Mammalian Genome 13, 524–534 (2002).
DOI: 10.1007/s00335-001-2137-4

Mammalian Genome
Incorporating Mouse Genome

Cytogenetic localization of 136 genes in the horse: comparative mapping with the human genome

Dragan Milenkovic,[1] Anne Oustry-Vaiman,[1] Teri L. Lear,[2] Alain Billault,[3] Denis Mariat,[1] François Piumi,[4] Laurent Schibler,[1] Edmond Cribiu,[1] Gérard Guérin[1]

[1]Institut National de la Recherche Agronomique, Centre de Recherches de Jouy, Laboratoire de Génétique biochimique et de Cytogénétique, Département de Génétique animale, 78352 Jouy-en-Josas Cedex, France
[2]M.H. Gluck Equine Research Center, Department of Veterinary Science, University of Kentucky, Lexington, KY 40546-0099, USA
[3]Molecular Engines Laboratories, 20 rue Bouvier, 75011 Paris, France
[4]Laboratoire mixte INRA-CEA de Radiobiologie et d'Etude du Génome, UMR 13.314, 78352 Jouy-en-Josas Cedex, France

Received: 16 October 2001 / Accepted: 2 May 2002

Abstract. The aim of this study was to increase the number of type I markers on the horse cytogenetic map and to improve comparison with maps of other species, thus facilitating positional candidate cloning studies. BAC clones from two different sources were FISH mapped: homologous horse BAC clones selected from our newly extended BAC library using consensus primer sequences and heterologous goat BAC clones. We report the localization of 136 genes on the horse cytogenetic map, almost doubling the number of cytogenetically mapped genes with 48 localizations from horse BAC clones and 88 from goat BAC clones. For the first time, genes were mapped to ECA13p, ECA29, and probably ECA30. A total of 284 genes are now FISH mapped on the horse chromosomes. Comparison with the human map defines 113 conserved segments that include new homologous segments not identified by Zoo-FISH on ECA7 and ECA13p.

During the last few years, a large effort has been made to develop comparative mapping studies between vertebrate genomes, especially between human and domestic animals. These studies have two major goals: 1) to identify major genes or QTL in "map-poor" species by the use of genetic data from high-density map genomes such as humans or mice, and 2) to improve our understanding of the evolutionary process of mammalian genomes. Synteny mapping and heterologous painting, or Zoo-FISH, have been used to reveal conserved chromosomes or chromosome segments between humans and other species, including the horse (Raudsepp et al. 1996; Chowdhary et al. 1998; Caetano et al. 1999), leading to a first global comparative map. However, both techniques are limited in their ability to discern gene order within conserved segments, and boundaries of homologous segments still must be defined more precisely to detect smaller genome rearrangements. Unlike microsatellites, coding sequences are well conserved across mammalian species, making them useful for comparing genomes between distantly related species (Nadeau and Sankoff 1998; O'Brien et al. 1999). Sequence homologies have become a useful tool for identification of genes and genome comparison for the human genome with goats, cattle, or pigs (Schibler et al. 1998b; Band et al. 2000; Pinton et al. 2000). These studies revealed modifications in gene order within conserved segments, as already observed between human and mouse (Carver and Stubbs 1997).

Correspondence to: G. Guérin; email: guerin@diamant.jouy.inra.fr

Genome mapping in the horse has exhibited rapid growth in recent years, including development of linkage maps (Lindgren et al. 1998; Guérin et al. 1999; Swinburne et al. 2000), synteny maps (Shiue et al. 1999), radiation hybrid panels (Kiguwa et al. 2000; Chowdhary et al. 2002), as well as shared resources including BAC libraries, genome scanning panels, and DNA for linkage mapping (http://www.uky.edu/AG/Horsemap.htm/). Likewise, but less rapidly, the cytogenetic map has also progressed with about 40 markers localized in 1998, 80 in 2000, and nearly 150 in 2001 [Godard et al. 2000; Mariat et al. 2001; Lindgren et al. 2001; Lear et al. 2001; Hirota et al. 2001; see also HORSEMAP database http://locus.jouy.inra.fr/ (INRA) and Horsebase http://www.thearkdb.org/ (Roslin Institute)]. Although Zoo-FISH identified 43 conserved segments between the human and horse genome (Raudsepp et al. 1996), a more detailed map was needed to anchor linkage groups and to facilitate positional cloning and candidate gene approaches. Therefore, we conducted the following study to increase the density of physically mapped genes for the horse and to increase the resolution of the comparative gene map of the horse relative to other species, especially humans.

Materials and methods

Extension of the BAC library. Our first horse BAC library, including almost 40,000 clones for a 1.5× genome coverage (Godard et al. 1998), was complemented with a new library and now consists of 108,288 clones. The method to construct the new library was the same as previously described (Godard et al. 1998; Schibler et al. 1998a; Brosch et al. 1998), except that DNA was prepared from peripheral blood leukocytes obtained from a stallion belonging to the Haras Nationaux. The average insert size was estimated at 100 kb, after *Not*I digestion followed by FIGE (Field Inverted Gel Electrophoresis) on 672 randomly chosen clones. Insert sizes ranged from 20 kb to 150 kb. These values define coverage of 3.4× genome equivalents and a theoretical probability of 0.96 to isolate a particular DNA sequence in the library, taking into account 9% of empty clones. The library was distributed in 1128 96-well plates pooled by groups of 24, constituting 47 superpools. Each pool was divided into 12 column pools (196 clones per pool), 8 row pools (288 clones per pool), and 24 plate pools (2304 clones per pool). DNA was extracted and distributed for PCR screening. Effectively, the library was prepared for PCR screening.

BAC library screening. The BAC library was screened with primers that we designed (Table 1) plus TOAST primers (Jiang et al. 1998) provided by C. André (UPR41 CNRS, Rennes, France). The primer pair for ACE was provided by I. Tammen (University of Sydney, Camden, Australia). Primers were first tested for PCR amplification on whole genomic equine DNA before screening the library (47 super-

Table 1. Primer sequences newly designed for genes in this study.

Symbol	Locus Name	Primer F	Primer R
ACE	Angiotensin I-converting enzyme (peptidyl-dipeptidase A) 1	GCCCAGGAGGATGTTTAAGGA	CTTGCCGTTGTAGAAGTCCCA
ATP1A1	ATPase, Na^+/K^+ transporting, alpha 1 polypeptide	TCGAGCTGCTGAGATCCTGG	CAGCTTGGATGCCATAAGCC
CRTL1	Cartilage linking protein 1	CTGTAATTAGGGGATCTGGGAGG	CTCTTTGTGGCGCTGGATGG
CSN1	Casein alpha S 1	CATTCTTGATTTCTCTCCTCA	ATGACACTTACAGAAGAAGCA
DMD	Dystrophin (muscular dystrophy, Duchenne and Becker types)	TCAGAGTGGTTACCACACCG	TGCCAATAAGTTGGCTGCTG
ELA-DRA	Major histocompatibility complex, class II, DR alpha	GAACCCTGGCGACTCAGGAG	AGGTTGGCTTTGTCCACAGC
ELN	Elastin	GCTGCAGCCGCTAAAGCAGCC	GGCCAGCGGCACCTGTGAAG
GUSB	Glucuronidase, beta	GCCAACGAAGCAGGTTGAAG	CTGGGTCCCTGTGGGGGTGG
LAMA3	Laminin, alpha 3	CAGTTTGCAAGGCTTAATTACACC	CAGGTGGGTAACCTCCAACA
LIF	Leukemia inhibitory factor (cholinergic differentiation factor)	CTTCCACGCCAACGGCACGG	GACAGCCCAGCTTCTTCTTCTGG
MGAT1	Mannosyl (alpha-1,3-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase	AGTCTGCAGGGCTTGTGCT	CTGGAACTTGCGGTGGTC
PDGFB	v-sis platelet-derived growth factor beta polypeptide	AGTTTGCACCTCTCCCTGCC	AGCTTTTTTGCAACATTTTCTGG
TYR	Tyrosinase (oculocutaneous albinism IA)	AATGCTCCTGGCTGTTTTGT	TTCTCTGGGGCACTCAAATC

pools and 44 pools for each positive super-pool). All PCR were performed under the same conditions: a 10-µl reaction volume with 40 ng of DNA, 0.05 units of GoldStar Taq polymerase (Eurogentec), 2 mM $MgCl_2$, 0.25 mM of dNTP, 1 pmole of each primer (Eurogentec) in a Perkin-Elmer Cetus 9600 or MJ Research PTC100 thermocycler. PCR conditions were as follows: denaturation 5 min at 94°C, followed by 35 cycles of 30 s at 94°C, 30 s at 55°C, 30 s at 72°C. PCR products were electrophoresed on 3% agarose gels for analysis. For primer pairs that failed to amplify a fragment in any super-pool at 55°C, the annealing temperature was lowered to 50°C. Conversely, when all pools produced products, the annealing temperature was increased to 60°C. The size of the amplified fragment with intra-exonic primers was compared with the expected fragment size to predict clone identity. This criterion was not used for inter-exonic TOASTs because of possible variation in intron size between species.

Positive clones were picked from the library, grown overnight in 3 ml of Luria-Bertani (LB) with 12.5 µl/ml chloramphenicol, and miniprepped by alkaline lysis. PCR was done on miniprepped BAC DNA with the same screening primers. PCR fragments were purified with a QIAquick PCR purification kit (Qiagen) and sequenced on an ABI 377 automated sequencer (Applied Biosystems). Homologies to sequences present in the databases were searched by using the BLAST algorithm at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov) to confirm the identity of the gene in the isolated clone.

Goat BAC clones. From previous work, 150 BAC clones had been isolated from the goat BAC library (Schibler et al. 1998a) and FISH mapped at the goat genome (Schibler et al. 1998b). These clones were also used in this study for mapping genes to the horse.

FISH mapping. BAC clones isolated from the horse BAC library were localized on metaphase chromosomes by fluorescent in situ hybridization (FISH), as described elsewhere (Godard et al. 1998; Lear et al. 1998). For heterologous hybridization with goat BAC clones, this protocol was modified as follows: BAC DNA (400 ng/ml) was labeled by nick translation with the BIONick Labeling System kit (GibcoBRL); after labeling, probes were resuspended in a hybridization mixture with 30% formamide to reduce stringency. Following hybridizations for 48–72 h, the slides were washed with 30% formamide. Probe detection was done with fluorescence-labeled avidin, and the chromosomes were stained with propidium iodide (1 µg/ml) and R-banded (Lemieux et al. 1992). The slides were screened with a Zeiss MC80DX microscope, and images were analyzed by using a CytoVision & Genus System (Applied Imaging). Chromosome identity was determined by comparison of the R-bands with the international horse chromosome banding standard (ISCNH, 1997).

Identification of conserved segments. Conserved segments were deduced by comparing the localization of orthologous genes for horse and human. These segments were defined either by one gene or by several genes presumed to occupy an uninterrupted chromosomal segment, in an identical order, in the two species. Segment boundaries were thus defined as resulting from an inversion in the orientation of the genes, an insertion of gene(s) from other segment(s), or a deletion within a segment, reflecting genome rearrangement since the descent from the common ancestor of humans and horses. In the case of overlapping FISH localizations within a horse chromosomal region, the human order of the genes was considered to minimize the number of breakages according to the principle of parsimony. Similarly, discrepancies between human localizations in different databases [Genome DataBase (http://www.gdb.org), Ensembl (http://www.ensembl.org), Human draft genome (http://genome.ucsc.edu)] were resolved to allow the best fit to the horse genome. All localizations were reported except those that created a new segment incompatible with Zoo-FISH defined by a single gene.

Results

BAC library screening. Out of 196 TOAST and 13 other primer pairs, 135 (65%) amplified horse genomic DNA. This was consistent with the proportion of amplification observed by using these tools for the canine genome (Priat et al. 1999). These primers were subsequently used to screen the whole BAC library. Ninety-eight out of the 135 primer pairs amplified one or a few super-pools. Primer sets were discarded for the 26 amplifying all super-pools and the 11 that did not amplify any super-pool. These 98 primer pairs, corresponding to 76 different genes, gave a clear PCR product that led to the identification of a clone.

BAC gene content for 76 clones was ascertained by sequencing the PCR product obtained from primers for the expected genes. Fifty-six sequences (73.7%) showed very strong homologies with the expected gene sequence (average 93%, with a mean length of 190 bp). GenBank accession numbers of the homologous sequences are shown in Table 2. For the inter-exonic primers, homology was found only with regions corresponding to the exons of the gene and not the intron region. Sequences of PCR products from 20 clones showed no homologies with the expected gene nor with any other gene, and the corresponding BAC clones were excluded from this report. In total, 56 clones containing the expected gene were identified.

Fish mapping. Fifty-six horse and 150 goat BAC clones were hybridized on horse chromosomes by FISH. One hundred and fifty-three clones, 54 horse and 99 goat BAC clones, gave a clear, single, interpretable signal. One horse BAC hybridized to multiple sites, and another, TGFB2, hybridized on two different chromosomes, ECA10 and ECA30. One hundred and twenty-eight localizations were consistent with Zoo-FISH results (Raudsepp et al. 1996). Another four were localized on ECA13p with no Zoo-FISH data available. Three genes,

Table 2. Cytogenetic localizations of the BAC clones. Horse and human localizations, according to GDB, are represented. H: horse BAC clones, G: goat BAC clones. Gene symbols and names were actualized according to currently available data.

BAC origin	Locus Symbol	Locus Name	Horse localization	Human localization	GenBank Accession Number
G	*OAT*	Ornithine aminotransferase (gyrate atrophy)	1p12.1	10q26	
G	*DNTT*	Deoxynucleotidyltransferase, terminal	1p14.1	10q23-q24	
G	*PAX2*	Paired box homeotic gene 2	1p14.3	10q24	
G	*RBP3*	Retinol-binding protein 3, interstitial	1p14-p15	10q11.2	
G	*ACTA2*	Actin, alpha 2, smooth muscle, aorta	1p15-p16	10q22-q24	
G	*COMT*	Catechol-O-methyltransferase	1p18	22q11.21-q11.23	
G	*MEF2A*	MADS box transcription enhancer factor 2, polypeptide A (myocyte enhancer factor 2A)	1q14-p15	15q26	
G	*CYP19*	Cytochrome P450, subfamily XIX (aromatization of androgens)	1q17.3-p21.1	15q21	
G	*TPM1*	Tropomyosin 1	1q17.3-p21.1	15q22	
G	*HEXA*	Hexosaminidase A (alpha polypeptide)	1q21.1-q21.2	15q23-q24	
G	*UBE3A*	Ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	1q21.3	15q11-q13	
G	*GRP58*	Glucose regulated protein 58kDa	1q23	15q15	
G	*SORD*	Sorbitol dehydrogenase	1q23	15q15	
G	*THBS1*	Trombospondine 1	1q23	15q15	
G	*MYH6*	Myosin, heavy polypeptide 6, cardiac muscle beta	1q25-q26	14q11.2-q13	
G	*MYH7*	Myosin, heavy polypeptide 7, cardiac muscle beta	1q25-q26	14q11.2-q12	
G	*TGM1*	Transglutaminase 1 (K polypeptide epidermal type I, protein-glutamine-gamma-glutamyltransferase)	1q27	14q11.2	
G	*MGAT2*	Mannosyl (alpha-1,6) glycoprotein beta-1, 2-N-acetylglucosaminyltransferase	1q29	14q21	
G	*STS-D29580*		2p12-p13	1p35-p36.1	
G	*RHD*	Rhesus blood group, D antigen	2p13-p14	1p34-p36.2	
G	*NPPA*	Natriuretic peptide precursor A	2p14-p12	1p36	
H	*SLC2A1*	Solute carrier family 2 member 1	2p16	1p31.3-p35	U89029
H	*FGG*	Fibrinogen, gamma	2q21	4q28	M10014.1
G	*IL2*	Interleukin 2	2q23	4q26-q27	
H	*CES2*	Carboxylesterase 2	3p14	16q22.1	X63323
H	*PDHA2*	Pyruvate dehydrogenase alpha 2	3q13-q14.1	4q22-q23	14762224
G	*F11*	Coagulation factor XI	3q21-q22.1	4q35	
G	*STS-A12501S*		3q22.1-q22.2	4q13-q22	
H	*CSN1*	Casein alpha S 1	3q22.1-q22.2	4q21.1	3860334
G	*CSN2*	Casein beta	3q22.2-q22.3	4q21	
G	*TXK*	Thioredoxin	3q24	4p12	
G	*HGF*	Hepatocyte growth factor (hepapoietin, scatter factor)	4q12	7q21.1	
G	*LAMB1*	Laminin beta 1	4q13	7q22	
H	*AKR1B1*	Aldehyde reductase	4q23-q24	7q35	J05474
G	*S100A6*	S100 calcium binding protein A6 (calcyclin) (CACY)	5p12	1q21	
G	*STS-L16464*		5p12	1q21	
H	*NFIA*	Nuclear factor I	5q12-q13	1p31.2-p31.3	U18761
H	*ATP1A1*	ATPase, Na^+/K^+ transporting, alpha 1 polypeptide	5q13	1p13	X16776.1
G	*TSHB*	Thyroid stimulating hormone beta	5q14-q15	1p13	
G	*NRAS*	v-ras neuroblastome RAS viral oncogene homolog	5q15	1p13.2	
G	*SGII*		6p12-p13	2q35-q36	
G	*FGF6*	Fibroblast growth factor 6	6q13	12p13	
H	*TRA1*	Tumor rejection antigen (gp96) 1	6q22	12q24.2-qter	Y09136
G	*IFNG*	Interferon, gamma	6q22-q23	12q14	
G	*APOA1*	Apolipoprotein A1	7p14.1	11q23.3	
G	*LDHA*	Lactate dehydrogenase A	7p14.3	11p15.1	
G	*LDLR*	Low density lipoprotein receptor	7p15-p14.3	19p13.2	
H	*TYR*	Tyrosinase (oculocutaneous albinism IA)	7q12	11q21	AF252540
G	*COX8*	Cytochrome c oxidase subunit VIII	7q13-q14	11q13	
G	*WT1*	Wilm's tumor	7q14-q16	11p13	
H	*HBB*	Globin beta	7q14-q15	11p15.4	M15389.1
H	*HBE1*	Globin epsilon	7q16	11p15.5	V00508
H	*POU2F2*	POU domain, class 2, transcription factor 2	7q16	19pter-qter	14756332
G	*OPCML*	Opioid-binding protein/cell adhesion molecule-like (OCAM)	7q16	11pter-qter	
H	*INSL3*	Leydig insulin-like hormone	7q17-p18	19p13-p12	X73636.1
H	*BDNF*	Brain-derived neurotrophic factor	7q19	11p13	U56638.1
H	*HPD*	4-hydroxyphenyl pyruvic acid dioxygenase	8p12-p13	12q14-qter	D31628
H	*LIF*	Leukemia inhibitory factor (cholinergic differentiation factor)	8p15	22q12	AF048827.1
G	*NOS1*	Nitric oxide synthase 1 (neuronal)	8p15-p16	12q24.2-q24.33	
G	*DSG2*	Desmoglein 2	8q14-q15	18q12.1-q12.2	
H	*LAMA3*	Laminin, alpha 3	8q14-q15	18q12.2	45577110
G	*SFTPC*	Surfactant, pulmonary-associated protein C	9p11	8p21	
G	*CRH*	Corticotropin releasing hormone	9p12-p13	8q13	
G	*DEFB*	Defensin beta	9q14	8p23	

(Continued on next page)

Table 2. Continued

BAC origin	Locus Symbol	Locus Name	Horse localization	Human localization	GenBank Accession Number
G	*ANK1*	Ankyrin 1, erythrocytic	9q15-q16	8p11.2	
G	*MYC*	v-*myc* avian myelocytomatosis viral oncogene homolog	9q16	8q24.12-q24.13	
G	*LHB*	Luteinizing hormone beta polypeptide	10p13	19q13.3	
H	*GPI*	Glucosephosphate isomerase	10p15	19q13.1	Z28402
G	*CGA*	Glycoprotein hormone, alpha polypeptide	10q23	6q23	
G	*MYH2*	Myosin, heavy polypeptide 2, skeletal muscle, adult	11p12	17p13.1	
H	*ACE*	Angiotensin I converting enzyme (peptidyl-dipeptidase A) 1	11p13	17q23	XM_044372
G	*ACACA*	Acetyl-coenzyme A carboxylase alpha	11q12	17q12	
G	*GAS*	Gastrin	11q12	17q21	
H	*SSTR2*	Somatostatin receptor	11q14.2-q14.3	17q24	D21338
G	*PNMT*	Phenylethanolamine N-methyltransferase	11q14.3-q15	17q21-q22	
G	*MAPT*	Microtubule-associated protein tau	11q15-q16	17q21	
H	*CHRM1*	Muscarin acetylcholine receptor I	12q14	11q13	M16406
G	*CBLN1*	Cerebellin 1 precursor	13p12	16q12.1	
H	*POR*	P450 (cytochrome) oxidoreductase	13p13	7q11.2	L33893
G	*MT2A*	Metallothionein 2A	13p14-p15	16q13	
H	*ELN*	Elastin	13q13	7q11.23	AF130761
G	*SLC6A3*	Solute carrier family 6 (neurotransmitter transporter dopamine) member 3	14q11-q12	5p15.3	
H	*MGAT1*	Mannosyl (alpha-1,3) glycoprotein beta-1,2-N-acetylglucosaminyltransferase	14q13	5q35	M57301
H	*FGF1*	Acidic fibroblast growth factor	14q21	5q31	M30490
G	*IL3*	Interleukin 3 (colony stimulating factor multiple)	14q21	5q23-q31	
H	*CRTL1*	Cartilage linking protein 1	14q27	5q13-q14	X78077.1
G	*HMGCR*	3-hydroxy-3-methylglutaryl-Coenzyme A reductase	14q27	5q13.3-q14	
G	*EN1*	Engrailed 1	15q13	2q13-q21	
H	*IL1B*	Interleukin-1 beta	15q13	2q13-q21	U92481.1
H	*MDH1*	Malate dehydrogenase	15q21.3	2p16-p23.3	U20353
H	*FSHR*	Follicle stimulating hormone receptor	15q23	2p16-p21	U73659.1
G	*POMC*	Proopiomelanocortin	15q25	2p23	
H	*APOB*	Apolipoprotein B	15q25-q26	2p23-p24	J02610
G	*UMPS*	Uridine monophosphate synthetase	16q12	3q13	
G	*TFDP2*	Transcription factor Dp-2 (E2F dimerization partner 2)	16q16	3q23	
H	*APEH*	Acylamino acidreleasing enzyme	16q21	3p21	NM_001640
G	*GPX1*	Glutathion peroxidase 1	16q21.1	3q11-q12	
G	*HRH1*	Histamine receptor H1	16q21.1	3p25	
G	*ACPP*	Acid phosphatase, prostate	16q21.3	3q21-q23	
G	*LTF*	Lactotransferrin	16q21.3	3q21-q23	
G	*RBP1*	Retinol-binding protein 1, cellular (CRBP)	16q22	3q21-q22	
G	*CPA3*	Carboxypeptidase A3 (mast cell)	16q22-q23.1	3q21.3-q25	
G	*NCK1*	Non-catalytic region of tyrosine kinase	16q22-q23.1	3q21	
G	*KNG*	Kininogen	16q23.3	3q27	
G	*AGTR1*	Angiotensin receptor 1	16q25	3q21-q25	
H	*ALOX5AP*	5-lipoxygenase-activating protein	17q14-q15	13q12	XM_015396
G	*BRCA2*	Breast cancer 2 early onset	17q22	13q12-q13	
G	*SGCG*	Sarcoglycan gamma 35kD dystrophin associated glycoprotein (DAGA4)	17q22-q23	13q12	
H	*TTN*	Titin	18q24	2q31	M97767
G	*MITF*	Microphthalmia-associated transcription factor	19q15	3p14.1-p12	
G	*SOX2*	Sex determining region-Y box 2	19q17-q18	3q26.3-q27	
G	*NDUFS8*	NADH dehydrogenase (ubiquinone) Fe-S protein 8 (23 kDa) (NADH-coenzyme Q reductase)	19q21	3q28	
H	*PROS1*	Vitamin K dependent protein S	19q25	3p11-q11	L31379
H	*ELA-DRA*	Major histocompatibility complex, class II, DR alpha	20q16-q21.1	6p21.3	L47174.1
H	*HSPA1A*	Heat shock 70 kDa protein 1	20q16-q21	6p21.3	NM_005346.2
H	*CSNK2B*	Casein kinase II beta subunit	20q21	6p21-p12	X56503
H	*GBC*	Compement factor B	20q21	6p21.3	M59240
G	*OLADRB*	Major histocompatibility complex, class II, DR beta (ovine)	20q21.1	6p21.3	
G	*STS-D29310*		20q21.1	6pter-qter	
G	*GSTA1*	Glutathione S transferase A1	20q21.2	6p12	
G	*BF*	B-factor properdin	20q21.3	6p21.3	
G	*STS-D38552*		21q13-q14	5q12	
H	*GHR*	Growth hormone receptor	21q17	5p12-p14	X54429
G	*NPR3*	Natriuretic peptide receptor C (ANPRC)	21q17	5p13-p14	
H	*PDYN*	Prodynorphin	22q15	20p13	K02268.1
H	*PRNP*	Prion protein	22q15	20p13	AF117312
G	*STS-D29077*		22q15	20p12	
G	*VLDLR*	Very low density lipoprotein receptor	23q15.3	9p24	
G	*COX4P1*	Cytochrome C oxidase subunit IV pseudogene 1	24q12-q13	14q21-qter	
H	*PI*	Antitrypsine alpha 1	24q15-q16	14q32.1	AF034077.2
H	*GGTA1*	Glycoprotein, alpha-galactosyltransferase 1	25q15-q16	9q33-q34	J04989
G	*KRTAP8*	Keratin associated protein	26q14-q15	ND	
G	*CRYAA*	Crystallin alpha A	26q17	21q22.3	
H	*MX1*	Myxovirus resistance 1	26q17	21q22.3	M65087

(Continued on next page)

Table 2. Continued

BAC origin	Locus Symbol	Locus Name	Horse localization	Human localization	GenBank Accession Number
G	*IGF1*	Insulin-like growth factor 1 (somatomedin C)	28q15	12q22-q23	
G	*CTSLL*	Cathepsin L-like	29q14	10q11.2-q21.1	
H	*TGFB2*	Transforming growth factor, beta 2	30q14	1q41	NM_003238.1
H	*ESR1*	Estradiol receptor 1	31p16	6q25.1	AF124093
G	*ALAS2*	Aminolevulinate delta synthase 2 sideroblastic hypochromic anemia	Xp14	Xp11.21	
H	*DMD*	Dystrophin (muscular dystrophy, Duchenne and Becker types)	Xp16	Xp21.3-p21.2	L0211.3
G	*STS-L08239*		Xq12	Xq11.21-q11.23	
G	*PGK1*	Phosphoglycerate kinase 1	Xq13-q16	Xq13.3	
G	*AR*	Androgen receptor	Xq15-q16	Xq11.2-q12	
G	*GLRA2*	Glycine receptor apha 2	Xq17	Xp22.1-p21.3	

LDLH, *POU2F2*, and *INSL3*, identified new segments containing more than one gene, undetected by Zoo-FISH. The remaining localizations were inconsistent with Zoo-FISH results and are not reported here, as they need further investigations. These three localizations were reported since they form new segments, containing more than one gene, undetected by Zoo-FISH. Overall, 136 genes were added to the horse cytogenetic map, for a total of 284 genes localized by FISH (Table 2 and Fig. 1). On average, nine genes were localized on each chromosome, although no genes were mapped to ECA27.

Discussion

Identifying economically important horse genes will be facilitated by an effective comparative gene map with humans. Several approaches contribute to the comparative map. A key study used chromosome painting to identify 43 conserved chromosomal segments between the horse and human genome (Raudsepp et al. 1996). Here, FISH mapping with horse and goat clones was used to refine chromosomal segments defined by synteny mapping and by chromosome painting. Goat BAC clones were found to be very useful for generating comparative mapping data. One can thus expect that clones from different species can be mapped on the horse genome, and even if only a proportion hybridizes (two-thirds in our case), they provide a good panel of interspecific anchor markers. But, compared with intraspecific clones, they cannot be used to produce genetic polymorphic markers absolutely necessary to detect genes of breeding interest.

This study confirmed the results from chromosome painting and increased the resolution of physical mapping by identifying chromosome arm homologies: one gene localized on ECA1p18 (HSA22q11) confirmed homology between ECA1pter and HSA22; genes localized on ECA19 confirmed homology between ECA19 and HSA3, ECA17 and HSA13, ECA21 and HSA5, ECA22 and HSA20, as well as a part of the complex rearrangements for ECA1. Support for homologies between ECA1 and regions of HSA2 and HSA12 was not found. No markers were mapped on ECA27, and one, tentatively mapped on ECA30, needs confirmation. HSA3 shows homology to the horse Chrs 16 and 19. Half of the distal part of HSA3p and most of HSA3q are homologous to ECA16, while the rest of the chromosome would be homologous to ECA19. Most of HSA5q is homologous to ECA14, while ECA21 shows homology to HSA5p and proximal HSA5q genes. ECA25 shows homology only to HSA9q genes, while ECA23 displays mostly HSA9p homologous genes except for one HSA9q segment.

A number of these localizations are significant because they represent either a first gene located on chromosomes or localizations that increase map resolution. Of particular interest are the four genes localized on ECA13p, two from HSA16q and two from HSA7q, allowing us to deduce homology between ECA13p and HSA7 and HSA16, homology that was undetected by Zoo-FISH. In one case, the human gene localization was used to remove an ambiguity due to gene family interferences. The gene fragment detected with PDHA1 primers had a sequence homology of 91% and was localized on ECA3. Human localizations (PDHA1 on HSAX; PDHA2 on HSA4) suggest that we have, in fact, selected and mapped the clone containing the PDHA2 gene. The molecular organization of the casein cluster, revealing that CSN1 and CSN2 were 15 kb apart (Milenkovic et al. 2002), helped us to refine the cytogenetic localization of these two genes to the unique overlapping band (3q22.2). This study also provided confirmation and increased resolution for previously mapped genes as for 11 genes assigned by synteny (Caetano et al. 1999).

A number of cytogenetic localizations obtained in this study were not in agreement with the Zoo-FISH results, a situation for which different explanations can be put forward. A clone may detect a fragment under the detection limit of the Zoo-FISH estimated to be 5–7 Mb (Hayes, 1995); it may contain one gene of a family, or a pseudogene, or an unrelated conserved fragment. Finally, the BAC clone may be chimeric. For these reasons, we adopted a very conservative position in not showing conflicting results involving a new segment defined by a single gene unless there was compelling or corroborating information. For example, ECA7 was known to be homologous to HSA11 except for two regions, one centromeric region painted by HSA19, and one telomeric to the p arm detected by C3 (Millon et al. 1993). Our results confirm the homology to HSA19 with the localization of LDLR near C3 and reveal a new segment containing POU2F2 and INSL3 on ECA7q17-q18, regions undetected by Zoo-FISH.

For another gene, TGFB2, the TGFB2-specific BAC clone hybridized to two different horse chromosomes, ECA10 and ECA30. These results could be the result of chimerism of the BAC or the existence of a pseudogene. In humans, TGFB2 maps to HSA1q41. Zoo-FISH studies indicated that ECA30 had homology to HSA1, while ECA10 had homology to HSA6 (Raudsepp et al. 1996). No other gene of the TGFB family was found compatible with an ECA10 (HSA6) mapping. Therefore, these data constitute preliminary evidence for homology of the region with HSA1q41 genes to ECA30.

A recent study mapping markers on Chr X by using radiation hybrid panel showed that gene order is conserved in the two species (Raudsepp et al. 2002). In the present study, we observe that the order of genes is the same except for LAMP2 and GLRA2, genes that were not localized on the RH panel. Comparison of the order of genes for Chr X between horse and goat showed that the order of genes is identical, except for F9, in the two species, including LAMP2 and GLRA2. Such intra-

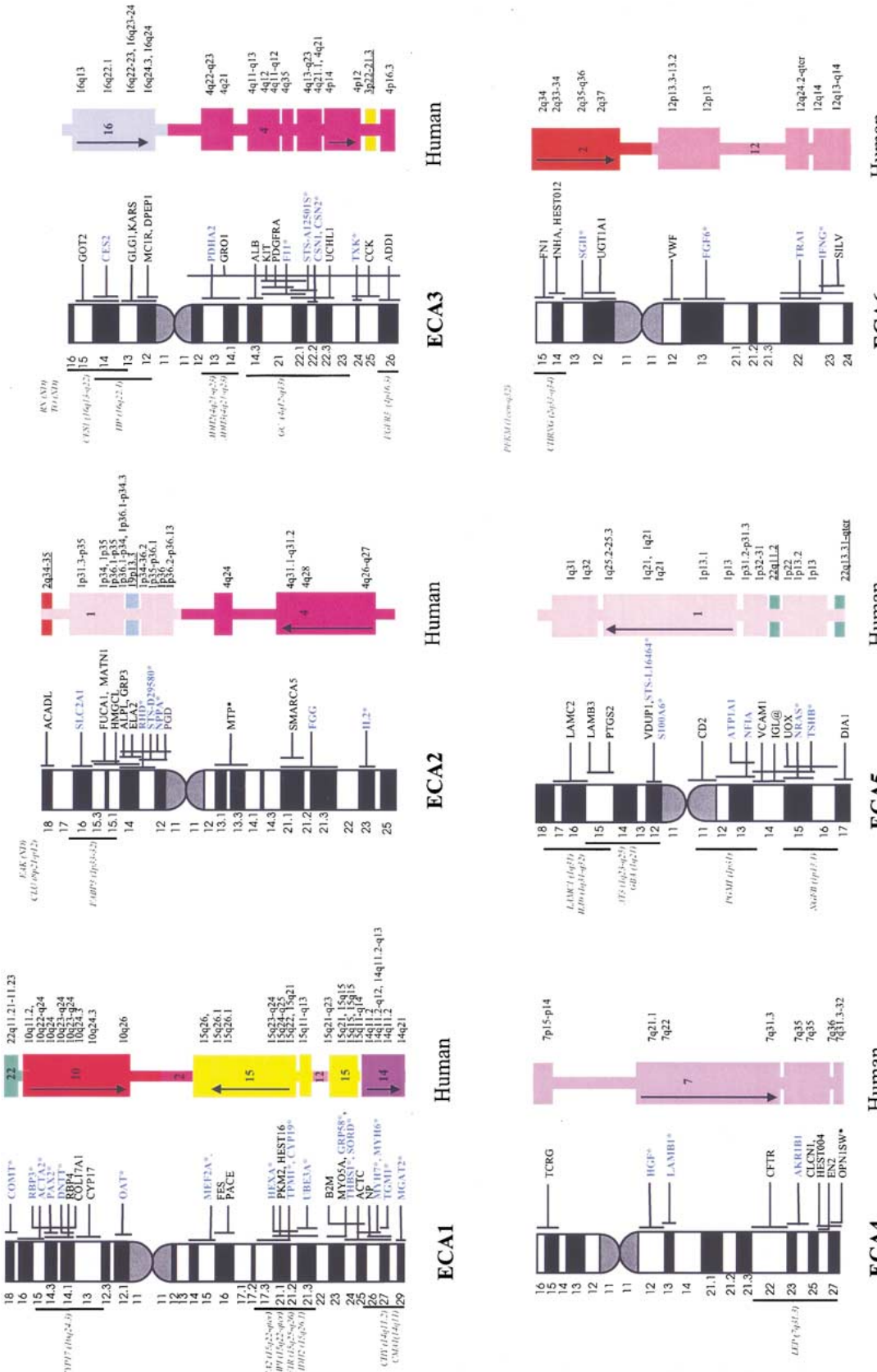

Fig. 1. Continued

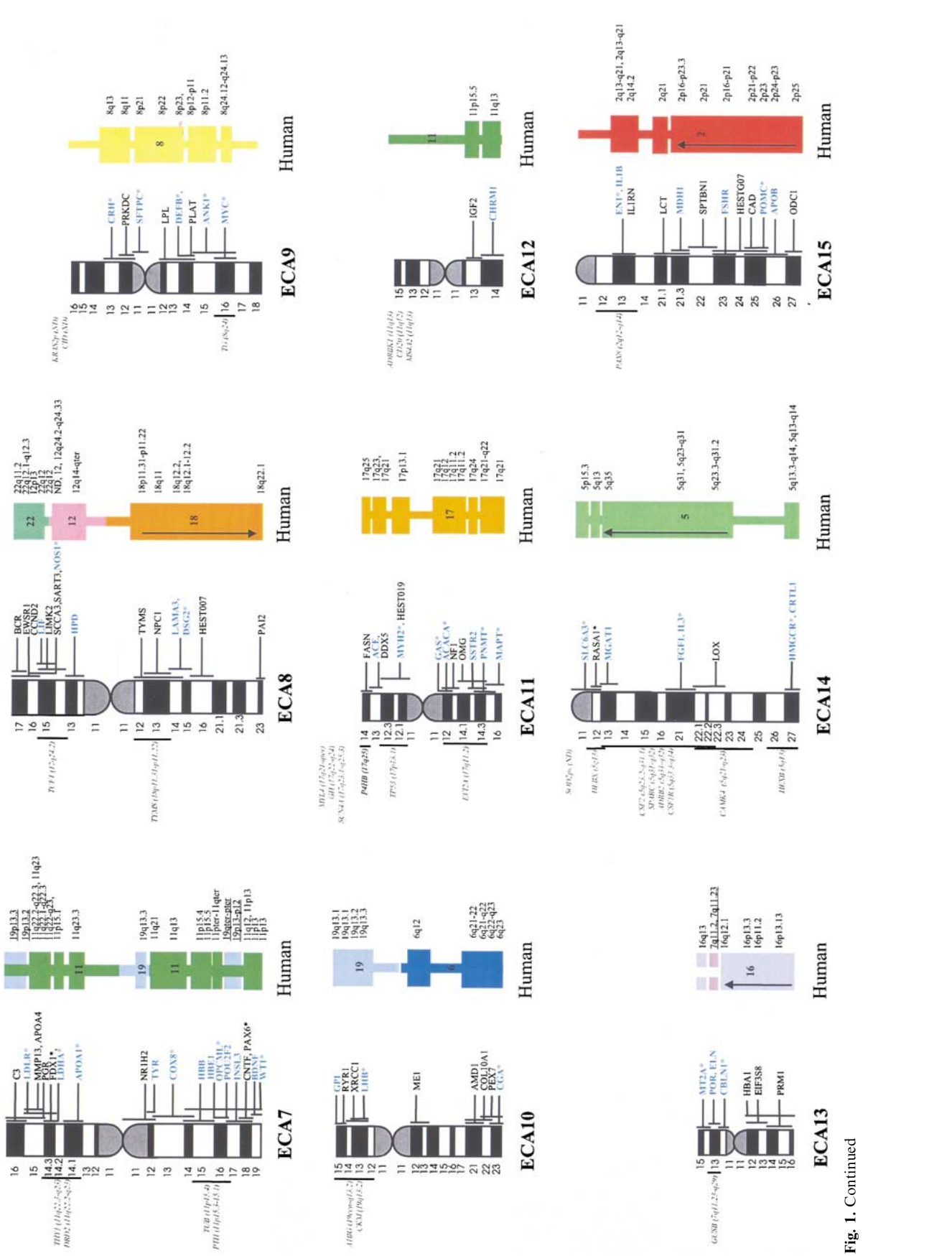

Fig. 1. Continued

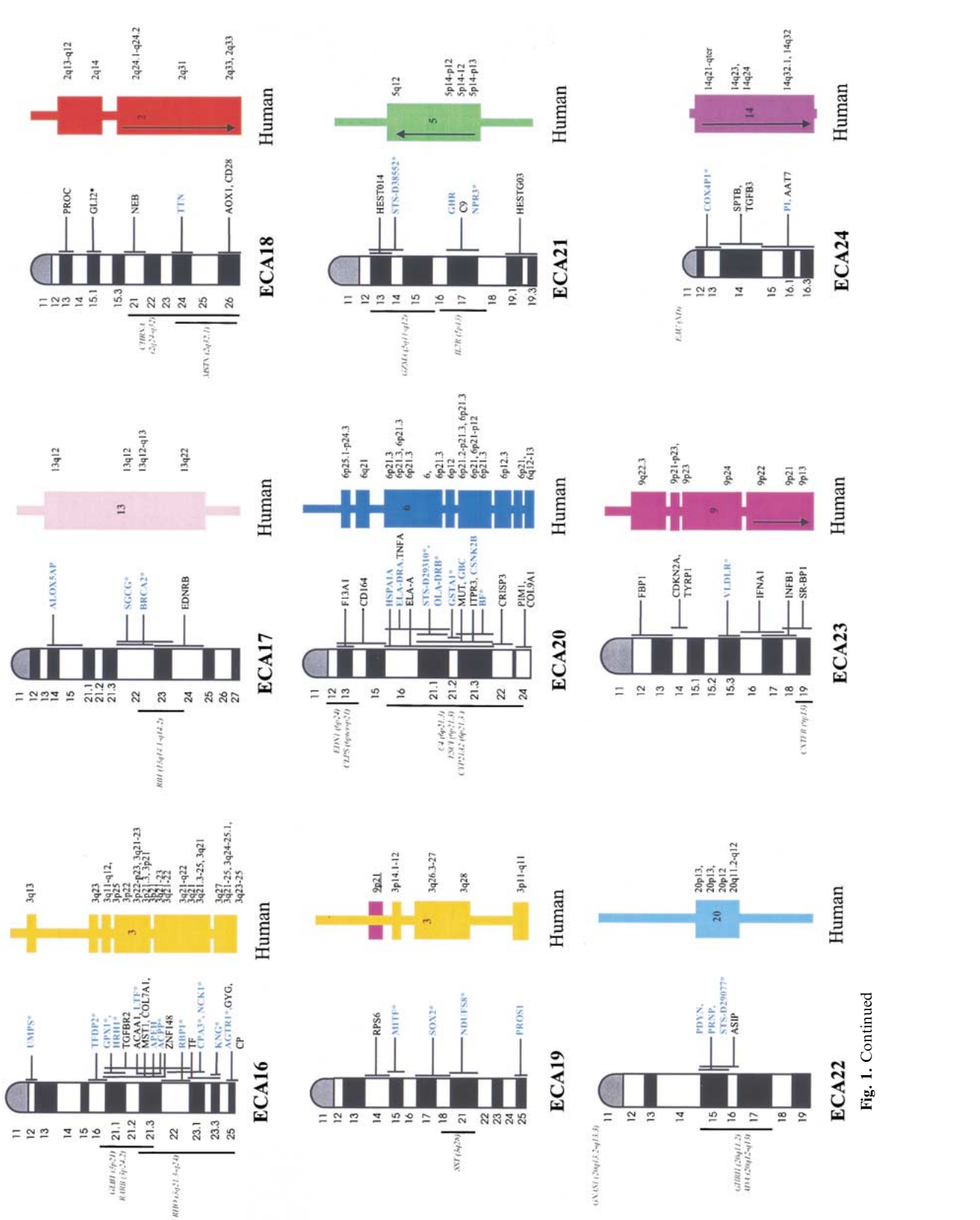

Fig. 1. Continued

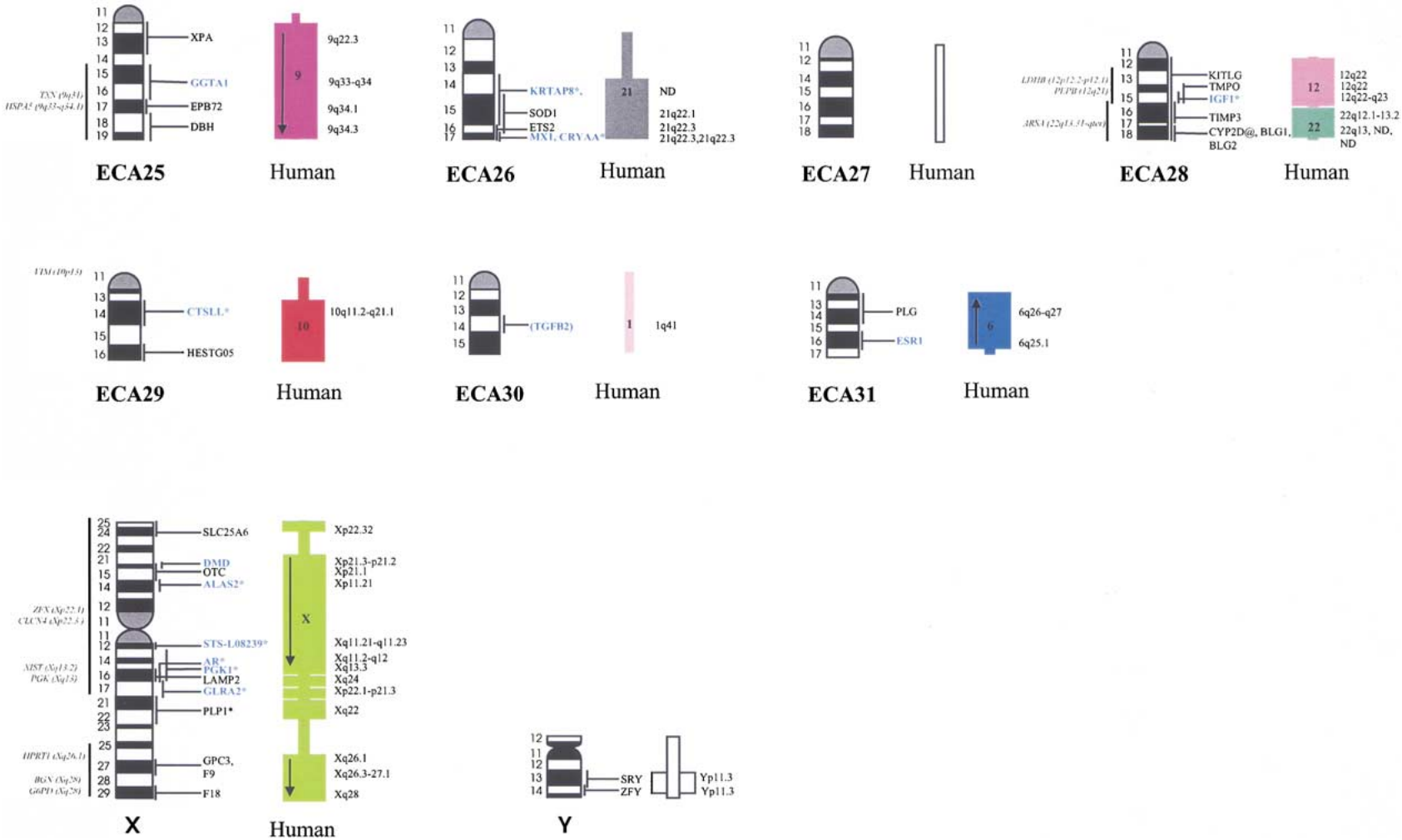

Fig. 1. Horse and human comparative map. The horse R-banding idiogram (ECA) is represented with the cytogenetic localizations (on the right). New localizations are drawn in blue; heterologous clones are marked with an asterisk. Somatic cell hybrid synteny data are represented on the left with their most likely localizations according to the compiled data. Human data are displayed further to the right. Vertical colored lines represent heterologous painting data and include the corresponding human chromosome numbers. Colored boxes show the conserved segments, with arrows indicating deduced segment orientation. Cytogenetic localizations from the Human Genome Database are represented on the right; horse/human inconsistencies are underlined.

chromosomal rearrangements were already observed in other species (Piumi et al. 1998; Schibler et al. 1998b; Band et al. 2000).

The conserved segments were identified as described in previous sections by using the most likely order of the genes. Although the resolution of FISH mapping for single genes is greater than that of Zoo-FISH mapping, the order of closely linked genes may be incorrectly identified, and small inversions may be missed. Linkage mapping or Radiation hybrid mapping has a greater power to determine gene order for closely linked genes. Indeed, cytogenetic locations are not always good indicators of genetic distance, as indicated by comparisons of the RH mapping data to cytogenetic locations for FASN, ACE, and MYL4 in the horse (Chowdhary et al. 2002). Likewise, in terms of comparative mapping, it is observed that two genes (ALOX5AP and SGCG) present on the same band in humans are far apart on Chr 17 of the horse. Finally, inconsistencies in map localizations of a number of genes exist between regularly updated human databases that may lead to changes in gene order.

Nevertheless, we do not believe that the total number of conserved fragments identified for the horse/human comparison will change dramatically. Certainly, more rearrangements may be detected, but by comparison with the situation in other, more completely mapped species, we may be approaching the limit. While the number of conserved segments predicted by Zoo-FISH was 48 between cattle and humans and 47 between pigs and humans, that number was increased by FISH to 105 and 84 respectively. In this study, the number of conserved segments between horse and humans increased from 43 to 113. This number of conserved fragments is in the same magnitude range as for cattle (105) and goats (107) (Band et al. 2000; Schibler et al. 1998b), while it is lower (84) for pigs (Pinton et al. 2000).

The number of gene markers FISH mapped on the horse genome is now 284. Homologous or heterologous FISH mapping of BAC clones containing genes was shown to be a relevant but limited approach for comparative mapping. Comparative mapping is an ongoing process, and marker position on the maps will certainly evolve, as shown by cytogenetic and sequence location inconsistencies still existing in human databases. Nevertheless, we believe that this work represents a significant contribution both to the horse and to comparative mapping. There are still large regions lacking sufficient data for effective comparison. In addition to cytogenetic mapping, this problem could be addressed by selecting genes, perhaps as equine expressed sequence tags (ESTs), systematically, evenly distributed over the human genome, and mapping them to the horse by different positioning techniques such as radiation hybrid mapping or linkage mapping. This overall integrated map is another step towards the identification of economically important genes in regions detected by positional cloning. Furthermore, the horse represents a unique family among animals with gene maps. Data from many families allow for increasingly comprehensive comparisons based on genome conservation and rearrangement. ''Phylogenomics'' is becoming a cornerstone to a better understanding of chromosomal evolution leading to diversity (Nadeau and Sankoff 1998; O'Brien et al. 1999).

Acknowledgments. We are grateful to S. Taourit for sequencing the PCR products and to C. André (UPR41 CNRS, Rennes, France) for kindly giving us the TOAST primers. This study was partially supported by the Service des Haras Nationaux and D. Milenkovic was supported by a fellowship from INRA. T.L. Lear's research is supported by a grant from the Morris Animal Foundation and the University of Kentucky Agricultural Experiment Station, published as paper 01-14-141. E. Bailey (M.H. Gluck Equine Research Center, University of Kentucky, Lexington, USA) is very gratefully acknowledged for critically reviewing the manuscript.

References

Band MR, Larson JH, Rebeiz M, Green CA, Heyen W et al. (2000) An order comparative map of the cattle and human genomes. Genome Res 10, 1359–1368

Brosch R, Gordon SV, Billault A, Garnier T, Eiglmeier K et al. (1998) Use of a Mycobacterium tuberculosis H37Rv bacterial artificial chromosome library for genome mapping, sequencing, and comparative genomics. Infect Immun 66, 2221–2229

Caetano AR, Shiue YL, Lyons LA, O'Brien SJ, Laughlin TF et al. (1999) A comparative gene map of the horse [Equus caballus]. Genome Res 9, 1239–1249

Carver EA, Stubbs L (1997) Zooming in to the human-mouse comparative map: genome conservation re-examined on a high-resolution scale. Genome Res 7, 1123–1137

Chowdhary BP, Raudsepp T, Frönicke L, Scherthan H (1998) Emerging patterns of comparative genome organization in some mammalian species as revealed by Zoo-FISH. Genome Res 8, 577–589

Chowdhary BP, Raudsepp T, Honeycutt D, Owens EK, Piumi F, et al. (2002) Construction of a 5000rad whole genome radiation hybrid panel in the horse and generation of a comprehensive map for ECA11. Mamm Genome 13, 89–94

Godard S, Schibler L, Oustry A, Cribiu EP, Guérin G (1998) Construction of a horse BAC library and cytogenetical assignment of 20 type I and type II markers. Mamm Genome 9, 633–637

Godard S, Vaiman A, Schibler L, Mariat D, Vaiman D et al. (2000) Cytogenetic localization of 44 new coding sequences in the horse. Mamm Genome 11, 1093–1097

Guérin G, Bailey E, Bernoco D, Anderson I, Antczak DF, et al. (1999) Report of the International Equine Gene Mapping Workshop: male linkage map. Anim Genet 5, 341–354

Hayes H (1995) Chromosome painting with human chromosome-specific DNA libraries reveals the extent and distribution of conserved segments in bovine chromosomes. Cytogenet Cell Genet 71, 168–174

Hirota K, Piumi F, Sato F, Ishida N, Guérin G et al. (2001) FISH assignment of two equine BAC clones containing SRY and ZFY. Anim Genet 32, 326–327

ISCNH (1997) Bowling AT, Breen M, Chowdhary BP, Hirota K, Lear T et al. (Committee) International System for Cytogenetic Nomenclature of the domestic Horse. Chromosome Res 5, 433–443

Jiang Z, Priat C, Galibert F (1998) Traced orthologous amplified sequence tags (TOASTs) and mammalian comparative maps. Mamm Genome 9, 577–587

Kiguwa SL, Hextall P, Smith AL, Critcher R, Swinburne J et al. (2000) A horse whole-genome-radiation hybrid panel: Chromosome 1 and 10 preliminary maps. Mamm Genome 11, 803–805

Lear TL, Breen M, Ponce de Leon FA, Coogle L, Ferguson EM et al. (1998) Cloning and chromosomal localization of MX1 and ETS2 to chromosome 26 of the horse (*Equus caballus*) Chromosome Res 6, 333–335

Lear TL, Brandon R, Piumi F, Terry RR, Guérin G et al. (2001) Mapping of 31 horse genes in BACs by FISH. Chromosome Res 9, 261–262

Lemieux N, Dutrillaux B, Viegas-Pequignot E (1992) A simple method for simultaneous R- or G-banding and fluorescence in situ hybridization of small single-copy genes. Cytogenet Cell Genet 59, 311–312

Lindgren G, Sandberg K, Persson H, Marklund S, Breen M et al. (1998) A primary male autosomal linkage map of the horse genome. Genome Res 8, 951–966

Lindgren G, Breen M, Godard S, Bowling AT, Murray J et al. (2001) Mapping of 13 horse gene by fluorescent in situ hybridization (FISH) and somatic cell hybrid analysis. Chromosome Res 9, 53–59

Mariat D, Oustry-Vaiman A, Cribiu EP, Raudsepp T, Chowdhary BP, et al (2001) Isolation, characterization and FISH assignments of horse BAC clones containing type I and type II markers. Cytogenet Cell Genet 92, 144–148

Milenkovic D, Martin P, Guérin G, Leroux C (2002) A specific pattern of splicing for the horse αS1-Casein mRNA and partial genomic characterization of the relevant locus. Genet Sel Evol (Paris), in press

Millon LV, Bowling AT, Bickel LA (1993) Fluorescence in situ hybridization of C3 and 18S rDNA to horse chromosomes. Proc. of the 8th North Am. Coll. on Dom. Anim. Cytogenet. and Gene Mapping, Guelph, 163

Nadeau JH, Sankoff D (1998) Counting on comparative maps. Trends Genet 14, 495–501

O'Brien SJ, Menotti-Raymond M, Murphy WJ, Nash WG, Wienberg J et al. (1999) The promise of comparative genomics in mammals. Science 286, 458–481

Pinton P, Schibler L, Cribiu EP, Gellin J, Yerle M (2000) Localization of 113 anchor loci in pigs: improvement of the comparative map for human, pigs and goats. Mamm Genome 11, 306–315

Piumi F, Schibler L, Vaiman D, Oustry A, Cribiu EP (1998) Comparative cytogenetic mapping reveals chromosome rearrangements between the X chromosomes of two closely related mammalian species (cattle and goats). Cytogenet Cell Genet 81, 36–41

Priat C, Zhihua H, Jiang Z, Renier C, André C et al. (1999) Characterization of 463 type I markers suitable for dog genome mapping. Mamm Genome 8, 803–813

Raudsepp T, Frönicke L, Scherthan H, Gustavsson I, Chowdhary BP (1996) Zoo-FISH delineates conserved chromosomal segments in horse and man. Chromosome Res 4, 218–225

Raudsepp T, Kata SR, Piumi F, Swinburne J, Womack JE et al. (2002) Conservation of gene order between horse and human X chromosomes as evidenced through radiation hybrid mapping. Genomics 79, 451–457

Schibler L, Vaiman D, Oustry A, Guinec N, Dangy-Caye AL, Billault A, Cribiu EP (1998a) Construction and extensive characterization of a goat Bacterial Artificial Chromosome library with threefold genome coverage. Mamm Genome 9, 119–124

Schibler L, Vaiman D, Oustry A, Giraud-Delville C, Cribiu EP (1998b) Comparative gene mapping: a fine scale survey of chromosome rearrangements between ruminants and human. Genome Res 8, 901–915

Shiue YL, Bickel L, Caetano A., Millon L, Clark R et al. (1999) A synteny map of the horse genome comprised of 240 microsatellite and RAPD markers. Anim Genet 30, 1–9

Swinburne J, Gerstenberg C, Breen M, Aldridge V, Lockhart L, et al. (2000) First comprehensive low-density horse linkage map based on two 3-generation, full-sibling, cross-bred horse reference families. Genomics 2, 123–134