J Mol Evol (1997) 44:327–335

JOURNAL OF MOLECULAR EVOLUTION

Higher-Order Organization of Subrepeats and the Evolution of Cervid Satellite I DNA

Charles Lee,* Dean R. Court, Charles Cho, Jennifer L. Haslett, Chyi-Chyang Lin**

Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta, Canada, T6G 2B7

Received: 28 May 1996 / Accepted: 24 October 1996

Abstract. Based on sequence analyses of 17 complete centromeric DNA monomers from ten different deer species, a model is proposed for the genesis, evolution, and genomic organization of cervid satellite I DNA. All cervid satellite I DNA arose from the initial amplification of a 31-bp DNA sequence. These 31-bp subrepeats were organized in a hierarchical fashion as 0.8-kb monomers in plesiometacarpalia deer and 1-kb monomers in telemetacarpalia deer. The higher-order repeat nature of cervid centromeric satellite DNA monomers accounts for their high intragenomic and intraspecific sequence conservation. Such high intraspecific sequence conservation validates the use of a single cervid satellite I DNA monomer from each deer species for interspecific sequence comparisons to elucidate phylogenetic relationships. Also, a specific 0.18-kb tandem duplication was observed in all 1-kb monomers, implying that 1-kb cervid satellite I DNA monomers arose from an unequal crossover event between two similar 0.8-kb ancestral DNA sequences.

Key words: Cervid satellite DNA — Centromere — Subrepeats — Monomers — Higher-order repeats

* *Present address:* Department of Pathology, Cambridge University, Tennis Court Road, Cambridge, United Kingdom CB2 1QP

** *Present address:* Cytogenetics Laboratory, 5B4.49 W.C. Mackenzie Health Sciences Centre, University of Alberta Hospitals, 8440—112 Street, Edmonton, Alberta, Canada T6G 2B7; e-mail cclin@gpu.srv.ualberta.ca

Correspondence to: C.-C. Lin

Introduction

Mammalian centromeric DNAs consist primarily of tandemly organized, repetitive DNA sequences (i.e., satellite DNA). Although more than one satellite DNA family may exist in the centromeres of a given mammalian species, usually a single centromeric satellite DNA family is predominant, accounting for 3% or more of the species' genome. Each centromeric satellite DNA family is usually characterized by restriction endonuclease-defined repeat units (monomers) of relatively consistent length. For example, primate alphoid DNA consists of 171-bp monomers (Rosenberg et al. 1978) and mouse major satellite DNA is defined by 234-bp monomers (Horz and Altenburger 1981). Although substantial nucleotide sequence variation is commonly observed among monomers of a satellite DNA family, some centromeric satellite DNA monomers can be organized in a hierarchical fashion to yield higher-order repeats (HORs) which have near-identical sequences within a given subfamily (reviewed in Willard and Waye 1987).

In deer, the prominent centromeric satellite DNA family is referred to as either major cervid centromeric satellite DNA or cervid satellite I DNA (because of its homology to bovine satellite I DNA) and is localized to the centromeric region of nearly all cervid acrocentric chromosomes. Interestingly, monomers of this DNA family vary in size between deer species. Those deer belonging to the paleontological division telemetacarpalia (Goss 1983) have their cervid satellite I DNA organized primarily as 1-kb monomers. Other deer, which have retained the more proximal remnants of the second

Table 1. Satellite DNA clones isolated from five deer species

Deer species	Clone	Clone description	Size	Specimen	Source	GenBank No.
Fallow deer	Dd-Pst1	*D*ama *d*ama *Pst*I Clone *1*	804 bp	1	Testis cell line	U53515
Wapiti	Ce-Msp1	*C*ervus *e*laphus c. *Msp*I Clone *1*	680 bp	2	Genomic DNA	U53516
White-tailed	Ov-Msp1	*O*docoileus *v*irginianus *Msp*I Clone *1*	990 bp	3	Genomic DNA	U53517
deer	Ov-Msp2	*O*docoileus *v*irginianus *Msp*I Clone *2*	748 bp	3		
North American	Aa-Msp1	*A*lces *a*lces *Msp*I Clone *1*	987 bp	4	Genomic DNA	U53518
moose	Aa-Msp2	*A*lces *a*lces *Msp*I Clone *2*	712 bp	4		
Mule deer	Oh-Msp1	*O*docoileus *h*emionus *Msp*I Clone *1*	991 bp	5	Liver	U53519
	Oh-Msp2	*O*docoileus *h*emionus *Msp*I Clone *2*	1,970 bp	6	Kidney	U55813
	Oh-Msp3	*O*docoileus *h*emionus *Msp*I Clone *3*	1,975 bp	6		U55814

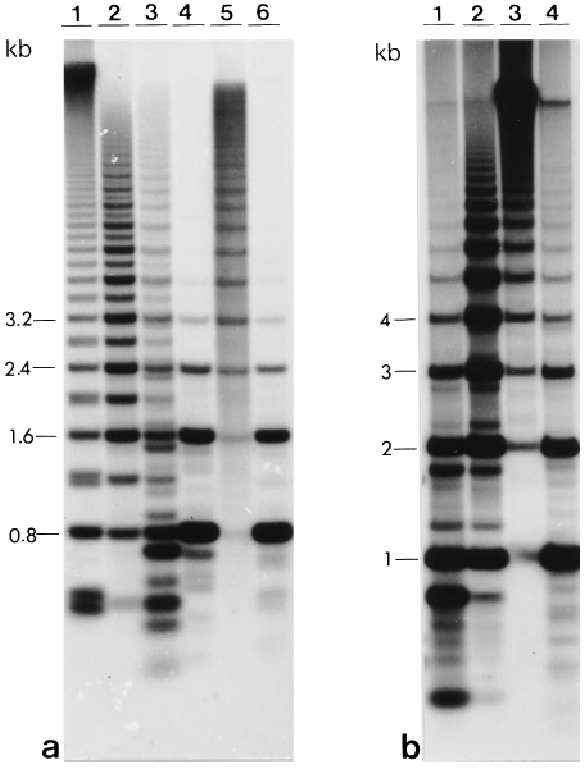


Fig. 1. Tandem organization of cervid satellite I DNA in the genomes of the fallow deer and the white-tailed deer. **a** A Southern blot consisting of fallow deer genomic DNA digested with *Bam*HI, *Hpa*II, *Msp*I, *Pst*I, *Rsa*I, and *Xba*I (lanes 1–6, respectively). DNA fragment sizes are indicated to the left of the figure, demonstrating the prominent 0.8-kb register in all six digestions. **b** A Southern blot consisting of white-tailed deer genomic DNA digested with *Msp*I, *Hpa*II, *Bam*HI, and *Xba*I (lanes 1–4, respectively). *Pst*I- and *Rsa*I-digested DNAs are not shown. DNA fragment sizes are indicated to the left of the figure, demonstrating the prominent 1-kb register in all four restriction enzyme digestions.

and fifth metacarpals (i.e., belonging to the plesiometacarpalia division), have their cervid satellite I DNA primarily organized as 0.8-kb monomers (Bogenberger et al. 1987). Higher-order repeats consisting of multimers of cervid satellite I DNA monomers have not been reported.

The presence of two distinct different-sized monomers in this single satellite DNA family is unusual. Previous sequence analyses on a number of cervid satellite

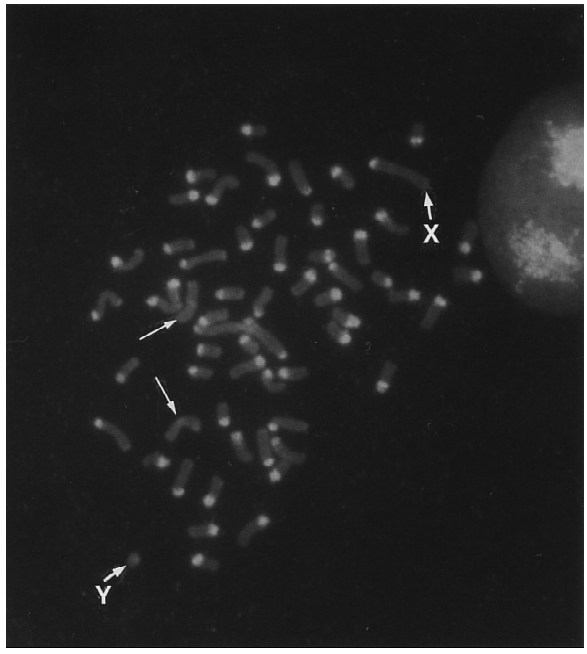
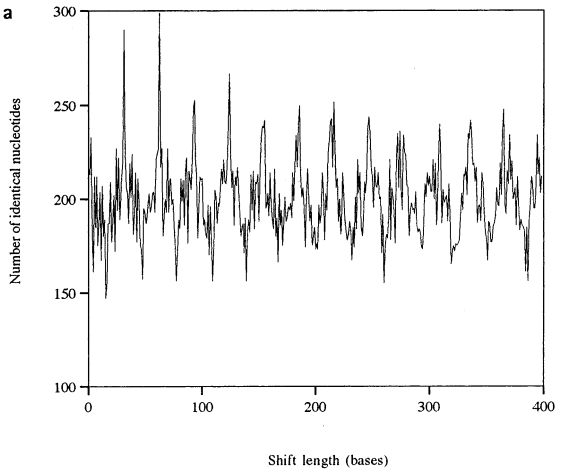


Fig. 2. Localization of the cervid satellite I DNA clone, Dd-Pst1, to metaphase chromosomes of a male European red deer by fluorescence in situ hybridization. The centromeric regions of all acrocentric chromosomes, including the X chromosome (denoted), show large, intense hybridization signals. A single pair of metacentric autosomes (indicated by *arrows*) and the Y chromosome (denoted) did not display such hybridization signals.

I DNA monomers revealed internal 31-bp periodicities (Lee and Lin 1996). In the present study, the 0.8-kb and 1-kb monomers are shown as HORs of their constituent 31-bp subrepeats. Furthermore, evidence is provided which indicates that 1-kb cervid satellite I DNA monomers were formed from a 0.18-kb tandem duplication within a 0.8-kb ancestral DNA sequence. Cumulatively, these data provide insights into the evolution of this satellite DNA and offer an explanation for the presence of two different-sized monomers in a single satellite DNA family.

Materials and Methods

Source of Genomic DNAs. Fallow deer (*Dama dama*) genomic DNA was prepared from an established testis cell line (specimen 1) obtained from the University of Ferrara, Italy. Specimens 2, 3, and 4 were



b

```
Consensus:TCCCTGCCTC AACTCGAGAG GAATCCTGAG T      Subtracted Nucleotides

     1   ↓.↓.AG..TA↓. ..TG.---.. .---.TC.T. .      C,G,A
    29   .G.A.CT.C. ..GA...-.. .G..G↓...A .        T
    60   C.....TGGA G..CAC.... A..GA.CT.. .
    91   ..T.C..... .T.G..CCC. ..GG...C.C A
   122   .TTT↓.AA-- .......... .T-..GC↓.. A        T,G
   152   ...A...T.. C.AAG.G..C ..TG.....C .
   183   --...-.T.G ..A↓G..T.. ..G...CAG. A        A
   212   ......TGG. .CG.G..A.. .G.C....G. .
   243   ↓....A-.... .TG.G....C .CG...CA.T .       C
   274   G..A....AA G.......GA ......C... G↓       TG
   307   ......--.G. .C..A.GC.. ..G.....↓. .       C
   337   -.G...AACA ...A..T.T↓ ..↓G..C--- -        CT,GG
   367   ....C...GT ....T....A T.TA..CC.. G
   398   .T.GG...G. .........A A..C.A.... A
   429   ↓.T..--.... G..AT....T ..GG.....T .       CT
   460   C.....↓..G TG-.GT.... C...T...T. C        A
   491   ..↓.AT..AA .G..-..A.. ..GC.T...T .        A
   523   ...T..ATGG C....C.... A..A....A.. A
   553   ..AG..T... .AA..TT... .G.....T.T G
   584   ..A...TAG. .T.A...A.. .GC...GTG. A
   615   C...AAA-.. .........T ..GAG..T.. .
   645   -....↓.TT.. G....A.... ....A.GA.C .       A
   676   .T..ACAAG. .C..GA.... ..GG.T.CTC .
   707   CAG..A.AG↓ .TG.-..... .G.C..... .         G
   738   .TGTG..... ..G.G..AT. ..--↓.C... A        A
   768   .G↓...---- -.......-.. T..GG.A↓.T A       T,G
   795   ↓......AG                                N
```

Fig. 3. The presence of 31-bp subrepeats in the Dd-Pst1 clone of the fallow deer. **a** A line graph showing increased DNA sequence similarities (peaks) in a 31-bp shift periodicity when the Dd-Pst1 clone is compared with itself. As the complete graph was a symmetrical image defined by a vertical axis of symmetry halfways across the graph, only the results for the first 400 shifts are presented. **b** Derivation of a consensus sequence from optimal sequence alignments of 31-bp subrepeats in the Dd-Pst1 clone. The nucleotide position of the beginning of each subrepeat is indicated to the left. The positions where nucleotides have been removed to optimize alignment of subrepeats are indicated by *small vertical arrows.* These subtracted nucleotides are listed to the right of each subrepeat and are separated by *commas. Dashes* were also occasionally inserted to improve alignments.

purified genomic DNA samples from unknown sexes of wapiti (*Cervus elaphus canadensis*), white-tailed deer (*Odocoileus virginianus*), and the North American moose (*Alces alces*), respectively. These three specimens were obtained from the Bovine Blood Typing Lab, Saskatchewan Research Council, Canada. Mule deer (*Odocoileus hemionus*) genomic DNAs were prepared from a liver sample of a 2-year-old male (specimen 5) and a kidney sample from a second male mule deer (specimen 6).

Cloning and Characterization of Cervid Satellite DNAs. All cervid satellite I DNA clones were obtained from prominent 0.8-kb or 1-kb ethidium-bromide-stained bands, which were observed after electrophoretic fractionation of various restriction-enzyme-digested genomic DNAs. DNA fragments were cloned into pUC19 plasmid vector and propagated in *E. coli* DH5α bacteria. To determine the genomic organization of each clone, ^{32}P-dCTP-labeled clones were used as probes to Southern blots containing endonuclease-digested genomic DNAs. Southern blot hybridizations, washings, and autoradiographic exposures were performed as previously described (Lin et al. 1991). All newly isolated cervid satellite I DNA clones were also biotin labeled for use in fluorescence in situ hybridization (FISH) experiments to red deer chromosomes. Chromosome preparations and FISH experiments were performed as previously described in Lee et al. (1994).

Each centromeric DNA clone was digested to produce smaller fragments for sequencing. All sequencing reactions were performed with dideoxy chain termination Sequenase kits (United States Biochemical Corporation) and read on an ABI DNA Sequencer (model 373). Multiple sequences were combined to determine the complete DNA sequence for each monomer clone. Clones Dd-Pst1, Ce-Msp1, Ov-Msp1, Aa-Msp1, and Oh-Msp1 were designated as representative monomers for the fallow deer, wapiti, white-tailed deer, North American moose, and mule deer, respectively.

DNA Sequence Analyses and Comparisons. Sequence analyses first involved the identification of 31-bp subrepeats in all representative monomer clones, using the method of Pluciennicczak et al. (1982). These methods are described in detail elsewhere (Lee and Lin 1996).

Intragenomic sequence conservation of cervid satellite I DNA monomers was determined from sequence comparisons of multiple clones from a white-tailed deer, a North American moose, and a mule deer. Mule deer clones Oh-Msp2 and Oh-Msp3 each contained two complete 1-kb monomers and therefore individual monomers in these clones were further differentiated by the suffix designation *a* or *b.*

Intraspecific sequence conservation of cervid satellite I DNA monomers could only be examined in the mule deer as it was the only deer species that had complete monomer clones isolated from more than one animal. For these comparisons, clone Oh-Msp1 from one animal was compared to each 1-kb monomer in clones Oh-Msp2 and Oh-Msp3 of a second animal. The sequence divergence of constituent 31-bp subrepeats was also examined through pair-wise comparisons of subrepeats in mule deer clones Oh-Msp1 and Oh-Msp2a.

Interspecific monomer sequence conservation of cervid satellite I DNA was determined from pair-wise sequence comparisons of the ten designated representative monomers (five newly isolated representative monomers from the present study and five previously characterized satellite DNA clones). The previously characterized clones included the Ce-Pst1 clone of the red deer (Lee and Lin 1996), the 1A clone of the Indian muntjac (Bogenberger et al. 1985), the C5 clone of the Chinese muntjac (Lin et al. 1991), the CCSatI clone of the roe deer (Scherthan 1991), and the Rt-Pst3 clone of the caribou (Lee et al. 1994).

Results

Initial Characterization of Newly Isolated Cervid Satellite I DNA Clones

Table 1 lists the nine newly isolated cervid satellite I DNA clones for which complete DNA sequences were obtained. These include one clone from a fallow deer, one clone from a wapiti, two clones from the white-tailed deer, two clones from a North American moose, and three clones from two mule deer. Among these clones, seven clones contained single monomers and two clones (Oh-Msp2 and Oh-Msp3) each consisted of two intact, adjacent monomers.

Southern blot hybridizations of each newly isolated cervid satellite I DNA clone to *Bam*HI-, *Hpa*II-, *Msp*I-, *Pst*I-, *Rsa*I-, and *Xba*I-digested genomic DNAs from the five deer species studied consistently revealed type A–like ladder patterns in a 0.8-kb register for the fallow

Table 2. Percent sequence divergence of 31-bp subrepeats in 1-kb monomers from mule deer*

		Oh-Msp1 subrepeats															
Oh-Msp2a Subrepeats		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	(26′)	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.
2	(27′)	N.C.	55.9	54.5	48.6	54.5	59.4	51.5	51.5	54.3	60.6	54.3	54.5	53.1	50.0	55.9	50.0
3	(28′)	N.C.	54.5	48.5	51.4	55.6	58.1	55.6	54.5	52.9	58.1	58.3	54.8	39.5	38.9	51.4	66.7
4	(29′)	N.C.	54.5	43.8	45.7	54.3	56.2	57.6	50.0	51.5	54.3	54.5	57.1	42.4	51.6	56.2	50.0
5	(30′)	N.C.	50.0	63.6	48.6	50.0	53.3	51.5	43.8	45.2	61.3	58.1	59.4	40.0	43.8	48.3	50.0
6	(31′)	N.C.	42.9	47.1	50.0	47.1	51.5	54.1	48.6	62.5	44.1	54.3	41.9	45.5	46.9	50.0	51.6
7	(32′)	N.C	54.5	40.0	45.9	45.2	48.6	53.1	40.0	50.0	48.5	55.9	52.9	39.4	48.4	66.7	54.8
8	(1′)	N.C. **a**	51.9	51.9	51.4	51.5	64.5	50.0	53.3	56.2	51.6	50.0	67.7	42.4	50.0	50.0	58.6
9	(2′)	N.C	**100.0**	53.1	61.3	42.9	52.9	52.6	58.8	50.0	47.1	48.5	54.5	51.4	51.6	57.6	48.5
10	(3′)	N.C	53.1	**100.0**	40.6	55.2	55.2	62.5	53.3	53.1	56.1	48.5	57.6	48.1	44.4	46.9	56.7
11	(4′)	N.C.	61.3	40.6	**100.0**	46.2	50.0	41.7	48.6	56.4	48.7	52.8	45.9	47.4	59.5	52.8	52.8
12	(5′)	N.C.	54.8	51.4	47.4	**100.0**	58.8	60.0	55.9	45.5	65.6	50.0	62.9	41.2	55.9	57.6	60.6
13	(6′)	N.C.	50.0	58.6	54.8	57.6	**96.7**	46.9	62.5	60.0	53.1	63.6	51.6	56.2	46.9	63.6	51.6
14	(7′)	N.C.	52.6	62.5	42.9	62.9	45.7	**96.8**	61.8	51.5	51.6	51.5	66.7	47.0	47.1	51.6	57.6
15	(8′)	N.C.	58.8	50.0	53.1	61.1	65.6	61.8	**100.0**	59.4	48.5	50.0	70.0	46.9	50.0	50.0	51.6
16	(9′)	N.C.	52.9	53.1	63.3	47.1	60.0	48.4	56.2	**96.7**	52.9	54.5	62.5	47.1	46.7	61.3	48.4
17	(10′)	N.C.	45.5	50.0	53.8	62.5	56.2	46.9	50.0	51.6	**96.8**	53.1	57.6	45.7	51.6	55.9	58.1
18	(11′)	N.C.	51.6	58.6	59.5	54.3	69.7	54.5	52.9	61.8	50.0	**90.3**	62.9	51.5	57.6	67.6	63.6
19	(12′)	N.C.	54.5	57.6	45.9	62.9	54.8	63.6	61.8	55.9	60.6	57.6	**100.0**	50.0	51.6	60.6	64.5
20	(13′)	N.C.	50.0	53.1	46.3	48.6	54.8	51.5	48.4	46.9	50.0	47.1	54.8	**90.9**	44.1	56.2	51.5
21	(14′)	N.C.	51.6	42.4	62.2	57.6	53.1	43.8	53.3	60.0	54.8	57.6	54.5	47.1	**86.7**	67.7	63.6
22	(15′)	N.C.	51.5	47.1	50.0	64.7	57.6	54.8	48.5	54.8	57.6	57.6	60.6	45.5	61.8	**93.9**	61.8
23	(16′)	N.C.	50.0	56.7	51.4	58.8	54.8	58.1	51.6	50.0	58.1	60.6	64.5	45.5	56.7	54.8	**96.6**
24	(17′)	N.C.	56.2	48.5	51.4	62.5	54.5	45.7	57.6	53.1	55.9	54.5	57.1	47.1	52.9	64.5	53.1
25	(18′)	N.C.	45.2	48.4	56.2	62.9	58.1	69.7	62.5	51.6	51.5	51.5	68.8	43.8	60.0	55.9	61.3
26	(19′)	N.C.	57.6	50.0	52.6	51.5	53.1	47.1	54.5	43.8	45.5	48.6	51.4	42.1	45.2	48.5	47.1
27	(20′)	N.C.	42.4	59.4	56.4	42.4	64.5	66.7	53.1	67.7	46.9	63.6	51.6	48.5	54.8	64.5	62.5
28	(21′)	N.C.	43.8	58.1	54.3	58.8	50.0	58.8	46.9	51.6	60.6	62.9	59.4	54.8	56.2	61.3 **b**	61.3
29	(22′)	N.C.	47.1	48.4	47.4	60.0	61.3	48.4	56.2	54.8	64.5	50.0	59.4	41.9	46.7	54.5	**70.0**
30	(23′)	N.C.	53.1	40.6	56.8	51.4	57.6	51.4	61.8	53.1	53.1	53.1	51.5	39.4	54.8	55.9	52.9
31	(24′)	N.C.	45.9	61.3	42.9	56.2	50.0	64.7	50.0	41.9	59.4	48.4	67.7	47.1	50.0	51.6	66.7
32	(25′)	N.C.	44.4	46.4	54.8	41.9	64.5	44.1	53.1	64.5	54.5	54.3	52.9	51.6	50.0	61.3	53.1
33	(26′)	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.

* *Key:* **a,** Sequence comparisons of corresponding subrepeats (bold boxes). **b, c,** Regions of 3 adjacent subrepeats sharing sequence similarity of approximately 70%. **d, e,** Regions of 2 adjacent subrepeats sharing sequence similarity of approximately 70%. Subrepeats in Oh-Msp2a are renumbered (in brackets) according to their near sequence identity to corresponding Oh-Msp1 subrepeats

deer and wapiti (e.g., Fig. 1a) and a 1-kb register for white-tailed deer, North American moose, and mule deer (e.g., Fig. 1b). These ladder patterns of hybridized fragments were characteristic of satellite DNA.

Fluorescence in situ hybridization experiments with each newly isolated satellite DNA clone consistently resulted in pronounced hybridization signals at the centromeric region of all red deer acrocentric chromosomes (e.g., Fig. 2). These chromosomal deposition patterns were similar to that previously observed with a cervid satellite I DNA clone from the red deer (Lee and Lin 1996).

DNA Sequence Analyses of Cervid Satellite I DNA

31-bp Subrepeats in Monomer Clones

Line graphs were produced with data from DNA sequence shifts and self-comparisons of all newly isolated representative monomers. Each graph exhibited ''in-frame'' peaks at approximately every 31 single-base shifts (e.g., Fig. 3a). This indicated the presence of 31-bp subrepeats in all monomers investigated. A consensus sequence was derived from the subrepeats in each representative monomer clone (e.g., Fig. 3b), and each consensus sequence exhibited a high degree of sequence similarity to the bovine 1.715-satellite-subrepeat consensus sequence.

The sequence divergence of 31-bp subrepeats, among monomers of a given deer species, was ascertained by comparing subrepeats in the mule deer clone Oh-Msp1 with subrepeats in clone Oh-Msp2a (isolated from a second mule deer). The results of all possible pair-wise sequence comparisons are shown in Table 2. Starting with subrepeat 9 of the Oh-Msp2a and subrepeat 2 of Oh-Msp1, all consecutive subrepeats in the Oh-Msp2a clone shared identical or near-identical sequence identity with

Table 2. Continued

Oh-Msp1 subrepeats																		
17	18	19	20		21		22	23	24	25	26		27	28	29	30	31	32	33
N.C.	N.C.	N.C.	N.C.	**c**	N.C.		N.C.	N.C.	N.C.	N.C.	N.C.	**a**	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.
51.5	51.5	54.5	48.5		**67.6**		60.6	51.4	51.6	60.6	57.6		**100.0**	54.3	50.0	62.9	55.9	54.5	54.5
48.4	58.3	41.9	50.0		58.8		**71.9**	56.2	59.4	48.5	52.9		51.5	**100.0**	54.5	59.4	50.0	53.6	62.1
41.9	51.4	51.6	45.5		43.8		63.6	**66.7**	52.9	58.8	54.5		50.0	54.5	**100.0**	57.1	51.6	54.5	54.8
51.6	51.5	38.7	50.0		62.1		62.5	45.5	60.6	48.3	53.1		62.5	60.0	55.9	**100.0**	48.4	50.0	53.1
48.6	43.8	50.0	66.7		54.3		51.4	40.0	44.1	53.1	54.1		55.9	50.0	53.1	50.0	**100.0**	48.4	44.0
56.2	48.5	51.4	48.5		45.5		43.8	40.0	51.6	54.3	50.0		46.9	45.7	58.1	45.2	51.4	**93.5**	51.7
54.5	61.3	50.0	53.1		58.1		67.7	47.2	51.6	45.5	45.2		54.5	61.3	50.0	54.3	45.2	45.5	**100.0**
56.2	45.2	62.5	48.5		54.3		43.8	62.5	48.5	40.6	52.8		55.9	54.5	54.5	50.0	47.1	54.5	46.4
46.9	46.7	50.0	59.4		57.1		59.3	47.1	59.4	42.9	46.9		58.6	46.9	45.5	60.0	47.1	53.6	51.9
54.3	56.2	52.6	59.5		51.4		45.7	55.6	42.9	48.6	54.3		48.6	51.4	54.3	51.4	50.0	45.9	50.0
62.5	56.0	50.0	45.5		60.6		58.8	52.6	58.8	48.6	47.1		55.9	54.3	54.3	53.1	47.1	46.9	51.7
50.0	58.1	54.5	64.5		53.1		59.4	48.6	47.1	61.3	54.5		56.2	54.8	59.4	54.8	51.5	50.0	68.0
45.9	61.5	42.9	60.6		58.8		50.0	50.0	63.6	41.9	51.5		52.9	55.6	57.6	54.3	54.1	50.0	53.3
58.1	68.0	54.5	53.1		46.9		54.5	63.9	48.5	51.5	48.5		51.5	55.9	51.5	45.5	43.8	47.2	55.6
56.2	43.8	43.8	61.3		48.4		54.5	50.0	45.2	58.1	45.2		52.9	52.9	54.1	45.2	59.4	50.0	61.3
61.8	62.5	42.4	51.5		58.8		65.6	52.9	54.8	54.3	53.1		59.4	58.1	55.9	57.6	45.5	42.4	51.7
62.5	53.1	48.4	65.6		62.9		57.6	55.9	59.4	55.9	57.6		57.1	62.9	60.6	57.6	54.3	52.9	59.4
48.6	64.0	51.5	51.6		62.5		57.6	51.4	58.1	50.0	48.4		54.5	54.8	58.3	59.4	41.9	52.9	71.4
45.5	41.4	40.6	55.9		53.1		54.5	45.5	57.6	51.6	50.0		54.5	45.2	48.6	44.4	47.2	43.8	45.5
59.4	50.0	48.4	59.4		59.4		56.7	58.8	48.4	59.4	58.8		52.9	50.0	61.3	47.1	54.5	56.2	56.7
58.1	54.8	48.4	51.6		61.3	**d**	55.9	51.4	57.6	54.8	54.8		58.1	55.6	55.6	51.5	47.1	57.6	45.5
56.2	61.3	45.7	59.4		61.3		**71.0**	52.6	70.4	53.1	58.1		54.5	63.6	52.9	51.6	44.4	53.1	61.5
96.8	57.1	54.8	50.0		56.2		50.0	**70.6**	51.6	60.0	48.5		54.5	45.2	41.9	51.5	51.4	58.8	57.7
60.6	**100.0**	50.0	51.5		54.8		57.1	52.9	63.0	51.5	48.5		51.5	58.3	51.4	51.5	43.8	45.5	58.6
54.8	50.0	**100.0**	45.5		41.9		46.9	50.0	40.6	54.8	51.4		54.5	42.4	50.0	41.9	50.0	55.6	51.7
53.1	52.9	38.7	**93.5**		52.8		52.9	50.0	58.1	57.6	64.7	**e**	51.5	50.0	56.2	42.9	67.6	51.6	57.6
54.8	58.1	41.9	58.8		**96.8**		54.5	47.4	51.5	50.0	55.9		**69.7**	58.8	43.8	59.4	54.3	40.6	54.8
51.6	57.6	48.4	51.5		56.2		**96.8**	58.8	63.0	58.1	51.6		63.6	**74.2**	64.5	64.5	52.9	43.8	64.3
74.2	57.1	51.5	50.0		50.0		56.2	**90.9**	48.4	47.1	48.5		52.9	53.1	61.3	41.9	41.2	54.5	51.6
51.5	**70.0**	41.7	48.6		57.6		62.5	48.6	**96.3**	50.0	38.7		53.1	61.8	54.3	57.6	45.5	53.1	55.6
58.1	51.5	51.6	57.6		50.0		56.2	62.2	55.6	**96.8**	51.6		58.8	48.5	55.9	51.6	58.8	51.4	53.3
N.C.	N.C.	N.C.	N.C.		N.C.		N.C.	N.C.	N.C.	N.C.	**100.0**		N.C.	N.C.	N.C.	N.C.	N.C.	N.C.	N.C.

subsequent corresponding subrepeats in the Oh-Msp1 clone (Table 2, region a). Excluding sequence similarities of corresponding subrepeats, pair-wise comparisons revealed sequence similarities from 38.7% to 74.2% with an average sequence identity of 53.4%. Certain juxtaposed subrepeats in Oh-Msp1 also shared approximately 70% sequence similarity to noncorresponding, adjacent subrepeats in Oh-Msp2a (Table 2, regions b–e). Figure 4 schematically illustrates those noncorresponding, juxtaposed subrepeats which share high sequence similarity. Segment b–c (encompassing nucleotides 466–650 of Oh-Msp1) shares 70% sequence similarity to segment b′–c′ (harboring nucleotides 869–985 and 1–63 of Oh-Msp2a). Since nucleotides 869–985 and 1–63 of Oh-Msp2a are analogous to nucleotides 651–837 of Oh-Msp1, nucleotides 466–837 of Oh-Msp1 could constitute a tandem duplication of 186 bp (See top panel of Fig. 4). Likewise, 70% sequence similarity is demonstrated between segments d–e and d′–e′, again substantiating the notion that nucleotides 466–837 of Oh-Msp1 represent a 186-bp tandem duplication.

Intragenomic Sequence Conservation

Sequence similarities between cervid satellite I monomers of an individual animal were determined for three deer species. Based on the sequence identity of clone Ov-Msp2 to the first 753 bp of Ov-Msp1 (both clones were isolated from the same white-tailed deer), an intragenomic sequence conservation of 95.6% was calculated. Mismatches between Ov-Msp1 and Ov-Msp2 consisted of 22 single-base substitutions as well as six single-base insertions, three single-base deletions, and a two-base deletion in Ov-Msp2 (Fig. 5a). Likewise, in a North American moose, the last 712 bp of clone Aa-Msp1 showed 97.1% sequence similarity with clone Aa-Msp2. Mismatches during this comparison included 20 single-base substitutions and one single-base insertion in Aa-Msp1 (Fig. 5b). Clones Oh-Msp2 and Oh-Msp3 provided the complete DNA sequences of four different 1-kb monomers from the same mule deer. All possible pair-wise comparisons between these four 1-kb monomers revealed sequence similarities from 96.4% to 99.6% with an average of 97.6% (Table 3). Approxi-

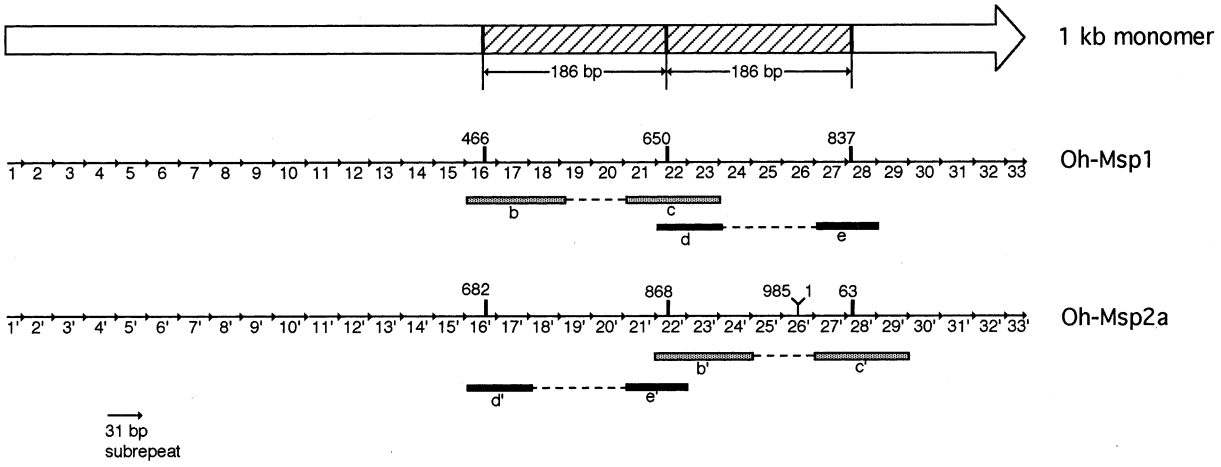


Fig. 4. Higher-order organization of 31-bp subrepeats and a 0.18-kb tandem duplication in the 1-kb monomers, Oh-Msp1 and Oh-Msp2a. Subrepeats 16–18 and 21–23 of Oh-Msp1 share approximately 70% sequence similarity with subrepeats 22′–24′ and 27′–29′ of Oh-Msp2a and are indicated by speckled bars b, c, b′, c′, respectively. Subrepeats 22–23 and 27–28 of Oh-Msp1 also share 70% sequence similarity with subrepeats 16′–17′ and 21′–22′ of Oh-Msp2a and are indicated by solid bars d, e, d′, e′, respectively. The position of nucleotides which define the 0.18-kb duplication are indicated in the Oh-Msp1 and Oh-Msp2a monomers. The tandem duplication is also illustrated in the top panel by *boxes with diagonal lines.*

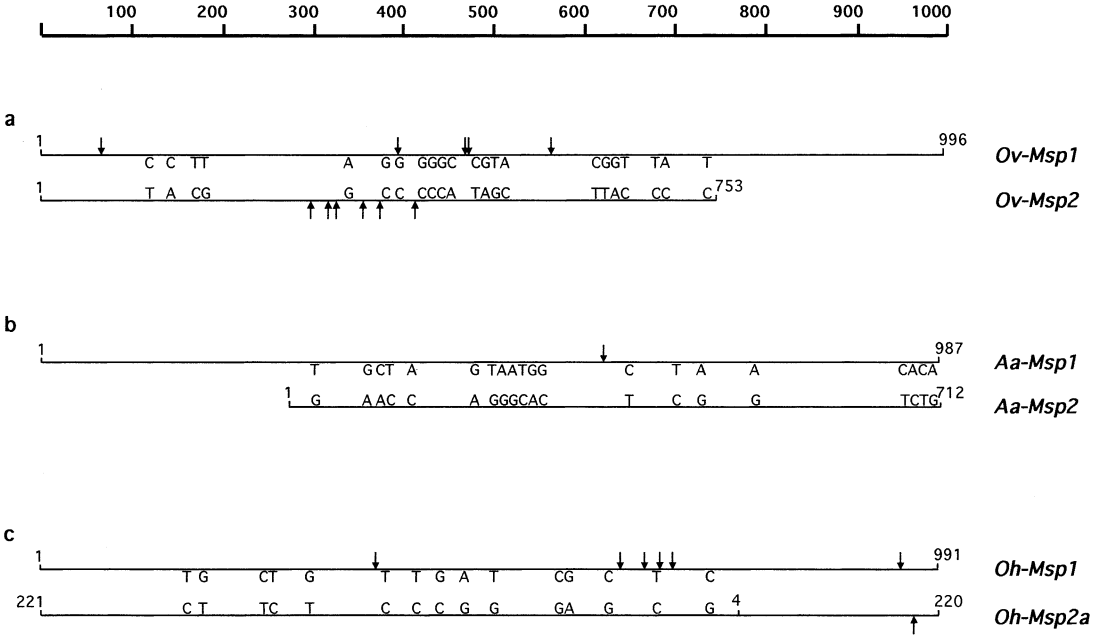


Fig. 5. Intragenomic and intraspecific monomeric sequence comparisons. **a** Intragenomic sequence comparisons of Ov-Msp1 to Ov-Msp2 and **b** Aa-Msp1 to Aa-Msp2. **c** Intraspecific sequence comparisons between Oh-Msp1 and Oh-Msp2a. Single-base substitutions are denoted by C, T, A, or G and single-base insertions are designated by *vertical arrows.* Certain nucleotide positions are also indicated for each monomer clone. The *scale* at the top of the figure provides relative base positions for the various nucleotide changes.

mately one-half of the mismatches observed during these sequence comparisons consisted of single-base substitutions and the other half of the mismatches were single-base insertions/deletions (Data not shown).

Intraspecific Sequence Conservation

DNA sequence similarities of over 95% were also consistently observed for intraspecific sequence comparisons between the Oh-Msp1 clone of one mule deer and the constituent monomers of the Ov-Msp2 and Ov-Msp3 clones from a second mule deer (Table 3). Approximately three-quarters of the mismatches observed during these comparisons consisted of single-base substitutions and one-quarter consisted of single-base insertions/deletions (e.g., Fig. 5c).

Interspecific Sequence Conservation

Alignment strategies for interspecific sequence comparisons between the representative monomer clones of

Table 3. Intragenomic and intraspecific sequence homologies

	Clones compared		
	Clone 1	Clone 2	Homology
Intragenomic:	Ov-Msp1	Ov-Msp2	95.8%
	Aa-Msp1	Aa-Msp2	97.1%
	Oh-Msp2a	Oh-Msp2b	99.6%
	Oh-Msp3a	Oh-Msp3b	96.4%
	Oh-Msp2a	Oh-Msp3a	97.1%
	Oh-Msp2a	Oh-Msp3b	97.3%
	Oh-Msp2b	Oh-Msp3a	97.4%
	Oh-Msp2b	Oh-Msp3b	97.7%
Intraspecific:	Oh-Msp1	Oh-Msp2a	97.4%
	Oh-Msp1	Oh-Msp2b	97.8%
	Oh-Msp1	Oh-Msp3a	97.1%
	Oh-Msp1	Oh-Msp3b	95.6%

ten different deer species are shown in Fig. 6. The comparison results are summarized in Table 4. Maximum sequence similarity was achieved between the 0.8-kb monomers and 1-kb monomers when a specific region, of approximately 186 bp, was first removed from each 1-kb monomer. Optimal sequence alignments with the CCSatI clone of the roe deer was achieved when the CCSatI clone was treated as a partial sequence of a 1-kb monomer. Specifically, a gap of 260 bases was required after nucleotide 730 and a 184-bp region (i.e., nucleotides 219–402) was removed from this clone prior to sequence comparisons. Relatively high interspecific sequence similarities were observed between clones Oh-Msp1 (mule deer) and Ov-Msp1 (white-tailed deer) (94.7%), Ce-Msp1 (wapiti) and Ce-Pst1 (red deer) (92.3%), Dd-Pst1 (fallow deer) and Ce-Pst1 (red deer) (87.8%), and Dd-Pst1 (fallow deer) and Ce-Msp1 (wapiti) (87.3%).

Discussion

Intragenomic and Intraspecific Sequence Conservation of Cervid Satellite I DNA

High intragenomic and intraspecific sequence conservation of cervid satellite I DNA monomers was established in the present study. The number of nucleotide changes observed between monomers were relatively small and appeared to contain an excess of nucleotide substitutions over deletions/insertions in some sequence comparisons (e.g., Fig. 5b,c). Furthermore, nucleotide changes appeared to be nonrandomly distributed in certain intragenomic sequence comparisons (e.g., Fig. 5a,b). However, these observations are based on a limited number of monomers from an animal/deer species. Additional cervid satellite I DNA monomer clones are required to ascertain whether indeed one form of mutation is favored and/or hotspots for mutations exist in these centromeric DNAs. Nevertheless, the observed high intraspecific sequence conservation warrants the use of a single cervid satellite I DNA monomer (from each deer species) for interspecific phylogenetic sequence comparisons.

According to Whitehead (1993), there are currently 41 known deer species in the world, of which one is probably extinct and six to seven species belong to the genus *Muntiacus.* Based primarily on distinct geographical distributions rather than specific marked physical differences, 196 subspecies of deer have now been proposed. Results from interspecific sequence comparisons of representative monomers from ten different deer species imply that the white-tailed deer and the mule deer are very closely related. This is consistent with other studies suggesting a close genetic relationship between these two *Odocoileus* species (e.g., Derr et al. 1991). In 1777, Erxleben identified the North American elk (wapiti) as the distinct species *Cervus canadensis* (Kurten and Anderson 1980). However, its similar morphological characteristics and ability to readily interbreed with the red deer have already led others to consider the wapiti as a subspecies of *Cervus elaphus.* Arguments have also been raised for the reclassification of the fallow deer as a species of the genus *Cervus* rather than as the separate genus *Dama.* The high degree of sequence similarity between cervid satellite I DNA clones from the red deer, fallow deer, and wapiti in the present study implies substantial genetic relatedness between these three deer species and hence substantiates such a reclassification.

Higher-Order Organization of 31-bp Subrepeats in Cervid Satellite I Monomer Clones

The presence of 31-bp subrepeats in isolated cervid satellite I monomer clones from ten different deer species (Bogenberger et al. 1985; Yu et al. 1986; Lee and Lin 1996; and the present study) suggests that these subrepeats exist in cervid satellite I DNA monomers of all deer species. These 31-bp subrepeats have often been undetected because of the extent of their intersubrepeat sequence variation. In the present study, detailed sequence comparisons among subrepeats of two mule deer monomers have revealed an average sequence divergence of 46.6%. However, corresponding subrepeats of each monomer exhibit near or complete sequence identity, suggesting that each cervid monomer represents a hierarchical organization of these constituent subrepeats. Higher-order repeats in mammalian centromeric satellite DNA families usually exhibit much more sequence homogeneity than their constituent basic repeat units (Willard and Waye 1987). Thus, in the context of these centromeric DNA monomers as higher-order repeats, it is not surprising that extremely high intragenomic and

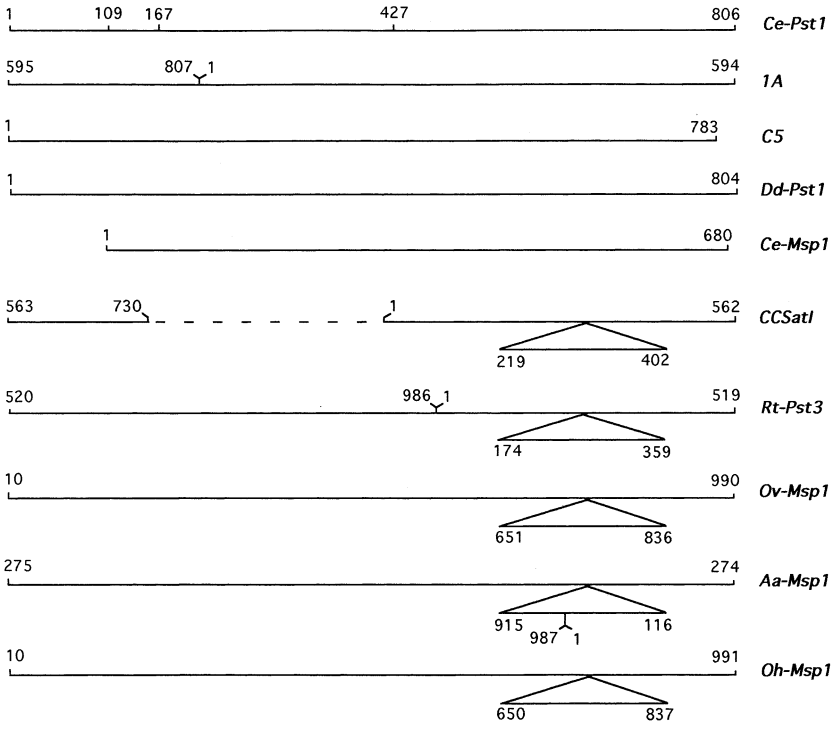


Fig. 6. Alignment strategies for interspecific sequence comparisons of ten representative cervid centromeric satellite I DNA monomer clones. Certain nucleotide positions for each monomer are indicated and a 260-bp gap in the CCSatI DNA sequence is represented by a *dotted line*. Specific 0.18-kb DNA sequences, which were removed from the 1-kb monomers to obtain maximum sequence similarities with 0.8-kb clones, are denoted by *triangles*.

Table 4. Interspecific sequence homology between ten representative centromeric satellite DNA clones

	1A	C5	Dd-Pst1	Ce-Msp1	CCSatI	Rt-Pst3	Ov-Msp1	Aa-Msp1	Oh-Msp1
Ce-Pst1	77.6	74.7	87.8	92.3	73.0	78.9	78.1	77.6	78.2
1A	—	84.9	71.9	77.9	69.5	75.7	73.9	74.8	74.4
C5		—	75.3	74.7	67.1	73.1	72.0	72.7	72.3
Dd-Pst1			—	87.3	73.2	76.5	77.4	77.2	77.4
Ce-Msp1				—	71.6	77.9	77.3	77.5	77.9
CCSatI					—	74.5	73.1	72.8	73.6
Rt-Pst3						—	84.8	76.3	85.2
Ov-Msp1							—	73.9	94.7
Aa-Msp1								—	74.3

intraspecific sequence similarities were observed among cervid satellite I DNA monomers. Similarly, 1.4-kb bovine satellite I monomers also likely represent higher-order repeats of diverged 31-bp subrepeats since less than 3% sequence divergence was reported between independently isolated monomer clones (Taparowsky and Gerbi 1982). Such a conservative nature for this and other mammalian centromeric higher-order repeats may imply the existence of structural and/or functional constraints on these centromeric DNA sequences.

Evolution of 1-kb Cervid Satellite I DNA Monomers from a 0.8-kb DNA Sequence

It is uncommon to find two distinct different-sized monomers within a single satellite DNA family. However, cervid satellite I DNA was shown to be primarily organized into 0.8-kb monomers in plesiometacarpalia deer and into 1-kb monomers in telemetacarpalia deer. Interspecific sequence comparisons between 0.8-kb and 1-kb monomers demonstrated that the additional 0.18-kb DNA is localized to a specific region within the 1-kb monomers (Fig. 6). Intraspecific sequence comparisons of subrepeats from different 1-kb monomers revealed that each 1-kb monomer contains adjacent 0.18-kb segments which share approximately 70% sequence similarity (Fig. 4). This leads to the postulation that 1-kb cervid satellite I DNA monomers are indeed derived from a 0.18-kb tandem duplication within an original 0.8-kb DNA sequence (Lee and Lin 1996), possibly by means of an unequal crossing over exchange.

Therefore, it is proposed that a primordial 31-bp DNA sequence was initially amplified in an ancestral species to bovids and cervids some 25 million years ago (Fig. 7). Approximately 26 tandemly arranged subrepeats (resulting from this initial amplification event) produced a 0.8-

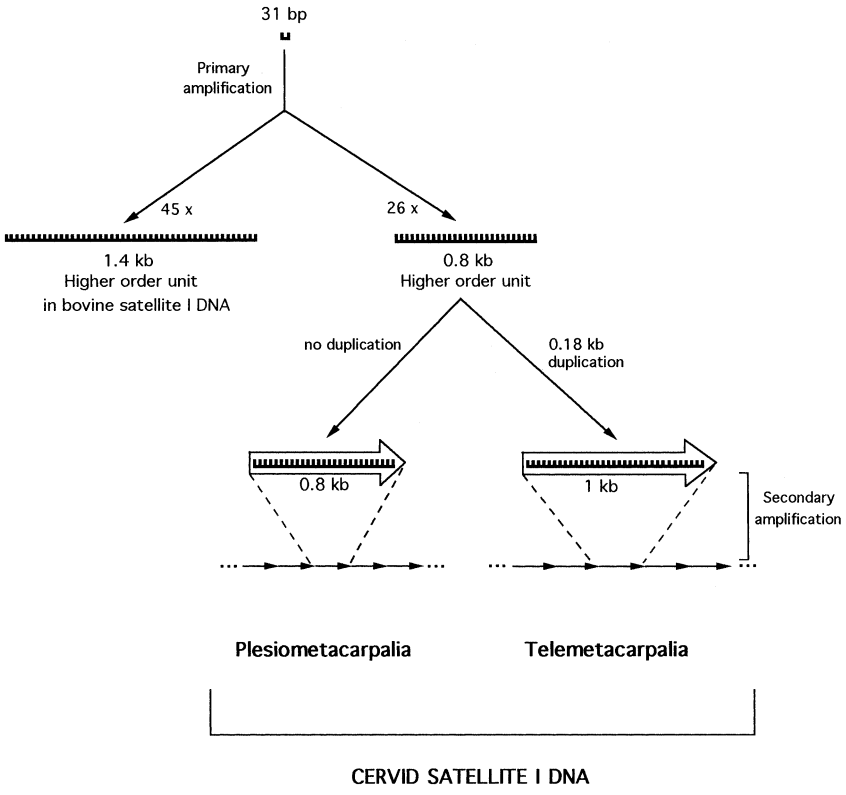


Fig. 7. Proposed genesis of cervid satellite I DNA. A 31-bp DNA sequence is amplified 26 times to produce a higher-order 0.8-kb DNA sequence in an ancestral deer species and 45 times to produce a higher-order 1.4-kb DNA sequence in bovine. In telemetacarpalia deer, a 0.18-kb tandem duplication occurs, resulting in a 1-kb DNA sequence. No duplication occurs in plesiometacarpalia deer, yielding a 0.8-kb DNA sequence. These DNA sequences were subsequently amplified to produce 0.8-kb and 1-kb monomers in present-day deer species.

kb DNA unit in a progenitor deer species. A 0.18-kb duplication then occurred in this sequence, resulting in the eventual amplification of a 1-kb monomer (higher-order repeat) in telemetacarpalia deer. In plesiometacarpalia deer, the 0.8-kb DNA unit did not experience this duplication and was consequently amplified as a 0.8-kb monomer (higher-order repeat).

Acknowledgments. The authors would like to thank Drs. Francesco Fontana and Michele Rubini of the University of Ferrara for providing the fallow deer cell line and Dr. Yves Plante for providing the DNA samples of the wapiti, white-tailed deer, and the North American moose. This work was supported by a Natural Sciences and Engineering Research Council (NSERC) of Canada operating grant to C.-C.L. C.L. was a recipient of an Alberta Heritage Foundation for Medical Research (AHFMR) studentship in support of a Ph.D. dissertation.

References

Bogenberger JM, Neumaier PS, Fittler F (1985) The Muntjak satellite IA sequence is composed of 31-bp-pair internal repeats that are highly homologous to the 31-base-pair subrepeats of the bovine satellite 1.715. Eur J Biochem 148:55–59

Bogenberger JM, Neitzel H, Fittler F (1987) A highly repetitive DNA component common to all cervidae: its organization and chromosomal distribution during evolution. Chromosoma 95:154–161

Derr JN, Hale DW, Ellsworth DL, Bickman JW (1991) Fertility in an F_1 male hybrid of white tailed deer (*Odocoileus virginianus*) × mule deer (*O. hemionus*). J Reprod Fertil 93:111–117

Goss RJ (1983) Deer antlers: regeneration, function, and evolution. Academic Press, New York, pp 42–51

Horz W, Altenburger W (1981) Nucleotide sequence of mouse satellite DNA. Nucleic Acids Res 9:683–696

Kurten B, Anderson E (1980) Pleistocene mammals of North America. Columbia University Press, New York, pp 309

Lee C, Lin CC (1996) Conservation of a 31 bp bovine subrepeat in centromeric satellite DNA monomers of *Cervus elaphus* and other cervid species. Chromosome Res 4:427–435

Lee C, Ritchie DBC, Lin CC (1994) A tandemly repetitive, centromeric DNA sequence from the Canadian woodland caribou (*Rangifer tarandus caribou*): its conservation and evolution in several deer species. Chromosome Res 2:293–306

Lin CC, Sasi R, Fan Y-S, Chen Z-Q (1991) New evidence for tandem chromosome fusions in the karyotypic evolution of Asian muntjacs. Chromosoma 101:19–24

Pluciennicczak A, Skowronski J, Jaworski J (1982) Nucleotide sequence of bovine 1.715 satellite DNA and its relation to other bovine satellite sequences. J Mol Biol 158:293–304

Rosenberg H, Singer M, Rosenberg M (1978) Highly reiterated sequences of simiansimiansimiansimiansimian. Science 200:394–402

Scherthan H (1991) Characterization of a tandem repetitive sequence cloned from the deer *Capreolus capreolus* and its chromosomal localisation in two muntjac species. Hereditas 115:43–49

Taparowsky EJ, Gerbi SA (1982) Sequence analysis of bovine satellite I DNA (1.715 gm/cm^3). Nucleic Acids Res 10:1271–1281

Willard HF, Waye JS (1987) Hierarchical order in chromosome-specific human alpha satellite DNA. Trends Genet 3:192–198

Whitehead GK (1993) The Whitehead encyclopedia of deer. Swan Hill Press, Shrewsbury, England, pp 238–242, 313–314

Yu LC, Lowensteiner D, Wong EFK, Sawada I, Mazrimas J, Schmid C (1986) Localization and characterization of recombinant DNA clones derived from the highly repetitive DNA sequences in the Indian muntjac cells: their presence in the Chinese muntjac. Chromosoma 93:521–528