Gene Mapping – The Texas A&M University Angleton Project

J.F. Taylor*
Scott K. Davis, J.O. Sanders, J.W. Turner,
J.W. Savell, R.K. Miller, S.B. Smith

Abstract

In this paper, we discuss the application of maps of linked markers to identify genes of economic importance to the livestock industries and summarize significant results to date. We describe our approach to identify genes associated with beef carcass merit in the "Angleton project." We consider the potential benefits and limitations of marker-assisted selection and discuss issues of additivity of gene effects, epistasis, genotype by environment interaction and linkage disequilibrium. Progress toward the construction of a bovine Bacterial Artificial Chromosome library and results from initial screenings of this library for single-copy DNA sequences are presented.

Introduction

The term "economic trait locus" (ETL) is used to denote loci that influence traits of economic importance, be they qualitative (e.g., affected by only a few genes, such as diseases, horns and coat color) or quantitative (e.g., affected by many genes, such as growth, meat tenderness and milk production). The term "quantitative trait locus" (QTL) is used to denote a locus that is one of many that influence a trait. Presumably, no one would attempt to map a gene influencing a quantitative trait that had no economic significance and so all QTLs are potentially ETLs. The distinction between QTLs and qualitative trait loci is not terribly important except that *a-priori*, there is no guarantee that QTLs of large effect influencing a particular trait will exist; whereas, for a qualitative trait, the method of inheritance is usually quite well understood before any attempt is made to map the responsible gene. This explains why the majority of genes mapped in livestock species to date have been for qualitative traits; however, QTLs have also been detected for a number of traits, such as carcass merit and milk production. The detection of QTLs still segregating for milk yield in dairy cattle after many generations of intense selection attests to the power of this approach.

*J.F. Taylor, Department of Animal Science, Texas A&M University, College Station TX 77845.

Reciprocal Meat Conference Proceedings, Volume 47, 1994.

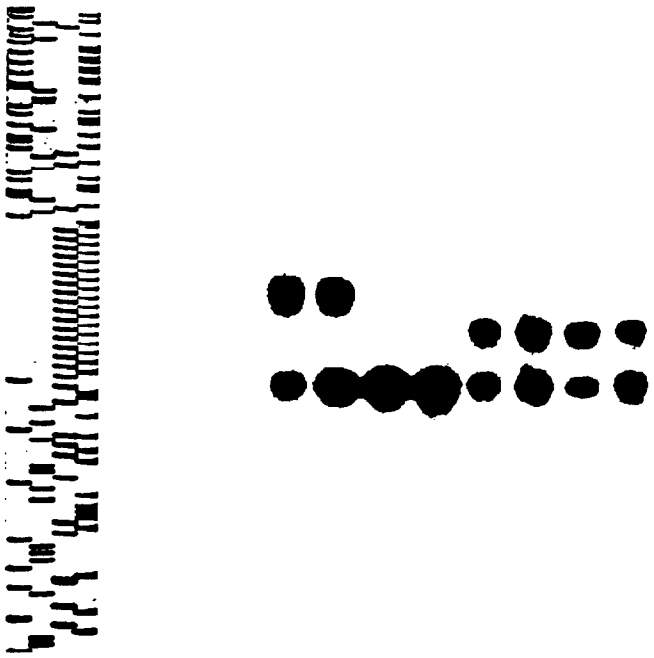
Markers

Historically, marker loci used by breeders and geneticists as selection aids were physiological, such as coat color and polled genes. A major limitation of these early studies was the lack of sufficient numbers of variable markers to saturate the genomes of the target species and thus genome research languished until the 1950's when erythrocyte antigens (blood groups) became available. In the 1970's, biochemical marker systems were augmented by the development of protein electrophoresis (isozymes). Both of these systems of markers were limited in their ubiquity and genomic distribution and along with certain of the blood group loci further suffered in utility due to their lack of polymorphism within breeds. It was not until the early 1980's that the advent of Restriction Fragment Length Polymorphisms (RFLPs) provided DNA-based markers that would allow the development of saturated maps of animal genomes.

Unfortunately, although these marker systems have been extremely useful in plants, they not been sufficiently polymorphic or uniformly distributed throughout the genome to allow for the development of highly-saturated genomic maps in animals. The solution came in the form of microsatellites, described by Weber and May (1989), which use the polymerase chain reaction to amplify a region varying for a simple dinucleotide or trinucleotide repeat. Figure 1 confirms the Mendelian inheritance of a microsatellite locus designated SRCRSP-4. The sire in lane 4 is homozygous for the 249 bp allele, the dam in lane 1 is heterozygous for the 249 bp and 259 bp alleles. The progeny in lane 2 inherited the 259 bp allele from the dam and the 249 bp allele from the sire; the progeny in lane 3 inherited one 249 bp allele from the dam and the other 249 bp allele from the sire. The dam in lane 8 is heterozygous for the 249 bp and 255 bp alleles; and the progeny in lanes 6, 7 and 8 all inherited the 255 bp allele from the dam and the 249 bp allele from the sire.

Microsatellites offer advantages over other types of genetic markers because microsatellites are ubiquitous and uniformly distributed throughout the genome. We estimate there to be a dinucleotide repeat on average every 30-50 kb in the bovine genome. Furthermore, the technique is repeatable, quick and cost-efficient. Microsatellite markers are not as useful as RFLPs for comparing gene maps across species, because the sequences detected are less well conserved. However, the relatively high level of polymorphism for microsatellites makes them ideal for the construction of linkage maps and the recently published sheep, pig and cow maps are based on microsatellite markers.

Figure 1
SRCRSP-4 Microsatellite Sequence and Genotypes Scored in Two Full-sib Families.



While markers and maps of linked markers can be used directly in multiple sire mating programs to identify the parentage of animals or for population genetic studies, the most interesting application of markers is as a ***tool*** to localize genes responsible for disease or for large differences in quantitative traits. In and of themselves, markers are not particularly interesting or valuable.

Species Linkage Maps

The degree to which markers on the same chromosome are inherited together is determined by their physical separation or map distance, which is measured in centimorgans (cM). Linkage analysis is used to determine map distances among linked markers from recombination events among families segregating for both markers. A saturated linkage map of markers is critical for the detection and localization of ETL to a chromosomal segment. The initial target has been to saturate each livestock species map to 20 cM (centimorgans), which would require about 180 evenly-spaced markers.

Following the discovery of microsatellites, there has been very rapid progress in the development of species genome maps. Bishop et al. (1994) published a map of 313 markers (280 were microsatellites) anchored to 24 of the 29 bovine autosomes and spanning 2,464 cM (about 88%) of the genome. Barendse et al. (1994) published a bovine map of 202 markers (144 were microsatellites) with a similar degree of genomic coverage. Rohrer et al. (1994) published a porcine map of 376 microsatellite and 7 RFLP loci spanning 1,997 cM. Crawford et al. (1994) published a map of 44 microsatellites and 8 RFLPs in the ovine. More than 500 microsatellite markers are currently available for bovine gene mapping studies and it is anticipated that this number will exceed 1,000 within the next 1 to 2 years. This number of markers should be sufficient to saturate the bovine map to an average of 5 cM, which should allow the localization of a number of ETLs to sufficiently small chromosomal regions to allow "chromosome walking" and "jumping" studies to begin in efforts to clone the responsible genes.

QTL Mapped by Markers

To date, the majority of ETL that have been mapped in livestock have been genes responsible for major qualitative traits. Georges et al. (1993a) have mapped the gene responsible for degenerative myeloencephalopathy (Weaver disease) in Brown Swiss cattle to chromosome 4 (TGLA116; Barendse et al. 1994). The gene responsible for this disease seems to have increased in frequency in the breed over time due to close linkage with a gene of large effect on milk production and selection for increased milk production within the breed. Although TGLA116 is estimated to lie only 3cM from the Weaver locus, the two loci are apparently in linkage equilibrium. This limits the application of this marker to determining the risk status of progeny from animals known *a priori* to be carriers for the condition.

Georges et al. (1993b) have mapped the gene responsible for the horn/polled condition to the centromere of bovine chromosome 1. Although there is no evidence that horned and polled cattle differ in productivity (Frisch et al. 1980), horned cattle are considered undesirable from the perspective of losses due to bruising and hide damage. The GMPOLL-1 and GMPOLL-2 markers of Georges et al. (1993b) map to approximately 15 cM from the horn/polled locus but do not flank the locus. Regardless of whether the marker and horn/polled loci are in linkage disequilibrium, these markers are probably too distant from the horn/polled locus to be useful predictors of heterozygote genotype status.

The Booroola *FecB* gene is a major gene in sheep which increases ovulation rate and litter size. Montgomery et al. (1993) have identified microsatellite markers that localize the *FecB* locus to ovine linkage group D which is homologous to a region on human chromosome 4q. The most closely linked marker, OarAE101, maps to 13 cM from the *FecB* locus. OarAE101 is probably too distant to be a useful predictor of genotype at the *FecB* locus; however, localization of this fertility gene will allow research to proceed towards the ultimate cloning of the gene.

The recently-discovered callipyge locus is responsible for muscular hypertrophy and associated leaness and improved feed efficiency in sheep. The gene is apparently distinct from the "double muscling" gene in cattle and the locus responsible for increased muscularity (and linked to PSE and halothane sensitivity) in swine. Cockett et al. (1994) have assigned the callipyge locus to ovine chromosome 18, using a battery of markers from bovine chromosome 21. Their placement of the callipyge locus is to the telomeric region of chromosome 18 in the interval between markers CSSM18 and TGLA122, 3 cM from CSSM18 and 17.5 cM from TGLA122.

There have been a number of studies that have found evidence of the existence of ETL of significant effect on milk production, growth and carcass traits in cattle which are inherited as quantitative traits. Beever et al. (1990) found associations

between markers and postweaning growth and carcass lean content in Angus cattle. Cowan et al. (1990) found an association between variation at the prolactin locus and milk production and Andersson-Eklund and Rendel (1993) found an association between variation at the amylase-1 locus and milk fat percentage. In fact, there has been more than 40 years of research into the detection of associations of linkage between blood group markers and ETL influencing milk production in cattle. The striking feature of this research is that a number of consistent associations between blood group loci and ETL of very large effect have been known for some time; however, to our knowledge, none of these markers have ever been employed in genetic improvement programs (Rocha et al., 1994).

Marker-Assisted Selection

Once DNA markers have been detected and their synteny and linkage relationships established, positional cloning studies can be designed involving families of animals segregating for the trait of interest to identify ETLs for marker-assisted selection schemes. The potential for success of such studies depends on:

1) The degree of saturation of the marker map, which determines the likelihood of detecting a QTL through a) having a marker in the region where the gene resides and b) the ability to detect the QTL, depending on its magnitude of observable effect,
2) The ability to use the marker to identify animals possessing the desirable QTL allele so that selection errors are avoided and
3) The cost effectiveness of the technology over more conventional breeding technologies.

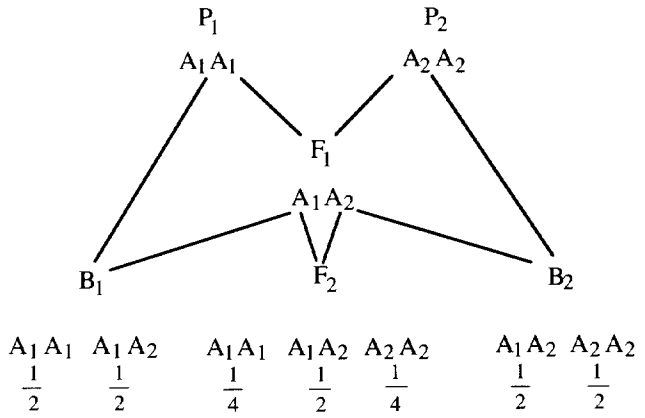
The degree of marker saturation of the various species linkage maps, particularly that of the bovine, is sufficient to detect ETL of useful but quite modest magnitude of effect. However, this is but the first (and possibly the easiest) step along the path towards implementing a marker-assisted selection (MAS) program. In the previous examples of mapped ETL, it should be clear that using markers to map an ETL to a chromosomal region does not invariably result in the identification (or cloning) of the responsible gene. In fact, a marker linked as closely as 3 cM may lie 3 million nucleotides away from the ETL and there may be 150 genes in this chromosomal interval. The process of "chromosome walking" and "jumping" from a linked marker to the gene itself requires the existence of large fragment DNA libraries (YACs or BACs) which can be aligned in a "contig," a lot of hard work, luck and the ability to both identify which clone contains the ETL and which of the genes within this clone is the ETL. This is not to say that it will be necessary to clone individual ETL to make use of them in breeding programs — but this may be necessary in some cases. This is also a necessary step to move from knowing that there is an effect due to an specific gene, to understanding why the gene causes an effect.

On a population basis, the utility of a marker to identify the desirable (or undesirable in the case of disease) allele(s) at an ETL depends on the degree of linkage disequilibrium between the marker and ETL. When two loci are in linkage equilibrium, the frequency of ETL alleles associated with each of the marker alleles is the same and the marker cannot distinguish the desirable ETL genotypes outside of a pedigree structure. This is apparently the case for the TGLA116 marker for the Weaver condition in Brown Swiss cattle, which is estimated to lie only 3cM from the Weaver locus (Georges et al. 1993a). In this case, linkage disequilibrium must be created between the loci (as in a family of progeny from a known carrier) before the marker can usefully be applied to identify the carrier progeny (with a level of confidence determined by the map distance between marker and disease loci). The amount of disequilibrium present between the marker and ETL loci depends on the map distance between them and the evolutionary history of the loci. If a recent mutation (or introgression) was responsible for the ETL allele, there is likely to be a significant degree of linkage disequilibrium and the marker will be useful on a population basis. Microsatellite loci evolve much more rapidly than ETLs and tend to be highly variable in livestock populations. Consequently, there may be useful levels of linkage disequilibrium between microsatellite and ETL loci; however, there appears to have been little research to date to elucidate this issue.

This discussion highlights an important distinction between the positional cloning and candidate gene approaches to the identification of ETLs. The positional cloning approach screens evenly-spaced markers along the species linkage map for associations with the ETL and has a high likelihood of localizing the ETL. However, once found, the marker may be in linkage equilibrium with the ETL in the target population, restricting its application to genotyping within segregating families until the ETL is ultimately cloned. On the other hand, the candidate gene approach, as exemplified by the case of bovine Calpastatin (CAST) which may be involved in determining beef tenderness, has a much reduced likelihood of identifying ETLs (depending on whether the researcher guessed correctly or incorrectly). However, if the candidate gene is indeed an ETL, it is more straightforward to identify (and design a test for) the genetic variation responsible for the desirable phenotype.

In their review of over 40 years of research into associations between blood group polymorphisms and milk production in cattle, Rocha et al. (1994) concluded that the enormous number of inconsistent findings between studies was, at least in part, due to: 1) inadequate experimental designs or statistical analyses, 2) trapping effects (presence or absence of ETL alleles across breeds, populations or experimental samples), 3) genetic architecture effects (epistasis) and 4) genotype by environment interactions. The last three points should also cause us some degree of concern for the implementation of MAS programs. If certain ETL alleles are "trapped" within breeds, will it be possible for us to introgress these alleles into other breeds (or species)? Many non-geneticists adhere to a reductionist view of genetics in which gene action is primarily additive. Rocha et al. (1994) refer to this as the "additive illusion" and point out that an ETL effect can be regarded as the "sensitivity coefficient" (holding the effects of all other loci constant) of an enzymatic step monitored by a marker. In reality, the process is highly non-linear and different genetic architectures will result in different ETL effects; to the extent that in some architectures, a locus may have no effect on a trait. Similarly, the magnitude of effect of an ETL may be dependent on the environment (and perhaps also the genetic

Figure 2
Angleton Project Mating Design.



architecture considering the term genotype in its broader context). Of course, Sewall Wright already knew all of this by 1939 when he said "Separate genes cannot be considered as good or bad. Only the genotype as a whole can be so classified (and even this is relative to environmental conditions)." While we are strong advocates for research into the detection and utilization of ETLs in MAS schemes, we do not believe that the task in front of us is going to be easy!

Materials and Methods

The Angleton Project

Texas A&M University and the Texas Agricultural Experiment Station have dedicated the Angleton Research Station and all of its cattle and technical resources to the development of a resource herd segregating for carcass merit traits. The Angleton Research Station has 218 Brahman x Hereford cows of breeding age that serve as recipients for a MOET program. This program is currently producing backcross and F_2 progeny according to a double reciprocal backcross design with Brahman and Angus parents as illustrated in Figure 2. The mating scheme includes four replicates of all possible reciprocal crosses as follows:

A♂ X A·B♀	A·B♂ X A♀
A♂ X B·A♀	A·B♂ X B♀
B♂ X A·B♀	B·A♂ X A♀
B♂ X B·A♀	B·A♂ X B♀

where A·B♀ denotes an F_1 female, derived from mating an Angus sire to a Brahman cow. This design requires a minimum of 16 bulls and 16 donor cows. Cows are implanted up to three times per breeding season to maximize the conception rate. The average pregancy rate per transfer from 392 transfers in 1991-92 was 57.9%. The MOET program is designed to run for five years to produce approximately 600 calves, with 15 to 20 full-sibs per family.

The goals of the project are to develop a linkage map of markers saturated to 20 cM in these cattle families and to screen these markers for associations with ETLs for carcass merit traits. The project has produced 394 progeny, there are 113 confirmed pregnancies, 477 frozen embryos and 149 head

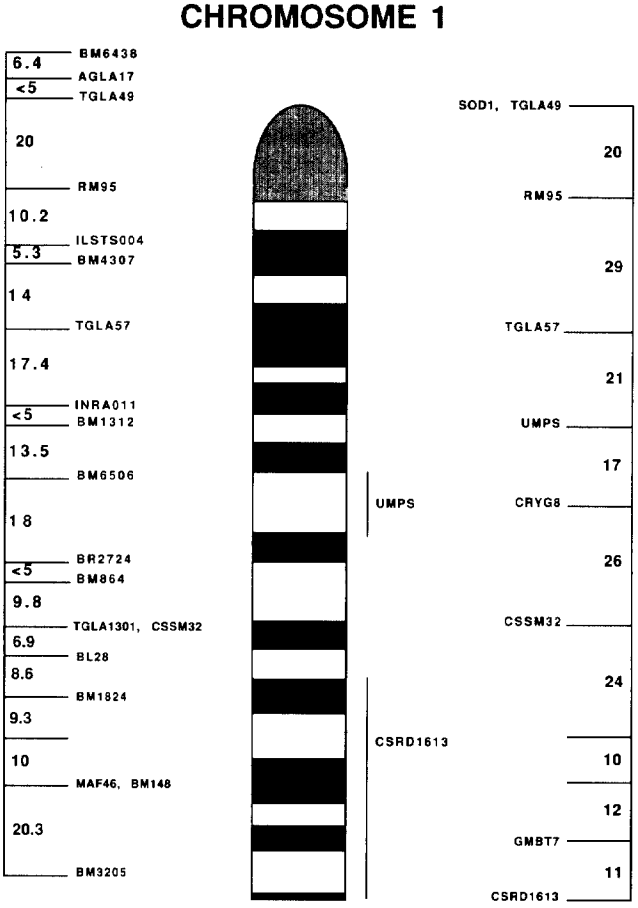
Table 1. Angleton Station Project Mating Scheme.

Family	Type	Calves	Pregnancies	Embryos	Slaughtered
1	3/4 Angus	17	0	15	0
2	"	18	0	15	4
3	"	1	0	0	1
4	"	8	8	12	0
5	"	10	9	8	6
6	"	14	7	33	9
7	"	20	0	11	0
8	"	14	9	31	0
9	"	16	1	7	10
10	"	11	11	5	5
11	"	21	0	0	19
12	"	14	0	11	14
13	"	10	10	18	5
14	"	9	7	0	1
15	3/4 Brahman	9	0	0	0
16	"	15	0	14	0
17	"	10	1	16	10
18	"	16	5	20	12
19	"	13	0	36	0
20	"	12	0	14	6
21	"	15	0	20	4
22	"	11	8	0	1
23	"	2	0	0	2
24	"	17	6	0	4
25	"	13	0	16	5
26	"	1	0	0	0
27	"	13	0	0	9
28	"	14	0	13	0
29	"	8	24	0	1
30	"	6	6	8	3
31	"	5	1	0	0
32	F_2	13	0	13	0
33	"	0	0	47	0
34	"	15	0	13	15
35	"	1	0	0	1
36	"	2	0	0	2
37	"	0	0	37	0
38	"	0	0	9	0
39	"	0	0	6	0
40	"	0	0	11	0
41	"	0	0	3	0
42	"	0	0	15	0
Total		394	113	477	149

have been slaughtered for carcass data. Table 1 indicates the distribution of animals by family type. While all progeny will be slaughtered, a DNA library will be maintained representing every animal involved in the breeding program to allow for additional analyses as new markers become available. Surplus embryos of each family will be stored to provide a resource for future studies.

The Angleton progeny are recorded for horn/polled status, coat color, coat speckling, structural, health, weight for age and growth characteristics. All progeny are carried through feedlot and carcass evaluation stages, with slaughter after

Figure 3
Sex Average Linkage Map of Bovine Chromosome 1.



about 150 days on feed. Carcass evaluation data are obtained at slaughter, describing maturity, marbling, quality grade, yield grade, fat thickness, ribeye area, percentage kidney-pelvic-heart fat and carcass weight. Tissue samples are brought to the Meats and Muscle Biology Laboratory at Texas A&M University for determination of extractable lipids, moisture content, protein content, collagen analysis, 9-10-11th rib dissection, Warner-Bratzler shear force, descriptive sensory analysis (taste panel), fragmentation index, calcium-dependent protease analysis, sarcomere length, fatty acid and cholesterol composition of *longissimus dorsi* and stearyl coA desaturase and fatty acid elongase activity in *longissiumus dorsi*. We are also recording the following measures for intermuscular and subcutaneous adipose tissue: rates of incorporation of acetate, glucose and palmitate; activity of fatty acid elongase and stearyl coA desaturase; fatty acid composition; cholesterol content; and cellularity.

Results and Discussion

Markers Scored

We have scored 11 blood group loci, 4 isozyme loci and 120 dinucleotide repeat microsatellites in 280 Angleton project grandparents, parents and progeny. These markers map to all of the bovine autosomes and saturate chromosomes 1, 2, 3, 5, 6, 21 and 23 to less than 20 cM. The linear distance accounted for by these markers is approximately 1,400 cM or 50% of the bovine genome. Figure 3 illustrates the sex average linkage map of bovine chromosome 1.

Qualitative Loci Mapped

The horn/polled condition, black/non-black coat color and the color speckling gene originating in the Gir breed are segregating within the Brahman backcross families of this experiment. We confirm the finding of Georges et al. (1993b) that the horn/polled locus maps to the centromere of bovine chromosome 1. While BM6438 was not informative in any of these families and AGLA17 was not scored, TGLA49 is linked to the horn/polled locus at a map distance of 24.3 cM and with a logarithm of odds (lod) score of 5.03. There were a total of 75 progeny in 6 informative families and all families generated positive lod scores. As with the analysis of Georges et al. (1993b), we assumed autosomal recessive inheritance of horns with complete penetrance of the homozygous recessive horned genotype and penetrances of 0 for the other genotypes.

The gene responsible for black versus red coat color in Angus cattle also segregated in these families generating black versus non-black (red, white, dun or brindled) coat color phenotypes. Assuming autosomal dominant inheritance of this trait, we have a microsatellite marker linked to coat color at a map distance of 13 cM and with a lod score of 5.54. The micro-satellite resides within a linkage group of microsatellites that have been assigned to two separate bovine chromosomes. We are in the process of scoring markers in the appropriate chromosomal regions of these markers in order to resolve the map position of the black/non-black locus.

No markers were found linked to the speckled coat color locus.

ETL Associations

Calpastatin down-regulates the calpain enzyme which is associated with the post-mortem tenderization of beef. High levels of calpastatin measured at day 1 post-mortem are associated with increased Warner-Bratzler shear forces of steaks cut from the *longissimus dorsi* aged to 14 days. A RFLP marker for CAST localizes the gene to a terminal region of bovine chromosome 7. However, it is not yet apparent whether this marker detects variation within CAST that is associated with variation in calpastatin activity or Warner-Bratzler shear force. We plan a collaborative study to score the CAST RFLP in the Angleton families; however, in the interim, we scored ILSTS006, a microsatellite locus which maps 10.7 cM (Bishop et al., 1994) from CAST. In a preliminary analysis involving a subset of data on 64 animals for which we had both genotypic and phenotypic data, we found no associations between ILSTS006 and either calpastatin or Warner-Bratzler shear force. This suggests that the CAST locus itself may not be the critical regulator of calpastatin activity and that other as yet unidentified loci may be responsible.

Phenotypes for slaughter weight, hot carcass weight, dressing percent, rib eye area, kidney-pelvic-heart fat, fat, skeletal

maturity, lean maturity, marbling, quality grade, yield grade, rib eye sarcomere diameter, moisture, fat % and calpastatin activity on the first 47 slaughter progeny were analyzed for associations with 55 markers. In 880 tests of statistical significance using a family design analysis, 28 associations were detected at less than the 1% significance level. Based on the Type I error rate, we would expect only 9 spurious associations and these results encourage us to believe that some of these associations will be validated when the data on the 153 cattle slaughtered to date are analyzed this summer. Of particular interest, we found one microsatellite to be associated with kidney-pelvic-heart fat, marbling, quality grade and yield grade in a breed-specific manner, where the presence of a Brahman allele at this marker reduced fatness relative to the presence of an Angus allele. A second microsatellite was associated with skeletal maturity, marbling and fat %. By the end of 1994, marker data will be available for every bovine chromosome and carcass data will be available from over 200 animals. At this point, we will be able to complete robust statistical analyses identifying the genes controlling carcass merit in beef cattle.

Bacterial Artificial Chromosome Library

Sufficient markers have now been placed within the framework of the bovine linkage map that chromosomal regions containing ETLs have begun to be identified. To move beyond this point, a library of large DNA clones must be available to researchers for physical mapping of key regions identified by genetic mapping and for the positional cloning of ETLs. We have initiated the construction of a bovine Bacterial Artificial Chromosome (BAC) library (Shizuya et al. 1992) targeted at a five-fold coverage of the bovine genome, an average insert size of 150 kb and a >95% probability of containing a specific single-copy sequence (PSCS). We have chosen to construct a BAC library rather than a Yeast Artificial Chromosome (YAC) library (Libert et al. 1993) because BACs offer the advantages of efficiency of construction, ease of handling, cloned DNA, very low chimerism and very low deletion rate relative to YACs. These properties outweigh the disadvantage of a smaller insert size for BACs. We have constructed 18,000 BAC clones with an average insert size of 142 kb, representing an 85% genomic coverage and with approximately a 60% PSCS.

A PCR systematic screening of the library for five single-copy loci was initially conducted at 14,400 clones corresponding to a 51% PSCS. This screen resulted in the isolation of two clones containing the ETH225 microsatellite locus, two clones containing the glucocerebrosidase gene and one clone containing the 3-β-hydroxy-5-ene steroid dehydrogenase gene. At this stage, the library did not contain the ETH1113 or BOVIRBP microsatellite loci or the butyryl cholinesterase gene. A more recent screening at 60% PSCS has yielded three clones containing the MHC I gene, two clones containing the MHC 21OH gene and 10 clones containing members of the MHC Heat Shock Protein family of genes. The library did not contain Butyrophilin or Calpastatin genes at this stage of development. Glucocerebrosidase and 3-β-hydroxy-5-ene steroid dehydrogenase map to different arms of human chromosome 1; and we have identified three polymorphic dinucleotide microsatellites within one clone, containing glucocerebrosidase and one polymorphic and one monomorphic microsatellite within the clone, containing 3-β-hydroxy-5-ene steroid dehydrogenase. These markers will allow us to map these loci by linkage within the framework of the bovine gene map.

Not only are the results of our BAC library screening consistent with the theoretical probabilities of containing single-copy sequences, but they reveal the power of this approach for mapping Type I anchor loci, either through use of the clones for fluorescent in-situ hybridization or through the identification of polymorphic microsatellite loci within the clones containing these loci. Although we have limited data to date, our success in identifying polymorphic microsatellites within each of the clones containing glucocerebrosidase and 3-β-hydroxy-5-ene steroid dehydrogenase supports both the ubiquity (a polymorphic microsatellite every 30-50 kb) and uniformity of distribution of microsatellites.

Conclusions

The sophisticated genetic tools now available will allow us to enhance our knowledge of physiology and of basic metabolic and biochemical mechanisms underlying the productive differences between individuals, lines and breeds. As our knowlege of these physiological, metabolic and biochemical processes increases, we will be able to achieve significant increases in animal productivity.

In 1939, Sewall Wright said "Suppose that we were given a reasonably complete map of all of the chromosomes, showing the location of all important genes affecting a certain character as well as of convenient marker genes. What could we do with it?" That day is fast approaching and we had better attempt some answers to this question.

ACKNOWLEDGEMENTS

J.F. Taylor and S.K. Davis were partially supported by a grant from the Beef Industry Council of the National Live Stock and Meat Board and by USDA CSRS grant No. 92-34267-6962.

References

Andersson-Eklund, L.; Rendel, J. 1993. Linkage between amylase-1 locus and a major gene for milk fat content in cattle. Animal Genetics 24:101.

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.; Hetzel. D.J.S. 1994. A genetic linkage map of the bovine genome. Nature Genetics 6:227.

Beever, J.E.; George, P.D.; Fernando, R.L.; Stormont C.J.; Lewin, H.J. 1990. Associations between genetic markers and growth and carcass traits in a paternal half-sib family of Angus cattle. J. Anim. Sci. 68:337.

Bishop, M.D.; Kappes, S.M.; Keele, J.W.; Stone, R.T.; Sunden, S.L.F.; Hawkins, G.A.; Toldo, S.S.; Fries, R.; Groz, M.D.; Yoo, J.; Beattie, C.W. 1994. A genetic linkage map for cattle. Genetics 136:619.

Cockett, N.E.; Jackson, S.P.; Shay, T.L.; Nielsen, D.; Steele, M.R.; Green, R.D.; Georges, M. 1994. Chromosomal localization of the *callipyge* gene in sheep (*Ovis aries*) using bovine DNA markers. Proc. Nat. Acad. Sci. (In press).

Cowan, C.M.; Dentine, M.R.; Ax, R.L.; Schuler, L.A. 1990. Structural variation around prolactin gene linked to quantitative traits in an elite Holstein sire family. Theor. Appl. Genet. 79: 577.

Crawford, A.M.; Montgomery, G.W.; Pierson, C.W.; Brown, T.; Dodds, K.G.; Sunden, S.L.F.; Henry, H.M.; Ede, A.J.; Swarbrick, P.A.; Berryman, T.; Penty, J.M.; Hill, D.F. 1994. Sheep linkage mapping: Nineteen linkage groups derived from the analysis of paternal half-sib families. Genetics:137:573.

Frisch, J.E.; Nishimura, H.; Cousins, K.J.; Turner, H.G. 1980. The inheritance and effect on production of polledness in four crossbred lines of beef cattle. Anim. Prod. 31:119.

Georges, M.; Dietz, A.B.; Mishra, A.; Nielsen, D.; Sargeant, L.S.; Sorensen, A.; Steele, M.R.; Zhao, X.; Liepold, H.; Womack, J.E.; Lathrop, M. 1993a. Microsatellite mapping of the gene causing Weaver disease in cattle will allow the study of an associated QTL. Proc. Natl. Acad. Sci. USA 90:1058.

Georges, M.; Drinkwater, R.; King, T.; Mishra, A.; Moore, S.S.; Nielsen, D.; Sargeant, L.S.; Sorensen, A.; Steele, M.R.; Zhao, X.; Womack, J.E.; Hetzel, J. 1993b. Microsatellite mapping of a gene affecting horn development in *Bos Taurus*. Nature Genetics 4:206.

Green, P.; Falls, K.; Crooks, S. 1990. Documentation for CRI-MAP, version 2.4 (3/26/90). Department of Genetics, Washington University School of Medicine. St Louis, MO p. 15.

Libert, F.; Lefort, A.; Okimoto, R.; Womack, J.; Georges, M. 1993. Construction of a bovine genomic library of large yeast artificial chromosome clones. Genomics 18:270.

Montgomery, G.W.; Crawford, A.M.; Penty, J.M.; Dodds, K.G.; Ede, A.J.; Henry, H.M.; Pierson, C.A.; Lord, E.A.; Galloway, S.M.; Schmack, A.E.; Sise, J.A.; Swarbrick, P.A.; Hanrahan, V.; Buchanan, F.C.; Hill, D.F. 1993. The ovine Booroola fecundity gene (FecB) is linked to markers from a region of human chromosome 4q. Nature Genetics 4:410.

Rocha, J.L.; Taylor, J.F.; Sanders, J.O.; Cherbonnier, D.M.; Lawlor, T.J. 1994. Blood groups and milk and type traits in dairy cattle: After forty years of research. J. Dairy Sci. (In preparation).

Rohrer, G.A.; Alexander, L.J.; Keele, J.W.; Smith, T.P.; Beattie, C.W. 1994. A microsatellite linkage map of the porcine genome. Genetics 136:231.

Shizuya, H.; Birren, B.; Kim, U.-J.; Mancino, V.; Slepak, T.; Tachiiri, Y.; Simon, M. 1992. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escheria coli* using an F-factor based vector. Proc. Natl. Acad. Sci. USA 89:8794.

Weber, J.L.; May, P.E. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet. 44:388.

Discussion

M. Koohmaraie: Two or three years ago, we published data showing that the calpastatin gene is polymorphic; this was the first indication that there is an association between polymorphism in the calpastatin gene and tenderness. We have found out, and obviously you folks have too, that there is no relationship, no association between the calpastatin polymorphism and tenderness. I just want to make it clear that just because there is no association between these two traits, that does not say anything about the fact that the protein is associated with meat tenderness.

J. Taylor: Yes, I would agree. I would also like to say that the polymorphism you are detecting is a "rough look" and is not highly variable in the population. So *a priori*, you may be in a position that your polymorphism may not be detecting the structural variation within the gene or the promoter-enhancer regions of the gene that is causing the differences in calpastatin enzyme activity that you measure. They want post-mortem. That's another possibility. That's why we have tried to see if calpastatin is in our library. If it is, we can pull it out of the library, find some microsatellites for the gene that should be more polymorphic, then come back and look closely at that one issue.

Koohmaraie: Suppose that four or five years from now, we come up in your reference families and our reference families with 4 or 5 markers for meat tenderness. Can those be applied to different families or are the markers only good for the reference families in which they were discovered?

Taylor: That comes back to two things. It comes back to linkage disequilibrium which I have been discussing here. Is there disequilibrium between the QTL and the marker right across the population? If there is, you will be able to use it across different families. You are not always going to be right, but you are going to be right more often than wrong. And it will be useful. If you don't have disequilibrium, you have to use the marker within families. The phase has to be established within each family before you can go ahead and use the marker. If that's the case (like Weaver, for example), there's a 3 cM marker right there for Weaver disease and I believe it is being used. I'm not sure if that's being run by anyone within families to detect carriers of the Weaver disease. But if you have that large a problem, what you must do is try to locate the gene. So you must go into that region (the 3 cMs) and try to figure out which gene is responsible for the Weaver disease.

I guess what I didn't explicitly say is that there's a lot of work involved. All you have to do is look at the resources that are being put into human gene mapping. For example, finding the gene responsible for Huntington's disease. You know the marker associated with it. I won't give you correct dates but I'll say 1979 was when the marker was discovered. But there was a period of many years between when the marker was identified and the gene itself was located.

C. Beatty: I think it's disequilibrium. I think if you have enough of a cluster of markers close to that particular locus (let's say 5 cMs or less) the chances of your being heterozygous in a number of families is quite high.

Taylor: The comment was well put to this audience (and any other audience). It is remarkably tedious work to specifically engineer or re-engineer the positional cloning efforts for any given interval that you need to undertake. It took literally 10 years of remarkably highly-focused research from some of the largest U.S. labs to move from a 2 cM linkage of a marker in the Huntington's region to the actual cloning of the disease. The physical genomics are quite complicated and almost all of the large efforts have been complicated by some property of genomics.

I'd like to give you another hopefully encouraging note: Human genetics has become almost naively enthusiastic about the possibility of mapping quantitative traits. Politically speaking, I guess it would be correct to just call humans ETL's, but there are a number of pharmaceutical companies interested in them, I assume, for the same reasons that would make them ETL's as well as QTL's. Imagine finding a gene for hypertension; it would be an ETL in humans. One thing that I would encourage you to consider is that by minor manipulations and mating structure, you can make powerful use of genetic maps, even of the density of 10 cM's, to localize regions that may have polygenes.

Simulation studies where we take inbred populations, or actually make backcrosses where we can look at inbred chromosomes on a background population, suggest that you can artificially generate linkage disequilibrium that will extend as large as 20 cMs. In fact, by manipulating the crosses, you may be able to find or at least identify the regions and traits. It won't necessarily diminish the effort to find them, or in the more outbred populations, to perhaps use them. But then you could use link markers in the segregating lines.

Beatty: As a general case, in terms of marker-assisted selection, you don't have to find the gene for us to establish value relative to what's already out there. You must be fortunate enough to have linkage disequilibrium present for what you're most interested in. Assuming that, then you must have markers close enough to the ETL that they consistently segregate with it. Then it would have to be heterozygous. Plus, the markers would have to have enough allelic variation associated with them. But they are useful in any number of animals that you would hopefully want to breed. Particularly, a number of sires that would be carrying this particular phenotype.

Taylor: I want to say that I'm being the devil's advocate here. I wouldn't be doing this research, investing this much time in my life trying to get this project done, if I didn't believe in it! I think there are going to be new technologies that the human genome initiative will produce for us, because they are going through exactly the same problems. They'll reach the point where they'll want to avoid going into this region that has 150 genes in it to try to localize a gene. There are new techniques like selectable fragment libraries where you can select the common elements or differences in a library, that I think are really going to accelerate that process. But we don't have them yet. I have a graduate student in my lab working on some of these ideas. If we can make these work, the jump between localizing a gene to a region of a chromosome and finding the gene itself is going to be much faster. It will be something that we can do in a couple of months. But we're not there yet.

K. Buetow: I think that in the human genome project side, we have been able to successfully convince our funding agents, who are benefactors, that it is remarkably inefficient and wasteful for each of us to generate these primary agents on an interval-by-interval basis. One of the huge things we have learned (or that has been now demonstrated) is that if the government or some other insightful group supplies the up-front money to generate the physical library, for example, the BAC contig is being discussed or a YAC contig, etc. In fact, the benefits will be dramatic but very difficult to demonstrate. We have learned that having all these reagents in the can when one wants to investigate something is a huge win, rather than each lab having to generate them *de novo*. So I encourage you to think of the long-term productivity of investment in these people's efforts, because it's demonstrated to be a big win in human genetics. One of the reasons it took 10 years to get the Huntington's disease gene is that all of the reagents and even the methods literally had to be invented, to go from that linkage to the locus. Now that we actually have these reagents in hand, one just dials up on the Internet to find out what YACS contain Genathon markers and actually what contig has already been constructed in that region when I make my linkage map. The availability of those regions saves thousands of hours and is estimated to save hundreds of thousands of dollars in terms of research time. So the up-front investment really pays off.

L. Orme: Let me change the subject a little. Ken, the question is for you. What are the moral issues associated with this human program and what do you foresee in the future?

Buetow: I think there are significant societal issues associated with the human genome initiative. I think that within the next few years, those people who are working with it will continue to be very sensitive to those issues. It's quite clear that shortly we will have the ability at birth to identify significant causes of morbidity, ranging anywhere from obesity to heart disease to cancer susceptibility. Literally, the work being done in my lab right now is to identify genes that will predict your risk of cancer. You will have that information at birth, or the information will be theoretically available at your birth. The consequences of knowing that information have both personal and professional ramifications, in terms of health insurance, employment, in terms of all those classes of questions.

As a cancer researcher, I tend to think this information is empowering at a certain level. Let me use an example I find instructive. I open my talk by saying that only one in ten smokers develops lung cancer. In fact, that is a true statement. The majority of people can probably smoke safely at some level; of course, you may get heart disease, emphysema or some-

thing else. But you probably won't get lung cancer if you smoke. However, one person out of ten has probably a very high chance of developing lung cancer!

Where we see these techniques as empowering is being able to share that information with people. It must be a powerful incentive to a person when you have told your audience "Don't smoke, you have a one in ten chance of developing lung cancer" and then continue to say "But you, sir or madam. If you smoke, yours is not a one-in-ten chance but a 90% to 99% chance of **developing** lung cancer!"

Now with that said, there are clearly issues that society will have to wrestle with. Who should have access to this information? Is it appropriate for insurance companies to discriminate by changing or increasing premiums based on a person's genetic constitution? Now it is acceptable to deny insurance to people with pre-existing conditions. We will have to make decisions as a society as to how this information should be used.

I believe there's a silver lining to this type of dark cloud. The one advantage we have in genetic testing is that we all have something bad. Population genetics tells us that at the simplest of levels, we all carry three recessive lethals. We all carry three genes that will be sufficient to **not have** viable offspring with an inappropriate mate. In addition, we all probably carry at least 6 to 12 other susceptibility factors that could make us more susceptible to lung cancer, heart disease, or one of the other diseases.

So if we do open up that type of discrimination (a real can of worms), we must recognize we'll be discriminating not only against the person who shouldn't be covered by insurance based on his/her genome type. We must be ready to say "In fact, I'd better be prepared to say that I won't have any insurance, or I won't be employed."

Maybe there will be enough tension that it won't be possible to discriminate against people since we are all in the same boat.

I think these are all issues that must be discussed. One of the reasons I am enthusiastic about speaking to this type of audience is that I think these issues don't get discussed among knowledgable people like yourselves; people who are familiar with genetics at a very sophisticated level, to communicate both the strengths and weaknesses of these sorts of techniques and approaches.

M. Dikeman: If we knew all this information, what would we do with it? You pretended that you really did not have an answer for that, which I'm not sure I believe. Would you take a shot at telling us what we could possibly do with this kind of information, and whether you think it might be more applicable to traditional genetic selection for animals, or maybe the development of transgenic-type animals?

Taylor: If I had that information, I'd have a tool that would answer some questions about genetic architecture. I'm talking about non-additive gene action as we've been discussing with the calpastatin markers. I think there's another gene that is "beating up" on calpastatin. So you have an interaction. So it depends on your genome type at both the calpastatin and another regulating locus. It would give us an answer to those things and also genotype by environment interaction.

What we've learned over the years of animal breeding is that with some loci, you do get a lot of genotype by environment interaction or a lot of epistasis. For other traits, you don't. If you think about horns, you have to be homozygous recessive for horns; you will have horns no matter what other genes you have, no matter what environment you are in.

So my first response to your question is that we must be looking for loci like that. If there are loci influencing tenderness, marbling or other traits that are economically important and really difficult for us to get at by any other method, they are the kind of traits we must look at. But don't blindly think "If I find this gene that has this effect in this breed, if I move that gene into Brahman it will work just as well in Brahman." It may not, there are just some things you must be aware of before you do those kinds of studies.

The second part of my response would be to ask if you gave all the cattlemen in the world a list of EPD's for 10,000 traits, what would they do with them? What would I do with them? I think it would come down to saying that we must try to do some kind of prioritization as to what really is most important. Let's try to find the genetic combinations that will work and start going about making them. I think it is an interesting question because it is like everything I have learned in research, the more afraid I become of what I don't know!

Koohmaraie: I'd like to make a closing comment. My only interest is tenderness. The heritable component of tenderness is only about 30%. This means that genetics is only 30%. So if we find the genes for it, do everything we can, we have probably only accounted for 30%. Seventy percent is environment and probably genotype by environment interaction. I don't know what you folks think about it.

Taylor: Markers will let you get at the non-additive components too. If you are real lucky, you can get a dominance effect, you can get epistatic effects potentially. So we may be able to get 50% of the total variation in tenderness taking a marker approach if we can identify all the genes, although we probably can't. You know that for milk production, 25% of the variation is genetic. For tenderness in cattle, 30% is due to the additive effects of genes. If you look at the changes we've made in milk production in 30 or 40 years, it is staggering. So it may be enough.