Beef Cattle Genomics: Promises from the Past, Looking to the Future

A.D. Herring[1], D.G. Riley[1], J.O. Sanders[1], P.K. Riggs[1], C.A. Gill[1]

[1]Department of Animal Science, Texas A&M University, College Station, TX

Introduction

The potential of genomics to better understand and improve traits in livestock species has been considered for many years. As with many fields of study, increased knowledge in the field brings the realization that the field is more complicated than originally believed. This is certainly true of genetics and genomics. The concept of what a gene is (and can be) has expanded greatly over the years due to improved technologies, experimentation and understanding. The same nucleic acid sequence may have widely variable expression in different body tissues, and can have variable expression in the same tissue across individuals. In some cases, "unusual" results from genetics research that may have been thought of as "noise" in unexplained variation in phenotypes have been found to be caused by non-traditional inheritance patterns such as imprinting (when genes may be expressed differently if inherited from male vs. female parent). As genomic technologies and understanding improve, these provide valuable resources for animal breeding, but they do not provide simple answers to complex cattle production problems. The goal of this paper is to briefly discuss the progress that has been made in the area of beef cattle genomics and to speculate about its potential impacts on beef cattle breeding in the immediate and longer-term future.

Brief History of Genetics and Animal Breeding

The field of genetics is young compared to many other sciences. The publication (late 1800s) and recognition (early 1900s) of Gregor Mendel's experiments lay the foundation of the field of genetics (the study of inheritance of traits). Application of livestock breeding and genetics has been practiced for thousands of years beginning with domestication processes, development of local populations, choice of parents and replacement animals, and eventually development of breeds. In the 1920s, Sewall Wright and Jay Lush performed important research projects and published several papers dealing with the potential implications of livestock breeding strategies and concepts. Lush's book Animal Breeding Plans published in the 1930s lay the groundwork for the field of animal breeding. Sanders (2012) reviewed many historical research projects relevant to beef cattle breeding in the Southern USA. In the 1960s, the emphasis on accurate recording of performance traits and pedigrees of beef cattle led to the establishment of the Beef Improvement Federation (BIF). Research by Chuck Henderson and others provided for mixed model statistical analyses that incorporated relationships among animals and contemporary group effects to lead to EPD development. During the 1960s and 1970s there was also much discussion about the "systems" concept of beef cattle production led by Tom Cartwright and others that included the need for balanced selection emphases to provide productive populations of animals, including the recognition of genotype x environment and other interactions. During this time several crossbreeding projects also documented the importance of hybrid vigor for many important traits.

The fundamental principles of livestock breeding have been based on the concepts that animals tend to resemble their parents and other close relatives for many traits, inbred animals tend to breed better than they perform (and conversely that outbred animals tend to perform better than they breed), and that outbred animals tend to outperform inbred animals for many traits. As our knowledge improves about the genetics and the underlying biology that affect traits, we become better able to explain what dictates performance and to predict it, but it does

not change the fundamentals of livestock breeding; these technologies do, however, greatly affect potential rate of change.

Another important concept that must be understood is that traits may be qualitatively inherited or quantitatively inherited. Qualitatively inherited traits are those affected by only one or two gene loci, the animals can be grouped into distinct classes, and typical environmental influences do not affect the phenotypes. Quantitatively inherited traits are those where many gene loci are involved, animals show a continuous distribution of phenotypes (as with a bell curve) and typical environmental influences do affect the phenotypes. Any DNA sequence that is associated with a particular phenotype or performance level can be a genetic marker. An area on a chromosome that influences a quantitatively inherited trait is called a quantitative trait locus (QTL). Cattle have 30 pairs of chromosomes, and the entire DNA sequence (called the genome) of cattle is approximately 3,000,000,000 base pairs (bp). There are four bases that DNA is composed of (A, C, G and T); this is the genetic alphabet.

Development of Genetic Markers

Several DNA-based techniques were developed in the 1970s and 1980s that gave rise to the concept of genetic markers for traits. The first of these genetic markers used enzymes to cut DNA at specific base sequences to yield different size fragments (RFLP, restriction fragment length polymorphism). Later, it was discovered that some small DNA sequences were repeated in variable units (such as CA sequence, etc., called microsatellites), and eventually single base changes in DNA (SNP, single nucleotide polymorphism) were discovered to be prevalent throughout the genome and became the main type of genetic marker used.

Several groups around the world have been and are currently working on cattle genetics projects to map genes. The term mapping refers to establishing a relationship of a gene or a genetic marker to another already documented group of genes/markers on a specific chromosome. Actual genes might be 1,000 to 100,000 bp in length. In the late 1980s, there were very few genes that had been mapped in cattle, and people began to establish genetic maps. This led to the identification of certain DNA sequences scattered throughout the genome that showed differences across animals, and that had potential to be used as genetic markers for phenotype differences.

New genetic markers have continued to be discovered by a variety of methods. Having genomic data from multiple animal species allows for comparison of gene sequences of known genes in other species like mice and humans, to look for related gene sequences in cattle. This is how the double muscling (myostatin) gene in cattle was found a few years ago from a similar phenotype reported in mice (Kambadur et al., 1997; McPherron et al., 1997).

Numerous genetic tests have been developed for qualitative traits such as black coat color, polled status, and many genetic diseases. The use of genetic markers for parentage determination is also well established. These tests have greatly aided in selection decisions. Commercial parentage testing for cattle ranges in price from $15 to $35 per animal. The need and interest to better understand and predict genetic influences on quantitatively inherited traits of economic importance led to concepts of marker-assisted selection.

Genetic Marker Validation

Validation refers to the process to determine if a genetic marker that has been deemed useful in one population of cattle is also useful in other populations. The National Beef Cattle Evaluation Consortium (NBCEC) is the group of universities that have been responsible for conducting national cattle genetic evaluations for many breed associations for several years. This group also provides validation for genetic markers to determine their usefulness in several cattle populations. More information about this group and the validation process is available at their web site: http://www.nbcec.org. A report in 2007 by Van Eenennaam et al. showed that commercial genetic tests could be useful across many groups of cattle. It is important to ask whether or not a commercially available genetic

test has been validated, and in what types of cattle. Some breed associations have developed breed-specific DNA tests and/or begun to incorporate genomic-enhanced EPDs that blend genomic information with traditional performance and pedigree measures.

Some of the historical efforts to identify genetic markers in cattle that we have been involved with are briefly discussed below.

TAMU Angleton Project
In 1990, a project at Texas A&M University was initiated to search for genetic markers in cattle related to growth and carcass traits. Calves were produced by backcross (F_1 bred back to purebred) matings involving three distinct sire and dam types: Angus (A), Bos indicus (B) (Brahman), and F_1 Angus-Bos indicus (Brahman or Nellore). Resulting progeny were 3/4 A-1/4 B, or 3/4 B-1/4 A. Over 500 calves were born from 1990 to 1996. There were 32 full sib families that were produced through embryo transfer. All calves were born to recipient dams that were approximately 1/2 Brahman and 1/2 British. All calves produced in this project were fed and killed to obtain feedlot and carcass data. This project was jointly funded by the Texas Agricultural Experiment Station, USDA and the Beef Checkoff program. There were over 80 QTL identified as being related to differences in carcass and growth traits, and 11 of these QTL were identified as influencing traits of shear force, taste panel tenderness, marbling and ribeye area. To determine if these markers were useful in other populations of cattle, a marker validation study financed through the Beef Checkoff Program was implemented and was called the Carcass Merit Project.

NCBA Carcass Merit Project
The carcass merit project was initiated in 1998 and involved four universities, 13 breed associations (14 breeds) and a private genetics company. This project was completed in 2003 and involved over 8,200 progeny of 300 different sires across the 14 breeds. All breed associations were invited to participate. Up to 10 bulls from each breed were designated as "DNA sires" and produced 50 progeny each; five of the 10 sires in each breed were

designated "sensory sires" and also had taste panel data collected on all 50 progeny. A total of 70 sires were used for the DNA analysis. There were 11 QTL evaluated across these breeds, and some of the validated markers led to commercialized tests. All of the evaluated QTL from the TAMU Angleton project appeared to be useful in explaining some of the differences for at least one trait in the NCBA carcass merit project, and many also seemed to be useful in explaining variation in several traits. Although all these QTL markers were significant in explaining differences in certain traits, this still does not guarantee that they all will be useful in all breeds or populations. It is highly likely that certain breeds will have specific gene combinations some places in the genome that are different from other breeds. This is the concept behind DNA marker validation.

Bovine Genome Project
Many people have heard of the human genome project, but don't realize that many species of plants and animals (including cattle) have also had "genome projects" to determine entire DNA sequences for those species. The bovine genome project was an international project funded through the National Institutes of Health (NIH), USDA, the State of Texas, as well as several other countries. This effort has greatly enhanced our ability to study cattle genetics and will do so for many years to come. A major result of the cattle genome project is that thousands of different SNP markers have been identified, and this has led to development of DNA-based platforms called "SNP chips" which can be used to search for QTL across the entire genome and come in a variety of formats such as 3,000, 50,000 or 770,000 SNP platforms. The SNP on these commercial platforms were chosen for their utility across a wide variety of beef and dairy breeds.

McGregor Genomics Project
In 2003 we initiated a project with the primary objective to study genomic influences on beef cow fertility and longevity. Secondary objectives of this project have been to evaluate feed intake and feed efficiency, temperament, carcass traits and health aspects. Cattle produced have been primarily F_2 and F_3 Nellore-Angus

crosses. To date we have produced approximately 1,200 cattle in this project. One particular concern that we have had is that some widely different results have been observed in regard to genomic influences on traits when different statistical approaches (that have reported in the scientific literature) are used.

Decreasing Costs of Obtaining Sequence Data
As DNA sequencing technology has improved, this has allowed the cost of genotyping to continually decrease on a per marker basis. Eggen (2012) stated that there had been a 100 million-fold decrease in DNA sequencing costs from 1990 to 2012. Microsatellite genotyping still costs about $1.25 per marker, whereas the cost per SNP genotype on the 770K high density chip is about $0.0002 (2/10 of a cent per genotype). This excludes DNA extraction costs.

The investment in the Human Genome Project (http://www.genome.gov/) took 13 years and cost $3.8 billion, and, caused many researchers to search for and develop technologies to obtain DNA sequence data with lower costs; the bovine genome project cost $53 million and was completed in 4 years. This was total cost (not just the sequencing). Each year the genome.gov publishes sequencing costs (can also see http://www.dnasequencing.org/history-of-dna), and the cost to sequence a genome was about $6,000 in early 2102. It was predicted by several of the next-generation sequencing companies that cost to sequence a genome would reach $1,000 by Fall 2012, and $100 per genome by Fall 2014. This would not include the costs to analyze and interpret the data, which becomes increasingly complicated as more sequence data per animal is incorporated into analyses. We need computer programmers that have understanding of the biology of our livestock species.

Marker-assisted Selection and Genomic Selection
Use of genetic markers offers the potential to get an idea about an animal's genetic makeup before a lot of time and expense are invested in that animal. Traits that are hard to measure such as feed intake and efficiency or traits measured later in life such as carcass quality and reproductive performance and longevity offer a lot of promise from use of genetic markers. Also to improve traits where the phenotype may be hard to accurately measure such as health where sub-clinical illness may go unrecognized, it could be very useful to have genetic markers. However, any trait could benefit from useful genetic tests if the genetic markers explain differences in phenotypes and/or breeding values. When the genomic information is incorporated into breeding programs, it is considered marker assisted selection. As a result, the ability to have increased accuracy of EPD in young animals based on genomic information has the potential to alter the rate of genetic change. The concept of genomic selection was proposed in the scientific literature in 2001 by Meuwissen et al. to represent incorporation of large-scale genomic information into estimation of breeding values (or EPD). As more animals can be genotyped due to continuing reduction in costs, there has been more interest and effort placed on genomic selection.

There are multiple ways that genomic selection is being incorporated, and these were reviewed (NBCEC, 2012), and the descriptions below are summarized from it. In many cases, a molecular breeding value (MBV) is calculated and that explains some percentage of the genetic variation for a trait. One approach is the MBV is combined with EPD as a type of index for a trait. The dairy industry has been quite proactive about incorporation of genomic selection in selection of sires and has utilized marker panel information as genomic relationship into genetic evaluation. The American Angus Association has utilized MBV as a correlated trait with the trait of interest to calculate EPD where increased accuracy occurs through the MBV accounting for more genetic variation in the trait. And, the American Simmental Association utilizes an approach where the MBV is treated like an external value (outside of the population as with an EPD of an animal of another breed).

There is no doubt that sequencing technology is a powerful tool for genomic evaluation, that it will continue to decrease in cost, and, that resultantly, its utilization for study of complex traits will continue. However, at this point, some

discussion of why current genomic selection approaches are likely to not provide the answers to all important cattle breeding questions is provided.

Non-traditional Genetic Influences (Epigenetics)

In recent years there has been much research and documentation about non-traditional genetics, termed epigenetics. This broad area refers to inheritance patterns that can differ even when the DNA sequence is the same. An individual has one genome but many epigenomes. That is, epigenetic modifications of the genome contribute to regulation of the genome so that more than 200 specialized types of cells function differently within an organism, even though each contains the same set of chromosomes (Rada-Iglesias and Wysocka, 2011). In many cases, different sets of genes are "turned on," or expressed, in each cell type. Genes, spelled out in the genome sequence, are transcribed into messenger RNA molecules (mRNA), that are then translated into proteins that function in the cell. Other types of RNAs perform regulatory functions. The set of all the RNAs transcribed in a cell is called the trancriptome. In order to understand the biology behind the networks of genes produced in specific types of cells and tissues, we are interested in utilizing all the information that we can gather about an individual – its genome, transcriptome, and its phenotype. This type of integrated approach will lead to a better understanding of the complex traits in animals such as cattle that are affected and controlled by networks of genes. (Similar idea reviewed in Kadarmideen and Reverter 2007)

Epigenetics was first proposed in the 1950s and is increasingly recognized as an important emphasis area in human health. Non-traditional inheritance patterns in cattle have been documented, particularly in Bos indicus-Bos taurus reciprocal crosses. This has been and continues to be a major interest area for our group (for instance Roberson et al., 1986; Amen et al., 2007), and production of reciprocal F_2 crosses in our McGregor Genomics Project is driven by this interest. We have initiated formal collaboration with researchers in Australia at the Queensland Alliance for Agriculture and Food Innovation and the University of Adelaide to widen the study of epigenetic influences in several Bos indicus-Bos taurus populations. González-Recio (2012) emphasized the need to recognize the broad area of livestock epigenetics and epigenomics for future research. Genomic technologies provide tools to better study epigenetic influences, but current genetic evaluations and genomic selection strategies do not consider it.

Systems Concepts about Beef Cattle Breeding and Production

Sewall Wright in 1939 said that both production as well as the economy of production must be considered in improvement of characters. He also asked if we were given a map of the chromosomes showing the locations of all important genes as well as convenient marker genes, what could we do with it? In a few years the cost of DNA sequencing will be low enough that entire genomes of animals may be used for analyses as compared to DNA markers. In 1970, Tom Cartwright stated there likely needed to be less emphasis on "breeding better cattle," and more emphasis on "better cattle breeding." Both of these statements have relevance today as we need to consider what genetic improvements are most needed in regard to beef cattle production, particularly in light of historically high feed prices and a reduced national cattle inventory.

Many people compare production outputs per unit of production as a measure of efficiency. Tremendous improvements have been made in these types of measures in U.S. agricultural species. The figures below highlight corn, milk production and beef production in the United States (data from USDA National Agricultural Statistics Service) from 1936 to 2012, and some of the specific values from selected years are provided in Table 1.
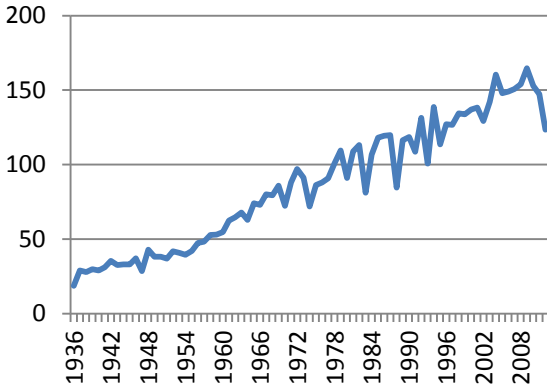
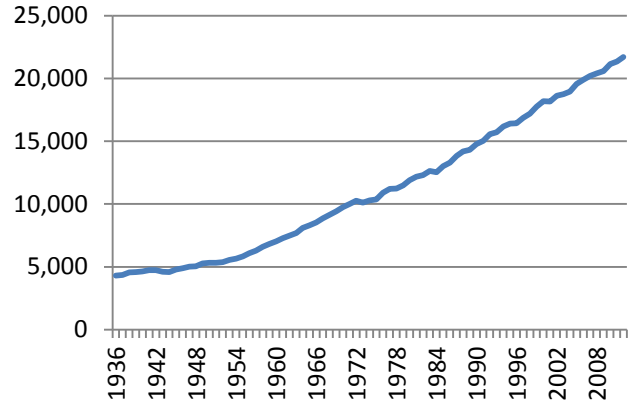**Figure 1.** U.S. corn yield (bu/acre) from 1936-2012

Figure 2. U.S. milk production (lb) per cow from 1936-2012

Improvements in these species have been from both genetic and environmental/management changes. We probably do not know what the upper limit is on production potential of corn per acre, milk per cow or beef per animal, but there are maximal limits. The profitability of production is and will be the primary factor that keeps producers operating. The role of genomic information has had a limited role as it has only recently become available in these species.

Table 1. Comparisons of production levels across time in some agricultural species

Time	Corn (bu/ac)	Milk (lb/cow)	Slaughter weight (lb/animal)	Beef produced (b) per cow inventory
1940	29	4622	905	195
1960	55	7029	1,004	316
1980	91	11,891	1,072	436
2000	137	18,197	1,219	632
2012	153*	21,697	1,277	672
Current* vs. 1940 (%)	528	469	141	345
Current* vs. 1980 (%)	168	182	119	154

*Current comparisons for corn based on 2010 year.

The limit on beef production per animal is also unknown because we do not know how big we can make cattle. The long-term genetic improvement in the context of a production system must be considered because unlike some artificial input such as amount or type of feed, use or not of implants, etc. the genetic change is hard to remove. The beef industry is sometimes criticized because it has not made as much improvement in production as crops such as corn or milk production, or other more "progressive" agricultural species.
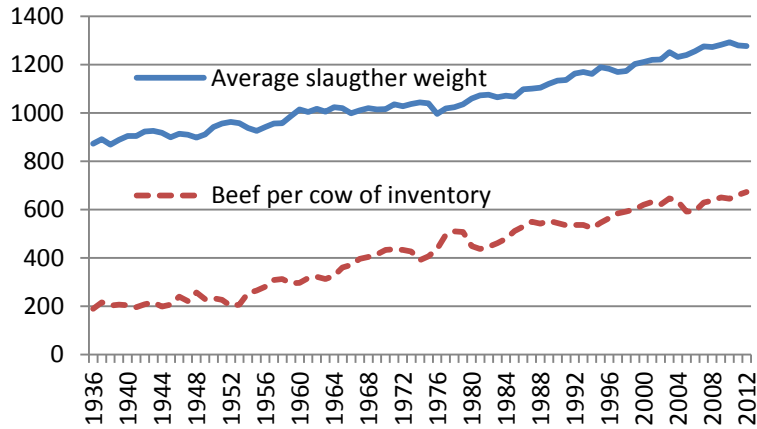
Figure 3. Average slaughter weight (all cattle classes) in lb and beef produced per cow of national inventory.

Figure 3 shows that the increase in average slaughter weight from 1936 to 2012 is much smaller that the increase in beef produced per cow of inventory because it is more of a trait of a system, as opposed to a trait of an animal. When evaluating beef production relative to the national beef cow inventory, the beef increased production from 1940 or 1980 bases is much more in line with increases in corn and milk production. Understanding the genetic/genomic influences on the component traits and there potential interactions that influence the overall desired trait of the production system should be more of a priority that individual traits. This is the bases of selection index theory that was also first proposed in the 1940s.

Several breed associations currently calculate and report index values on their respective animals based on EPDs. Van Eenennaam et al. (2011) stated that DNA information could be valuable for seedstock when properly incorporated into genetic evaluations and need to be incorporated into multiple-trait indices. As more information is understood about gene expression and functional gene networks (groups of genes that produce certain related biological function(s)) it is likely that this type of information will add to DNA sequence-based marker utilization and produce a new type of genomic selection than exists today

Summary
The purpose of this paper/presentation is not to be negative about genomic technologies; these tools possess huge potential for improved understanding of genetics and biology associated with all economically important traits. The purpose, however, is to emphasize that beef cattle breeding needs to be based on balanced, systems based approaches, and there are currently available tools that allow more focused single-trait selection potential than ever before if someone chooses to go that route. The cost of genomic technologies will continue to decrease, and, therefore it will become more economically attractive to utilize.

Genomic technologies offer powerful evaluation tools, but with large-scale genomic selection how can we be sure that: (1) we are truly changing what we think we are changing, and (2) be sure that we are only changing what we think we are changing? There are not certain answers to these traits with genomic selection and current knowledge. Research in human health is taking an integrated approach to make use of whole genome sequence information as well as gene expression based on messenger RNA (called the transcriptome) and protein products (the proteome) information to understand genetically complex diseases and health traits. Because of the reduction in

genotyping costs, huge populations are now being genotyped somewhat routinely (>100,000 individuals) as very large sample size allows for detection rare genetic variants. The integrated approach to incorporate multiple types of genomic information also holds much potential in cattle and other livestock populations.

As there is improved understanding about the relationships involving DNA sequence, total genomic influences (and interactions) and traits of economic importance in livestock, both the procedures of their use and the recommendations about their use will also change. This has also been the case with utilization of EPDs. Cattle producers need to maintain balanced focus on producing functional cattle that work for their own and their customers' production and marketing environments. Use of genomic information can help these things continue/improve, or, can completely disrupt them.

References

Amen, T.S., A.D. Herring, J.O. Sanders, and C.A. Gill. 2007. Evaluation of reciprocal differences in *Bos indicus* x *Bos taurus* backcross calves produced through embryo transfer: I. Birth and weaning traits. J. Anim. Sci. 85:365-372.

Cartwright, T.C. 1970. Selection criteria for beef cattle for the future. J. Anim. Sci. 30:706-711

Eggen, A. 2012. The development and application of genomic selection as a new breeding paradigm. Animal Frontiers 2:10-15.

González-Recio, O. (2012). Epigenetics: a new challenge in the post-genomic era of livestock. Frontiers in Genetics 2:1-4.

Kadarmideen, H.N., and A. Reverter. 2007. Combined genetic, genomic and transcriptomic methods in the analysis of animal traits. CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources 2: No. 042 doi: 10.1079/PAVSNNR20072042. http://www.cababstractsplus.org/cabreviews

Kambadur, R., M. Sharma, T. P. Smith, and J. J. Bass. 1997. Mutations in myostatin (GDF8) in double-muscled Belgian Blue and Piedmontese cattle. Genome Res. 7: 910-916.

McPherron, A. C., and S. J. Lee. 1997. Double muscling in cattle due to mutations in the myostatin gene. Proc Natl Acad Sci U S A. 94: 12457-12461.

Meuwissen, T. H., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819-1829.

NBCEC. 2012. National Beef Cattle Evaluation Consortium (NBCEC) Wite paper – Delivering genomics technology to the beef industry. http://www.nbcec.org.

Sanders, J.O. 2012. Historical Overview of Animal Breeding and Genetics Research in the South. Proceedings Beef Improvement Federation 44th Annual Research Symposium & Meeting, Houston, TX pp 1-11.

Rada-Iglesias, A., and J. Wysocka. 2011. Epigenomics of human embryonic stem cells and induced pluripotent stem cells: insights into pluripotency and implications for disease. *Genome Medicine* 2011, 3:36. http://genomemedicine.com/content/3/6/36

Roberson, R.L., J. O. Sanders, and T. C. Cartwright. 1986. Direct and maternal genetic effects on preweaning characters of Brahman, Hereford and Brahman-Hereford crossbred cattle. J. Anim. Sci. 63:438-446.

Van Eenennaam, A.L., J. Li, R.M. Thallman, R.L. Quaas, M.E. Dikeman, C.A. Gill, D.E. Franke, and M.G. Thomas. 2007. Validation of commercial DNA tests for quantitative beef quality traits. J. Anim. Sci. 85:891-900.

Wright, S. 1939. Genetic principles governing the rate of progress of livestock breeding. J. Anim. Sci. 18-26.