Beyond EPDs - Genomics: Practical and Economic Considerations

Ronnie D. Green
National Program Leader
USDA / ARS
Washington, DC

Introduction

Genetic improvement in livestock has a truly amazing history, with the beef industry having been blessed with many of the major scientific innovations that have occurred along the way. In recent years, it has been nearly impossible to miss seemingly daily news reports about exciting discoveries in the new field of molecular genetics and genomics. While most of these reports have focused on the unraveling of the human genome and its implications for human health, there has been significant spillover into plant and animal agriculture as well. At times over the past 15 years, it has seemed to many that this new and exciting field would hold all of the immediate answers to breeding better beef cattle. Today we stand on the brink of having the DNA sequence of the cow genome completed and made publicly available. What will this mean to the beef industry? The objectives of this presentation are to: 1) provide a rigorous overview of the field of genomics as related to genetic improvement of beef cattle, from historical, current, and future perspectives; and 2) provide some insight into what the future system of genetic evaluation will look like with the coming addition of molecular genetic information.

A Brief History of Beef Cattle Genetic Evaluation

It is believed that cattle were domesticated over 5,000 years ago. Only in the last few hundred years has the human race applied systematic animal breeding programs to these amazing animals to mold them into more specific roles – i.e. meat, milk, or draft. Today the number of distinct cattle breeds numbers in the hundreds across the world.

In the U.S., our cattle industry quickly developed as segregated into dairy and beef sectors. By the dawn of the 20th century, the beef part of this industry had essentially become made up of three breed populations – Aberdeen Angus, Hereford, and Shorthorn. It is unlikely that our ancestors engaged in the beef business at that time – for most of us our grand- or great-grandparents - would have been able to predict the dramatic changes that would take place in the next 100 years.

The first half of the 20th century was an immensely prolific time in agricultural science. Arguably, the most dramatic discoveries were actually in the fields of genetics and statistics. During the 1920s and 1930s, the field of population genetics came of age – primarily as a means of quantifying and describing Darwin's writings from the late 1800s. The emerging leaders of this field helped to describe the concepts of genes, gene loci, chromosomes, and cellular reproduction. They were also instrumental in establishing the field of biometrics – statistics as applied to biological phenomena. These early statisticians developed much of the underlying theory used broadly in science today. What most people do not know is that they originally were geneticists trying to describe how populations of animals change over generations! Also, at the same time there were pioneering scientists who had the foresight to develop populations of beef cattle upon which they began to practice selection and inbreeding – ones like the Miles City Hereford lines that gave us the Line 1 of today.

Scientists also made what seemed to be an unrelated, but extremely valuable, discovery in plant genetics during this same time period. Scientists observed that when two unrelated lines of germplasm were crossed – or "hybridized" – the resulting

crossbred progeny had better performance than the expected average of the parents. The concept of heterosis between lines was born – and with it the seed industry and crop agriculture was revolutionized. At the time, livestock breeders did not see any great benefit from this phenomenon – but as we now know, that would dramatically change later.

The post-WWII era was a particularly exciting time for livestock genetic improvement, as it was in many fields. The 1940s saw some of the greatest minds to ever grace the study of livestock genetic improvement at their prime. Jay Lush, who many refer to as the modern day father of animal breeding, was busy defining with co-workers Lanoy Hazel and Gordon Dickerson the concept of the "selection index" and "breeding value". The field of biometrics had matured to the point where it was now possible to determine from experimental populations that performance for traits affecting production could be measured – and that many of these traits appeared to be heritable.

In 1953, James Watson and Francis Crick presented for the first time in the scientific literature the molecular structure of the genetic code – i.e. DNA. Combined with the theories of genes and heritable variation of traits, it was now possible to visualize how these genetic differences at the gene level might one day be exploited for genetic improvement.

Also, in the 1950s, two significant events occurred which would forever change the nature of cattle breeding. The first was that artificial insemination techniques matured to an adoptable level for cattle breeders – especially dairy producers. Coupled with the institution of the Dairy Herd Improvement programs of USDA a bit earlier, volumes of data began to accumulate matching pedigrees to milk production records. At the same time, computing technology was beginning to surface as a usable tool – even though it was rudimentary to what we now have today. Dairy cattle breeders had enough foresight, however, to understand the power of coupling quantitative genetics theory to artificial insemination and as a result genetic evaluation as an applied science was born. Now, through all of the technological and computing improvements of the last 45 years, we have seen that this works – to the tune of almost 100% improvement in milk yield per cow!

Thankfully, the plans of the dairy industry did not go unnoticed by beef cattle breeders. The American Angus Association and the American Hereford Association quickly established performance recording programs for their breeders – focusing initially on 205 day weaning weights within herds. A few years later in 1968, some visionary cattle breeders, including Sally Forbes, Frank Baker, Jim Brinks, Bob deBaca, and others, formed an organization called the Beef Improvement Federation. This organization was instituted to take on the task of developing uniform guidelines for performance recording programs, the same task that it still performs 37 years later. One of the initial visions of this group was that it would soon be able to develop methodology to compare animals across herds – making the standardization of performance recording critical.

The late 1960s and early 1970s was the next time of great change in beef cattle breeding. Two things occurred somewhat simultaneously – the importation of semen from a number of continental European breeds of cattle and the next generation of computing technology coming of age. As a result of their higher growth rates, size, and muscularity, a number of these breeds quickly took a strong foothold in the beef cattle seedstock industry – especially Simmental, Limousin, and Charolais. As the American Simmental Association took its first steps, it carefully studied the performance recording movement and was quick to the chase to be the first group to recommend that they should attempt to take advantage of the improvements in genetic prediction methodology, artificial insemination, and computing technology to compute and make publicly available the first "across herd" comparisons. They did so, using what was called a "sire model" developed at Iowa State University, by Richard Willham in 1972. This allowed the prediction of "estimated breeding values" (EBVs) for the growth traits by tying herds together through a reference sire network. The era of true beef cattle evaluation was now born. Shortly thereafter, maternal grandsires were added to the evaluation framework – allowing "maternal" weaning

weight EBVs to be added.

At the same time, it was clear that much more information was needed for beef producers to effectively sort out the widening levels of genetic variation available to them for commercial production. Additionally, producers discovered that hybrid vigor was indeed possible – and very economically beneficial – when many of the new breeds were bred to the available Hereford and Angus cows. Crossbreeding and hybrid vigor seemed to have a place at the table. Fortunately, USDA's Agricultural Research Service saw the need for scientific data in this area. As a result, the U.S. Meat Animal Research Center (MARC) at Clay Center, NE, was born and within a short time initiated two monumental projects – the Germ Plasm Evaluation (GPE) program led by Larry Cundiff and the Germ Plasm Utilization (GPU) program led by Keith Gregory. At the same time the Fort Robinson station was being closed, and the selection lines of cattle there were relocated to MARC – becoming the third piece of the puzzle led by Bob Koch. Over the next 30 years, this collective effort produced the fundamental body of knowledge now used world-wide to understand genetic variation, and how to effectively use it in beef cattle production.

The 1980s were a true time of transition for beef cattle breeding. Computing technology had now matured to the level where statistical methodology developed by a dairy geneticist named Charles Henderson, in the 1950s, could be applied to beef and dairy performance data – so called BLUP (Best Linear Unbiased Prediction) methodology. Scientists worked out the kinks and were successful in using these methods to compute for the first time what we now know as EPDs – ***Expected Progeny Differences*** within breeds. These new genetic evaluation tools were significantly more powerful and accurate to allow breeders to sort not only bulls – but also cows — than the previously used EBVs from the sire/maternal grandsire model approach. Over the ensuing 20 years we would be the benefactors of continual refinement in genetic prediction methodology, including more accurate predictions as well as a plethora of new traits added to the evaluation pipeline. We would even see the MARC GPE populations serve another useful role, when in the early 1990s, data from the breeds evaluated in the GPE project coupled with breed genetic evaluation data, were used to develop an "across-breed" adjustment process allowing commercial producers for the first time to compare bulls across not only herds, but also across breeds.

The other monumental event in the 1980s was the unleashing of a new field of science collectively referred to as "genomics". This term was first used in 1986 to collectively describe the scientific discipline of mapping, sequencing, and analyzing genomic level DNA information. A technology called "polymerase chain reaction," developed in 1987 by Kary Mullis in California, literally unleashed the forces of research into the genetic code of plants and animals. It had only taken 34 years to go from understanding the structure of DNA to being able to start the process of deciphering the meaning of the code!

This somewhat exhaustive and comprehensive history lesson has been presented here to intentionally bring light to the fact that the process of getting to today's state of the art beef cattle breeding has not been easy, or achieved quickly. One could argue that 100 years in the bigger picture of 5,000+ years of domesticated livestock production is a drop in the bucket. However, most of us would still argue that those 100 years have been a monumental and unprecedented effort. ***As we enter the era of molecularly aided genetic improvement – we must be careful to remember the big picture, and that while these new tools are fascinating and almost unbelievable to many of us, they are simply the next pieces of the puzzle in a long process of continual refinement and improvement as beef cattle breeders.***

What is Genomics and Why is it So Difficult?

Genomics in the simplest terms is the study of the DNA complement of a given species. There are a number of sub-categories of the broader field including structural genomics, comparative genomics, and functional genomics.

Structural genomics is a collection of research strategies and tools used to better understand the organization and content of the genetic code. Cattle, as is the case for most mammals, have a rather large genome – in the order of 3 billion individual base pairs. These strings of bases are organized into pieces called chromosomes. In cattle there are a total of 29 pairs of non-sex determining chromosomes plus sex-determining X and Y. The genes that affect traits are dispersed across the chromosomes, with a given gene located at a specific site on a particular chromosome. Genes are the units of the DNA that individually encode specific protein products used in the body either as building blocks or metabolites. Our best estimate is that the number of cattle genes is somewhere in the order of 30,000 to 40,000. We have long known that the basis for the genetic variation we observe in cattle performance is variation in the DNA code of individual genes contributing to complex quantitative traits – what you may have learned in basic genetics as "alleles" – or different forms of genes. The goal of structural genomics is to develop a complete enough understanding of how the genome is organized so that we can begin to locate and understand these DNA level variations – what the scientists call polymorphisms.

As molecular genetics tools became available to lab scientists in the late 1980s, researchers began the arduous process of genetic mapping. Because they were unable at that time to know what the base sequence of the DNA code was, they had to use a somewhat "black-box" approach to identify locations on the chromosomes that might contain genes affecting these traits. This process, called linkage mapping, took advantage of DNA polymorphisms called microsatellite markers, a type of variation found readily throughout the genome. Using one of the laws of genetic inheritance that had been well defined for many years called linkage, these markers could be used to identify regions of the genome in the same vicinity where they occurred that seemed to affect differing levels of performance, as well as identify where these markers were located in proximity to one another on the chromosomes. In 1994, the first genetic linkage maps of cattle were published by USDA-ARS scientists from US MARC and Australian CSIRO scientists. Today, these linkage maps, combined with what are known as radiation hybrid maps, are quite well defined with a total of over 9,000 individual markers identified and localized to chromosomes. An excellent example of the level of information contained in these maps can be viewed on the USDA-ARS MARC web-site (http://www.marc.usda.gov/genome/htmls/LinkageMap.jsp).

The availability of the first linkage maps allowed researchers to begin the search for regions of the genome harboring genes containing polymorphisms causing differences in performance for economically important traits – what have become known in the jargon as quantitative trait loci (QTLs). This research, conducted at several locations in the U.S., Australia, New Zealand, and Canada first required the establishment of cattle resource populations that would have a high probability of having different copies of the genes on an individual animal's maternal versus paternal chromosome. Consequently, a number of these resource populations were designed targeting different classes of traits – although most were focused on carcass and end product attributes initially. Typically, these populations were made by crossing widely divergent breeds to make F_1 sires who were subsequently mated to cows of one of the original breeds to produce progeny. This allowed the researchers to detect differing alleles having large effects from these QTLs. The markers in the gene map were then the "tags" which were inherited with these different alleles through genetic linkage that allowed the pinpointing of the chromosomal locations of the QTLs.

A number of these resource populations were formed at the MARC. Over the course of the last 7 to 8 years, these populations have been utilized to identify over 25 QTLs affecting a wide variety of traits on 11 different chromosomes. Other research groups have also identified a number of QTLs, principally the Angleton population at Texas A&M funded primarily by the beef checkoff and the CRC/MRC project in Australia. In the Texas project ten QTLs were identified for various carcass traits. The results of these projects were exciting and stimulated a considerable amount of attention in the beef industry in the mid to late 1990s. Unfortunately, as is too often the case, in the rush to

find the silver bullet, the immediate promise of genomics was clearly oversold. The identification of QTLs was only the first "baby step" in the process to bringing these results to a practicable technology.

The easiest way to comprehend the enormity of the task of "mapping a gene" is to think in terms of needles and haystacks. Consider each of the thirty chromosomes of the bovine genome to be round hay bales – some bigger than others. Then consider that these QTLs that had been identified had each been determined to be in one of the haystacks and to have been in a particular region of the original windrow before the baler rolled up the bale. While the scope of where we were looking for the gene has been dramatically narrowed, the fact remained that we still were faced with literally looking for a needle within a haystack. The individual gene causing a QTL effect was in an area surrounded by millions of pieces of grass along a rather long stretch of windrow. And worse yet, because the researchers did not have available to them the DNA sequence in those regions, they were required to come up with indirect ways to narrow the scope any further. Many of the folks involved in the early research vastly underestimated the task.

Fortunately for the cattle genomics community, the U.S. government placed a high priority through its human medical research arm – the National Institutes of Health (NIH) – on deciphering the human genetic code. The idea was very similar to what has been described in this paper for cattle – except that in this case the target was to develop new ways to combat human disease / improve human health. Initially, many of the same approaches of linkage mapping were used in human genomics, with the additional twist that model organisms were intensely studied as proxies for man – principally the laboratory mouse and rat. This was possible because as we began to be able to see small regions of DNA code, the similarities between species were remarkably high – usually in the 90% or higher level. Scientists also observed that while the arrangement of the pieces of the genetic puzzle was not the same across species, large regions of the genome had been conserved throughout evolutionary time. This now allowed the opportunity to take information from species being studied with very large research budgets in comparison to cattle to infer what might be the case in cattle. This approach – broadly called "comparative mapping," has since been very effectively used to identify a number of the genes that we now know in livestock.

Perhaps the best example of the use of comparative genomics in livestock is the gene causing double muscling in continental European breeds. The condition of double muscling has been a curiosity in cattle breeding for many years. Breeds such as the Belgian Blue, and to a lesser extent Limousin, Charolais, and Simmental, are clearly different from other breeds in terms of their muscle:bone ratio. Scientists have comprehensively studied this phenomenon to determine how the condition could be favorably used in lean beef production systems and to attempt to understand the underlying physiology causing the muscle hypertrophy. When the advances of molecular biology occurred in the 1980s and 90s, this trait was one of high interest in early QTL studies, as it appeared in many ways to be caused by one, or very few, genes. Scientists in several groups used the approach described above to use markers and linkage maps to localize the chromosomal region containing the double muscling gene. Tim Smith and Eduardo Casas at MARC and Michel Georges at the University of Liege, in Belgium, working independently, were able to localize the QTL effect to bovine chromosome 2 in the mid-1990s. They then went to work to "fine map" the gene by looking at additional markers in this region. They also went to the maps of the other species to look for "candidate genes" that might fall into the regions of those genomes corresponding to that region of bovine chromosome 2. While they were making progress in pinpointing the specific gene causing the double muscling effect by finding markers more closely linked, there was still no clear picture of the specific gene.

The search for the double muscling gene took a strange twist, however, when human geneticists reported in the scientific literature a gene called "myostatin" that had been identified in mice as having a huge impact on muscle development and quantity. These researchers had noted that when this gene was deactivated (so called "knock-out" mice), the mice

developed muscle hypertrophy – just as observed in Belgian Blue cattle. And, furthermore, the myostatin gene mapped to a region of the human genome syntenic to that of bovine chromosome 2! Consequently, researchers were able to pinpoint the bovine myostatin gene to chromosome 2, and to identify a single nucleotide switch from guanine to adenine at codon 313 in the gene that caused the double muscling effect. Since the original bovine publication in 1997, the myostatin gene has been further studied with a total of 13 different polymorphisms found in this gene, several that are specific to various breeds expressing varying degrees of muscle hypertrophy. This is but one example of the power of "comparative mapping" to elucidate the underlying genes of importance in these QTL effects.

A handful of genes have been mapped in cattle through the "QTL-search followed by comparative mapping / fine mapping" approach. Two genes have been identified affecting carcass quality in cattle, both in Australia, that are now being marketed publicly under the label of "GeneStar" by a company called Genetic Solutions. GeneStar Marbling™ is the trade name for a genetic test of the thyroglobulin gene that has been shown to affect degree of marbling. GeneStar Tenderness™ is the tradename for a genetic test of the calpastatin gene that has been shown to affect meat tenderness. Genetic markers of the u-calpain gene have also been identified at MARC and have been shown to reflect a difference of between 0.5 and 0.8 lb of Warner-Bratzler shear force between alternate forms of the gene. Leptin, a protein important in energy metabolism first identified in the 1980s, has also been mapped in cattle by Canadian researchers, and is being studied to determine its effectiveness as a selection tool for regulating feed intake and energy metabolism in dairy cattle as well as altering carcass composition in beef cattle. Diacylglycerol acetyltransferase (commonly called DGAT) has been mapped by Michel Georges in dairy cattle and seems to have an effect on fat deposition in milk. Several new genes are now entering the commercial pipeline in Australia and the U.S., including somatostatin and a retinoid receptor gene, both affecting marbling.

While it has been very exciting to see the discovery of these few genes and see how they can be used to make genetic improvement, there are hundreds of other QTLs that have been identified for various traits that have not been successfully fine mapped to the gene level. Additionally, it is important to keep in mind that most of the economically important traits we seek to improve in cattle breeding are complex quantitative traits under the control of many genes simultaneously, and their interactions both with one another as well as the production environment. The genes with the largest effects will tend to explain only 10% or less of the genetic variation in these traits, which aside from myostatin is the case with all of the others described above thus far.

An example of the challenge confronted by genomics researchers "post QTL identification" is the story of the original gene mapping project funded by the Beef Checkoff and Texas A&M University at the Angleton station. As mentioned earlier, this project was conducted in the first half of the 1990s by constructing a resource population made by crossing Brahman to Angus followed by backcrossing to the two breeds along with some F_2 matings to produce full-sib families via embryo transfer. The result of this effort was that by the mid-1990s a series of QTLs had been identified through linkage mapping, including a high potential marker for marbling along with several for various meat quality attributes associated with tenderness. However, once the initial mapping effort was completed, the researchers and the beef industry were left with an uphill battle to try to determine if these markers would be meaningful for selection within breeds, or if they simply were reflective of the differences between the two widely divergent Angus and Brahman breeds for carcass traits. Thus, a second large-scale project was initiated in 1998 by the Beef Checkoff to "validate" these markers across the major breeds of beef cattle in the U.S., in cooperation with many of the breed associations. Each breed was asked to facilitate the collection of a minimum of 50 progeny from each of 10 sires chosen to represent the diversity in their breed. These sires and their progeny were then genotyped for the ten most important markers identified in the Angleton population (3 for shear force, 2 for sensory taste panel, 3 for marbling, and 2 for retail yield traits). This required a monumental effort to coordinate the

project and collect all of the phenotypic data from these animals. That project has now been completed (14 years after the initial launch of the Angleton project) – and while it has shown that some of the QTLs are segregating within other breeds, the use of this technology still has considerable limitations.

The experiences of the past couple of decades have led to the inescapable conclusion that progress in using genomic tools in beef cattle breeding (as well as in other livestock species) will be painfully slow using the approaches detailed above. Fortunately, the playing field is currently experiencing a transformation because of the bovine genome project.

Sequencing of the Human Genome to Sequencing Cows??

The human genetics community quickly recognized that if progress in building new tools through genomics for human health applications was to occur expeditiously, infrastructure needed to be built right up front. Linkage maps, QTL searches, comparative mapping, and some fine mapping were useful, but extremely inefficient, timely, and high in cost. Thus, in the last half of the 1990s, the National Institutes of Health, through its National Human Genome Research Institute, built a plan for sequencing the human genome, along with the highly used lab species of the mouse and rat. The project became broadly known as the "Human Genome Project" and involved a network of "sequencing centers" contracted to do high-throughput sequencing (i.e. determination of the DNA base code) of the human genome. These centers were at Baylor College of Medicine, MIT, Washington University in St. Louis, and the Sanger Centre in the UK. At the same time a scientist named Craig Venter came up with a different and novel approach for DNA sequencing called "whole genome shotgun sequencing" that he predicted would be much faster and more efficient than the approach being used by NHGRI. What transpired over the next few years was an ongoing debate and competition between the federal effort (i.e. NHGRI) and the privately funded parallel effort (i.e. Craig Venter through his new company known as Celera Genomics). An initial rough draft of the human genome sequence was completed in 2001, followed by a complete, finished sequence in April 2003, fifty years after Watson and Crick's initial elucidation of the double-stranded helical nature of DNA! The Human Genome Project was not cheap (in the billions rather than millions of dollars), but is widely believed by many to be the most important scientific project in the history of mankind to date. Obviously, as evidenced by the number of breakthrough discoveries occurring now on a routine basis, that may in fact prove to be true. It will be extremely exciting to see how the next decade unfolds in human medicine as a result.

The cattle, poultry, and swine industries, however, also have been placed in a position to reap huge rewards from the infrastructure built by NHGRI to sequence the human genome. In order to build the most comprehensive infrastructure to capitalize on the human genome for discoveries in human health, NHGRI launched down a path in 2002 of supporting the sequencing of a number of other genomes. These have been chosen to most highly leverage the investment in human genomics, as based on comparative mapping and medical model species use. Fortunately, the cow has been widely used as a model species in a number of areas for human medicine, especially in the area of reproductive physiology. As a result, the agricultural community developed a "partnership" approach in 2003 with NHGRI to move forward the sequencing of livestock genomes. There have been a number of strong voices that have moved these efforts forward, too many to mention here. The result, however, is that in March 2004 the draft sequence of the chicken genome was completed and released at Washington University and even more exciting to the beef industry is that the sequencing of the bovine genome was launched at Baylor College of Medicine's Human Genome Sequencing Center in Houston in December 2003. We recently were successful in garnering the funds to launch the sequencing of the swine genome in late 2005.

The bovine genome sequencing effort is expected to yield an 8-fold coverage sequence map of the genome by December 2005 with a cost of $53M. The funding sources of the effort include NHGRI ($25M), USDA ($11M), the state of Texas ($10M), Genome Canada ($5M), Australia and New Zealand

($1M each), and the national, Texas, and South Dakota beef councils ($0.8M). This follows an initial investment of over $4M to develop the scaffolding, called a bacterial artificial chromosome (BAC) map, invested by an international consortium of ten laboratories in seven countries, led by USDA-ARS. The animal providing the DNA for the sequencing project is a Line 1 Hereford female from the USDA-ARS long-term linebreeding and selection project at the Fort Keogh Livestock and Range Research Lab at Miles City, MT. This animal was selected to provide a higher chance of producing a high-quality sequence assembly because she carries an inbreeding coefficient of over 40%. All sequence information is being deposited in the public domain, through the NIH's National Center for Biotechnology Information (NCBI), as it is completed, allowing all researchers around the globe to have access to spurn forward developments.

In October 2004, the first draft assembly (3.3-fold sequence coverage) of the bovine genome was announced by the project team. As of March 2005, the sequencing had commenced to the 6-fold coverage level and light sequencing had been completed on a panel of animals representing the Holstein, Jersey, Angus, Limousin, Brahman, and Norwegian Red breeds to allow detection of new single nucleotide polymorphisms (SNPs). The process of validating a set of 20,000 of these SNPs has been initiated and will be carried out on a wide panel of breeds to evaluate genetic diversity of the world cattle population during the summer of 2005. Additionally, the Hereford female used for the sequencing project and one of her progeny recently supplied a wide array of tissues to the project team to allow development of full-length cDNAs for the study of gene expression in functional genomics projects over the coming years. Efforts over the remainder of the project will focus on the remaining 2-fold sequence coverage, to be done from a minimum tiling path of BAC clones and then development of gene predictions from the sequence information and gene annotation. Planning has also commenced by the research community on what infrastructure needs to be laid into place to fully capitalize on having this huge volume of new information available.

The availability of the genome sequence is expected to speed gene discovery by a factor of 100 fold! ***The fact that this effort will result in the sequence of the genome being made available in the public domain further adds to the impact of the investment as it will spurn on discoveries for the public good faster.*** We truly do live in exciting times.

What Will be the Practical Applicability of Genomics for Beef Cattle Breeding?

As genomic tools develop over the coming years, what can beef cattle producers expect to see as a result? Will DNA selection tools essentially replace breed genetic evaluation programs / EPDs as we know them today? Will we no longer need to worry about collecting expensive performance data? Will we essentially be able to know the genetic value of a calf in utero? Will we be able to predict the perfect range cow for a given production environment, sort that cow out with genomics, and then mass clone her? At various points in the past decade, there have been people who have painted the picture of the future by answering yes to all of these questions. What is the practical truth?

As genomics technology matures in the coming decade, we will see an explosion of genes that are identified for various traits. However, will that information give us all of the answers? Hopefully the message has been clearly delivered in this presentation that the science of genetic improvement in a business as multi-faceted as beef production is very complex. It is easy to predict that as we identify many of the genes underlying variation in performance for traits, we will identify more questions than we do answers. Some of those are likely to be:

1. **What is the function of these genes in the physiology of the animal and how is this function altered by changes in the production environment?** We are now routinely talking about the next big opportunity area of livestock genetics research being in "functional genomics".

2. **How do the various genes impacting**

economically important traits interact with one another at the genome and proteome level?

3. How many animals within a population (ie a herd or a breed) need to be genotyped for these gene tests in order to get ***enough*** information?

4. **Can we combine phenotypic performance information with gene level DNA information to come up with "DNA-enhanced EPDs"?**

5. How will the free enterprise system embrace this technology – i.e. what is the best business model to capitalize on these advances?

6. How will the cost of this technology be borne by the industry? One cannot expect the genetics or commercial sectors of the beef industry to pay $50 or more per test for a lot of genes to identify the top sires as has been proposed in the initial ventures of gene testing into the public marketplace. The value capture of this technology is likely to require a new type of business model than anything we have seen previously in cattle genetics.

Unfortunately, as human beings, we have not been granted the wisdom of our creator. However, some things do appear to be clear for the future. Genomics will provide revolutionary advances in our ability to genetically improve ***and*** better manage beef cattle, just as EPDs and other technologies have done before. Gene level information will add to performance information to give us more accurate EPDs at earlier stages of an animal's life for many traits. We may be able to add EPDs to our genetic evaluation system using DNA tools that we have not been able to afford before. But, in the end, this technology will be judged by the marketplace in an industry with historically slim margins, and will only be successful if it is priced relative to value delivered. Current commercial efforts to market tests on a gene by gene basis at $75 or more a pop are not sustainable. Entepreneurs will need to be cognizant of the fact that such a price tag may need to deliver the molecular picture for an entire segment of performance (i.e. end product value, cow herd input costs, etc.) rather than individual components on an individual gene basis.

Lastly, it is important to point out that much of the research and development in big areas like genetic improvement has historically been required to be done with public funding. The future of genetic improvement in beef cattle will still need to rely on this approach. This means that the cattle industry must be active in supporting and encouraging research that will contribute to increased efficiency of high quality beef products using environmentally sustainable production systems. For example, while private industry may choose to focus efforts on developing genetic tools to allow improvement in end product quality – for example tenderness – it will be difficult or impossible for the same to be done on a trait like feed efficiency or cow herd maintenance requirements or reproductive rate. These are the areas that we will need to place full effort upon in the future of our public research efforts in order to continue to solve problems of importance to the industry and the consumers of our products. There is much work yet to be done by beef cattle breeding and genetics researchers – thus we need to be attracting and educating our best young minds into this area. Furthermore, the trends we have seen in more recent years of decreasing investments in traditional population genetics based programs must be reversed given that we are now close to coming back full circle to needing these folks to help the gene jockeys interpret their new data.

Wouldn't it be a blast to be able to see where it all goes in the next 100 years? My gut instinct tells me that we likely would not have been bold enough in our predictions.

Acknowledgements

The author acknowledges the helpful assistance of colleagues in developing the information in this presentation including:

USDA-ARS: Steve Kappes, Larry Cundiff, Mark Thallman, John Keele, Tim Smith, Roger Stone, Eduardo Casas, Curt Van Tassell, Cal Ferrell, Leland Ellis

University Researchers: John Pollak (Cornell), Daniel Pomp (University of Nebraska), Noelle Cockett (Utah State University), Clare Gill (Texas A&M University), Richard Gibbs and George Weinstock (Baylor College of Medicine), Larry Schook and Harris Lewin (University of Illinois)

Kellye Eversole (Alliance for Animal Genomics Research)

Jane Peterson (NHGRI, NIH)

Notes:

Notes: