Where We Are Going with Genomics and Genetic Improvement:

Matt Spangler, University of Nebraska-Lincoln and Alison Van Eenennaam, University of California-Davis

Historical Summary

It has become clear that that real-time changes in genetic merit estimates, also known as Expected Progeny Differences (**EPD**), are not well accepted by many in the beef industry. Increasing the accuracy of prediction much earlier in life does not allow for an extended time horizon to justify changes in EPD based on empirical evidence (e.g., visualizing differences in offspring). This, coupled with changes in prediction methodology that have taken longer to implement than desired, have led many to question the validity of genomic selection and the entire infrastructure for National Cattle Evaluation (**NCE**). However, we contend that tremendous progress has been made in a relatively short amount of time, and that the scientific issues we currently face pale in comparison to the issues created by an overall lack of cohesion in the beef sector and the general reluctance to adopt technology.

Before we prognosticate about what the future may hold, let us take a look back to see how we got to the point at which we are currently. As different breed associations began including genomic information (primarily genotype data from the 50,000 (**50K**) single nucleotide polymorphism (SNP) chip) into their NCE, the nuances related to methodology for doing so increased. The method used by the American Angus Association (AAA) was first proposed by Kachman (2008) and used by MacNeil et al. (2010) in their prototype evaluation. This became known as the "correlated trait approach" and assumed that the linear combination of the 50K SNP genotypes known as the Molecular Breeding Value (**MBV**) could be fitted as a correlated indicator trait in existing multiple-trait models. A primary benefit of this was the familiarity of the concept to breed associations. It also allowed for genomic information to influence the predictions of animals in the pedigree that were not genotyped.

As other breeds began to include genomic information into their NCE, "new" methods of doing so were being developed. It is important to note that the choice of inclusion method was arguably based on the genetic service provider (entity that conducted NCE) and not through purposeful model comparison. The majority of breeds that followed implemented a blending (indexing) approach whereby the MBV and EPD were indexed together to produce a "genomically enhanced" EPD (**GE-EPD**). This has primarily been done post evaluation and consequently only impacted the prediction of the genotyped animal. This created the largest difference between blending and the correlated trait approach. All of these methods are essentially variations on the same two-step theme; estimate the SNP effects using a large data set of genotyped and phenotyped animals from the same "breed" to train the MBV and then fit them into NCE. Since 2009, many breeds have made tremendous investments in this technology and currently offer GE-EPD. Table 1 represents counts of genotyped animals as of fall 2016.

Table 1. Number of Animals Genotyped, use of Low-Density (LD) Panels for Imputation, Method of Incorporation of Genomic Data into National Cattle Evaluation (NCE), and Genotyping Service Provider by Breed.

Breed	Samples Included in NCE[1]	LD Panel to Impute	Method[2]	Service Provider[3]
Angus	264,656	Yes	Corr	G,Z
Hereford	~23,000	Yes	Blend	G
Red Angus	22,791	Yes	Blend	G,Z
Charolais	2,454	No	Corr	G
Gelbvieh	10,162	Yes	Blend	G
Limousin	3,340	Yes	Blend	G
Simmental	32,629	Yes	Blend	G
Shorthorn	~1,000	Yes	Blend	G
Brangus[4]	3,909	Yes	ssGBLUP	G,Z
Santa Gertrudis[4]	3,160	No	ssGBLUP	G

[1]These are the number of either high density or low density samples included into NCE. Some breeds have access to additional genotyped animals for training and research purposes. These counts do not include legacy 384 SNP panels, although for Angus these are being included (n=26,282; unknown number of duplicates with higher density panels).

[2]Corr=A correlated trait approach; Blend = post-evaluation indexing; ssGBLUP = single-step Genomic Best Linear Unbiased Prediction.

[3]G=GeneSeek; Z=Zoetis

[4]Updates as of Oct. 2016

We know that the inclusion of genomic information into NCE can add accuracy to EPD, particularly for young animals. The benefit of increased accuracy, and perhaps the salesmanship associated with being a technology adopter, has spurred rapid genotyping in several breeds as evidenced by the counts in Table 1. The availability and use of low-density (**LD**) panels (fewer SNP and less expensive) has also aided in the increased rate of genotyping but more importantly in the fraction of animals within a contemporary group that are being genotyped. The ad hoc selective genotyping strategy whereby only the "best" animals were genotyped was undoubtedly a disservice to NCE and genomic selection. Early prediction equations were built based on a highly-selected subset of animals and as a consequence bias was introduced. The ability to affordably genotype entire contemporary groups can resolve this issue. However, the technology will need to continue to decrease in cost toward commodity based pricing before the strategy of genotyping every animal becomes pervasive. We suspect collective bargaining, with all breed associations engaged and on the same side of the table, could help to drive the cost of genotyping down.

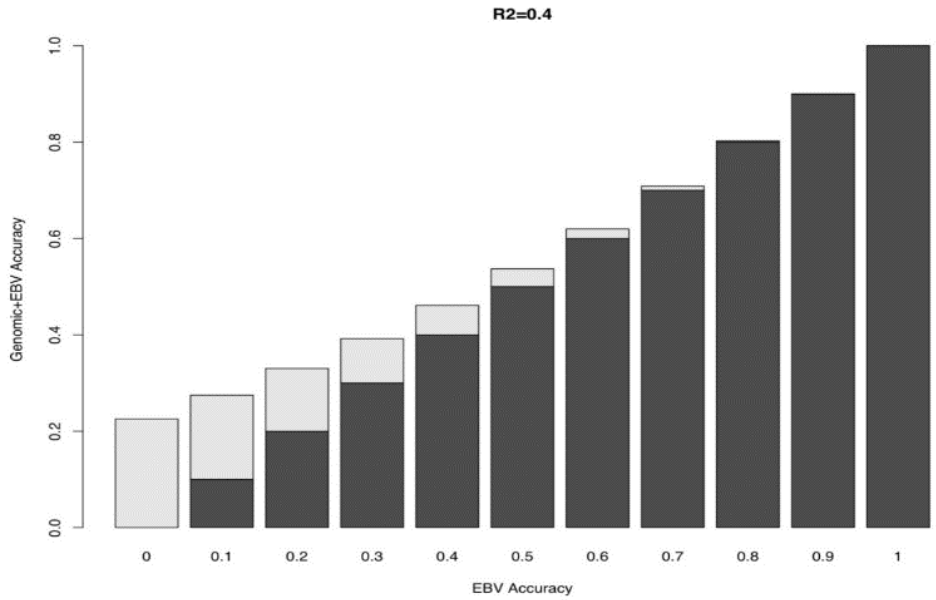

Figure 1. Increase in EPD Accuracy Resulting from Inclusion of Genomic Information That Explains 40% of the Genetic Variation (GV; squared genetic correlation). The darker portion of the bars shows the EPD accuracy before the inclusion of genomic information and the lighter colored portion shows the increase in accuracy after the inclusion of the MBV into the EPD calculation. As the %GV explained by the genomic information increases, the increase in EPD accuracy becomes larger.

The use of LD SNP panels presented a critical question to breed associations relative to the enabling of imputation going forward in time. For example, if the use of LD panels (e.g. 20,000 SNP) becomes commonplace and the target SNP density for inclusion into NCE is of higher density (e.g. 50K) how would breed associations ensure that imputation could be performed given that the animals with actual HD genotypes would become further and further removed from the target population? This issue made breed associations decide to re-genotype critical animals (e.g. prominent sires) with a high density panel to ensure imputation could proceed with an acceptable degree of accuracy. This re-genotyping is often subsidized by the breed association. To our knowledge, the choice of which animals to re-genotype is done by setting a relatively arbitrary threshold relative to the number of daughters in production, calves registered, accuracy of a certain EPD, etc. Gains in efficiency could be achieved by using more advanced criteria, such as calculating the relatedness of the genotyping candidate to the rest of the population and approximating the gain in imputation accuracy that could be achieved by re-genotyping an individual. It is unclear if the resulting cost savings would justify the added effort of doing so.

The methods of inclusion of genomic information into NCE have been relatively static over the past 5 years despite a considerable amount of research. The ongoing research is focused on the commercial scale implementation of various "single-step" methods, some that allow variable weighting of different SNP effects and others that do not. Presumably, all beef breed associations will move to one of these two methods (software provided by University of Georgia or Theta Solutions) by early July of 2017 (this timeline is a prediction and not a guarantee). This is an exciting evolution, particularly given some breed associations currently use software for

NCE that is decades old and that has undergone several "patches" along the way to keep it afloat. It will also be beneficial to move away from the possibility of using MBV or "percentile ranks" based on MBV alone as selection criteria.

It will also benefit any breeds that are using differing prediction equations from multiple labs, which creates confusion at best and at worst adds unwanted heterogeneity and differing sources of bias to post blended EPD. However, these are not truly single step methods, and they still rely on imputation to a common density. In some cases, this means data are cast aside if, for example, the target density is the 50K and animals are genotyped with a higher density. Although data are being cast aside, we suppose it is arguable that information may not be lost given panels with a density higher than 50K have not yielded substantial increases in accuracy.

Priority Areas

Despite the tremendous progress that has been made over the past several years, challenges still exist. Although not an exhaustive list, we believe the three bullet points below summarize the current hurdles to further refine genomic selection.

1) Improve the portability of genomic predictions.

It is well known that the accuracy of genomic predictions erode as the target population becomes more distantly related to the training population. This can occur over time, but perhaps the larger concern is across differing breeds. Kachman et al., (2013) clearly illustrated that a 50K-based genomic predictor for weight (either yearling or weaning) that was trained in Angus was not predictive in Red Angus. The issue of robustness, or portability, of genomic predictions across breeds is critical for three reasons: 1) Not all breeds will have the resources to adequately estimate markers effects for all traits, 2) prediction in non-pedigree commercial populations will remain elusive unless this issue is resolved, and 3) The transfer of genomic information from research settings for novel traits to industry populations will not occur otherwise.

Initially there existed hope that simply increasing the marker density (e.g., going from 50,000 SNP to 770,000 SNP) would alleviate this problem. It did not. Simply adding more markers actually reduces statistical power when the number of genotyped animals does not increase proportionally and results suggest that any gains in using higher density panels are negligible at best both within and across breeds. The new hope comes from a growing body of whole-genome sequence information. The omnipresent thought is that "functional" variants can be identified from sequence data and used to construct lower-density panels. If truly functional variants are identified, they should not be subject to the reliance on linkage disequilibrium (non-random association of alleles at different loci in a given population) that plagues the use of markers alone, and thus they should be valuable predictors across populations. This is easier said than done.

First, we must develop a system of categorizing DNA variants that provides more resolution than is currently used (e.g., a variant is classified as important—but how important is it, and what is the evidence of this?). This will help narrow the list of candidate variants. Secondly, given a pragmatic view of how well we can identify functional variants, we must refine the methodology we use to estimate genomic predictors. For example, there is evidence that haplotype based models may be more robust in admixed populations in terms of prediction accuracy and resolution of QTL locations (e.g., Schweer et al., 2016). Additionally, encouraging results using

identified variants from whole genome sequence information that are contained on the new GGP-F250 panel are becoming available (e.g., Snelling et al., 2017). In this paper, Snelling and colleagues reported that 293 variants explained 36% of birth weight genetic variation in the Germ Plasm Evaluation project (GPE) at the US Meat Animal Research Center, and molecular breeding values trained using GPE effects had genetic correlations with birth weight in other populations ranging from 0.25 to 0.44. Similar correlations were obtained from a subset of SNP that contained only 11 variants. Genetic correlations between birth weight and genotypes for the single most significant variant in GPE were between 0.17 and 0.34 in the independent populations. Although we have a considerable amount of work yet to do, the incorporation of biological information into our predictions of genetic merit using genomic data seems encouraging.

2) Improve phenotypic data recording for traits that are commercial industry profit drivers.

The principal reasons that genomic selection has worked well in dairy cattle are that Holstein is a homogeneous population, and the selection objective focuses primarily on sex-limited traits (e.g., milk production). In contrast, the majority of EPD in the beef industry are not sex-limited and represent phenotypes, or indicators of the desired phenotype, that can be collected on bulls at or before 12 months of age. Exceptions include heifer pregnancy and measures of sustained cow fertility (e.g., stayability). Although fertility EPDs do exist in some form for several beef breed associations, the information content is not sufficient. This is due to a combination of factors including the lowly heritable nature of these traits, and both the quantity and quality of data reported. In other words, phenotypic data collection needs to be ramped up.

There are other traits that are economically relevant to the commercial industry that are either sparsely collected or non-existent in current breed association databases. Traits such as disease susceptibility (Bovine Respiratory Disease (BRD), pinkeye, etc.), carcass data including primal yields, mature cow weights, male fertility, cow feed intake, and the list could go on. The majority of these phenotypes exist in the commercial sector (cow/calf, feedlot, and packer) and are collected in some form every day. To fully capitalize on genomics, we must exploit the data that exists in our industry and ensure it enters into NCE. It is obvious that breeds without a solid NCE foundation cannot make use of genomics, and for many traits of economic importance all breeds fall into this category. Unfortunately, these are the traits for which genomics could help the most—those that are expensive to collect, collected on older animals, or sex-limited. It is not the entity that genotypes the largest number of animals, or the entity that first implements single step genomic selection that will win the NCE race. Rather it is the entity that is able to fully exploit commercial level data in genetic prediction that will gain the most from genomic selection.

3) Improve the understanding and utilization of genetic selection tools.

Psychology might be a better degree to hold to solve this issue as compared to quantitative genetics. The implementation of genomic selection is only advantageous if breeders, particularly nucleus breeders, believe in and utilize traditional EPD and selection indices. Moreover, commercial producers must value increases in EPD accuracy as a means of mitigating risk. If these two qualifications are not met, genomic selection in beef cattle is futile.

Currently, there are people in leadership positions who believe publishing the actual MBV is valuable, in addition to publishing GE-EPD. This illustrates that continued educational efforts

relative to genomic selection and the outputs of NCE have somehow fallen short. Perhaps part of this can be attributed to the survey findings of Weaber et al. (2014) regarding where beef producers seek genetic selection information. Interestingly, Weaber and colleagues reported that unpaid consultants, such as neighbors or friends, were most frequently (38.9%) identified by respondents as valuable sources of breeding and genetics information followed by veterinarians (29.7%), extension professionals (29.5%), seedstock producers (27.7%), internet search (18.9%), farm supply or feed store staff (18.1%), breed association personnel (14.7%), AI stud personnel (11.7%), popular press sources (9.3%) and paid consultants (2.1%). These results suggest that it is important to educate not only traditional information providers (veterinarians and extension educators), but also commercial producer peers and their seedstock suppliers about genetic and breeding principles as these entities are often consulted as trusted sources of genetic selection information (Weaber et al., 2014).

The traditional vehicle for outreach has been face-to-face delivery of educational and written material. These delivery approaches are generally targeted towards increasing knowledge and awareness. Unfortunately, despite decades of effort using these two traditional approaches to outreach, little has been accomplished relative to attitude and behavior changes. Survey results suggest that upwards of 70% of U.S. beef cattle producers in the commercial sector do not utilize genetic merit estimates, EPD, as their primary selection criterion (e.g., Weaber et al. (2014)).

Using the thesis that current adoption of fundamental genetic selection tools by bull buyers is archaic, and that traditional means of outreach have not been able to penetrate the beef industry such that behavior changes have occurred, a new approach was deemed necessary to ensure technology adoption of emerging tools like genomics. A hands-on approach where beef cattle producers could 'learn by doing' was trialed to augment traditional outreach vehicles. Moreover, this approach leant itself to training beef cattle producers and breed association personnel to be effective educators themselves. The latter point is critical given the general lack of outreach personnel in the United States that are trained in quantitative genetics/genomics.

In 2009, an integrated effort between the National Beef Cattle Evaluation Consortium (NBCEC), the University of Nebraska, and the 7 largest beef breeds in the U.S. (Angus, Hereford, Red Angus, Charolais, Gelbvieh, Limousin, and Simmental) was initiated in an effort to develop an educational program centered on genomics and to build a resource population for the development and evaluation of genomic predictors and related methodology. These 7 breed associations 'nominated' seedstock producers (n=24) in the Northern Plains region of the U.S. to participate in the project. Initially, producers agreed to provide hair samples on all 2009 born bull calves. These animals were genotyped with a reduced assay for weaning weight and post-weaning gain. The SNP discovery for this reduced assay occurred in the Cycle VII population at US Meat Animal Research Center (**USMARC**). Given the early focus on weight traits as proof of concept, the project was named the Weight Trait Project (**WTP**).

In subsequent years, producer-owned herd bulls were genotyped with the 50K, and MBV and marker-assisted EPD were provided back to producers for growth and carcass traits. The MBV were trained using currently available genotypes in the NBCEC database using both within-breed and across-breed training sets. All genotypes generated were provided to the respective breed associations to aid in the development of training sets that would eventually be used to generate MBV that were included into NCE.

As part of the WTP, a two-day meeting has been held annually at the USMARC, with the first day focused on short (approx. 20 minute) presentations accompanied by brief (2-page) handouts. Talks on the first day of the meeting have focused on the current status of genomic selection in beef cattle, novel trait discovery and, in more recent years, considerations related to selection for improved feed utilization. All talks have been recorded and posted at www.beefefficiency.org. All attendees of the first day meeting were asked to complete an anonymous survey indicating levels of knowledge gained and any likely behavior changes as a result of the presentations. They were also asked to provide an indication of numbers of beef cattle they owned or for which they directly influenced management decisions. On the second day, activities centered on project aims and results, and upcoming project activities. This forum allows for direct industry feedback from progressive seedstock producers and breed association personnel related to the direction of genomics research and issues of technology adoption.

The impact of an outreach program is best evaluated by changes in behaviors and practices of targeted producers and the industry at large. Of the 7 beef breeds represented in the integrated project, all have implemented GE-EPD. The WTP arguably aided in developing the framework for these breeds to develop a training population and empowered a group of seedstock producers to educate their peers relative to the advantages of genomic selection.

A survey was conducted by Spangler et al. (2011) to gauge changes in knowledge, practices, and behavior; the survey was sent to participants in the WTP. The 17 respondents indicated that collectively they own 20,125 beef cows. Increases in knowledge were rated from 0 (none) to 4 (significant). Mean survey results were 1.5, 2.8, 2.0, 3.4, 2.4, 2.7, 2.8, and 2.9 for EBV, genomics terminology, parentage verification, marker assisted selection, across breed genomic predictions, whole genome selection and panel development, test validation, and accuracy improvement of EBV, respectively. Producers indicated adoption of methods to improve the following production practices: making mating decisions (40%), efficient use of DNA technology (75%) and selection (bull buying) decisions (47%). Mean responses for changes in behavior (1 = none; 5 = very likely) were 3.9, 3.8, 4.3, and 4.6 for making more informed selection decisions, better educating their clientele, feeling comfortable with terminology, and desiring to stay abreast of DNA technology, respectively.

A critical outcome of this integrated project is the development of a forum for researchers, breed association personnel, and seedstock producers to continue a dialogue regarding genomic technology, implementation methods for genomic selection, and discovery for novel traits. Consequently, these types of activities are likely better suited at generating behavior change than classical extension talks during an industry-sponsored meal.

Emerging Technologies in Beef Cattle Breeding

Genome Editing

Genome editing is a category of new methods that can be used to precisely edit or change the sequence of DNA or the genetic code. As the name “genome editing” suggests, these technologies enable researchers to add, delete, or replace letters in the genetic code. In the same way that spell check identifies and corrects single letter errors in a word or grammar errors in a sentence, gene editing can be used to identify and change the letters that make up the genetic code (i.e. DNA) within an individual.

Gene editing has many potential applications. For example, it can be used to correct diseases and disorders that have a genetic basis. It could also be used to change a less desirable form of a gene (called an allele) to a more desirable allele without the need to introgress (repeatedly backcross) or bring in that allele through outcrossing with an animal that carries the desirable allele. Therefore gene editing is really more like precision breeding where breeders can introduce the specific sequences that they would like to select for using gene editing technologies.

Gene editing is different from "traditional" genetic engineering. Continuing with the analogy of a word processor, genetic engineering enables a gene sequence of "foreign DNA" to be "cut and pasted" from one species to another; typically the location where the new DNA sequence inserts into the genome is random. Gene editing can add, delete, or replace a series of letters in the genetic code at a very precise location in the genome.

Without the addition of template DNA, the double stranded breaks created at a precise location in the genome by the nucleases are repaired by the cell's natural DNA repair mechanism in a process called "nonhomologous end joining" (NHEJ; Figure 2). This typically results in single nucleotide changes, deletions or small (1-2 nucleotide) insertions at the DNA cut site. In this case, although the location of the cut site is very precise, the exact change that occurs when the DNA is repaired is random and so a number of different outcomes representing minor sequence changes are possible.

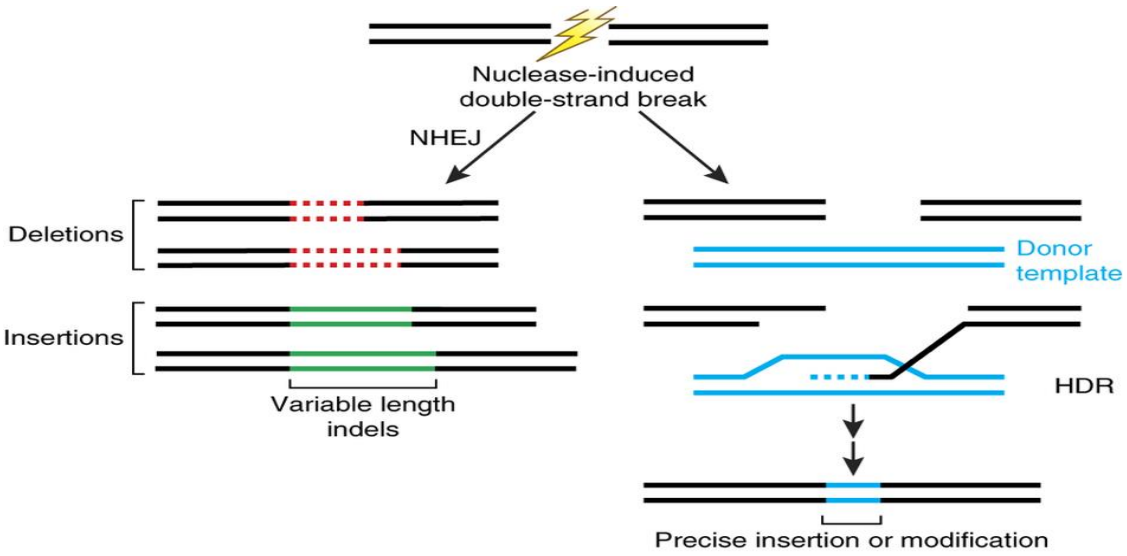

Figure 2. Nuclease induced double-strand breaks (DSBs) can be repaired by "nonhomologous end joining" (NHEJ) or "homology directed repair" (HDR) pathways. Imprecise NHEJ-mediated repair can produce variable-length insertion and deletion mutations at the site of the DSB. HDR-mediated repair can introduce precise point mutations or insertions from a single-stranded or double-stranded DNA donor template (blue). Image from (Sander and Joung, 2014).

Supplied with a nucleic acid template, however, the double stranded breaks initiated by the nucleases are repaired via the cell's "homology directed repair" (HDR; Figure 2) pathway whereby the template dictates the sequence resulting from the repair, allowing the introduction of the DNA sequence dictated by the template into the host genome. Such changes might range

from nucleotide-specific changes, to large (whole gene) insertions or substitutions depending upon the template. The end result of this maybe a targeted SNP edit (e.g. the nucleotide A at a given location in the genome is deliberately replaced by T), the replacement of one naturally occurring allele with another naturally occurring allele at a targeted genetic locus in that species, or the introduction of a novel DNA sequence as encoded by the template at the target location in the genome.

The currently available set of gene editing tools (zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered regulatory interspersed short palindromic repeats (CRISPRs) associated system) have been used for a relatively small number of livestock applications to date. Several recent reviews describe the potential to use these tools in food animals for agricultural purposes (Bosch et al., 2015; Laible et al., 2015; Murray and Maga, 2016; Tan et al., 2016), and include detailed descriptions of their mechanics and relative efficiencies.

Gene editing has been used to produce genetically hornless Holstein dairy cattle by replacing the Holstein "horned" allele with the naturally-occurring Angus "polled" allele at the gene that is responsible for horn development (Carlson et al., 2016), and to generate pigs with a single base deletion in a gene that may confer resilience to African Swine Fever Virus (Lillico et al., 2016). Recently a paper was published showing that gene edited pigs were protected from porcine respiratory and reproductive syndrome (PRRS) virus, a particularly devastating disease of the global pork industry (Whitworth et al., 2016). It has also been used to introduce changes in the *myostatin* gene in sheep and cattle (Proudfoot et al., 2015). As the Latin origin of the word *myostatin* (muscle/stop) suggests, turning off this gene results in muscle growth. Naturally-occurring mutations in this gene have historically been selected by conventional animal breeders and are the genetic basis for the "double muscled" phenotype that is seen in cattle breeds like the Belgian Blue, and the "bully" phenotype in whippet dogs.

In this way, gene editing really mimics the natural processes that form the basis of selective breeding programs, and for that matter, natural selection. Breeders work with the genetic variation that exists within a species, and that genetic variation ultimately arises from naturally-occurring mutations. Although the word "mutation" sounds negative, it simply refers to variations in DNA sequences. These variations, or mutations, are responsible for virtually all genetic differences which exist between individuals, such as having blue eyes instead of brown.

Although different mammals have many of the same genes, many people do not appreciate that the genetic code that makes up those genes differs among animals of different breeds, and even among animals within the same breed. In fact, with the exception of identical twins, there are literally millions of DNA sequence variations between two individuals of any species. For example, an enormous number of genetic variants have accumulated within cattle since the advent of domestication and selective breeding due to the naturally-occurring processes that lead to a small number of mutations each generation. In one recent analysis of whole-genome sequence data from 234 taurine cattle representing 3 breeds (Daetwyler et al., 2014), more than 28 million variants were observed, including insertions, deletions and single nucleotide variants. Most of these mutations are silent and have no impact on traits of importance to breeding programs. Occasionally, such mutations result in a genetic condition such as red or black coat color or an undesirable disease condition such as dwarfism.

Sequence Data

Some of the large-scale genomic and sequencing projects have revealed a number of single nucleotide polymorphisms (SNPs) and haplotypes in which one naturally-occurring allele results in superior performance to that observed to be associated with an alternative allele. Consequently, an animal's genome could theoretically be edited to the superior allele at one or more genomic locations. To date, targeting different genes simultaneously has allowed bi-allelic modification of as many as three genes at once. Multiple favorable alleles are rarely found in a single individual, and gene editing offers an advantage over conventional selection by efficiently increasing the frequency of desirable alleles in an individual, or even an entire breed, by moving naturally-occurring alleles without also bringing along all of the unwanted alleles that come along with conventional crosses to introduce a desired allele. This is referred to as "linkage drag" and is used to describe the (usually undesirable) effects of alleles at genes located adjacent to the allele we are trying to introgress. If a desirable allele for trait X lies close to an undesirable gene affecting trait Y, you will want to "break" the linkage drag – that is, separate the good allele from the bad.

In order for gene editing to be an important factor for genetic change, it must integrate smoothly into conventional animal breeding programs and reliably edit the germline of single cell zygotes that will form the breeding stock of the next generation. Gene editing could theoretically be applied to many different traits in livestock, including known fertility impairing haplotypes, and to correct known Mendelian genetic defects, in conjunction with conventional selection methods to continue making progress towards a defined selection objective. It also provides a means by which the discovery of causative SNPs (Quantitative Trait Nucleotides; QTNs) through sequencing projects and the information obtained from various genome wide association studies (GWAS) could be translated into valuable genetic variation for use in animal breeding programs. In one simulation study, response to selection was improved four-fold after 20 generations as a result of the combined use of gene editing and traditional genomic selection (Jenko et al., 2015). At best, gene editing will be used to complement conventional breeding programs; it will not replace them.

Although these methods offer many advantages, it is important to understand that hundreds, if not thousands, of different genes and their interactions impact complex traits. As a result, not all of the genes that influence these traits have been identified, so the sequences of the desirable alleles are not always known. For now, it is likely that relatively large effect loci and known targets will be the focus of editing in efforts to correct genetic defects or decrease disease susceptibility. The backbone of breeding programs will continue to be conventional selection in which selection for many small effect loci that impact complex traits will contribute to the breeding objective.

There have been a multitude of genome wide association studies (GWAS) performed over the past decade on all manner of traits, and large scale whole genome sequencing projects. Yet despite all of this information there are few obvious targets for editing at the current time, other than those associated with simple qualitative traits where one allele has a known affect (e.g. polled). As we develop a more sophisticated understanding of gene networks and quantitative trait variation, additional targets will likely be identified. In the future we may use editing to introduce specific alleles into maternal lines, without diluting the genetics that makes them superior maternal lines. We may even make maternal lines homozygous at certain alleles, and terminal lines homozygous at alternative alleles so that every mating results in a heterozygous individual with maximal heterosis. It may be used to ensure maternal lines have adequate

carcass merit by making targeted edits in loci associated with meat quality, yet continue to excel in maternal traits.

To emphasize it relative role in a breeding program, we can envision breeding programs as an ice-cream sundae as shown in Figure 3. Genome editing can be analogized as the cherry on top of all of the other components that are part of genetic improvement programs in the beef industry.

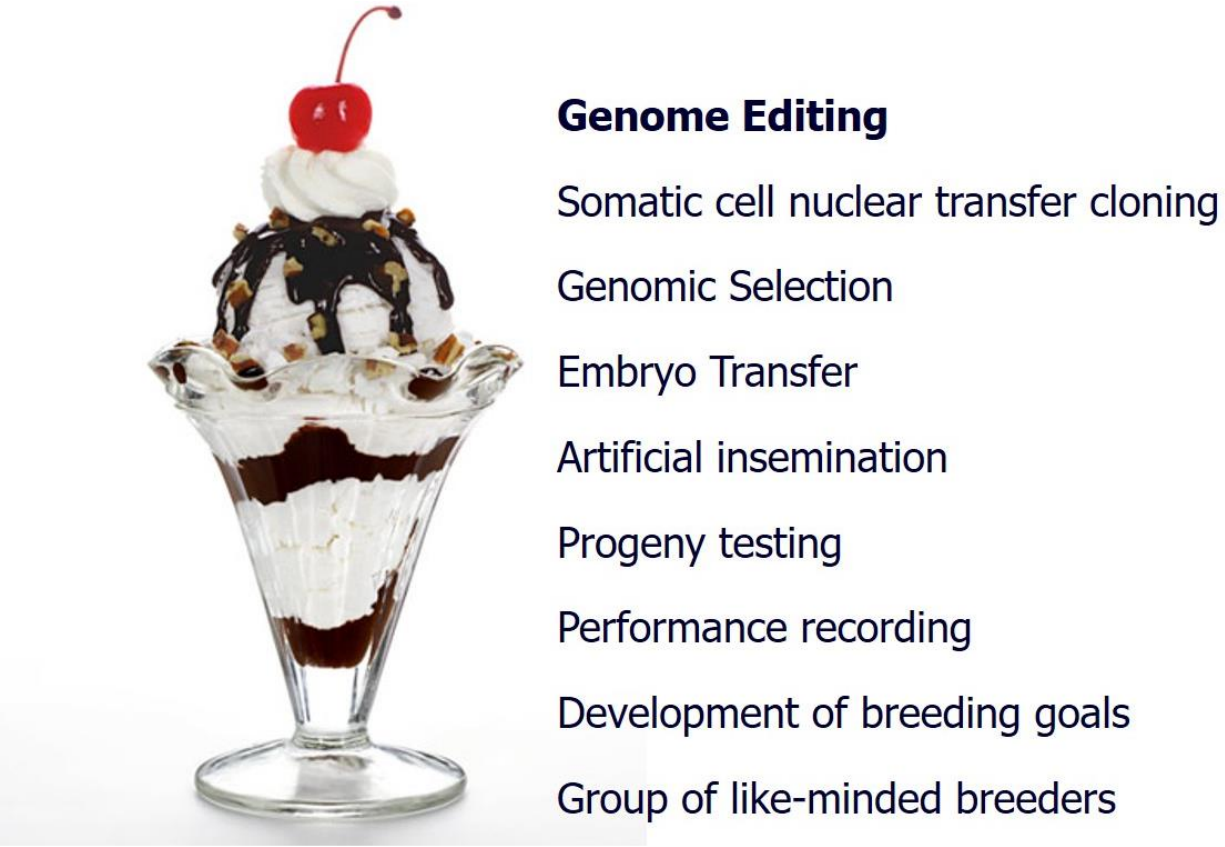

Figure 3. Schematic representation of genetic improvement programs in beef cattle. There are many requisite and interacting components that must be in place to drive genetic improvement.

Combining Advanced Reproductive Technologies with New Breeding Methods

It is perhaps underappreciated how much assisted reproductive technologies (ART) such as ovum-pick up and in vitro production (OPU-IVP) are being combined with the use of genomic selection (GS) in beef cattle breeding. While GS can decrease the generation interval in conventional cattle breeding by allowing for the more accurate genetic evaluation and use of young bulls, the expected benefits of combining GS and OPU-IVP far exceed the benefits achieved by either GS or OPU-IVP alone due to the very large reduction in generation interval (Kadarmideen et al., 2015).

In 2013 the global bovine embryo market reached 1,275,874 embryos, of which 40.6% (517,587 produced embryos) were IVP embryos. Brazil was responsible for 70.8% (366,517) of these IVP of embryos. In 2013, 45.7% (167,452 embryos) were obtained from dairy donors (88.6% from *Bos taurus* females) and 54.3% (199,065 embryos) from beef cattle (86.8% from *Bos indicus*

females; Viana et al., 2015; University of São Paulo, Brazil; unpublished data) (Kadarmideen et al., 2015).

Given there are so many OPU-IVP embryos being used in cattle breeding, editing may also have a role in reducing genetic lag. Genetic lag is defined as the time it takes for any genetic improvement made in the selection program of the top tier of the breeding pyramid (i.e. the nucleus seedstock sector) to trickle down to commercial sector. If genome editing can be reliably used to produce the desired edits in developing embryos, it could be routinely used to introduce useful genetic variants into newly fertilized embryos that are going to be part of an embryo transfer program.

In livestock to date, the primary method to deliver nuclease-mediated genetic changes has been cell culture followed by somatic cell nuclear cloning (SCNT). This method is advantageous because it allows for genotyping and/or screening of the gene edited cell line before it is transferred into the enucleated oocyte. This ensures that only the desired edits are made. On the downside, SCNT is associated with well-documented drawbacks such as early embryonic losses, postnatal death, and birth defects. Figure 4 shows how editing could fit into a selection program using advanced reproductive technologies combined with genomic selection (Van Eenennaam, 2017).

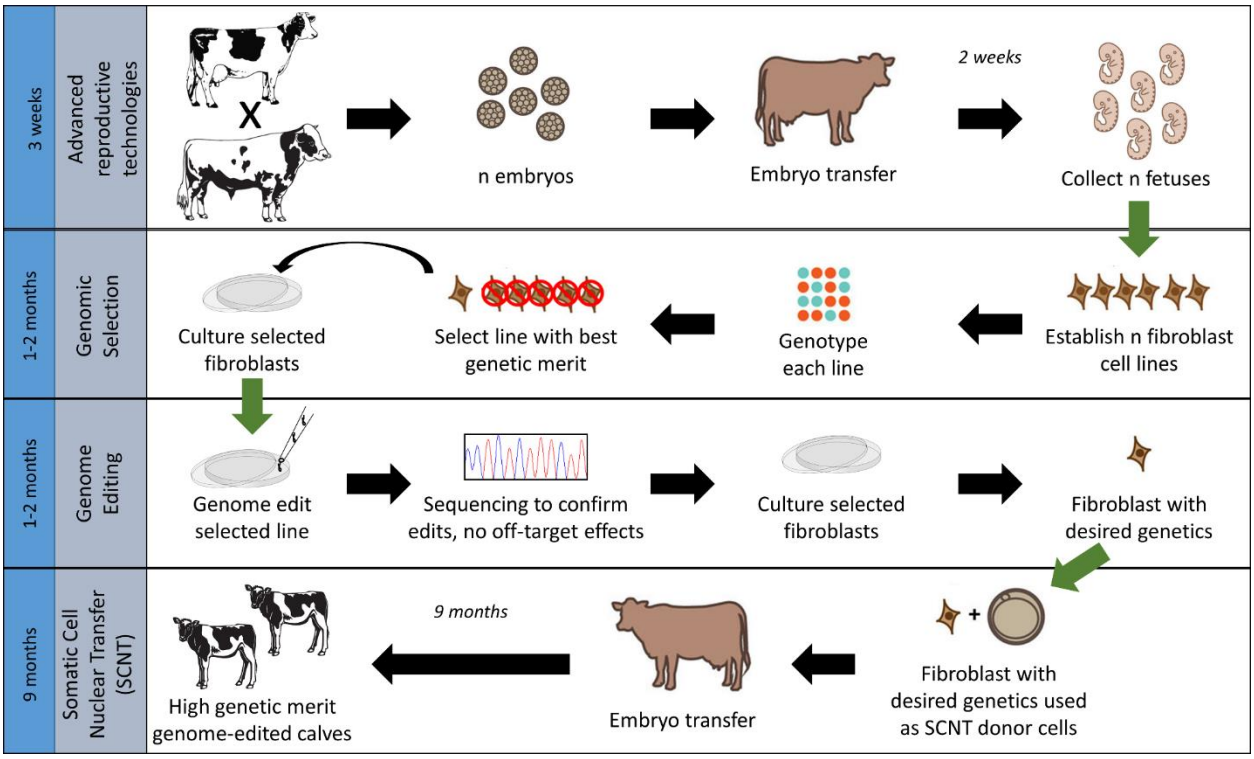

Figure 4. Production of high genetic merit calves using a range of biotechnologies and showing where gene editing might fit into the process. Collection of day 21-23 early stage embryos and the establishment of cell lines from them allows rapid determination of genetic merit for a large number of candidate embryos, the best of which would be selected for subsequent editing. Image from (Van Eenennaam, 2017)

Alternatively, editing of single-cell zygotes offers an approach to introduce edits directly into the next generation; however, the disadvantage is that not all of the embryos will be correctly edited. Despite this, direct editing is more desirable than SCNT since edited embryos gave a two times higher pregnancy rate, and fewer embryos are required, on average, to achieve the desired result (Tan et al., 2016). Direct editing of zygotes has successfully been used to knock-in entire interspecies allele substitutions (Peng et al., 2015; Lillico et al., 2016). Issues with mosaicism, meaning that some of the cells are edited and some are not as the edit occurred only in a subset of cells after the embryo began dividing, have been associated with this method, but researchers are developing approaches to edit the one cell embryo prior to the first cell division to minimize this problem.

How will gene editing be regulated?

Animal breeding *per se* is not regulated by the federal government, although it is illegal to sell an unsafe food product regardless of the breeding method that was used to produce it. Gene editing as a tool does not necessarily introduce any foreign genetic DNA or "transgenic sequences" into the genome, and many of the changes produced would not be distinguishable from naturally-occurring alleles and variation. As such, many applications will not fit the classical definition of "genetic engineering". For example, many edits are likely to edit alleles of a given gene using a template nucleic acid dictated by the sequence of a naturally-occurring allele from the same species. As such there will be no novel recombinant DNA (rDNA) sequence present in the genome of the edited animal, and likewise no novel phenotype associated with that sequence. It is not evident what unique risks might be associated with an animal that is carrying such an allele given the exact same sequence and resulting phenotype would be observed in the breed from which the allele sequence was derived.

In January 2017, the FDA expanded the scope of its "Guidance for Industry #187" for producers and developers of genetically improved animals and their products to address animals whose DNA has been intentionally altered through use of genome editing techniques. The new guidance (Food and Drug Administration, 2017) entitled, "**Regulation of Intentionally Altered Genomic DNA in Animals**" triggers mandatory, pre-market FDA new animal drug approval of ANY "intentionally altered genomic DNA" sequence in an animal. This altered DNA sequence trigger seems to be aimed squarely at breeder intention and human intervention in the DNA alteration.

The guidance states that "*intentionally altered genomic DNA may result from random or targeted DNA sequence changes including nucleotide insertions, substitutions, or deletions*"; however, it clarifies that selective breeding, including random mutagenesis followed by phenotypic selection, are not included as triggers. The new FDA Guidance contends that "*a specific DNA alteration is an article that meets the definition of a new animal drug at each site in the genome where the alteration (insertion, substitution or deletion) occurs. The specific alteration sequence and the site at which the alteration is located can affect both the health of the animals in the lineage and the level and control of expression of the altered sequence, which influences its effectiveness in that lineage. Therefore, in general, each specific genomic alteration is considered to be a separate new animal drug subject to new animal drug approval requirements.*" So every SNP is potentially a new drug, if associated with an intended alteration.

To put this in perspective, as was mentioned earlier, whole-genome sequence data from 234 taurine cattle representing 3 breeds revealed > 28 million variants comprising insertions, deletions and single nucleotide variants (Daetwyler et al., 2014). A small fraction of these mutations have been selected owing to their beneficial effects on phenotypes of agronomic importance. None of them is known to produce ill effects on the consumers of milk and beef products, and few impact the well-being of the animals themselves. In other words, there are a lot of SNP variations when comparing two healthy animals.

What is not clear is how developers are meant to determine which alterations are due to their "intention" and which result from spontaneous *de novo* mutations that occur in every generation. Certainly breeders can sequence to confirm the intended alteration especially if they are inserting a novel DNA sequence, but how can they determine which of the random nucleotide insertions, substitutions, or deletions are part of the regulatory evaluation, and which are exempt as they occurred spontaneously due to random mutagenesis. And if there is risk involved with the latter, why are only the random mutations associated with intentional modifications subject to regulatory evaluation? And what if the intended modification is a single base pair deletion (meaning the regulatory trigger would be the absence of a SNP) – something that is not there?

Many proposed gene editing applications will result in animals carrying desirable alleles or sequences that originated in other breeds or individuals from within that species (e.g. hornless Holsteins were edited to carry the Celtic polled allele found in breeds like Angus). As such, there will be no novel combination of genetic material or phenotype. The genetic material will also not be altered in a way that could not be achieved by mating or techniques used in traditional breeding and selection. It will just be done with improved precision and minus the linkage drag of conventional introgression.

It does not make scientific sense to regulate hornless dairy calves differently to hornless beef calves carrying the exact same allele at the polled locus (Carroll et al., 2016). Nor does it make sense to base regulations on human intent rather than product risk. Regulatory processes should be proportional to risk and consistent across products that have equivalent levels of risk.

There is a need to ensure that the extent of regulatory oversight is proportional to the unique risks, if any, associated with the novel phenotypes, and weighed against the resultant benefits. This question is of course important from the point of view of technology development, innovation and international trade, as well as the ability of the animal breeding community to use genome editing.

Currently there is only a single genetically engineered animal containing a heritable rDNA construct approved for food consumption anywhere in the world. In December 2015 the United States Food and Drug Administration (FDA) approved the AquAdvantage salmon for human consumption, although it is still not commercially available in the United States until the FDA publishes labeling guidelines for the fish. In 2016 Health Canada gave approval for the AquAdvantage salmon to be produced, sold and consumed in Canada. Animal breeders are therefore painfully aware of the chilling impact that regulatory gridlock can have on the deployment of potentially valuable breeding techniques. While regulation to ensure the safety of

new technologies is necessary, in a world facing burgeoning animal protein demands, overregulation of safe breeding methods is an indulgence that global food security can ill afford.

Conclusion

A plethora of technologies are currently at hand, with more to come. Our charge to the industry is to effectively make use of them towards improved animal populations. Animal breeding programs should position themselves to capitalize on a combination of advanced biotechnologies such as genomic information and advanced reproductive technologies to accelerate the rate of genetic gain. Ultimately these biotechnologies complement the genetic improvement that can be accomplished using traditional selection techniques and, if judged acceptable, offer an opportunity to synergistically accelerate beef cattle genetic improvement. Perhaps the bigger challenge is to improve the understanding and utilization of genetic selection tools both among those making selection decisions in the beef cattle industry, and in those groups seeking to influence public opinion. Many animal breeding goals have the potential to address sustainability challenges including improved animal well-being, efficiency and reduced environmental footprint. Something we would argue aligns with the shared, common values of a large segment of both cattle producers and the consuming public.

Acknowledgements

This work was supported by Biotechnology Risk Assessment Grant Program Competitive Grant no. 2015-33522-24106 and Agriculture and Food Research Initiative Competitive Grant no. 2015-67015-23316 from the National Institute of Food and Agriculture/U.S. Department of Agriculture.

References

Bosch, P., Forcato, D.O., Alustiza, F.E., Alessio, A.P., Fili, A.E., Olmos Nicotra, M.F., Liaudat, A.C., Rodriguez, N., Talluri, T.R., Kues, W.A. 2015. Exogenous enzymes upgrade transgenesis and genetic engineering of farm animals. Cell. Mol. Life Sci. 72: 1907-1929. doi: 10.1007/s00018-015-1842-1

Carlson, D. F., Lancto, C.A., Zang, B., Kim, E.S., Walton, M., Oldeschulte, D., Seabury, C., Sonstegard, T.S., Fahrenkrug, S.C. 2016. Production of hornless dairy cattle from genome-edited cell lines. Nat. Biotechnol. 34: 479-481. doi: 10.1038/nbt.3560

Carroll, D., A. L. Van Eenennaam, J. F. Taylor, J. Seger, and D. F. Voytas. 2016. Regulate genome-edited products, not genome editing itself. Nat. Biotechnol. 34: 477-479. doi: 10.1038/nbt.3566

Daetwyler, H. D., Capitan, A., Pausch, H., Stothard, P., van Binsbergen, R., Brøndum, R.F., Liao, X., Djari, A., Rodriguez, S.C., Grohs, C., Esquerré, D., Bouchez, O., Rossignol, M-N., Klopp, C., Rocha, D., Fritz, S., Eggen, A., Bowman, P.J., Coote, D., Chamberlain, A.J., Anderson, C., Van Tassell, C.P., Hulsegge, I., Goddard, M.E., Guldbrandtsen, B. 2014. Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat. Genet. 46: 858-865. doi: 10.1038/ng.3034

Food and Drug Administration. 2017. Draft Guidance for Industry #187 Regulation of Intentionally Altered Genomic DNA in Animals.

https://www.fda.gov/downloads/AnimalVeterinary/GuidanceComplianceEnforcement/GuidanceforIndustry/UCM113903.pdf (Accessed 1 May 2017.)

Jenko, J., Gorjanc, G., Cleveland, M.A., Varshney, R.K., Whitelaw, C.B.A., Woolliams, J.A., Hickey, J.M. 2015. Potential of promotion of alleles by genome editing to improve quantitative traits in livestock breeding programs. Genet. Sel. Evol. 47: 1-14. doi: 10.1186/s12711-015-0135-3

Kachman, S.D., M. L. Spangler, G. L. Bennett, K. J. Hanford, L. A. Kuehn, E. J. Pollak, W. M. Snelling, R. M. Thallman, M. Saatchi, and D. J. Garrick. 2013. Comparison of within and across breed trained molecular breeding values in seven breeds of beef cattle. Genet. Sel. Evol. 45:30. doi: 10.1186/1297-9686-45-30

Kachman, S. 2008. Incorporation of marker scores into national cattle evaluations. Proc. 9th Genetic Prediction Workshop, Kansas City, MO, p. 88-91.

Kadarmideen, H.N., Mazzoni, G., Watanabe, Y.F., Strobech, L., Baruselli, P.S., Meirelles, F.V., Callesen, H., Hyttel, P., Ferraz, J.B.S., Nogueira, M.F.G. 2015. Genomic selection of in vitro produced and somatic cell nuclear transfer embryos for rapid genetic improvement in cattle production. Anim. Reprod. 12: 389-396.

Laible, G., J. Wei, and S. Wagner. 2015. Improving livestock for agriculture - technological progress from random transgenesis to precision genome editing heralds a new era. Biotechnol. J. 10: 109-120. doi: 10.1002/biot.201400193

Lillico, S. G., Proudfoot, C., King, T.J., Tan, W., Zhang, L., Mardjuki, R., Paschon, D.E., Rebar, E.J., Urnov, F.D., Mileham, A.J., McLaren, D.G., Whitelaw, C.B.A. 2016. Mammalian interspecies substitution of immune modulatory alleles by genome editing. Sci. Rep. 6: 21645. doi: 10.1038/srep21645

MacNeil, M. D., Nkrumah, J.D., Woodward, B.W., Northcutt, S.L. 2010. Genetic evaluation of Angus cattle for carcass marbling using ultrasound and genomic indicators. J. Anim. Sci. 88: 517. doi: 10.2527/jas.2009-2022

Murray, J. D., and E. A. Maga. 2016. Genetically engineered livestock for agriculture: a generation after the first transgenic animal research conference. Transgenic Res. 25: 321-327. doi: 10.1007/s11248-016-9927-7

Peng, J., Wang, Y., Jiang, J., Zhou, X., Song, L., Wang, L., Ding, C., Qin, J., Liu, L., Wang, W., Liu, J., Huang, X., Wei, H., Zhang, P. 2015. Production of Human Albumin in Pigs Through CRISPR/Cas9-Mediated Knockin of Human cDNA into Swine Albumin Locus in the Zygotes. Sci. Rep 5: 16705. doi: 10.1038/srep16705

Proudfoot, C., Calrson, D.F., Huddart, R., Long, C.R., Pryor, J.H., King, T.J., Lillico, S.G., Mileham, A.J., McLaren, D.G., Whitelaw, C.B., Fahrenkrug, S.C. 2015. Genome edited sheep and cattle. Transgenic Res. 24: 147-153. doi: 10.1007/s11248-014-9832-x

Sander, JD, Joung, J.K. 2014. CRISPR-Cas systems for editing, regulating and targeting genomes. Nat. Biotechnol. 32(4): 347-355. doi: 10.1038/nbt.2842

Schweer, K.R., S.D. Kachman, L.A Kuehn, and M.L. Spangler. 2016. Comparison of SNP and haplotype models for genome-wide association studies for feed efficiency traits in crossbred beef cattle. Proc. International Society of Animal Genetics, Salt Lake City, UT p 147.

Snelling, W.M., S.D. Kachman, G.L. Bennett, M.L. Spangler, L.A. Kuehn, and R.M. Thallman. 2017. Functional SNP associated with birth weight in independent populations identified with a permutation step added to GBLUP-GWAS. J. Anim. Sci. (Aaccepted).

Spangler, M.L., E.J. Pollak, G.L. Bennett, K.J. Hanford, S.D. Kachman, L.A. Kuehn, W.M.

Snelling, and R.M. Thallman. 2011. Enhancing knowledge and technology adoption in a misunderstood discipline: The weight trait project. J. Anim. Sci. 88: E Suppl. 2. (Abstr.)
Tan, W., C. Proudfoot, S. G. Lillico, and C. B. Whitelaw. 2016. Gene targeting, genome editing: from Dolly to editors. Transgenic Res. 25(3): 273-87. doi: 10.1007/s11248-016-9932-x
Van Eenennaam, A. L. 2017. Genetic modification of food animals. Curr. Opin. Biotechnol. 44: 27-34. doi: 10.1016/j.copbio.2016.10.007
Weaber, R.L., J.E. Beever, H.C. Freetly, D.J. Garrick S.L. Hansen, K.A. Johnson, M.S. Kerley, D.D. Loy, E. Marques, H.L. Neibergs, E.J. Pollak, R.D. Schnabel, C.M. Seabury, D.W. Shike, M.L. Spangler and J.F. Taylor. 2014. Analysis of US Cow-Calf Producer Survey Data to Assess Knowledge, Awareness and Attitudes Related to Genetic Improvement of Feed Efficiency. In Proc. 10th World Congress on Genetics Applied to Livestock Production. (Abstr.)
Whitworth, K. M., Rowland, R.R., Ewen, C.L., Trible, B.R., Kerrigan, M.A., Cino-Ozuna, A.G., Samuel, M.S., Lightner, J.E., McLaren, D.G., Mileham, A.J., Wells, K.D., Prather, R.S. 2016. Gene-edited pigs are protected from porcine reproductive and respiratory syndrome virus. Nat. Biotechnol. 34: 20-22. doi: 10.1038/nbt.3434