Cytogenetic mapping of 25 goat mammary gland expressed sequence tags (ESTs) (cytogenetic localisation, mammary gland cDNA, comparative mapping)

Fabienne Le Provost, Laurent L. Schibler, A. Oustry-Vaiman, Patrice Martin, E.P. Cribiu

Genet. Sel. Evol. 32 (2000) 311–320

Original article

Cytogenetic mapping of 25 goat mammary gland Expressed Sequence Tags (ESTs)

Fabienne LE PROVOST, Laurent SCHIBLER, Anne OUSTRY-VAIMAN, Patrice MARTIN, Edmond P. CRIBIU*

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

(Received 16 August 1999; accepted 11 January 2000)

Abstract – Today, there is a shift towards a positional candidate approach in the molecular identification of genes. This study reports on an Expressed Sequence Tags (ESTs) mapping initiative in goats, based on sequence information gathered from a previous mammary gland cDNA systematic sequencing project. A total of 25 novel genes was localised cytogenetically on 16 goat chromosomes. Six of these ESTs were found to map to cattle milk QTL regions. These results made it possible to assess the use of ESTs as a shortcut to the molecular identification of some QTLs and as a valuable tool for comparative mapping.

cytogenetic localisation / mammary gland cDNA / comparative mapping / goat

Résumé – Localisation cytogénétique de 25 ESTs caprins issus de glande mammaire. L'identification moléculaire des gènes s'oriente actuellement vers une stratégie de candidats positionnels. Cette étude, réalisée chez la chèvre, décrit la localisation cytogénétique d'étiquettes (ESTs) obtenues précédemment par séquençage d'ADNc de glande mammaire. Au total, 25 nouveaux gènes ont ainsi été localisés sur 16 chromosomes caprins. Six de ces gènes ont été cartographiés dans des régions de QTL laitiers bovins. Ces résultats ont permis d'évaluer les possibilités d'utilisation des ESTs, d'une part, pour l'identification moléculaire des QTL et d'autre part, comme outil de cartographie comparée.

localisation cytogénétique / ADNc de glande mammaire / carte comparée / chèvre

1. INTRODUCTION

A major objective in livestock genome research is to construct comprehensive maps which make it possible to identify loci affecting economically important

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

traits. Detection of these regions is expected to lead to more efficient breeding schemes, particularly marker-assisted selection (MAS), and should improve the accuracy and intensity of selection. Genetic maps constructed in livestock species may be sufficient for detecting genes and QTLs in 5–10 cM intervals and subsequently for starting MAS projects. However, it appears that their efficient use in breeding in most cases should require the identification and cloning of the underlying genes. This could be achieved by standard positional cloning, taking advantage of YAC and BAC libraries. This first step should however be combined with the search for transcribed sequences in these regions using for instance cDNA fishing and exon trapping. In humans, a systematic Expressed Sequence Tags (ESTs) mapping strategy [1, 16] has been developed to efficiently address this crucial step of finding candidate genes for the trait of interest. Thus, there is a rapid shift towards positional candidate cloning, which relies on the availability of a pool of expressed sequences mapped to the same chromosomal region as the disease gene.

In livestock, the use of high resolution comparative mapping [14, 18] makes it possible to take advantage of the spectacular progress achieved in the human and mouse genome projects. In fact, since a virtually complete human transcript map can be expected in the near future [3], the most efficient strategy for positional cloning in livestock could be the comparative positional candidate gene approach [17].

Since the human gene map is already under construction, it would be of great interest to develop limited ESTs project for livestock species, in order to establish tissue-specific expression maps which could provide a shortcut to the molecular identification of some genes or QTLs. In view of the economic importance of milk production traits in ruminants, the mammary gland, which is involved in the synthesis and secretion of specific milk components, was used to generate ESTs in goats [8]. ESTs could provide candidates for known milk QTLs in cattle. We now report on an EST mapping initiative in goats based on sequence information gathered from the cDNA project mentioned above. In our study, 25 novel genes were localised by FISH.

2. MATERIALS AND METHODS

Oligonucleotide primers were designed for a subset of 42 mammary gland ESTs described previously [8]. Our primer pairs were selected in the 3′ region to produce 115–410 bp amplified fragments (Tab. I). Clones were isolated from a goat BAC library using PCR, and DNA was prepared as described previously [15]. FISH mapping experiments were achieved after random priming in the presence of biotin 11 dUTP as described by Bahri-Darwich *et al.* [2]. Goat metaphase spreads were obtained from a 59 XY rob (6;15) primary cell line [5], cultured in the presence of 5′ bromodeoxyuridine during the S phase of the cell cycle [6]. R-banded chromosomes were identified and classified according to the recommendations of the Texas meeting [12].

Human homologues of goat ESTs were searched in the Genbank database using both FASTA (GCG package) and BLAST[1]. Human gene localisations

[1] http://www.ncbi.nlm.nih.gov

were obtained either directly from the FASTA/BLAST output, from the Human Gene Map 98[2], or from the UniGene database[3].

3. RESULTS AND DISCUSSION

3.1. Isolation of BAC clones and cytogenetic localisation

Forty-two primer pairs were designed from goat mammary gland 3′ ESTs obtained previously by random sequencing of oligo-dT primed cDNA [8]. An identical PCR product size was observed between cDNA and genomic DNA in 33 cases, suggesting that it is most probably the last exon which was amplified. No genomic DNA amplification or a different PCR product size was obtained for 9 (21%) primer pairs. Although this percentage appears to be high, contrasting with only 8% when using primers defined in a single exon of goat genes, it is lower than with human ESTs (67%) [7,14]. This could either be the result of primers designed in two consecutive exons (with an intron too large to be amplified in our PCR conditions) or of primers encompassing an exon/intron junction.

A goat BAC library was screened with the remaining 33 primer pairs amplifying genomic DNA, two of which were shown to amplify a repeated sequence. We attempted to map the 31 BAC clones thus isolated by using FISH. This gave rise to 25 cytogenetic localisations on goat chromosomes. For the seven remaining BACs, no specific double spot was detected among a high background noise level, probably because of an excess of repeated sequences in these BACs. The chromosomal locations of the 25 goat mammary gland ESTs are displayed in Table I and Figure 1.

3.2. QTL region targeting using mammary gland ESTs

Six of the 25 ESTs (3–29revR, 2–21revR, 5–17revR, 6–37rev, 4–43 and 17–48) were localised within the 21 previously identified QTL regions affecting milk yield, fat yield or length of productive life [4,13,19]. A statistical analysis based on the hypothesis that these QTL regions covered 57 out of 300 chromosome bands showed however that this distribution could arise randomly. Therefore, using random ESTs generated from a given tissue may not be an efficient way of finding candidates for particular functions. QTL region targeting should be improved when considering only ESTs expressed specifically in the tissue of interest. Nevertheless, caution should be taken with such an approach; indeed "non-specific" expressed genes are involved in mammary gland development and differentiation (STAT5 [9] and prolactin receptor [11]) and may thus influence milk production traits.

[2] http://www.ncbi.nlm nih.gov/genemap98
[3] http://www.ncbi.nlm nih.gov/UniGene

Table I. Goat mammary gland ESTs used for chromosomal localisation and for comparative mapping with humans.

EST	Accession number	Primer pair sequence	PCR product (bp)	Goat chromosomal localisation	Human homology (Accession number)	Sequencing % identity (length in bp)	Human putative Chromosomal localisation
1-14	X73543	01-GGAGGCGGCCAATGAAGAG 02-CAGACCGTCTTCCCAGTGG	168	20q12	Steroid receptor RNA activator (AC005214)	79 (227)	05q*
3-29revR	X73711	01-GTTGGACTGAGACACTAGTG 02-GCAAGCTCTAGACTTTGGTG	180	1q13	DNA sequence from clone 86D1 (AL034349)	67 (345)	06*
5-49rev	X73730	01-CAGGTTCCTTGCTGTCCTG 02-CATACCCTGCTCCCCACC	230	21q12	SFPQ STSG40825 (H79209)	87 (220)	1p31-33*
5-50revR	X73733	01-GCCTGCAGCCCTGGCAGAC 02-GCACCGGCAAGAGCTCCC	186	Xq11-12	cDNA clone IMAGE (N36819)	84 (292)	10
7-04rev	X73739	01-ACCTGTCAATTGCAGGGC 02-AGGAACAAGCTAGGGACC	360	25q21			
8-16R	X73745	01-GGCTGATCCTCAGACCCAC 02-CGTCTGGAGGAGAACTGTC	300	3q37	SNAP23 (AA885306)	91 (137)	15
8-41rev	X73744	01-CCGGTCCTCTGATCCCATC 02-CTGTTCCCCATGCGGAGCC	245	Xq33	Neutrophic tyrosine kinase receptor-related 1 (AI336174)	84 (215)	01/11/15
11-24	X73757	01-GTGTTCCAGGCTAGGCCC 02-GAATCGAGGCCTGCTTGCC	210	6q15	EST (AL042400)	70 (438)	unknown
13-01revR	X73765	01-CTTGGATGCAGGAGCCTG 02-CCAGTCCACACCCTTGAC	300	10q33	DNA sequence from PAC (AL021406)	72 (219)	20p12*
15-03	X73772	01-CACCCTCTCCTGCCAGTC 02-CTTCTATCGGCGGTTCTAG	213	20q12	cDNA clone IMAGE (W63718)	84 (125)	14
15-11	X73774	01-GGACGTGTTTCTGTGGCAG 02-CACAGCCATGCACTTGGG	214	1q33			
17-48	X73781	01-GCCATCTGCCCTGGGGAC 02-CGGAGCCACCAGACCTCTC	266	13q24			

18-12rev	X73790	01-GGCCAGAGCTTGGAGTTTG 02-GTGCCCTGAACGAGGCC	287	8q26	mRNA for KIAA0769 protein (AB018312)	85 (458)	11q13*
1-17	X73544	01-GCCTAAGTAGTTAGAGGGG 02-CTAGACACCAGCTTCCAAG	238	22q24	cDNA clone IMAGE (AA688401)	84 (298)	1q12-31
1-43revR	X73738	01-CGTGTTGAGACTGGGTGTG 02-GCGGCAGTAGGGTTGCTTC	159	26q13-14	cDNA clone IMAGE EST (STSG29704)	69 (296) 88 (131)	4p16 3* 11*
2-21revR	X73706	01-GCATCTCCTTCCCTTCAGAG 02-GTCACGAGCACCACACCCC	313	5q31			
2-25rev	X73707	01-GGGTTTCACTGTGTTCAGC 02-GCCCTAGCATAAGAGGAGG	162	10q34	cDNA clone IMAGE (AI536639)	72 (141)	**14/21q**
3-49revR	X73715	01-GGAAGACACCTGGGGTTTAG 02-GTCTAGCGTTCCCACAGCTG	220	8q28	AIB1 clone 1049G16 (AL034418)	65 (291)	20q12-13.2*
4-07R	X73717	01-CACTGGTACCATCAAGGCC 02-GTATAGGCACACAGGAAGAG	280	10q24+21q23			
4-43	X73726	01-GAGTGGCCAGGCTGGATGAC 02-GCTACACAGCCAAGAGGAGG	215	10q22	cDNA clone IMAGE (N52503)	83 (289)	**15q21***
5-17revR	X73727	01-GAAGGGTACGCACTGTTCTG 02-CCTTGCTAACTGCAGCCAA	165	17q24			
6-23	X73735	01-GAGTACAACACGGAACAAG 02-CTTAGTTGGCCCTCTCGATC	226	4q23	EST similar to Hsp40 NIB1821 (AA027766)	89 (191)	**7q31***
6-37rev	X73734	01-CCACCACGGGCTCGGAC 02-GGGGACTTTCTTTCCCAGTAC	410	13q24	Ubiquitin-conjugating enzyme E2 variant 1 (U39360)	93 (483)	**20q12***
9-10	Z71920	01-GATGAACTGGATGCTGCAAC 02-CCGGGACTGGTCGAATGAG	115	27q12	Ribosomal protein L37 WI8561 (L11567)	89 (137)	**8p11***
9-12	Z71921	01-GTGGACTCTGATGCTACAAG 02-GTTTCCTTGTATTTTAGTGCTC	122	5q35	Chromodomain helicase DNA binding protein 4 (AC006064)	98 (68)	**12p13.3***

Goat chromosomal localisations corresponding to QTL regions affecting milk production are underlined. Human localisations were obtained directly (shown with asterisk) or by the UniGene cluster (others). Human localisations consistent with previous comparative mapping results are written in bold.

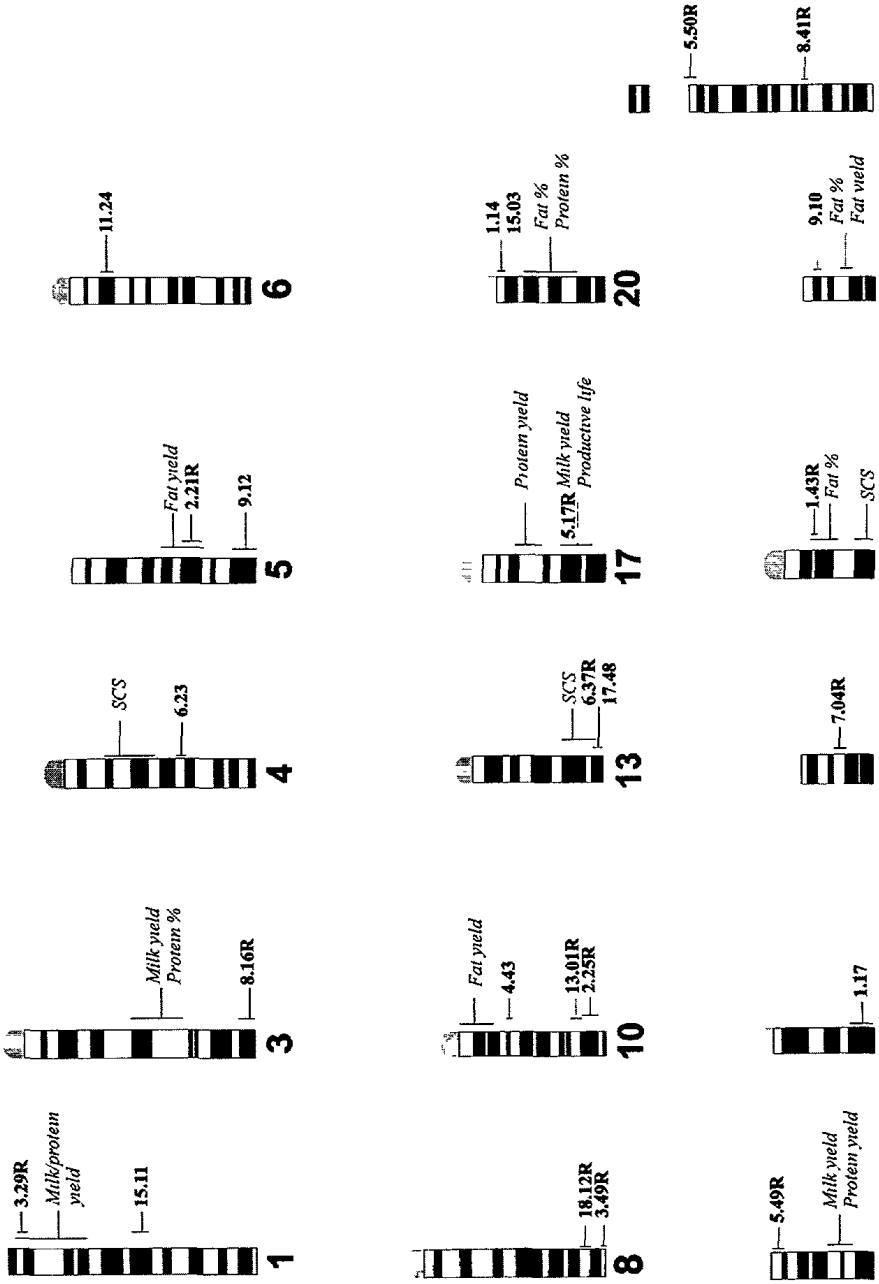

Figure 1. Localisations of goat ESTs on standard R banding karyotypes. The 25 ESTs (in bold) are spread on 16 chromosomes, seven of which carry two or three ESTs. QTL regions for milk production are italicised.

3.3. Comparative mapping between humans and goats

Sequence alignments between goat ESTs and nucleotide sequences in human databases made it possible to identify putative homologues of 19 goat ESTs, with identities ranging between 65 and 98% (Tab. I). The six remaining ESTs displayed no match in public databases and may thus represent novel sequences. Eighteen of the putative homologues have previously been localised in humans: 12 by either FISH or RH, whereas the six remaining localisations were deduced from other ESTs belonging to the same UniGene cluster. Surprisingly, these human gene localisations were consistent with previous comparative mapping results in only seven cases, two of which more accurately delineate conserved segments between humans and goats. Indeed, EST 4–43 confirms the homology of Chir 10q22–23 and Hsap 15 previously obtained using Zoo FISH. Moreover, EST 6–23 made it possible to define a new conserved segment (from LAMB1 to 6–23) between Chir 4q22-q23 and Hsap 7q21.1-q31 (Fig. 2).

In the other cases, localisations were inconsistent with both somatic hybrids and Zoo FISH results. This could be accounted for either by new human-goat chromosomal homologies, erroneous goat localisations, or identification of human paralogues rather than orthologues. The discovery of new conserved synteny seems to be unlikely, given the resolution level of the ruminant-human comparative map [14]. In addition, this study has shown that only 10% of goat BACs isolated with ruminant microsatellites or genes led to inconsistent localisations either with genetic or comparative maps, probably due to chimerism, pseudogenes or gene families. Thus, our study may comprise at most four incorrect goat gene localisations. Conversely, certain human localisations should be considered with circumspection because UniGene clusters may contain some paralogues, as illustrated by the three human locations corresponding to the cytochrome C1 gene[4]. In this study, six putative human homologue localisations were deduced from those of other ESTs belonging to the same UniGene cluster. Likewise, EST 8–41rev showed homologies with human ESTs belonging to the same UniGene cluster and localised on three chromosomes (Tab. I). Therefore, some of the data listed in the Table I are likely to be erroneous due to the difficulty to identify paralogues on the basis of short sequences (200–500 bp). Since ESTs are generated in a cDNA 3′ untranslated region, they may be less conserved during the course of evolution. Consequently, two orthologous genes may exhibit a low percentage of sequence similarity (*e.g.* less than 70% between mice and humans [10], making it difficult to discriminate between the homologous locus (orthologues) and other genes of the same family (paralogues). Therefore, although human ESTs have been shown to be useful in comparative mapping when used in an interspecific way [7], our results suggest that the use of species-specific ESTs in comparative mapping is presently more hazardous. The human genome project will soon lead to completing the entire sequence, which should make it possible to accurately identify gene families and thus improve the detection of gene homologies.

[4] http://www.ncbi.nlm.nih.gov/UniGene

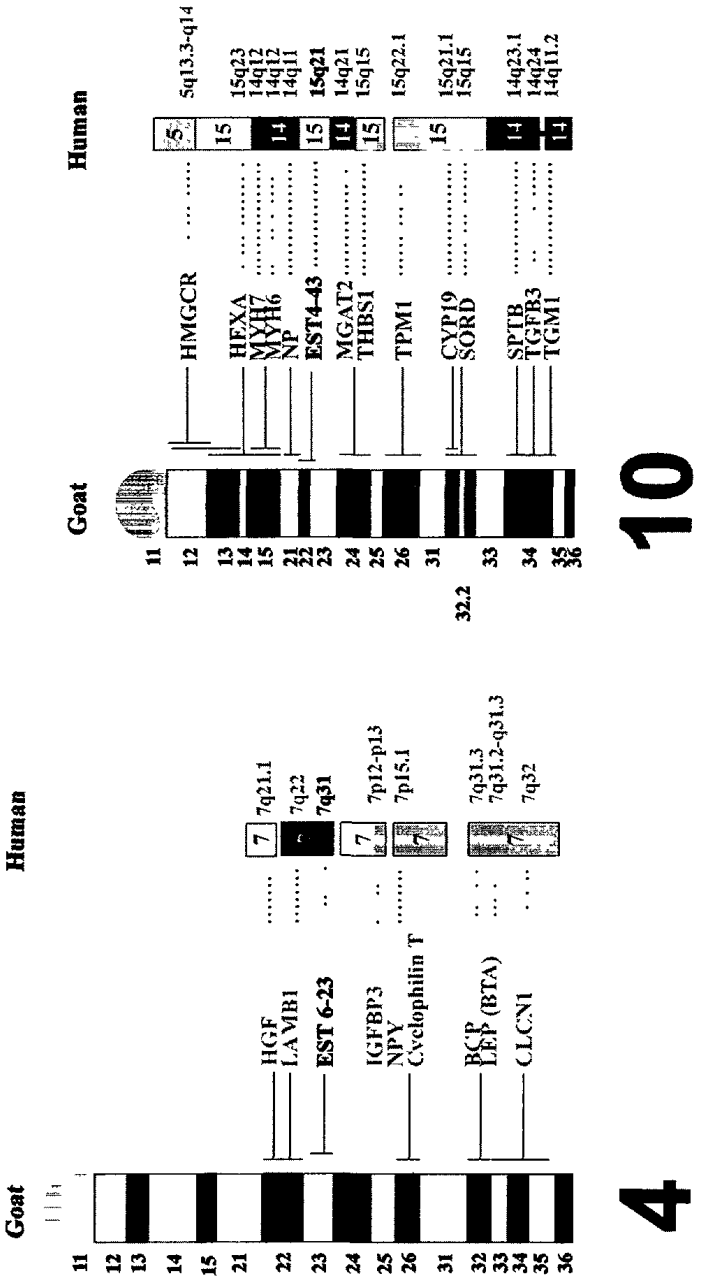

Figure 2. Human-Goat comparative mapping of Chir4 and Chir10. Previously localised genes are drawn in gray on the right of the standard R banding karyotype. Newly localised ESTs are in bold. Corresponding human conserved segments are represented by boxes.

ACKNOWLEDGEMENTS

We would like to thank K. Lynch for revising the English version of the manuscript (the INRA's translation unit, U.C.D., Jouy-en-Josas, France).

REFERENCES

[1] Adams M.D., Kelley J.M , Gocayne J.D., Dubnick M., Polymeropoulos M.H., Xiao H., Merril C.R., Wu A., Olden B., Moreno R.F., Kerlavage A.R., McCombie W.R., Venter J.C., Complementary DNA sequencing: expressed sequence tags and human genome projet, Science 252 (1991) 1651–1656.

[2] 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) 261–265.

[3] Deloukas P., Schuler G., Gyapay G , Beasley E., Soderlund C., Rodriguez-Tome P., Hui L., Matise T., McKusick K., Beckmann J , Bentolila S., Bihoreau M., Birren B., Browne J., Butler A. *et al.*, A physical map of 30,000 human genes, Science 282 (1998) 744–746.

[4] Georges M., Nielsen D., Mackinnon M., Mishra A., Okimoto R , Pasquino A., Sargeant L., Sorensen A., Steele M., Zhao X., Womack J., Hoeschele I., Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing, Genetics 139 (1995) 907–920.

[5] Guillemot E., Gary F., Berland H.M., Durand V., Darré R., Cribiu E.P., Cytogenetic investigation in Saanen and Alpine artificial insemination bucks. Identification of a Robertsonian translocation, Genet. Sel. Evol. 23 (1991) 449–454.

[6] Hayes H., Petit E , Dutrillaux B., Comparison of RBG-banded karyotypes of cattle, sheep and goats, Cytogenet. Cell. Genet. 57 (1991) 51–55.

[7] Lahbib-Mansais Y., Dalias G., Milan D., Yerle M , Robic A., Gyapay G., Gellin J., A successful strategy for comparative mapping with human ESTs: 65 new regional assignments in the pig, Mamm. Genome 10 (1999) 145–153.

[8] Le Provost F., Lépingle A., Martin P., A survey of the goat genome transcribed in the lactating mammary gland, Mamm. Genome 7 (1996) 657–666.

[9] Liu X., Robinson G.W., Wagner K.U., Garrett L., Wynshaw-Boris A., Hennighausen L., Stat5a is mandatory for adult mammary gland development and lactogenesis, Genes Dev. 11 (1997) 179–186.

[10] Makalowski W., Zhang J., Boguski M.S., Comparative analysis of 1196 orthologous mouse and human full-length mRNA and protein sequences, Genome Res. 6 (1996) 846–857.

[11] Ormandy C.J., Camus A., Barra J., Damotte D., Lucas B., Buteau H., Edery M., Brousse N., Babinet C., Binard N., Kelly P., Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse, Genes Dev. 11 (1997) 167–178.

[12] 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.

[13] Ron M., Band M., Yanai A., Weller J.I., Mapping quantitative trait loci with DNA microsatellites in a commercial dairy cattle population, Anim. Genet. 25 (1994) 259–264.

[14] 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) 901–915.

[15] 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) 119–124.

[16] Wilcox A.S., Khan A.S., Hopkins J.A., Sikela J.M., Use of 3′ untranslated sequences of human cDNAs for rapid chromosome assignment and conversion to STSs: implications for an expression map of the genome, Nucleic Acids Res. 19 (1991) 1837–1843.

[17] Womack J.E., Kata S.R., Bovine genome mapping: evolutionary inference and the power of comparative genomics, Curr. Opin. Genet Dev. 5 (1995) 725–733.

[18] Yang Y.P., Womack J.E., Parallel radiation hybrid mapping: a powerful tool for high-resolution genomic comparison, Genome Res. 8 (1998) 731–736.

[19] Zhang Q., Boichard D., Hoeschele I., Ernst C., Eggen A., Murkve B., Pfister-Genskow M., Witte L.A., Grignola F.E., Uimari P., Thaller G., Bishop M.D., Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree, Genetics 149 (1998) 1959–1973.