BARD

FINAL REPORT

PROJECT NO. US-1687-89C

Mapping the Bovine Genome Using DNA Level Markers

J. Womack, M. Soller, A. Friedman

1994

Date October 28, 1994

BARD
P.O. Box 6
Bet-Dagan, ISRAEL

BARD Project No. US-1687-89RC

Title: Mapping the Bovine Genome Using DNA Level Markers

Investigator's

James E. Womack
Morris Soller

Institutions

Texas A&M University
The Hebrew University

Project's starting date: Aug. 10, 1990

Type of Report: Final

Signatures:

Principal Investigator

A handwritten signature in cursive script, appearing to read "James E. Womack", written over a horizontal line.

Institution's Authorizing Official

A handwritten signature in cursive script, appearing to read "Anne Hoffman", written over a horizontal line.

II.	TABLE OF CONTENTS	Page
III.	Abstract	3
IV.	Objectives of Proposal	3
V.	Relevant Data, Methodology, Results and Discussion	3
	A. Data	3
	B. Methodology	4
	C. Results and Discussion	6
	1. Somatic cell mapping of genes of the bovine immune system	6
	a. T-cell receptor CD3 and CD8	6
	b. Immunoglobulin genes	7
	2. Somatic cell mapping and in situ mapping of the bovine uridine monophosphate synthase gene.	8
	3. Genetic characterization and mapping of the interferon gene family in cattle.	9
	4. Mapping of other Type I anchor loci in cattle	9
	a. Visual transduction genes	9
	b. RASA	9
	c. Uroplakins	9
	d. Prion proteins	10
	5. PCR based assignments of Type I loci to the physical map	10
	6. Mapping microsatellites	11
	7. A genetic linkage map of the bovine genome.	11
VI.	Cooperation: Summary of Work of Israel Group	12
	1. Development of microsatellite markers	13
	2. Single-parent segregant pools for allocation of markers to a specific chromosomal region in outcrossing species.	16
	3. Chromosome 1 specific library	28
	4. PCR genotyping on milk samples	30
	5. Early PCR amplification test for identifying chimerism in female calves co-twin to a male calf.	32

6. Determining relative microsatellite allele frequencies in pooled milk samples in the presence of shadow bands.	34
VII. Evaluation of research achievements	46
VIII Publications	47

III Abstract

The US laboratory provided somatic cell genetic analysis of new genetic markers developed in both the US and Israel to place them on chromosomes prior to linkage analysis. In addition, we contributed our reference families to an international reference panel that is now being used globally.

More than 200 markers consisting of RFLPs and microsatellites have been placed on this international linkage map, recently published with both Principal Investigators Womack and Soller among the co-authors (Barendse, et al., 1994). The families are the bovine equivalent of the human C.E.P.H. families, distributed worldwide for testing the largest possible number of markers in the same meiotic products. The map covers 85% of the cattle genome and is a landmark accomplishment in cattle genetics.

The US laboratory has concentrated on techniques to place conserved gene sequences (type I markers) on both the synteny linkage maps. We have been successful in mapping single strand conformational polymorphisms (SSCPs), several of which appear in the published linkage map. These are important reference loci for comparative mapping, providing the potential for extrapolating information from the human and mouse maps directly to cattle. While type II markers (most microsatellites) are extremely valuable for rapidly developing in linkage map, they are not particularly useful for comparative mapping. Without type I markers on the linkage map, the cattle map cannot benefit from the wealth of data coming from genomic analysis of other mammals.

IV. Objectives of Proposal

1. To extend the bovine gene map, with particular emphasis on uncovering locus-specific VNTR probes and mapping them in hybrid somatic cells.
2. To enlarge the physical map of the bovine genome by synteny mapping, particularly with respect to polymorphic markers and cloned genes of known function, and to relate the bovine synteny groups to the human gene map.
3. To establish reference families and initiate studies aimed at mapping DNA-level genetic markers with respect to one another.

V. Relevant Data, Methodology, Results and Discussion

A. Data

All data generated in the U.S. laboratory has been published and is attached to this report in the form of reprints. The various tables and figures will not be repeated in this section. The results and a discussion of their significance are, however, summarized by topic below.

B. Methodology

Two key methodologies were utilized throughout this project:

Somatic cell genetics

Bovine leukocytes were fused with hypoxanthine phosphoribosyltransferase-deficient rodent CHO, E= 36, and thymidine kinase- (LMTK⁻) cell lines, and the resulting hybrid cell lines have been described previously. Genomic DNA was extracted from bovine, rodent, and bovine-rodent cells according to established protocols.

Plasmids containing probes of interest were transformed and isolated according to standard procedures. The plasmid inserts were isolated by digestion with the appropriate restriction enzymes and purified from agarose gel by electroelution. All cDNA probes were labeled with ³²p-dCTP to specific activities of 10⁸ cpm/μg by the random primer method.

Southern hybridization. The hybrid somatic cell DNAs were prehybridized at 42°C for 2-4 h in 5 X Standard sodium citrate, 10 X Denhardt's 0.05 M Na₃PO₄ (pH 6.7), 500 μg/ml sheared salmon sperm DNA, 5% dextran sulfate, 0.1% sodium dodecyl sulfate (SDS) and 50% formamide. Hybridizations were at 42°C for 18-20 h in 5 X SSC, 1 X Denhardt's, 0.02 M phosphate buffer (pH 7.0), 100 μg/ml sheared salmon sperm DNA, 10% dextran sulfate, 0.5% SDS and 50% formamide. Final wash was to stringency of 2 X SSC at 65°C for all probes. Filters were placed against Kodak XAR-5 film with one Dupont Cronex Lighting Plus intensifying screen at 70°C for 1-5 days.

PCR primers were designed from published bovine sequences and synthesized in the Department of Biology (Texas A&M University). A PCR cocktail consisting of 1.0 μM upper and lower primers and 0.2 mM dNTPs, 10 mM Tris pH 9.0, 50 mM KCl, 1.5 mM MgCl₂, 0.01% gelatin (w/v), 0.1% Triton X-100, and 1-1.25 units of *Taq* polymerase (Promega, Madison, Wis.). The cocktail was aliquoted into tubes with 100 ng of bovine CHO, E36, LMTK⁻, or hybrid cell DNA per reaction and overlaid with approximately 100 μl of sterile mineral oil. The reaction cycled for 1 min at 94°C, 2 min at an optimized annealing temperature, and 2 min at 72°C for 25-35 cycles in a Perkin-Elmer Cetus thermocycler. The primer pairs were optimized for amplification at the highest temperature possible. This temperature allowed preferential amplification of the bovine target sequence with little or no amplification of the rodent sequence. The subsequent reaction products were electrophoresed on a 3% Nuseive, 1% Seakem Me (FMC, Rockland, Me.) agarose gel containing ethidium bromide and were photographed.

Presence or absence of bovine-specific bands on Southern blots or PCR amplification product was tested in 20 clones from a well-characterized, bovine-rodent somatic cell panel. Segregation profiles of bovine-specific products were compared with those of previously assigned markers for each of the 31 bovine syntenic groups. Statistical analysis based on correlation was performed between the tested locus and each of the representative markers ($p < 0.02$).

In Situ Hybridization

Fibroblast cell lines established from skin biopsies of domestic cattle (*Bos taurus* L.) were grown to confluence in DMEM medium supplemented with 10% fetal bovine serum, 0.1 mM nonessential amino acids, 0.1 mM MEM vitamins, 2 mM L-glutamine, 1 mM sodium pyruvate, 100 units penicillin, 100 mg streptomycin, and 2.5 µg fungizone. Confluent fibroblasts from five T25 flasks were subcultured into five T75 flasks, and 5-bromodeoxyuridine was added to a final concentration of 45 µg/ml. Approximately 24 h later, cells were washed twice with Hank's balanced salt solution, and fresh thymidine-supplemented medium (10^{-5} M) was added for 6 h. Mitotic cells were shaken off into prewarmed (37°C) 0.075 M KCl for 25 min, and collected and fixed in several changes of methanol acetic acid (3:1). Cell suspensions were kept at 4°C until use, and chromosome preparations were made in standard fashion. Slides were air dried overnight and aged for 1.5 h at 55°C prior to hybridization.

Biotin-11 dUTP was incorporated into the plasmids of interest by nick translation with Oncor® Non-Isotopic Probe Labeling Kit S4089. Biotinylated plasmid was separated from unincorporated nucleotides with a Sephadex G-50 spin column followed by ethanol precipitation. BRL Blu-GENE Nonradioactive Nucleic Acid Detection System 8279SA was used to detect 0.1-, 1-, 10-, and 100-pg quantities of labeled plasmid dotted onto nitrocellulose filters. Biotinylated plasmid was used in FISH experiments if 1 pg could be detected.

In situ hybridization and fluorescein detection were carried out using Oncor® Chromosome in situ Kit S1370CF. Slides were hybridized with 15 ng of biotinylated plasmid in 30 µl of hybridization solution (Oncor® Hybrisol VI) under sealed coverslips for 16 h at 37°C, and then washed in 50% formamide, 2 X SSC at 43°C for 30 min, with four additional washes in 2 X SSC (43°C). Hybridized probe was detected following three rounds of avidin-fluorescein isothiocyanate (FITC) and anti-avidin antibody amplification.

Post-hybridization QFH-banding was induced by adding Hoechst 33258 to Oncor® antifade mountant at a final concentration of 1 µg/ml. Chromosomes were photographed with Kodak Technical Pan film 2415 (ASA 1280-1600), with an Olympus Vanox-T microscope equipped for epifluorescence (Olympus® filter set U: excitation 300-400 nm). FITC hybridization signals on Hoechst 33258-counterstained chromosomes were photographed at ASA 400 with Olympus filter set B (excitation 375-490 nm). Film was processed in Kodak Dektol developer for 3 min, washed in several changes of deionized water, and placed in Kodak Fixer for 4 min. Chromosomes were printed on Kodak RC III Polycontrast paper at a contrast level of 3.

QFH-banded chromosomes were compared with QFH-banded karyotypes arranged according to the domestic cow standard.

C. Results and Discussion

1. SOMATIC CELL MAPPING OF GENES OF THE BOVINE IMMUNE SYSTEM

T-cell receptor CD3 and CD8

T-cell receptors (Tcr) and CD3 molecules are part of the supergene family of immunoglobulin-like molecules which share structural homology. This supergene family also includes CD8 and Thy-1 antigens, and major histocompatibility genes. The Tcr is associated with CD3 in a multisubunit complex, the Tcr component recognizing distinct antigens in association with the major histocompatibility complex (MHC) and the CD3 components being involved in signal transduction. Tcr α and β chains, linked by a disulfide-bridge, are expressed on the majority of mature T cells of adult animals, while the Tcr γ/δ heterodimer is found on the surface of some peripheral effector T lymphocytes with specificity for MHC. In addition to the two polymorphic heterodimers, there are at least five nonpolymorphic chains in the Tcr-CD3 complex. These are CD3- γ , CD3- δ , CD3- ϵ , and two ζ chains (or ζ and η chains).

The genes encoding *a*, *b*, *g*, and *d* subunits of the Tcr, *CD3G*, *CD3D*, and *CD3E* have been mapped to human and mouse chromosomes. *TCRA* is located on human chromosome 14q11.2 and on mouse chromosome 14 near the gene for purine nucleoside phosphorylase. *TCRB* is located on human chromosome 7q35 and on mouse chromosome 6, *TCRG* is also located on human chromosome 7 sublocalized to q15, but is not syntenic with *Tcrb* in mouse where it is located on chromosome 13. The *TCRD* gene is located on human chromosome 14 overlapped with the *Tcra* gene. The mouse *d* locus, located on chromosome 14, is also overlapped with the *Tcra* gene.

The genes for *CD3G*, *CD3D*, and *CD3E* are linked and found on human chromosome 11q23 and on mouse chromosome 9. The primary structures of *CD3G* and *CD3D* genes show a high degree of sequence similarity and are separated by 1.4-1.6 kilobases (kb) of DNA. The *CD8* gene has been mapped to human chromosome 2 and mouse chromosome 6.

Little is known of the genomic locations of the *Tcr* and *CD3* genes in domestic animals. Herein, we describe the use of bovine cDNA probes to map *Tcr*, *CD3*, and *CD8* genes to bovine syntenic groups. We also provide additional insights into the evolution of the human, mouse, and bovine chromosomes on which these markers are located.

Bovine genes encoding T-cell receptor, CD3, and CD8 molecules have been mapped to syntenic groups using bovine x rodent hybrid somatic cells. T-cell receptor α and δ chains were assigned to bovine syntenic group U5, and the β and γ genes were syntenic with each other and with markers on U13. *CD3E* and *CD3D* genes were syntenic with each other and located to bovine syntenic group U19. *CD8* was most concordant with markers of syntenic group U16, although the concordancy was only 85% and the assignment must be regarded as tentative. The comparative gene maps of human chromosome 7, bovine syntenic group U13, and mouse chromosomes 6 and 13 suggest extensive evolutionary conservation.

Immunoglobulin Genes

Antibodies, which are produced by mature B-cells after stimulation by an antigen, are an integral part of the mammalian immune system. The basic mammalian antibody structural unit consists of four chains: two heavy chains, and two light chains. Each chain is additionally divided into two regions; an N-terminal variable (V) region, which is responsible for antibody binding, and a constant (C) region that determines antibody effector function. Antibodies are coded by immunoglobulin genes, which, through DNA recombination provide the antibody diversity necessary to efficiently combat the complexity of environmental antigens.

The immunoglobulin (Ig) heavy chain genes code for five different classes of antibodies in mammals, IgM, IgD, IgG, IgA, and IgE, which are coded by the genes *IGHM*, *IGHD*, *IGHG*, *IGHA*, and *IHGE*, respectively. While the immunoglobulins of mouse, human, and other species have been well characterized as to structure, organization, and expression, those of cattle have not been subject to such intense molecular scrutiny.

The immunoglobulin genes have not been genetically characterized as thoroughly in cattle as in other mammals, particularly humans and mice. Comparative genes mapping in mammals suggests that the bovine immunoglobulin heavy chain genes, *IGHG4* and *IGHM* might be syntenic with the *FOS* oncogene. Interestingly, however, when these genes are assigned to bovine syntenic groups utilizing a panel of bovine: hamster hybrid somatic cells, *IGH* genes were shown to be syntenic with the *FES* oncogene rather than *FOS*. In this study *IGH* and *FES* were assigned to *Bos taurus* chromosomal 21 while *FOS* was assigned to chromosome 10. In addition, bovine-specific immunoglobulin-like sequences were observed in the hybrid somatic cells, and one, *IGHML1* was mapped to bovine syntenic groups U16. The probes used for somatic-cell mapping were also used to screen a small number of cattle of several different breeds for restriction fragment length polymorphisms. *IGHG4* and *IGHM* were shown to be highly polymorphic, while *FOS* and *FES* were not.

2. SOMATIC CELL MAPPING AND IN SITU MAPPING OF THE BOVINE URIDINE MONOPHOSPHATE SYNTHASE GENE

Uridine monophosphate synthase (UMPS) catalyzes the final two steps of de novo pyrimidine synthesis, converting orotic acid to uridine 5' monophosphate. In Holstein and Red Holstein cattle, deficiency of UMPS (DUMPS) is inherited as a monogenic autosomal recessive trait. Heterozygous animals are phenotypically normal but have decreased UMPS activity in a number of tissues, whereas the homozygous genotype is lethal in utero. The cDNA for UMPS has recently been cloned in cattle. A C → T point mutation in codon 405 results in a premature stop codon in DUMPS cattle, providing a DNA test for the detection of heterozygous bulls prior to entry into artificial insemination programs.

UMPS has previously been mapped to human Chromosome (Chr) 3q13, Chinese hamster Chr 4p2, and sheep Chr 1q. Using somatic cell hybrid analysis and fluorescence in situ hybridization (FISH), we have assigned UMPS to cattle syntenic group U10 and sublocalized the gene to bovine Chr (BTA) 1 band 31. This represents the first assignment of a U10 marker to cattle chromosomes.

Specific hybridization between a 1.5 kb bovine UMPS probe and a 4.3-kb band in the bovine control and 11 of 20 hybrid somatic cell DNAs was observed following Southern blotting and autoradiography of hybrid cell and control DNA. The segregation of UMPS was 100% concordant with SOD1, the marker for bovine syntenic group U10. The assignment of UMPS to U10 further extends the proximal border of homology between U10 and HSA 3q because UMPS has been localized to HSA 3q13.

A biotinylated BUMPSg2 clone produced bright fluorescent signals that were symmetrically located on paired chromatids following hybridization to bovine metaphase chromosomes. The intensity of fluorescent signals and the symmetric location on both chromatids allowed hybridization signals to be discriminated from nonspecific background. No secondary sites of hybridization were identified. Post-hybridization QFH banding localized UMPS to bovine Chr (BTA) 1 band 31.

This assignment is significant in that U10 was previously unassigned to a bovine chromosome and thus tentatively permits other U10 marker (including APP, C9, CBS, CD18, COL6A1, COL6A2, CP, CRYA1, ETS2, FIM3, GAP43, GART, IFNAR, IFNBR, PFKL, PRGS, S100B, SI, SOD1, SMPP, SST, and TF) to be placed on this chromosome.

TaqI restriction fragment length polymorphisms (RFLPs) for UMPS have been reported in cattle. Genetic linkage between UMPS and other DNA markers in the Cattle Genome Database has produced a preliminary genetic map for 12 loci on BTA1 covering 168 cM and anchored the genetic and physical maps for this chromosome.

3. GENETIC CHARACTERIZATION AND MAPPING OF THE INTERFERON GENE FAMILY IN CATTLE

Multiple, superimposed Type I interferon (IFN) restriction fragments were resolved following 72-92h of horizontal electrophoresis. Restriction fragment length polymorphisms (RFLPs) for α IFN (*IFNA*) β IFN (*IFNB*), ω IFN (*IFNW*) and trophoblast IFN (*IFNT*) genes were identified in *HindIII* *EcoRI* and *TaqI* digestions from 313 cattle. RFLPs with codominant segregation in cattle pedigrees were considered alleles, and 19 distinct polymorphic Type I IFN loci (5 *IFNA*, 4 *IFNB*, 8 *IFNW* and 2 *IFNT*) were identified. Allele frequencies and observed heterozygosity values were calculated for each locus and several loci were considered highly informative for linkage analysis. Bovine IFN gene numbers (10 *IFNA*, 6 *IFNB*, 20 *IFNW* and 6 *IFNT*) were estimated from the number of polymorphic loci plus additional monomorphic hybridizing bands present in *EcoRI* and *HindIII* digestions. Physical linkage of the Type I IFN gene families on bovine chromosome 8 was demonstrated by pulsed field gel electrophoresis (PFGE). Hybridization of two or more IFN probes to similarly sized PFGE fragments suggested the tentative gene family order: *IFNA/IFNW-IFNT-IFNB*. These studies provide a basis for the development of more detailed genetic and physical maps of the bovine Type I IFNS.

The previous assignment of bovine α -(*IFNA*) and β -(*IFNB*) interferon gene families to syntenic group U18 was confirmed with additional cDNA probes and a

bovine-rodent hybrid somatic cell panel representing all 29 bovine autosomal syntenic groups. Fluorescent in situ hybridization (FISH) localized these genes to bovine Chromosome (Chr) 8 band 15 and demonstrates that with biotinylated plasmids, as few as five tandemly arrayed sequences can be detected by conventional fluorescent microscopy. This technique can be applied to physical mapping of other multicopy genes in domestic animals.

Bovine α - (IFNA), β - (IFNB), ω - (IFNW), trophoblast (IFNT), and γ - (IFNG) interferon (IFN) cDNA probes were hybridized to bovine and ovine DNA samples digested with a panel of 7 restriction enzymes to characterize the hybridization patterns and genetic variability of IFN genes in these two species. Sixty percent of probe-enzyme combinations identified RFLPs in cattle compared to 46% of the combinations in sheep. The presence of polymorphic bands in each probe-enzyme combination was compared between species. Bovine IFNA and IFNW RFLPs were observed in cattle with all 7 enzymes while ovine IFNA and IFNW RFLPs were found with 5 and 2 enzymes, respectively. IFNG RFLPs were observed with *Hind* III (61% of cattle screened), *Taq*I (28% of sheep) and *Pst*I (56% of sheep). IFN hybridization profiles were more complex in sheep, but RFLPs were more common in cattle. The most useful enzyme to screen for IFN RFLPs was found to *Hind* III in cattle and *Taq*I in sheep.

4. MAPPING OF OTHER TYPE I ANCHOR LOCI IN CATTLE

Visual Transduction Genes

To establish syntenic relationships of phototransduction genes, we have mapped the genes encoding the α -, β -, and γ -subunits of rod cGMP phosphodiesterase (PDE) (PDEA, PDEB, PDEG), the α -subunit of cone PDE (PDEA2), and the rod cGMP-gated channel (CNCG) to bovine syntenic groups. The rod cGMP PDE α -, β -, and γ -subunit genes map to bovine syntenic groups U22, U15 (chromosome 6), and U21 (chromosome 19), respectively. The rod cGMP-gated channel gene also maps to syntenic group U15, and the bovine cone α -subunit gene maps to U26 (chromosome 26). With the exception of the cone PDE α -subunit gene, which has not been mapped in other mammals, all of these genes have been assigned to conserved chromosomal regions shared among bovine, human, and mouse.

RASA

The bovine gene for the p21^{ras} protein activator (*RASA*) includes in its 5' untranslated region a (TG)_n repeat. Analysis of this (TG)_n repeat by PCR amplification of genomic DNA revealed a four-allele polymorphism. A cDNA probe was used to assign *RASA* to the region 2.4-qter of bovine Chromosome (Chr) 7 by in situ hybridization. PCR analysis in our a panel of somatic hybrid lines allowed the assignment of *RASA* to the unassigned syntenic group 22 (U22) and thus localizes U22 on Chr. 7.

Uroplakins

The asymmetric unit membrane (AUM) of the apical surface of a mammalian urinary bladder epithelium contains several major integral membrane proteins,

including uroplakins IA and IB (both 27 kDa), II (15 kDa), and III (47 kDa). These proteins are synthesized only in terminally differentiated bladder epithelial cells. They are encoded by separate genes and, except for uroplakins IA and IB, appear to be unrelated in their amino acid sequences. The genes encoding these uroplakins were mapped to chromosomes of cattle through their segregation in our panel of bovine x rodent somatic cell hybrids. Genes for uroplakins IA, IB, and II were mapped to bovine (BTA) Chromosomes (Chrs) 18 (*UPK1A*), 1 (*UPK1B*), and 15 (*UPK2*), respectively. Two bovine genomic DNA sequences reactive with a uroplakin III cDNA probe were identified and mapped to BTA 6 (*UPK3A*) and 5 (*UPK3B*). We also collaborated in the mapping of genes for uroplakins IA and II in mice, to the proximal regions of mouse Chr 7 (*Upk1a*) and 9 (*Upk2*), respectively, by analyzing the inheritance of restriction fragment length variants in recombinant inbred mouse strains. These assignments are consistent with linkage relationships known to be conserved between cattle and mice. The mouse genes for uroplakins IB and III were not mapped because the mouse genomic DNA fragments reactive with each probe were invariant among the inbred strains tested. Although the stoichiometry of AUM proteins is nearly constant, the fact that the uroplakin genes are unlinked indicates that their expression must be independently regulated. Our results also suggest likely positions for two human uroplakin genes and should facilitate further analysis of their possible involvement in disease.

Prion protein

Brains affected by the progressive neurological disease bovine spongiform encephalopathy (BSE) contain scrapie-associated fibrils and the protease-resistant isoform of prion protein. The gene encoding the normal host prion protein (PRNP) has been mapped to human chromosome 20 and mouse chromosome 2 with the hamster cDNA probe PEA974. Using this probe and our panel of bovine-rodent hybrid somatic cells, we mapped PRNP to bovine syntenic group U11 (100% concordancy). PRNP restriction fragment length polymorphisms (RFLPs) were detected with five of six enzymes (*Bgl*II, *Eco*RI, *Hind*III, *Msp*I and *Taq*I) in sheep, in contrast to one of 16 enzymes (*Hinc*II) in cattle. Codominant segregation of the bovine *Hinc*II RFLP was demonstrated in six backcross pedigrees. While PRNP RFLPs are tightly linked to scrapie incubation period, and consequently susceptibility or resistance to disease in rodents and sheep, the relationship between the PRNP RFLPs and BSE incubation period in cattle has not been determined.

5. PCR BASED ASSIGNMENTS OF TYPE I LOCI TO THE PHYSICAL MAP

Syntenic comparisons among distantly related species suggest that large chromosomal regions are conserved with mammals. Because of this chromosomal conservation, it is possible to use the extensive gene maps of human and mouse to speed the development of gene maps in other species. Inferences can be made about information derived from the mapping of loci in one species to their location in another. The development of the domestic cow gene map has benefited from a comparative mapping approach. If the borders of the evolutionarily conserved regions can be determined, then mapping information from mouse and human will be immediately useful in the bovine gene map. Likewise, syntenic information from cattle can be used to make predictions regarding syntenic organization in human and mouse.

The syntenic characterization of cattle has been primarily through Southern analysis of bovine-rodent DNAs. Recently, PCR has been used extensively in the development of physical and linkage maps of mouse and man. Here we demonstrate its value as a tool in bovine gene mapping.

The polymerase chain reaction (PCR) has been combined with hybrid somatic cell technology to extend the bovine physical map. Eight bovine loci-glycoprotein hormone alpha (*CGA*), coagulation factor X (*F10*), chromogranin A (*CHGA*), low-density lipoprotein receptor (*LLR*), human prochymosin pseudogene (*CYM*), oxytocin (*OXT*), arginine-vasopressin (*ARVP*), and cytochrome oxidase c subunit IV pseudogene (*COXP*)-were assigned to bovine syntenic groups with this approach. *CCA* was assigned to bovine syntenic group U2, *F10* to U27, *CHGA* to U4 [bovine Chromosome (Chr) 21], *LDLR* to U22, *CYM* to U6, *OXT* and *ARVP* to U11, and *COXP* to U3 (bovine Chr 5). Seven of these genes, *CGA*, *F10*, *CHGA*, *LDLR*, *OXT*, *ARVP*, and *CYM*, further delineate regions of chromosomal conservation on human Chrs 6, 13, 14, 19, 20, and 1, respectively. *CHGA*, *OXT*, and *ARVP* are unmapped in the mouse. Comparative mapping predicts the mouse *CHGA* will map to Chr 12, and mouse *OXT* and *ARVP* will map to mouse Chr 2. Furthermore, human *CYM* is predicted to be sublocalized to 1p32-q21. The primers developed for these eight loci will be useful for the development of hybrid somatic cell panels in the future as well as establishing a collection of bovine expressed sequence tags.

6. MAPPING MICROSATELLITES

A partial plasmid library with bovine genomic inserts of about 500 basepairs was screened with a (dC-dA)_n (dG-dT)_n oligonucleotide probe for the repeated nucleotide motif (CA)_n. Eleven positive clones (0.3% of all colonies screened) were discovered and were subsequently isolated and sequenced. Eight microsatellite loci were analysed, one with eight alleles, one with seven alleles, three with six alleles, one with three alleles and two with two alleles. Six of these microsatellites were mapped by PCR-analysis in our panel of somatic hybrid lines.

7. A GENETIC LINKAGE MAP OF THE BOVINE GENOME

A cattle genetic linkage map was constructed which marks about 90% of the expected length of the cattle genome. Over 200 DNA polymorphisms were genotyped in cattle families which comprise 295 individuals in full sibling pedigrees. One hundred and seventy-one loci were found linked to one other locus. Twenty-nine of the 30 chromosome pairs were represented by at least one of the 36 linkage groups. Less than a 50 cM difference was found in the male and female genetic maps. The conserved loci on this map show as many differences in gene order compared to human as if found between humans and mice. The conservation is consistent with the patterns of karyotypic evolution found in the rodents, primates and artiodactyls. This map will be important for localizing quantitative trait loci and provides a basis for further mapping.

This international effort will continue. Both US and Israel investigators will continue to contribute to the growth and development of this primary linkage map of the bovine genome.

BARD project - US-1687-89RC

Mapping the Bovine Genome Using DNA Level Markers

Work of the Israel group

The objectives of the Israel portion of the BARD research program were to contribute to development of the bovine genome map, and to the development of methods for useful application of the bovine genome map to problems in livestock production. In the context of bovine genome mapping we have: (1) Developed and mapped a group of new microsatellite markers, (2) Developed a procedure ("Single Sire Cosegregant Analysis") for the screening of new markers, directed to any chosen chromosomal region, and (3) Prepared a bovine chromosome-1 specific library. In the context of map applications, we have: (4) Shown that microsatellite markers can be genotyped on milk samples; a procedure which greatly reduces the costs of sample collection in dairy cattle, (5) Shown that freemartin and fertility status of females born co-twin to a male can be determined by genotyping milk or hair-root follicles, using the PCR reaction and Y-chromosome specific primers; a procedure which can rescue about 10% of co-twins for productive life as dairy cows, and (6) Developed a model for estimation of microsatellite allele frequencies in pooled DNA or milk samples, in the presence of shadow bands; a procedure which can allow very efficient determination of linkage between microsatellite markers and economic trait loci. Each of these studies will now be described in detail.

1. Development Of Microsatellite Markers

The purpose of this research was to prepare a group of microsatellite markers, as our contribution to development of a bovine genome map based primarily on these markers.

METHODS AND RESULTS

In the first stage of the research the GenBank data repository was searched for known bovine sequences that included the poly(TG) microsatellite. Five sequences were chosen, and primer sequences for PCR amplification were synthesized. These sequences served primarily for development of laboratory methodologies, including a method for ^{35}S end-labelling of the primers which provided sharp definition of alleles and safer work conditions. Length polymorphism of microsatellite alleles was determined by denaturing acrylamide gel electrophoresis of PCR products. Although a number of polymorphisms were uncovered, these were not followed through for mapping, since the total number of GenBank microsatellites was small, and other laboratories involved in the bovine genome mapping endeavor were also exploiting this source of microsatellites.

Simultaneously with the GenBank work, a group of anonymous markers, not associated with known genes were developed. For this purpose two short-sequence bovine genome libraries were developed (200-600 bp and 800-2000 bp) in pBS plasmid. Including both libraries a total of 52,000 colonies were screened using ^{32}P labelled (TG)_n oligonucleotide. Positive signals were obtained from 250 colonies. Of these, 160 colonies were lifted and examined further for size of insert and presence of poly(TG) tracts. 102 of the shorter inserts (<1200 bp) containing poly(TG) tracts were taken for sequencing. Of these, 42 contained a poly(TG) tract with $n > 12$. Nine of these inserts were found to contain flanking SINE sequences and were not examined further. 24 microsatellites were chosen for amplification. Of these, 9 did not give PCR products.

The remaining 15 microsatellites (termed: "HUJ microsatellites") were examined for polymorphism on the parents of the international panel of bovine reference families. With two exceptions (one diallelic and one triallelic), all of the HUJ microsatellites were polyallelic, showing four to ten alleles. Heterozygosity in the parents of the reference families was greater than 50% in all cases, with even the triallelic microsatellite showing a heterozygosity of 69%.

The 15 HUJ microsatellites described above, and two microsatellites from the GenBank screening were mapped to synteny groups using a panel of 27 rodent/bovine somatic-hybrid cell lines. The panel was developed and provided by Prof. Jim Womack, the USA collaborator in this study. Syntenic mapping was carried out by PCR amplification of DNA from each of the panel lines using the primers specific to each microsatellite. Presence of PCR product of the proper size, was scored as a positive reaction. The results for each microsatellite were compared to the standard pattern defining each of the syntenic groups and the microsatellites were assigned to the syntenic group with which they showed the highest concordance. In this way, 15 of the HUJ microsatellites were mapped to 12 syntenic groups.

Fourteen of the HUI microsatellites (not including the diallelic microsatellite) were mapped by segregation linkage analysis using the international bovine mapping reference family panel. This panel includes four large half-sib families established at CSIRO (Australia), one large full-sib family and one large half-sib family from Texas A & M, and two large full-sib families from ILRAD (Kenya). In all, the reference families included 295 parents and offspring. Results of individual genotyping of parents and offspring for the various HUI microsatellites were sent to CSIRO for incorporation in the Cattle Genotypic Database (CGD) and were examined for linkage associations by the CRIMAP program. Most markers were readily mapped to one locus at least in the same synteny group, and in accordance to the synteny mapping results. Two markers, HUII77 and HUI246, turned out to be linked to one another. HUII77 had been assigned to U6 by synteny mapping with a concordance score of 92%, while HUI246 had 100% concordance to the same synteny group. Linkage between the two markers provided confirmation of HUII77 syntenic assignment to U6.

The syntenic assignment of HUI175 was doubtful: amplification in one of the somatic cell hybrid lines was unclear, and as a result HUI175 could have been assigned to either of two alternative synteny groups. Genetic linkage of HUI175 to the mapped CSIRO marker CSSM6 provided the relevant data for assigning HUI175 to synteny group U12. Assignment of HUI113 to the synteny group U25, together with its linkage to CSSM36 provided more evidence to the reality of the controversial synteny group U25. HUI121 was obviously X-linked since all males showed a single allele, and only females were heterozygous. Nevertheless, this marker could not be shown to belong to the defined X synteny group on either of the two somatic hybrid panels (from Texas A&M and INRA). However, on linkage mapping it proved indeed to be linked to the sex-linked GenMark marker TGLA128.

On the other hand, HUI625, very well defined as belonging to synteny group U1, did not show linkage to any other marker in the CGD, when examined during the initial steps of establishment of the linkage map. This is attributed to lack of saturation of the map, and indeed HUI625 was later shown to be linked to additional markers on chromosome 16.

DISCUSSION

The full CGD map of the bovine genome has recently been published (Barendse et al., 1994). On this map, for the first time, all synteny groups have been assigned to their chromosomes, and mapped markers/genes are present on all of the 30 chromosome pairs of the bovine genome. In this map, eleven of the markers studied here are mapped to 10 different chromosomes, namely chromosomes 3,4,8,11,13,16,17,22,27, and X. Three markers were mapped at a later stage (Table 1.).

HUII77 and HUI246, two linked markers were mapped in the 71 map unit interval between TGLA263 and RM38, on chromosome 3, contributing to refining this region of the map and to linking together both ends of the chromosome. The markers HUI673, HUI614 and HUI113 have terminal locations on chromosome 4, 16, and 27, respectively, adding 11, 32 and 11 map units to their respective chromosomes. HUI121 is one of only two markers that were mapped to the sex-specific region of chromosome X.

Table 1. GenBank accession numbers and map location of the HUU microsatellites. The map location is indicated as chromosomal, or syntenic assignment, according to data available. When the markers were assigned a "D-number", these are indicated.

Marker	GenBank	Map location
HUU614	M93653	D16S8
HUU616	M88250	D13S6
HUU625	M93004	Chr16
HUU673	G00407	Chr4
HUU121	M97201	ChrX
HUU174	G00404	Chr8
HUU223	M93647	D17S5
HUU225	G00405	U22
HUU246	M93645	D3S3
HUUI13	M96346	Chr27
HUUI29	G00406	U26
HUUI75	M96347	D11S30
HUUI77	M96348	D3S4
HUUVI74	M96344	Chr11

2. Single-Parent Segregant Pools for Allocation of Markers to a Specified Chromosomal Region in Outcrossing Species

Once economic trait loci (ETL) have been located to a particular chromosomal region, fine mapping of these loci, will require saturation mapping of the chromosomal region. Thus, methods for rapidly screening and developing anonymous markers for specific chromosomal regions are required. Such methods can also serve to fill in any remaining gaps in the primary species genome maps. Bulked cosegregant analysis is a method, developed for utilization in F₂ mapping-populations derived from crosses between fully inbred lines, for rapidly allocating unmapped genetic markers to a specific chromosomal region. In this procedure the individuals of the mapping population are separated into two groups, homozygous for alternative parental genotypes, at an index marker, known to be located in the chromosomal region of interest. DNA samples of the individuals in each marker genotype group are then pooled. The two pooled samples will each be homozygous for the alternative homologous parental chromosomal regions at the index marker; and will be close to homozygosity for alternative parental chromosomal regions in the near vicinity of the index marker. Consequently, any marker closely linked to the index marker, and for which the parental lines differ, will be at very different frequencies in the two pools. These differences in frequency can be detected by genotyping the two pools. Furthermore, by applying RAPD techniques to the alternative pools, polymorphic bands located in the same chromosomal region can readily be detected. These can then be cloned and converted into locus specific markers. Thus, cosegregant analysis can serve, not only to allocate existing markers to a given region, but also to develop new markers in the given region.

With appropriate modification cosegregant pools could also serve as a useful tool in outbreeding populations as well, if individuals from only a single family are pooled (Dunnington et al., 1990; Micheltmore et al., 1991). In this study we derive power and tests of significance for single-parent cosegregant analysis in large full-sib, single-sire backcross, and single-sire half-sib families. Large single-sire backcross and half-sib families are presently available as part of the international chicken and bovine reference family panels, respectively. We illustrate the utilization of a large single-sire bovine family, part of the international bovine reference family panel, for single-sire cosegregant analysis, using known index and linked microsatellite loci.

THEORY

A marker, heterozygous in a given individual, can serve as an index marker, by which the progeny of the chosen individual can be divided into two groups, according to the marker allele received from their common parent. Segregant pools for the index marker, are formed by separately pooling DNA samples of the progeny in each of the groups. Half of the chromosomes in each segregant pool will derive from the common parent. With respect to these chromosomes, the two pools will differ systematically for the chromosomal regions at and flanking the index marker. All other common-parent chromosomes will be distributed at random between the two segregant pools. Chromosomes, derived from the mate(s) of the common parent, including those carrying the index marker, will be distributed at random between the two segregant pools.

Given an unmapped marker, then, which is also heterozygous in the common parent, the expected difference in densitometric frequency of allele 1 of the test marker, in segregant pool 1 (f_{11}) and densitometric frequency of test marker allele 1 in segregant pool 2 (f_{12}), will be

$$D = f_{11} - f_{12}$$

f_{11} is calculated from the densitometric density of allele 1 in pool 1, relative to the summed densities of alleles 1 and 2 only. Any other alleles present in the pool, will have been contributed by the mates and are not included. Similarly, f_{12} will be calculated from the densitometric density of allele 1 in pool 2, relative to the summed densities of alleles 1 and 2 only.

The densitometric density of a given allele in a given pool will stand in direct proportion to the number of occurrences of that allele in the given pool. On this basis, of the $2N$ chromosomes contributed to a segregant pool by the N individuals that it comprises, N derive from the common parent, and N from the mates. The number of chromosomes carrying allele 1 derived from the common parent will have a simple binomial distribution with parameters N and r , where r is the proportion of recombination between index marker and test marker. Thus, among the N chromosomes derived from the common parent, it is expected that a proportion $(1-r)$ and r will carry allele 1, and allele 2, respectively, with variance, $VC = r(1-r)/N$.

Among the N chromosomes contributed by the mates, it is expected that s_1N and s_2N will be allele 1 and allele 2, respectively, where s_1 and s_2 are the frequencies of allele 1 and allele 2 among the mates of the common parent; while $(1-k)N$ will be of alleles other than s_1 or s_2 , where $k = s_1 + s_2$. Consequently, the number of chromosomes carrying allele 1 derived from the mates, will have a more complex distribution. This is a consequence of the fact that if there are other alleles present among the mates, in addition to alleles 1 and 2, the overall number of chromosomes carrying alleles 1 and 2, N^* , is itself a binomially distributed random variable with parameters N and k , expectation $E(N^*) = kN$, and variance $k(1-k)N$. Since N will typically be quite large (>20), this distribution will approach normality, and N^* will generally vary only slightly from its expectation. Consequently, the distribution of the number of mate-derived chromosomes carrying allele 1 contributed by the mates will closely approximate a binomial distribution with parameters kN and s_1 . Thus, among the N^* chromosomes derived from the mates, it is expected that a proportion $s_1^* = s_1/k$ and $s_2^* = s_2/k$ will carry allele 1 and allele 2, respectively, with variance, $V_M = s_1^*s_2^*/kN$.

The expectation of f_{11} will clearly equal the average of the expectations of the common parent and mate components, weighted by N and N^* , the number of chromosomes derived from each source, respectively. Thus, it can readily be shown that:

$$E(f_{11}) = [N(1-r) + kN_s 1^*] / (N + kN) = (1-r + ks_1^*) / (1+k)$$

with binomial variance

$$V_B(f_{11}) = (V_C + k^2V_M) / (1+k)^2 = [r(1-r) + ks_1^*s_2^*] / N(1+k)^2$$

where,

N is the number of individuals in a segregant pool,

r is the proportion of recombination between index marker and test marker,

s_1 and s_2 are the frequencies of allele 1 and allele 2 among the mates of the common parent,

$$k = s_1 + s_2, \text{ and}$$

$$s_1^* = s_1/k, s_2^* = s_2/k$$

Similarly, the expectation of f_{12} will be

$$E(f_{12}) = (r + ks_1^*)/(1+k)$$

with binomial variance $V_B(f_{12}) = V_B(f_{11}) = V_B$

If the mates derive from an inbred line, s_1 and s_2 will equal 0 or 1. If the mate is a single individual, as in a large full-sib family, s_1 and s_2 will equal 0, 0.5, or 1.0, as the case might be. If the mates are a group of individuals, as in a large half-sib family, s_1 and s_2 will be the frequencies of the common-parent alleles among these individuals. In this case, if alleles 1 and 2 are the only marker alleles found among the mates, then $s_2 = 1 - s_1$.

The expected difference in frequency of allele 1 in pool 1, and allele 1 in pool 2, will be

$$E(D) = E(f_{11}) - E(f_{12}) = (1-2r)/(1+k)$$

The expected error variance of this difference, V_D , will be,

$$V_D = Vf_{11} + Vf_{12}$$

where, Vf_{11} and Vf_{12} are the expected error variances in the estimates of f_{11} and f_{12} , respectively.

Both of these have the same expectation, so that

$$V_D = 2Vf_{11}$$

Vf_{11} has two components. One, denoted V_T , is due to technical error in densitometric determination of relative allele frequencies. The other, denoted V_B , is due to sampling variation of allele numbers due to Mendelian segregation and gamete sampling, as derived above.

V_T depends on experimental procedures, and will equal the variance between f_{11} estimated from replicated pools. V_T has two components. V_{TA} , derived from variation in genotyping and densitometry of the same pool; and V_{TP} , derived from variation in making up the pool. Both of these components are unbiased, so that if the densitometry is based on, say, m replicate pools, each genotyped n times, $V_T = V_{TP}/m + V_{TA}/mn$.

Taking both components into account, then, we have

$$V_D = 2Vf_{11} = 2V_T + 2V_B$$

Test for linkage

The test for linkage will be a t-test for significance of the observed difference, D , in relative frequency of one of the common-parent alleles of the unmapped marker, in the two segregant pools, relative to its standard error, $V_D^{1/2}$. The error variance, V_D is constructed on the null hypothesis ($r=0.5$, absence of linkage), so that on the simplifying assumption that $s_1^* = s_2^* = 0.5$ as well, the binomial component of f_{11} becomes,

$$V_B = 1/4(1+k)N$$

Thus, when s_1 and s_2 are the only alleles found in the mate gamete pool, $V_B = 1/8N$; when s_1 and s_2 are absent from the gamete pool, $V_B = 1/4N$; and when s_1 and s_2 make up, say, half of the mate gamete pool, $V_B = 1/6N$.

Power of the test

The power of the test, as a function of progeny numbers and assumed proportion of recombination between unmapped marker and index marker, will be $1 - \beta$, where β , the Type II error, is the area under the standardized normal curve from $-\infty$ to Z_β , and

$$Z_\beta = D/V_D^{1/2} - Z_{\alpha/2}$$

where,

α = Type I error, and

V_D is constructed according to specific hypotheses of linkage and test-marker allele frequencies in the mate gamete pool.

EXPERIMENTAL

An experimental test of the proposed procedure was carried out utilizing the large half-sib family of sire 0101 of the Texas/CSIRO/ILRAD panel, which comprises 44 progeny of this sire out of three dams.

Markers and pools

Two pairs of markers were chosen in order to test for the ability to detect linkage in single-sire segregant pools. Each pair consisted of one index marker (HUJI13 and HUJ223, respectively), and one linked test marker (CSSM36, and CSSM33, respectively). HUJI13 and CSSM36 map to chromosome 27; HUJ223 and CSSM33 map to chromosome 17. Sire 0101 was heterozygous for each of the four markers.

The pools were constructed using DNA samples of the 44 half-sib progeny of sire 0101. These animals had previously been genotyped for all four of the above markers, as part of the routine linkage mapping of these markers in the reference family panel. The animals were then divided into alternative pools, according to the sire allele at each of the two index markers. Segregant pools were formed for markers HUJI13, CSSM36 and HUJ223. Each pool was prepared in duplicate.

Genotyping and densitometry of the segregant pools

Radioactive PCR amplification was carried out on each of the pools, using the usual protocols, separately for each of the four markers. Two independent amplifications were carried out on each of the replicate pools, so that $m=n=2$. Each segregant pool was thus independently amplified four times for each marker. The PCR products were run on a denaturing polyacrylamide gel until allele resolution was satisfactory for densitometric readings.

The dry gel was exposed to an X-ray film for 2 to 3 days. The intensity of each sire allele band in each pool amplification was measured using a laser densitometer. Frequency of sire marker allele 1 in each pool was then calculated as

$$F_{11} = D_{11}/(D_{11}+D_{12})$$

$$F_{12} = D_{12}/(D_{12}+D_{22})$$

where,

D_{ij} are the actual densitometric values for the i th sire marker allele in the j th segregant pool ($i=1, 2; j=1,2$).

RESULTS

Table 1 shows anticipated power of sire cosegregant analysis, as a function of sample size, proportion of recombination between index marker and unmapped marker, test marker allele frequency among the mates, and technical error variance (after replication). For technical error variance after replication within the range obtained in this and other experiments ($V_T=0.0016$) power is essentially equal to 1.0, even for segregant pools consisting of relatively few individuals ($N=20$ or 30) (data not shown). It is only when technical error variance is very high ($V_T=.01$) that power drops below 1.0. In this case, technical error, rather than binomial sampling, is the main source of error so that power is little affected by size of the segregant pool. At this level of technical error, the binomial contribution of segregation of sire alleles from the mates

also becomes an important source of error. Nevertheless, even at this level of technical error variance, power remains very high for close linkage ($r=0.05$ or 0.10) and even for looser linkage ($r=0.20$) when sire alleles are not common among the mates ($k=0.0$ or 0.5).

Table 2 shows the genotypes of sire 0101 and of his three mates for the four markers utilized in the experimental study. In all cases at least one of the sire alleles was also found among the mates. In no instance, however, did any of the mates carry both of the sire alleles.

Table 3 shows the distribution of the two sire alleles among the segregant groups, according to linkage group and marker, separately for alleles derived from the sire, for alleles derived from the mates, and totalled over both sources. Considering alleles derived from the sire, for index marker and linked test marker sire allele 1 and sire allele 2 were invariably in great excess in their respective segregant groups. This was not found for unlinked markers. Compare, for example, sire alleles 1 and 2 in segregant group 11-1 for marker 12 (linked) and marker 21 (unlinked). In contrast, sire alleles derived from the mates were distributed in a roughly equal manner between the segregant pools for each index marker. In all cases, the strong differential distribution of sire alleles for the index and linked markers, and the roughly equal distribution of sire alleles for the unlinked markers, remained when the total of sire and dam contributions were considered.

HUJ223 and HUJ113 segregant pools were genotyped with HUJ223 and the linked marker CSSM33. Each pool was prepared in two independent replicates and each replicate was amplified twice. Although not all lanes were readable, there were two to four densitometric estimates for each marker in each segregant pool in which it was tested. Pooled over all markers and segregant pools, the variance component between duplicate amplifications, VTA, was .0004, and the variance component between different pools, VTP, was .0001.

Table 4 shows actual allele frequencies for the longer sire allele in the segregant pools, and the corresponding frequencies as estimated by densitometry. For those cases in which allele frequencies were estimated by densitometry as well, there was generally a good correspondence between differences in actual frequencies, and differences in estimated frequencies. The differences in actual frequencies of the longer sire allele in the alternative segregant pools, ranged from .32 to .78 for index or linked marker, while the differences in frequencies as estimated by densitometry ranged from 0.29 to 0.72. For unlinked markers, actual differences in allele frequencies between alternative pools ranged from .00 to .23, and estimated frequencies from .01 to .11. Taking $N=20$, $k=.5$, $s_1*s_2=.25$, and $V_r=.0002$ (assuming replication of pools and amplifications), the standard error of the difference between alternative segregant pools, as estimated by densitometry was $V_D^{.05}=0.13$. Thus, In all cases, the densitometrically estimated differences between segregant pools for index and linked markers would have been judged highly significant statistically, while those for unlinked markers would have been judged nonsignificant.

DISCUSSION

The results of this study show that both in theory and in practice, single-parent segregant pools can provide a convenient means of rapidly assigning new markers to a particular chromosomal region in outbreeding populations. This extends the field of application of this procedure to essentially all species of biological or agricultural interest, including all experimental and agricultural animals, and fruit trees and other outcrossing plants. All that is required is a large family derived on one side at least, from a single, common parent. The mates of the common-parent, can be individuals taken from an inbred line, a single individual, or some number of individuals, as required by restraints imposed by the biology or husbandry of the species.

Source of families for segregant analysis

For a given single parent, segregant pools can be formed only with respect to index markers that are heterozygous in the given parent, and will be useful only for test markers that are also heterozygous in that parent. Thus, in order to enable segregant pools to be constructed for any desired chromosomal region on the one hand, and in order for the common parent of segregant pool to also be heterozygous for the test marker, on the other, it would seem useful to have a number of large single-parent families available. When interest develops in a particular chromosomal region, the common parents of the various families would be genotyped with respect to a battery of potential index markers, known to be located in the chromosomal region of interest. Once one or more of the common parents were identified as heterozygous for index markers in the region, their progeny would be genotyped with respect to the index markers, and allocated to segregant pools. Test markers would then be genotyped on the common parents used for the segregant pools, to identify those common parents in which they were heterozygous, and would then be genotyped with quantitative densitometry, on the corresponding segregant pools.

For bovine gene mapping, the TEXAS/CSIRO/ILRAD international bovine reference family panel contains two large families, with 44 and 48 progeny respectively. These families have already been genotyped with respect to all markers mapped in this panel, and DNA samples of the individual animals in each family are widely available. Thus, they form a convenient resource for construction of single-parent segregant families for most bovine chromosomal regions. Should additional single-sire families be required, they can readily be constructed by obtaining semen or blood samples of any widely used A.I. sire, and milk samples of 50 to 100 of his daughters.

Similarly, for chicken gene mapping, the East Lansing and Compton chicken reference families can serve as convenient sources for single-sire segregant pools. Each of the families has been genotyped against a wide panel of markers, so that segregant pools for any chromosomal region can be readily produced. Each of these families is headed by a single F1 male, produced by crossing a single male (East Lansing: Red Jungle Fowl; Compton: inbred White Leghorn) to a single female of a highly inbred White Leghorn line. The reference families themselves were produced by backcrossing the F1 male to a single female from the inbred female parental line. Here too, additional families can be produced readily by crossing a single male to one or more inbred or outbred females.

Technical design considerations

Although the different types of mate (inbred line, single, multiple) have somewhat different power as a result of their effect on frequency of common-parent test-marker alleles in the mate gamete pool, the results of the theoretical analyses show that all mate options have excellent power for test markers within 10 cM of the index marker. In most cases, whether for filling in gaps in the map, or for saturating a particular map region, single-parent cosegregant analysis would be employed to detect markers within this range. Thus, the choice of mate-type can be dictated by the husbandry and biology of the species, without too much effort expended on producing the most suitable sort. Indeed, any small differences in power according to type of mate can readily be accommodated by increasing the size of the families. A special advantage of a large full-sib family, is that either of the parents can serve as common parent for forming segregant pools. This may be somewhat more convenient in some circumstances than preparing two independent single-parent families.

Technical experimental considerations

Technical error components of variance obtained in this study ($VTA=.0002$, $VTP=.0001$), were less than those that we have obtained in other studies in our laboratory. Experience with densitometric estimation of microsatellite marker allele frequencies in pooled DNA samples in our laboratory (Khatib et al., 1994; Lipkin et al., 1994), lead to estimates of error variance components in the range $VTA=.0004$, and $VTP = .0016$ to $.0036$. Based on the latter values, VT for a single amplification of a single segregant pool would equal $.0020$ to $.0040$; VT for single amplifications of duplicate pools would equal $.0010$ to $.0020$; and VT for duplicate amplification of duplicate pools would be marginally less, $.0009$ to $.0019$. However, even when a value of $VT=.002$ ($m=2$, $n=2$), in place of the $.0002$ found in this study is used, the standard error of the difference between segregant pools increases only marginally (from 0.13 to 0.15), and does not change the conclusions with respect to statistical significance of the observed results.

Single-parent segregant pools as a means of developing new markers

Michelmore et al., (1991) and Giovanonni et al. (1991) screened segregant pools for polymorphism using random amplified polymorphic DNA (RAPD) primers. Both groups were successful in identifying RAPD polymorphisms in the target chromosomal regions. In both cases a polymorphic RAPD band was detected as present in one segregant pool, and absent from the alternative segregant pool. This all-or-none behavior is to be expected with F2 segregant pools, each consisting of only a few individuals derived from a cross between inbred lines. In single-parent segregant pools, however, even when a RAPD band is in tight linkage to one of the alleles at the index marker, the same RAPD band can be expected to be also present, albeit at differing frequency, depending on its frequency in the mate gamete pool, in both of the alternative segregant pools. Information on quantitative amplification of RAPD bands is not available, so it is not clear whether in this case a difference in band frequency will be detected by densitometry. In the event that amplification is not quantitative, segregant pools formed in full-sib F2 families, may provide a situation in which there is all-or-none presence of a RAPD polymorphism closely linked to the index marker, in alternative segregant pools. Also, backcross families to an inbred line, as found for the chicken reference family panels, can also provide all-or-none separation with respect to closely linked markers in the two segregant pools, for bands that derive from the non-recurrent parent only.

Table 1. Power of single-parent cosegregant analysis, as a function of proportion of recombination between index marker and linked marker (r), proportion of common parent alleles among the mates (k), number of individuals in each segregant pool (N), and technical error variance (VT).

N	VT	r=0.05			r=0.10			r=0.20		
		k=0	0.5	1.0	k=0	0.5	1.0	k=0	0.5	1.0
20	.0016	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98
	.01	1.00	0.98	0.86	1.00	0.93	0.77	0.93	0.72	0.52
30	.0016	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
	.01	1.00	0.98	0.87	1.00	0.94	0.78	0.96	0.75	0.53
40	.01	1.00	0.98	0.87	1.00	0.95	0.79	0.97	0.76	0.54
60	.01	1.00	0.99	0.88	1.00	0.96	0.79	0.98	0.78	0.55
100	.01	1.00	0.99	0.88	1.00	0.96	0.80	0.98	0.79	0.55

Table 2. Designation and distribution among the sire and three dams of a large half-sib family, of the markers and alleles utilized for single-sire cosegregant analysis. Linkage groups are designated, 1, 2. Markers within linkage groups are designated 1, 2. Alleles found in the sire are designated 1,2. Alleles found in dams only are designated, M. Proportion of sire alleles among the dams are designated s1 and s2

Linkage group	Marker	Sire	Dam1	Dam2	Dam3	s1	s2
1	11	1,2	1,M	2,M	1,M	.33	.17
	12	1,2	1,M	2,M	1,M	.33	.17
2	21	1,2	M,M	1,M	1,M	.33	.00
	22	1,2	1,M	M,M	M,M	.17	.00

Linkage groups 1 and 2 correspond to chromosomes 27 and 17, respectively. Markers 11, 12, 21, 22 correspond to HUI13, CSSM36, HUI223 and CSSM33, respectively.

Table 3. Distribution of the two sire alleles (1 and 2), among segregant groups formed in a large half sib family, according to linkage group, marker and derivation of the allele: from the sire, from the dams and total. The number of progeny in each segregant group is shown under Total (in parentheses).

Segregant group	Linkage group 1				Linkage group 2			
	Marker 11		Marker 12		Marker 21		Marker 22	
	1	2	1	2	1	2	1	2
<u>Derived from the sire</u>								
11-1	27	0	23	4	16	11	16	11
11-2	0	17	4	13	9	8	9	8
12-1	22	4	26	0	18	8	18	8
12-2	4	14						
21-1	15	10	17	8	25	0	25	0
21-2	10	9	9	10	0	18	0	18
<u>Derived from the Dams</u>								
11-1	11	4	8	5	10	0	2	0
11-2	9	1	8	1	7	0	4	0
12-1	8	3	7	3	9	0	4	0
12-2	8	2	7	2	8	0	2	0
21-1	12	3	9	4	9	0	1	0
21-2	7	2	7	2	7	0	5	0
<u>Total</u>								
11-1(27)	38	4	31	9	26	11	18	11
11-2(17)	9	18	12	14	16	8	13	8
12-1(26)	30	7	33	3	27	8	22	8
12-2(16)	12	16	7	20	16	10	10	10
21-1(25)	27	13	26	12	34	0	26	0
21-2(19)	17	11	16	12	7	18	5	18

Each segregant group is designated according to index marker according to which the groups were formed, and sire allele within the index marker used in defining each segregant group. For example, segregant group 11-1 comprised all progeny of the half-sib family that received allele 1 from their sire, at marker 11.

Table 4. Actual frequency (A) of the longer sire allele among the segregant groups and frequency as estimated by densitometry (E) of pooled DNA samples of the segregant groups, according to relation of test marker to index marker: Index, test marker is the index marker; Linked, test marker is in same linkage group as index marker; Unlinked, test marker is in a different linkage group than the index marker. The markers in each category are shown (in parentheses). Actual and estimated difference in allele frequency of the sire allele in the two segregant pools for each index marker are also shown (Bold).

Segregant pool	Index		Linked		Unlinked			
	A	E	A	E	A	E	A	E
	11		12		21		22	
11-1	.10	.23	.22	ND	.70	.67	.38	.28
11-2	.68	.52	.54	ND	.67	.68	.38	.17
Difference	.58	.29	.32		.03	.01	.00	.11
	12		11		21		22	
12-1	.09	ND	.19	.27	.77	ND	.27	ND
12-2	.74	ND	.57	.57	.61	ND	.50	ND
Difference	.65		.46	.30	.16		.23	
	21		22		11		12	
21-1	1.00	1.00	.78	.72	.32	.31	.31	ND
21-2	.28	.22	.00	.00	.40	.35	.43	ND
Difference	.72	.72	.78	.72	.08	.04	.12	

3. Chromosome I specific library.

Chromosome specific libraries are a means of obtaining markers specific for particular chromosomal regions. We explored the possibility of constructing such a library using a rodent/bovine somatic hybrid cell line carrying a single bovine chromosome only, as a source of genomic DNA for a large insert library; and a unique bovine SINE sequence as a probe for identifying clones in the library carrying bovine chromosomal fragments.

METHODS

DNA from somatic hybrid cell line COB4.5 developed at Texas A&M and provided by Prof. J. Womack, our USA collaborator was used as a source of DNA for library construction. Hybridization patterns obtained with this line indicate that the bovine chromosomal material that it carries derive primarily from bovine chromosome 1. DNA from this line was digested to completion with the restriction enzyme BamHI. The digestion products were run in a 0.7% low-melting point agarose gel. Fragments in the 10 to 23 kb size range, representing 60 to 70% of the genome were cut out of the gel. The size-selected fragments were ligated to BamHI restricted Emb13 arms (Stratagene) using T4 DNA ligase (USAB).

The plated library was screened for the presence of (AC)_n repeats, using ³²P end-labelled (TG)₁₀ oligonucleotide probe; and for bovine DNA inserts using total bovine genomic DNA ³²P-labelled by random priming, or one or both of two different ³²P end-labelled oligonucleotide probes: ANB12 (5'CGTGGGGTCCC-AAAGAGTCG), present, with some sequence variation, in the monomer region common to all of the bovine SINE repeats; and ANB18 (5'GTCGGCAGAGTAATGTCTTTGC), with high sequence homology to the PstI SINE repeat.

RESULTS

The titer of the library was at 2.4×10^6 pfu/ml. The total size of the library was 1.2×10^6 phage particles, and was estimated to represent 9-10 genome equivalents.

Screening of 6.8×10^4 plaques, using total genomic bovine DNA as a probe, uncovered 4.8×10^3 bovine phages, representing 7% of the total library. This proportion of clones in the library hybridizing to bovine DNA is consistent with the expected proportion of cloned bovine sequences from the somatic hybrid cell line, on the assumption that the size of bovine chromosome 1 (the largest bovine chromosome) can be estimated at 2×10^8 bp (approximately 6%) of the bovine genome.

For isolation of bovine specific microsatellite clones, the library was plated twice; each time using 1.5 genome equivalents (1.6×10^5 phage). The first plating was double screened, using (TG)₁₀ for the first screening and ANB12/ANB18 simultaneously for the second screening. Cross hybridizing colonies were lifted. The second plating was screened with ANB12 only, and hybridizing colonies were lifted. The selected plaques of both platings were plated individually, and submitted to a second double screening with (TG)₁₀ and either ANB12/ANB18 (first plating) or ANB12 (second plating). From these screenings, 14 clones were isolated that hybridized both to (TG)₁₀ and to ANB12 and/or ANB18.

The selected phage clones were grown on XL1 (MRA) E. coli host, in 10 ml cultures. DNA from the phage was extracted and restricted with BamHI, BamHI/BglII, and HindIII restriction enzymes. The restriction products were run in an 0.7% agarose gel, and (TG)₁₀ in-gel hybridization was carried out. Twelve of the 14 clones presented a (TG)₁₀ hybridizing fragment. Seven of these were further examined by using them as probes in Southern analysis of total bovine and hamster DNA digested with HindIII or EcoRI restriction enzymes.

Of these seven putative bovine-insert clones, three gave specific bands on hybridization to bovine DNA blots, and did not hybridize to the hamster DNA blots. These are considered to carry bovine-specific unique-DNA inserts. Two clones gave non-specific smears with bovine DNA and did not hybridize to hamster DNA. These are considered to carry bovine specific repetitive inserts. Two clones gave non-specific smears with both bovine and hamster DNA. These are considered to carry repetitive inserts of a class found in both bovine and hamster DNA.

FUTURE PLANS

The results of the clone analysis show that the library contains bovine inserts, and that the screening method used is able to isolate bovine inserts from the library. Although we do not have specific plans to continue actively with further characterization and exploitation of this library at the present time, we have sent a copy of the library to the laboratory of Jay Hetzel at Molecular Animal Genetics Centre, CSIRO Division of Tropical Animal Production, University of Queensland, Australia. They, in turn, have lifted some clones using a double screening procedure similar to that outlined above, and have sent them to the laboratory of Barbara Schwenger at Institute für Tierzucht und Vererbungsforschung, Tierärztliche Hochschule Hannover, Germany, for microsatellite sequencing using procedures for direct sequencing of microsatellite flanking regions using microsatellite primers. If they confirm that chromosome 1 specific clones can be derived from the library, we will make the library generally available to the bovine genome mapping community.

4. PCR Genotyping on Milk Samples

In cattle peripheral blood leukocytes are the usual source of DNA for genotyping. However, obvious technical difficulties exist in collection of blood samples from large numbers of individuals among widely separated herds. This limits the application of marker-based methods for genetic analysis and for genetic improvement of economic traits in dairy cattle.

Bovine milk contains a large, albeit highly variable number of somatic cells (predominantly leukocytes), ranging from 10^4 to 10^7 /ml, depending on the status of the cow. Use of these cells would facilitate obtaining DNA samples from cows. Collection of milk is part of the daily routine and does not require trained personnel or special help. Furthermore, under some milk recording schemes, milk samples of individual cows in many herds are collected monthly and conveniently sent to a central laboratory where they are tested for protein, fat, lactose, and somatic cell content.

In this study we show that milk samples can serve as a substrate for direct amplification of specific DNA sequences using the PCR.

MATERIALS AND METHODS

During routine monthly milk recording, milk samples (25 or 50 ml) were taken from 31 Israeli-Holstein cows. Milk sample preservative tablets were added to each sample, and the samples were kept at 4 C until DNA extraction or treatment prior to the PCR. Somatic cell counts, based on DNA fluorescent dyeing were measured at the central laboratory of the Israeli Cattle Breeders Association. One milliliter of each sample was set aside for PCR amplification. The remainder served for DNA extractions.

DNA extraction was carried out according to the procedure described for semen. For direct PCR milk somatic cells were generally lysed by incubation of the samples in boiling water. Aliquots containing the equivalent of 100 to 2000 cells were taken as a substrate for PCR amplification. Six different primer pairs were used.

RESULTS

DNA precipitated visibly and was readily collected in 16 of the 31 samples. Total DNA yield per sample ranged from 11 to 1101 μ g. Cell counts per milliliter for these samples ranged from 2.4×10^4 to 5.4×10^6 . The DNA yield depended on the total number of cells in the milk sample (i.e., on cell count per milliliter multiplied by 25 or 50, depending on sample size). The regression of DNA yield on total number of cells in the sample, calculated only for the 16 samples for which a DNA precipitate was collected, gave a threshold of roughly 17×10^6 cells for obtaining visible DNA precipitate, which is crucial for successful extraction. Only 2 of 15 sample with a greater total number of cells failed to yield a visible precipitate, but DNA was obtained with fewer cells in only 1 of 13 samples.

All of the primer pairs yielded large amounts of amplification products of the expected length with all of the DNA and milk samples. The products were the same in length and approximate quantity for milk, for DNA extracted from milk, and for DNA extracted from blood. The efficiency of direct PCR on milk samples was not reduced by storage of the sample. PCR on somatic cells from milk samples was readily able to uncover polymorphisms of the microsatellite type.

DISCUSSION

When milk was used as a source of DNA, extractions were generally successful only when a total of 17×10^6 cells or more were used, yielding 11 to 1100 μg of DNA. Thus, milk is a less reliable source of DNA than is blood. Large numbers of DNA extractions for routine genotyping demand a consistent and standard protocol. The present dependence on somatic cell count and the need to adjust the volume of the sample according to this count complicates the procedure and limits the advantage of milk over blood as a source of DNA.

In contrast, milk with preservative added, refrigerated or frozen for 14 to over 200 days, proved to be a reliable source of DNA for PCR amplification. The stability of the samples over a relatively long period, and the need for only a minute quantity of milk per reaction (a few microliters at the most) make milk a preferred source of DNA for genotyping using the PCR. Given suitable apparatus, the PCR procedure can readily be adapted for automated genotyping in a central milk laboratory.

The use of milk for genotyping using the PCR, has a truly liberating effect on all activities that have to do with marker-based genetic analysis of quantitative traits in dairy cattle and subsequent marker-assisted selection. For example, a full manpower year of onerous work is required for collection and extraction of DNA from 1000 blood samples. With milk samples and PCR genotyping, collection of a comparable sample would be carried out primarily by the milk control officers, with some technical supervision, and DNA extraction is not required.

5. Early PCR Amplification Test For Identifying Chimerism In Female Calves Co-Twin To A Male In Cattle.

In cattle, over 90% of twin pregnancies result in anastomoses of the blood systems of the embryos, followed by establishment of cells from one twin in the other. Although innocuous in same sex twins, in mixed sex twins the consequences are female freemartinism, with resulting complete sterility, and reproductive problems up to and including complete sterility in the male. Loss of freemartin females as potential replacements for the dairy herd reduces the scope for genetic improvement through female-side selection. There are also direct economic losses due to reduced fertility of the male twin, when the latter is used for artificial insemination. While the loss of the actual freemartin female cannot be avoided, it has been shown that when cell chimerism is absent the female is fertile. Thus, the fertile 10% of females, most of which are also culled because of the high probability of sterility, can be retained, and bull fecundity problems can be avoided, by an early test for the occurrence of cell chimerism.

In this study, we confirm that cell chimerism, as shown by PCR amplification of a Y chromosome fragment in female blood leukocytes and hair roots, is indeed accompanied by freemartinism as judged by clinical inspection, and that in the absence of cell chimerism, as evidenced by lack of Y-fragment amplification, the female is completely fertile.

MATERIALS AND METHODS

The study included 20 newborn female calves, and three 13-month old heifers, co-twins to males and a single singleton heifer with abnormal internal sex organs. A pregnant heifer and two adult fertile cows, all recorded as co-twins to males, were also examined. A normal singleton female and a normal singleton male served as controls. All animals were of Israel Holstein origin.

Peripheral blood samples were collected in vacuum tubes with EDTA. DNA was extracted from part of the samples, while the leukocytes from the others served directly as a substrate for PCR. For examination of hair root follicles, hair was plucked with the roots, and 30-60 roots were separated using a 3x power magnifying glass, and transferred to TE. The cells were lysed using the same protocol as for direct lysis of leukocytes.

The primer pair Bov97M1 (GATCACTATACATACACCACTC) AND BOV97M2 (AAGGCTATGCTACACAAATTCTG) was used to amplify 114 bp from the BOV97M sequence of the bovine Y chromosome (Miller and Koopman, 1990).

The three 13-month old twin heifers and the one singleton heifer were inspected twice in two weeks by rectal palpation of the reproductive organs. Since this kind of inspection is possible only after one year, the 20 newborn calves were not checked.

RESULTS

PCR amplification of a Y chromosome fragment yielded positive results in blood samples from a normal singleton male and from two 13-month old heifers born co-twin to a male. On clinical examination the two 13-month old heifers were found to be freemartins. PCR amplification of blood leukocyte samples of the 20 newborn female calves co-twins to males yielded 17 positive and three negative results. Although these animals were not examined clinically for freemartinism, the proportion of negative animals (155) is not far from the customary figure of 10% normal heifers born co-twin to males.

PCR amplification yielded negative results in blood samples obtained from a normal singleton female and from one pregnant heifer and two fertile cows, the last three born co-twin to males.

Amplification from hair roots similarly gave positive results for a normal single male, for a clinically diagnosed freemartin female born co-twin to a male, and for eight newborn females. Six of the newborn female calves were also identified as Y-positive by the PCR blood test. Negative results were obtained for a normal singleton female and for a singleton heifer with abnormal internal sex organs.

DISCUSSION

In the present study, the positive Y-amplification found in the freemartin heifers confirms the association between cell chimerism and freemartinism. Conversely, the lack of Y-amplification in the three fertile females co-twins to males (the pregnant heifer and the two adult fertile cows) confirms that when cell chimerism is shown to be absent by this method, the female can be expected to be fertile. Thus, Y-amplification based on blood leukocytes or hair root cells, can serve as a rapid, sensitive and convenient means of identifying potentially fertile females born co-twin to a male. It has been shown that different-sex twins usually have the same proportion of cells of male and female origin. Consequently, absence of chimerism in a female twin, can be taken as an indication of absence of chimerism in her male co-twin. This can be important in identifying those males that will subsequently not be affected by reduced fertility.

6. Determining Relative Microsatellite Allele Frequencies in Pooled Milk Samples in the Presence of Shadow Bands

A major application of genomic maps is for purposes of determining linkage between marker loci and loci affecting traits of economic importance (Economic Trait Loci, ETL). This is necessary for application of marker assisted selection in genetic improvement. In animal populations, marker-ETL linkage can be determined by analyzing the progeny of single individuals, heterozygous at marker loci and ETL. Since the power of analyses based on single families is limited, "Daughter" and "Grand-daughter" designs have been devised in which data are pooled across half-sib families. Most ETL are quantitative in nature. Consequently, the sample size required for establishing marker-ETL linkage is large. However, a statistical development, termed "selective genotyping" significantly reduces the number of offspring that need to be genotyped, with only very minor loss in power. In dairy cattle, the test for marker-ETL linkage under this procedure is based on the quantitative value of marker genotypes found in offspring at the two extreme phenotypic tails of large half-sib families.

Theoretical studies in our laboratory show that the power of selective genotyping, based on allele frequency differences estimated from pooled DNA samples of the selected groups, can be equivalent to power based on individually genotyping the individuals making up the samples; depending on the error variance of allele frequency estimation in the pooled sample. The crucial question for utilization of DNA pooling is the magnitude of experimental error obtained in densitometric estimation of allele frequencies from pooled DNA. Close estimates of allelic frequencies can be obtained in pooled DNA samples for microsatellites having a tetranucleotide repeat motif (such microsatellites do not show artifactual "shadow bands"). We have shown that the same is true for microsatellites having a dinucleotide repeat motif (which do show marked leading shadow bands), when shadow bands and main bands of the various alleles in the pool do not overlap. Allelic variation at most dinucleotide microsatellites in cattle, however, is such that the shadow bands and main bands of the various alleles in the pool overlap. In this case, simple densitometric values for the main bands of the alleles do not provide accurate estimates of allele frequency in the pool.

We describe here development and initial validation of a model which provides estimates of microsatellite allele frequencies in pooled DNA or milk samples, in the presence of overlapping shadow bands. Initial application of the model showed generally good accuracy. Technical error was greater than obtained in the absence of overlapping shadow bands, but still of a magnitude allowing useful application of DNA pooling for linkage mapping of ETL, given reasonable replication.

MATERIALS AND METHODS

Samples

All genotyping was based on milk samples. Samples were obtained for 80 Israel Holstein cows, all half-sib daughters of a single A.I. sire, Amir. For each cow in the sample, an individual milk sample was collected during routine monthly milk recording in Israeli commercial herds. Somatic cell counts of these samples were carried out at the Central laboratory of the Israeli Cattle Breeders Association at Bitan Aharon.

PCR genotyping

PCR genotyping was carried out for two poly(TG) microsatellite markers, HUJ616 (product length 210 to 230 bp), and HUJ177 (product length 122 to 132 bp) (Barendse et al., 1994).

Individual genotyping

Each cow was genotyped with respect to both of the above microsatellites. For individual genotyping, an aliquot of milk containing 10^5 somatic cells was suspended in 15 ml saline. Aliquot size was adjusted according to the somatic cell count of the individual milk sample. The milk suspension was centrifuged for 10 min at 1000 x g and resuspended in 1 ml TE, giving a final concentration of 100 cell/ml. The resuspended cells were lysed by incubation in boiling water for 5 min, at 55 C for 5 min, and again in boiling water for 5 min. 500 cells (5 ml) were used for individual PCR reactions.

Pool genotyping

Composition of pools. In making up the pools, 30 ml of cell suspension, prepared as for individual genotyping (i.e., 3000 cells) were taken for each cow, giving a total of 60,000 cells per pool. Each pool was prepared in two independent replicates, and each preparation was amplified twice. In preliminary experiments, in which various total cell numbers were used in the PCR reaction mixtures, inconsistent results were obtained when the number of cells per reaction mixture was small (~500). Consequently, in the definitive experiments, a total of 5000 cells from each pool were taken for each PCR reaction mixture.

Three pairs of pools were constructed for each of the two markers HUJ616 and HUJ177 (a total of 12 pools in all). Each pool contained a variety of marker alleles. Two of these were the two alleles of the sire, henceforth denoted "sire alleles", although some of the dams also carried these alleles as well. For each marker, the shorter of the two sire alleles is denoted, S; the longer of the two sire alleles is denoted, L. Since the dams provided half of the marker alleles present in each pool, the total frequency of the sire alleles in the pool, varied from 0.7 to 0.8, depending on the frequency of the sire alleles among the dams. The remainder consisted of alleles, other than the sire alleles, that were present in the dams.

By appropriate choice of the individual cows making up each pool, frequencies of the S and L marker alleles derived from the sire were varied in a systematic manner between the pairs of pools, as detailed in Tables 1 and 2. The two pools of each pair, although including a more or less different group of cows, were both constructed so as to comprise approximately the same frequency for the two sire alleles. The first set of pools constructed for marker HUJ177 gave seemingly anomalous results. In order to clarify the source of the anomaly, the pools were made up a second time, taking aliquots from the original resuspended milk somatic cell suspensions, and genotyped a second time.

Genotyping. Composition of the reaction mixture, cycling, and densitometry were as for individual genotyping.

Calculations

Relative intensity of shadow bands. In order to determine the relative intensity of the shadow bands, as compared to the main band, individual cows were chosen which were either homozygous for a given allele, or for which, if heterozygous, the two alleles were well separated with little overlap of shadow bands. Relative intensity, R_{ij} , of the i^{th} shadow band for the marker allele of length j , was calculated as the densitometric value for the given shadow band, divided by the densitometric value for the main band of that allele.

The linear regression of R on allele length was calculated separately for first, second and third shadow bands, pooling R values for a given shadow-band number over all alleles of the two markers.

Actual and relative frequency of marker alleles in the pool. The actual frequency, F_j , of the marker allele of length j , in a given pool, as determined from the individually determined genotypes of the N individuals making up the pool, were calculated as

$$F_j = n_j/2N$$

where,

n_j = the number of occurrences of the marker allele of length j , among the genotypes of the N animals making up the pool.

The relative frequencies of S and L, F_{RS} and F_{RL} , were calculated as

$$F_{RS} = F_S/(F_S + F_L)$$

$$F_{RL} = F_L/(F_S + F_L)$$

where,

F_S = the actual frequency of S in the pool, and

F_L = the actual frequency of L in the pool

Note, that alleles other than S and L that are present in the pool (derived from the dams of the individuals making up the pool), are not included in the calculation of F_{RS} and F_{RL} .

Predicted relative densitometric intensity of the short and long alleles of the sire

Given the individual genotypes of the individuals making up the pool, so that F_S and F_L are known, and taking into account shadow bands and differential amplification, the following model was used to predict the relative densitometric intensities of the main bands of S and L, P_S and P_L ,

$$P_S = I_S / (I_S + I_L)$$

$$P_L = I_L / (I_S + I_L)$$

where, I_S and I_L , the predicted densitometric intensities of the main bands of S and L, respectively, are:

$$I_S = F_S + \sum_{i=1}^3 SF_{S+i} R_{i,S+i}, \quad (i \text{ runs from } 1 \text{ to } 3)$$

$$I_L = F_L + \sum_{i=1}^3 SF_{L+i} R_{i,L+i}$$

where,

$S+i$ and $L+i$ are the markers alleles having i repeats more than the S and L alleles, respectively.

Actual and relative densitometric frequency of marker alleles in the pool

The actual densitometric frequencies of the marker allele of length j , D_j , in a given pool, as determined from the individually determined genotypes of the N individuals making up the pool, were calculated as

$$D_j = d_j / Sd_j$$

where,

d_j = the densitometric reading of the marker allele of length j .

The relative densitometric frequencies of S and L, D_{RS} and D_{RL} , were calculated as

$$D_{RS} = D_S / (D_S + D_L)$$

$$D_{RL} = D_L / (D_S + D_L)$$

where,

D_S = the densitometric frequency of S in the pool, and

D_L = the densitometric frequency of L in the pool

Note, that alleles other than S and L that are present in the pool (derived from the dams of the individuals making up the pool), are not included in the calculation of D_{RS} and D_{RL} .

Estimating the relative frequency of the short and long alleles of the sire from their relative densitometric frequencies

Given the relative densitometric frequencies of S and L in the pool, D_{PS} and D_{PL} , and taking into account shadow bands, the following model was used to obtain estimates, F_{ES} and F_{EL} , of the relative frequencies of S and L in the pool

$$F_{ES} = D_{CS} / (D_{CS} + A_{SL} D_{CL})$$

$$F_{EL} = D_{CL} / (D_{CS} + D_{CL})$$

where, D_{CS} and D_{CL} , the corrected densitometric intensities of the main bands of S and L, respectively, are:

$$D_{CS} = D_S - \sum_{i=1}^3 SD_{S+i} R_{i,S+i}, \quad (i \text{ runs from } 1 \text{ to } 3)$$

$$D_{CL} = D_L - \sum_{i=1}^3 SD_{L+i} R_{i,L+i}$$

where,

$S+i$ and $L+i$ are the markers alleles having i repeats more than S and L, respectively.

RESULTS

Individual genotyping

With the markers used in this study and two others (data not shown), two milk samples were found, each of which showed more than two alleles. Both came from cows registered as singletons, and both were from the same herd. These samples are therefore interpreted as instances of mixed milk samples. With the same four markers, seven cows out of the eighty did not carry any paternal alleles, and therefore could not have been the daughters of their assigned sire. After deducting the mixed milk samples and the samples from cows with wrongly assigned paternity, 71 daughters of Amir remained. The total proportion of sire alleles among the daughters was .70 for HUI616 and .82 for HUI77, depending on the frequency of the sire allele among the dams of the daughters.

Shadow bands

The relative intensity of shadow bands as a function of allele repeat number, was investigated for markers HUI616 and HUI77, both of which showed strong shadow bands. Allele repeat number for HUI77 extended from $n=11$ to $n=16$, and for HUI616 from $n=13$ to $n=24$. Thus, taken together the two markers covered a range of repeat numbers extending from $n=11$ to $n=24$. The total number of data points for the first shadow band was 108. The total number of data points for the second and third shadow bands were somewhat less (80 and 43, respectively), since instances where shadow bands overlapped main bands or other shadow bands were not included.

Relative intensity of the shadow bands decreased strongly from first to second, much less so, if at all, from second to third. For all three shadow bands, there was a clear positive relationship of relative intensity on allele repeat number. The linear regression equations for the three shadow band orders, were:

$$R_1 = -.06 + .040n \quad (r=.95)$$

$$R_2 = -.36 + .041n \quad (r=.99)$$

$$R_3 = -.66 + .056n \quad (r=.97)$$

In a small number of cases, genotyping was carried out both on DNA extracted from milk samples as well as directly on the milk samples themselves. The relative intensity of the shadow bands was the same in both procedures (data not shown).

Pool densitometry

Table 1 shows, for marker HUJ616, the densitometric intensity values for the main band of the short sire allele (S) and for all bands located above this band in the gel. The corresponding values are given in Table 2 for marker HUJ177. In both cases, the long sire allele (L) is separated by one repeat only from the short allele so that densitometric values for the two sire alleles are directly adjacent in the tables. Considering duplicate amplifications of the same pool, although the absolute intensities can vary as much as two-fold, as in pool D2 of marker HUJ616 (pool 2 of sample D), or pools G2 and L2 of marker HUJ177, the relative proportions of the various bands within a given lane remain quite similar. When frequency of the S and L alleles in the pooled sample were equal (pools C, D, I, J), densitometric intensity of S was greater than that of L. This is probably due to the strong contribution of the first shadow band of L to the main band of S. A major exception is found for pool H1 of marker HUJ177 (Table 2) for which densitometric intensity of L was more than twice that of S. In order to obtain some idea as to the possible source of this exceptional result, the entire set of pools for this marker were prepared and genotyped again. The densitometric results, including the exceptional results of pool H1, corresponded very closely to those obtained in the first set of pools (data not shown). Thus, we attribute the anomalous values of pool H1 to some error in the somatic cell counts assigned to one or more of the milk samples included in this particular pool, rather than to technical error in constructing the pools or carrying out the amplification. For this reason, pool H1 of this marker was not included in subsequent calculations of technical error variance.

In spite of the large magnitude of differences in actual allele frequencies of the various pools, relative densitometric intensities of S (Column D in Tables 1 and 2), varied over a rather narrow range. Nevertheless, there was a clear overall tendency for relative densitometric intensity of S to increase with increase in the actual relative allelic frequency of the short sire allele in the pool. The estimated relative frequency of S (Column E in tables 1 and 2), derived from the densitometric band intensities according to the model described in the methods section was quite consistent within duplicate amplification of the same pool, but showed considerable variation between duplicate pools of the same sample, and between pairs of samples having the same constructed

actual allele frequencies. In this case, however, the range of variation of estimated frequencies of S, extended over the full range of values shown by the actual allele frequencies of this allele across the various samples.

Table 3 shows the actual (F_s) and estimated (F_{es}) relative allele frequencies, and observed (D_s) and predicted (D_{ps}) relative densitometric intensities of S, by marker and sample. For marker HUI616 there is a close correspondence between observed and predicted relative densitometric intensities. This is generally true for marker HUI77 as well, but in this case there are a number of discrepancies (e.g., pools G,H,J). In all cases, however the correspondence between the mean predicted value and the mean observed value, for pairs of samples having the same constructed allele frequencies was extremely close. For allele frequencies, the correspondence between actual frequency in the pool and estimated frequency, although generally close, exhibits a number of rather large discrepancies (e.g., pools D, G, H). Nevertheless, there were only two pairs of samples (B and D for HUI616, and H and I for HUI77) which would have been ranked incorrectly according to their estimated allele frequencies. The actual differences in allele frequencies between these misranked samples were .08 for samples B and D, and .12 for samples H and I. There were no instances of rank reversal of estimated allele frequencies, for samples which differed in actual allele frequency by 0.20 or more. For estimated allele frequencies, as well, the correspondence between mean estimated allele frequency value, and actual allele frequency value, for pairs of samples having the same constructed allele frequencies was quite close, particularly for HUI77. For these mean values there were no instances of rank reversal for any of the sample pairs, even when actual allele frequency differences were of the order of .10.

For each sample, the deviation of the mean estimated relative allele frequency of S from the actual frequency of S in the sample was calculated. Similarly, for each possible pair of samples, the deviation of the difference based on the estimated allele frequencies in the samples, and the difference based on the actual allele frequency of the sample was calculated. In both cases, the distribution was symmetrically centered on zero (data not shown), showing that the estimated allele frequencies are not biased with respect to the actual allele frequencies.

Table 4 shows standard deviations between estimated allele frequencies obtained (i) from duplicate amplifications of the same pool, (ii) from mean values (over duplicate amplifications) of replicate pools of the same sample, and (iii) from mean values (over replicate pools) of samples constructed to have approximately the same allele frequencies. Variance components obtained from these standard deviations are also shown. The variance components between duplicate amplifications (V_A) and between replicate pools (V_P), are of a magnitude which can be readily reduced to manageable proportions by replication. The variance components between pairs of samples (V_S), is relatively small. This is fortunate, since in many experiments it is generally not feasible to obtain pairs of samples, except by artificially splitting an existing sample. From these variance components, we obtain a technical standard error of .082, for estimates of allele frequencies from a single amplification of a single pool of a single sample; of .054, for estimates of allele frequencies from single amplifications of replicate pools of a single sample; of .049, for duplicate amplifications of replicate pools

of a single sample; of .044, for single amplifications of replicate pools of a pair of samples from the same population; and of .040, for single amplifications of four samples from the same population, as proposed in this study.

DISCUSSION

Differential amplification of microsatellite marker alleles is a commonly reported phenomenon, and was also observed in this study. However, detailed examination of differential amplification, showed that it was not strong over the allele size range observed. It was also found that including a factor for differential amplification in the model did not improve the correspondence between estimated allele frequencies and actual allele frequencies in the pools, while it increased technical error variance. This is probably due to the fact that including additional factors in the model increases the "modeling" component of technical error. For this reason, we did not include differential amplification in the model presented here. The fact that the observed deviation of estimates of allele frequency from actual allele frequencies in the samples was centered about zero, implies that ignoring differential amplification did not bias the estimates in any consistent way.

The technical standard deviation of .045, obtained between duplicate amplifications of a single sample, is about twice that obtained in Khatib et al. (1994), for allele frequency estimation in the absence of overlapping shadow bands. Part, perhaps all, of this increase is certainly due to the numerical manipulations of the model. Thus, it is not clear whether this can be reduced substantially by increased experience with the pooling technique. The technical standard deviation of .062 obtained between replicate pools (after subtracting error due to duplicate amplifications), is also about twice that obtained in Khatib et al. (1994). It is not clear why this component should be so large, since error variation introduced by the model would presumably be included in the standard deviation between duplicate amplifications. Thus, we believe that further experience in working with milk samples, and with the pooling technique in general, should be able to reduce this error component.

Referring to Figure 3 of Darvasi and Soller (1994), shows that technical standard error 0.04, anticipated with the degree of replication planned in this proposal would have little effect on statistical power, when power is high; but would reduce power by about one-third (from .74 to .50) as compared to DNA pooling in the absence of technical error, when power is only moderate. However, due to the very large sample sizes that can be handled by means of the pooling technique, the effective power of the proposed experiment should be at least equivalent to that which could be obtained in practice with individual selective genotyping. Furthermore, although not yet quantified in detail, it can be anticipated that selective DNA pooling would be part of an overall sequential mapping program, in which selective DNA pooling would be only the first stage, followed by independent replication of those markers showing significant effects, and by individual genotyping of those markers that continued to show significant effects following replication. In such a sequential procedure, a relatively high Type I error would be employed in the DNA pooling stage of the research. This should also contribute to a lower overall Type II error, than implied by the technical standard error.

Table 1. Marker HUI616. Observed intensities of main band of the shorter paternal allele (n=19) and all higher bands, for all amplifications (A) of all pools (P) and samples (S). Relative proportion of the sire short allele (RP) based on direct densitometry (D, and estimated using the model (E). In parentheses under "S": actual RP in the sample. Main band of longer paternal allele, n=20.

S	P	A	Number of repeats in band (n)						RP	
			19	20	21	22	23	24	D	E
A (.43)	1	1	5.62	4.59	2.57	2.76	2.05	1.20	.55	.44
		2	4.47	3.93	1.89	1.87	1.70	1.10	.53	.38
	2	1	3.85	3.38	1.81	1.74	1.69	1.07	.53	.39
		2	2.30	1.98	0.91	1.06	0.86	0.77	.54	.33
B (.44)	1	1	3.45	2.81	1.40	1.45	1.13	0.76	.55	.41
		2	2.96	2.45	1.21	1.22	1.28	0.64	.55	.45
	2	1	Not done							
		2	3.39	2.56	1.29	1.44	1.25	0.97	.57	.43
C (.52)	1	1	1.67	1.19	0.63	0.69	0.78	0.33	.58	.56
		2	1.80	1.03	0.53	0.48	0.57	0.31	.64	.61
	2	1	1.44	1.03	0.35	0.35	0.46	0.28	.58	.47
		2	1.79	1.14	0.55	0.64	0.63	0.47	.61	.52
D (.52)	1	1	3.84	3.23	1.77	1.71	1.36	1.03	.54	.39
		2	3.65	3.49	1.78	1.65	1.32	1.06	.51	.30
	2	1	2.61	1.65	0.90	0.83	0.88	0.77	.61	.52
		2	4.68	3.23	1.52	1.44	1.13	1.05	.54	.46
E (.68)	1	1	4.41	2.19	1.62	1.85	1.44	1.08	.67	.65
		2	4.72	2.61	2.09	2.10	1.99	1.31	.64	.66
	2	1	4.22	2.57	1.84	1.98	1.66	1.12	.62	.60
		2	4.36	2.94	1.94	2.18	1.97	1.46	.60	.52
F (.62)	1	1	5.12	3.28	2.45	2.78	2.49	1.10	.61	.65
		2	3.94	3.12	2.61	2.66	2.43	0.98	.56	.61
	2	1	6.02	4.17	2.52	2.38	2.26	0.98	.59	.58
		2	6.79	4.79	3.17	3.45	3.58	1.23	.59	.63

Table 2. Marker HUJII77. Observed intensities of main band of the shorter paternal allele (n=13) and all higher bands, for all amplifications (A) of all pools (P) and samples (S). Relative proportion of the sire short allele (RP) based on direct densitometry (D, and estimated using the model (E). In parentheses under "S": actual RP in the sample. Main band of longer paternal allele, n=14.

S	P	A	Number of repeats in band				RP	
			S=13	L=14	15	16	D	E
G (.38)	1	1	9.75	11.84	3.96	1.49	.45	.28
		2	8.09	10.27	2.86	1.97	.44	.24
	2	1	11.68	14.66	5.54	5.19	.44	.25
		2	5.97	7.90	2.44	2.53	.43	.21
H (.38)	1	1	21.01	47.86	20.12	7.44	.31	-.10
		2	15.35	34.35	12.76	4.99	.31	-.09
	2	1	26.13	20.49	7.95	4.46	.56	.49
		2	19.16	15.26	6.57	3.36	.56	.49
I (.50)	1	1	14.65	11.09	4.34	1.94	.57	.53
		2	15.76	11.13	6.38	4.35	.59	.56
	2	1	15.77	16.91	12.93	15.31	.48	.37
		2	15.07	15.80	8.08	8.50	.49	.35
J (.50)	1	1	21.88	18.38	9.24	4.61	.54	.48
		2	15.76	10.34	4.56	3.81	.60	.56
	2	1	20.44	12.75	2.91	2.00	.62	.55
		2	26.14	15.84	8.79	4.78	.62	.61
K (.63)	1	1	19.54	10.99	7.58	1.93	.64	.67
		2	23.54	8.16	6.26	2.61	.74	.80
	2	1	19.84	11.57	7.74	2.94	.63	.65
		2	13.37	10.35	7.41	2.82	.56	.55
L (.63)	1	1	21.60	13.02	4.04	2.41	.62	.58
		2	22.12	11.62	4.43	2.77	.66	.63
	2	1	16.83	11.13	5.99	2.94	.60	.58
		2	8.62	5.63	2.46	1.74	.61	.56

Table 3. Relative proportion of the sire short allele, according to marker allele. Actual (F) and estimated (E) proportions; Observed (D) and predicted (P) band intensities. Means of replicate pools of the same sample (4 amplifications) and means of sample pairs (8 amplifications).

HUJ616					HUJII77				
Sample	F	E	D	P	Sample	F	E	D	P
A	.43	.38	.55	.54	G	.38	.24	.44	.50
B	.44	.43	.54	.56	H	.38	.49	.56	.50
C	.52	.54	.59	.60	I	.50	.45	.53	.54
D	.52	.42	.58	.57	J	.50	.55	.60	.55
E	.68	.61	.66	.63	K	.63	.67	.64	.62
F	.62	.62	.60	.59	L	.63	.59	.62	.62
A+B	.43	.41	.54	.55	G+H	.38	.37	.50	.50
C+D	.52	.48	.58	.58	I+J	.50	.50	.56	.55
E+F	.65	.61	.63	.61	K+L	.63	.63	.63	.62

Table 4. Standard deviations (SD) of the estimated relative sire short allele frequencies: Between duplicate amplifications of the same pool (s^2_A), between means of duplicate pools of the same sample (s^2_R), and between samples constructed to have the same paternal allele frequencies (s^2_S). Standard deviations are pooled over the two markers. Total degrees of freedom (d.f.) of estimates are shown. The corresponding variance components (V_A , V_R , V_S) estimated from these standard deviations, are also shown.

Source	d.f.	SD	Component	Variance
s^2_A	24	.045	V_A	.00198
s^2_R	12	.069	V_R	.00380
s^2_S	5	.058	V_S	.00097

VII. EVALUATION OF ORIGINAL OBJECTIVES

We have been very successful in that a linkage map of hypervariable markers, mostly PCR based, is now in publication and the original objective of a 20 cM map is within a few markers of reality. Locus specific VNTRs in the form of microsatellites were identified and mapped. This map will be the foundation for the growing bovine linkage map since the database and reference families are being used by many laboratories around the world. Both the U.S. and Israel investigators have played a major role in this organized international activity. We will continue to fill gaps in the developing linkage map assisted by somatic cell genetics and to add conserved sequences to the linkage map to facilitate comparative mapping. Our ultimate goal of one conserved type I marker every 20 cM will not likely be reached for 2 or 3 years. Work will continue toward that objective, however, with the momentum generated from this project.

Publications (US)

Dietz, A.B., Neibergs, H.L. and Womack, J.E.: Assignment of eight loci to bovine syntenic groups using PCR: extensive of a comparative gene map. *Mammalian Genome* 3:106-111, 1992.

Tobin-Janzen, T.C. and Womack, J.E.: Comparative mapping of IGHG1, IGHM, FES, and FOS in domestic cattle. *Immunogenetics* 36:157-165, 1992.

Lei, L., Teal, A., Bensaid, A., Dunlap, S., Dietz, A.B., and Womack, J.E.: Somatic cell mapping of the T-cell receptor CD3 complex and CD8 genes in cattle. *Immunogenetics* 36:224-229, 1992.

Gallagher, D.S., Womack, J.E., Baehr, W., and Pittler, S.J.: Syntenic assignments of visual transduction genes in cattle. *Genomics* 14, 699-706, 1992.

Ryan, A.M., Gallagher, D.S., and Womack, J.E.: Syntenic mapping and chromosomal localization of bovine a and b interferon genes. *Mamm. Genome* 3:575-578, 1992.

Eggen, A., Solinas-Toldo, S., Dietz, A.B., Womack, J.E., Stranzinger, G., and Fries, R.: *RASA* contains a polymorphic microsatellite and maps to bovine syntenic group U22 on chromosome 7q2.4-qter. *Mammalian Genome* 3:559-563, 1992.

Steffen, P., Eggen, A., Dietz, A.B., Womack, J.E., Stranzinger, G. and Fries, R.: Isolation and mapping of polymorphic microsatellites in cattle. *Animal Genetics* 24:121-124, 1993

Ryan, A.M. and Womack, J.E.: Type I interferon genes in cattle: restriction fragment length polymorphisms, gene numbers and physical organization on bovine chromosome 8. *Animal Genetics* 24:9-16, 1993.

Ryan, A.M. and Womack, J.E.: Somatic cell mapping of the bovine prion protein gene and restriction fragment length polymorphism studies in cattle and sheep. *Animal Genetics* 24:23-26, 1993.

Ryan, A.M. and Womack, J.E.: Hybridization profiles and restriction fragment length polymorphisms for bovine and ovine interferon genes. *Animal Biotechnology* 4:11-30, 1993.

Ryan, A.M., Womack, J.E., Yu, J., Lin, J.-H., Wu, X.-R., Sun, T.-T., Clarke, V., and D'Eustachio, P. Chromosomal localization of uroplakin genes of cattle and mice. *Mammalian Genome* 4:656-661, 1993

Modi, W.S., Gallagher, D.S., and Womack, J.E.: Molecular organization and chromosomal localization of six highly repeated DNA families in the bovine genome. *Animal Biotechnology* 4:143-161, 1993.

Barendse, W., Armitage, S.M., Kossarek, L.M., Shalom, A., Kirkpatrick, B.W., Ryan, A.M., Clayton, D., Li, L., Neibergs, H.L., Zhang, N., Grosse, W.M., Weiss, J., Creighton, P., McCarthy, F., Ron, M., Teale, A.J., Fries, R., McGraw, R.A., Moore, S.S., Georges,

M., Soller, M., Womack, J.E., and Hetzel, D.J.S.: A genetic linkage map of the bovine genome. *Nature Genetics* 6:227-235, 1994.

Ryan, A.M., Gallagher, D.S., Jr., Schöber, S., Schwenger, B., and Womack, J.E.: Somatic cell mapping and in situ localization of the bovine uridine monophosphate synthase gene (UMPS). *Mammalian Genome* 5:46-47, 1994.

Ryan, A.M. and Womack, J.E.: Assignment of the proopiomelanocortin (POMC) gene to BTA 11 and RFLPs in cattle and sheep. *Animal Genetics* 25:58, 1994.

Ryan, A.M., Womack, J.E., Lin, J.-H., and Sun, T.-T.: *TaqI* RFLPs at the bovine uroplakin II locus (UPK2). *Animal Genetics* 25:58, 1994.

Ryan, A.M., Womack, J.E., Wu, X.-R., and Sun, T.-T.: *TaqI* and *BglII* RFLPs for the bovine uroplakin III loci UPK3A and UPK3B. *Animal Genetics*, 25:64, 1994.

Ryan, A.M. and Womack, J.E.: Somatic cell mapping of the mitochondrial aconitase gene (ACO2) to bovine chromosome 5. *Animal Genetics* 25:123, 1994.

Ryan, A.M. and Womack, J.E.: Rapid communication: A *PstI* restriction fragment length polymorphism in the ovine *Aconitase* 2 gene. *J. Anim. Sci.* 72:1076, 1994.

Publications (ISRAEL)

Soller M (1990) Genetic mapping of the bovine genome using DNA-level markers with particular attention to loci affecting quantitative traits of economic importance. *J. Dairy Sci.* 73:2628-2646.

Shalom A, Soller M, Friedmann A (1992). Dinucleotide repeat polymorphism at the bovine HUJ616 locus. *Animal Genetics* 24: 327.

Shalom A, Soller M, Friedmann A (1992). Dinucleotide repeat polymorphism at the bovine HUJ614 locus. *Animal Genetics* 24:327.

Shalom A, Soller M, Friedmann A (1992). Dinucleotide repeat polymorphism at the bovine HUJ625 locus. *Animal Genetics* 24: 328.

Shalom A, Mosig MG, Barendse W, Friedmann A, Soller M (1993). Dinucleotide repeat polymorphism at the bovine HUJ246, HUJ1177, HUJ223, HUJVI74 and HUJI75 loci. *Animal Genetics* 25: 56

Lipkin E, Shalom Anne, Khatib H, Soller M, Friedmann A (1993). Milk as a source of deoxyribonucleic acid and substrate for the polymerase chain reaction. *J. of Dairy Science* 76:2025-2032.

Lipkin E, Tikoschinsky Y, Arbel R, Sharoni D, Soller M, Friedmann A (1993). Early PCR amplification test for identifying chimerism in female calves co-twin to a male in cattle. *Animal Biotechnology* 4:195-201.

Barendse W, Armitage SM, Kossarek L, Kirkpatrick, BW, Ryan AM, Shalom A, Clayton D, Li L, Neibergs H, Nan Z, Grosse M, Creighton P, McCarthy F, Ron M, Soller M, Fries R, McGraw RA, Moore SS, Teale A, Georges M, Womack JE, Hetzel DJS (1994). A preliminary map of the bovine genome. *Nature Genetics* 6:227-235.

Shalom A, Mosig MG, Barendse W, Soller M, Friedmann A (1994). Dinucleotide repeat polymorphism at the bovine HUJ673, HUJ121, HUJ174, HUJ113 and HUJI29 loci. *Animal Genetics* (submitted).