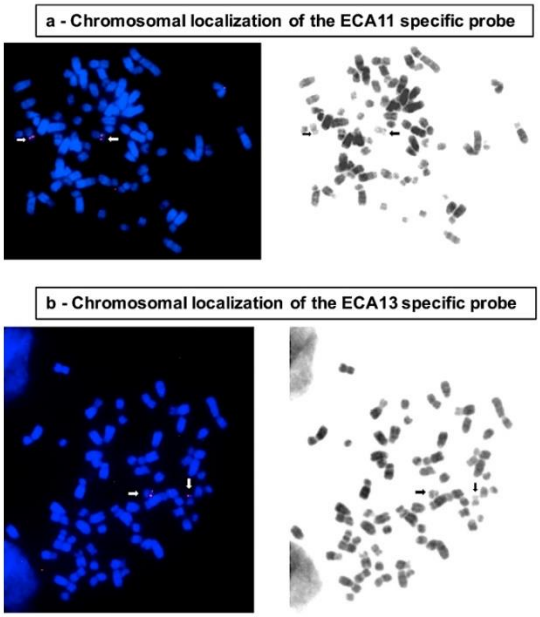


Figure 1. Chromosomal localization by fluorescence in situ hybridization (FISH) of the bacterial artificial chromosome (BAC) probes used. (**a**) FISH with the ECA11 specific BAC probe. Chromosome 11 is marked by the arrows. (**b**) FISH with the ECA13 specific BAC probe. Chromosome 13 is marked by the arrows. Probe localization is marked by the red FISH signals in the images on the left. Chromosomes are counterstained with 4′, 6′-Diamidino-2-phenylindole hydrochloride (DAPI). Reverse DAPI banding is shown in the images on the right.

To evaluate the frequency of aneuploid cells, the number of fluorescence spots corresponding to each probe was counted by visual inspection on interphase nuclei; numbers differing from two indicated an aneuploid condition. Since the frequency of chromosome specific aneuploidy and micronuclei was expected to be low, even minor differences in the efficiency or the stability of different fluorophores would impair the results. Therefore, to minimize experimental error, separate FISH experiments using the same fluorophore (Cy3) were carried out with probes for chromosome 11 and chromosome 13. Figure 2 shows examples of the results of FISH experiments. The number of signals was then counted on interphase nuclei hybridized with the chromosome 11 or with the chromosome 13 specific probe.

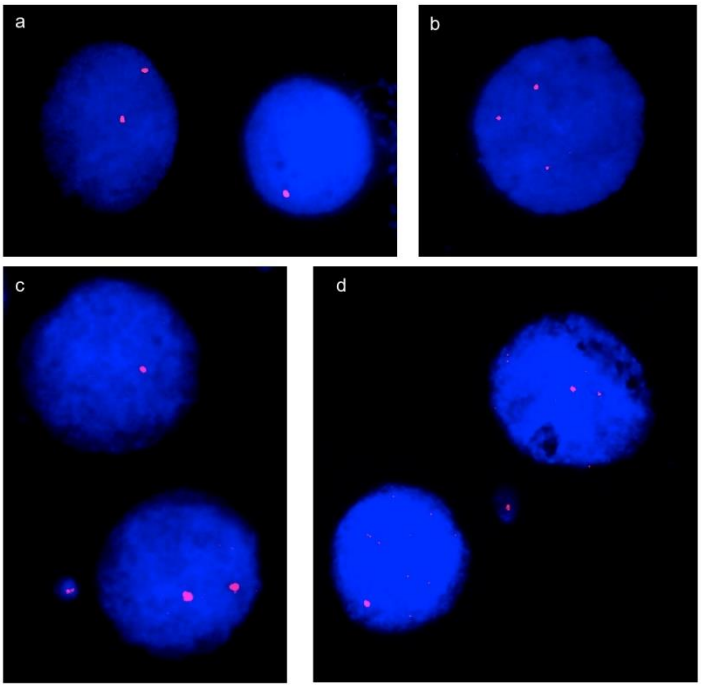


Figure 2. FISH on interphase nuclei and on micronuclei. (**a**) A nucleus, disomic for chromosome 11, is flanked by a monosomic one. (**b**) Nucleus trisomic for chromosome 13. (**c**) Bi-nucleated Cyt-B treated cell with a disomic nucleus, a monosomic nucleus, and a chromosome 11 positive micronucleus. (**d**) Bi-nucleated Cyt-B treated cell with a disomic nucleus, a monosomic nucleus, and a chromosome 11 positive micronucleus.

Samples of 4000 interphase nuclei from three replicate experiments for chromosome 11 and 13, respectively, were analyzed under control conditions. The results of the analysis are reported in Table 1. Statistical analysis of the difference in the number of nuclei aneuploid for chromosome 11 or 13 was performed with the χ^2 test. The mitotic behavior of the two chromosomes was comparable, because the difference between the number of cells aneuploid for chromosome 11 (1,90) and for chromosome 13 (1,98) was not statistically significant, and the standard error of the replicates is reported in the table. Table 1 also shows that the percentages of nullisomic, monosomic, and trisomic nuclei for the two chromosomes were similar. A few nuclei with four signals were observed: 0.40% and 0.36% for chromosome 11 and for chromosome 13, respectively. However, since this assay does not allow determination of whether these nuclei correspond to tetrasomic or tetraploid conditions, they were not included in the analysis of aneuploid cells.

Table 1. Interphase FISH under normal conditions.

Chromosome	Total Number of Nuclei	Diploid Nuclei (%)	Aneuploid Nuclei (% ± SE)	Type of Aneuploidy	Number (%)
ECA11	4000	3924 (98,10)	76 (1.90 ± 0.196)	*nullisomy*	5 (0,12)
				monosomy	30 (0,75)
				trisomy	41 (1,02)
ECA13	4000	3921 (98,03)	79 (1.98 ± 0.329)	*nullisomy*	8 (0,20)
				monosomy	32 (0,80)
				trisomy	39 (0,98)

Horse fibroblasts were then treated with 100 nM nocodazole for 48 h, and the number of aneuploid nuclei was counted in samples of about 1000 nuclei from two replicates for chromosomes 11 and 13, respectively, under control and stress conditions (Table 2). Nocodazole is a well described antimitotic agent that interferes with the function of spindle and cytoplasmic microtubules by binding to tubulin. The use of this drug was aimed at amplifying the difference, if any, in segregation fidelity between the satellite-free ECA11 centromere and the satellite-based ECA13 centromere and also at identifying possible differences in their sensitivity to conditions perturbing cell division.

A highly significant increase in the number of aneuploid nuclei was observed for both chromosomes after drug treatment (χ^2 test, p-value 3×10^{-3} and 1×10^{-4}, respectively). However, the difference in the number of nuclei aneuploid for chromosomes 11 or 13, as assayed by the χ^2 test, was not significant either under control conditions or after drug treatment. This result indicates that satellite-free and satellite based horse centromeres are equally sensitive to this antimitotic drug.

Table 2. Interphase FISH under mitotic stress.

Chromosome	Total Number of Nuclei	Diploid Nuclei (%)	Aneuploid Nuclei (% ± SE)	Type of Aneuploidy	Number (%)
Control					
ECA11	1107	1084 (97,92)	23 (2.08 ± 0.014)	*nullisomy*	0
				monosomy	11 (0,99)
				trisomy	12 (1,08)
ECA13	1117	1098 (98,30)	19 (1.70 ± 0.057)	*nullisomy*	0
				monosomy	11 (0,98)
				trisomy	8 (0,72)
Nocodazole 100 nM					
ECA11	1045	1000 (95,70)	45 (4.31 ± 0.085)	*nullisomy*	0
				monosomy	13 (1,24)
				trisomy	32 (3,01)
ECA13	1023	976 (95,40)	47 (4.60 ± 0.042)	*nullisomy*	0
				monosomy	12 (1,17)
				trisomy	35 (3,42)

3.2. Micronucleus Assay and Cytokinesis-Block MicroNucleus (CBMN) Assay

The micronucleus assay is a mutagenic test for the detection of small membrane-bound DNA fragments (i.e., micronuclei in the cytoplasm of interphase cells) [39,40]. Centric and acentric chromosome fragments as well as whole chromosomes unable to migrate to one pole during anaphase can be included into micronuclei. Two mechanisms, chromosome breakage and disturbance of chromosome segregation, may lead to the formation of micronuclei.

To test whether the formation of micronuclei is influenced by the presence/absence of centromeric satellite DNA, we performed the micronucleus assay using the chromosome specific BAC probes described above. Also, in this assay, to minimize the experimental error, micronuclei containing chromosome 11 or chromosome 13 were counted in independent experiments in which both probes were red labeled with Cy3.

In the same samples of 4000 nuclei previously analyzed for aneuploidy, we calculated the frequency of spontaneously occurring micronuclei. The results are reported in Table 3. The frequencies of spontaneously occurring micronuclei (1.6 and 1.4%) were comparable to those previously observed in human primary fibroblasts [41,42], and no significant difference in the number of chromosome 11 or chromosome 13 containing micronuclei was observed, as determined by the Fisher's exact test. The standard error calculated in the three replicates is reported in the table.

Table 3. Spontaneously occurring micronuclei under control conditions without cytokinesis-block.

Chromosome	Total Number of Nuclei	Total Number of Micronuclei (% ± SE)	Micronuclei Containing the Chromosome (% ± SE)	Micronuclei not Containing the Chromosome (%)
ECA11	4000	63 (1.6 ± 0.098)	7 (11.1 ± 1.143)	56 (88,9)
ECA13	4000	55 (1.4 ± 0.127)	6 (10.9 ± 2.390)	49 (89,1)

As for the aneuploidy assay, we then analysed micronuclei in cells treated with nocodazole. To make sure that the cells being scored have completed mitosis during nocodazole treatment we performed the Cytokinesis-Block MicroNucleus (CBMN) assay [43,44]. In this assay scoring is specifically restricted to once-divided bi-nucleated (BN) cells, after blocking cytokinesis with cytochalasin-B (Cyt-B), an inhibitor of microfilament ring assembly required for the completion of cytokinesis.

Cells were treated with 100 nM nocodazole for 48 h. During the last 24 h Cyt-B was added to the cultures. Two replicates were performed for each condition and in each experiment a total of 2.000 BN cells was scored. The results of the CBMN assay are reported in Table 4.

Table 4. Cytokinesis-Block MicroNucleus assay under mitotic stress.

Chromosome	Total Number of BN Cells	Total Number of Micronuclei (% ± SE)	Micronuclei Containing the Chromosome (% ± SE)	Micronuclei not Containing the Chromosome (%)
Control				
ECA11	2000	31 (1.6 ± 0.002)	6 (19.4 ± 0.019)	25 (80,6)
ECA13	2000	32 (1.6 ± 0.000)	6 (18.8 ± 0.000)	26 (81,2)
Nocodazole 100 nM				
ECA11	2000	75 (3.8 ± 0.004)	11 (14.7 ± 0.0004)	64 (85,3)
ECA13	2000	65 (3.3 ± 0.009)	10 (15.4 ± 0.0180)	55 (84,6)

After nocodazole exposure, a highly significant increase in the total number of micronuclei was observed (χ^2 test, p-value 3×10^{-5} and 9×10^{-4}, respectively), thus demonstrating the effectiveness of the treatment. When the numbers of chromosome 11 and chromosome 13 containing micronuclei were compared, no difference between the two chromosomes was observed under control conditions or after exposure to nocodazole (due to the small number of chromosome positive micronuclei observed, the Fisher's exact test was used in this case). Figure 2c,d show examples of bi-nucleated cells containing micronuclei positive for chromosome 11 and chromosome 13, respectively.

The results of the micronucleus assay confirmed that the mitotic behavior of the two chromosomes was comparable and that the absence of satellite DNA at the centromere did not affect its sensitivity to the spindle inhibitor nocodazole. These results are in agreement with those obtained by interphase

nuclei analysis. Thus, two independent molecular-cytogenetic approaches demonstrated that, in the horse fibroblast system, segregation fidelity is not influenced by the presence of satellite DNA sequences at the centromere.

3.3. Chromatid Cohesion Analysis

Literature data report that aurora B-kinase, which regulates chromosome-spindle attachments, mislocalizes at pathological human satellite-free centromeres [45,46]. Aurora B is part of a protein network, the chromosomal passenger complex (CPC), which regulates proper chromatid segregation. In the hypothesis that, in the horse system, the absence of satellite DNA at the centromere could affect CPC function, thus perturbing chromatid cohesion during cell division, we analyzed samples of 200 metaphase spreads from control fibroblast cultures and from cultures treated with 100 nM nocodazole. Again, FISH with the chromosome 11 and the chromosome 13 probes was performed in order to compare the behavior of the two centromeres. We never observed chromatid cohesion defects for any chromosome in either control samples or under mitotic stress conditions.

4. Conclusions

The species belonging to the genus *Equus* are characterized by karyotypes where chromosomes with canonical satellite-based centromeres are present together with chromosomes with satellite-free centromeres [6,14,29–34]. This exceptional biological model offers the opportunity to directly investigate, in a natural environment, the behavior of these differently organized centromeres.

Here, we analyzed the mitotic behavior of the satellite-free centromere of ECA11 and compared it with that of the canonical, satellite-based, ECA13 centromere. Our results demonstrated that the segregation accuracy of these two chromosomes is similar, thus suggesting that satellite DNA is dispensable for transmission fidelity.

Sequence analysis of the centromere of horse chromosome 11 showed that no motif resembling a CENP-B box is present (unpublished results). Consequently, our results suggest that, in the horse system, the CENP-B protein does not play a central role in chromosome segregation.

The function of satellite DNA at the centromere is a matter of debate, literature data suggesting that centromeric and/or pericentromeric repeated DNA sequences create the chromatin environment needed for sister chromatid cohesion and for kinetochore recruitment [5,12,13,47]. Indeed, the large majority of vertebrate centromeres contain highly repeated DNA sequences [1,2]. The widespread presence of repeated DNA at natural centromeres suggests that there is a positive selection for this kind of arrangement. The results presented here demonstrate that, at least in the horse, the absence of satellite DNA at the centromere does not affect chromosome segregation. Therefore, we postulate that satellite DNA may play roles other than the maintenance of segregation fidelity. One possibility based on our previous results on centromere sliding [33,34] is that satellite DNA may contribute to constrain the borders of the functional centromeric domain within non-coding genomic regions.

Author Contributions: A.R. carried out the experiments and contributed to manuscript drafting and figure preparation. M.B. and A.M. contributed to cell culturing and chromosome preparation. F.M.P. contributed to cytogenetic analysis. S.G.N. carried out probe isolation and amplification and contributed to data analysis. E.G. designed and supervised experiments and participated to manuscript preparation. E.R. conceived the study, supervised the experiments and wrote the manuscript.

Funding: This work was supported by grants from: Consiglio Nazionale delle Ricerche (CNR Progetto Bandiera Epigenomica); Ministero dell'Istruzione dell'Università e della Ricerca (MIUR-PRIN); Dipartimenti di Eccellenza Program (2018–2022)—Dept. of Biology and Biotechnology "L. Spallanzani," University of Pavia (to A.R., M.B., F.M.P., S.G.N., E.G. and E.R.).

Acknowledgments: We thank Antonella Lisa, Computational Analysis, Institute of Molecular Genetics, National Research Council, Pavia, for the helpful suggestions in performing the statistical analysis of the results.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. 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] [PubMed]
2. Plohl, M.; Mestrovic, N.; Mravinac, B. Centromere identity from the DNA point of view. *Chromosoma* **2014**, *123*, 313–325. [CrossRef] [PubMed]
3. Voullaire, L.E.; Slater, H.R.; Petrovic, V.; Choo, K.H. A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: Activation of a latent centromere? *Am. J. Hum. Genet.* **1993**, *52*, 1153–1163. [PubMed]
4. Choo, K.H. Centromere DNA dynamics: Latent centromeres and neocentromere formation. *Am. J. Hum. Genet.* **1997**, *61*, 1225–1233. [CrossRef] [PubMed]
5. 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]
6. 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]
7. 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]
8. 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]
9. Gong, Z.; Wu, Y.; Koblizkova, A.; Torres, G.A.; Wang, K.; Iovene, M.; Neumann, P.; Zhang, W.; Novak, P.; Buell, C.R.; et al. Repeatless and repeat-based centromeres in potato: Implications for centromere evolution. *Plant Cell* **2012**, *24*, 3559–3574. [CrossRef]
10. Giulotto, E.; Raimondi, E.; Sullivan, K.F. The Unique DNA Sequences Underlying Equine Centromeres. *Prog. Mol. Subcell. Biol.* **2017**, *56*, 337–354. [CrossRef]
11. Quenet, D.; Dalal, Y. A long non-coding RNA is required for targeting centromeric protein A to the human centromere. *Elife* **2014**, *3*, e03254. [CrossRef] [PubMed]
12. Rosic, S.; Kohler, F.; Erhardt, S. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J. Cell Biol.* **2014**, *207*, 335–349. [CrossRef] [PubMed]
13. 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]
14. 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]
15. Masumoto, H.; Masukata, H.; Muro, Y.; Nozaki, N.; Okazaki, T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J. Cell Biol.* **1989**, *109*, 1963–1973. [CrossRef] [PubMed]
16. Miga, K.H.; Newton, Y.; Jain, M.; Altemose, N.; Willard, H.F.; Kent, W.J. Centromere reference models for human chromosomes X and Y satellite arrays. *Genome Res.* **2014**, *24*, 697–707. [CrossRef] [PubMed]
17. 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]
18. Drinnenberg, I.A.; Henikoff, S.; Malik, H.S. Evolutionary Turnover of Kinetochore Proteins: A Ship of Theseus? *Trends Cell. Biol.* **2016**, *26*, 498–510. [CrossRef]
19. Farr, C.J.; Bayne, R.A.; Kipling, D.; Mills, W.; Critcher, R.; Cooke, H.J. Generation of a human X-derived minichromosome using telomere-associated chromosome fragmentation. *Embo J.* **1995**, *14*, 5444–5454. [CrossRef]
20. Heller, R.; Brown, K.E.; Burgtorf, C.; Brown, W.R. Mini-chromosomes derived from the human Y chromosome by telomere directed chromosome breakage. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 7125–7130. [CrossRef]

21. Raimondi, E.; Balzaretti, M.; Moralli, D.; Vagnarelli, P.; Tredici, F.; Bensi, M.; De Carli, L. Gene targeting to the centromeric DNA of a human minichromosome. *Hum. Gene Ther.* **1996**, *7*, 1103–1109. [CrossRef] [PubMed]
22. Harrington, J.J.; Van Bokkelen, G.; Mays, R.W.; Gustashaw, K.; Willard, H.F. Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat. Genet.* **1997**, *15*, 345–355. [CrossRef] [PubMed]
23. Masumoto, H.; Nakano, M.; Ohzeki, J. The role of CENP-B and alpha-satellite DNA: De novo assembly and epigenetic maintenance of human centromeres. *Chromosome Res.* **2004**, *12*, 543–556. [CrossRef] [PubMed]
24. Henikoff, J.G.; Thakur, J.; Kasinathan, S.; Henikoff, S. A unique chromatin complex occupies young alpha-satellite arrays of human centromeres. *Sci. Adv.* **2015**, *1*. [CrossRef] [PubMed]
25. Nakano, M.; Cardinale, S.; Noskov, V.N.; Gassmann, R.; Vagnarelli, P.; Kandels-Lewis, S.; Larionov, V.; Earnshaw, W.C.; Masumoto, H. Inactivation of a human kinetochore by specific targeting of chromatin modifiers. *Dev. Cell* **2008**, *14*, 507–522. [CrossRef] [PubMed]
26. Ohzeki, J.; Bergmann, J.H.; Kouprina, N.; Noskov, V.N.; Nakano, M.; Kimura, H.; Earnshaw, W.C.; Larionov, V.; Masumoto, H. Breaking the HAC Barrier: Histone H3K9 acetyl/methyl balance regulates CENP-A assembly. *Embo J.* **2012**, *31*, 2391–2402. [CrossRef] [PubMed]
27. Molina, O.; Vargiu, G.; Abad, M.A.; Zhiteneva, A.; Jeyaprakash, A.A.; Masumoto, H.; Kouprina, N.; Larionov, V.; Earnshaw, W.C. Epigenetic engineering reveals a balance between histone modifications and transcription in kinetochore maintenance. *Nat. Commun.* **2016**, *7*, 13334. [CrossRef] [PubMed]
28. Sullivan, B.A.; Karpen, G.H. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.* **2004**, *11*, 1076–1083. [CrossRef]
29. 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]
30. Piras, F.M.; Nergadze, S.G.; Magnani, E.; Bertoni, L.; Attolini, C.; Khoriauli, L.; Raimondi, E.; Giulotto, E. Uncoupling of satellite DNA and centromeric function in the genus Equus. *PLoS Genet.* **2010**, *6*, e1000845. [CrossRef]
31. Raimondi, E.; Piras, F.M.; Nergadze, S.G.; Di Meo, G.P.; Ruiz-Herrera, A.; Ponsa, M.; Ianuzzi, L.; Giulotto, E. Polymorphic organization of constitutive heterochromatin in Equus asinus (2n = 62) chromosome 1. *Hereditas* **2011**, *148*, 110–113. [CrossRef] [PubMed]
32. Nergadze, S.G.; Belloni, E.; Piras, F.M.; Khoriauli, 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] [PubMed]
33. 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] [PubMed]
34. 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]
35. Montefalcone, G.; Tempesta, S.; Rocchi, M.; Archidiacono, N. Centromere repositioning. *Genome Res.* **1999**, *9*, 1184–1188. [CrossRef] [PubMed]
36. 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]
37. Faas, B.H.; Cirigliano, V.; Bui, T.H. Rapid methods for targeted prenatal diagnosis of common chromosome aneuploidies. *Semin. Fetal Neonatal Med.* **2011**, *16*, 81–87. [CrossRef]
38. Leeb, T.; Vogl, C.; Zhu, B.; de Jong, P.J.; Binns, M.M.; Chowdhary, B.P.; Scharfe, M.; Jarek, M.; Nordsiek, G.; Schrader, F.; et al. A human-horse comparative map based on equine BAC end sequences. *Genomics* **2006**, *87*, 772–776. [CrossRef]
39. Kirsch-Volders, M.; Elhajouji, A.; Cundari, E.; Van Hummelen, P. The in vitro micronucleus test: A multi-endpoint assay to detect simultaneously mitotic delay, apoptosis, chromosome breakage, chromosome loss and non-disjunction. *Mutat. Res.* **1997**, *392*, 19–30. [CrossRef]
40. Fenech, M. The in vitro micronucleus technique. *Mutat. Res.* **2000**, *455*, 81–95. [CrossRef]

41. Rudd, N.L.; Hoar, D.I.; Williams, S.E.; Hennig, U.G. Genotype and the cryopreservation process affect the levels of aneuploidy and chromosome breakage in cultured human fibroblasts. *Genome* **1989**, *32*, 196–202. [CrossRef] [PubMed]
42. Schmidt-Preuss, U.; Roser, M.; Weichenthal, M.; Rudiger, H.W. Elevated frequencies of micronuclei in cultured fibroblasts after freezing and thawing. *Mutat. Res.* **1990**, *241*, 279–282. [CrossRef]
43. Kirsch-Volders, M.; Sofuni, T.; Aardema, M.; Albertini, S.; Eastmond, D.; Fenech, M.; Ishidate, M.; Kirchner, S.; Lorge, E.; Morita, T.; et al. Report from the in vitro micronucleus assay working group. *Mutat. Res.* **2003**, *540*, 153–163. [CrossRef]
44. Fenech, M. Cytokinesis-block micronucleus cytome assay. *Nat. Protoc.* **2007**, *2*, 1084–1104. [CrossRef] [PubMed]
45. Liu, D.; Vader, G.; Vromans, M.J.; Lampson, M.A.; Lens, S.M. Sensing chromosome bi-orientation by spatial separation of aurora B kinase from kinetochore substrates. *Science* **2009**, *323*, 1350–1353. [CrossRef] [PubMed]
46. Bassett, E.A.; Wood, S.; Salimian, K.J.; Ajith, S.; Foltz, D.R.; Black, B.E. Epigenetic centromere specification directs aurora B accumulation but is insufficient to efficiently correct mitotic errors. *J. Cell Biol.* **2010**, *190*, 177–185. [CrossRef]
47. Shang, W.H.; Hori, T.; Martins, N.M.; Toyoda, A.; Misu, S.; Monma, N.; Hiratani, I.; Maeshima, K.; Ikeo, K.; Fujiyama, A.; et al. Chromosome engineering allows the efficient isolation of vertebrate neocentromeres. *Dev. Cell* **2013**, *24*, 635–648. [CrossRef]