

# Centromere Evolution: Digging into Mammalian Primary Constriction

Giuliana Giannuzzi\*, Claudia Rita Catacchio\* and Mario Ventura  
*University of Bari,  
Italy*

## 1. Introduction

Mammalian cell division requires even and complete distribution of chromosome complement in daughter cells. At this cell cycle stage, segregation fidelity is critical in order to prevent aneuploidy. Chromosomes attach by a proteinaceous bridge called kinetochore to the spindle apparatus (Rieder & Salmon, 1994). The chromosomal locus at which kinetochore is organized and chromatids align is represented by the centromere, historically and cytologically defined as *primary constriction*. Throughout the mammalian order, as well as more generally among all higher eukaryotes, a few distinctive features describe the centromeric chromatin, such as the incorporation of the centromeric protein CENP-A, a histone H3 variant (Sullivan, B. A. & Karpen, 2004; Sullivan, K. F. et al., 1994), and the CENP-B, an essential component of the mitotic chromosome scaffold (Masumoto et al., 1989).

Phylogenetic studies show that, in contrast to the high conservation of the chromosome segregation machinery, the primary sequence of centromeric DNA has undergone rapid evolution resulting strikingly different even among closely related mammalian species. These studies suggest that specialized chromatin structure is more critical in centromeric function rather than the presence of specific sequences (Torras-Llort et al., 2009), revealing a role of epigenetic mechanisms on centromeric identity and function (Karpen & Allshire, 1997).

In this chapter we review the variable DNA sequence that forms mammalian centromeres: the satellite DNA. The first pilot studies on mammal satellite DNA date back to 1960s and were accomplished on the mouse, guinea pig, and bovine genomes. Since then, plentiful analyses have been developed for the characterization of satellite DNA in several mammals, particularly in human. These studies enabled the formulation of different theories aiming to explain the rapid evolution of alphoid DNA in primates and clarify a possible role of these sequences in centromeric function. Therefore, this chapter provides an overview of the state of the art in the field of mammalian centromeric DNA organization and evolution.

## 2. Organization of centromeric satellite DNA in non-primate mammals

Mammalia is a class of air-breathing vertebrate animals, characterized by the possession of hair, three middle ear bones, a four-chambered heart, and mammary glands functional in

---

\* Equally contributing

mothers with young. The class Mammalia is divided into two subclasses: the Theria, comprising the infraclasses Eutheria (the placentals) and Metatheria (the marsupials), and the Prototheria, comprising the order Monotremata of mammalian species that lay eggs (Prasad et al., 2008).

The centromeres of all Therian species examined consist of long arrays of head-to-tail tandemly repeated DNA families, the satellite DNA. One exception to this general rule is the sequence of newly formed centromeres (the centromere of horse chromosome 11 is an example), which is devoid of satellite DNA, thus demonstrating that centromeres can stably function over million years and many generations in the absence of satellite DNA (Wade et al., 2009). However, the acquisition and maintenance of satellite DNA is an obligated fate for all mammalian centromeres, since all mature centromeres possess satellite sequences. (Piras et al., 2010; Wade et al., 2009).

Centromeric satellite DNA sequences have been characterized in almost all mammalian orders. Here, we report exclusively Therian mammal centromeric satellite, because it has never been isolated from any species of the subclass Prototheria. Recently, an attempt to isolate centromeric satellite DNA of the Prototherian mammal platypus (order Monotremata) failed, suggesting that in this species centromeres are not enriched in satellite sequences (Alkan et al., 2010). Studies of other Prototherians such as the echidna might elucidate whether the lack of satellite DNA at centromeres is a platypus feature or it is common in the entire subclass.

The paragraph reports the state of the art about the sequence and organization of centromeric satellite DNA in mammalian species of the Eutherian orders Rodentia, Lagomorpha, Cetartiodactyla, Perissodactyla, Carnivora, Chiroptera, Cingulata, and Proboscidea and of the infraclass Metatheria, including the orders Diprotodontia and Didelphimorphia. The description of centromeric satellite DNA in Primates, to which the human species belongs, has a dedicate section since primate satellite sequences are the best characterized, both structurally and functionally, and plentiful information and studies have been developed in this field.

Despite its conserved function, the centromeric satellite is extraordinarily variable in the repeat unit length, sequence, organization, and relative quantity in respect to the total DNA even among closely related mammalian species. This singularity is known as the "centromere paradox" (Henikoff et al., 2001). In fact, the repeat unit length of the mammalian centromeric satellites ranges from 7 bp in the red-necked wallaby to 2.3 kb in the domestic cattle (**Table 1**). Exception to the high variability is the CENP-B box, a 17 bp sequence motif that is shared and conserved in all centromeric satellite families involved in the centromeric function (Masumoto et al., 1989). However, although this apparent lack of any rule among the satellite sequences, there is an evolutionary conserved pattern of sequence arrangement and organization (Sunkel & Coelho, 1995).

The study of centromeric satellite in different mammalian orders has contributed to the knowledge of centromeric satellite organization and evolution. Each mammalian order has its peculiarity, with some orders best revealing a particular aspect of the centromeric satellite organization. For examples in few primates the centromere is arranged in higher-order repeat (HOR) and in Cetartiodactyla satellite DNA is distributed among chromosomes according to the position of the centromere (acrocentric vs. (sub)metacentric), resulting in a similar centromeric satellite composition and organization among chromosomes that share the same centromere position.

Infraclass	Superorder	Order	Species	Common name	Karyotype	Satellite	Repeat unit size	Chromosomes	Localization	CENP-B box
Eutheria	Euarchontoglires	Rodentia	<i>Mus musculus</i>	house mouse	2n=20	minor major MS3 MS4 Ymin	120 bp 234 bp 150 bp 300 bp 60-61 or 121 bp	autosomes, X autosomes, X - - Y	centromere pericentromere centromere centromere centromere	yes no yes yes no
			<i>Oryctolagus cuniculus</i>	domestic rabbit	2n=44	Rsat I Rsat II Rsat IIE	375 bp 585 bp 582 bp	subgroup of autosomes subgroup of autosomes subgroup of autosomes	centromere centromere centromere	no no no
		Cetartiodactyla	<i>Bos taurus</i>	domestic cattle	2n=60	I or 1.715 II or 1.722 III or 1.706 IV or 1.709	1400 bp 700 bp 2350 bp 3220 bp	autosomes few autosomes most autosomes few autosomes	- centromere centromere pericentromere	- - - -
				domestic sheep	2n=54	I or 1.714 II or 1.723	820 bp 700 bp	autosomes autosomes, X	centromere pericentromere	- -
	Laurasiatheria	Perissodactyla	<i>Equus caballus</i>	domestic dog	2n=78	major	221 bp	60 out of 64	centromere	-
				grey fox	2n=66	-	23 bp	58 out of 64	centromere	-
				common pipistrelle	2n=44	-	80 bp	-	centromere	-
				Kuhl's pipistrelle	2n=42	-	80 bp	-	centromere	-
			domestic horse	2n=64	ECAcons71 ECA3cons221 ECAcons70 ECAconsE0 ECA3cons451 ECAcons421+424	221 bp 221 bp 419 bp 450 bp 451 bp 475 bp	24 out of 64 60 out of 64 62 out of 64 60 out of 64 60 out of 64 60 out of 64	centromere centromere centromere centromere centromere centromere	yes yes yes yes yes yes	
			domestic dog	2n=78	CFA-SAT	237 bp	-	centromere	-	
Xenarthra	Proboscidea	<i>Dasypus novemcinctus</i>	armadillo	2n=64	DNCons173	173 bp	all chromosomes	centromere	yes	
			wallaroo	2n=56	LAFcons842-936	1220 bp	all chromosomes	centromere	yes	
Metatheria	Australidelphia	Diprotodontia	red kangaroo	2n=20	#6 satellites	-	6 different subgroups	centromere	-	
			red-necked wallaby	2n=16	Mrb-sat1 Mrb-sat23	342 bp 178 bp	X, Y all chromosomes	centromere and pericentromere centromere and pericentromere	no no	
			wallaroo	2n=16	Mrb-B29	7 bp	X, Y, 2p	centromere and pericentromere	no	
			short-tailed opossum	2n=18	MDCcons528	528 bp	8 out of 18	centromere	yes	

Table 1. Satellite families at the centromeres of non-primate mammals. - not available data.

We start this excursus of mammalian centromeres with the description of centromeric satellites isolated from mammals other than primates of the superorder Euarchontoglires. This superorder includes the grandorder Euarchonta and the clade Glires. The grandorder Euarchonta (“true ancestors”) comprises extant mammals belonging to the orders Primates, Scandentia and Dermoptera; the clade Glires includes the orders Rodentia and Lagomorpha. We illustrate the centromeric satellites of mammals of the orders Rodentia and Lagomorpha, since satellite DNA from mammals of the orders Scandentia and Dermoptera has never been characterized.

## 2.1 Rodentia

The satellite sequences of several Rodentia mammals have been investigated: *Mus musculus* (the house mouse), four *Acomys* species (the spiny mice) (Kunze et al., 1999), *Microtus chrotorrhinus* (the rock vole) (Modi, 1992), *Gerbillus nigeriae* (the Nigerian gerbil) (Volobouev et al., 1995), and a chromosome 2-specific centromeric DNA repeat from *Cricetulus griseus* (the Chinese hamster) (Fatyol et al., 1994). The *Microtus* and *Cricetulus* species show satellite sequences with the longest monomer size (>2.5 kb) reported until now among mammals. Here, we focus on the centromeric satellites of the house mouse, the most studied Rodentia species.

The centromere of all mouse telocentric chromosomes consists of two highly conserved, tandemly repeated sequences discovered by isopycnic centrifugation in CsCl gradient: the minor and major satellites (Pardue & Gall, 1970; Wong & Rattner, 1988). Minor satellite DNA represents 0.5–1% of the mouse genome and comprises an AT-rich 120 bp monomer that contains the CENP-B box motif (Yoda et al., 1992). The minor satellite array occupies 300–600 kb of the terminal region of all mouse telocentric chromosomes (the autosomes and the X chromosome) and coincides with the centric constriction and function (Kipling et al., 1991). It does not show evidence of higher-order organization. Major satellite DNA is a more abundant, AT-rich, 234 bp tandem repeat (Horz & Altenburger, 1981), organized into long arrays of 240–2000 kb (Vissel & Choo, 1989). It represents 5–10% of the total genomic DNA and shows pericentromeric localization, adjacent to the minor satellite. The major satellite monomer is internally repetitious and consists of eight 28-mer or 30-mer subrepeats, with a set of three related 9 bp ancestral sequence motifs (Horz & Altenburger, 1981). The minor satellite sequence is divided into two similar halves, which are further subdivided into quarter repeats. However, it does not have a primordial sequence similar to the 9 bp motif of the major satellite (**Fig. 1**).

The minor and major satellite sequences are highly conserved across the centromeres of all mouse telocentric chromosomes: the minor satellite monomers share a mean pairwise identity of 95% (Kalitsis et al., 2006); the major satellite monomers share a mean deviation from the consensus sequence of 3.9% (Vissel & Choo, 1989). This high degree of sequence conservation argues strongly for frequent recombinational exchanges between nonhomologous telocentric chromosomes driving sequence homogenization at mouse centromeres (Kalitsis et al., 2006; Vissel & Choo, 1989).

Besides these two AT-rich satellites, the major and minor satellites, two further GC-rich satellites have been identified at mouse centromeres: the Mouse Satellite 3 (MS3) and the

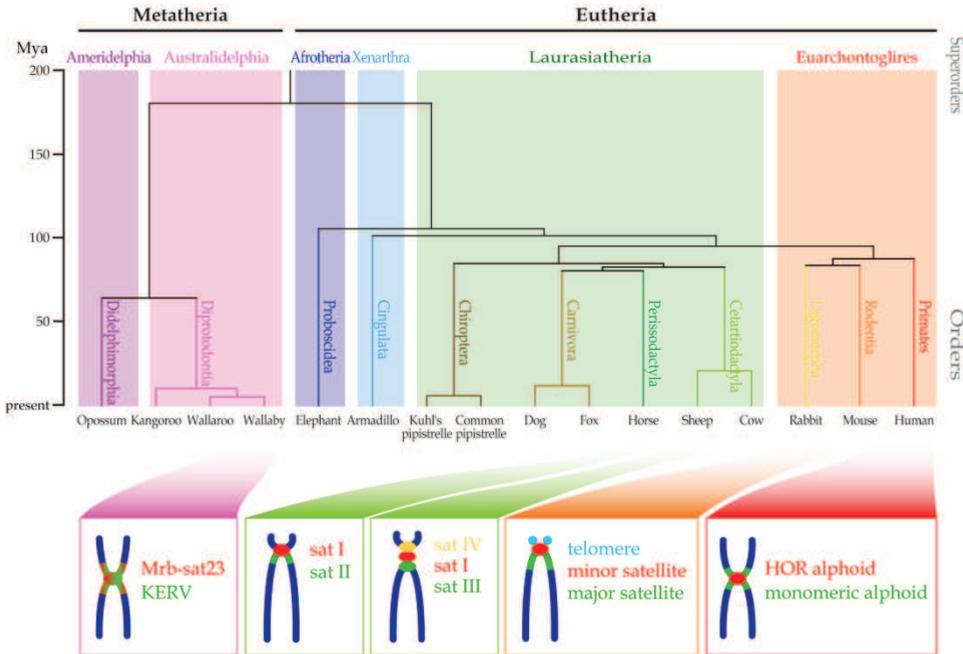


Fig. 1. Organization of centromeric satellites at mammalian centromeres. *top* Phylogenetic relationships among mammalian species discussed in the chapter with satellite DNA at their centromeres. *bottom* Spatial organization of satellite families at the centromere and pericentromere of -from right to left- human chromosomes, mouse telocentric chromosomes, cow autosomes, sheep autosomes, and wallaby autosomes. Chromosome arms are blue coloured. In the wallaby there is the co-existence of a satellite (Mrb-sat23) and a non-satellite element (KERV).

Mouse Satellite 4 (MS4). The MS3 monomer is 150 bp long and accounts for 2.2% of the total DNA; the MS4 monomer is 300 bp long and accounts for 1.6% of the total DNA. Both monomers contain the CENP-B box. Divergence between MS3 sequences is 0.7% due to single-nucleotide changes (Kuznetsova et al., 2005).

Mouse chromosome Y centromeric sequence has been recently investigated (Pertile et al., 2009), highlighting several features never described before at mouse centromeres. Mouse chromosome Y centromere has a chromosome-specific sequence and a unique multimeric HOR organization. It comprises a 90 kb array of an AT-rich and minor satellite-like tandem repeat (Ymin), with distant homology (76.8%) to mouse minor satellite. The Ymin satellite is closely associated with the kinetochore. It has a HOR sequence organization devoid of transposable elements, with a unit size of 2.3 (*Mus musculus molossinus*) or 1.6 (*Mus musculus domesticus*) kb, and a sequence identity among the repeated units of 99–100%. The HOR unit has a remarkably complex structure, consisting of an amalgam of highly diverged (<70% mean pairwise identity) monomers with a periodicity of 60–61 or 121 bp. The majority of the monomers form progressively larger and less diverged repeating HOR subunits, the largest comprising 2.4 copies of an 840 bp periodicity repeat (with 95% pairwise identity) that spans

the greater part of the HOR unit domain. The singularity of chromosome Y centromere argues for an intrachromosomal mode of sequence homogenization and an isolated evolution (Smith, 1976).

## 2.2 Lagomorpha

Ekes and collaborators described for the first time the satellite DNA families of a mammal belonging to the order Lagomorpha, the domestic rabbit, *Oryctolagus cuniculus* (Ekes et al., 2004). They found two major centromeric satellite DNA sequences, named Rsat I and Rsat II, which are not related to each other, and a divergent Rsat II-related subfamily, Rsat III. The Rsat I monomer has an average length of 375 bp, whereas repeat units Rsat II and Rsat III are ~585 bp long. These satellites do not provide a complete coverage of the rabbit complement, since seven autosome pairs and the sex chromosomes do not contain any of these satellites. Rsat I, Rsat II and Rsat III satellites are each distributed in variable amounts at the centromeres of a subgroup of rabbit chromosomes, with some chromosomes containing both Rsat I and Rsat II, or Rsat II and Rsat III. Part of Rsat I and Rsat II satellites shows a dimeric organization. However, further studies are required to isolate the rabbit sequences that constitute the centromeres devoid of the three known satellites, and to elucidate the higher-order repeat organization of rabbit satellites.

## 2.3 Cetartiodactyla

The Laurasiatheria is the evolutionarily closest superorder to the Euarchontoglires, and comprises the orders Cetartiodactyla, Perissodactyla, Carnivora, Chiroptera, Eulipotyphla, and Pholidota. The isolation of satellite DNA from a mammal of the last two orders has never been reported. Among mammals of the order Cetartiodactyla (cetaceans and even-toed ungulates), we describe the centromeric satellites isolated and characterized from two species of the family Bovidae: the domestic cattle (*Bos taurus*) of the subfamily Bovinae, and the domestic sheep (*Ovis aries*) of the subfamily Caprinae.

The cattle satellite I DNA (the 1.715 family) is a 1.4 kb tandem repeat that comprises 6-9% of the total genomic DNA (Kurnit et al., 1973). It constitutes the centromeric heterochromatin of all autosomes, but not of the sex chromosomes (Plucienniczak et al., 1982; Taparowsky & Gerbi, 1982) (**Fig. 1**). However, the primitive form of bovine X chromosome is acrocentric and has satellite I sequences at centromere (Chaves et al., 2005). Three additional bovine satellites, satellite II, III and IV, were localized at centromeric and pericentromeric regions of autosomes, respectively: satellite II is mostly localized at the autosomal centromeres; satellite III is present on most autosomes; satellite IV is present on less than half the autosomes (Kopecka et al., 1978; Kurnit et al., 1973). None of them is present on the sex chromosomes. The repeat unit of satellite III is 2,350 bp long and consists of two related and homogeneous 23-mer tandem subrepeats, the Pvu and the Sau motives, respectively (Pech et al., 1979). The three bovine satellites I, III and IV, or a subgroup of them, are organized on autosomes following always the same order: p-ter-sat IV-sat I-sat III-q (Chaves et al., 2003) (**Fig. 1**).

Two repetitive DNA families, satellite I (Buckland, 1983; Reisner & Bucholtz, 1983) and satellite II (Buckland, 1985) are the major components of the sheep centromeric and pericentromeric heterochromatin, respectively (Burkin et al., 1996; D'Aiuto et al., 1997) (**Fig.**

1). The sheep satellite I DNA (the 1.714 family) has a repeat unit of 820 bp and, as the bovine satellites, constitutes the centromeric heterochromatin of all autosomes, but not of the sex chromosomes (Buckland, 1983; Burkin et al., 1996; Chaves et al., 2000; Chaves et al., 2005). However, the amount and organization of satellite I DNA differ among the autosomes, with a lower amount at the centromeres of the banded chromosomes 1, 2, and 3, particularly of chromosome 1 (Burkin et al., 1996; D'Aiuto et al., 1997). The sheep satellite II DNA has a 700 bp monomer and constitutes the pericentromeric heterochromatin of all chromosomes with the exception of the Y chromosome (Burkin et al., 1996; D'Aiuto et al., 1997). Differently from sheep satellite I DNA, the satellite II family has a more variable chromosomal distribution: few acrocentric chromosomes are devoid and it is present at the centromeres of sheep metacentric and X chromosomes in large amounts (Burkin et al., 1996).

The bovine and ovine satellite I families show 70% sequence similarity and both consist of a degenerated 31 bp GC-rich tandem subrepeat (Novak, 1984; Reisner & Bucholtz, 1983). The presence of this 31 bp motif across the entire length of the satellite I repeat suggests that its present structure could have arisen from a tandemly amplification of an ancestral ~31 bp unit. The 31-mer motif sequence has been found also in other bovine satellites, like satellite III (Plucienniczak et al., 1982; Taparowsky & Gerbi, 1982), and in the deer, muntjac, and pronghorn centromeric satellites (Bogenberger et al., 1985; Denome et al., 1994; Lee, C. & Lin, 1996), arguing that the amplification of the ~31 bp unit may have occurred in their common ancestral.

## 2.4 Perissodactyla

The order Perissodactyla (odd-toed ungulates) comprises the family Equidae, with eight living species all belonging to the genus *Equus*: two horses (*E. caballus* and *E. przewalskii*), two Asiatic donkeys (*E. kiang* and *E. hemionus*), one African donkey (*E. asinus*), and three zebras (*E. grevyi*, *E. Burchelli*, and *E. zebra*). Despite the *Equus* species can be crossbred and diverged recently, sharing a common ancestor about 2–3 million years ago, their karyotypes differ extensively and their satellite DNA has evolved rapidly (Wijers et al., 1993). Moreover, during the evolution of the genus *Equus*, centromere repositioning, the shift along the chromosome of the centromere without structural chromosome rearrangements, has occurred frequently (Carbone et al., 2006). It implies that several evolutionary new centromeres and ancestral now inactive centromeres are present in the *Equus* karyotypes.

Several satellite families have been identified in the horse genome with centromeric localization, suggesting a great diversity and variability in structure and organization of horse centromeric sequences. The horse major satellite accounts for 5–10% of the total genome and has a repeat unit of 221 bp. It is localized at the centromeric regions of 30 pairs of chromosomes and is missing at the centromere of chromosomes 2 and 11, both submetacentric. Repeat units share a sequence identity of 90–100% and have no internal repeat structure (Piras et al., 2010; Wijers et al., 1993).

In 1995, another horse satellite family with a repeat unit of 23 bp was isolated and localized at the centromeres of acrocentric but not metacentric horse chromosomes (Broad et al., 1995b). This pattern reminds the satellite distribution in Cetartiodactyla that occurs according to the position of the centromere along the chromosome.

Two further horse satellite families with a repeat unit of ~80 bp were identified, with a likely centromeric localization (Broad et al., 1995a). Alkan and colleagues extracted six distinct

satellite consensus sequences in the *E. caballus* genome of 221, 221, 419, 450, 451, and 475 bp, respectively. Their FISH hybridization patterns included the centromeres of all or a part of horse chromosomes except chromosome 11 (Alkan et al., 2010). All horse centromeres have either one or more than one satellite whereas horse chromosome 11 is the only one lacking any satellite (Alkan et al., 2010; Piras et al., 2010; Wade et al., 2009).

## 2.5 Carnivora

Few studies have been accomplished on centromeric satellites in species belonging to the order Carnivora. In 1988 and 1989, the major centromeric satellite of the domestic dog (*Canis familiaris*) and of the grey fox (*Urocyon cinereoargenteus*), two species of the order Carnivora that diverged from a common ancestor 10–12 million years ago (Wayne et al., 1997), have been investigated in regard to their sequence and localization. Chromosomes of the domestic dog and grey fox are primarily acrocentric. The dog satellite monomer is 737 bp long and has a GC-content of 51%; the grey fox satellite monomer is 880 bp long and has a GC-content of 54% (Fanning, 1989). Recently dog fosmid clones containing satellite DNA were mapped to the centromere of a different subgroup of dog chromosomes. These heterogeneous patterns support the existence of a complex patchwork organization of satellites at dog centromeres, similar to horse centromeric sequence organization (Alkan et al., 2010).

## 2.6 Chiroptera

Bat genomes (order Chiroptera) are characterized by low DNA content, with a size approximately 50–87% the size of other eutherian genomes (Burton et al., 1989). Centromeric satellite DNA has been isolated from two bat species of the genus *Pipistrellus*, family Vespertilionidae, suborder Microchiroptera: the common pipistrelle (*Pipistrellus pipistrellus*) and the Kuhl's pipistrelle (*Pipistrellus kuhli*). Satellite DNA of the common bat represents approximately 3% of the whole genome and is organized in tandem repeats with a monomer size of 418 bp. The monomer units are highly similar, with a sequence identity of 95–100% and few base-substitutions randomly spread along the sequence. The common bat satellite has an AT-content of 62% and contains a putative CENP-B box motif. It is localized at the pericentromeric constitutive heterochromatin of all the autosomes and X chromosome, but it is absent from the Y chromosome (Barragan et al., 2003). *Pipistrellus kuhli* satellite represents approximately 5% of the total genomic DNA. The monomer unit is 1100 bp long and contains the CENP-B box as well as subrepeats, palindromes, and AT-rich tracts. The monomers group into two clusters (Fantaccione et al., 2005). Both pipistrelle satellites are absent in the genomes of other bat species analyzed, thus revealing that they might be species-specific.

## 2.7 Cingulata and Proboscidea

The study of centromeric satellite families in mammalian species belonging to the superorders Xenarthra and Afrotheria of the clade Eutheria has started very recently. In 2010, Alkan and colleagues included in their list of sequenced mammalian genomes to analyze in regard to centromeric satellite sequences, the armadillo (*Dasypus novemcinctus*), a species of the superorder Xenarthra, order Cingulata, and the African elephant (*Loxodonta*

*africana*), a species of the superorder Afrotheria, order Proboscidea. Using the RepeatNet algorithm, they extracted a 173 bp satellite consensus from the armadillo genome and a 1220 bp satellite consensus from the African elephant genome, and localized both satellites at the centromere of all chromosomes of the corresponding species (Alkan et al., 2010).

## 2.8 Diprotodontia and Didelphimorphia

The infraclass Metatheria comprises the marsupial mammals. Among the extant marsupials, species of the order Diprotodontia, family Macropodidae, and a species of the order Didelphimorphia were investigated in regard to their centromeric DNA organization.

The family Macropodidae (kangaroos, wallabies, and wallaroos), one of the largest in the infraclass Metatheria, exhibits an extraordinary karyotypic diversity and plasticity (Eldridge & Close, 1993; Hayman, 1990; Rofe, 1978). The Macropodidae centromere is a highly dynamic locus and has played a pivotal role in the karyotype evolution of this family, since all rearrangements, including centromere repositioning, fissions, fusions (Robertsonian translocations), pericentric inversions, and translocations, involved the centromeric loci as breakpoints (O'Neill et al., 2004; Rens et al., 2003).

Marsupial satellites are characterized by an uneven distribution among the centromeres of the different chromosomes. The centromeric satellite families isolated from each marsupial are also present in other marsupials analyzed, often in a different amount and with a different localization and distribution.

In 1981, Venolia and Peacock isolated a major satellite from the wallaroo (*Macropus robustus*) genome. It accounts for about 10% of the total DNA and localizes at the centromere of all chromosomes in different amounts and at the nucleolus organizer region (Venolia & Peacock, 1981). This satellite was localized in other *Macropus* species. In *M. rufus* and *M. rufogriseus* it is present mainly on the X chromosome at large, non-centromeric blocks, and in the region of the nucleolus organizer. In *M. rufogriseus* the satellite also occurs on the Y chromosome and in *M. rufus* at the centromere of four acrocentric autosomes.

Six different satellite DNA fractions have been isolated from the genome of the red kangaroo (*Macropus rufus*), each accounting for 1–3% of the total DNA. These satellites localize at the centromeres with each heterochromatic centromeric block differing in the amount and distribution of these satellites, as well as at interstitial regions and X chromosome telomeric heterochromatin (Elizur et al., 1982).

The red-necked wallaby (*Macropus rufogriseus*) karyotype has a distinctive feature: its chromosomes harbour an exceptional amount of centric and pericentric heterochromatin (Hayman & Martin, 1974), comprising almost 30% of the genome (Bulazel et al., 2006). They have unusually lengthened pericentromeric regions that are up to half the length of the chromosome, with the functional centromere restricted to a discrete point location within the larger region. In 2006, Bulazel and collaborators isolated three satellite families, named Mrb-sat1, Mrb-sat23, and Mrb-B29, from the red-necked wallaby genome. These satellites constitute the large centromeric and pericentromeric regions of the wallaby chromosomes and show a different chromosomal distribution. Mrb-sat23 constitutes the centromeric core as well as the large pericentric heterochromatic region of all chromosomes and is present in tandem arrays at all centromeres of most *Macropus* species (**Fig. 1**). In *M. rufogriseus*, Mrb-

sat23 experienced large-scale amplifications as it resides over the entire Y chromosome and is spread throughout the extensive X chromosome pericentromere. The presence of a CENP-B binding-competent domain on the Y of a marsupial suggests that ancestral mammalian sex chromosomes utilized CENP-B to differentiate centromere location and that the loss of CENP-B protein binding and CENP-B box DNA on the Y are derived when found within eutherian mammals. In this species the centromeric satellite constitution differs between the autosomes and the sex chromosomes: all autosomes have sequences of only Mrb-sat23, whereas the X and Y chromosomes harbour sequences of three satellites, Mrb-sat1, Mrb-sat23, and Mrb-B29, in different amounts (Bulazel et al., 2006).

Besides the satellite sequences, an active retroviral element, the Kangaroo Endogenous Retrovirus (KERV), is localized at the centromere and pericentromere in the genus *Macropus* (Fig. 1). It is considered thereof a major constituent of *Macropus* active and latent centromeres since it has undergone amplification at this locus (Ferreri et al., 2011; Ferreri et al., 2005; Ferreri et al., 2004; O'Neill et al., 1998). Particularly, in *M. rufogriseus* KERV is localized at the centromere of all autosomes, but it is absent or present in low copy number at the centromere of the sex chromosomes (Ferreri et al., 2011; Ferreri et al., 2004). Recently, Alkan and colleagues reported the centromeric satellite of the short-tailed opossum (*Monodelphis domestica*), a marsupial species of the order Didelphimorphia. They identified a 528 bp satellite, that is an LTR/ERV1 element, and localized it at the centromere of four homologous opossum chromosomes (Alkan et al., 2010). Such finding of a retroviral element at the centromeres of the short-tailed opossum, a marsupial belonging to a superorder different from the one of the Macropodidae marsupials, suggests that the use of a retroviral element as centromeric satellite might be ancestral in the infraclass Metatheria.

### 3. The centromeric alpha satellite in primates

The order Primates belongs to the subclass Theria, infraclass Eutheria (Fig 1). It includes an ancient group of which the size (the number of species) is still ambiguous. In fact, depending on whether some closely related groups are considered to be varieties of the same species or not, most taxonomic classifications refer to a range of 230–270 species.

The order Primates includes two suborders: Strepsirhini (including six families) and Haplorhini. Strepsirhini has a rhinarium, the tapetum lucidum, a bicornuate uterus, a toothcomb (with the exception of the Aye-aye), and a toilet-claw for grooming. On the other hand, the so-called “higher primates” compose the suborder Haplorhini, in turn divided in two main hyporders, the Anthropoidea, including the Platyrrhini (New World monkeys) and the Catarrhini (Old World monkeys and apes), and the Tarsiiformes (Goodman et al., 1989).

Primate predominant class of centromeric DNA is made up of long stretches of repeats consisting of 171 bp and AT-rich monomers, tandemly reiterated in a head-to-tail configuration (Vissel & Choo, 1987; Waye & Willard, 1987) commonly referred to as “alphoid satellite” DNA (Manuelidis, 1978). These sequences have been identified throughout the order Primates, including great apes, Old World and New World monkeys (Alves et al., 1994; Musich et al., 1980; Willard & Waye, 1987), with the exception of the suborder Strepsirhini (Lee, H. R. et al., 2011). Alpha satellite DNA is the most abundant repetitive DNA in all primate species studied, making up to 3–5% of each chromosome; it is pancentromeric in primate chromosomes and appears to be distinctive of the primate lineage (Vissel & Choo, 1987; Waye & Willard, 1987).

### 3.1 Great apes

At human centromeres alphoid DNA extends for ~250 kb up to ~5 Mb (Wevrick & Willard, 1989; Willard, 1990) and is known to exist in two forms designated (1) monomeric arrays and (2) higher-order repeats (HORs). In the monomeric arrays the 171 bp monomers lack further sequence structure and the head-to-tail configuration provides directionality to each satellite block. A higher-order repeat is rather composed of arrays in which a defined number of monomers has been homogenized as a unit which, in turn, is tandemly repeated many times to span several megabases, resulting in multiple copies of an alphoid multimer (Willard & Wayne, 1987). At human centromeres monomeric alphoid satellites flank HORs (Fig. 1). Sequence analysis revealed that monomers within a higher-order repeat unit share as low as 72% average pairwise sequence identity however, adjacent individual higher-order repeated units are 98–100% identical (Rudd & Willard, 2004). Higher-order alpha satellite within an array is extremely homogeneous and appears to be uninterrupted by other (non-satellite) DNA sequences (Tyler-Smith & Brown, 1987; Warburton & Willard, 1990). In contrast, monomeric alpha satellite is more heterogeneous in sequence and is extensively interspersed with non-alpha-satellite sequences such as transposable elements (Guy et al., 2003; Kazakov et al., 2003; Schueler et al., 2001).

The human higher-order repeating structures can shape alphoid subfamilies. These subfamilies are classified according to whether they are specific for a single chromosome or shared by a small group of chromosomes (Choo et al., 1991). As a consequence, (i) some centromeres contain only the chromosome-specific subfamily, while others possess several distinctive alphoid subfamilies (Choo et al., 1991) organized in discrete and homogenized physical blocks (Schueler et al., 2001); (ii) multiple domains located on one chromosome may belong to the same or to different suprachromosomal families (Choo et al., 1991). Human alpha satellite sequences have been historically grouped into five suprachromosomal families (SFs) initially according to the higher-order repeat unit length, revealed by restriction site periodicity, and then founded on sequence-based phylogenetic analyses (Alexandrov, I. et al., 2001; Alexandrov, I. A. et al., 1993; Iurov Iu et al., 1988) (Table 2). Human SFs 1–3 likely derived from an ancestral sequence by interchromosomal exchange (Waye & Willard, 1986). The reconstruction of the ancestral monomer sequence revealed that it originated the actual two phylogenetic homology groups of monomers: J1-D2-W4-W5 belong to group A, J2-D1-W1-W2-W3 compose group B (Alexandrov, I. A. et al., 1993). Moreover, there are several subsets that appear to form two more homogenous families, which are characterized by an array of equally related monomers (SF 4) and an irregular alternation of two different types of monomers (SF 5) (Alexandrov, I. et al., 2001; Alexandrov, I. A. et al., 1993). SF 4 consensus monomer, M1, is closely related to D2 and W4 monomeric types of SF 2 and SF 3; also, consensus sequences R1 and R2 (SF 5) clearly belong to subset A and B, respectively. SF 4 is related both structurally and in sequence to the African green monkey alpha component, of which the consensus sequence is closely associated to the phylogenetic homology group A; besides, consensus sequences derived from groups A and B resulted to be nearly identical to R1 and R2. This shows that the amplification of the ancestral sequences that gave rise to the two homology groupings of alphoid monomers happened at the very beginning of primate evolution and that SF 5 consensus monomers may represent the ancestral form of primate A-B satellite. Among the primate lineages, the W1-W5 pentamers and D1-D2 dimers are present in gorilla, orangutan, chimpanzee, and human, while J1-J2 sequences are present in gorilla, chimpanzee, and human (Alexandrov, I. A. et al., 1993).

SF	Repetitive unit	Phylogenetic homology groups	Consensus monomers	Chromosomes
1	dimer	A-B	J1, J2	1,3,5,6,7,10,12,16,19
2	dimer	B-A	D1, D2	2,4,8,9,13,14,15,18,20,21,22
3	pentamer	B-B-B-A-A	W1,W2,W3,W4, W5	1,11,17,X
4	monomer	A	M1	5,7,13,14,15,19,21,22,Y
5	monomer	A-B	R1,R2	1,2,3,5,6,7,9,11,12,13,14,16,18,19,20,22

Table 2. Structure, periodicity, and localization of human suprachromosomal families. SF, suprachromosomal family; data from Alexandrov et al., 2001.

Among individuals, array length of a single satellite can be highly polymorphic. There is an extensive variation of centromere size between nonhomologous as well as homologous human chromosomes. For example, alpha-satellite array length of human chromosomes X and Y varies almost three times: in the case of chromosome X from 1380 to 3730 kb (Mahtani & Willard, 1990); in the case of chromosome Y from 285 to 1020 kb (Tyler-Smith & Brown, 1987).

Cross-hybridization studies of alphoid sequences among great apes show that the higher-order sequence organization of the chromosome X alpha-satellite subset has been conserved among closely related species (orthologous evolution) (Durfy & Willard, 1990). However, similar relationships were not found for the other subfamilies. In fact, experimental evidence suggests that the majority of human-derived chromosome-specific alphoid satellite DNA probes does not recognize orthologous chromosomes in great apes (Samonte et al., 1997).

### 3.2 Catarrhini

The first species in which primate centromeric sequence was identified is the African green monkey (AGM) (*Chlorocebus sabaeus*). The AGM is an Old World monkey, belonging to the family Cercopithecidae, genus *Chlorocebus*. In 1971, Maio and colleagues identified the so-called component alpha DNA as the homogenous fraction composing the 20% of the total AGM genomic DNA, showing a behaviour similar to the mouse satellite DNA in renaturation kinetics experiments (Maio, 1971). Furthermore, two other Cercopithecidae species have been investigated in their centromeric main component: the Rhesus monkey (*Macaca mulatta*) and the baboon (genus *Papio*). Both their karyotypes show 20 pairs of autosomes plus the two sex chromosomes.

In the three species examined, the centromeric satellite DNA is the most abundant repetitive sequence, since it comprises up to 24% of the AGM and from 8 to 10% of the Rhesus monkey and baboon genomes (Musich et al., 1980). Besides, in the AGM genome, the tandemly repeated centromeric satellite monomeric unit is 172 bp long, while its length is doubled to 340 bp in the baboon and the nucleotide divergence among the satellite sequences of the two species is ~10% (Singer & Donehower, 1979) (Fig. 2). At Rhesus monkey centromeres, the underlying repetitive alphoid DNA is a 343 bp dimer (Fig. 2). The similarity between the two 171 and 172 bp component monomers is as much as 70%. Nevertheless, few high-identity regions have been detected, conserved not only between the two Rhesus monomers but also among all known Cercopithecidae monomers. Homologies between the macaque consensus alphoid sequence and the baboon, AGM and human alphoid sequences are >98%, 81%, and <70%, respectively (Musich et al., 1980).

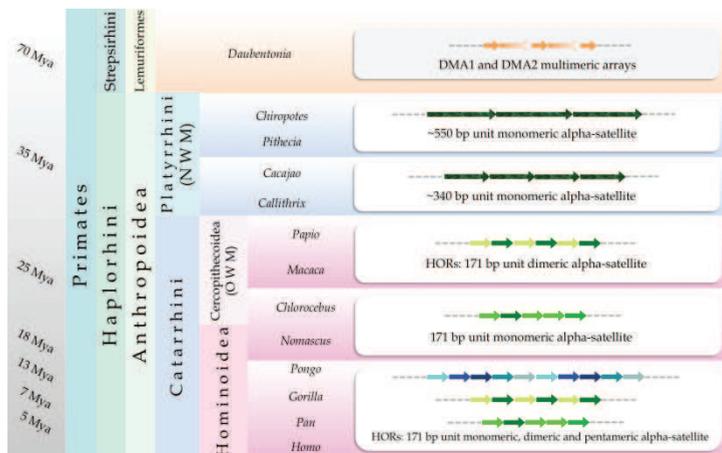


Fig. 2. Centromeric satellite organization in primates. *left* Simplified phylogenetic relationships among primate species. Estimated divergence times are reported in million of years. *right* Schematic organization of the alphoid arrays at primate centromeres. Arrows represent monomeric units. All Anthrooidea families have alphoid DNA at centromeres. For the genus *Daubentonia* (orange arrows) a model of the centromeric DMA1 and DMA2 satellites is proposed.

Among Catarrhini, a species belonging to the superfamily Hominoidea, family Hylobatidae, has been, lately, deeply characterized in its centromeric satellite: the white-cheeked gibbon (*Nomascus leucogenys*). In *N. leucogenys* alphoid monomers are 171 bp long and show four different hybridization patterns: telomeric, centromeric, telomeric-centromeric and Y-chromosome specific (Fig. 2). *N. leucogenys* is then the first primate species analysed in which alphoid DNA is detected at “ectopic” regions (telomeric and interstitial sites); though, the authors speculate that the different mapping pattern may be likely due to different sequence organizations rather than to site-specific sequence divergence. In fact, the gibbon karyotype is known to be extensively rearranged against the ancestral primate karyotype, leading to the hypothesis that alphoid sequences at telomeric level could represent the rest of the evolutionary fissions within centromeric breakpoints (Cellamare et al., 2009).

### 3.3 Platyrrhini

The group Platyrrhini includes five families: Callitrichidae, Cebidae, Aotidae, Pitheciidae, and Atelidae. Few centromeric sequences of several Platyrrhini species have been studied: three components of the family Pitheciidae (Alves et al., 1994) plus the common marmoset (Cellamare et al., 2009).

Centromeric sequence analysis on the black bearded saki (*Chiropotes satanas*) of the genus *Chiropotes* and the Rio Tapajós saki (*Pithecia irrorata*) of the genus *Pithecia* family Pitheciidae, have abundant amounts of satellite with monomeric unit of 539 and 559 bp, respectively. The difference in size between the two sequences is mainly due to 14 contiguous bases of which almost half consists in GA dinucleotides suggesting a strand slippage mechanism as a cause of the expansion.

The 539 bp alphoid satellite in *Chiropotes satanas* consists of four 170 bp subunits of which the third is incomplete. *C. satanas* alphoid DNA strongly hybridizes to *Pithecia*, while hybridization to the black-headed uakari (*Cacajao melanocephalus*) is much less intense, thus suggesting a satellite content loss in this species rather than a higher sequence divergence, since the divergence time between *Chiropotes* and *Cacajao* is thought to be about 5 Myr (millions years), while the divergence time between *Chiropotes/Cacajao* and *Pithecia* is 8 Myr (Schneider et al., 1993). Moreover, in *Cacajao melanocephalus* a substantial proportion of the satellite mass is composed by a 340 bp alphoid monomer, while the ~550 pb monomer constitutes a small subset. Nevertheless, it is likely that the ~550 bp monomer arose from an array of 340 bp repeats, leading to the conclusion that the ancestor of the *Pitheciini* harbored both structures (Alves et al., 1998) (Fig. 2).

Insights of the centromeric satellite among New World monkeys come from the genus *Callithrix*, more precisely from the common marmoset (*Callithrix jacchus*). Cellamare and collaborators characterized *C. jacchus* centromeric sequences by several cytogenetic and molecular approaches. Their analysis showed that, like for the above-mentioned New World monkeys, the alpha satellite monomer in this species is 340 bp long (Fig. 2). Thus, it is likely that two of these ancestral monomers fused and no further homogenization occurred between the two halves. The similarity between the first and second monomer is reported to be 40–50% and no results were obtained from sequence comparisons with great apes (human, chimpanzee, and gorilla), nor Old World monkeys (macaque) alphoid sequences (Cellamare et al., 2009).

### 3.4 Strepsirhini

Recently, the centromeric sequence of the aye-aye has been characterized (*Daubentonia madagascariensis*). Aye-aye centromeres are composed of two different, AT-rich, CENP-A-associated classes of repetitive DNA, termed DMA1 and DMA2, ~146 and ~268 bp long, respectively. DMA1 and DMA2 are often adjacent to one other at aye-aye centromeres and are completely unrelated to alphoid DNA in sequence composition though including a highly divergent CENP-B box. Moreover, their sequence analysis revealed significant homology values in the first 100 bp of both monomers, thus indicating that the two satellite classes share an evolutionary history (Lee, H. R. et al., 2011) (Fig. 2).

### 3.5 The evolution of alphoid DNA

Alphoid DNA in primates evolves rapidly (Mahtani & Willard, 1990; Wevrick & Willard, 1989); alpha-satellite DNA monomers evolve through a non-independent mechanism named *molecular drive* (Charlesworth et al., 1994; Smith, 1976; Stephan, 1986), a stochastic process in which mutations can accumulate, spread quickly through a repeat family, and fix in a population (Dover, 1982). In this evolutionary process mutations are homogenized throughout members of the satellite DNA family, and fixed within a species (Dover, 1982). Although Schindelhauer and Schwarz favour gene conversion as an explanation for both intrachromosomal and inter-homologue homogenization (Schindelhauer & Schwarz, 2002), only unequal crossover can explain the generation and maintenance of a multimeric higher-order repeat length, the extensive spread of sequence variants across megabases, and the rapid fall in sequence identity documented at the edge of the centromeric array.

Wu and Manuelidis proposed a two-step evolutionary process for the formation of tandem duplication arrays: after homology between two sequences is created, an unequal crossover might occur and thus result in dimer formation from divergent monomers; subsequently an amplification of the dimer into long tandem arrays might occur by subsequent unequal crossovers (Wu & Manuelidis, 1980). Eventually, a subset of monomers might be homogenized together to form HOR unit in which former monomers constitute subunits (Warburton & Willard, 1990; Willard & Wayne, 1987). As the number of tandem repeats increases so will frequency of unequal crossovers between them. By this mechanism, variant nucleotides can be spread along tandem repeats at a rate much faster than, and independent of, the mutation rate (Choo, 1990). The two major types of phylogenetically distinct alpha satellite DNA existing in great apes are a consequence of the homogenization process: multimeric, higher-order repeats of ~171 bp units form centrally located, chromosome-specific alphoid domains (class A and B alphoid monomers), flanked by domains of more heterogeneous monomeric alpha-satellite from which they have evolved (class A alphoid monomers) lacking any further organization.

The final outcome of *molecular drive* is concerted evolution, exhibiting higher identity between HORs within a species than with the orthologous array in other species (Dover, 1982; Rudd et al., 2006; Willard & Wayne, 1987), thus explaining the great diversity (24% divergence) seen between the human and chimpanzee alphoid regions on chromosomes 21 and 22. Given the neutral mutation rate of 0.13% per Myr and the estimated divergence of human and chimpanzee lineages 6-8 Mya (million years ago), those levels of divergence would have been absolutely unexpected. The extreme variability of HOR length among individuals may then be explained by the fact that unequal crossovers between higher-order repeat units will occur more frequently than between monomeric units because of the exceptionally high homology among HORs (Willard & Wayne, 1987). Besides, these mechanisms appear to proceed in a localized, short-range fashion that leads to the formation of large domains of sequence identity, rather than among intra- or interchromosomal repeats (Dover, 2002), thus resulting in the chromosome-specific alphoid subdomains. As a consequence, adjacent monomers display a higher degree of sequence similarity (Durfy & Willard, 1989; Roizes, 2006; Schindelbauer & Schwarz, 2002; Willard & Wayne, 1987) and monomers at array ends show lower identity due to the low efficiency of homogenization mechanisms at the edges of the satellite array (Schueler et al., 2005).

Unequal crossover events of alpha satellite arrays may represent both interchromosomal and intrachromosomal structural modifications. In the first case they will give rise to suprachromosomal families of higher-order alpha satellite (Alexandrov, I. A. et al., 1993; Wayne & Willard, 1986), while in the latter they will result in chromosome-specific arrays of higher-order alpha satellite (Durfy & Willard, 1989; Schindelbauer & Schwarz, 2002; Schueler et al., 2001; Willard & Wayne, 1987). In summary, the adjacent organization of higher-order and monomeric alpha satellite, as well as the fact that lower primates have only monomeric alpha satellite at their centromeres (Alves et al., 1994; Musich et al., 1980; Rosenberg et al., 1978), supports the hypothesis that higher-order alpha-satellite evolved from ancestral arrays of monomeric alpha-satellite and subsequently transposed to the centromeric regions of all great ape chromosomes (Alexandrov, I. et al., 2001; Kazakov et al., 2003; Schueler et al., 2001; Warburton et al., 1996). This is further confirmed by the age gradient revealed by L1 elements in alphoid regions. The theory is that after the insertion of

active LINES that disrupts the centromeric periodicity, thus compromising the centromere function, an expansion of alphoid DNA occurs in order to compensate this unrest (Schueler et al., 2001; Shepelev et al., 2009). As a consequence, the analysis of LINES can be used to deduce and calculate the age of different satellite blocks. These studies reveal that the most distal alpha-satellite domain is the oldest, with an age gradient advancing proximally through the satellite region.

Finally, it is clear that monomeric alpha satellite present within the pericentromeric regions of human chromosomes predates higher-order arrays of alpha satellite and thus may represent direct descendants of the ancestral primate centromere sequence. Thus, monomeric alphoid arrays are likely the remnants of the centromeres of our primate ancestors, once active and homogenous, that have been replaced by HOR sequences that are a much more efficient substrate for homogenization.

In the evolution of the order Primates, the 171 bp repeat unit seems to be the starting point. Two 171 bp monomers were firstly amplified together as a dimer, then in the Platyrrhini lineage, the two monomers began to accumulate differences due to the decrease of homogenization mechanisms, thus forming the specific New World monkeys ~342 bp monomeric unit (variation of this unit generated the 550 bp in *Chiropotes* and *Cacajao*). In the Catharrhini ancestor, instead, the ~171 bp dimer continued to be amplified by unequal crossover, thus forming the dimeric structure common to all the centromeres as reported in macaque and baboon. Moreover, in the superfamily Hominoidea the 171 bp monomer amplified and diverged in monomeric arrays in gibbon and in higher-order repeats in orangutan, gorilla, chimpanzee, and human.

#### **4. Methods for centromeric sequence analysis: History and news**

The study and analysis of repetitive elements such as centromeric satellite are particularly complicated due the low-variation and high-copy-number nature of these sequences. The development of specific techniques and the application of conventional assays have both contributed in satellite DNA analysis over the last decades.

Elective methods for the analysis of repetitive DNA sequences were the measurement of DNA renaturation kinetics (Britten et al., 1974), the isopycnic centrifugation in gradients of CsCl and CsSO<sub>4</sub> (Szybalski, 1968), and hybridization techniques. The application of these techniques to the analysis of mammalian genomes revealed the existence of several DNA fractions, which differed in their physical and chemical properties. In particular, they allowed the identification and separation of repeated DNA sequences taking advantage of the different copy number and density, respectively, from the bulk of the whole DNA. In fact, the term "satellite" derives from the satellite bands in the genomic DNA profile in CsCl density-gradient centrifugation assay. Southern blot analyses have been successfully applied to discover and examine the tandem-repetitive nature of satellite DNA, with cleavage sites for a number of different restriction endonucleases spaced at highly regular intervals. Further, the development of sequencing strategies has allowed its characterization at the nucleotide level. Most of these historical techniques are still widely used for centromeric satellite sequence analysis, in combination with the newly developed ones.

Further advances came from cytogenetics and molecular cytogenetics that provided some insights in the location and organization of satellite sequences. C-banding has been used to specifically stain the constitutive heterochromatin, including centromeric regions; the use of

satellite sequences as probes first in ISH (*in situ* hybridization) and then in FISH (Fluorescent *in situ* hybridization) experiments, has allowed the definition of their centromeric and pericentromeric localization and distribution among chromosomes. Thus, the application of such techniques clarified the role of satellite DNA as the main constituent of mammalian centromeric DNA. Besides the structural characterization of centromeric heterochromatin, several functional studies of its protein constituents have been developed. The availability of the CREST antiserum, containing a mixture of antibodies against constitutive centromeric proteins, and of the anti-CENPA, anti-CENPB, anti-CENPC, and anti-CENPE specific antibodies, has allowed the localization of the centromeric functionality through immunofluorescence assays. The combination of satellite-FISH and immunofluorescence has definitely showed the co-localization of satellite DNA and centromeric functionality. The awareness of the binding between centromeric DNA and proteins has suggested the use of anti-CENP antibodies to isolate centromeric competent DNA from the bulk of the total DNA through chromatin immunoprecipitation (ChIP) assays. The subsequent FISH-localization, cloning, and sequencing of the isolated DNA allow the full characterization of centromeric competent DNA.

The recent development of high-throughput sequencing technologies has greatly increased the number of organisms with a sequenced genome. However, in genome sequencing projects of mammalian species, the assembly and characterization of centromeric regions cannot be directly achieved due to their repetitive and complex nature. In fact, in each human and other primate genome chromosome assembly there is no sequence in the existing gap between the p and q arms (Rudd & Willard, 2004). Nonetheless, whole-genome shotgun (WGS) sequence reads are not completely inapplicable for the characterization of centromeric regions and satellite DNA sequences. In fact, two computational methods were recently developed to isolate and characterize the centromeric repeats from WGS sequence data: HORdetect (Alkan et al., 2007) and RepeatNet (Alkan et al., 2010). HORdetect recovers alpha satellite sequences and predicts higher-order repeat structure in primate sequencing projects; RepeatNet allows the identification of higher-order repeat structures with no a priori information about the consensus, being thus applicable to any sequenced organism. This method has expanded the knowledge of mammalian satellite DNA and isolated for the first time centromeric satellites from species of the orders Cingulata, Proboscidea, and Didelphimorphia, previously reported in this chapter, although further efforts are needed to better characterize the isolated sequences.

## 5. Conclusion

The overall data collected so far allows the postulation of several considerations regarding centromeric satellite DNA in mammals. Several aspects are quite common among mammalian species, whereas others are shared among a subset of species, in some cases among evolutionarily distant ones, displaying an example of convergent evolution phenomena for satellite DNA. The main features described are summarized in **Table 1**.

Satellite DNA shows high variability across mammalian taxa in monomer size, nucleotide sequence, and quantity relative to the total genome. Although the satellite sequence is not evolutionarily conserved, there are recurrent elements among satellite sequences of closely related species. It is clear that satellite DNA follows a concerted evolution mechanism, so monomers are more similar to monomers of the same species than to monomers of other species. The amount of centromeric and pericentromeric satellite DNA is highly variable as well, and it is not related to the total genome size of the species. In fact, bats have a low C

value but do not show a small relative amount of satellite DNA, as it was thought before its evaluation: the percentage of satellite DNA relative to the total DNA, 3–5%, is very similar to the data of other mammals (**Table 1**).

Quite common features of satellite DNA are: i) the presence of internal direct and inverted subrepeats (Bogenberger et al., 1985; Lee, C. & Lin, 1996; Zhang & Horz, 1984); ii) the presence of the same satellite DNA families at the centromeres of autosomes and the X chromosome, but not at the Y chromosome centromere, like in mouse, sheep, and bat; iii) a greater divergence of the Y centromere sequence in comparison to the other centromeres, like in mouse and primates. Sheep, swine, and horse have a different satellite content in acrocentric and metacentric chromosomes. In two evolutionarily distant mammals, the marsupial *Macropus rufogriseus* and the rodent *Microtus chrotorrhinus*, giant sex chromosomes derived from a large block of heterochromatin at the centromeric and pericentromeric regions have been observed. Finally, evidence of a different rate at which autosomes and sex chromosomes accumulate and dissipate centromeric material has been found in cervid deer, muntjac, and in the genus *Macropus* (Bulazel et al., 2006; Li et al., 2005; Lin & Li, 2006), with the retention for longer periods of time of tandem arrays of ancestral satellites in the sex chromosomes that are not found in the autosomes.

The centromeric satellite is valuable as phylogenetic marker to establish the evolutionarily relationships among species, when they are not found or are ambiguous in the fossil record or other data (Saffery et al., 1999). The centromeric satellite was used as a phylogenetic marker in regard to two different aspects: its nucleotide sequence and its chromosome localization and distribution. The analysis of satellite DNA sequence, thanks to its high divergence rate and rapid evolution during speciation, was used to define the evolutionarily relationships among closely related species that diverged recently. An example is provided by the analysis performed on the spiny mice satellites (Kunze et al., 1999). On the other hand, the comparison of satellite DNA *in situ* hybridization patterns and the study of the nature and amplification of the satellite DNA families on the autosomes and the X chromosome allowed to infer phylogeny and increase the resolution of the evolutionary tree of the Artiodactyla (Chaves et al., 2000; Chaves et al., 2005; Modi et al., 1996). In the family Equidae, the phylogeny of four *Equus* chromosomes was reconstructed by centromere and satellite DNA localization (Alkan et al., 2010; Wade et al., 2009). Moreover, the analysis of satellite DNA sequence, organization, and chromosome distribution, in conjunction with karyotype analysis, is a valuable tool to measure species relationships while also elucidating important aspects of both genome and repetitive sequence evolution.

All the data collected up to date in mammals suggest that centromere satellite sequences are neither necessary nor sufficient for centromere function, and that repetitive DNA is more likely a consequence than a source of centromere function. Nevertheless, the pancentromeric presence of satellite DNA on all mammalian centromeres clearly indicates that a repetitive, ordered, and homogenous sequence is important for centromere maintenance during the evolution of the species.

## 6. References

- Alexandrov, I., Kazakov, A., Tumeneva, I., Shepelev, V. and Yurov, Y. (2001). Alpha-satellite DNA of primates: old and new families. *Chromosoma*, Vol. 110, No. 4, (Aug 2001), pp. (253-66)

- Alexandrov, I.A., Medvedev, L.I., Mashkova, T.D., Kisselev, L.L., Romanova, L.Y. and Yurov, Y.B. (1993). Definition of a new alpha satellite suprachromosomal family characterized by monomeric organization. *Nucleic Acids Res*, Vol. 21, No. 9, (May 11 1993), pp. (2209-15)
- Alkan, C., Cardone, M.F., Catacchio, C.R., Antonacci, F., O'Brien, S.J., Ryder, O.A., Purgato, S., Zoli, M., Della Valle, G., Eichler, E.E. and Ventura, M. (2010). Genome-wide characterization of centromeric satellites from multiple mammalian genomes. *Genome Res*, Vol. 21, No. 1, (Jan 2010), pp. (137-45)
- Alkan, C., Ventura, M., Archidiacono, N., Rocchi, M., Sahinalp, S.C. and Eichler, E.E. (2007). Organization and evolution of primate centromeric DNA from whole-genome shotgun sequence data. *PLoS Comput Biol*, Vol. 3, No. 9, (Sep 2007), pp. (1807-18)
- Alves, G., Seuanez, H.N. and Fanning, T. (1994). Alpha satellite DNA in neotropical primates (Platyrrhini). *Chromosoma*, Vol. 103, No. 4, (Jul 1994), pp. (262-7)
- Alves, G., Seuanez, H.N. and Fanning, T. (1998). A clade of New World primates with distinctive alphoid satellite DNAs. *Mol Phylogenet Evol*, Vol. 9, No. 2, (Apr 1998), pp. (220-4)
- Barragan, M.J., Martinez, S., Marchal, J.A., Fernandez, R., Bullejos, M., Diaz de la Guardia, R. and Sanchez, A. (2003). Pericentric satellite DNA sequences in *Pipistrellus pipistrellus* (Vespertilionidae; Chiroptera). *Heredity*, Vol. 91, No. 3, (Sep 2003), pp. (232-8)
- Bogenberger, J.M., Neumaier, P.S. and Fittler, F. (1985). The muntjak satellite IA sequence is composed of 31-base-pair internal repeats that are highly homologous to the 31-base-pair subrepeats of the bovine satellite 1.715. *Eur J Biochem*, Vol. 148, No. 1, (Apr 1 1985), pp. (55-9)
- Britten, R.J., Graham, D.E. and Neufeld, B.R. (1974). Analysis of repeating DNA sequences by reassociation. *Methods Enzymol*, Vol. 29, No. 0, (1974), pp. (363-418)
- Broad, T.E., Ede, A.J., Forrest, J.W., Lewis, P.E., Phua, S.H. and Pugh, P.A. (1995a). Families of tandemly repeated DNA elements from horse: cloning, nucleotide sequence, and organization. *Genome*, Vol. 38, No. 6, (Dec 1995a), pp. (1285-9)
- Broad, T.E., Forrest, J.W., Lewis, P.E., Pearce, P.D., Phua, S.H., Pugh, P.A. and Stewart-Scott, I.A. (1995b). Cloning of a DNA repeat element from horse: DNA sequence and chromosomal localization. *Genome*, Vol. 38, No. 6, (Dec 1995b), pp. (1132-8)
- Buckland, R.A. (1983). Comparative structure and evolution of goat and sheep satellite I DNAs. *Nucleic Acids Res*, Vol. 11, No. 5, (Mar 11 1983), pp. (1349-60)
- Buckland, R.A. (1985). Sequence and evolution of related bovine and caprine satellite DNAs. Identification of a short DNA sequence potentially involved in satellite DNA amplification. *J Mol Biol*, Vol. 186, No. 1, (Nov 5 1985), pp. (25-30)
- Bulazel, K., Metcalfe, C., Ferreri, G.C., Yu, J., Eldridge, M.D. and O'Neill, R.J. (2006). Cytogenetic and molecular evaluation of centromere-associated DNA sequences from a marsupial (Macropodidae: *Macropus rufogriseus*) X chromosome. *Genetics*, Vol. 172, No. 2, (Feb 2006), pp. (1129-37)
- Burkin, D.J., Broad, T.E. and Jones, C. (1996). The chromosomal distribution and organization of sheep satellite I and II centromeric DNA using characterized sheep-hamster somatic cell hybrids. *Chromosome Res*, Vol. 4, No. 1, (Jan 1996), pp. (49-55)

- Burton, D.W., Bickham, J.W. and Genoways, H.H. (1989). Flow-cytometric analyses of nuclear DNA content in four families of neotropical bats. *Evolution*, Vol. 43, 1989, pp. (756-765)
- Carbone, L., Nergadze, S.G., Magnani, E., Misceo, D., Francesca Cardone, M., Roberto, R., Bertoni, L., Attolini, C., Francesca Piras, M., de Jong, P., Raudsepp, T., Chowdhary, B.P., Guerin, G., Archidiacono, N., Rocchi, M. and Giulotto, E. (2006). Evolutionary movement of centromeres in horse, donkey, and zebra. *Genomics*, Vol. 87, No. 6, (Jun 2006), pp. (777-82)
- Cellamare, A., Catacchio, C.R., Alkan, C., Giannuzzi, G., Antonacci, F., Cardone, M.F., Della Valle, G., Malig, M., Rocchi, M., Eichler, E.E. and Ventura, M. (2009). New insights into centromere organization and evolution from the white-cheeked gibbon and marmoset. *Mol Biol Evol*, Vol. 26, No. 8, (Aug 2009), pp. (1889-900)
- Charlesworth, B., Sniegowski, P. and Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature*, Vol. 371, No. 6494, (Sep 15 1994), pp. (215-20)
- Chaves, R., Adegas, F., Heslop-Harrison, J.S., Guedes-Pinto, H. and Wienberg, J. (2003). Complex satellite DNA reshuffling in the polymorphic t(1;29) Robertsonian translocation and evolutionarily derived chromosomes in cattle. *Chromosome Res*, Vol. 11, No. 7, 2003, pp. (641-8)
- Chaves, R., Guedes-Pinto, H., Heslop-Harrison, J. and Schwarzacher, T. (2000). The species and chromosomal distribution of the centromeric alpha-satellite I sequence from sheep in the tribe Caprini and other Bovidae. *Cytogenet Cell Genet*, Vol. 91, No. 1-4, 2000, pp. (62-6)
- Chaves, R., Guedes-Pinto, H. and Heslop-Harrison, J.S. (2005). Phylogenetic relationships and the primitive X chromosome inferred from chromosomal and satellite DNA analysis in Bovidae. *Proc Biol Sci*, Vol. 272, No. 1576, (Oct 7 2005), pp. (2009-16)
- Choo, K.H. (1990). Role of acrocentric centromere satellite DNA in Robertsonian translocation and chromosomal non-disjunction. *Mol Biol Med*, Vol. 7, No. 5, (Oct 1990), pp. (437-49)
- Choo, K.H., Vissel, B., Nagy, A., Earle, E. and Kalitsis, P. (1991). A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. *Nucleic Acids Res*, Vol. 19, No. 6, (Mar 25 1991), pp. (1179-82)
- D'Aiuto, L., Barsanti, P., Mauro, S., Cserpan, I., Lanave, C. and Ciccarese, S. (1997). Physical relationship between satellite I and II DNA in centromeric regions of sheep chromosomes. *Chromosome Res*, Vol. 5, No. 6, (Sep 1997), pp. (375-81)
- Denome, R.M., O'Callaghan, B., Luitjens, C., Harper, E. and Bianco, R. (1994). Characterization of a satellite DNA from *Antilocapra americana*. *Gene*, Vol. 145, No. 2, (Aug 5 1994), pp. (257-9)
- Dover, G. (1982). Molecular drive: a cohesive mode of species evolution. *Nature*, Vol. 299, No. 5879, (Sep 9 1982), pp. (111-7)
- Dover, G. (2002). Molecular drive. *Trends Genet*, Vol. 18, No. 11, (Nov 2002), pp. (587-9)
- Durfy, S.J. and Willard, H.F. (1989). Patterns of intra- and interarray sequence variation in alpha satellite from the human X chromosome: evidence for short-range homogenization of tandemly repeated DNA sequences. *Genomics*, Vol. 5, No. 4, (Nov 1989), pp. (810-21)

- Durfy, S.J. and Willard, H.F. (1990). Concerted evolution of primate alpha satellite DNA. Evidence for an ancestral sequence shared by gorilla and human X chromosome alpha satellite. *J Mol Biol*, Vol. 216, No. 3, (Dec 5 1990), pp. (555-66)
- Ekes, C., Csonka, E., Hadlaczky, G. and Cserpan, I. (2004). Isolation, cloning and characterization of two major satellite DNA families of rabbit (*Oryctolagus cuniculus*). *Gene*, Vol. 343, No. 2, (Dec 22 2004), pp. (271-9)
- Eldridge, M.D. and Close, R.L. (1993). Radiation of chromosome shuffles. *Curr Opin Genet Dev*, Vol. 3, No. 6, (Dec 1993), pp. (915-22)
- Elizur, A., Dennis, E.S. and Peacock, W.J. (1982). Satellite DNA sequences in the red kangaroo (*Macropus rufus*). *Aust J Biol Sci*, Vol. 35, No. 3, (1982), pp. (313-25)
- Fanning, T.G. (1989). Molecular evolution of centromere-associated nucleotide sequences in two species of canids. *Gene*, Vol. 85, No. 2, (Dec 28 1989), pp. (559-63)
- Fantaccione, S., Pontecorvo, G. and Zampella, V. (2005). Molecular characterization of the first satellite DNA with CENP-B and CDEIII motifs in the bat *Pipistrellus kuhli*. *FEBS Lett*, Vol. 579, No. 11, (Apr 25 2005), pp. (2519-27)
- Fatyol, K., Cserpan, I., Praznovszky, T., Kereso, J. and Hadlaczky, G. (1994). Cloning and molecular characterization of a novel chromosome specific centromere sequence of Chinese hamster. *Nucleic Acids Res*, Vol. 22, No. 18, (Sep 11 1994), pp. (3728-36)
- Ferreri, G.C., Brown, J.D., Obergefell, C., Jue, N., Finn, C.E., O'Neill, M.J. and O'Neill, R.J. (2011). Recent amplification of the kangaroo endogenous retrovirus, KERV, limited to the centromere. *J Virol*, Vol. 85, No. 10, (May 2011), pp. (4761-71)
- 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*, Vol. 96, No. 3, (May-Jun 2005), pp. (217-24)
- Ferreri, G.C., Marzelli, M., Rens, W. and O'Neill, R.J. (2004). A centromere-specific retroviral element associated with breaks of synteny in macropodine marsupials. *Cytogenet Genome Res*, Vol. 107, No. 1-2, (2004), pp. (115-8)
- Goodman, M., Koop, B.F., Czelusniak, J., Fitch, D.H., Tagle, D.A. and Slightom, J.L. (1989). Molecular phylogeny of the family of apes and humans. *Genome*, Vol. 31, No. 1, (1989), pp. (316-35)
- Guy, J., Hearn, T., Crosier, M., Mudge, J., Viggiano, L., Koczan, D., Thiesen, H.J., Bailey, J.A., Horvath, J.E., Eichler, E.E., Earthrowl, M.E., Deloukas, P., French, L., Rogers, J., Bentley, D. and Jackson, M.S. (2003). Genomic sequence and transcriptional profile of the boundary between pericentromeric satellites and genes on human chromosome arm 10p. *Genome Res*, Vol. 13, No. 2, (Feb 2003), pp. (159-72)
- Hayman, D.L. (1990). Marsupial cytogenetics. *Aust J Zool*, Vol. 37, (1990), pp. (331-349)
- Hayman, D.L. and Martin, P.G. (1974). Mammalia I: Monotremata and Marsupialia. *Animal Cytogenetics*. John., B. 4.
- Henikoff, S., Ahmad, K. and Malik, H.S. (2001). The centromere paradox: stable inheritance with rapidly evolving DNA. *Science*, Vol. 293, No. 5532, (Aug 10 2001), pp. (1098-102)
- Horz, W. and Altenburger, W. (1981). Nucleotide sequence of mouse satellite DNA. *Nucleic Acids Res*, Vol. 9, No. 3, (Feb 11 1981), pp. (683-96)
- Iurov Iu, B., Mitkevich, S.P. and Aleksandrov, I.A. (1988). [Molecular cytogenetic research on the polymorphism of segments of the constitutive heterochromatin in human chromosomes]. *Genetika*, Vol. 24, No. 2, (Feb 1988), pp. (356-65)

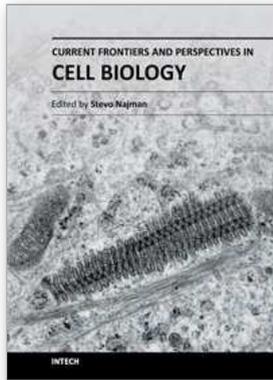
- Kalitsis, P., Griffiths, B. and Choo, K.H. (2006). Mouse telocentric sequences reveal a high rate of homogenization and possible role in Robertsonian translocation. *Proc Natl Acad Sci U S A*, Vol. 103, No. 23, (Jun 6 2006), pp. (8786-91)
- Karpen, G.H. and Allshire, R.C. (1997). The case for epigenetic effects on centromere identity and function. *Trends Genet*, Vol. 13, No. 12, (Dec 1997), pp. (489-96)
- Kazakov, A.E., Shepelev, V.A., Tumeneva, I.G., Alexandrov, A.A., Yurov, Y.B. and Alexandrov, I.A. (2003). Interspersed repeats are found predominantly in the "old" alpha satellite families. *Genomics*, Vol. 82, No. 6, (Dec 2003), pp. (619-27)
- Kipling, D., Ackford, H.E., Taylor, B.A. and Cooke, H.J. (1991). Mouse minor satellite DNA genetically maps to the centromere and is physically linked to the proximal telomere. *Genomics*, Vol. 11, No. 2, (Oct 1991), pp. (235-41)
- Kopecka, H., Macaya, G., Cortadas, J., Thiery, J.P. and Bernardi, G. (1978). Restriction enzyme analysis of satellite DNA components from the bovine genome. *Eur J Biochem*, Vol. 84, No. 1, (Mar 1978), pp. (189-95)
- Kunze, B., Traut, W., Garagna, S., Weichenhan, D., Redi, C.A. and Winking, H. (1999). Pericentric satellite DNA and molecular phylogeny in *Acomys* (Rodentia). *Chromosome Res*, Vol. 7, No. 2, (1999), pp. (131-41)
- Kurnit, D.M., Shafit, B.R. and Maio, J.J. (1973). Multiple satellite deoxyribonucleic acids in the calf and their relation to the sex chromosomes. *J Mol Biol*, Vol. 81, No. 3, (Dec 15 1973), pp. (273-84)
- Kuznetsova, I.S., Prusov, A.N., Erukashvily, N.I. and Podgornaya, O.I. (2005). New types of mouse centromeric satellite DNAs. *Chromosome Res*, Vol. 13, No. 1, (2005), pp. (9-25)
- Lee, C. and Lin, C.C. (1996). Conservation of a 31-bp bovine subrepeat in centromeric satellite DNA monomers of *Cervus elaphus* and other cervid species. *Chromosome Res*, Vol. 4, No. 6, (Sep 1996), pp. (427-35)
- Lee, H.R., Hayden, K.E. and Willard, H.F. (2011). Organization and molecular evolution of CENP-A-associated satellite DNA families in a basal primate genome. *Genome Biol Evol*, (Aug 9 2011),
- Li, Y.C., Cheng, Y.M., Hsieh, L.J., Ryder, O.A., Yang, F., Liao, S.J., Hsiao, K.M., Tsai, F.J., Tsai, C.H. and Lin, C.C. (2005). Karyotypic evolution of a novel cervid satellite DNA family isolated by microdissection from the Indian muntjac Y-chromosome. *Chromosoma*, Vol. 114, No. 1, (May 2005), pp. (28-38)
- Lin, C.C. and Li, Y.C. (2006). Chromosomal distribution and organization of three cervid satellite DNAs in Chinese water deer (*Hydropotes inermis*). *Cytogenet Genome Res*, Vol. 114, No. 2, (2006), pp. (147-54)
- Mahtani, M.M. and Willard, H.F. (1990). Pulsed-field gel analysis of alpha-satellite DNA at the human X chromosome centromere: high-frequency polymorphisms and array size estimate. *Genomics*, Vol. 7, No. 4, (Aug 1990), pp. (607-13)
- Maio, J.J. (1971). DNA strand reassociation and polyribonucleotide binding in the African green monkey, *Cercopithecus aethiops*. *J Mol Biol*, Vol. 56, No. 3, (Mar 28 1971), pp. (579-95)
- Manuelidis, L. (1978). Chromosomal localization of complex and simple repeated human DNAs. *Chromosoma*, Vol. 66, No. 1, (Mar 22 1978), pp. (23-32)
- 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*, Vol. 109, No. 5, (Nov 1989), pp. (1963-73)
- Modi, W.S. (1992). Nucleotide sequence and genomic organization of a tandem satellite array from the rock vole *Microtus chrotorrhinus* (Rodentia). *Mamm Genome*, Vol. 3, No. 4, 1992, pp. (226-32)
- Modi, W.S., Gallagher, D.S. and Womack, J.E. (1996). Evolutionary histories of highly repeated DNA families among the Artiodactyla (Mammalia). *J Mol Evol*, Vol. 42, No. 3, (Mar 1996), pp. (337-49)
- Musich, P.R., Brown, F.L. and Maio, J.J. (1980). Highly repetitive component alpha and related allopolyploid DNAs in man and monkeys. *Chromosoma*, Vol. 80, No. 3, 1980, pp. (331-48)
- Novak, U. (1984). Structure and properties of a highly repetitive DNA sequence in sheep. *Nucleic Acids Res*, Vol. 12, No. 5, (Mar 12 1984), pp. (2343-50)
- O'Neill, R.J., Eldridge, M.D. and Metcalfe, C.J. (2004). Centromere dynamics and chromosome evolution in marsupials. *J Hered*, Vol. 95, No. 5, (Sep-Oct 2004), pp. (375-81)
- 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*, Vol. 393, No. 6680, (May 7 1998), pp. (68-72)
- Pardue, M.L. and Gall, J.G. (1970). Chromosomal localization of mouse satellite DNA. *Science*, Vol. 168, No. 937, (Jun 12 1970), pp. (1356-8)
- Pech, M., Streeck, R.E. and Zachau, H.G. (1979). Patchwork structure of a bovine satellite DNA. *Cell*, Vol. 18, No. 3, (Nov 1979), pp. (883-93)
- Pertile, M.D., Graham, A.N., Choo, K.H. and Kalitsis, P. (2009). Rapid evolution of mouse Y centromere repeat DNA belies recent sequence stability. *Genome Res*, Vol. 19, No. 12, (Dec 2009), pp. (2202-13)
- Piras, F.M., Nergadze, S.G., Magnani, E., Bertoni, L., Attolini, C., Khoriantuli, L., Raimondi, E. and Giulotto, E. (2010). Uncoupling of satellite DNA and centromeric function in the genus *Equus*. *PLoS Genet*, Vol. 6, No. 2, (Feb 2010), pp. (e1000845)
- Plucienniczak, A., Skowronski, J. and Jaworski, J. (1982). Nucleotide sequence of bovine 1.715 satellite DNA and its relation to other bovine satellite sequences. *J Mol Biol*, Vol. 158, No. 2, (Jun 25 1982), pp. (293-304)
- Prasad, A.B., Allard, M.W., Program, N.C.S. and Green, E.D. (2008). Confirming the phylogeny of mammals by use of large comparative sequence data sets. *Mol Biol Evol*, Vol. 25, No. 9, (Sep 2008), pp. (1795-808)
- Reisner, A.H. and Bucholtz, C.A. (1983). Apparent relatedness of the main component of ovine 1.714 satellite DNA to bovine 1.715 satellite DNA. *EMBO J*, Vol. 2, No. 7, 1983, pp. (1145-9)
- Rens, W., O'Brien, P.C., Fairclough, H., Harman, L., Graves, J.A. and Ferguson-Smith, M.A. (2003). Reversal and convergence in marsupial chromosome evolution. *Cytogenet Genome Res*, Vol. 102, No. 1-4, 2003, pp. (282-90)
- Rieder, C.L. and Salmon, E.D. (1994). Motile kinetochores and polar ejection forces dictate chromosome position on the vertebrate mitotic spindle. *J Cell Biol*, Vol. 124, No. 3, (Feb 1994), pp. (223-33)
- Rofe, R.H. (1978). G-banded chromosomes and the evolution of Macropodidae. *Aust Mammal*, Vol. 2, 1978, pp. (52-62)

- Roizes, G. (2006). Human centromeric alphoid domains are periodically homogenized so that they vary substantially between homologues. Mechanism and implications for centromere functioning. *Nucleic Acids Res*, Vol. 34, No. 6, (2006), pp. (1912-24)
- Rosenberg, H., Singer, M. and Rosenberg, M. (1978). Highly reiterated sequences of SIMIANSIMIANSIMIANSIMIANSIMIAN. *Science*, Vol. 200, No. 4340, (Apr 28 1978), pp. (394-402)
- Rudd, M.K. and Willard, H.F. (2004). Analysis of the centromeric regions of the human genome assembly. *Trends Genet*, Vol. 20, No. 11, (Nov 2004), pp. (529-33)
- Rudd, M.K., Wray, G.A. and Willard, H.F. (2006). The evolutionary dynamics of alpha-satellite. *Genome Res*, Vol. 16, No. 1, (Jan 2006), pp. (88-96)
- Saffery, R., Earle, E., Irvine, D.V., Kalitsis, P. and Choo, K.H. (1999). Conservation of centromere protein in vertebrates. *Chromosome Res*, Vol. 7, No. 4, (1999), pp. (261-5)
- Samonte, R.V., Ramesh, K.H. and Verma, R.S. (1997). Comparative mapping of human alphoid satellite DNA repeat sequences in the great apes. *Genetica*, Vol. 101, No. 2, (1997), pp. (97-104)
- Schindelbauer, D. and Schwarz, T. (2002). Evidence for a fast, intrachromosomal conversion mechanism from mapping of nucleotide variants within a homogeneous alpha-satellite DNA array. *Genome Res*, Vol. 12, No. 12, (Dec 2002), pp. (1815-26)
- Schneider, H., Schneider, M.P., Sampaio, I., Harada, M.L., Stanhope, M., Czelusniak, J. and Goodman, M. (1993). Molecular phylogeny of the New World monkeys (Platyrrhini, primates). *Mol Phylogenet Evol*, Vol. 2, No. 3, (Sep 1993), pp. (225-42)
- Schueler, M.G., Dunn, J.M., Bird, C.P., Ross, M.T., Viggiano, L., Program, N.C.S., Rocchi, M., Willard, H.F. and Green, E.D. (2005). Progressive proximal expansion of the primate X chromosome centromere. *Proc Natl Acad Sci U S A*, Vol. 102, No. 30, (Jul 26 2005), pp. (10563-8)
- 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*, Vol. 294, No. 5540, (Oct 5 2001), pp. (109-15)
- Shepelev, V.A., Alexandrov, A.A., Yurov, Y.B. and Alexandrov, I.A. (2009). The evolutionary origin of man can be traced in the layers of defunct ancestral alpha satellites flanking the active centromeres of human chromosomes. *PLoS Genet*, Vol. 5, No. 9, (Sep 2009), pp. (e1000641)
- Singer, D. and Donehower, L. (1979). Highly repeated DNA of the baboon: organization of sequences homologous to highly repeated DNA of the African green monkey. *J Mol Biol*, Vol. 134, No. 4, (Nov 15 1979), pp. (835-42)
- Smith, G.P. (1976). Evolution of repeated DNA sequences by unequal crossover. *Science*, Vol. 191, No. 4227, (Feb 13 1976), pp. (528-35)
- Stephan, W. (1986). Recombination and the evolution of satellite DNA. *Genet Res*, Vol. 47, No. 3, (Jun 1986), pp. (167-74)
- 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*, Vol. 11, No. 11, (Nov 2004), pp. (1076-83)
- Sullivan, K.F., Hechenberger, M. and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J Cell Biol*, Vol. 127, No. 3, (Nov 1994), pp. (581-92)

- Sunkel, C.E. and Coelho, P.A. (1995). The elusive centromere: sequence divergence and functional conservation. *Curr Opin Genet Dev*, Vol. 5, No. 6, (Dec 1995), pp. (756-67)
- Szybalski, W. (1968). Use of cesium sulfate for equilibrium density gradient centrifugation. *Methods Enzymol*, Vol. 12B, 1968, pp. (330-360)
- Taparowsky, E.J. and Gerbi, S.A. (1982). Sequence analysis of bovine satellite I DNA (1.715 gm/cm<sup>3</sup>). *Nucleic Acids Res*, Vol. 10, No. 4, (Feb 25 1982), pp. (1271-81)
- Torras-Llort, M., Moreno-Moreno, O. and Azorin, F. (2009). Focus on the centre: the role of chromatin on the regulation of centromere identity and function. *EMBO J*, Vol. 28, No. 16, (Aug 19 2009), pp. (2337-48)
- Tyler-Smith, C. and Brown, W.R. (1987). Structure of the major block of alphoid satellite DNA on the human Y chromosome. *J Mol Biol*, Vol. 195, No. 3, (Jun 5 1987), pp. (457-70)
- Venolia, L. and Peacock, W.J. (1981). A highly repeated DNA from the genome of the wallaroo (*Macropus robustus robustus*). *Aust J Biol Sci*, Vol. 34, No. 1, 1981, pp. (97-113)
- Vissel, B. and Choo, K.H. (1987). Human alpha satellite DNA--consensus sequence and conserved regions. *Nucleic Acids Res*, Vol. 15, No. 16, (Aug 25 1987), pp. (6751-2)
- 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*, Vol. 5, No. 3, (Oct 1989), pp. (407-14)
- Volobouev, V., Vogt, N., Viegas-Pequignot, E., Malfoy, B. and Dutrillaux, B. (1995). Characterization and chromosomal location of two repeated DNAs in three Gerbillus species. *Chromosoma*, Vol. 104, No. 4, (Dec 1995), pp. (252-9)
- Wade, C.M., Giulotto, E., Sigurdsson, S., Zoli, M., Gnerre, S., Insland, F., Lear, T.L., Adelson, D.L., Bailey, E., Bellone, R.R., Blocker, H., Distl, O., Edgar, R.C., Garber, M., Leeb, T., Mauceli, E., MacLeod, J.N., Penedo, M.C., Raison, J.M., Sharpe, T., Vogel, J., Andersson, L., Antczak, D.F., Biagi, T., Binns, M.M., Chowdhary, B.P., Coleman, S.J., Della Valle, G., Fryc, S., Guerin, G., Hasegawa, T., Hill, E.W., Jurka, J., Kiialainen, A., Lindgren, G., Liu, J., Magnani, E., Mickelson, J.R., Murray, J., Nergadze, S.G., Onofrio, R., Pedroni, S., Piras, M.F., Raudsepp, T., Rocchi, M., Roed, K.H., Ryder, O.A., Searle, S., Skow, L., Swinburne, J.E., Syvanen, A.C., Tozaki, T., Valberg, S.J., Vaudin, M., White, J.R., Zody, M.C., Lander, E.S. and Lindblad-Toh, K. (2009). Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*, Vol. 326, No. 5954, (Nov 6 2009), pp. (865-7)
- Warburton, P.E., Haaf, T., Gosden, J., Lawson, D. and Willard, H.F. (1996). Characterization of a chromosome-specific chimpanzee alpha satellite subset: evolutionary relationship to subsets on human chromosomes. *Genomics*, Vol. 33, No. 2, (Apr 15 1996), pp. (220-8)
- Warburton, P.E. and Willard, H.F. (1990). Genomic analysis of sequence variation in tandemly repeated DNA. Evidence for localized homogeneous sequence domains within arrays of alpha-satellite DNA. *J Mol Biol*, Vol. 216, No. 1, (Nov 5 1990), pp. (3-16)
- Waye, J.S. and Willard, H.F. (1986). Molecular analysis of a deletion polymorphism in alpha satellite of human chromosome 17: evidence for homologous unequal crossing-over

- and subsequent fixation. *Nucleic Acids Res*, Vol. 14, No. 17, (Sep 11 1986), pp. (6915-27)
- Waye, J.S. and Willard, H.F. (1987). Nucleotide sequence heterogeneity of alpha satellite repetitive DNA: a survey of alphoid sequences from different human chromosomes. *Nucleic Acids Res*, Vol. 15, No. 18, (Sep 25 1987), pp. (7549-69)
- Wayne, R.K., Geffen, E., Girman, D.J., Koepfli, K.P., Lau, L.M. and Marshall, C.R. (1997). Molecular systematics of the Canidae. *Syst Biol*, Vol. 46, No. 4, (Dec 1997), pp. (622-53)
- Wevrick, R. and Willard, H.F. (1989). Long-range organization of tandem arrays of alpha satellite DNA at the centromeres of human chromosomes: high-frequency array-length polymorphism and meiotic stability. *Proc Natl Acad Sci U S A*, Vol. 86, No. 23, (Dec 1989), pp. (9394-8)
- Wijers, E.R., Zijlstra, C. and Lenstra, J.A. (1993). Rapid evolution of horse satellite DNA. *Genomics*, Vol. 18, No. 1, (Oct 1993), pp. (113-7)
- Willard, H.F. (1990). Centromeres of mammalian chromosomes. *Trends Genet*, Vol. 6, No. 12, (Dec 1990), pp. (410-6)
- Willard, H.F. and Waye, J.S. (1987). Chromosome-specific subsets of human alpha satellite DNA: analysis of sequence divergence within and between chromosomal subsets and evidence for an ancestral pentameric repeat. *J Mol Evol*, Vol. 25, No. 3, (1987), pp. (207-14)
- Wong, A.K. and Rattner, J.B. (1988). Sequence organization and cytological localization of the minor satellite of mouse. *Nucleic Acids Res*, Vol. 16, No. 24, (Dec 23 1988), pp. (11645-61)
- Wu, J.C. and Manuelidis, L. (1980). Sequence definition and organization of a human repeated DNA. *J Mol Biol*, Vol. 142, No. 3, (Sep 25 1980), pp. (363-86)
- Yoda, K., Kitagawa, K., Masumoto, H., Muro, Y. and Okazaki, T. (1992). A human centromere protein, CENP-B, has a DNA binding domain containing four potential alpha helices at the NH2 terminus, which is separable from dimerizing activity. *J Cell Biol*, Vol. 119, No. 6, (Dec 1992), pp. (1413-27)
- Zhang, X.Y. and Horz, W. (1984). Nucleosomes are positioned on mouse satellite DNA in multiple highly specific frames that are correlated with a diverged subrepeat of nine base-pairs. *J Mol Biol*, Vol. 176, No. 1, (Jun 15 1984), pp. (105-29)



## **Current Frontiers and Perspectives in Cell Biology**

Edited by Prof. Stevo Najman

ISBN 978-953-51-0544-2

Hard cover, 556 pages

**Publisher** InTech

**Published online** 25, April, 2012

**Published in print edition** April, 2012

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Giuliana Giannuzzi, Claudia Rita Catacchio and Mario Ventura (2012). Centromere Evolution: Digging into Mammalian Primary Constriction, Current Frontiers and Perspectives in Cell Biology, Prof. Stevo Najman (Ed.), ISBN: 978-953-51-0544-2, InTech, Available from: <http://www.intechopen.com/books/current-frontiers-and-perspectives-in-cell-biology/centromere-evolution-digging-into-mammalian-primary-constriction>

# **INTECH**

open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.