15

77

BARD

FINAL REPORT

PROJECT NO. IS-3103-99

**Determination of Allele Frequencies for Quantitative
Trait Loci in Commercial Animal Populations**

J.I. Weller, H.A. Lewin, M. Ron

2004

**BARD Final Scientific Report
Cover Page****Date of Submission of the report:****BARD Project Number:** IS-3103-99CR**Project Title:** Determination of Allele Frequencies for Quantitative Trait Loci in Commercial Animal Populations**Investigators****Principal Investigator (PI):** Joel I. Weller**Co-Principal Investigator (Co-PI):** Harris A. Lewin**Collaborating Investigators:** Micha Ron**Institutions****ARO, The Volcani Center****University of Illinois****ARO, The Volcani Center** **Keywords** *not* appearing in the title and in order of importance. Avoid abbreviations.

Genetic mapping, genetic markers, DNA-microsatellites, candidate genes, marker-assisted selection

Abbreviations commonly used in the report, in alphabetical order: **BTA**, *Bos taurus* autosome; **ETL**, economic trait loci; **HAS**, *Homo sapiens* autosome.**Budget:** IS: \$159,000

US: \$141,000

Total: \$300,000

Signature
Principal InvestigatorSignature
Authorizing Official, Principal Institution

BARD Final Scientific Report

Cover Page

Publication Summary (numbers)

	Joint IS/US authorship	US Authors only	Israeli Authors only	Total
Refereed (published, in press, accepted)	2	2	9	13
Submitted, in review, in preparation			1	1
Invited review papers				0
Book chapters				0
Books				0
Master theses			1	1
Ph.D. theses		1	1	2
Abstracts	2	2	4	8
Not refereed (proceedings, reports, etc.)			2	2

Postdoctoral Training: List the names and social security/identity numbers of all postdocs who received more than 50% of their funding by the grant.

Jun Heon Lee, Ph.D., SS# 325-98-9756

Cooperation Summary (numbers)

	From US to Israel	From Israel to US	Together, elsewhere	Total
Short Visits & Meetings	1	1	2	4
Longer Visits (Sabbaticals)	0	2	0	2

Description of Cooperation:

Dr. Ron spent a sabbatical at the U of IL from August, 2002, through April, 2003. Miri Cohen, an ARO doctorate student, visited the U of IL from October, 2003, through January, 2004. Dr. Weller visited the U of IL in August, 2002. Dr. Lewin visited the ARO in October, 2003. Drs. Lewin and Ron met in August, 2002, at the ISAG meeting at Göttingen, Germany, and in January, 2003, at the Plant and Animal Genome conference at San Diego, USA.

Patent Summary (numbers)

	Israeli inventor (s) only	US inventor (s) only	Joint IS/US inventors	Total
Submitted	0	0	0	0
Issued (allowed)	0	0	0	0
Licensed	0	0	0	0

Table of Contents

	Page
Abstract	4
Achievements	5
Details of cooperation	7
List of publications	8
Appendix - "In press" publications	10
Factors affecting incorrect paternity assignment in the Israeli Holstein population.	11
Cloning and characterization of <i>FAM13A1</i> - a gene near a milk protein QTL on BTA6	25

Abstract

Individual loci affecting economic traits in dairy cattle (ETL) have been detected via linkage to genetic markers by application of the granddaughter design in the US population and the daughter design in the Israeli population. From these analyses it is not possible to determine allelic frequencies in the population at large, or whether the same alleles are segregating in different families. We proposed to answer this question by application of the "modified granddaughter design", in which granddaughters with a common maternal grandsire are both genotyped and analyzed for the economic traits. The objectives of the proposal were: 1) to fine map three segregating ETL previously detected by a daughter design analysis of the Israeli dairy cattle population; 2) to determine the effects of ETL alleles in different families relative to the population mean; 3) for each ETL, to determine the number of alleles and allele frequencies.

The ETL on *Bos taurus* autosome (BTA) 6 chiefly affecting protein concentration was localized to a 4 cM chromosomal segment centered on the microsatellite *BM143* by the daughter design. The modified granddaughter design was applied to a single family. The frequency of the allele increasing protein percent was estimated at 0.63 ± 0.06 . The hypothesis of equal allelic frequencies was rejected at $p < 0.05$. Segregation of this ETL in the Israeli population was confirmed.

The genes *IBSP*, *SPP1*, and *LAP3* located adjacent to *BM143* in the whole genome cattle-human comparative map were used as anchors for the human genome sequence and bovine BAC clones. Fifteen genes within 2 cM upstream of *BM143* were located in the orthologous syntenic groups on HSA4q22 and HSA4p15. Only a single gene, *SLIT2*, was located within 2 cM downstream of *BM143* in the orthologous HSA4p15 region. The order of these genes, as derived from physical mapping of BAC end sequences, was identical to the order within the orthologous syntenic groups on HSA4: *FAM13A1*, *HERC3*, *CEB1*, *FLJ20637*, *PP2C-like*, *ABCG2*, *PKD2*, *SPP1*, *MEPE*, *IBSP*, *LAP3*, *EG1*, *KIAA1276*, *HCAPG*, *MLR1*, *BM143*, and *SLIT2*. Four hundred and twenty AI bulls with genetic evaluations were genotyped for 12 SNPs identified in 10 of these genes, and for *BM143*. Seven SNPs displayed highly significant linkage disequilibrium effects on protein percentage ($P < 0.0001$) with the greatest effect for *SPP1*. None of SNP genotypes for two sires heterozygous for the ETL, and six sires homozygous for the ETL completely corresponded to the causative mutation. The expression of *SPP1* and *ABCG2* in the mammary gland corresponded to the lactation curve, as determined by microarray and QPCR assays, but not in the liver. Anti-sense *SPP1* transgenic mice displayed abnormal mammary gland differentiation and milk secretion. Thus *SPP1* is a prime candidate gene for this ETL.

We confirmed that *DGAT1* is the ETL segregating on BTA 14 that chiefly effects fat concentration, and that the polymorphism is due to a missense mutation in an exon. Four hundred Israeli Holstein bulls were genotyped for this polymorphism, and the change in allelic frequency over the last 20 years was monitored.

A method was developed to estimate ETL effects in an animal model analysis of the complete commercial population even though only a small fraction of the population was genotyped. This method gave accurate results on simulated data, but not on the actual data for *DGAT1*.

This project demonstrated that ETL allelic frequencies can be estimated even if the actual gene responsible has not been identified. This result has major implications for marker assisted selection in farm animals. We have shown that the same alleles for *DGAT1* segregating in the Dutch Holstein population are also segregating in the Israeli Holstein population.

Achievements

The daughter design analysis of the Israeli Holstein population was completed. Eleven Israeli Holstein families including 5221 cows were analyzed for eight economic traits; milk, fat and protein production, fat and protein percentage, somatic cell score (SCS), herd-life, and female fertility. The cows were genotyped for 73 microsatellites with maximum spacing between markers of 53 cM. There were 86,304 informative genotypes. Significance was determined by controlling the false discovery rate at 0.4. In order to perform interval mapping for the families with significant contrasts, 21 additional markers were genotyped on BTA 2, 7, and 27. The bootstrap confidence intervals for gene effect did not include zero for protein percent on BTA 2; and fat yield, protein yield, and somatic cell score on BTA 7. Estimated economic trait locus (ETL) heterozygosity was 33%, which is consistent with the hypothesis that only two alleles are segregating with unequal allele frequency.

Nine Israeli Holstein sire families with 2984 daughters were analyzed for ETL effects on BTA 6 for five milk production traits by a daughter design. Twelve microsatellites spanning positions 0 to 122 cM of BTA6 were genotyped. Two sires were segregating for a locus affecting protein and fat percentage near position 55 cM with an estimated substitution effect of 0.18% protein, which is equivalent to one phenotypic standard deviation. This locus was localized to a confidence interval of 4 cM. One of these sires was also heterozygous for a locus affecting milk, fat, and protein production near the centromere. The hypothesis of two segregating loci was verified by multiple regression analysis, and interval mapping of the canonical variables. A third sire was heterozygous for a locus affecting milk and protein percentage near the telomeric end of the chromosome.

The modified granddaughter design was developed to obtain estimates of ETL allele frequencies in dairy cattle populations, and to determine ETL genotypes for both homozygous and heterozygous grandsires. The method is based on determining the ETL allele passed from grandsires to their maternal granddaughters using haplotypes consisting of several closely linked genetic markers. This method was applied to simulated data of ten grandsire families, each with 500 granddaughters, and a ETL with a substitution effect of 0.4 phenotypic standard deviations; and to actual data for a previously analyzed ETL in the center of BTA6, with substitution effect

of one phenotypic standard deviation on protein percent. In the simulated data the standard error for the estimated ETL substitution effect with four closely linked multiallelic markers was only 7% greater than the expected standard error with completely correct identification of ETL allele origin. The method estimated the population ETL allelic frequency as 0.64 ± 0.07 , as compared to the simulated value of 0.7. In the actual data, the frequency of the allele that increases protein percent was estimated as 0.63 ± 0.06 . In both data sets the hypothesis of equal allelic frequencies was rejected at $p < 0.05$.

The genes *IBSP*, *SPP1* and *LAP3* located adjacent to *BM143* in the whole genome cattle-human comparative map were used as anchors for the human genome sequence and bovine BAC clones. Fifteen genes within 2 cM upstream of *BM143* were located in the orthologous syntenic groups on HSA4q22 and HSA4p15. Only a single gene, *SLIT2*, was located within 2 cM downstream of *BM143* due to a "gene desert" in the orthologous HSA4p15 region. The order of these genes, as derived from physical mapping of BAC end sequences, was identical to the order within the orthologous syntenic groups on HSA4: *FAM13A1*, *HERC3*, *CEB1*, *FLJ20637*, *PP2C-like*, *ABCG2*, *PKD2*, *SPP1*, *MEPE*, *IBSP*, *LAP3*, *EG1*, *KIAA1276*, *HCAPG*, *MLR1*, *BM143*, and *SLIT2*. Four hundred and twenty AI bulls with genetic evaluations were genotyped for 12 SNPs identified in 10 of the 16 genes, and for *BM143*. Seven SNPs displayed highly significant linkage disequilibrium effects on protein percentage ($P < 0.0001$) with the greatest effect for *SPP1*, but none of the SNP genotypes of two sires heterozygous for the ETL, and six sires homozygous for the ETL completely corresponded to the causative mutation. The expression of *SPP1* and *ABCG2* in the mammary gland corresponded to the lactation curve, as determined by microarray and QPCR assays, but not in the liver. Furthermore, anti-sense *SPP1* transgenic mice displayed abnormal mammary gland differentiation and milk secretion (Nemir et al., 2000, *J Biol Chem.* 14; 969-976). Thus *SPP1* is a prime candidate gene for the ETL on BTA6 affecting % protein.

We found a population-wide linkage disequilibrium on BTA 14 between microsatellite *ILSTS039* and *DGATI*, the quantitative trait locus affecting milk production traits, in the Israeli Holstein population. A total of 394 bulls were genotyped for both *DGATI* and *ILSTS039*, and 1747 cows were genotyped for *ILSTS039*. The *ILSTS039* allele termed "225" and the *DGATI* K allele (substitution of a lysine residue with alanine) were associated with decreased milk production, and increased fat production and fat and protein percent. The number of "225" *ILSTS039* and K *DGATI* alleles per individual were the same for 80% of the bulls genotyped.

From the effects associated with cows homozygous for the 225 allele, the effect of the quantitative trait locus appears to be approximately codominant. The substitution effect was 0.16% fat. Genotype probabilities for the quantitative gene were determined for the entire Israeli Holstein milk-recorded population, including 507,725 cows and 1442 bulls, using segregation analysis. Overall frequency of the allele that increased fat percent was 8.9% in cows and 15.5% in bulls. The frequency of this allele decreased from 1981 until 1990, from 15 to 5%, and since has increased to 10%. The effects estimated on the population-wide analyses of both cows and bulls were similar to the effect associated with *DGAT1* in the daughters of genotyped bulls. Modified animal model evaluations were computed for the entire population with the effect of this gene included in the model. The correlations between the modified and standard animal model evaluations for all traits were > 0.99 .

A method was developed to estimate ETL effects in an animal model analysis of the complete commercial population even though only a small fraction of the population was genotyped. Sire quantitative gene genotype was determined by the regression of the daughter genetic evaluations on their paternal markers. Marker allele effects were included in an animal model analysis of the simulated populations. The algorithm of Whittaker et al. (1996) was used to estimate gene effects and location. Estimates for both effect and location of the quantitative gene were nearly unbiased. Cow genetic evaluations were always more accurate by the model proposed than by a standard animal model. The increase of genetic gain due to marker-assisted selection of young sires, as compared to dam selection was between 2 to 15%. Type I error rate did not appreciably affect selection decisions.

Details of Cooperation

Dr. Ron spent a sabbatical at the University of Illinois from August, 2002, through April, 2003. He worked on microarray analysis in mice. Miri Cohen, an ARO doctorate student, visited the University of Illinois from October, 2003, through January, 2004, and worked on real-time PCR. Dr. Weller made a short visit to the University of Illinois in August, 2002, to report on the current status of the Israeli results, and Dr. Lewin made a short visit to the ARO in October, 2003, to summarize the results. Drs. Lewin and Ron met in August, 2002, at the ISAG meeting at Göttingen, Germany, and in January, 2003, at the Plant and Animal Genome conference at San Diego, USA.

List of Publications

- Ron, M., Kliger, D., Feldmesser, E., Seroussi, E. Ezra, E., and Weller, J. I. (2001) Multiple QTL analysis of bovine chromosome 6 in the Israeli Holstein population by a daughter design. *Genetics* **159**; 727-735.
- Israel, C. and Weller, J. I. (2002) Estimation of quantitative trait loci effects in dairy cattle populations. *J. Dairy Sci.* **85**; 1285-1297.
- Rodriguez-Zas, S. L., Southey, B. R., Heyen, D. W. and Lewin, H. A. (2002) Detection of quantitative trait loci influencing dairy traits using a model for longitudinal data. *J. Dairy Sci.* **85**; 2681-2691.
- Rodriguez-Zas, S. L., Southey, B. R., Heyen, D. W. and Lewin, H. A. (2002) Interval and composite interval mapping of somatic cell score, milk yield and components in dairy cattle. *J. Dairy Sci.* **85**; 3081-3091.
- Weller, J. I., Weller, H., Kliger, D., and Ron, M. (2002) Estimation of quantitative trait locus allele frequency via a modified granddaughter design. *Genetics* **162**; 841-849.
- Ron, M., Domochofsky, R., Golik, M., Seroussi, E., Ezra, E., Shturman, C., and Weller, J. I. (2003) Analysis of vaginal swabs for paternity testing and marker-assisted selection in cattle. *J. Dairy Sci.* **86**; 1818-1820.
- Weller, J. I., Golik, M., Seroussi, E., Ezra, E., and Ron, M. (2003) Population-wide analysis of a QTL affecting milk-fat production in the Israeli Holstein population. *J. Dairy Sci.* **86**; 2219-2227.
- Ashwell, M. S., Heyen, D. W., Sonstegard, T. S., Van Tassell, C. P., Da, Y., VanRaden, P. M., Ron, M., Weller, J. I., and Lewin, H. A. (2004) Detection of quantitative trait loci affecting female fertility and milk production in ten Dairy Bull DNA Repository families. *J. Dairy Sci.* **87**; 468-475.
- Cohen, M., Reichenstein, M., Everts-van der Wind, A., Heon-Lee, J., Shani, M., Lewin, H. A., Weller, J. I., Ron, M., and Seroussi, E. (2004) Cloning and characterization of *FAM13A1*-a gene near a milk protein QTL on BTA6: evidence for population-wide linkage-disequilibrium in Israeli Holsteins. *Genomics*, pre-published online at: <http://authors.elsevier.com/sd/article/S0888754304000722>.
- Israel, C. and Weller, J. I. (2004) Effect of type I error threshold on marker-assisted selection in

dairy cattle. *Livest. Prod. Sci.* **85**; 189-199.

- Ron, M., Feldmesser, E., Golik, M., Tager-Cohen, I., Kliger, D., Reiss, V., Domochofsky, R., Alus, O., Seroussi, E., Ezra, E., and Weller, J. I. (2004) A complete genome scan of the Israeli Holstein population for quantitative trait loci by a daughter design. *J. Dairy Sci.* **87**; 476-490.
- Weller, J. I. and Soller, M. (2004) An analytical formula to estimate confidence interval of qtl location with a saturated genetic map as a function of experimental design. *Theor. Appl. Genet.* (In press).
- Weller, J. I., Feldmesser, E., Golik, M., Tager-Cohen, I., Domochofsky, R., Alus, O., Ezra, E., and Ron, M. (2004) Factors affecting incorrect paternity assignment in the Israeli Holstein population. *J. Dairy Sci.* **87**; (In press).