MOLECULAR APPROACHES TO GENETIC IMPROVEMENT

R. Mark Thallman
U.S. Meat Animal Research Center
Agricultural Research Service, U.S. Department of Agriculture
Clay Center, NE 68933

Benefits of Genetic Testing

- Obtain evaluations earlier in the life cycle.
- Increase accuracy of selection, especially for traits that are expensive to measure, sex-limited, or measured postmortem.
- Capture more benefit from each phenotype that is measured.
- Increase opportunity to select for traits with antagonistic genetic relationships (e.g., birth weight and growth rate).

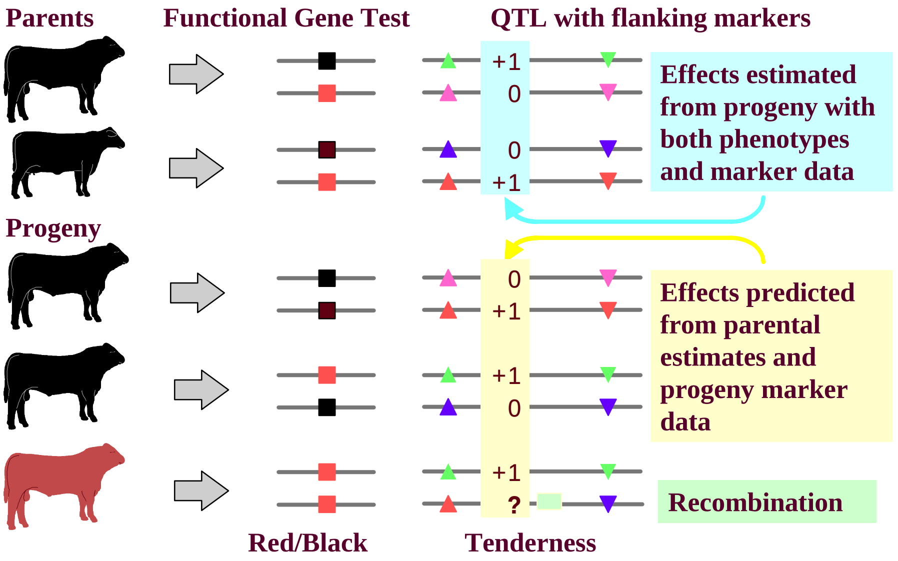
Realistic Expectations

- DNA testing can increase the amount of information that each phenotype contributes.
- DNA testing can reduce the number of phenotypes needed, but DNA testing can not replace phenotypes.
- DNA testing will probably make cattle breeding more complicated, not easier.

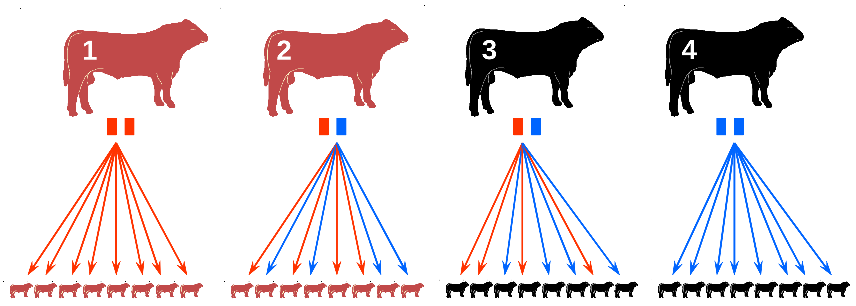
Types of Genetic Tests that Could Be Available

- Linked Markers – highly polymorphic markers within a few cM of, and generally assumed to be in linkage equilibrium with, the functional polymorphism
- Functional Tests – polymorphisms that directly affect phenotypes of interest
- Association Tests – polymorphisms in high linkage disequilibrium (and generally physically close to) one or more functional polymorphisms that affect phenotypes

Functional Tests vs Linked Marker Tests



Linked Marker Tests



Advantages of Linked Marker Tests

- They are relatively inexpensive to develop
- We have considerable experience in using them, at least in QTL detection experiments
- They tend to be highly polymorphic

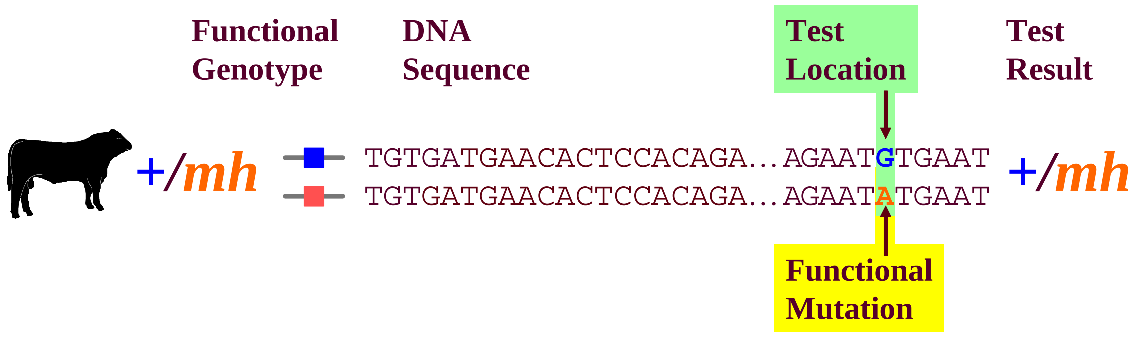
Challenges in Applying Linked Markers

- Must establish phase between markers and QTL within each family
- Must estimate QTL effects within family
- Therefore, marker data must be collected on a substantial number of individuals in each family
- The required statistical analysis is computationally demanding
- Breeders cannot easily interpret the test results themselves

Commercialization of Linked Markers Depends On:

- Testing technology that is inexpensive enough to be applied on a whole-herd basis
- Existence of the infrastructure to incorporate the information into NCE
- Not likely to be used widely

Functional Tests



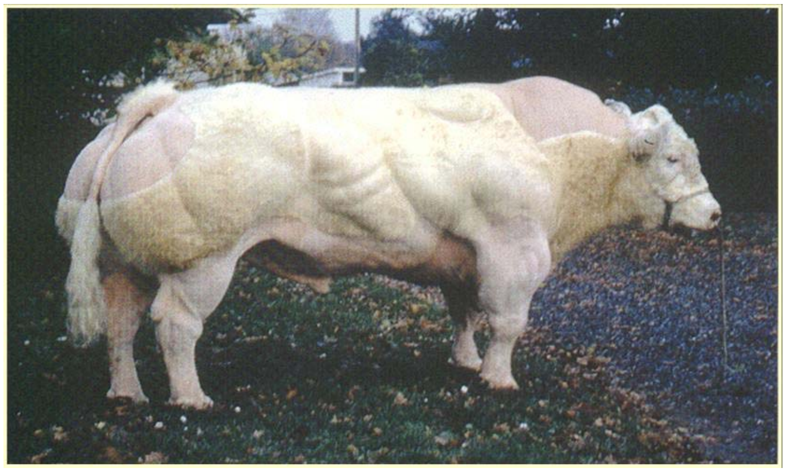
Advantages of Functional Tests

- Test results are easier to interpret
- Does not require extensive testing of relatives
- *We don't have as much experience using them in livestock, so their disadvantages are not as apparent as those of linked markers.*

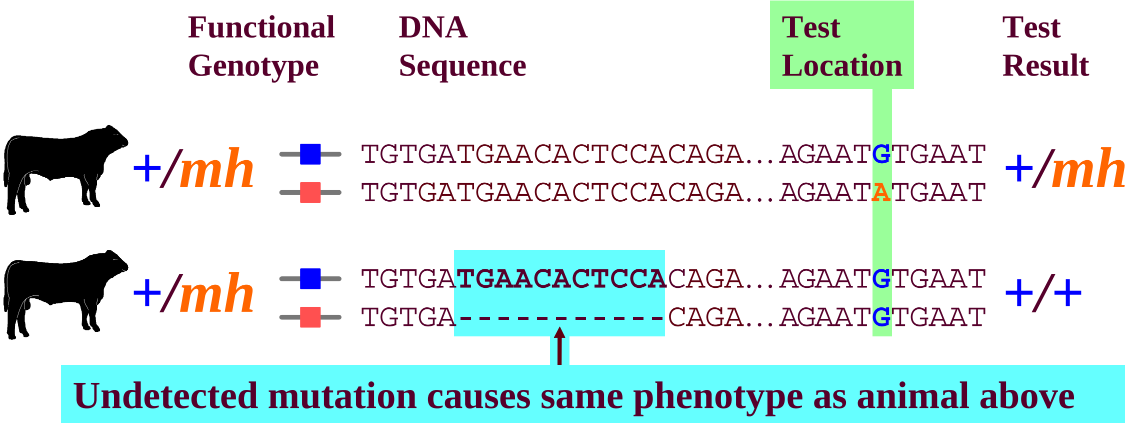
Challenges in Applying Functional Tests

- Very expensive to develop
- Some will not detect all of the functionally different alleles in the population.
- Different size of effect exists between breeds or production systems

Myostatin is an example in which a functional test is available.



Undetected Functional Alleles



Undetected mutation causes same phenotype as animal above

- These undetected functional alleles are likely to be common and will only be detected if phenotypes continue to be collected and associated with DNA test results.

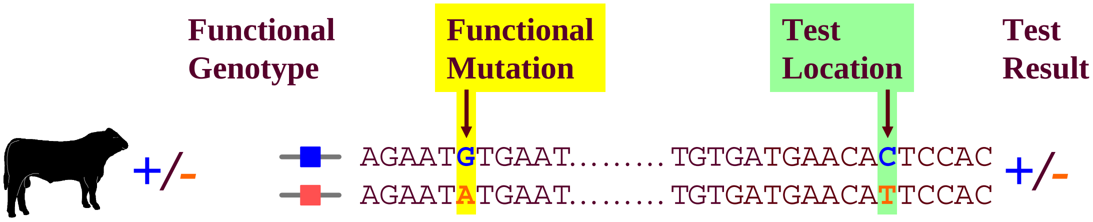
Consequences of Undetected Functional Alleles

- May result in decreased accuracy of the test, depending on the frequency of the undetected alleles
- The inaccuracies may be infrequent, but very large, and are likely to go undetected for a long time. When they are recognized, the consequences could be considerable
- Could cause underestimation of the effect and degree of dominance of the gene in populations with undetected alleles

Different Sizes of Effects Between Breeds will Present Challenges in Incorporating Test Data into NCE

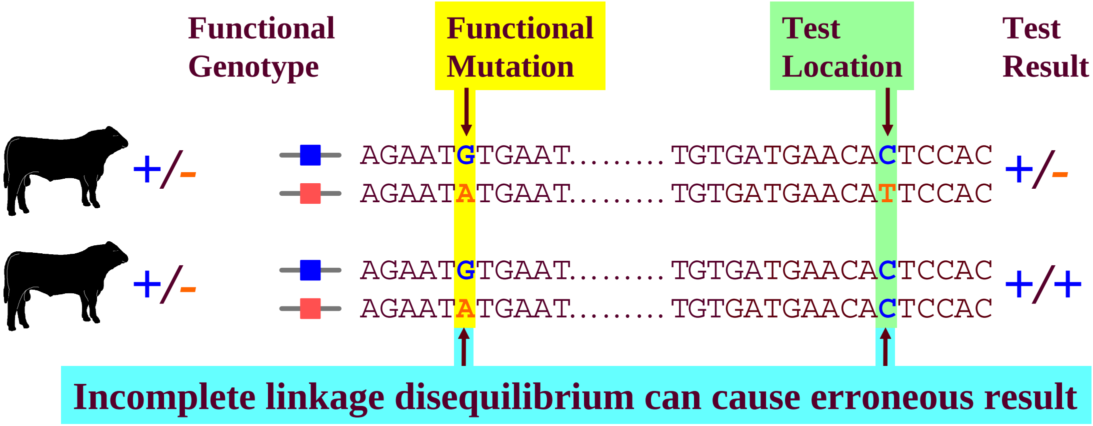
Genotype	+/+	-/-
Breed A	+5	-5
Breed B	+8	-8

Association Tests



- Association is due to linkage disequilibrium
- Share many of the advantages and disadvantages of functional tests
- May be very difficult to distinguish from functional tests
- Easier to develop because it is not necessary to prove that the polymorphism being tested causes the effect on phenotype

Challenge in Applying Association Tests



- When this situation is detected, more SNP should be added to the test

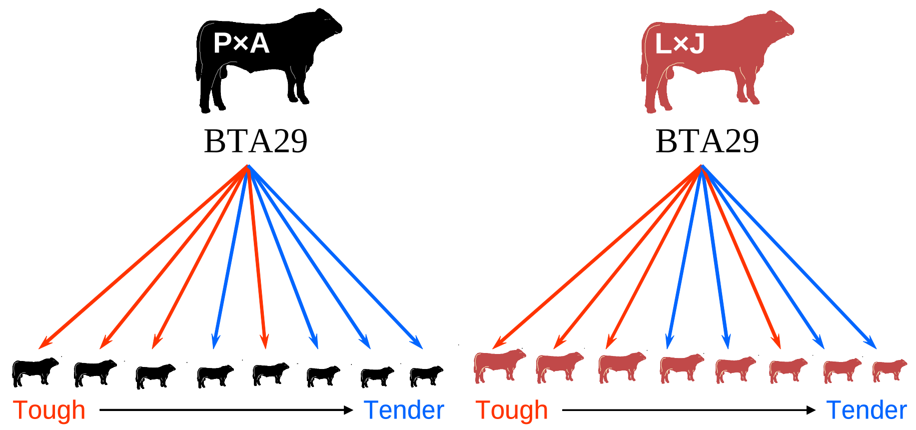
Which Type of DNA Test Will Be Most Widely Used in the Near Future?

- Association tests

Case Study: μ-Calpain

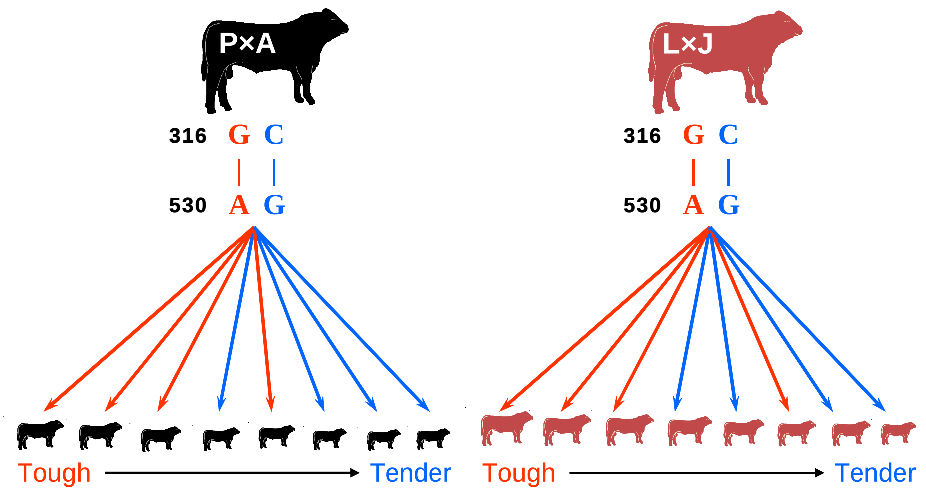
- QTL for Warner-Bratzler Shear Force on BTA29
- μ-calpain gene located under QTL peak
- μ-calpain is a proteolytic enzyme that plays a key role in postmortem tenderization of meat
- U.S. Meat Animal Research Center
- Tim Smith, Eduardo Casas, Roger Stone, Brent Page, Stephen White, Mohammad Koohmaraie
- μ-calpain was first discovered as a QTL on BTA29 in large paternal half sib families sired by a Piedmontese × Angus F_1 bull at MARC and a Limousin × Jersey F_1 bull in New Zealand.

BTA29 QTL



- The μ-calpain gene was sequenced in both of these F_1 bulls and they were found to have the same genotypes at the only two SNP that resulted in amino acid changes.

SNP Association



Possible Test Results Can be Viewed from Two Perspectives

SNP 530

SNP 316

	GG	AG	AA
CC	CC GG	CC AG	CC AA
CG	CG GG	***CG*** ***AG***	CG AA
GG	GG GG	GG AG	GG AA

Haplotype

Haplotype (316 530)

	C G	G G	G A	C A
C G	CC GG	CG GG	***CG*** ***GA***	CC GA
G G	GC GG	GG GG	GG GA	***GC*** ***GA***
G A	***GC*** ***AG***	GG AG	GG AA	GC AA
C A	CC AG	***CG*** ***AG***	CG AA	CC AA

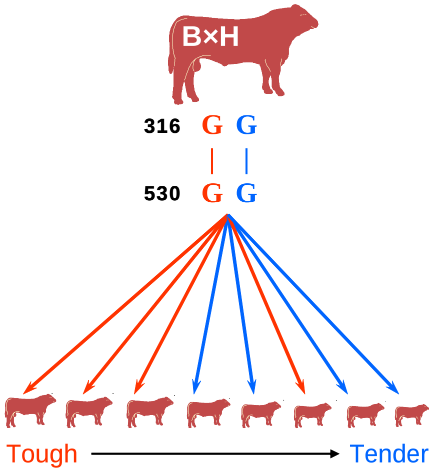
- Interpretation of the results becomes considerably more complicated with more than one SNP.

Two Possible Models for Observed Intermediate Haplotype Effect

Avg. Effect of Haplotype	Three Functional Alleles	Mixture of Two Functional Alleles
Tender	**316 C** **530 G**	**316 C** **530 G**
Intermediate	**316 G** **530 G**	**316 G G** **530 G G**
Tough	**316 G** **530 A**	**316 G** **530 A**
Too Rare to Estimate Well	**316 C** **530 A**	**316 C** **530 A**

A Challenge and an Opportunity

- A Brahman × Hereford F$_1$ bull was heterozygous for the QTL, but homozygous at the test loci.
- Red vs blue alleles in progeny are determined by flanking microsatellites.
- This situation suggests rather strongly that the G-G Haplotype is actually associated with a mixture of functional alleles.
- It was not the result we were looking for, but it provided an opportunity to make the test more powerful.



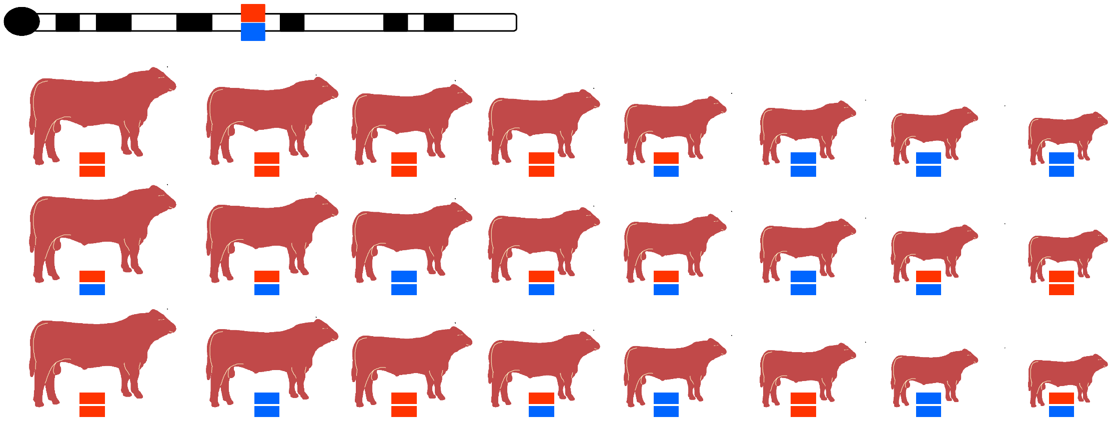
Why Did We Discover this Opportunity to Improve the Test?

- Because we looked
- Because we had an appropriate population in which to find it

We Are Learning a Great Deal From the μ-Calpain DNA Test

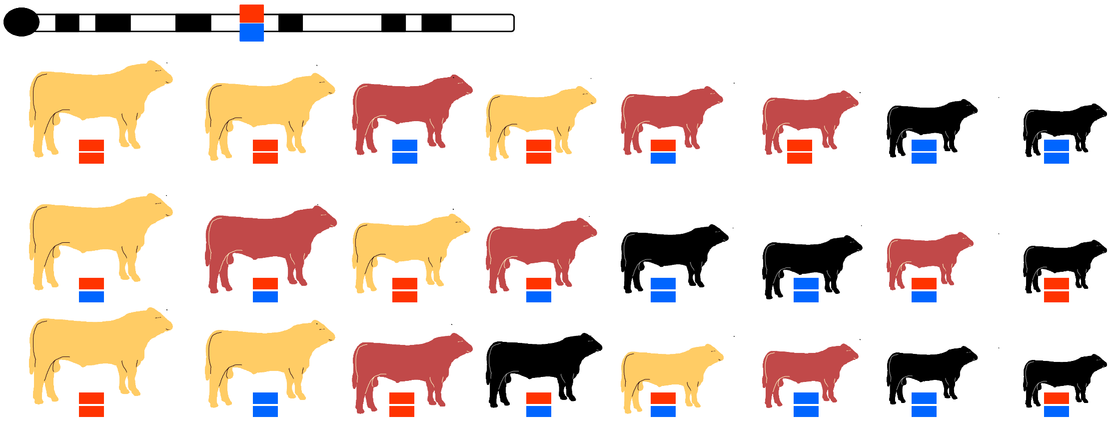
- Multiple SNP present some challenges.
- The test works very well as it is.
- I am confident that it will work even better when we add a few more SNP to it.

QTL Detection by Association (Linkage Disequilibrium)



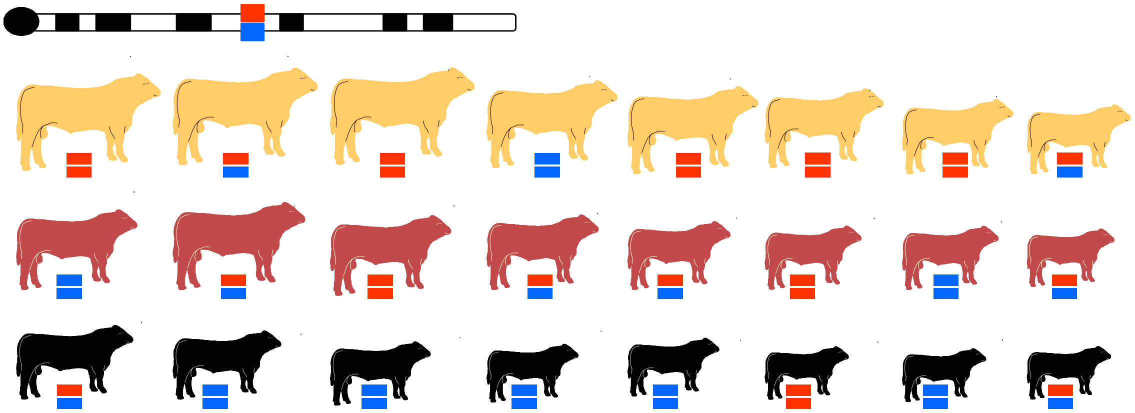
- In this figure, animals with red alleles tend to be larger than those with blue alleles. One possible explanation is that allele color is in linkage disequilibrium with a gene influencing size.
- This type of analysis usually assumes that animals are unrelated, so it is difficult to control for factors besides the QTL.

QTL Detection by Association (Linkage Disequilibrium)



- It is often desirable to use multiple breeds to increase the inference space in this type of project.
- In this figure, animals with red alleles certainly are larger than those with blue alleles.

QTL Detection by Association (Linkage Disequilibrium)



- However, when the cattle from the previous figure are sorted by breed, it is clear that there is no within-breed association between allele and size.
- The overall association between allele and size does not imply that the marker locus is in proximity to a gene that affects size.
- If breeds differ substantially for a phenotype, any marker for which the breeds differ in allele frequency will show an association.
- Therefore, it is critical to adjust for breed and pedigree in association studies.

Importance of Proper Design of Association Studies

- As illustrated in the previous set of diagrams, ignoring genetic effects (breed or pedigree) in association studies can easily lead to the conclusion that a polymorphism has a useful genetic effect, when in fact, it does not.
- A common tactic in association studies is to use populations in which pedigree (and perhaps even breed composition) is unknown. Then the lack of information is used as an excuse for ignoring these genetic effects in the analysis.
- The justification often used is that the goal is to be able to use the test on animals with unknown genetics, but this is faulty logic. In reality, we want the association to be due to linkage disequilibrium, not to some nebulous population stratification.
- In beef cattle, the most common abusive situation is tests on “feedlot cattle.” Pens of feedlot cattle usually fall into at least one of the following situations: 1) they are of heterogeneous breed composition and it is not recorded or 2) they are ranch-raised pens with paternal half-sib groups of substantial size.
- In dairy cattle, a similar, yet different situation occurs. It is essentially impossible to have large groups of “unrelated commercial Holstein cows,” but association studies are sometimes performed on populations in which such claims are made in order to avoid accounting for pedigree in the analysis. Similar, even if less extreme, examples of ignoring relevant pedigree information can be found in beef cattle studies.
- It is not unreasonably difficult to find populations of cattle with known breed composition and pedigree. In fact, most cattle that have extensive phenotypic information meet that criteria and evaluation of DNA tests is most effective when conducted on cattle with phenotypes for a number of traits.
- The cost per animal of DNA testing is no greater for properly designed and analyzed populations than for poor ones, but the information generated can be much greater and free of bias.
- Evaluating DNA tests in such “real world situations” makes for impressive sound bytes, but it is not the most productive use of resources. A population with known breed composition and pedigree can easily be analyzed ignoring those factors (for the real world application), but those factors can not be added to the analysis of a population in which they are unknown.
- If a polymorphism has a real effect that is functional or due to tight linkage disequilibrium, it should be present even if genetic effects are ignored. In fact, failure to find an effect in a “real world” population with sufficient numbers of each genotype would imply that a test does not work, although the statistical power would be better (because residual variance would be lower) if genetic effects were included. However, the converse is not true; if a significant association is detected in the absence of breed and pedigree effects, it is difficult or impossible to distinguish whether the effect is due to population stratification (which is not very useful) or to either linkage disequilibrium or a functional effect (which are useful).
- It can be argued legitimately that associations that are due to population stratification could be useful for predicting the performance of commercial animals with unknown genetics, for example, in a feedlot. The basis for the

argument is that the polymorphisms differ in frequency between breeds and, hence, are useful as predictors of breed composition. However, it seems obvious that a large number of such markers would be required to account for a substantial fraction of phenotypic variance and that the same fraction of variance could be accounted for by a much smaller number of markers that are, or are in linkage disequilibrium with, a functional polymorphism. The latter polymorphisms may also be useful in estimating breed composition. The number of markers is especially important in the feedlot application because of the current laboratory cost per polymorphism relative to the limited potential value per animal of genetic information in the feedlot. Furthermore, markers with associations due only to population stratification have no value for testing seedstock for selection purposes. It seems logical to focus on tests that can be used in both markets.

- In spite of these comments, when used properly linkage disequilibrium is a powerful tool and association tests will certainly contribute to beef cattle improvement. They just need to be evaluated in appropriate populations with appropriate models.

Recommendations for Associations Discovered Through Linkage Disequilibrium

- Validation should include checking for segregation in large families in which flanking microsatellite (or SNP) markers have been scored in the progeny.
- The progeny should segregate according to the prediction of the association test for the parents(s) (Transmission Disequilibrium Test).
- Should be a routine part of the protocol, just like testing DNA markers for Mendelian segregation.

Prioritization of Resources in Developing New Tests and Refining Existing Tests

- Need high-throughput systems for converting QTL into association tests
- Need a new generation of resource populations
- Microsatellite markers will continue to be an important resource, especially in developing robust association and/or functional tests based on SNPs.
- Converting QTL to association tests should be a high priority.
- Refining existing tests should be done primarily when opportunities for improvement are obvious. In the future, this will become a higher priority.
- In general, spending considerable resources to convert association tests into functional tests is likely to be less productive than the above approaches.

Analysis of Multiple-SNP DNA Test Data

- Fitting haplotype effects seems more reasonable genetically than fitting effects of multilocus genotypes.
- In many situations, it may be more reasonable to fit haplotypes as random effects.
- Multiple–trait haplotype models may be useful.

Which DNA Tests Should Be Included in NCE?

- A number of questions about a DNA test should be answered before deciding whether to use it or not, but the most important question is “Has it been independently validated?”

Conclusions

- DNA testing is quite challenging and not as simple as it first appears.
- DNA tests should be viewed as very fluid systems.
- The potential benefits are enormous.
- It will eventually become widespread in cattle breeding.
- It will probably eventually become widespread in cattle management.