Genet. Sel. Evol. 32 (2000) 599–619

© INRA, EDP Sciences

Original article

Isolation of subtelomeric DNA sequences labelling sheep and goat chromosome ends

Daniel VAIMAN[a], Ana BRUNIALTI[a], Mohamed BENSAADA[b], Céline DERBOIS[a], Anne VAIMAN[a], Allan CRAWFORD[c], Philippe METEZEAU[b], Edmond P. CRIBIU[a,*]

[a] Laboratoire de génétique biochimique et de cytogénétique, Département de génétique animale, Institut national de la recherche agronomique, 78352 Jouy-en-Josas Cedex, France

[b] Laboratoire de cytométrie analytique, Institut Pasteur, 25 rue du Dr Roux, 75014 Paris, France

[c] AgResearch Molecular Biology Unit, University of Otago, P.O. Box 56 Dunedin, New Zealand

(Received 3 May 2000; accepted 15 September 2000)

Abstract – Two techniques that make it possible to isolate telomere DNA are presented, using sheep as an example. The first technique is based upon the screening of a sheep BAC library with PCR amplified DNA segments preserved from high-power laser beam irradiation. Twenty-three BACs hybridising to 13 subtelomeric regions in sheep and goats were obtained (out of 27 in the sheep complement), of which 13 recognised more than one region, telomeric or not. Twenty-three microsatellites were isolated from these BACs and 22 were genetically mapped on the sheep international genetic map, always consistently with the cytogenetical localisation in 17 cases out of 22. These results are discussed. The second technique is based upon the selective cloning of subtelomeric enriched DNA. Preliminary results were obtained by this approach.

telomeres / FISH / sheep / goats

Résumé – Isolement de séquences d'ADN subtélomériques chez les ruminants. Deux techniques permettant d'isoler l'ADN télomérique sont présentées, le mouton servant d'exemple. La première technique est basée sur le criblage d'une banque de BACs ovins avec des segments d'ADN amplifiés par PCR préservés de l'irradiation par un laser à forte puissance. Vingt trois BACs s'hybridant à 13 régions sub-télomériques (parmi les 27 du génome ovin) ont été obtenus, dont treize reconnaissent plus qu'une région unique, télomérique ou non. Vingt-trois microsatellites ont été isolés à partir des BACs et 22 cartographiés génétiquement sur la carte ovine internationale, de façon cohérente avec la localisation cytogénétique dans 17 cas sur 22. Ces résultats sont discutés. La seconde technique est basée sur le clonage

* Correspondence and reprints
E-mail: cribiu@biotec.jouy.inra.fr

sélectif d'ADN enrichi en régions sub-télomériques. Des résultats préliminaires ont été obtenus par cette approche.

télomères / FISH / ovin / caprin

1. INTRODUCTION

Chromosome ends constitute a complex problem for pluricellular organisms. The reason for this lies in DNA replication, which, after unwinding the DNA helix is initiated with an RNA primer followed by the polymerisation of an Okazaki fragment. This RNA primer is degraded later, locally leaving a single-stranded DNA molecule. Inside a chromosome, DNA polymerase fills the hole that is generated and the extremities are ligated; in contrast to the mammalian chromosome end, a single strand overhang of 50–100 nucleotides persists [11]. Therefore, at each cell division, the telomere end shortens, at least in all somatic and non-cancerous cells, where the telomerase gene is not expressed. The mammalian chromosome end consists of 10 to 20 kb of TTAGGG repeats — very recently shown to be organised in a loop structure [9] — whose length varies according to the species and the number of cell generations since birth. Most knowledge about mammalian telomeres has been collected in mice and humans. However, it would be interesting to obtain data about telomeric regions in other mammalian species, particularly those of economic interest such as ruminants, where recently developed genomic tools and applications could benefit from such knowledge. For instance, the construction of genetic maps of ruminants has been an intense activity of animal geneticists in the last few years [2,4,19]. However, the true coverage of linkage groups along chromosome length has been addressed only statistically, and no tool is truly available today to measure the distance between the last marker of a linkage group and the physical chromosome end. Furthermore, the development of telomeric probes could help assess the role of cryptic chromosome translocations associated with infertility in sheep. Finally, since the evolution of mammalian genomes involves the fusion and fission of ancestor chromosome segments, telomeric probes developed in one species could be hybridised in another to detect remnants of ancient telomeric regions inside chromosomes. Since the interspecific hybridisation of large DNA probes such as BACs or YACs work well, such an approach would be easy to implement, as long as the probes are available.

Ruminant telomeric sequences are composed of the TTAGGG paradigm sequence [10], present at the end of all chromosomes. By contrast, subtelomeric sequences have been shown in humans to be characteristic of one or a few chromosomes [7,8], suggesting that the same situation could be observed in other species. Until now, however, to the best of our knowledge, no effort has been devoted to clone this type of subtelomeric chromosomal regions in domestic species. In this study, our aim was to develop molecular techniques making it possible to isolate such sequences in sheep. Two original approaches have been elaborated and will be presented, together with some preliminary results suggesting that as in humans, subtelomeric sequences are present on a few chromosomes in ruminants. The telomeric BACs obtained were subcloned in order to isolate microsatellites. These new markers were incorporated on the international genetic map, making it possible to ascertain the genetic position of the markers.

2. MATERIALS AND METHODS

2.1. Cell lines

A primary sheep cell line SBK (54, XY) was used to prepare metaphase chromosomes prior to laser irradiation. Cells were grown to confluency in RPMI medium with 10% foetal calf serum, and plated at low density on 3.5 cm plates before laser treatment. For FISH analysis, a primary fibroblast cell culture derived from a 57, XX sheep goat hybrid was used [6], making it possible to simultaneously map the FISH probes in both species.

2.2. Laser treatment of sheep metaphase chromosomes and PCR amplification

Sheep metaphase chromosomes were spread on a plastic membrane overlaying a 2 cm plate. The metaphases were visualised under a light microscope and a high power laser beam was moved along the chromosome with the exception of the terminal regions according to Metezeau and coworkers [13]. After destruction of the majority of the chromosomes, the remaining DNA material from 10 metaphases was recovered on the plastic film and amplified by DOP-PCR using the 6-MW primer, CCGACTCGAGNNNNNNATGTG [17], and called DATS for DOP-Amplified Telomeric Sequences. PCR was carried out in a MJ thermocycler for a total of 30 cycles. After 5 min at 95 °C, 5 low temperature cycles were performed for 1 min at 94 °C, 1.5 min at 30 °C and 3 min extension at 72 °C with a ramping of 3 min between 30 °C and 72 °C. The samples were then subjected to 25 further cycles of 1 min at 94 °C, 1 min at 62 °C and 3 min at 72 °C, this last time being increased by 1 s per cycle. The resulting material was then labelled by random priming using a Boehringer labelling kit and α-^{32}PdATP.

2.3. Hybridisation to sheep BACs

Five thousand sheep BAC clones plated on two LB-agar 22 cm $\times$ 22 cm trays with 12.5 $\mu g \cdot mL^{-1}$ chloramphenicol constituted the basic biological material. The BACs were obtained as previously described [18]. Duplicate Hybond N+ membranes (Amersham) were hybridised with the telomeric probe. Positive clones were ascertained after gridding by a second hybridisation step.

2.4. FISH analysis

After BrdU incorporation during the late S phase, air-dried chromosomal preparations were obtained using standard procedures [5]. Protocols used for FISH analysis were previously described [1]. Chromosomal DNA was counterstained and R-banded by the direct fluorescent technique described by Lemieux *et al.* [12]. Competition was performed using 10 μL of sonicated goat DNA (200 ng $\cdot$ μL^{-1}) and the prehybridisation was carried out for 2 h 30 min. The slides were screened with a Leica fluorescence microscope and photomicrographs were taken with a Fujichrome 400 Asa colour slide film. Chromosomes were identified using the Texas nomenclature [14].

2.5. Microsatellite isolation and genotyping

BAC DNA was prepared by miniprep as previously described [16]. After complete digestion by *Sau*3A1, BAC fragments were ligated with T4 DNA ligase (Biolabs) in 20 μL of 1X ligase buffer, into 50 ng of the pGEM4Z vector (Promega), linearised by *Bam*H1 and dephosphorylated using CIAP (Boehringer). Ligations were then diluted to 100 μL, and 1 μL was electroporated in 20 μL DH10B laboratory prepared electrocompetent cells. Each transformation was plated on a 10 cm LB agar dish in the presence of 200 $\mu g \cdot mL^{-1}$ ampicillin. After overnight growth, the clones were screened with a mixture of (TG)12 / (TC)12 oligonucleotides fluorescently labelled with the 3′-DIG labelling kit (Boehringer Manheim). Positive clones were checked by screening after gridding, miniprepped and sequenced using a 377 ABI automatic sequencer. Primers were designed in order to obtain an amplification product in the 100–200 bp range. Genotyping was carried out on ovine reference families [4] by PCR in the presence of α-^{33}P dATP, using Taq DNA polymerase from Promega or BRL, as previously described [19]. After a denaturation of 5 min at 95 °C, PCR was carried out for 30 cycles (94 °C 15 s, 58 °C 15 s and 72 °C 20 s) in an MJ thermocycler. Samples were loaded on denaturing 5% acrylamide gels, run at 1 500 volts for 2 to 3 h. The gels were then dried and autoradiographed.

2.6. Construction of a telomere-enriched library

Total blood was collected on EDTA from sheep by punction in the jugular vein, and lymphocytes were prepared after lysis of the red cells by NE treatment as previously described [16]. For each plug, 50 μL of NE containing 10^7 cells were mixed with 50 μL of 1% low melting agarose and poured in a 100 μL mold. After solidification, the agarose plug was incubated twice in the lysis buffer (100 mM EDTA, 1% N-lauryl-sarkosyl, and 1 $mg \cdot mL^{-1}$ Proteinase K) at 55 °C for 12 h and 4 h. Plugs were then rinsed twice in TE 10 mM–1 mM 30 min). The plugs were then equilibrated 20 min on ice in 2 mL 1X *Bal*31 buffer (Promega). The plugs were transferred to 2 mL 1X *Bal*31 buffer with 10 μL (30 u) of *Bal*31. Incubation was carried out at 30 °C for 30 min, 120 min and 400 min. In this last case, an additional 10 μL of enzyme was added after 120 min. Plugs were then melted at 65 °C for 15 min, and treated with β-agarase (Boehringer) for 1 h at 45 °C. One phenol/chloroform extraction was then performed by very gentle mixing of the two phases. The upper phase was collected with a cut tip and drop-dialysed for 1 h against 1X *Eco*R1 buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl and 10 mM $MgCl_2$) using 0.025 μ dialysing membranes (Millipore). After dialysis, the drop was collected in an Eppendorf tube and 3 μL *Eco*R1 (30 u) was added for overnight digestion at 37 °C. The DNA was then phenol/chloroform extracted, precipitated and resuspended in TE 10–1 at 160 $ng \cdot \mu L^{-1}$, based upon agarose gel estimation. For vector preparation, pGEM4Z was first totally digested with *Sma*1, phenol/chloroform extracted, precipitated and totally digested by *Eco*R1. The double-digested vector was loaded on a preparative 1% agarose gel, and the vector was shortly visualised under 320 nm UV light, cut out of the gel purified by the phenol/liquid nitrogen technique, precipitated and resuspended in TE

10–1 at 50 ng · μL^{-1}. Ligation was performed in 20 μL 1X Biolabs ligase buffer, between 1 μL *Sma*1/*Eco*R1 vector (50 ng), 1 μL of telomere-enriched insert (160 ng), with 80 u of T4 DNA ligase (Biolabs, 0.2 μL), for 36 h at 8 °C. Ligation products were diluted 5 times and 1 μL was electroporated in 20 μL home made electrocompetent DH10B cells. Transformed bacteria were plated on LB-agar plates with 200 $\mu g \cdot mL^{-1}$ ampicillin, and screened with a $(CCCTAA)_7$ oligonucleotide γ-^{32}P-end-labelled with T4 polynucleotide kinase (Eurogentec). Positive clones were gridded and re-screened together with a positive and negative control {$(TTAGGG)_7$ and $(CCCTAA)_7$ oligonucleotides, respectively}. After this second screening step, positive clones were partially sequenced from both sides of the insert with an ABI377 sequencer.

2.7. BAC library screening and FIGE

Screening was performed by PCR as previously described [16,18]. BAC DNA was prepared by miniprep [16] for further use in FISH or subcloning experiments. Size evaluation of BAC inserts was carried out by Field Inverted Gel Electrophoresis on a FIGE mapper apparatus (BioRad). Briefly, 400 ng (about 1/5 of a miniprep) of BAC DNA was digested by *Not*1 in 20 μL for 2–3 h and loaded on a 1% agarose gel together with the MidRange and LowRange DNA molecular size marker (BioLabs). The run was performed in the presence of Ethidium Bromide at 14 °C for 18 h at 110 V N/S, 170 V S/N with a linear time gradient from 3 s to 15 s.

3. RESULTS

3.1. BAC screening with DATS

BAC plates were leftovers from the construction of the sheep BAC library [18]. The BACs in these plates were not included in the main library because their individual sizes (evaluated by Field Inverted Gel Electrophoresis) were not consistent with the expected size from the size selection experiment, suggesting that they might have been subjected to internal rearrangements. Three 22 cm × 22 cm trays (about 1 700 white colonies per tray) were thus exploited by screening with subtelomeric sequences. Sixty BACs amongst 5 000 yielded a positive signal after hybridisation with the DATS and were used in FISH experiments on a sheep/goat hybrid cell line [6]. All the BACs hybridised to the chromosomes of both species at homologous locations. Twenty-three resulted in at least one telomeric or subtelomeric signal (Tabs. I and III, Figs. 1 and 2). Amongst these, 10 BACs hybridised to only one chromosome location, four to three chromosome locations while the 9 other resulted in two signals. Not all the chromosome ends were recognised by the 23 BACs, but several BAC yielded hybridisation signals at the same location. The labelled chromosome ends were 1p-1q, 2q, 3q, 7, 8, 9, 17, 11, 16, 20, 22, 26 using the sheep nomenclature, indicating that 13 extremities amongst 30 were labelled by the isolated BACs. The 14q14 (corresponding to goat or bovine 18q14) internal (non-telomeric) chromosomal region was labelled by two BACs (BAC13 and

Table I. FISH mapping results/library screening with PCR markers.

BAC number	Marker name (EPCDV are microsatellites)	Address in the ordered library	Localisation with BACs from the ordered library	Consistency Genetic *vs.* Cytogenetic
46	*EPCDV001*	251e12	1q46(telomeric)	OK
8	*EPCDV002*	91c3	17q26(telomeric)	OK
8	*EPCDV003*	74e10	24q14	OK
8	*EPCDV004*	74e10	24q14	OK
14	*EPCDV005*	461c6		nd
16	*EPCDV006*	multiple bands		nd
17	*EPCDV007*	346d5	2p12	
2	*EPCDV008*	4g8	9q12	OK
8	*EPCDV009*	91c3	17q26(telomeric)	OK
15	*EPCDV010*	37g7 and 46g7	1p43(subtelomeric)	OK
15	*EPCDV011*	29b5 and 31e6	14q13/14q13-14	OK
13	*EPCDV012*	31e6	14q13-14	OK
18	*EPCDV013*	101f6	1q46(telomeric)	OK
19	*EPCDV015*	252g7	24q21(telomeric)	OK
26	*EPCDV016*	436c8	8q27(telomeric)	inconsistent
14	*EPCDV017*	160a6	12q21	
15	*PCDV018*	149f5 and 164f5	15q12/3q34	OK
15	*EPCDV019*	447c8	14q14	OK
24	*EPCDV020*	272e7	22q25(telomeric)	OK
13	*EPCDV021*	158d5	1p37(subtelomeric)	OK
15	*EPCDV022*	158d5	1p37(subtelomeric)	OK
21	*EPCDV023*	33d2	11q32(subtelomeric)	OK
be1	BE1	31g4	all centromeres	irrelevant
be2	BE1	237b9	all centromeres	irrelevant
be7	BE1	151g6	9q34(telomeric)	irrelevant
be8	BE1	165g7	12q16	irrelevant
be9	BE1	189f9	1p37(telomeric)	irrelevant
be10	BE1	575b6	No signal	irrelevant
be3	BE1	590h11	10q12	irrelevant
be4	BE1	596h12	1p12	irrelevant
be11	BE1	603e6	1p31	irrelevant
be5	BE1	629a11	3q23-25	irrelevant
be6	BE1	648a11	xq38(telomeric)	irrelevant

BAC15). Several other internal chromosome regions were recognised by one BAC (Tab. I).

3.2. Genetic analysis

Twenty-three microsatellites called *EPCDV001* to *EPCDV023* were isolated from 17 out of the 23 telomeric BACs (Tabs. I and II, Fig. 2). No microsatellite could be identified for six BACs. Primers were designed for each marker and

Table II. Characteristics of the microsatellites of the study (telomeric in gray boxes). (*continued on the next page*)

Microsatellite	Primer names	Primer sequences	Repeated unit	Size of the amplification product	No. of alleles in the IMF
EPCDV001	EPCDV001.1	GTTGCCTCTCAAAGTGCCTC	(TG)18	180 bp	5
	EPCDV001.2	GGCCTTATGGAAGATTTATGCT			
EPCDV002	EPCDV002.1	CTACAACTAAGACCAAGCAC	(AC)11	140 bp	2
	EPCDV002.2	CTTTATGATTTGCCCCATATC			
EPCDV003	EPCDV003.1	CAAGAACACTGCTGGAATCAC	(TG)20	155 bp	7
	EPCDV003.2	GAGAGGTCTTGTGACCTATGA			
EPCDV004	EPCDV004.1	CAAGGGGTAACAAAGAGTCA	(CA)10	120 bp	4.1 ⇒ 2
	EPCDV004.2	GAAAGGTAGGCAGCAAGAG			4.2 ⇒ 3
EPCDV005	EPCDV005.1	GACTTAGCCTATAGCATGAGA	(CA)3(TA)(CA)6	165 bp	Monomorphic
	EPCDV005.2	CAAAGCCTACGTGTATTCAGAG			
EPCDV006	EPCDV006.1	GTCCTCCTGGTTTGCAAAC	(CA)14	185 bp	4
	EPCDV006.2	CTACTGTCTCTTGCTGGGT			
EPCDV007	EPCDV007.3	GATCTGAAACGTGAAGGGTG	(TG)9	130 bp	2
	EPCDV007.2	GCACTCTAGTATTCTTGCCA			
EPCDV008	EPCDV008.1	GACTTTCCAAGAGCTAAGCG	(GT)12	130 bp	9
	EPCDV008.2	GATCTCCTCTAAGCTCACAC			
EPCDV009	EPCDV009.1	GAGCTTTAAGGGTCCACGA	(GT)6(GA)(GT)10	145 bp	5
	EPCDV009.2	CTTAACCTTCTGGAGGTGAC			
EPCDV010	EPCDV010.3	GCGCACTGCTCTGTCTCC	(CT)16(CA)6	140 bp	6
	EPCDV010.2	CTGACTTGGTCGAACTTGAG			
EPCDV011	EPCDV011.1	GCATTGAGAAAAGAGACTTGC	(TG)17	160 bp	9
	EPCDV011.2	CACTATGGGAAACAGTATGGA			
EPCDV012	EPCDV012.1	GCAGAGATAATCAGAGCTGC	(TG)24	140 bp	11
	EPCDV012.2	CTAAGTAAGACCTGGCTCCT			

Table II. *continued.*

Microsatellite	Primer names	Primer sequences	Repeated unit	Size of the amplification product	No. of alleles in the IMF
EPCDV013	EPCDV013.1	CCATCTGTGCTGTCAGTCA	(TG)15	140 bp	5
	EPCDV013.2	GAGCCTGGTGGGTTATAGT			
EPCDV015	EPCDV015.3	CGTGATTGTCCTGCTTGGA	((CA)4-14)4	120 bp	15.1 ⇒ 4
	EPCDV015.4	AGGAAAATCACGTCAATGTGT			15.2 ⇒ 6
EPCDV016	EPCDV016.1	CTTCCCCTTCATGCATTCTTG	(TG)16	190 bp	No amplification
	EPCDV016.2	CAGTGTGGTATCTAATCCAGC			
EPCDV017	EPCDV017.1	GCACCCCAAGGTTCATTACA	(TG)17	200 bp	6
	EPCDV017.2	CATCCTTGTTGCTGCATATTG			
EPCDV018	EPCDV018.1	GATTCAGCCACTATGGGAAAC	(CA)15	210 bp	No amplification
	EPCDV018.2	CAAGCTTGCATGCCTGCAGGT			
EPCDV019	EPCDV019.1	CCAGAAATGCCTTCCTTCTG	(TG)14	140 bp	8
	EPCDV019.2	GTCTCTGAGGAGTCACATCT			
EPCDV020	EPCDV020.1	GTGTTTACCATTTAGAAATTCC	(AC)16	185 bp	3
	EPCDV020.2	GCCACAGATGACAGGATTG			
EPCDV021	EPCDV021.1	GATCTGCGCACTGCTCTGTC	(CT)15	135 bp	7
	EPCDV021.2	GGCTGACCTGGTCGAACTTG			
EPCDV022	EPCDV022.1	CTGGGATGAGAGAGTTCTTAG	(GT)17	165 pb	7
	EPCDV022.2	GCTCACGACTGAATGACTGA			
EPCDV023	EPCDV023.1	CATCTCCTCCCTCCTAACA	(TG)22	115 bp	7
	EPCDV023.2	CAAGAATGGTTTCTAGGGGC			

Figure 1. Fluorescent *in situ* hybridisation of BAC 17, labelling ovine chromosome 8q28 and the goat homeologous chromosome 9q28.

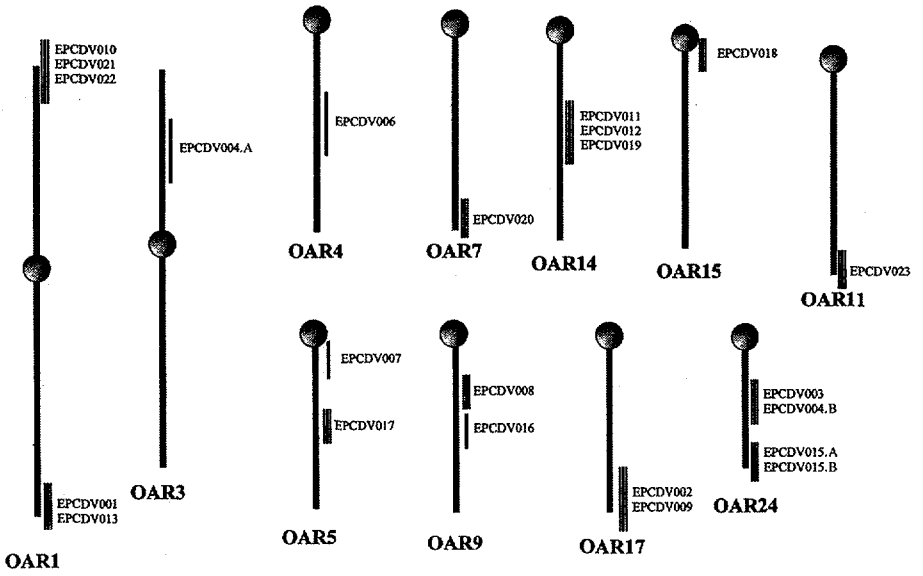
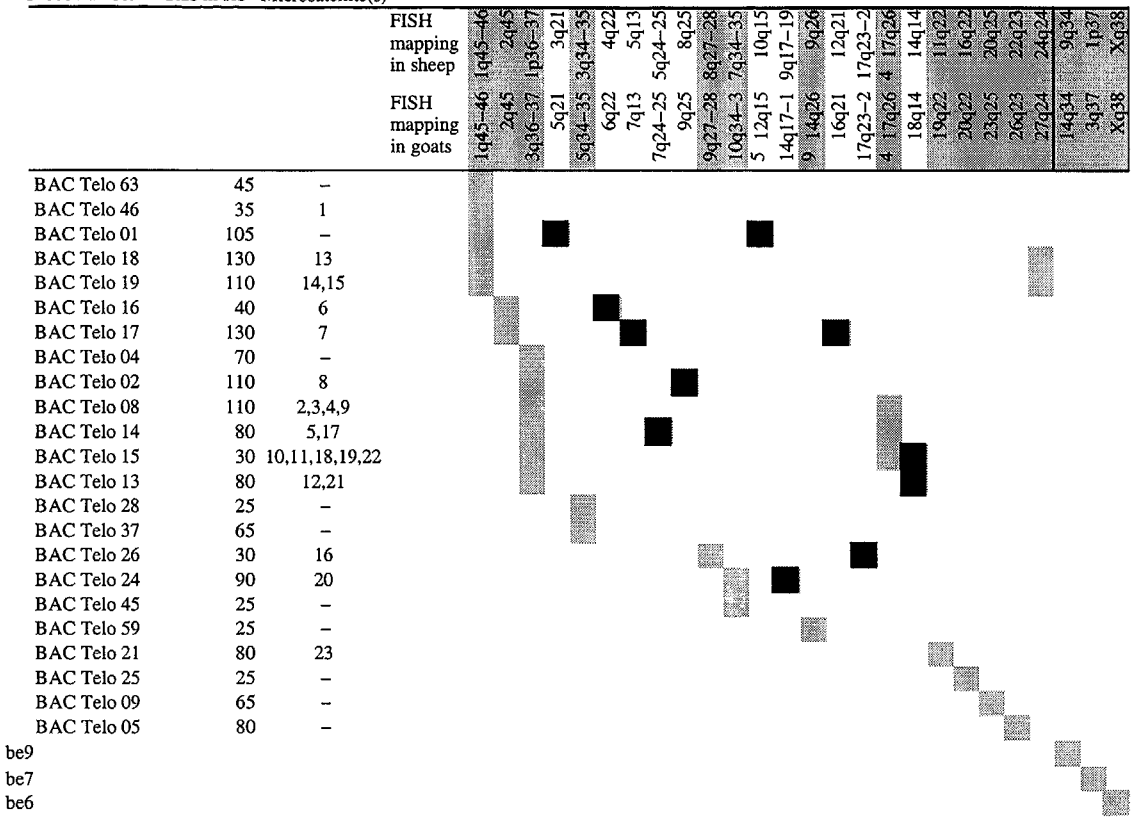


Figure 2. A schematic view of the genetic localisation of microsatellite markers on ovine chromosomes. Double lines indicated markers that have been mapped consistently with cytogenetical data either from the first localisation or after the BAC isolated from the ordered library with the microsatellite primers (see text).

called EPCDVXXX.1 and EPCDVXXX.2. The markers were then genotyped on the international panel of sheep reference families. Genotypes were analysed, making it possible to position them relative to other markers of the sheep International Mapping Flock (IMF) daTab.ase (Tab. III). All were mapped except *EPCDV005* which appeared homozygous on all the parental DNAs, and *EPCDV014*, whose radioactive amplification pattern was too difficult to score (Fig. 2, Tab. IV). Two markers, *EPCDV004* and *EPCDV015* yielded two polymorphic PCR amplification products of different sizes that could be interpreted in a Mendelian transmission. For *EPCDV015*, the two microsatellites were positioned at the end of OAR24, suggesting the existence of a duplication event at this chromosome end. *EPCDV004A* and *4B*, as *EPCDV015A* and *15B* resulted in two signals of different sizes that could be genotyped independently. The distance of 3 centimorgans between *15A* and *15B* (Tab. IV) suggests that duplication probably occurred at the end of ovine chromosome 24. *EPCDV003* and *EPCDV004* originated from the same BAC, the co-localisation of *EPCDV003* and *EPCDV004B* to OAR24 indicates that only the *4B* genotypes are specific of the relevant locus. The other signal 4.A that was biallelic presumably corresponds to an RAPD-type marker. In several cases, the cytogenetical and genetic locations were not consistent, suggesting a high level of chimerism. For instance, BACs 16 and 17 mapped to the telomeric region of sheep chromosome 2q (homeologous to goat chromosomes 2q, Tab. III); however, microsatellites *EPCDV006* and *EPCDV007* were genetically mapped to OAR4 and OAR5 (see Fig. 2). Microsatellites were used for screening in

Table III. Cytogenetic localisations of BACs from the two experiments (telomeric in light gray boxes, non telomeric in black boxes).



order to assess the chimerism of the BACs used in the screening experiment, compared to the bulk of the library which has been stored and ordered [18]. BAC addresses are reported in Table I. Contrarily to the first series of BACs obtained after the screening with DATs, all these BACs from the ordered library resulted in a unique signal in the FISH experiments (Tab. I).

Table IV. Two-points linkage tests.

Marker	Locus	Min Rec	Max Rec	Max Lod	θ	Lod
EPCDV001	*HH36*	8	123	26.355	0.06	22.769
	BMS2263	3	100	25.115	0.03	22.405
	URB14H	17	121	19.173	0.12	15.915
	McM357	9	98	18.794	0.08	15.815
	UWCA46	26	96	9.277	0.22	6.971
	BMS599	14	71	9.072	0.17	7.047
	BM1824	34	93	6.187	0.28	3.534
	MAF4	38	94	5.326	0.30	3.034
EPCDV002	*UCDO48*	0	8	2.408	0.00	2.107
EPCDV003	*ILSTS43*	6	119	27.174	0.05	24.161
	BMS2843	13	132	24.648	0.09	21.939
	BP28	19	147	24.325	0.11	21.007
	BM4005	11	120	23.029	0.08	20.317
	BM737	23	149	22.391	0.13	18.774
	BMS744	19	120	17.762	0.14	14.450
	INR206	11	93	16.060	0.11	13.648
	B1864A	3	53	11.777	0.05	10.269
	ILSTS102	34	122	11.439	0.22	8.153
	JMP29	32	105	8.900	0.23	5.908
	SPN	12	63	8.256	0.16	7.052
	PRM2	3	39	7.950	0.07	7.648
	ALDOA	17	70	7.526	0.20	5.431
	EP0042	0	22	6.623	0.00	6.021
	BMS514	20	70	6.388	0.23	3.535
	UCDO12	7	42	6.023	0.14	5.119
EPCDV004.A	*INR108*	0	34	10.235	0.00	9.633
	BM746	1	36	9.142	0.03	8.236
	TGLA77	1	36	9.142	0.03	8.236
	FCB129	1	35	8.853	0.03	7.961
	TGLA58	4	33	5.634	0.11	4.781
	BMS710	4	33	5.634	0.11	4.781
	BM8118	1	22	5.137	0.04	4.534
	BMS1915	3	28	5.052	0.10	4.148
	BMS460	3	27	4.795	0.10	3.898
	CP34	6	31	4.016	0.16	3.115
	TGLA67	2	20	3.712	0.09	3.411

Marker	Locus	Min Rec	Max Rec	Max Lod	θ	Lod
EPCDV004.2	*EPCDV003*	0	22	6.623	0.00	6.021
	BMS2843	2	20	3.712	0.09	3.110
	ILSTS43	1	16	3.466	0.06	2.864
	INR209	0	11	3.311	0.00	3.010
	INR206	0	11	3.311	0.00	3.010
	INR129	0	11	3.311	0.00	3.010
EPCDV006	*MAF70*	3	116	29.741	0.03	27.010
	McM144	8	85	16.152	0.09	14.082
	CRSP13	7	64	11.445	0.10	9.639
	RM67	4	54	11.138	0.07	9.633
	BM6437	7	54	8.923	0.12	6.815
	MAF50	4	42	7.945	0.09	6.740
	IL6_PM	12	58	7.144	0.17	5.063
	BMS1237	4	38	6.907	0.10	6.001
EPCDV007	*TGLA176*	0	58	17.460	0.00	15.955
	McM380	2	56	13.682	0.03	12.168
	RD2138	3	48	10.397	0.06	8.892
	URB051	0	26	7.827	0.00	7.225
	IL5	1	24	5.702	0.04	5.100
	BM9289	0	18	5.419	0.00	4.816
	TG303	3	29	5.309	0.09	4.405
	RM6	1	20	4.576	0.05	3.974
	IL3	4	26	3.915	0.13	3.011
EPCDV008	*MAF33*	15	146	26.804	0.09	23.529
	BM4630	13	124	22.576	0.10	19.302
	BL1009	24	142	20.183	0.14	16.868
	CSPS11	16	113	17.831	0.12	15.419
	HH29	8	88	16.940	0.08	14.836
	ILSTS11	17	99	13.928	0.15	11.589
	MS1678	16	91	12.605	0.15	10.801
	CSSM55	16	80	10.114	0.17	7.423
	ILSTS8	11	67	9.699	0.14	7.303
	RD240	40	113	7.880	0.26	5.214
	CSSM66	41	112	7.436	0.27	4.829
	McMA10	21	76	7.191	0.22	4.796
	BMS740	37	95	5.727	0.29	2.992
EPCDV009	*VNTR11*	11	113	21.197	0.09	18.491
	ILSTS58	23	120	15.656	0.16	12.721
	McMA20	19	107	14.723	0.15	12.085
	BM8125	22	98	11.295	0.18	8.972
	BM1862	6	57	10.360	0.10	7.995
	BM7136	34	112	9.538	0.24	6.696
	HUJ223	16	71	8.157	0.18	6.951

Marker	Locus	Min Rec	Max Rec	Max Lod	θ	Lod
	TG322	21	80	7.981	0.21	5.977
	URB002	28	91	7.626	0.24	5.382
	EPCDV002	0	25	7.526	0.00	6.924
EPCDV010	*BMS2833*	9	107	21.175	0.08	19.065
	INR197	12	109	19.437	0.10	16.728
	BMS835	23	116	14.761	0.17	11.754
	TGLA127	5	54	10.325	0.09	8.596
	INRA88	5	50	9.280	0.09	7.173
	RM309A	24	92	9.236	0.21	6.629
	HU1177	36	103	7.313	0.26	4.638
	BMS2145	12	58	7.144	0.17	5.062
	HH51	39	103	6.495	0.28	3.582
	BM4301	36	98	6.473	0.27	3.816
	McMA41	37	95	5.727	0.28	3.076
	BM3020	34	87	5.216	0.29	2.906
EPCDV011	*BMS2213*	4	152	38.882	0.03	35.547
	MT2	18	140	23.228	0.11	19.911
	S1915b	16	132	22.535	0.11	19.524
	RD270	18	133	21.497	0.12	18.487
	UWCA28	14	94	14.421	0.13	11.112
	INR138	28	119	13.166	0.19	9.866
	HAUT14	36	125	11.308	0.23	8.042
	MS2355	18	82	9.631	0.18	6.927
	TEX10	6	53	9.336	0.10	7.576
	McMA17	21	84	8.789	0.20	6.390
	TGLA357	35	109	8.666	0.24	5.959
	McM133	21	83	8.586	0.20	6.181
	BM7109	40	107	6.882	0.28	4.041
	INRA63	48	119	6.768	0.30	3.555
	CSRD232	44	105	5.586	0.30	3.057
EPCDV012	*BMS2213*	2	162	44.678	0.01	41.350
	EPCDV011	6	151	36.200	0.04	32.886
	MT2	20	147	23.695	0.12	20.383
	RD270	14	129	23.146	0.10	20.435
	S1915b	14	126	22.379	0.10	19.669
	UWCA28	10	96	17.525	0.09	14.209
	INR138	23	116	14.761	0.17	11.754
	McMA17	17	100	14.160	0.15	11.450
	TEX10	5	67	13.788	0.07	11.713
	HAUT14	39	138	12.746	0.22	9.174
	TGLA357	32	120	11.783	0.21	9.076
	McM133	23	87	8.618	0.21	6.215
	BMS2355	17	73	8.151	0.19	5.751

Marker	Locus	Min Rec	Max Rec	Max Lod	θ	Lod
	BM7109	46	109	5.725	0.30	3.243
EPCDV013	*URB14H*	5	148	36.493	0.03	33.475
	McM357	3	94	23.388	0.03	20.980
	HH36	7	102	21.526	0.06	18.809
	BMS2263	11	97	17.074	0.10	14.966
	EPCDV001	19	109	15.185	0.15	12.475
EPCDV015.A	*EPCDV015.2*	3	105	26.558	0.03	24.446
	RM24L	4	49	9.796	0.08	9.190
	ELN	13	63	7.776	0.17	6.573
	EL01	2	33	7.207	0.06	6.302
	ALDOA	9	47	6.136	0.16	4.947
	BM737	33	77	3.931	0.31	2.038
EPCDV015.B	*EPCDV015.1*	3	105	26.558	0.03	24.446
	RM24L	3	59	13.447	0.05	12.543
	ELN	14	62	7.111	0.19	5.699
	EL01	6	38	5.634	0.14	4.431
	BM737	31	81	5.022	0.28	2.787
	ALDOA	11	46	5.016	0.19	4.112
	HH72	6	32	4.241	0.16	3.940
	BP28	37	80	3.513	0.32	1.804
	UCDO12	0	11	3.311	0.00	3.010
EPCDV016	*BMS1967*	3	80	19.381	0.04	17.868
	URB024	6	70	13.762	0.08	12.257
	BM3215	7	63	11.189	0.10	9.985
	McMA42	1	38	9.720	0.03	8.511
	McM494	17	72	7.942	0.19	6.159
	BM4208	19	68	6.358	0.22	4.886
	RD275	3	29	5.309	0.10	3.852
EPCDV017	*BMS2258*	7	123	27.295	0.05	24.276
	BMS792	3	93	23.101	0.03	20.993
	CRSP14	20	86	9.614	0.19	7.512
	TG137	23	91	9.423	0.20	7.115
	INR192	0	26	7.827	0.00	7.225
	McM527	34	99	7.203	0.26	4.519
	BM1853	3	36	7.147	0.08	5.677
	UCDO37	0	11	3.311	0.00	3.010
EPCDV018	*BR3510*	8	52	7.830	0.13	6.624
	BMS1004	14	45	3.721	0.24	2.259
	BMS1MDR	0	12	3.612	0.00	3.311
	McMA16	0	11	3.311	0.00	3.010
	UCDO37	0	11	3.311	0.00	3.010
	BM4025	16	45	3.118	0.27	1.968

Marker	Locus	Min Rec	Max Rec	Max Lod	θ	Lod
	McMA49	0	10	3.010	0.00	2.709
EPCDV019	*EPCDV012*	3	115	29.451	0.03	27.023
	BMS2213	1	102	28.561	0.01	26.454
	EPCDV011	4	96	22.809	0.04	20.702
	MT2	13	105	17.746	0.11	15.337
	RD270	10	84	14.462	0.11	12.653
	S1915b	12	87	13.922	0.12	11.815
	McMA17	11	72	10.885	0.13	9.080
	HAUT14	22	96	10.871	0.19	8.463
	UWCA28	4	50	10.063	0.07	7.953
	TEX10	3	44	9.303	0.07	7.823
	TEX10	3	42	8.760	0.07	7.284
	TG357	22	81	7.804	0.21	5.999
	INR138	18	70	7.128	0.21	5.327
EPCDV020	*BMS2721*	4	71	15.795	0.05	14.286
	TG272	4	53	10.869	0.07	9.364
	McM156	7	56	9.421	0.11	8.216
	BMS2614	12	63	8.256	0.16	6.751
	BMS1620	9	56	8.214	0.14	6.720
	AE64	8	50	7.354	0.14	6.150
	MS2466	5	41	6.980	0.11	6.077
	ILSTS20	7	42	6.023	0.14	4.819
	McM139	7	41	5.790	0.15	4.299
	ILSTS5	6	38	5.634	0.14	5.031
	BMS823	5	31	4.537	0.14	3.935
	FNTBEB	2	23	4.499	0.08	2.998
EPCDV021	*EPCDV010*	7	127	28.405	0.05	25.692
	INR197	14	104	16.857	0.12	14.148
	BMS2833	12	96	16.150	0.11	14.344
	BMS835	21	108	13.943	0.16	11.233
	PPT	3	54	12.054	0.05	10.849
	RM309A	21	86	9.200	0.20	6.887
	HU1177	32	97	7.448	0.25	5.048
	TG127	6	45	7.330	0.12	5.935
	BMS2145	11	53	6.512	0.17	4.728
	HH51	36	96	6.145	0.28	3.507
EPCDV022	*EPCDV010*	4	125	31.089	0.03	28.078
	EPCDV021	9	110	21.974	0.08	19.258
	BMS2833	11	105	19.123	0.09	16.707
	INRA197	11	95	16.566	0.10	13.854
	BMS835	22	109	13.685	0.17	10.392
	TG127	5	52	9.801	0.09	7.816
	PPT	4	48	9.529	0.08	8.324

Marker	Locus	Min Rec	Max Rec	Max Lod	θ	Lod
	INRA88	6	50	8.577	0.11	6.175
	RM309A	23	81	7.443	0.22	5.146
	BMS2145	13	54	5.853	0.20	3.529
	HH51	41	97	5.080	0.31	2.033
EPCDV023	*McM120*	24	146	21.119	0.14	18.186
	CSSM15	22	129	18.229	0.15	14.986
	CSSM65	7	86	17.209	0.08	15.698
	MB087	2	66	16.551	0.03	13.541
	TAU	2	66	16.551	0.03	13.541
	McMA24	7	74	14.035	0.09	12.828
	CRSP6	38	140	13.498	0.22	10.127
	McMA29	19	101	13.355	0.16	10.150
	BM17132	16	77	9.453	0.17	7.063
	BMS501	1	30	7.413	0.03	5.606
	ETH3	6	45	7.330	0.12	5.824
	RD281h	16	67	7.315	0.20	4.757

3.3. Cloning of sheep subtelomeric sequences

The enrichment protocol was based upon a selection by the vector of inserts carrying *Eco*R1 cohesive ends on one side and blunt ends on the other side (Fig. 3). Two thousand clones of telomere enriched DNA were plated, and screened with the $(TTAGGG)_7$ probe, making it possible to select 18 positive clones (about 1%). The size of the DNA insert was evaluated by agarose gel electrophoresis after *Hin*dIII digestion and was comprised between 1 and 15 kb (average $\sim$ 4 kb). Partial sequencing was carried out from both sides and revealed telomeric repeats on the reverse sequencing side (*Sma*1) for a clone called BE1. Primers were designed from the *Eco*R1 sequence in order to screen the sheep BAC library. The screening with BE1 primers indicated that 20 superpools were positive amongst the 39 of the library, suggesting that BE1 primers recognise repetitive elements in sheep DNA. Whilst it was not possible to identify BAC addresses from the strongest responses in the superpool, increasing the PCR specificity by adding 2% formamide and raising the temperature to 60 °C made it possible to recover 11 different BACs from the ordered BAC library (be1 to be11, Tab. I). These positive BAC clones were then mapped by FISH (Tabs. I and III). Two resulted in centromeric signals, one did not give any clear response and five were inside chromosome arms. The three others (be6, be7) yielded strong signals at the telomeres of chromosomes X, 9 and 1, respectively.

4. DISCUSSION

In this study, two techniques are presented that allow the identification of subtelomeric sequences in sheep, where such sequences have not been cloned

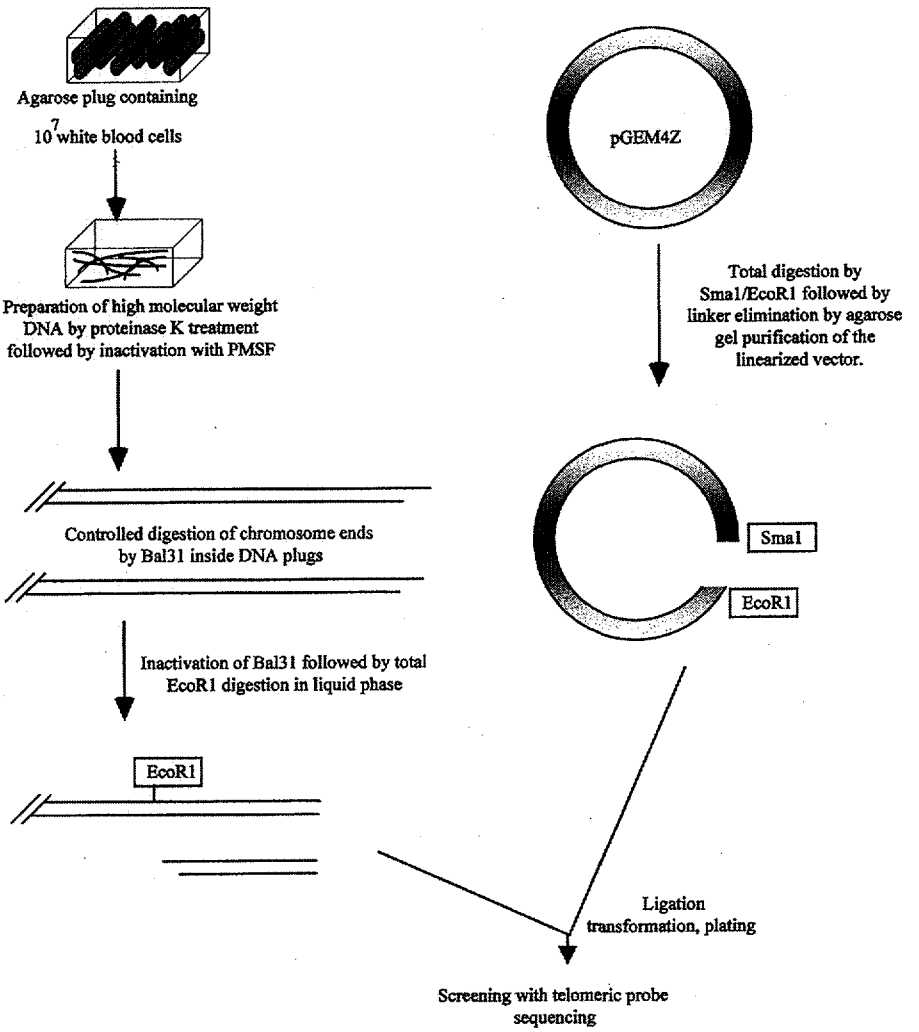


Figure 3. A scheme of the telomere enrichment procedure. Lymphocytes were included in low melting agarose prior to DNA extraction with Proteinase K. The enrichment is based upon the fact that after treatment, unfragmented chromosomes will result in *Eco*R1/*Eco*R1 fragments or in blunt/*Eco*R1 fragments, if they encompass a telomeric region. This technique results in a 100 X enrichment in such sequences over an *Eco*R1 library (see text).

before. The first technique was based upon the screening of a sheep BAC library with telomeric and subtelomeric material preserved from high-energy laser light and DOP-PCR amplified. Since this technique does not give clear information about the distance between the BACs obtained and the true physical chromosome end, another complementary approach was designed. Chromosome ends

have already been cloned in humans using a modified YAC vector containing a unique telomere, where the second chromosome telomere is provided by the genomic DNA [3]. By contrast, the direct cloning of telomere sequences is impossible in classical BAC vectors because the genomic DNA is prepared by partial digestion of high molecular weight DNA. For instance, in our sheep library, *Hin*dIII has been used as the cloning enzyme. Consequently, all the BAC inserts end with a *Hin*dIII restriction site, implicating an impossibility to clone the chromosome end, made of TTAGGG repeats. To circumvent this problem, we decided to identify inside the ordered sheep BAC library [18], the BAC(s) containing the last *Hin*dIII site of the chromosome. For this, a telomere-enriched library was constructed, with clones large enough to encompass the last *Hin*dIII site. The particularity of this library is that the telomeric side of the clones is blunt ended and shortened by the *Bal*31 exonuclease, while the other side is terminated by an *Eco*R1 site, generally rarer than *Hin*dIII sites in mammalian genomes. There is thus a good probability that the last *Hin*dIII site will be comprised in the telomere-enriched DNA, implying that a side of the insert is present in the BAC library. *Bal*31 is used in such a way that not all the telomeric (TTAGGG) repeats are digested, making it possible to screen the enriched library by hybridisation with specific telomeric sequences. The sequence from the *Eco*R1 side makes it possible to choose primers, to screen the sheep BAC library, and to obtain the BAC terminating the chromosome. About 1% of positive clones could be identified in the TED, making it possible to evaluate the enrichment achieved in this library. Indeed, a medium sheep chromosome contains 100 Mb of DNA, which when digested by *Eco*R1 (1 site every 4 000 bp of DNA) yield on average 25 000 fragments, of which two (about 0.01%) are telomeric. Consequently, the protocol designed in this paper results in a 100 X enrichment, and could be efficiently applied to other organisms, where subtelomeric regions deserve further analysis.

BAC clones were identified by the first technique and 23 actually mapped to subtelomeric regions. The hybridisation of 37 other BACs to internal chromosome regions can be explained either by the presence of telomere-like sequences inside chromosomes, or by the non-complete destruction of the non-telomeric sequence by the laser treatment. Often, the 23 BACs hybridising a telomeric region recognised more than one telomere, and sometimes also some non-telomeric regions. These results have been confirmed by the genetic analysis performed after genotyping the microsatellites isolated in the sheep reference families [4]. The hybridisation to different chromosomal regions suggests several possible explanations: either the BACs used were unusually chimerical (whilst it is assumed that the chimerism rate in BACs is around 5%), or some subtelomeric sequences were homologous between different chromosomes, leading to different hybridisation signals. FIGE analysis of the BACs from the first experiment (Tab. I) made it possible to estimate insert sizes between 25 kb and 130 kb, with an average of 69 kb (±35 kb). This figure is about half the average insert size of the clones in the library [18], suggesting that the subset used for screening with DATs was qualitatively different from the rest of the library. A further proof was obtained by screening the ordered library with the different microsatellites isolated, when they yielded hybridisation signals on different chromosomes, such as *EPCDV010* and *EPCDV011*, both isolated

from BAC 15, but genetically mapped to chromosomes 1p and 14, respectively. In all these cases, the two microsatellites belonged to different BACs from the ordered library, substantiating the possibility of a high level of chimerism of the leftover BACs used for screening with DATs. The intrachromosomal regions 14q14 labelled independently by BACs 13 and 15 might be interpreted as the remnant of a "fossil" telomeric region. Actually, the region roughly corresponds to a breakpoint between two human chromosome regions, 16q13 and 16q24 and between mouse regions 8.40 and 8.61, suggesting a possible position for an evolutionary breakpoint [15]. To summarise, this first approach to isolate subtelomeric DNA sequences, was impaired by the low quality of the BACs screened that displayed a high level of chimerism. This was illustrated at the cytogenetic level since different signals were visible in FISH, but also at the genetic level since when different microsatellites were isolated from several of these BACs they happened to be linked to markers from different chromosome locations. By contrast, in the ordered sheep BAC library, the BACs contain very few chimeric inserts. This explains why when the microsatellites originating from a BAC from the first series were used to screen the ordered sheep BAC library by PCR, each microsatellite resulted in one address.

The telomere enrichment technique presented in this paper could be an efficient way to provide telomeric sequences from a wide range of species. The primers from probe BE1 made it possible to identify 11 different BACs including two labelled centromeres, suggesting a high content in repetitive sequences (satellite-type), while three clearly labelled telomeric regions. This approach could be validated in the future by isolating new BE probes. The mapping of microsatellites originating from this material will presumably position them at the ultimate extremity of the chromosomes. They could thus constitute the best markers for following chromosome erosion after cloning experiments in ruminants.

REFERENCES

[1] Bahri-Darwich I., Vaiman D., Olsaker I., Oustry A., Cribiu E.P., Assignment of bovine synteny groups U27 and U8 to R-banded chromosome 12 and 27, respectively, Hereditas 120 (1994) 261–5.

[2] Barendse W., Vaiman D., Kemp S.J., Sugimoto Y., Armitage S.M., Williams J.L., Sun H.S., Eggen A., Agaba M., Aleyasin S.A., Band M., Bishop M.D., Buitkamp J., Byrne K., Collins F., Cooper L., Coppettiers W., Denys B., Drinkwater R.D., Easterday K., Elduque C., Ennis S., Erhardt G., Li L., *et al.*, A medium-density genetic linkage map of the bovine genome, Mamm. Genome 8 (1997) 21–8.

[3] Brown W.R., Molecular cloning of human telomeres in yeast, Nature 338 (1989) 774–6.

[4] Crawford A.M., Dodds K.G., Ede A.J., Pierson C.A., Montgomery G.W., Garmonsway H.G., Beattie A.E., Davies K., Maddox J.F., Kappes S.W., *et al.*, An autosomal genetic linkage map of the sheep genome, Genetics 140 (1995) 703–24.

[5] Cribiu E.P., Matejka M., Darre R., Durand V., Berland H.M., Bouvet A., Identification of chromosomes involved in a robertsonian translocation in cattle, Genet. Sel. Evol. 21 (1989) 555–560.

[6] Cribiu E.P., Matejka M., Denis B., Malher X., Étude chromosomique d'un hybride chèvre × mouton fertile, Génét. Sél. Evol. 20 (1988) 379–386.
[7] Cross S., Lindsey J., Fantes J., McKay S., McGill N., Cooke H., The structure of a subterminal repeated sequence present on many human chromosomes, Nucleic Acids Res. 18 (1990) 6649–57.
[8] Cross S.H., Allshire R.C., McKay S.J., McGill N.I., Cooke H.J., Cloning of human telomeres by complementation in yeast, Nature 338 (1989) 771–4.
[9] Griffith J.D., Comeau L., Rosenfield S., Stansel R.M., Bianchi A., Moss H., de Lange T., Mammalian telomeres end in a large duplex loop, Cell 97 (1999) 503–14.
[10] Gu F., Hindkjaer J., Primed *in situ* labeling (PRINS) detection of telomeric (CCCTAA)n sequences in chromosomes of domestic animals, Mamm. Genome 7 (1996) 231–232.
[11] Klobutcher L.A., Swanton M.T., Donini P., Prescott D.M., All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3′ terminus, Proc. Natl. Acad. Sci. U S A 78 (1981) 3015–9.
[12] Lemieux N., Dutrillaux B., Viegas-Pequignot E., A simple method for simultaneous R or G banding by fluorescence *in situ* hybridization of small single copy genes, Cytogenet. Cell Genet. 59 (1992) 311–312.
[13] Metezeau P., Kiefer H., Lamy G., Delamare G., Analytic and preparative laser scanning cytometry, Pathol. Biol. (Paris) 41 (1993) 276–80.
[14] Popescu C.P., Long S., Riggs P., Womack J., Schmutz S., Fries R., Gallagher D.S., Standardization of cattle karyotype nomenclature: Report of the committee for the standardization of the cattle karyotype, Cytogenet. Cell Genet. 74 (1996) 259–261.
[15] Schibler L., Vaiman D., Oustry A., Giraud-Delville C., Cribiu E.P., Comparative gene mapping: A fine-scale survey of chromosome rearrangements between ruminants and humans, Genome Res. 8 (1998) 901–15.
[16] Schibler L., Vaiman D., Oustry A., Guinec N., Dangy-Caye A.L., Billault A., Cribiu E.P., Construction and extensive characterization of a goat bacterial artificial chromosome library with threefold genome coverage, Mamm. Genome 9 (1998) 119–24.
[17] Telenius H., Carter N.P., Bebb C.E., Nordenskjöld M., Ponder B.A.J., Tunnacliffe A., Degenerate oligonucleotide primer PCR: general amplification of target DNA by a single degenerate primer, Genomics 13 (1992) 718–725.
[18] Vaiman D., Billault A., Tabet-Aoul K., Schibler L., Vilette D., Oustry-Vaiman A., Soravito C., Cribiu E.P., Construction and characterization of a sheep BAC library of three genome equivalents, Mamm. Genome 10 (1999) 585–7.
[19] Vaiman D., Schibler L., Bourgeois F., Oustry A., Amigues Y., Cribiu E.P., A genetic linkage map of the male goat genome, Genetics 144 (1996) 279–305.

To access this journal on line:
www.edpsciences.org