Growth Trait Pilot Project

The following document is in two parts. Part one describes a project that is being undertaken by the US MARC to apply the 50K SNP panel to discovery of QTL associated with economically relevant traits to the beef industry. Part two, is to validate this discovery in seedstock herds around the US and this component is the one we are seeking collaboration from the breed associations and their membership.

Part One: Discovery and 1st stage validation, collaborative effort of research institutions (and potentially commercial DNA testing companies).

Overview: A consortium of research organizations, iBMAC (Illumina, Beltsville, Missouri, Alberta, Clay Center) developed a high-density SNP genotyping assay that contains over 50,000 SNPs for researching the bovine genome. This tool has greatly energized the effort for discovery of tests beef producers can use for selection purposes. There are large datasets of animals with very rich phenotypic profiles at US MARC and elsewhere being genotyped for discovery with this large SNP panel. As well, US MARC has genotyped influential industry bulls with this panel in a project referred to as the "2000 Bull Project". Your breed association is participating in that program.

Despite the enthusiasm of the research community for this tool, we recognize that we are not fully equipped with the knowledge, methodology or the phenotypic resources to fully exploit its use. We feel there needs to be an industry wide and industry supported effort to collaborate in developing this knowledge base and to develop resources large enough to challenge the full potential of this panel.

We propose to begin by focusing on the weight traits that are routinely included in National Cattle Evaluation: birth weight, weaning weight, yearling weight, and milk (or more properly, the maternal component of weaning weight). The reasons for starting with these traits include:

There are many more animals with 50K SNP genotypes that also have these traits recorded than any other set of traits. Therefore, it is the set of traits for which we are likely to have a success the soonest and from which we can learn the most.

We can use 50K chip data on high accuracy AI sires as a means to validate the accuracy of any predictions that are developed from the existing phenotypic datasets.

It will be the best set of traits with which to explore how well DNA tests will work across a variety of breeds, including breeds that were not included in the discovery data sets and across the different environments of the U.S.

There has been less private investment in development of commercial DNA tests for these traits.

Contrary to popular opinion, there is value in DNA tests for these traits.

DNA tests could make concurrent selection for genetically antagonistic traits (especially birth weight and growth rate) substantially more effective.

DNA tests could provide higher accuracy EPDs on young bulls at the time of bull sales (especially relevant for yearling weight and milk).

Pilot Project:

To demonstrate the benefits of a collaborative approach, we propose a pilot project focused on the growth traits. This project will have multiple phases to include discovery and validation and be organized as follows:

Discovery: USMARC (possibly in collaboration with the Australian Beef CRC, the University of Guelph, and the University of Alberta) will analyze data from animals with genotypes for the Ilumina BovineSNP50 and phenotypes for growth traits (birth weight, weaning weight, and yearling weight).
From these analyses, a variety of approaches to identifying sets of SNP that account for the most possible genetic variance for the target traits will be used.
In identifying these sets of markers, SNP that affect birth weight with minimal or opposite effect on growth and SNP that affect growth with minimal or opposite effect on birth weight will be preferentially included to improve the ability to select for antagonistically correlated traits.
Other entities with appropriate populations with 50K SNP chip data could also be included in the discovery process at this point.

1st Validation Stage: identify additional data sets on which 50K SNP chip data exists with which to validate the original subsets of SNP and evaluate the effectiveness of various strategies for identifying SNP and estimating their effects jointly.

Potential contributors to this stage could include:

- The USMARC SNP data on approximately 2,000 industry AI sires (excluding sires of the discovery population) of 16 breeds with 50K SNP chip data.
- DNA testing companies
- Universities or other research populations with appropriate data
- The international collaborators could fit into the project at this phase instead of the discovery phase. Provided that the international collaborators contributed to the discovery phase independently, they could validate each other's SNP sets at this stage.

This stage would be done by providing formulas for MBVs to the population owners, with each population owner performing the analysis.

Panel Selection and Development Stage:
After having examined the results of the first validation stage, a subset of the 50K SNP chip is selected and one or more MBV formulas (restricted to the selected set of SNP) for each trait is developed.

The new formulas are then checked against the various populations used in 1st stage validation to ensure that the restricted subset of SNP is sufficiently predictive.

Assuming the subset is sufficient, a low-cost assay for this subset of SNP is developed. This also assumes that funds for this step can be obtained.

If the subset does not account for sufficient genetic variance, the number of SNPs in the subset will be increased.

Part Two: Validation in the US seedstock industry and investigation into genotype by environmental (GxE) interactions.

Overview: The NBCEC, breed associations and their membership will develop an industry resource to validate the DNA tests developed in part one that will allow for assessment of the efficacy of the test across numerous environments. It is envisioned that the minimum SNP panel to be run on these animals will be one of 96 to 384 SNPs as described above. However, contingent on funding larger panels may be run. 2nd validation stage:

Collaborating breed associations will identify herds in the geographic region environmentally compatible with that of US MARC (Nebraska and surrounding states). This resource will be used to validate the discovery from part one "free" from any complicating factor associated with differences in environment. The target is to collect tissue samples and obtain access to phenotypes on roughly 1,200 purebred calves from each breed (born in spring 2009) in this region from large, influential seedstock producers with excellent data recording in several breeds. It is hoped that each herd will wean 200 or more calves in 2009 to reduce the number of herds in the project.

Producer will be asked to provide, through their respective breed association, access to their weight records, pedigree information and EPDs on these calves as well as the necessary information to form appropriate contemporary groups. This will be done by providing the breed association with the proper documentation to allow for sharing this information with the research team on this project.

NBCEC will provide the producers with the materials necessary for DNA collection (the choice being that of the producer; hair, ear notches, blood cards) and organize the database for those samples and animal id's.

Each breed associations will provide the project with $5000 to allow for genotyping. It is projected that this contribution will cover the costs of genotyping about ½ of the animals in this part of the project.

If funded for 2009 - 2010, the NBCEC will provide the remaining dollars to complete genotype this resource.

The data will be analyzed to provide an independent and stringent validation of the SNP chosen for the assay and the various MBV formulas.

Genotype by Environment Interaction (GxE):

In addition to the large scale, multiple breed validation represented in the 2nd validation stage, we propose to pursue knowledge of the potential for GxE for the molecular breeding values by expanding the regions from which the seedstock herds are selected. Challenging environments that may include fescue regions, hot and humid regions, hot and dry regions, high altitude etc. will be identified.

As with the 1st validation phase:

Collaborating breed associations will identify herds each geographic region. The target is to collect tissue samples and obtain access to phenotypes on roughly 1,000 purebred calves from each breed (born in spring 2009) in these regions.

Producer will provide access to their weight records, pedigree information and EPDs on these calves as well as the necessary information to form the appropriate contemporary groups of these calves. This will be done by providing the breed association with the proper documentation to allow for sharing this information with the research team on this project.

NBCEC will provide the producers with the materials necessary for DNA collection (the choice being that of the producer, hair, ear notches, blood) and organize the database for those samples and animal id's.

We will submit a grant to USDA to explore GxE and in that grant we will request funds to run the smaller panel on all calves in these regions.

Resource: This industry resource will represent 10,000 animals used in the 2nd stage validation and approximately 30,000 calves in the outlying regions.