J Mol Evol (1998) 46:240–244

JOURNAL OF MOLECULAR EVOLUTION

An Alphoid-Like Satellite DNA Sequence Is Present in the Genome of a Lacertid Lizard

T. Capriglione,[1] M.G. De Santo,[1] G. Odierna,[1] E. Olmo[2]

[1] Dipartimento di Biologia Evolutiva e Comparata, Università di Napoli ''Federico II,'' Via Mezzocannone 8, Napoli, Italy
[2] Instituto di Biologia e Genetica, Facoltà di Scienze, Università di Ancona, Via delle Brecce Bianche, Ancona, Italy

Received: 6 February 1997 / Accepted: 14 May 1997

Abstract. A *PstI* DNA family was isolated from the genome of a lacertid, *Lacerta graeca.* The 185-bp monomeric unit (pGPS) was cloned and hybridized to DNAs and chromosomes of several lacertid species. The data showed that pGPS hybridizes to the (1) centromeric or pericentromeric heterochromatin of almost all the chromosomes of *L. graeca* and (2) genomic DNA of species phylogenetically related and unrelated to *L. graeca.* The presence of pGPS even in species immunologically apart more than 30 million years suggests that this repeated family might be either very ancient or have been conserved during evolution due to its functional role. The latter hypothesis might be supported by the results of sequence analysis which showed some homology with both several alphoid sequences of primates and the CDEIII centromeric sequence of yeast. Segments of the satellite sequence are similar to the mammalian CENP-B box. These observations suggest that pGPS might have a role in determining the centromeric function in lacertid lizards.

Key words: Reptiles — Satellite DNA — Centromere

Introduction

The centromeres of eukaryotes appear as structures playing a fundamental role in the distribution of the chromosome set during mitosis and meiosis. They perform this function by a specific chromosome attachment site, the kinetochore, to which spindle microtubules are directly attached.

The kinetochore region of the centromere has been widely characterized in several organisms, and attempts have been made to identify both the DNA sequences and the proteins involved in the determination of the structure and/or function of this region (Rieder 1982).

The centromeric DNA sequence has been determined in *S. cerevisiae:* it is a 120-bp-long sequence showing a 25-bp element, CDEIII, which is fundamental for yeast centromeres to function and, moreover, is evolutionarily conserved in higher eukaryotes as well (Grady 1992; Ragghianti et al. 1995). However, in these organisms, the centromeric region has a much more complex organization, being often flanked with highly repeated DNA. Particular attention has been attached to the study of these sequences, especially the human alphoid satellite family.

Alphoid satellite DNAs, together with other satellite centromeric components, seem to have a role in the primary constriction of humans and other mammals (Haaf et al. 1992, 1995; Tyler-Smith et al. 1993), since they appear to be associated with centromere antigens (Masumoto et al. 1989). These alphoid DNAs are probably highly conservative, since Grellet et al. (1986) have identified an alphoid-like sequence in the genome of a higher plant.

Over the last years, in our laboratory, we have tried to characterize the highly repeated DNA fraction present in the genome of some lacertids, since this fraction seems to

Correspondence to: T. Capriglione

have been the most variable in this group (Olmo et al. 1988; Capriglione et al. 1991, 1994).

We have already isolated three families of repeated DNA (pLCS, pLHS, pSHS) from several species of lacertid lizards; due to their varying degree of conservativeness, they can provide useful information on the phylogeny of this family (Capriglione 1995).

This paper describes a new family of repeated DNA isolated from the genome of the lacertid lizard *Lacerta graeca;* due to the characteristics of this sequence, it might be involved in determining the centromeric function.

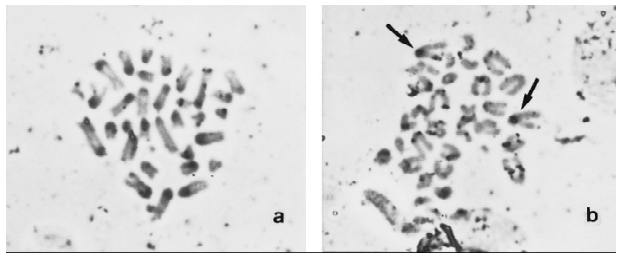

Fig. 1. Localization of the pGPS satellite DNA sequence to the chromosome of *L. graeca* **(a)** and *P. tiliguerta* **(b)** by DIG-labeled *in situ* hybridization. The *arrows* in **(b)** indicate the largest chromosome pairs with a centromeric localization.

Materials and Methods

Animals

Lacerta graeca and *Algyroides moreoticus* specimens were a kind gift of Dr. Mayer. Specimens of *Podarcis sicula, Podarcis muralis, Podarcis taurica, Podarcis tiliguerta, Lacerta viridis*, and *Lacerta bedriagae* were collected by Dr. V. Caputo. *Lacerta vivipara* and *L. saxicola* animals were kindly provided by Dr. Larissa Kupriyanova.

DNA Analysis, Cloning and Sequencing

Whole genomic DNA was purified from blood and/or gonads using the method previously described by Capriglione et al. (1994). An amount of 10 μg of whole genomic DNA of *L. graeca* was digested with the enzyme *PstI* at 2 U/μg. The fragments resulting from digestion were run on a 1.2% agarose gel. The 185-bp monomeric unit was purified from a low melting agarose gel by phenol extraction and cloned in a *PstI*-restricted pUC 18 plasmid. This clone has been indicated as pGPS.

For Southern blotting, *PstI*-digested DNAs of several lacertid species were transferred to Hybond-C membranes (Southern 1975) and hybridized to a pGPS probe (20 ng/ml), labeled with DIG-dUTP using a random primed DNA labeling kit (Boehringer Mannheim). Hybridization and immunological detection were carried out under the conditions suggested by the Boehringer Mannheim protocol. *In situ* hybridization was performed as previously described by Capriglione et al. (1994). The nucleotide sequence of four clones was determined by the Sanger method (Sanger et al. 1977), using the Boehringer Mannheim kit. Alphoid sequences were taken from GENEMBL sequence data base. Homology searches were done with the FASTA program in the University of Wisconsin GCG software package.

pGPS nucleotide sequence have been deposited in GenBank (accession number AF001405).

Results

A 185-bp *PstI* fragment was found to be the main component of the centromeric region in *L. graeca* chromosomes. *In situ* hybridization (Fig. 1a) experiments revealed that the pGPS clone, containing the satellite DNA, hybridized in the centromeric region of all *L. graeca* chromosomes; the two smallest among them, however, showed a faint hybridization signal. In *Podarcis tiliguerta*, a species considered to be phylogenetically more than 10 million years apart from *L. graeca* (Mayer and Lutz, 1990), the same clone appeared to be still conserved. However, in this species, the *PstI* satellite family was mostly localized in the pericentromeric region, except in one of the largest chromosome pairs, where pGPS preserved a centromeric localization (Fig. 1b).

The presence of this highly repeated DNA family was checked out in the *PstI*-restricted genome of several species belonging to different genera of the same group (Fig. 2).

The main ladder of bands was present and conserved in all the genomes analyzed, even in unrelated species (except *L. saxicola*); the low-molecular-weight bands, however, showed a fainter hybridization signal, if any. This result indicates a certain degree of divergence in the sequences of pGPS during evolution.

The data obtained by sequence analysis highlighted that pGPS shares the presence of AT-rich blocks with other centromeric satellites so far isolated both in lacertids and in other vertebrates. However, the consensus sequence of pGPS (Fig. 3) had the same percentage of AT and GC, unlike other satellite DNAs isolated from the genome of lacertid lizards, which are usually richer in AT. Interestingly, homology analysis showed that a region of pGPS shares a 72% identity with the CDEIII element of *S. cerevisiae* centromeres (Clarke and Carbon 1985) (Fig. 4). Furthermore, it is to be pointed out that cytosine sequences at position 13–14, the mutation of which dramatically alters the function of the yeast centromere, were preserved.

Three regions sharing up to 59% of similarity to the consensus sequence for mammalian CENP-B box were also conserved (Fig. 5). CENP-B box is the binding site of a helix-loop-helix highly conserved mammalian centromere protein (Cooke et al. 1990; Sullivan and Glass 1991).

Similarities were also found by comparing pGPS data with a data set of alphoid sequences. Homology with some of these sequences was found: in particular, in some regions, up to 61% identity with alphoid chr4, chr6, and chr17 of humans. Though to a lower degree, identity was also observed with alphoid DNA families of *Cebus apella* and *Cercopithecus aethiops* (Fig. 6).

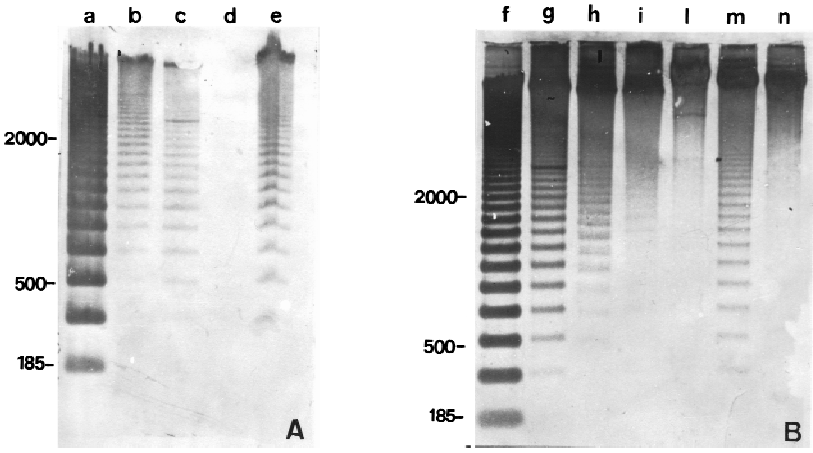

Fig. 2. Southern blots of *PstI*-digested genomic DNAs from several lacertid species of the same **(A)** or different genera **(B),** hybridized to DIG-labeled pGPS. a, f, *L. graeca;* b, *L. bedriagae;* c, *L. vivipara;* d, *L. saxicola;* e, *L. viridis;* g, *A. moreoticus;* h, *P. sicula;* i, *P. muralis;* l, *P. taurica;* m, *P. tiliguerta;* n, *P. hispanica.*

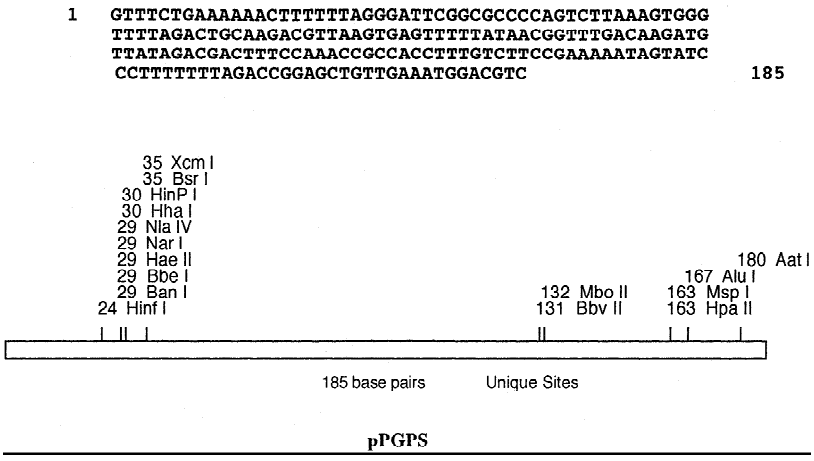

Fig. 3. Consensus sequence and restriction map of pGPS.

```
pGPS        CGCCACCTTTGTCTTCCGAAAAATAGTATCCCTTTTTTTTAGACCGGAGCTGTTGAAATGG      180
CDE III       TGTTTTTGNTTTCCGAAAANNNAAA
```

Fig. 4. Comparison of pGPS sequence with CDEIII of yeast A region between nucleotides 123 and 148 with a 72% homology is present.

```
pGPS        GTTTCTGAAAAAACTTTTTTAGGGATTCGGCGCCCCAGTCTTAAAGTGGGTTTTAGACTG      60
CENP-B                                                  CTTCGTTGGAAACGG

            CAAGACGTTAAGTGAGTTTTTATAACGGTTTGACAAGATGTTATAGACGACTTTCCAAAC     120
            GA                                                  CTTCGTTGGAAAC

            CGCCACCTTTGTCTTCCGAAAAATAGTATCCCTTTTTTTAGACCGGAGCTGTTGAAATGG     180
            GGGA                                    CTTCGTTGGAAACGGGA
```

Fig. 5. Comparison of pGPS sequence with the mammalian CENP-B box. Two regions showing 53% and one region showing 58% homology occur along the sequence.

Discussion

The PstI satellite family isolated by us displays a different behavior from those previously studied. The occurrence of the satellites so far studied in lacertids is restricted to species from related genera (pLCS), congeneric species (pLHS), or a single species (pSHS) in *Lacerta saxicola,* (Capriglione et al. 1991, 1994). In contrast, by hybridization experiments, pGPS appeared conserved in the genome of a large number of species irrespective of the degree of relationship. In fact, not only was it present and hybridized in species considered related to *L. graeca,* as those belonging to the genera Algyroides and Podarcis, but also in such species considered unrelated to it, as *L. vivipara* and *L. viridis.* The lack of positivity in *L. saxicola* confirms a result already reported by us (Capriglione 1995): this species seems to be a kind of outgroup with peculiar genomic characteristics shared only by the other species of the so-called ''*saxicola* group,'' which so far have not been exhaustively investigated.

By *in situ* hybridization experiments, the *PstI* family

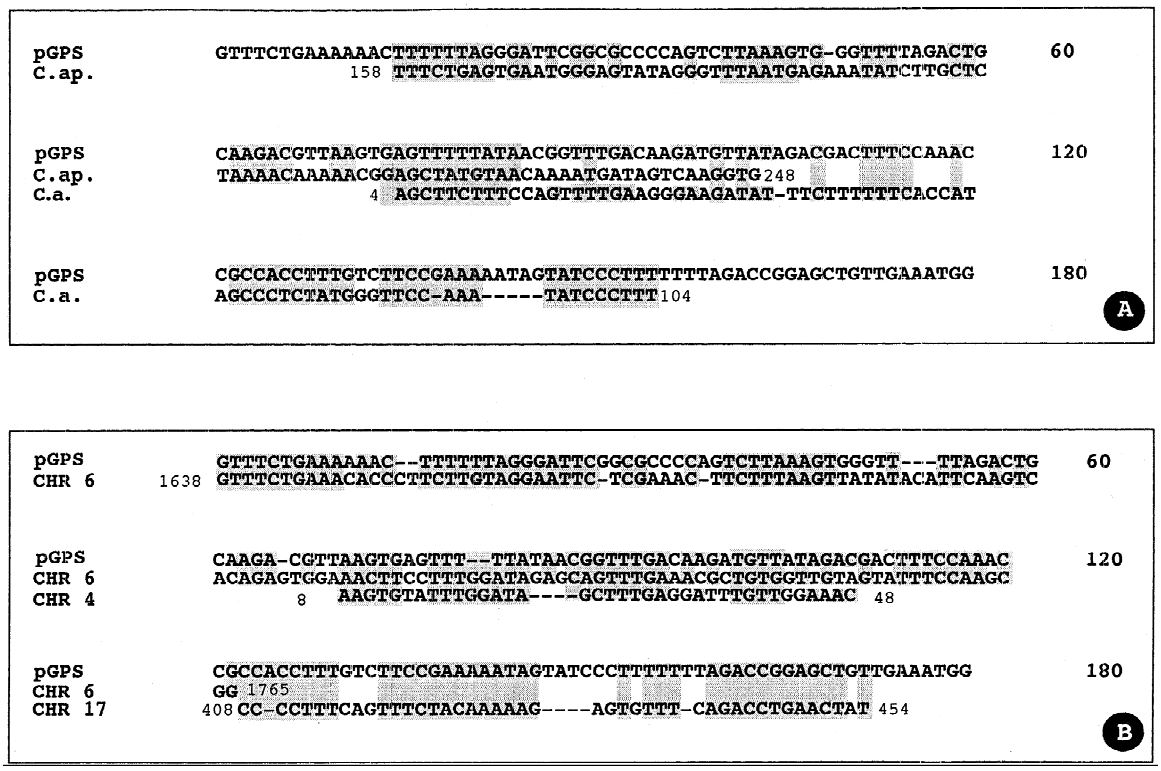

Fig. 6. Comparison of pGPS sequence with a data set of primate alphoid sequences. **(A)** C.ap = *Coebus apella;* C.a. = *Cercopithecus aetiops.* **(B)** Alphoid sequence of human chromosome 6 (CHR6), 4 (CHR4), and 17 (CHR17).

was localized on the centromeres of all *L. graeca* chromosomes. Clusters of satellite were present on *P. tiliguerta* chromosomes too, though mainly localized in the pericentromeric regions. It appears centromerically located only in one of the largest pairs.

Another satellite DNA (pLHS) has been previously localized by *in situ* hybridization in the pericentromeric regions of *P. tiliguerta* (Capriglione et al. 1994). Analysis of homology showed that pGPS shares, in reverse, a 55.3% identity in 85-bp overlap with pLHS sequences of the *Hind III* repetitive family. Though further data would be necessary, we might hypothesize that, as in other vertebrates (John and Miklos 1987), in reptiles too, more highly repeated DNA families might concur in forming constitutive heterochromatin at or near the centromeres. These sequences would show varying amplification according to the species.

Variation in the amount and/or composition of highly repeated DNAs located centromerically or pericentromerically might have contributed to preserving reproductive isolation among related species during evolution (Miklos 1985). However, some of them are evolutionary conserved, but their role, if any, is unknown (Garrido Ramos et al. 1994, 1995; Batistoni et al. 1996); some might be mostly required for centromeric function, as it is probably the case of human alphoid sequences (Willard 1990; Tyler-Smith et al. 1993).

For the characteristics of the sequence and its conservativeness, the satellite described by us could be a good candidate for this functional role. In fact, it is noteworthy that, in its consensus sequence, there is a region with a 72% overlap with the CDEIII element of the yeast centromere. This 25-bp element has resulted in being fundamental to a proper functioning of the yeast centromere (Ng and Carbon 1987; Lechner and Carbon 1991).

Base-specific mutagenesis induced by Na bisulfite in CDEIII reduces centromeric function; in particular, alterations of cytosine at position 14 can abolish or dramatically alter centromeric function in chromosome III of yeast (Clarke and Carbon 1985). In this regard, it is very interesting that this base is conserved in a correct position in the area of sequence overlap in pGPS. Moreover, upstream from this CDEIII-like region, there is a region showing a certain similarity to the human CENP-B box, the 17 *cis*-acting sequence that binds centromere protein B.

Finally, a possible involvement of this satellite in determining centromere functions is also suggested by the homology with the alphoid sequences of humans and neotropical primates. In fact, though more than one satellite concur in forming the centromere of human chromosomes, the alphoid sequences with a 171-bp repeated motif appear as those most probably involved in determining centromeric functions (Masumoto et al. 1989; Ikeno et al. 1994; Sugimoto et al. 1994). Alphoid sequences with a varying amplification are present on the

centromeres of all human chromosomes. On the basis of their characteristics, they were divided into three families (Alexandrov et al. 1988)

The sequence studied by us shows some homology with the human alphoid sequences belonging to all the three families. In view of the phylogenetic relationships between reptiles and mammals, pGPS might be considered a sequence very similar to a hypothetical ''ancestral alphoid'' sequence, from which the various alphoid families of mammals would be probably derived. During evolution, several regions of this alphoid ancestor might have undergone amplification and divergence to varying degrees, though preserving their functional role.

To conclude, there is evidence suggesting that pGPS might have a role in determining the centromeric function in lacertid lizards. A deeper insight would be required, in particular, to detect whether this sequence also shares the functional features of the main centromeric components with other eukaryotes.

References

Alexandrov IA, Mitkevich SP, Yurov YB (1988) The phylogeny of human chromosome specific alpha satellites. Chromosoma 96:443–453

Batistoni R, Pesole G, Marracci S, Nardi I (1995) A tandemly repeated DNA family originated from SINE-related elements in the European Plethodontid salamanders (Amphibia, Urodela). J Mol Evol 40:608–615

Capriglione T, Cardone A, Odierna G, Olmo E (1991) Evolution of a centromeric satellite DNA and phylogeny of lacertid lizards. Comp Biochem Physiol 100B:641–645

Capriglione T, Cardone A, Odierna G, Olmo E (1994) Further data on the occurrence and evolution of satellite DNA families in the lacertid genome. Chrom Research 2:327–330

Capriglione T (1995) Satellite DNA and phylogeny of lacertid lizards. In: Llorente GA, Montori A, Santos X, Carretero MA (eds) Scientia herpetologica. Societas Herpetologica Espanola, Barcelona, p 68

Clarke L, Carbon I (1985) The structure and function of yeast centromeres. Ann Rev Genet 19:29–56

Cooke CA, Bernat RL, Earnshaw WC (1990) CENP-B: a major human centromere protein located beneath the kinetochore. J Cell Biol 110:1475–1488

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

Garrido-Ramos MA, Jamilena M, Lozano R, Ruiz Rejòn C, Ruiz Rejòn M (1995) The EcoRI centromeric satellite DNA of the Sparidae family (Pisces, Perciformes) contains a sequence motive common to other vertebrate centromeric satellite DNAs. Cytogen Cell Genet 71:345–351

Grady DL, Ratliff RL, Robinson DL, McCanlies EC, Meyne J, Moyzis RK (1992) Highly conserved repetitive DNA sequences are present at human centromeres. Proc Natl Acad Sci USA 89:1695–1699

Grellet FG, Delcosso D, Panabieres F, Debeny M (1986) Organization and evolution of a higher plant alphoid-like satellite DNA sequence. J Mol Biol 187:495–507

Haaf T, Materg A, Wienberg J, Ward DC (1995) Presence and abundance of CENP-B box in great ape subsets of primate-specific α-satellite DNA. J Mol Evol 41:487–491

Haaf T, Warburton PE, Willard HF (1992) Integration of human a-satellite into Simian chromosomes: centromere binding and disruption of normal chromosome segregation. Cell 70:681–696

Ikeno M, Masumoto H, Okazaki T (1994) Distribution of CENP-B boxes reflected in CREST centromere antigenic sites on long-range α-satellite DNA arrays of human chromosome 21. Human Molec Genet 3(8):1245–1257

John B, Miklos GLG (1987) The eukaryote genome in development and evolution. Allen & Unwin, London

Lechner J, Carbon J (1991) A 240 kd multisubunit protein complex, CBF3, is a major component of the budding yeast centromere. Cell 64:717–725

Masumoto H, Kenji S, Tuneko O (1989) Alphoid satellite DNA is tightly associated with centromere antigens in human chromosomes throughout the cell cycle. Exp Cell Res 181:181–196

Mayer W, Lutz D (1990) Albumin-immunologische Untersuchungen zur Phylogenese der Gattung *Algyroides* und ihrer systematischen Position gegenüber der Sammelgattung *Lacerta* (Reptilia: Sauria: Lacertidae). Zool Anz 1/2:99–105

Miklos GLG (1985) Localized highly repetitive DNA sequences in vertebrate and invertebrate genomes. In: RJ Mac Intyre (ed) Molecular evolutionary genetics. Plenum Press, New York, p. 241

Ng R, Carbon J (1987) Mutational and in vitro protein-binding studies on centromere DNA from *Saccharomyces cerevisae.* Mol Cell Biol 7:4522–4534

Olmo E, Odierna G, Capriglione T, Merciai BM (1988) Heterochromatin and genome composition in lacertid lizards. Comp Bioch Physiol 89b n 41:1–4

Ragghianti M, Guerrini F, Bucci S, Mancino G, Hotz H, Uzzell T (1995) Molecular characterization of a centromeric satellite DNA in the hemiclonal hybrid frog *Rana esculenta* and its parental species. Chrom Research 3:497–506

Rieder (1982) The formation, structure and composition of the mammalian kinetochore and kinetochore fiber. Int Rev Cytol 79:1–58

Sanger F, Nicken S, Coulson AR (1977) DNA sequencing with chain-terminating method inhibitors. Proc Natl Acad USA 74:5463–5467

Southern EM (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol 98:503–517

Sugimoto K, Furukawa K, Himeno M (1994) Functional cloning of centromere protein B (CENP-B) box-enriched alphoid DNA repeats utilizing the sequences-specific DNA binding activity of human CENP-B *in vitro.* Chrom Research 2:453–459

Sullivan RF, Glass GA (1991) CENP-B is a highly conserved mammalian centromere protein with homology to the helix-loop-helix family of proteins. Chromosoma 100:360–370

Tyler Smith C, Oakey RJ, Larin Z, Fisher RB, Crocker M, Affare NA, Ferguson-Smith MA, Muenke M, Zuffardi O, Jobling MA (1993) Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes. Nature Genet 5:368–375

Willard HF (1990) Centromeres of mammalian chromosomes. Trends Genet 6:410–416