Genomics **66,** 123–134 (2000)
doi:10.1006/geno.2000.6207, available online at http://www.idealibrary.com on IDEAL®

First Comprehensive Low-Density Horse Linkage Map Based on Two 3-Generation, Full-Sibling, Cross-Bred Horse Reference Families

June Swinburne,[1,*] Cornelia Gerstenberg,[2] Matthew Breen,[1] Victoria Aldridge,[1] Lesley Lockhart,[1] Eliane Marti,[3] Doug Antczak,[4] Marcia Eggleston-Stott,[5] Ernie Bailey,[6] Jim Mickelson,[7] Knut Røed,[8] Gabriella Lindgren,[9] Wim von Haeringen,[10] Gérard Guérin,[11] Jon Bjarnason,[12] Twink Allen,[2,13] and Matthew Binns[1]

[1]*Animal Health Trust, Newmarket, Suffolk CB8 7UU, United Kingdom;* [2]*T.B.A. Equine Fertility Unit, Newmarket, Suffolk, CB8 9BH, United Kingdom;* [3]*Division of Immunogenetics, Institute for Animal Breeding, University of Berne, CH-3012 Berne, Switzerland;* [4]*Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, New York 14853;* [5]*Veterinary Genetics Laboratory, University of California, Davis, Davis, California 95616;* [6]*Department of Veterinary Science, University of Kentucky, Lexington, Kentucky 40546;* [7]*College of Veterinary Medicine, University of Minnesota, 1988 Fitch Avenue, St. Paul, Minnesota 55108;* [8]*Department of Morphology, Genetics, and Aquatic Biology, Norwegian College of Veterinary Medicine, P.O. Box 8146, N-0033 Oslo, Norway;* [9]*Evolutionary Biology Centre, Department of Evolutionary Biology, Uppsala University, Norbyvagen 18 D, S-752 36 Uppsala, Sweden;* [10]*Dr van Haeringen Laboratorium b.v., P.O. Box 408, Wageningen 6700 AK, The Netherlands;* [11]*INRA Centre de Recherches de Jouy, 78350 Jouy-en-Josas, France;* [12]*Holar i Hjaltadal, 551 Saudarkrokur, Iceland; and* [13]*Department of Clinical Veterinary Medicine, University of Cambridge, Cambridge CB3 0ES, United Kingdom*

Received December 27, 1999; accepted March 10, 2000

Two 3-generation full-sibling reference families have been produced and form a unique resource for genetic linkage mapping studies in the horse. The F_2 generations, now comprising 61 individuals, consist of 28- to 32-day-old embryos removed nonsurgically from two pairs of identical twin mares. The same stallion sired all F_2s such that the two full-sibling families are half-sibling with respect to each other. The families are crossbred to maximize levels of heterozygosity and include Arabian, Thoroughbred, Welsh Cob, and Icelandic Horse breeds. Milligram quantities of DNA have been isolated from each embryo and from blood samples of the parents and grandparents. The families have been genotyped with 353 equine microsatellites and 6 biallelic markers, and 42 linkage groups were formed. In addition, the physical location of 85 of the markers is known, and this has allowed 37 linkage groups to be anchored to the physical map. The inclusion of dams in the genotyping analysis has allowed the generation of a genetic map of the X chromosome. Markers have been assigned to all 31 autosomes and the X chromosome. The average interval between markers on the map is 10.5 cM, and the linkage groups collectively span 1780 cM. The results demonstrate the benefits for horse linkage mapping studies of genotyping on these unique full-sibling families, which comprise relatively few individuals, by the generation of a comprehensive low-density map of the horse genome.

* To whom correspondence should be addressed at Genetics Section, Animal Health Trust, Lanwades Park, Kentford, Newmarket, Suffolk, CB8 7UU, UK. Telephone: (0)1638 750659. Fax: (0)1638 750794. E-mail: jswinbur@hgmp.mrc.ac.uk.

INTRODUCTION

From prehistoric to modern times, the horse has played a vital role as a beast of burden in many fields as diverse as agriculture and warfare. In many developing countries, the horse is still the major form of transport for people and their possessions, whereas in the western world it has become an animal of sport and pleasure. In particular, flat racing has raised the value of individual Thoroughbreds to the level where an unproven yearling has commanded a price tag of $13.1 million (Seattle Dancer, sold in 1985 at auction in Kentucky). Clearly the pedigree of such a horse is a major factor in establishing such high esteem, and an increased understanding of the inherited factors that underpin valued traits will therefore be of interest to the horse-breeding industry. The generation of a genetic linkage map of the horse is the first step toward dissecting inherited factors that affect conformation, soundness, and ability, in a scientific and methodical manner. In addition, a genetic map will aid in the study of, and development of tests for, inherited defects of the horse including rhabdomyolysis (“tying up”), laryngeal hemiplegia (“roaring”), osteochondrosis dessicans, insect bite dermal hypersensitivity (“sweet itch”), and chronic obstructive pulmonary disease. Many of these conditions are likely to be polygenic, eventually requiring the development of high-density maps and other molecular tools to evaluate fully the component genetic loci involved.

Until recently, progress in this field was minimal, as marker isolation and characterization were slow. The first equine microsatellite markers were published

0888-7543/00 $35.00

only 7 years ago (Ellegren *et al.*, 1992), followed by reports detailing initial linkage observations (Marklund *et al.*, 1994) and physical position assignments for a limited number of markers (Breen *et al.*, 1997; Goddard *et al.*, 1997). More recently, the development of a mouse–horse somatic cell hybrid panel and its use to synteny map a substantial number of genetic markers provided the first assignment of large numbers of markers to horse chromosomes (Shiue *et al.*, 1999). Preliminary linkage maps that describe initial attempts to construct genetic linkage maps of the horse genome were also published (Lindgren *et al.*, 1998; Guérin *et al.*, 1999). These studies describe the analysis of 140 and 161 genetic markers, respectively, on large half-sibling families. These maps are incomplete, however, as they do not position linkage groups on the X chromosome or on every autosome. In contrast, the present study describes a much larger panel of genetic markers mapped onto a reference family that has proven to be highly efficient.

Much thought was given to the development of an effective horse reference family for use in this study. In human genetic mapping, for example, 40 core families with an average of 8.3 offspring, most with all four grandparents and collectively known as the CEPH families, have been extensively used (Dausset *et al.*, 1990). For mouse mapping, several collections of intercross and backcross families have been used to assemble different maps (Copeland *et al.*, 1993; Dietrich *et al.*, 1994).

The number of offspring per pregnancy, the duration of gestation, and the generation interval largely determine the ease with which suitable reference families can be assembled for different species. In this regard, cattle and horses appear to present similar problems, as they both usually produce single offspring following a relatively long gestation. In practice, however, the two species do differ significantly, in that it has been possible to superovulate cows and generate large full-sibling families for linkage studies by transferring multiple embryos into synchronized recipient cows (Barendse *et al.*, 1994). In contrast, practical superovulation has not yet been carried out successfully in horses.

Another approach used frequently with cattle involves genotyping large half-sibling families, where each of several males is mated with a large number of females to produce the number of offspring required for linkage analysis. This approach is also applicable to horses and was the basis for the two previously published horse linkage maps (Lindgren *et al.*, 1998; Guérin *et al.*, 1999). Half-sibling families suffer the disadvantages that large numbers of individuals must be typed, and the resulting map is based solely on recombination in the male and cannot therefore generate a map for the X chromosome.

Monozygotic twins have proven to be a useful tool for genetic studies in many species although they do not occur naturally in the horse. The experimental production of monozygotic horse twins by splitting the embryo has also proven technically difficult compared to other species (Allen and Pashen, 1984), due to the combination of the long oviducal transport time of the equine embryo (Battut *et al.*, 1997), the small size of the inner cell mass of the equine blastocyst (Skidmore *et al.*, 1989), and the toughness and elasticity of the glycoprotein capsule that surrounds the equine conceptus between days 6.5 and 23 after ovulation (Betteridge, 1989). Consequently very few pairs of identical twin horses have been generated (Allen and Pashen, 1984; Skidmore *et al.*, 1989; McKinnon *et al.*, 1989). Nevertheless, by using such identical twin mares for reference family production, it is possible to double the number of full-sibling offspring produced in a given time, since the offspring from each twin can be regarded genetically as full-siblings. The still-spherical equine conceptus is easily recoverable between 28 and 33 days of gestation by nonsurgical uterine lavage, and the same mare may be returned to estrus and remated within 3–5 days by a single intramuscular injection of a prostaglandin F analogue given at the time of conceptus recovery. Since sufficient DNA can be extracted from a 28- to 32-day-old embryo to perform thousands of PCRs, the use of monozygotic twin embryos in this manner provided a very practical way in which to generate sufficient genetic material within a relatively short period of time to form a pedigree for horse genetic linkage studies.

This report describes a family generated in this way, consisting of embryos produced by two pairs of monozygotic twin mares mated to one stallion. Three hundred fifty-nine genetic markers were used, 353 of which are microsatellites. Most of the markers used were dinucleotide microsatellites, chosen because of their high levels of polymorphism and the ease with which they can be scored in PCR-based assays. On analysis, 344 proved autosomal in origin and 15 were X-linked markers. Three hundred thirty-four markers (93%) linked to at least one other marker in a total of 42 linkage groups; 25 markers did not show significant linkage to another marker at greater than LOD 3. Thirty-seven of these linkage groups (consisting of a total of 323 markers) were physically assigned to chromosomes, while five linkage groups (consisting of 11 markers) were not physically assigned. The average distance between markers on the map is 10.5 cM (the range being 0–39 cM), and the total length of the linkage groups is 1780 cM.

MATERIALS AND METHODS

Production of conceptuses. Two pairs of monozygotic twin mares that had been produced by embryo bisection and transfer (Allen and Pashen, 1984; Skidmore *et al.*, 1989) were available for use as dams. An Icelandic Horse stallion, loaned by the Holar College of Agriculture in Iceland, was used as the sire. During estrus the mares were monitored by teasing with an active stallion, measurement of progesterone concentrations in daily jugular vein serum samples, and transrectal ultrasound scanning of the uterus and ovaries. When a

dominant Graffian follicle reached 3.5 cm in diameter, the mare was mated every second day until ovulation was detected by the disappearance of the follicle and a rise of serum progesterone concentrations to >1 ng/ml. Pregnancy was diagnosed ultrasonographically between days 14 and 16 after ovulation (Simpson *et al.*, 1982) and monitored weekly thereafter.

Between days 28 and 32, transcervical videoendoscopic hysteroscopy was carried out, as described by Allen and Bracher (1992), to visualize the conceptus, and the placental membranes were ruptured by a sharpened catheter passed through the working channel of the endoscope. A large-bore flexible polythene tube was then passed through the cervix, and the collapsed membranes and embryo were flushed out of the uterus using sterile phosphate-buffered saline (PBS).

The mare was given an intramuscular injection of cloprostenol (Estrumate, Schering-Plough Animal Health, Middlesex, UK) to induce luteolysis and remated during the ensuing estrus, which occurred 3–10 days later.

DNA isolation. The collapsed conceptus was examined under a stereomicroscope, and the embryo in its amnion was dissected free and placed in a dish of clean PBS. The amnion was opened with sterile scissors, and the embryo was washed in another aliquot of clean PBS before it was placed in a serum tube and frozen at −20°C.

Subsequently embryos were thoroughly disrupted in Griffiths tubes in 4 ml Nucleon B buffer, and DNA was prepared using a Nucleon II kit (Scotlabs, Strathclyde, UK) following the manufacturer's instructions except that additional chloroform extractions were carried out until the DNA-containing supernatant appeared clear. Following ethanol precipitation, DNA was spooled onto a glass rod and air-dried for 10 min before resuspension in 1.5 ml MQ water. The DNA remaining in the supernatant was pelleted by centrifugation, washed in 70% ethanol, vacuum dried, and resuspended in 500 μl MQ water. DNA was isolated from the parents and grandparents in the family using the same Nucleon II kit with 100 ml blood collected into EDTA. Again spooled DNA was resuspended in 1.5 ml MQ water, and the pelleted DNA from the supernatant was resuspended in 500 μl MQ water. The concentration of the DNA was estimated by measuring absorbance at 260_{nm}.

Genotyping of microsatellites. Equine microsatellites were genotyped on the reference family (initially comprising grandparents, parents, and 41 offspring) using approximately 40 ng DNA for each reaction. Twelve of the microsatellites were from the StockMarks Horse Paternity PCR Typing Kit (PE Applied Biosystems, Foster City, CA) and were used according to the manufacturer's instructions. The remaining markers included those both previously published and unpublished. A list of the published markers used in this study can be found at http://www.aht.org.uk/genetics/table2.html, and their details can be viewed at http://www.ri.bbsrc.ac.uk/cgi-bin/arkdb/browsers/browser.sh?species=horse and http://locus.jouy.inra.fr/cgi-bin/horsemap/Horsemap/main.pl; the details of unpublished markers can be found in Table 1. Most primer pairs either were pooled into PCR multiplex groups of two, three, four, or five pairs to increase efficiency and reduce costs or were PCR amplified separately and subsequently pooled prior to electrophoresis. The majority of microsatellite primer pairs included one member that was fluorescently labeled at the 5′ end with 6-FAM, TET, HEX, TAMRA, JOE, or NED (PE Applied Biosystems) while for the others 0.6 μM fluorescently labeled dUTP (R110) (PE Applied Biosystems) was included in the PCR. PCRs were performed in microtiter plate format using 0.2-ml thin-walled strip-tubes or plates. The reactions were carried out in 15-μl reaction volumes, generally using the $MgCl_2$ concentrations and annealing temperatures suggested in the relevant marker publication. All PCRs were performed using 1.5 units AmpliTaq GOLD (PE Applied Biosystems), 1× GeneAmp PCR buffer II (PE Applied Biosystems), and 200 μM each dNTP. Most primers were used at a final concentration of 0.67 μM, although several that had been previously optimized in multiplex groups required adjusted concentrations. An MJ Tetrad PCR cycler and a general PCR program with variable annealing temperature (T_a) were used for all PCR amplifications: 95°C for 10 min, followed by 30 cycles of (95°C for 30 s, T_a for 30 s, 72°C for 1 min), and then 72°C for 10 min. Reactions containing fluorescent dUTP were purified using spin columns of Bio-Gel P-10 Gel (Bio-Rad) formed in MultiScreen GV filtration plates (Millipore). Reactions containing fluorescently labeled primers did not require purification. Reactions were stored at −20°C for a maximum of 2 weeks before they were analyzed using an ABI 377 automated sequencer according to the manufacturer's instructions. Filter sets were chosen according to the fluorescent dyes in the reactions. Dilution of PCR products was performed prior to loading, and this was determined empirically for each primer pair. Genotyping data were collected and initially analyzed using Genescan (PE Applied Biosystems). Genotyper (PE Applied Biosystems) was then used, and alleles were called manually, with each allele being assigned an integer value as appropriate.

Biallelic marker typing. The equine single-nucleotide polymorphisms MT-A, MT-B, and MT-D were typed on the reference family using genetic bit analysis (Nikiforov *et al.*, 1994). The MC1R and TXN genes were mapped by PCR-RFLP. MC1R PCR products amplified using the primers EXT1 and EXT4 (Marklund *et al.*, 1996) were digested with *Taq*I and analyzed on 3% agarose gels. Equine TXN sequences were amplified with TXN CATS primers (Lyons *et al.*, 1997), digested with *Msp*I, and analyzed on 2% agarose gels.

DRB was mapped by SSCP analysis using the primers DRB2a and DRB2b (Fraser and Bailey, 1996).

Data handling and two-point linkage analysis. Pedigree information and genotypes were entered and stored in the Cyrillic 2.1 database (Chirwell Scientific, Oxford, UK). For linkage analysis, data were exported from Cyrillic in Crimap format and initially analyzed using the two-point option of Crimap. This software highlighted any cases of non-Mendelian inheritance, which were very infrequent and usually caused by incorrect scoring or data transfer. Only three examples—in markers ASB22, COR041, and COR065 (COR041 and COR065 are the same locus, see Discussion)—that appear to be genuine cases of novel mutation were encountered.

Linkage groups were generated that were significant at LOD 3 and assigned to chromosomes via markers previously physically mapped by fluorescence *in situ* hybridization (FISH). Most physical assignments were from previously published data. Other assignments were either via markers isolated from cosmid clones (Swinburne *et al.*, in press) or via markers originally isolated from small-insert libraries but for which a BAC was identified and FISH mapped (Lindgren *et al.*, submitted for publication).

Linkage groups that were not assigned to a chromosome by physical mapping were otherwise assigned by synteny mapping data (Shiue *et al.*, 1999; Hopman *et al.*, 1999; Murphie *et al.*, 1999; Ruth *et al.*, 1999; Tallmadge *et al.*, 1999a,b). Where two or more linkage groups were assigned to the same chromosome, genotyping of an additional 20 embryos was undertaken for selected markers to try to link the groups at LOD >3.

Map construction. The genotyping data were separated for each chromosome, and multipoint analysis programs in CRIMAP were used to generate sex-averaged linkage maps for each chromosome. Initially the BUILD option was used at a lod threshold of 3 to produce a framework order of reliably positioned markers. BUILD was then used at LOD 2 and then LOD 1 to suggest the probable order of the other markers with respect to the framework. The positions of markers that could not be ordered at a significance level of LOD 1 are depicted with a dotted line to the right of the map in Fig. 2. The ALL option was then used, starting with the framework order obtained from BUILD, to enter markers sequentially one-by-one in decreasing order of significance as suggested by BUILD. This confirmed the order, obtained in BUILD, of the less informative markers. The FLIPS option was then used at a significance level of LOD 3 to test further the order of adjacent markers. The FIXED option was used to calculate the genetic distance between the markers in their final order.

Sex-averaged map distances were generated for all autosomes and female-specific map distances for the X chromosome.

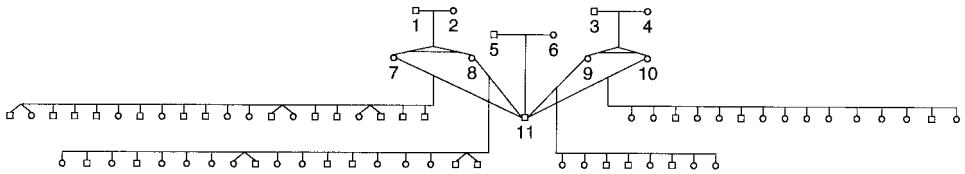

FIG. 1. Pedigree structure of the horse reference family used in this study. Two pairs of monozygotic twin mares (Cordelia and Regan, Ophelia and Desdemona) were used together with a single stallion (Thinur) to generate two families of full-sibling horse embryos. Females are depicted by circles, and males are depicted by squares. All twin embryos retrieved were found to be dizygotic in origin. 1, unnamed Arabian; 2, Concha, Thoroughbred; 3, Othello, Welsh Cob; 4, Queenslander, Polish Warmblood; 5, Reynir fra Holum, Icelandic Horse; 6, Thorf, Icelandic Horse; 7, Regan; 8, Cordelia; 9, Desdemona; 10, Ophelia; 11, Thinur.

RESULTS

Production of the Reference Family

In preliminary experiments, conceptuses were removed from the mares and dissected *in vitro* to extract the embryo from the enveloping membranes. It was found that 3–5 mg DNA could be extracted from each 28- to 32-day-old embryo and importantly, that in PCR assays the DNA produced showed no sign of being contaminated with maternally derived DNA. This quantity of DNA is sufficient for approximately 10,000 PCR assays, thereby providing adequate DNA to carry out extensive genetic linkage studies. The generation of multiple conceptuses from the two pairs of identical twin mares was therefore begun.

The first two conceptuses were removed in July 1995, and from then until November 1999, DNA was isolated from a total of 61 embryos, of which 38 were produced from the mares Regan and Cordelia and 23 from the mares Ophelia and Desdemona. The structure of the family is shown in Fig. 1, together with the names, breeds, and sex of each individual involved. All twin embryos retrieved were dizygotic.

Genotyping

All equine microsatellite markers that were known to the authors during the period of this study were genotyped on this family. This included many markers that were contributed prepublication from a number of international collaborating laboratories (see Acknowledgments). Initially, markers were typed on the stallion and on one of each pair of the twin mares, to determine which markers were informative in the family. These informative markers were then genotyped on all grandparents, parents, and 41 F_2 embryos. Of 498 markers screened, 359 proved to be polymorphic under these criteria, 88 were noninformative, and 51 did not produce sufficiently clear results to be acceptable. These last rejected markers either did not amplify a product or amplified multiple products that could not be unequivocally scored. HTG13 amplified two loci, one of which was scored. A small minority of markers produced spurious bands in addition to amplification at the microsatellite locus, but these were scored provided that the true alleles could be distinguished easily. Several markers segregated one null allele in the reference family, and the null alleles were given an arbitrary number during scoring. HMS16 and LEX055 segregated two null alleles in the reference family. Other markers were discarded after genotyping as their inheritance did not appear to be Mendelian and could not be explained by the presence of null alleles or novel mutations. The two pairs of identical twins served as internal controls to monitor possible typing errors; on no occasion were these twins scored with different genotypes, and so we are confident that the level of typing error is low.

Two-Point Analysis

The 359 polymorphic markers for which results were obtained comprised 353 microsatellites and 6 biallelic markers and their details are presented at http://www.aht.org.uk/genetics/table2.html and in Table 1. The majority of microsatellites were dinucleotide repeats, 3 were trinucleotide repeats, and 1 was a tetranucleotide repeat. A total of 334 markers (93%) were significantly linked to at least 1 other marker at a significance level of LOD 3, forming 42 linkage groups.

These linkage groups were compared with the synteny groups generated by somatic cell-hybrid analysis (Shiue *et al.*, 1999; Hopman *et al.*, 1999; Murphie *et al.*, 1999; Ruth *et al.*, 1999; Tallmadge *et al.*, 1999a,b), and only one contradiction was encountered. This involved marker HMS16, which had been assigned to synteny group UCD6, but which showed significant linkage to a marker in the group assigned to chromosome 2 (A14: $\theta = 7\%$, LOD 5.27). On further examination, it appears possible that the synteny data are not conclusive, as HMS16 is only tentatively syntenic with one marker in UCD6 (A. T. Bowling, Davis, CA, Aug. 1999, pers. comm.). Several linkages were generated at LOD 3–4 that were assumed to be spurious, as overwhelming evidence existed to contradict these suggested linkages. Such misleading linkages are statistically to be expected in such a large data set.

One linkage group, comprising 15 microsatellites, is X chromosome linked, highlighting a significant advantage of full-sibling families. Characterization of X chromosome-linked markers also provided a sex assignment for each embryo.

Assignment of Linkage Groups to Chromosomes

Thirty-seven of the linkage groups were assigned to chromosomes by means of 85 physically assigned

TABLE 1

Marker Information (Unpublished Markers)

Marker name	Chromosome assignment	PCR product range	FISH assignment	FISH reference	Number of informative meioses	Forward primer sequence	Reverse primer sequence	Source
1CA01	1	151–153			27	TTCAAAAGCGGAAGAGCATT	CATGCTGTGGTAGCATCCC	AF043198
1CA12	1	100–106			63	GGGAGTGGTGATTACTTCTTGC	TAGCCGTGAGAAGGTGTGTG	AF043199
1CA16	1	115–123			84	TCACTGGGGGGTATATGCAT	GATCCTACTCCACCTGAAGTGG	AF043200
1CA18	1	140–142			59	AGAAAGCAAGTTCGCTAGATGG	AGTTCCCAAGGAATGTGTGTG	AF043201
1CA20	1	112–122			61	CTAAGCAGGTTCCCTATCATGG	TCCACTACACAGGAAAACGAA	AF043202
1CA25	1	202–207			23	TCCAATTTTCCCCAATGGTA	CTGCATTTTGACAATGGTGG	AF043205
1CA30	1	128–132			40	TGGGGAGGGGGTTGTTCTAG	GCTCTCCACTGCATGCATAA	AF043209
1CA32	1	101–107			86	AGTTACCAAATGTCGCATTGC	TTCATCTGTAAAATGGGCAGG	AF043210
1CA41	1	237–247			86	CTGGTCAGGCCTATTACCCA	AGATATTGGGGGCGGAAG	AF043214
1CA43	1	122–126			63	ATGGCATGATTTGCTTCTCC	ATGGAAACAACCTAAATGTCCA	AF043215
AHT7	25	121–133			52	CCTTAGATCCGAGAAGGAGA	GAAGCCTCACTCCATCCAGG	
ASB29	4	123–135			27	CTGGCCCATAAAAAAACACTG	TGTATGGTTGTCAGCTCAAACC	X93543
COR101	18	232–272			86	CCATATAAGTGCACGCCTTC	TGGACCTTGAGGGTATGATG	D. Antczak
COR102	10	155–226			64	GAAGAGGCAGAGGATTCACA	GTGCTCCCCTAAACCTCAGT	D. Antczak
COR103	14	167–169			59	GGGAGTGTGTCCAGTTTGTC	CCAGATAAAGCCCAAATCCT	D. Antczak
COR104	14	177–188			84	GGGAGTGTGTCCAGTTTGTC	CCAGATAAAGCCCAAATCCT	D. Antczak
COR105	17	174–192			68	TTTCCTCATTGCTTCCTGAG	CCCAAGGTCTGTCTTGCTCTC	D. Antczak
HP12	29	132–148			84	Skow *et al.*, manuscript in preparation		L. Skow
HP13	23	134–138			47	Skow *et al.*, manuscript in preparation		L. Skow
HP27	1	136–152	1p15–p16	T. Lear, Kentucky, July 1999, pers. comm.	84	Skow *et al.*, manuscript in preparation		L. Skow
HMS4	Unassigned	106–112			25	GCAAACAACTGTGCAATAGAT	AATTCATCCCAATTGGCTGG	G. Guerin
RC01	17	137–145			85	Mashima *et al.*, manuscript in preparation		S. Mashima
RC07	15	155–157			20	Mashima *et al.*, manuscript in preparation		S. Mashima
RC08	31	196–204			39	Mashima *et al.*, manuscript in preparation		S. Mashima
RC09	1	129–133			26	Mashima *et al.*, manuscript in preparation		S. Mashima
RC11	6	127–129	6q13	Lindgren *et al.*, submitted	60	Mashima *et al.*, manuscript in preparation		S. Mashima
RC12	29	90–108			126	Mashima *et al.*, manuscript in preparation		S. Mashima
UCDEQ39	11	232–256			52	CTGCCGCTTTCCAAGACT	CTTGGGTCGAGTAATAGC	U25168
UCDEQ66	Unassigned	141–145			26	GAAAGAGGAAAAGAGCCAGAGT	ACACGCTCACCCACACAT	U25169
UCDEQ428	X	125–137			65	CTTTTCCCCGAACCTCCTAC	TTGGATGCTCCGAGAAGAGT	U67407
UM032	14	141–149			63	AAATGGTCAGCCTCTCCTC	TGTCTCTCTAGTCCCACTCCTC	J. Mickleson
UM041	1	105–115			47	TGCCCTTCCATGAACAGAC	TCCCTCTCTCTCTCTCCTTCTC	J. Mickleson
VIASH21	31	247–249	31q14–q15	Lindgren *et al.*, submitted	21	AAATGATAACGCCAAGTGCTCT	ATGTGAGTGCCAGCTTGTGAT	L10926

Note. Additional information can be found at http://www.aht.org.uk/genetics/table2.html, and all details of these markers can be found at http://www.ri.bbsrc.ac.uk/cgi-bin/arkdb/browsers/browser.sh?species=horse and http://locus.jouy.inra.fr/cgi-bin/horsemap/Horsemap/main.pl. Any marker that is not assigned to a chromosome via linkage to FISH or synteny mapped markers is labeled "unassigned," and none of these markers has shown significant linkage (LOD >3) to any other marker. Markers COR103 and COR104 are two linked loci that are amplified by the same primer pair. The locations of physically positioned markers are given together with relevant references. The number of informative meioses produced with each marker on the full-sibling family is given. The primer sequences are given for unpublished markers where possible, and the source of the marker is cited. Markers prefixed "1CA" were developed from sequences produced from a chromosome 1-specific library deposited in GenBank.

markers. Two of these markers (Lex025 and VHL20) were physically assigned to chromosome 30 by FISH mapping them onto a trisomic cell line (Bowling *et al.*, 1997a). Six inconsistencies were encountered in relation to the physical map assignments of markers belonging in the same linkage group, and we suggest that these fall into two categories. First, some assignments had apparently been made incorrectly due to the misidentification of similar chromosomes. Second, it appears that some clones are chimeric with the microsatellite marker originating from a minor fragment in the chimera. In particular, it seems that equine phage libraries contain a high proportion of chimeric clones. Three of these markers (AHT44, ASB14, and ASB38) have been synteny mapped (Shiue *et al.*, 1999), and these data concur with the linkage data in suggesting that their published FISH positions were incorrect. The inconsistencies are indicated at http://www.aht.org.uk/genetics/table2.html and in Table 1.

Confirmatory evaluation of the correct assignment of some physically mapped markers was performed, especially on those involving the small acrocentric chromosomes 24–31, which are more difficult to identify cytogenetically. This involved FISH mapping several small groups of markers to confirm to which autosomes each marker belonged.

Five linkage groups, consisting in total of 11 markers, did not contain a physically mapped marker, and their chromosome assignment is cited on the basis of synteny mapping data (Shiue *et al.*, 1999; Hopman *et al.*, 1999; Murphie *et al.*, 1999; Ruth *et al.*, 1999; Tallmadge *et al.*, 1999a,b). These five groups showed linkage to others on the same chromosome at significance levels of LOD 0.9–2.1.

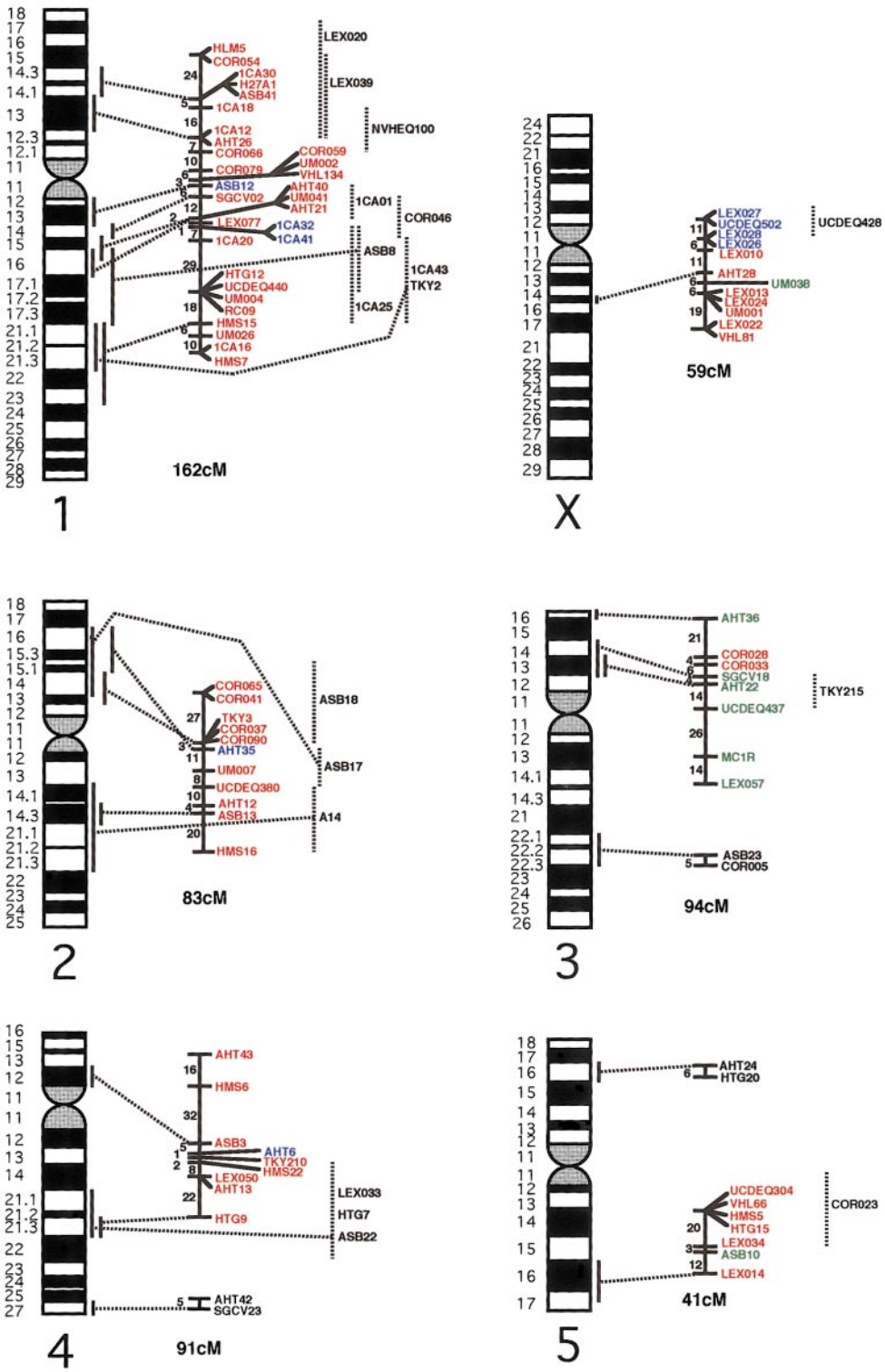

FIG. 2. Sex-averaged genetic maps of the horse autosomes and female-specific map of the X chromosome. Idiograms of each chromosome together with band numbering are taken from the standardized karyotype of the horse (Bowling *et al.*, 1997b). Markers whose physical positions are known are anchored on the idiogram (solid vertical bars adjacent to the idiogram). Linkage groups have also been assigned to chromosomes on the basis of synteny information (Shiue *et al.*, 1999; Hopman *et al.*, 1999; Murphie *et al.*, 1999; Ruth *et al.*, 1999; Tallmadge *et al.*, 1999a,b) and these markers are identified with an asterisk at http://www.aht.org.uk/genetics/table2.html and in Table 1. Markers whose order was significant at a lod threshold of 3 are shown in red, those whose order was significant at a lod threshold of 2 are shown in blue, and those whose order was significant at a lod threshold of 1 are shown in green. Markers whose order could not be determined at a lod threshold of 1 are shown to the right of the maps, and their most likely position is depicted with a broken line. Markers with no recombination between them are shown with a branched marker position. The map length (in centimorgans) is indicated beneath each chromosome map. The distances between markers (in centimorgans) are indicated between marker positions.

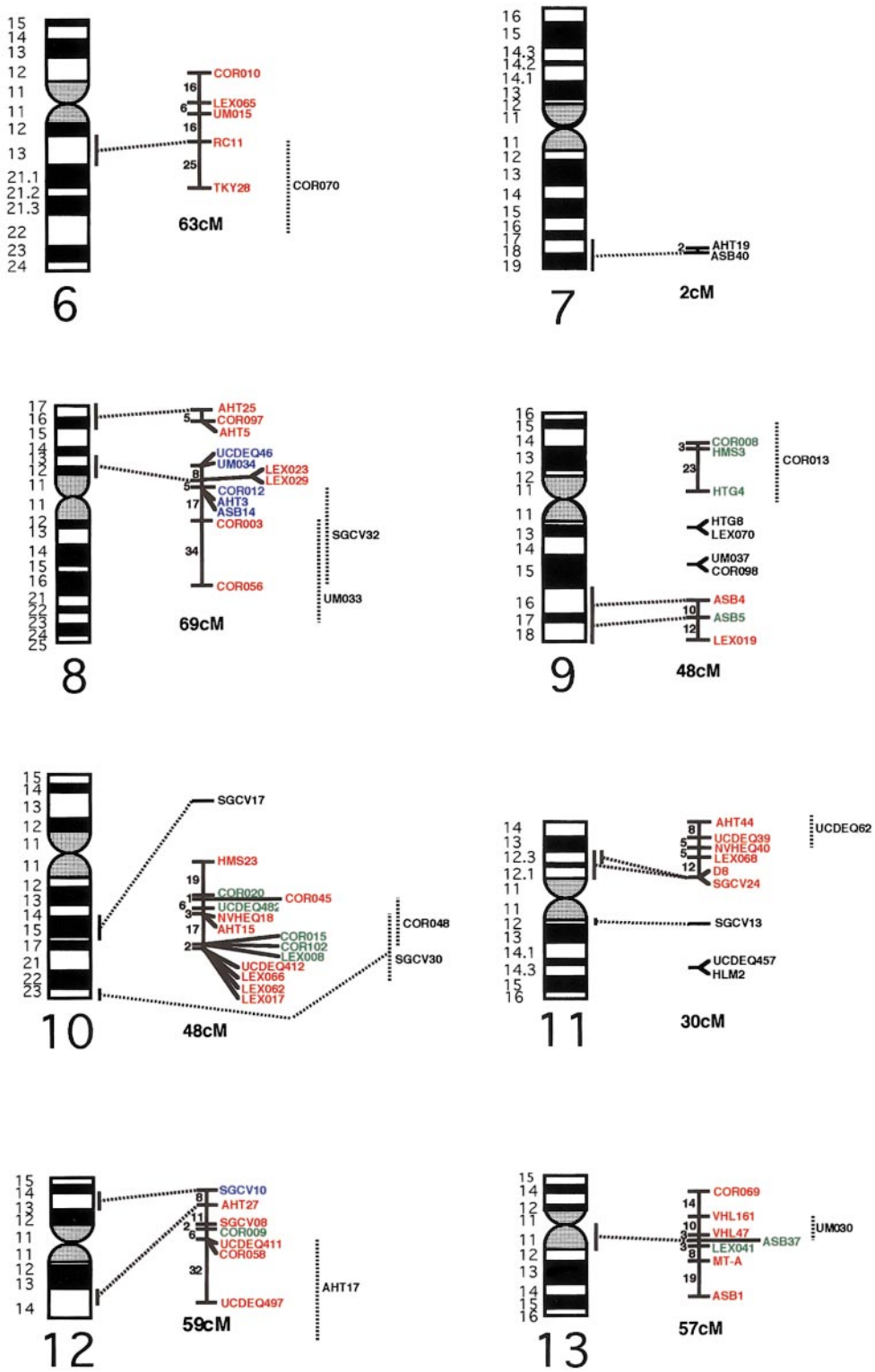

FIG. 2—*Continued*

Map Construction

Multipoint analysis was used to generate sex-averaged maps for each of the 31 autosomes and a female-specific map for the X chromosome. These are depicted in Fig. 2. The linkage groups on each chromosome vary in length from 0 cM comprising 2 markers (ECA 9,11,29) to 162 cM comprising 39 markers (ECA 1) and collectively are 1780 cM in length. The average distance between markers on the assembled map is 10.5 cM. Six markers that belong to linkage groups (ASB29, F18, LEX003, LEX007, NVHEQ29, and TKY222) are

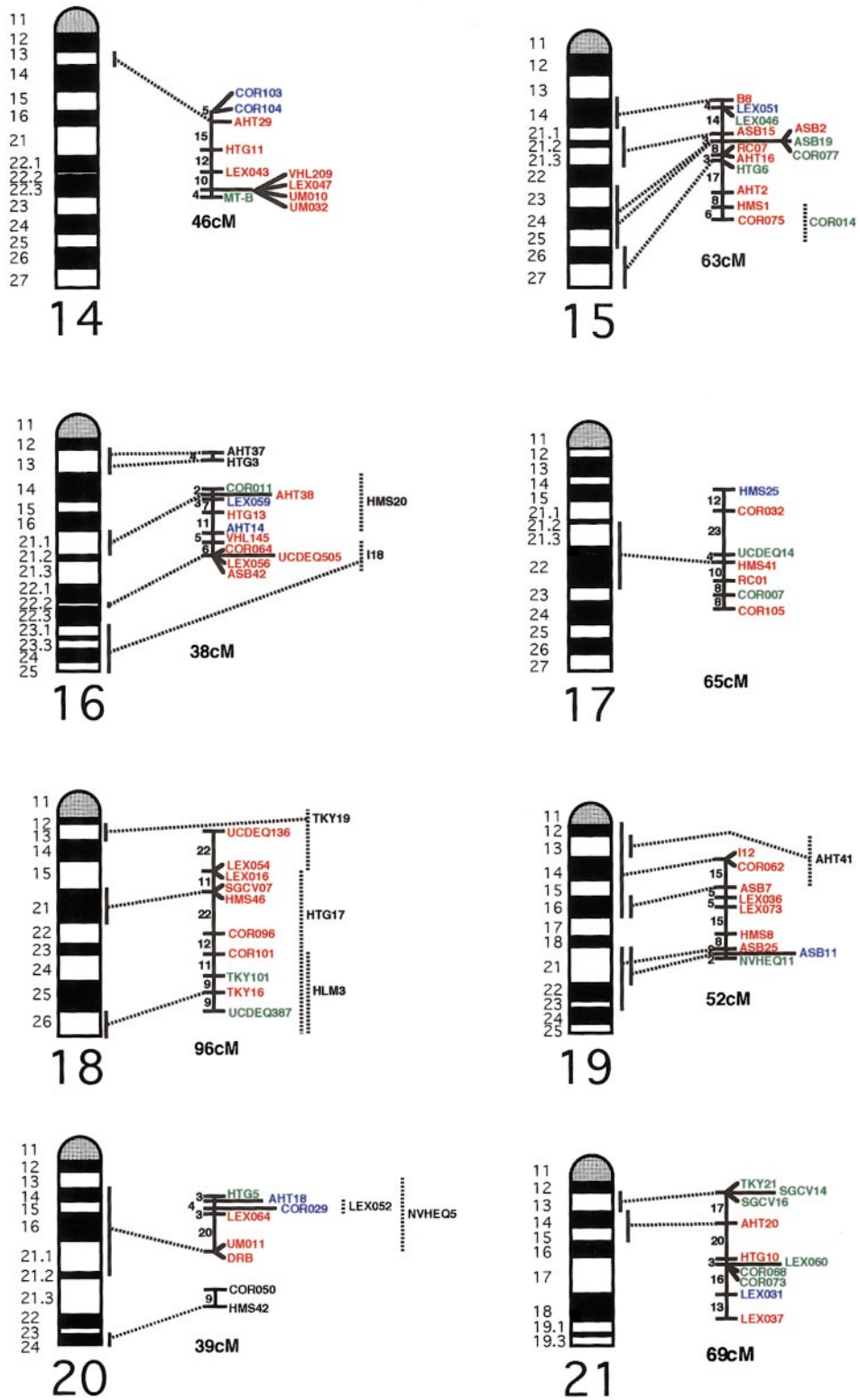

FIG. 2—*Continued*

not depicted on the map, as their position could not be localized to a specified region.

DISCUSSION

We present the first genetic map of the horse to assign markers to all 31 autosomes and the X chromosome. This map has been generated by genotyping 359 polymorphic markers, the majority of which are dinucleotide microsatellites, onto a novel reference pedigree consisting of two full-sibling three-generation families (Fig. 1).

The Icelandic Horse stallion (Thinur) used to gener-

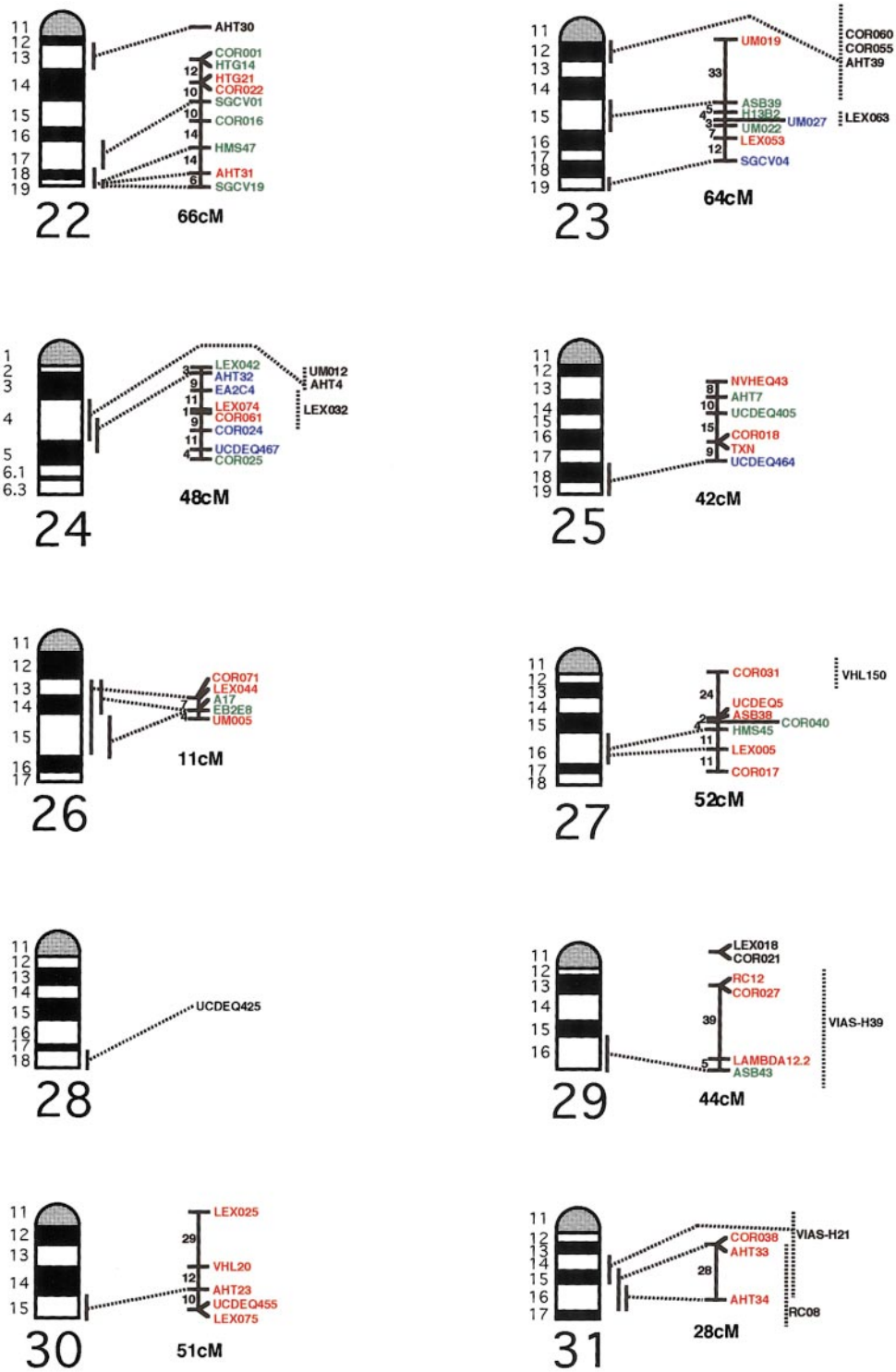

FIG. 2—*Continued*

ate all F_2s is a breed most likely to be maximally diverged from other breeds of horse. The original Viking settlers introduced the Icelandic Horse to Iceland, where it remained geographically isolated for 900 years, due to a ban on horse importation during this period. The two pairs of twin mares (Cordelia and Regan, Ophelia and Desdemona) are Arabian × Thoroughbred and Polish Warmblood × Welsh Cob, respectively. Mixing so many various breeds within the pedigree was an attempt to maximize heterozygosity. DNA was isolated from all five parents in the family and from all but one of the grandparents, which had died. The unique structure of this family, produced by the removal of embryos from the two pairs of monozygotic twin mares, allowed the construction of a genetic linkage map with maximum efficiency. Furthermore, the

use of genetically full-sibling individuals to perform linkage analysis greatly reduced the number of animals required to generate a robust map, compared to a half-sibling family approach. For example, only 51 individuals were genotyped for the majority of markers analyzed in this study compared to 271 and 458 in the studies by Lindgren *et al.* (1998) and Guérin *et al.* (1999). The previous maps are male-specific linkage maps and as such might be expected to be shorter than a female-specific or sex-averaged map (Haldane, 1922). Nonetheless, this map is a substantial advancement on the preceding versions. Although limitations exist in the approach taken in the present study—some markers cannot be incorporated into the map because they are not polymorphic, no phenotypic data can be obtained from the embryos, the quantity of DNA obtained from the embryos is finite—these limitations are insignificant in comparison to the advantages of the genetic material obtained.

The genetic linkage map generated with this pedigree comprises 42 linkage groups (consisting of 334 markers), 37 of which are physically anchored onto the horse genome. A high proportion of linked markers (176; 53%) have been ordered at a lod threshold of >3 (depicted in red in Fig. 2). An additional 88 markers (26%) have been ordered at lod thresholds of >2 and >1 (blue and green, respectively, in Fig. 2). It is anticipated that the 47 markers (14%) that could not be ordered with this degree of certainty (shown to the right of the maps in Fig. 2) will be integrated into the order as additional markers are incorporated in the future.

A strategic position has been reached now that genetic markers are assigned to all 31 of the horse autosomes and the X chromosome. However, the linkage groups assigned to 10 chromosomes (ECA 6, 7, 13, 14, 17, 25, 28, 29, 30, and X) still have only one physically mapped marker assigned to them, and the development of additional physically mapped markers will be valuable. The isolation and subsequent FISH mapping of BAC clones that contain other markers in these linkage groups could achieve this. As more physically mapped markers are added, the assignment of additional linkage groups to chromosomes will be possible, and the orientation of linkage groups in relation to the chromosome will become clearer. The isolation of microsatellites from the microdissected terminal ends of chromosome arms would be another valuable addition to the map.

The linkage groups generated here concur with the analysis of a synteny panel tested with a subset of these markers (Shiue *et al.*, 1999; Hopman *et al.*, 1999; Murphie *et al.*, 1999; Ruth *et al.*, 1999; Tallmadge *et al.*, 1999a,b). The synteny groups A, B, C, and D correspond to linkage groups assigned by recent FISH mapping assignments to ECA 27, 31, 28, and 6, respectively. Synteny group UCD6 is now acknowledged to belong to ECA 8, rather than ECA 6.

In addition, the map described here does not contradict the chromosome assignments made for any markers described in the two previous mapping reports (Lindgren *et al.*, 1998; Guérin *et al.*, 1999), other than the exchange of groups assigned to ECA 6 and 8 (Lindgren *et al.*, 1998). The ordering of all markers that these other studies have in common is not always identical, but this could not be expected in maps derived from different data sets as the quantity of information obtained for each marker will vary.

Each of four pairs of markers depicted on the map, with 0% recombination between each pair, exhibits the same pattern of alleles in the family. The sequences for these pairs (COR065 and COR041, COR055 and COR060, LEX024 and UM001, LEX062 and LEX066) were compared and found to be identical. Other marker pairs with 0% recombination between them are thought to be closely linked markers where the number of meioses examined here is not sufficient to separate them.

The total genetic distance covered by the linkage groups is 1780 cM. This map is therefore much more extensive than the maps of 679 and 936 cM described in previous studies (Lindgren *et al.*, 1998; Guérin *et al.*, 1999). Since 93% of the markers link to another marker at a significance level of LOD >3, we suggest that a significant proportion of the horse genome is covered by this map. Assuming conservatively that each linkage group ($n = 42$) and each unlinked marker ($n = 25$) could detect linkage over an extra 5 cM on each side, we estimate that this map covers 2450 cM. This approaches the total map length estimated from chiasma counts in horses (Scott and Long, 1980) of 2720 cM (Lindgren *et al.*, 1998).

The incorporation of further markers to extend map coverage across the length of all chromosomes should now be used to develop the equine linkage map. The development of multiplex groups comprising approximately 300 evenly spaced markers would then allow efficient whole-genome scanning. As this work progresses, the horse genetics community should make sure that pedigrees of animals that segregate interesting diseases and traits are available for testing.

In addition, to exploit data available from "map-rich" species such as human and mouse, horse radiation hybrid mapping should be undertaken. The mapping of expressed sequence tags on a radiation hybrid panel will allow the alignment of evolutionarily conserved genes in the horse with those in human or mouse. Homologous regions identified by this comparative mapping will enable the identification of positional candidate genes in the horse that may be responsible for interesting horse traits.

The implementation of this work during subsequent years will ensure that significant progress will be made toward the identification of, and development of screening tests for, genes responsible for important traits in the horse.

ACKNOWLEDGMENTS

We thank the Horserace Betting Levy Board for continued support of the Horse Genetic Linkage Mapping Programme and the Childwick Trust for their generous support for genetic research in the horse. We gratefully acknowledge the Dorothy Russell Havemeyer Foundation, Inc., for providing support for the workshops that facilitated the collaboration required for this undertaking. We also thank Dr. Janet Ziegle (Perkin–Elmer) for the provision of fluorescently labeled oligonucleotide primers and Anna Bristow (Animal Health Trust) for technical support. We thank all of those laboratories that donated primer pairs prior to publication—namely, James A. Baker Institute for Animal Health, Cornell University (Ithaca, NY); INRA Centre de Recherches de Jouy (Jouy-en-Josas, France); M. H. Gluck Equine Research Center, Department of Veterinary Science, University of Kentucky (Lexington, KY); Veterinary Genetics Laboratory, University of California, Davis (Davis, CA); Swedish University of Agricultural Sciences, (Uppsala, Sweden); College of Veterinary Medicine, University of Minnesota, St. Paul, MN); School of Veterinary Medicine, Texas A&M University (College Station, Texas); Norwegian College of Veterinary Medicine (Oslo, Norway); Dr. van Haeringen Laboratorium (Wageningen, The Netherlands); Australian Equine Blood Typing Research Laboratory, University of Queensland (Brisbane, Australia). We thank the Holar Agricultural College in Saudarkrojur, Iceland, for the generous donation of Thinur to the TBA Equine Fertility Unit.

REFERENCES

Allen, W. R., and Bracher, V. (1992). Videoendoscopic evaluation of the mare's uterus. III. Findings in the pregnant mare. *Equine Vet. J.* **24:** 285–291.

Allen, W. R., and Pashen, R. L. (1984). Production of monozygotic (identical) horse twins by embryo micromanipulation. *J. Reprod. Fertil.* **71:** 607–613.

Barandese, W., Armitage, S. M., Kossarek, L. M., Shalom, A., Kirkpatrick, B. W., Ryan, A. M., Clayton, D., Li, L., Neibergs, H. L., and Zhang, N. (1994). A genetic linkage map of the bovine genome. *Nat. Genet.* **6**(3)**:** 227–235.

Battut, I., Colchen, S., Fieni, F., Tainturier, D., and Bruyas, J. F. (1997). Success rates when attempting to nonsurgically collect equine embryos at 144, 156 or 168 hours after ovulation. *Equine Vet. J. Suppl.* **25:** 60–62.

Betteridge, K. J. (1989). The structure and function of the equine capsule in relation to embryo manipulation and transfer. *Equine Vet. J. Suppl.* **8:** 92–100.

Bowling, A. T., Millon, L. V., and Dileanis, S. (1997a). Physical mapping of genetic markers to chromosome 30 using a trisomic horse and evidence for maternal origin of the extra chromosome. *Chromosome Res.* **5:** 429–431.

Bowling, A. T., Breen, M., Chowdhary, B. P., Hirota, K., Lear, T., Millon, L. V., Ponce de Leon, F. A., Raudsepp, T., and Stranzinger, G. (1997b). International system for cytogenetics nomenclature of the domestic horse. *Chromosome Res.* **5:** 433–443.

Breen, M., Lindgren, G., Binns, M. M., Norman, J., Irvin, Z., Bell, K., Sandberg, K., and Ellegren, H. (1997). Genetical and physical assignments of equine microsatellites—First integration of anchored markers in horse genome mapping. *Mamm. Genome* **8:** 267–273.

Copeland, N. G., Jenkins, N. A., Gilbert, D. J., Eppig, J. T., Maltais, L. J., Miller, J. C., Dietrich, W. F., Weaver, A., Lincoln, S. E., and Steen, R. G. (1993). A genetic linkage map of the mouse: Current applications and future prospects. *Science* **262:** 57–66.

Dausset, J., Cann, H., Cohen, D., Lathrop, M., Lalouel, J-M., and White, R. (1990). Centre d'Etude du Polymorphisme Humain (CEPH): Collaborative genetic mapping of the human genome. *Genomics* **6:** 575–577.

Dietrich, W. F., Miller, J. C., Steen, R. G., Merchant, M., Damron, D., Nahf, R., Gross, A., Joyce, D. C., Wessel, M., and Dredge, R. D. (1994). A genetic map of the mouse with 4,006 simple sequence length polymorphisms. *Nat. Genet.* **7**(2)**:** 220–245.

Ellegren, H., Johansson, M., Sandberg, K., and Andersson, L. (1992). Cloning of highly polymorphic microsatellites in the horse. *Anim. Genet.* **23:** 133–142.

Fraser, D. G., and Bailey, E. (1996). Demonstration of three *DRB* loci in a domestic horse family. *Immunogenetics* **44:** 441–445.

Godard, S., Vaiman, D., Oustry, A., Nocart, M., Bertaud, M., Guzylack, S., Meriaux, J. C., Cribiu, E. P., and Guérin, G. (1997). Characterisation, genetic and physical mapping analysis of 36 horse plasmid and cosmid-derived microsatellites. *Mamm. Genome* **8:** 745–750.

Guérin, G., Bailey, E., Bernoco, D., Anderson, I., Antczak, D. F., Bell, K., Binns, M. M., Bowling, A. T., Brandon, R., Cholewinski, G., Cothran, E. G., Ellegren, H., Förster, M., Godard, S., Horin, P., Ketchum, M., Lindgren, G., McPartlan, H., Mériaux, J. C., Mickelson, J. R., Millon, L. V., Murray, J., Neau, A., Røed, K., Sandberg, K., Shiue, Y-L., Skow, L. C., Stott, M., Swinburne, J. E., Valberg, S. J., Van Haeringen, H., Van Haeringen, W. A., and Ziegle, J. (1999). Report of the International Equine Gene Mapping Workshop: Male linkage map. *Anim. Genet.* **30:** 341–354.

Haldane, J. B. S. (1922). Sex ration and unisexual sterility in hybrid animals. *J. Genet.* **12**(2)**:** 101–109.

Hopman, T. J., Han, E. B., Story, M. R., Schug, M. D., Aquadro, C. F., Bowling, A. T., Murray, J. D., Caetano, A. R., and Antczak, D. F. (1999). Equine dinucleotide repeat loci COR001–COR020. *Anim. Genet.* **30:** 225–226.

Lindgren, G., Sandberg, K., Persson, H., Marklund, S., Breen, M., Sandgren, B., Carlsten, J., and Ellegren, H. (1998). A primary male autosomal linkage map of the horse genome. *Genome Res.* **8:** 951–966.

Lindgren, G., Breen, M., Swinburne, J. E., Mariat, D., Sandberg, K., Guérin, G., Ellegren, H., and Binns, M. M. Anchoring equine linkage groups by FISH mapping BAC clones containing microsatellite markers. Submitted for publication.

Lyons, L. A., Laughlin, T. F., Copeland, N. G., Jenkins, N. A., Womack, J. E., and O'Brien, S. J. (1997). Comparative anchor tagged sequences (CATS) for integrative mapping of mammalian genomes. *Nat. Genet.* **15:** 47–56.

Marklund, S., Ellegren, H., Eriksson, S., Sandberg, K., and Andersson, L. (1994). Parentage testing and linkage anaylsis in the horse using a set of highly polymorphic microsatellites. *Anim. Genet.* **25:** 19–23.

Marklund, L., Johansson Moller, M., Sandberg, K., and Andersson, L. (1996). A missense mutation in the gene for melanocyte-stimulating hormone receptor (*MC1R*) is associated with the chestnut coat color in horses. *Mamm. Genome* **7:** 895–899.

McKinnon, A. O., Carnevale, E. M., Squires, E. L., Carrey, N. J., and Seidel, G. E. (1989). Bisection of equine embryos. *Equine Vet. J. Suppl.* **8:** 126–128.

Murphie, A. M., Hopman, T. J., Schug, M. D., Aquadro, C. F., Bowling, A. T., Murray, J. D., Caetano, A. R., and Antczak, D. F. (1999). Equine dinucleotide repeat loci COR021–COR040. *Anim. Genet.* **30:** 235–237.

Nikiforov, T. T., Rendle, R. B., Goelet, P., Rogers, Y. H., Kotewicz, M. L., Anderson, S., Trainor, G. L., and Knapp, M. R. (1994). Genetic bit analysis: A solid phase method for typing single nucleotide polymorphisms. *Nucleic Acids Res.* **22:** 4167–4175.

Ruth, L. S., Hopman, T. J., Schug, M. D., Aquadro, C. F., Bowling, A. T., Murray, J. D., Caetano, A. R., and Antczak, D. F. (1999). Equine dinucleotide repeat loci COR041–COR060. *Anim. Genet.* **30:** 320–321.

Scott, I. S., and Long, S. E. (1980). An examination of chromosomes in the stallion (*Equus caballus*) during meiosis. *Cytogenet. Cell Genet.* **26:** 7–13.

Shiue, Y-L., Bickel, L. A., Caetano, A. R., Millon, L. V., Clark, R. S., Eggleston, M. L., Michelmore, R., Bailey, E., Guérin, G., Godard, S., Mickelson, J. R., Valberg, S. J., Murray, J. D., and Bowling, A. T. (1999). A synteny map of the horse genome comprised of 240 microsatellite and RAPD markers. *Anim. Genet.* **30:** 1–9.

Simpson, D. J., Greenwood, R. E. S., Ricketts, S. W., Rossdale, P. D., Sanderson, M., and Allen, W. R. (1982). Use of ultrasound echography for early diagnosis of single and twin pregnancy in the mare. *J. Reprod. Fertil. Suppl.* **32:** 431–439.

Skidmore, J., Boyle, M. S., Cran, D., and Allen, W. R. (1989). Micromanipulation of equine embryos to produce monozygotic twins. *Equine Vet. J. Suppl.* **8:** 126–128.

Swinburne, J. E., Lockhart, L., Aldridge, V., Marti, E., Breen, M., and Binns, M. M. Characterisation of twenty-five new physically mapped horse microsatellite loci: AHT24–48. *Anim. Genet.,* in press.

Tallmadge, R. L., Hopman, T. J., Schug, M. D., Aquadro, C. F., Bowling, A. T., Murray, J. D., Caetano, A. R., and Antczak, D. F. (1999a). Equine dinucleotide repeat loci COR061–COR080. *Anim. Genet.* **30:** 462–463.

Tallmadge, R. L., Evans, K. G., Hopman, T. J., Schug, M. D., Aquadro, C. F., Bowling, A. T., Murray, J. D., Caetano, A. R., and Antczak, D. F. (1999b). Equine dinucleotide repeat loci COR081–COR100. *Anim. Genet.* **30:** 470–471.