Promotion of Alleles by Genome Engineering

John B. Cole

Address: Animal Genomics and Improvement Laboratory, Henry A. Wallace Beltsville Agricultural Research Center, Agricultural Research Service, United States Department of Agriculture, 10300 Baltimore Avenue, Beltsville, Maryland 20705-2350 USA

Correspondence: John B. Cole. Telephone: +1-301-504-8665. E-mail: john.cole@ars.usda.gov.

Abstract: Traditional genetic selection programs based on pedigree and performance information have been used to improve livestock populations for decades. The introduction of high-density single nucleotide polymorphism genotyping about 10 years ago supported increased rates of gain through more accurate prediction of genetic merit earlier in life. Recent continued technological advances enable the routine use of genetic engineering and gene editing tools in livestock research and, increasingly, production systems. Livestock geneticists have responded by proposing new breeding schemes that combine traditional selection methodology with these new tools to substantially increase rates of genetic gain while reducing harmful effects due to decreased heterozygosity. Genetic improvement strategies based on gene drives have the potential further increase rates of gain but pose risks that may not be acceptable to the public. Intense debate about the use of these technologies in the animal food chain are being driven by regulatory agencies and consumer advocates, and it is not clear if genetically modified animals will be acceptable to consumers. This review focuses on the application of genetic engineering and genome engineering tools to livestock population improvement through the management of genetic load and the promotion of desirable alleles in the population associated with both monogenic and polygenic traits. Limitations of the current technology, such as limited knowledge of true causal variants, are discussed, as are regulatory and consumer acceptance issues.

Keywords: gene editing, genome engineering, quantitative traits, recessive disorders

Review Methodology: Existing recent reviews on the subject have served as a useful input. More recent literature has been searched using Google Scholar. Search terms used: gene edited cattle; gene editing in livestock; genetic engineering in livestock; transgenic cattle; transgenic livestock. Published and unpublished information from researchers in the field has been accessed through personal communication.

Purpose of this Review

The objective of this review is to discuss the use of modern genome engineering tools to improve the genetic merit of livestock populations for monogenic and polygenic traits.

Genetic Improvement of Livestock
Both traditional genetic selection, which is based on the performance of relatives, and genomic selection, which uses single nucleotide polymorphism genotypes to directly track DNA inherited from parents, have been used very successfully to improve the genetic merit of many traits in a number of different populations. The most successful application of these technologies may be in dairy cattle [1], but they also have been used to improve many other livestock (e.g., [2,3]) and plant [4] species.

Traditional genetic selection programs have focused on improving the average genetic merit of animals in the population each generation. For example, the number of cows in the US national dairy herd has decreased from 23.67 million in 1940 to 9.26 million in 2014, while average annual fat yield has increased from 216 kg to 489 kg over the same time period. Much of this improvement in productivity is due to selection, with genetic gain averaging almost 2.4 kg of fat per year between 1960 and 2015 (Figure 1), roughly half of the total improvement in fat yield observed over that time period.

**Figure 1.** Genetic trend for predicted transmitting ability (PTA) for fat yield of US Holstein bulls (solid red line) and cows (dashed blue line) between 1957 and 2015.


This is done by selecting parents that have high genetic merit for traits of interest to produce the next generation of animals. The breeder’s equation, shown below, describes how different aspects of traits under selection affect the rate of genetic gain in a population (e.g., [5]).
\[ \Delta G_{\text{year}} = \sqrt{\text{reliability}} \times \text{selection intensity} \times \sqrt{\text{genetic variance}} / \text{generation interval} \]

In this equation, “\(\Delta G_{\text{year}}\)” is the annual rate of genetic change in the population, “reliability” is a measure of the precision with which an individual’s genetic merit is estimated, “selection intensity” is a measure of how selectively the parents of the next generation are chosen, “genetic variance” is the proportion of variation among animals in the population that is attributable to genetic differences, and “generation interval” is the average age of parents when their offspring are born. The “reliability” and “selection intensity” terms are the easiest to manipulate in traditional breeding programs, subject to market constraints. Genetic engineering and genomic selection provide opportunities to increase rates of gain through the “genetic variance” and “generation interval” terms, as well. These distinctions are not absolute; for example, generation interval can be reduced in a traditional breeding program if breeders are willing to accept lower reliabilities.

**Genetic and Genome Engineering of Livestock**

Introgression of an allele into a population by traditional breeding is typically a very slow process. The use of genetic and genome engineering technologies can dramatically increase the rate of introgression of desirable alleles [6]. In the discussion that follows, “genetic engineering” refers to the insertion of exogenous DNA into an animal’s genome, while “genome engineering” refers to manipulation of an animal’s own DNA.

**Genetic Engineering Through Transgenesis**

The first transgenic livestock were created by microinjecting DNA coding for a metallothionein-I-human growth hormone fusion gene into embryos of pigs, rabbits, and sheep [7]. Since those first successes, a number of transgenic animals and fish have been developed (see [8] and [9] for recent reviews). The production of transgenic animals remains laborious and expensive even though the microinjection of zygotes has been replaced by somatic cell nuclear transfer [8]. Animals have been modified both for food purposes, such as increased health and longevity (e.g., [10]), and non-food purposes, such as the expression of desirable products in their milk (e.g., [11,12]).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Source</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>β- and κ-casein</td>
<td>Cattle</td>
<td>Enhancement of milk composition and processing efficiency</td>
<td>[13]</td>
</tr>
<tr>
<td>Lactoferrin</td>
<td>Human</td>
<td>Innate host (immune) defense</td>
<td>[14]</td>
</tr>
<tr>
<td>Lysostaphin</td>
<td><em>S. simulans</em></td>
<td>Resistance to <em>S. aureus</em> mastitis</td>
<td>[10]</td>
</tr>
<tr>
<td>Myostatin</td>
<td>Knockout</td>
<td>Increased muscle growth</td>
<td>[15]</td>
</tr>
</tbody>
</table>

**Table 1.** Some examples of cattle genetically engineered for agricultural and pharmaceutical production.
There are both technical [17] and ethical [18,19] challenges associated with the widespread use of transgenic animals in livestock production, although the use of animal bioreactors to produce biomedical materials remains appealing (e.g., [20,21]). In addition, it remains expensive to generate transgenic animals because efficiency is low and control over transgene integration remains poor. Given such technical and economic limitations, transgenic livestock are likely to be used only in limited, biomedical settings.

**Genome Engineering Through Precision Gene Editing**

A number of tools are now available for editing eukaryotic genomes, including clustered regularly interspaced short palindromic repeats (CRISPR; [22]), transcription activator-like effector nucleases (TALEN; [23]), and zinc finger nucleases (ZFN; [24]). Edits can include the deactivation (knock-out) of genes, the alteration of short segments of DNA, and the insertion of new genes. The latter application means that the line between traditional genetic engineering and genome engineering is somewhat blurry since both tools can be used to introduce DNA from one species into another. However, gene editing refers to a specific suite of tools for making targeted changes to DNA, while genome engineering broadly refers to a wide array of strategies for using gene editing in concert with breeding strategies or other technologies to make population-wide changes [25].

Current gene editing approaches are based on the use of site-directed nucleases to precisely introduce double stranded breaks (DSB) at predetermined locations in the genome (e.g., [26]). There are two distinct pathways for repairing DSB. The underlying principle is that host repair factors will congregate at the site of a DSB, where they will repair the DNA. In the absence of a guide sequence to provide a template, the broken double-stranded sequence will be repaired by the non-homologous end-joining pathway (NHEJ), a process known as targeted mutagenesis, which often introduces deletions or additions of a few base pairs and can result in gene knock-outs [27]. However, if a nucleic acid template (donor DNA) is provided, then the repair enzymes will use the donor DNA as a guide for precise repair by homologous recombination (HR). The HR pathway can be used to precisely add, delete, or replace letters in the genetic code at the location of the break [27]. The ZFN-, TALEN-, and CRISPR-based approaches are briefly described below, but interested readers should consult a more comprehensive source such as [26] for additional details. Van Eenennaam [28] provides a comprehensive description of gene editing combined with somatic cell nuclear transfer in livestock production systems (see Figure 2 of that paper).

**Zinc finger nucleases.** Artificial restriction enzymes can be constructed by joining zinc finger domains, which bind to 3 to 6 nucleotide triplets, to nucleases that cut DNA [26]. The limited length of these domains can result in off-target breaks. The Fok1 enzyme that is commonly used as the nuclease domain must be present as a dimer in order to induce double-stranded breaks, which means that multiple ZFN must be used to target non-palindromic sites.
Alleles can be knocked-out by using individual ZFN to target single base pairs, while multiple ZFN can be used to cut out a large piece of DNA. If a homologous (guide) template is provided then DNA can be inserted at the break site using homology-directed repair. There are many commercial providers of ZFN-based tools, making their use relatively simple and affordable. The greatest obstacle to the use of ZFN is the construction of sufficiently specific target domains. This tool has been used to produce gene-edited cattle [29] and pigs [30].

**Transcription activator-like effector nucleases.** More precise targeting of single base pairs can be achieved using transcription activator-like effector molecules ligated to nucleases to produce TALEN. Their principal advantage over ZFN is that the DNA-binding domains in TALE are composed of a series of 33- to 35-amino-acid repeats, resulting in greater specificity. The specific base targeted by the TALE is flanked by target sequence so that non-palindromic DNA may be targeted using a single TALEN, rather than multiple ZFN (e.g., [31]). This technology has been used to produce many gene-edited animals, including cattle [32], chickens [33], and pigs [30].

**Clustered regularly interspaced short palindromic repeats.** Clustered regularly interspaced short palindromic repeats. The construction of novel ZFN and TALEN can be time-consuming, labor-intensive, and costly. which depend on protein-DNA interactions, the CRISPR-Cas9 system is derived from RNA-based defense systems found in bacteria [34]. The Cas9 protein uses a sequence of guide RNA (gRNA) to identify DNA sequences of interest, such as genes. The gRNA binds to the complementary DNA sequence and a double-stranded break is made. Genes can then be knocked-out by NHEJ, or “faulty” genes can be repaired or new genes inserted using HR. However, an important limitation of CRISPR-Cas9 is that editing targets must be upstream of a protospacer adjacent motif (PAM), a 3 to 5 nucleotide motif that serves as a binding signal for the Cas9 protein. Alternative PAM have been identified in different species of bacteria (e.g., [35–37]), and Cas9 has been engineered to recognize a wider array of motifs [38,39]. CRISPR-Cas systems have several advantages over ZFN- and TALEN-based systems, including simplicity of designing target sequences (it is easy to design gRNA targets), efficiency (plasmids encoding Cas proteins and gRNA can be microinjected directly into embryos) [40], and the ability to easily multiplex edits [41]. This approach has been used to produce gene-edited cattle [42], goats [43], and pigs [44].

There is great interest in the use of gene editing to improve animal health (e.g., [45]) and several studies have produced genetically modified cattle that are able to resist common diseases (Table 1). Additionally, gene editing may be an effective tool for reducing the frequency of genetic disorders in livestock populations or eliminating those disorders altogether (e.g., [46–48]), and a recent series of simulation studies showed that gene editing also has the potential to improve rates of genetic gain for quantitative traits [49,50].

Table 2. Examples of gene-edited cattle (adapted from [28]).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Type</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta-lactoglobulin</td>
<td>Knockout</td>
<td>Elimination of milk allergen</td>
<td>[51,52]</td>
</tr>
</tbody>
</table>
Applications of Genetic and Genome Engineering in Livestock Breeding Programs

Traits of interest in livestock breeding programs are commonly grouped into two groups: those controlled by a single gene (monogenic traits) and those controlled by many genes acting together (polygenic traits). There are also cases, such as susceptibility to some diseases, in which a phenotype is influenced by many loci but a major gene also accounts for a large proportion of the genetic variance. Genetic and genome engineering are applied differently to monogenic versus polygenic traits, as described below. Some applications proposed in the following discussion are not yet possible given the current state of the technology, and may never be feasible. Those cases will be noted, but the discussion will focus on possible benefits in the long term. Technical limitations and other concerns will be discussed in the section titled “Future outlook: opportunities and obstacles”.

Monogenic Traits

Modern genetics began with the work of Mendel on simple traits in peas (*Pisum sativum*) [61], such as flower color and seed appearance. While many traits of economic importance in livestock are controlled by many genes acting together, many monogenic traits are important for animal health and welfare.

Polled. Not all recessive alleles of interest in livestock populations are lethal, with the polled locus (OMIA ID: 000483-9913) in cattle as a prominent example. The polled condition in cattle is due to the action of a dominant allele [62] with a very low frequency – less than 1% – in the US Holstein population, and there is interest in increasing its frequency to improve animal welfare. However, polled Holstein bulls have lower average genetic merit than horned bulls and there have historically been relatively few polled bulls available. Gene editing can be used to produce polled calves from high-genetic-merit bulls; two healthy, homozygous polled Holstein calves, Spotigy and Buri, were produced by TALEN-stimulated homology-dependent repair [32]. Several studies have shown that that the frequency of the polled allele increases much faster using genome engineering than traditional breeding alone [46,48,63].

Figure 2. Observed allele frequency of the Holstein horned locus in a simulated population for traditional breeding (“noedits”) and four different gene-editing technologies (“CRISPR”, “TALEN”, “ZFN”, and a hypothetical “perfect” method) over 20 years when the top 1% of bulls and no cows are edited (adapted from [46]). The methods differ in their rates of efficiency [30,64], which is why they produce different allele frequencies.

<table>
<thead>
<tr>
<th>Trait</th>
<th>Transgene/Method</th>
<th>Description</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lysostaphin</td>
<td>Resistance</td>
<td>S. aureus mastitis</td>
<td>[53]</td>
</tr>
<tr>
<td>Lysozyme Transgene</td>
<td>Resistance</td>
<td>S. aureus mastitis</td>
<td>[54]</td>
</tr>
<tr>
<td>Myostatin Knockout</td>
<td>Increased</td>
<td>muscle yield</td>
<td>[55,56]</td>
</tr>
<tr>
<td>NRAMP1 Insertion</td>
<td>Resistance</td>
<td>tuberculosis</td>
<td>[54]</td>
</tr>
<tr>
<td>POLLED Substitution</td>
<td>Animals born</td>
<td>without horns</td>
<td>[32,57]</td>
</tr>
<tr>
<td>PRNP Knockout</td>
<td>Prion resistance</td>
<td></td>
<td>[58,59]</td>
</tr>
<tr>
<td>SP110 Transgene</td>
<td>Resistance</td>
<td>tuberculosis</td>
<td>[60]</td>
</tr>
</tbody>
</table>
Slick. The slick hair gene (OMIA ID: 001372-9913), an autosomal dominant locus, confers increased thermotolerance through a short, thin hair coat. The slick gene has been introgressed from the Senepol breed, where it was first observed, into US Holsteins [65]. A series of convergent mutations in the prolactin and prolactin receptor genes [66,67] are responsible for the phenotype. The slick gene is an attractive target for genome engineering, because an intra-species allele transfer can be used to produce clones of high genetic merit bulls that also are slick. Daughters of such edited bulls would have lower rectal temperatures and respiration rates, higher sweating rates, and smaller decreases in milk yield under heat stress conditions than daughters of wild type bulls [65]. Application of gene editing to slick provides clear advantages to cows, consumers, and farmers.

Management of genetic load. Recessive disorders have been identified in livestock populations since animal breeding programs began, and more than 500 such disorders have been catalogued in cattle alone [68]. In the past, test matings were used to identify carriers of recessive disorders as part of progeny-test programs [69], but most recessive mutations were identified after a carrier bull sired many daughters and had sons used for artificial insemination (e.g., bovine leukocyte adhesion deficiency [70], complex vertebral malformation [71], and deficiency of uridine monophosphate synthase [72]). Newly arisen or previously unknown
recessives also can spread quickly through a population through the use of popular bulls before routine screening is possible because such defects are not directly observable (e.g., CWC15 in the Jersey breed [73]).

Genomic tools have enabled the detection of new recessive alleles that have harmful effects on fertility [74–77], many of which act early in gestation and could not previously be distinguished from failed matings. MacArthur et al. [78] estimated that human genomes contain approximately 100 loss-of-function mutations, and about 20 completely inactivated genes. It seems reasonable to assume that other mammals with genomes of similar sizes may carry a similar genetic load. It is important to remember, however, that human populations are not subject to culling for fitness as livestock are. The apparent increase in the rate of discovery of genetic defects does not mean that mutations are increasing in frequency, it simply reflects our improved ability to detect such changes in the genome.

CRISPR-Cas9-based gene editing already has been used in human embryos to repair a mutation in MYBPC3, which causes hypertrophic cardiomyopathy [79], and similar approaches could be used to repair genetic defects in bulls with high genetic merit. Cole [46] has proposed that gene editing can be used to eliminate deleterious recessive alleles from dairy cattle populations, and showed that the rate of embryonic loss (death) decreases rapidly when the top 1% of bulls in the population are edited. Johnson et al. [47] used simulation to determine how best to use gene editing to manage genetic load, concluding that selection against carriers may be the best approach to managing recessive alleles in the short term, but that gene editing may be more effective in the long-term. However, Van Eenennaam and Kinghorn [80] found that selection against carriers may resulted in reduced rates of genetic gain. Another embryo-based strategy is to edit all IVF embryos and then genotype embryo biopsies to both confirm the success of the edit and identify the embryo with the highest genetic merit [81]. However, it may not make economic sense to edit individual embryos when IVF and embryo genotyping can be used to identify another embryo from the same flush with similar genetic merit that is free of known genetic defects. As the number of known genetic defects increases, it may not be possible to identify defect-free embryos from a flush, and gene editing may be necessary in order to produce animals free of known defects.

**Validation of causal variants.** Gene editing can be used to quickly and affordably validate causal variants associated with certain traits, such as early embryonic loss. For example, a loss-of-function mutation in the CWC15 gene is thought to be causal for embryonic losses associated with the JH1 haplotype in the Jersey breed [73]. The biological mechanism has not been elucidated and the association has not been validated in vitro, although there are time-dependent differences in gene expression for CWC15 [82]. CRISPR-based knockouts have been used to successfully identify embryonic lethals in other species (e.g., [83–85]).

**Polygenic Traits**

The infinitesimal model, first proposed by Fisher [86], is the foundation of quantitative genetics [87] and provides a mathematical framework based on the idea that phenotypes are influenced
by a large number of loci, each with a small effect. Cole et al. [88] concluded that Fisher's model holds for many traits routinely evaluated in dairy cattle populations. Genomic selection is a special case of marker-assisted selection in which a dense marker set is distributed across the genome [89,90], and the SNP used for prediction may be in linkage with the true causal variants, rather than causal variants themselves. The use of advanced reproductive technologies in conjunction with marker-assisted selection was proposed by Georges and Massey [91] as a way to substantially increase rates of genetic gain, and a cost-effective implementation of a similar scheme recently has been demonstrated [92].

Promotion of alleles by gene editing. In a recent paper, Jenko et al. [49] described a system that they call promotion of alleles by genome editing (PAGE), in which gene editing is used to edit hundreds of alleles to fix favorable alleles, thereby increasing the rate of genetic gain and the asymptote of genetic gain. Inbreeding increases as the number of edits increases, and as the number of animals edited decreases, although the authors suggest that this is not a case for concern because favorable alleles are being fixed rather than deleterious alleles. In the most ambitious scenario considered – 100 QTN edited in each of the top 5 sires – cumulative genetic gains were more than 4 times greater than in the baseline scenario of genomic selection with no editing. A more modest doubling was observed when the 20 loci with the largest effects were edited in all 25 bulls in the population each generation, a much more plausible scenario than editing 100 loci in a single animal. It should be noted that the number of unique loci edited was smaller than the total number edited because the same QTN could be present in many animals. Performing more edits on a smaller number of sires resulted in greater response to selection than performing fewer edits on more bulls when editing resources were limited. However, this also results in substantially greater inbreeding. The authors note that the potential of PAGE to dramatically increase rates of genetic gain is based on the fact that it permits the selection of favorable alleles independently of the haplotypes, chromosomes, and individuals that carry them. Ultimately, PAGE overcomes the barrier represented by the limited number of recombination events during meiosis in gametogenesis.

In an independent study of PAGE, Simianer et al. [93] concluded from their own study of PAGE that realistic gains are almost certain to be much smaller than those predicted by Jenko et al. [49], on the order of 11.6% more than GBLUP, rather than ~400%. Perhaps the most important difference between the two studies is that Jenko et al. [49] assumed that true QTN were known and could be edited directly, while Simianer et al. [93] used ridge regression to estimate SNP effects and edited the markers with the largest effects. In that case, the 20 loci with the largest effect were edited within each sire. Given that true QTN are unknown in our populations, and we are unlikely to identify them without some error, it seems plausible that the initial study of PAGE over-estimated rates of gain that can be expected in practice. These results underscore the importance of assumptions when comparing simulation studies. However, that simply means that PAGE will produce lower rates of gain than originally predicted, not that the approach is without value. The challenge for practitioners will be to balance predicted benefits against the cost of implementation. It could be the case that alternative breeding schemes, such as those using , will be more cost-effective in practice.
Accepting, for the sake of argument, that low-cost gene editing of many alleles in single individual is possible, the loss of genetic variance in the population is of great concern. If the objective of selection is to fix all favorable variants, gene editing can help us do that quickly, and precision edits could be made for the purpose of maintaining variation in the population. This is potentially an important advantage over classical genetic selection that operates on haplotypes, the lengths of which are limited by the rate of recombination. However, if the problem of identifying causal variants is formidable (as discussed below in the section titled “Challenges to the adoption of PAGE”), then the challenge of identifying what variants in the population should be maintained because of potential future is more formidable still. Genotype-by-environment effects also may favor some variants in one place or under one management system, but different variants in others. Even if the edited regions of the genome are free of undesirable alleles there is no guarantee that the unmodified parts of the genome will be, and the intense within-family selection driven by genomic selection will only increase. This would result in a loss of variation across the genome, not only in edited regions. Ultimately, dairy breeding programs might then come to resemble swine or poultry programs, in which crosses are made among inbred lines. It also would ultimately favor large breeding companies with the financial resources necessary to produce gene edited animals over individual registered breeders and small genetics firms that currently can compete in the marketplace.

**Maintenance of genetic variation.** While early research suggested that inbreeding in dairy cattle would decrease under genomic selection [94], it appears to have increased in practice [95,96]. There is a substantial body of literature on the deleterious effects of inbreeding on livestock performance (e.g., [97,98]), and many schemes have been proposed to control levels of inbreeding in both traditional [99] and genomic [100] breeding programs. Recessive and partially recessive deleterious alleles appear to be the principal drivers of inbreeding depression [101], suggesting that genome engineering may be a useful tool for reducing genetic load through targeted editing of known deleterious alleles. Immune gene regions also could be targeted for editing to ensure that low effective population sizes from intensive selection are not accompanied by loss of allelic diversity. Thompson-Crispi et al. [102] found that cell- and antibody-mediated immune responses are under genetic control, and identified several genomic regions [103] that are likely targets for genome engineering. It seems likely that increased homozygosity has harmful effects in some parts of the genome, and neutral or beneficial effects in others. One option for making PAGE schemes more appealing is to include edits for the purpose of increasing heterozygosity in those regions where it is known to be beneficial.

**Gene stacking.** Gene stacking [104] is the process by which two or more transgenes are accumulated in a single individual, often through intercrossing of transfected lines. While several approaches can be used for this purpose, all are time-consuming and labor-intensive. In plants, where the technique is most commonly used, the accumulation of three or four transgenes in a single line is considered very successful. New tools may increase the efficiency of gene stacking in some species of plants [105], but similar resources are not yet available for use with mammalian genomes. Fischer et al. [106] have recently used gene stacking in
conjunction with gene editing and bacterial- and phage artificial chromosome vectors to produce pigs for xenotransplantation whose genomes contain several transgenes, as well as knockouts. In concept, the use of gene editing to simultaneously modify several loci is simpler than gene stacking, but many technical challenges related to off-target insertions and unintended mutations remain [107,108].

Cole and VanRaden [109] described a conceptually similar approach that could be used to accumulate in one (e.g., Bos taurus) individual the 29 autosomes in the population having the highest genetic merit. When haplotypes are sampled at random during gametogenesis there are $2^{29}$ possible combinations of chromosomes, and there are many more when recombination is considered. Given than haplotypes segregate independently, there is no way to produce animals with a specific set of haplotypes short of crossing completely inbred lines. This challenge, on a smaller scale, also affects gene stacking programs, particularly when genes to be stacked are on different chromosomes and assort independently.

**Gene drive.** First proposed by Burt [110] as a means of controlling natural populations, such as mosquitoes which transmit malaria, gene drives make use of DNA repair mechanisms to ensure that mutations on one chromosome are copied onto the homologous chromosome. Gonen et al. [111] have suggested that gene drives could be incorporated into livestock breeding programs that include gene editing. The advantage of such approach is that edited alleles would reach fixation more quickly, ensuring homozygosity of the favorable allele in all descendants of the edited individual regardless of the genotype of the other parent. Their simulation found that gene editing produced 1.95 times as much gain as selection alone, gene drives achieved 1.43 times more gain that gene editing alone, and gene drives produced 2.8 times as much gain as selection alone. Allele frequencies reached fixation much faster with either genomic technology than selection alone, although gains from adding gene drives to gene editing were small. The rate of increase in population average inbreeding, already concerning under selection, increased dramatically when gene drives were used only in a small number of top bulls. The broader the portfolio of bulls used, the lower the impact on inbreeding. These results show that gene drives have the potential to amplify the benefits of genome editing in livestock breeding by reducing the time needed to fix favorable alleles. Cumulative genetic gain also is increased because, once the first group of edited alleles is fixed, the resources needed to edit the first group can be allocated to the alleles with the next-smallest effects.

There is considerable concern about the risks associated with using gene drives (e.g., [112]), and regulators may prohibit its use in livestock in the absence of a more compelling argument than increased rates of genetic gain. However, Zentner et al. [113] note that a number of biological mechanisms, such as inbreeding and naturally occurring variation, can interfere with the effectiveness of gene drives, which may limit the appeal and effectiveness of those tools in livestock populations.

**Future outlook: opportunities and obstacles**
Challenges to the adoption of PAGE

What loci should be edited? A critical issue not addressed by studies to date is that of which alleles should be edited. While this is understandable given that the purpose of the initial work was to demonstrate the utility of the PAGE concept, it must be addressed at the time of implementation. Weller et al. [114] used an *a posteriori* granddaughter design that included 52 grandsire families with 9,178 sons to identify QTL in the US Holstein population, identifying 30 significant QTL. In a follow-up study with 71 grandsire families with 14,246 sons, Wiggans and Weller [115] identified 56 QTL, 29 of which were confirmed from their earlier report. Recently, they used data from the 1000 Bull Genomes Project and other sources to determine concordance between QTL identified from SNP data and quantitative trait nucleotides (QTN) [116]. They found complete or almost-complete concordance only for stature on chromosome 14 and daughter pregnancy rate on chromosome 18. These results underscore the difficulty of identifying the hundreds of editing targets needed for the most effective PAGE strategies. The human genetics community faces similar challenges and has developed a set of guidelines to guide research into causal variants [117] that could be adapted to the needs of livestock research. Given enough time, international efforts such as the Functional Annotation of Animal Genomes project [118] may identify true variants that can be used as targets for editing, but the community is relatively small and constrained by financial limitations, limiting its ability to rapidly make progress.

It is also likely that the problem of identifying loci for editing is of greater magnitude than commonly assumed. Visscher et al. [119] concluded that, “it is the cumulative effect of many loci that underlies susceptibility to disease”, which they reiterated in a subsequent review [120]. Gianola et al. [121] also noted, based on a simulation study, that models commonly used for GWAS do not properly account for linkage disequilibrium, resulting in spurious or misleading results. In a simulation study, Jenko et al. [122] examined the power of using a population of 1 million sequenced animals to identify causal variants, but found that only a small proportion of true variants (2.5–4.8%) were discovered. These results are broadly in agreement with those of VanRaden et al. [123], who found that the addition of sequence variants to SNP used for genomic prediction provided only small gains in reliability because nearby markers already account for the effect of causal variants. Hickey et al. [50] proposed a scheme for testing alleles to differentiate between causal and non-causal alleles based on identification of candidate causal alleles using large-scale GWAS, followed by editing of desirable causal alleles into sire lines for testing in progeny. This allele testing strategy could be used in concert with the allele specific expression approach described by Khansefid et al. [124] to precisely identify true causal variants, which could then be rapidly introgressed into populations using accelerated breeding schemes (e.g., [92,125]).

These results underscore a critical point: gene editing is unlikely to result in substantial improvement of complex traits because 1) we lack the knowledge of direct effects of, and interactions among, individual loci needed to identify targets for editing, 2) our widely used statistical models may not be sufficient for identifying true causal variants, and 3) it seems implausible to assume that simultaneous, side-effect-free editing of hundreds of loci will ever
be feasible. However, gene editing could be useful for the improvement of monogenic traits, either through correction of genetic defects or promotion of desirable alleles, such as polled.

**Limits to editing technology.** Several technical challenges with gene editing technology must be overcome before PAGE can be implemented. The most notable of these is the development of low-cost tools for multiplexing edits; Hickey et al. [50] argue that this may be the easiest challenge to address, but the difference between 20 and 250 edits is substantial, and there is reason for considerable skepticism that this is an easily solved problem. Recent studies have identified off-target insertions and unintended mutations [107,108] associated with CRISPR editing, raising the possibility that the number of possible simultaneous edits will remain low to avoid the accumulation of uncontrolled changes. In such a case, the effectiveness of PAGE-based strategies will more closely resemble gene stacking (discussed below) or more traditional marker-assisted breeding schemes (e.g., [126]).

**Selection limits.** Cole and VanRaden [109] have argued that there is no evidence yet that dairy cattle are nearing selection limits, but such limits surely exist. In addition, non-genetic factors such as animal health and feed intake impose limits on phenotypic responses to genetic selection. High-performing animals currently face many challenges, most notably in the transition from pregnancy to lactation, and there will be a point at which it makes more sense to maintain a somewhat larger herd of animals with slightly lower genetic potential than to continue selecting for, e.g., greater fat and protein production.

**Unanticipated uses of technology.** The success of animal breeding programs depends on good-faith participation by many individuals (e.g., [98]). Technologies such as gene editing are inherently value-neutral and can be used for the mutual benefit of all, as well as for the specific benefit of only one party. For example, if lost-cost, multiplex gene editing is eventually realized then individuals could be edited so that their genome contains the most favorable alleles for each of the markers in the SNP panels used for genomic evaluation, regardless of the animal’s true genetic background. This requires that population-specific SNP effects are known, but approximate values can be back-calculated using publicly available genomic breeding values if an individual has access to a large enough library of genotypes. The resulting animals would receive very high genomic evaluations and would have high apparent value in the marketplace but would not provide the level of genetic improvement suggested by their genomic evaluations. However, if there is no “signature” of the editing process that unambiguously prove that such a genotype resulted from human intervention rather than chance then confidence in the system is undermined to benefit one at the expense of the community. The solution to this problem is probably to develop systems that incentivize desirable behavior, but we should not be blind to the existence of bad actors.

**Regulatory considerations**

While some genetically modified and gene-edited products recently have reached the U.S. marketplace [128,129], uncertainty about the manner in which gene-edited plant and animal products will be regulated remains a substantial concern [130]. The AquaAdvantage salmon,
genetically engineered for rapid growth, was finally approved for sale following a twenty-year
review by federal regulatory agencies [128]. Much of the discussion at the 2016 Large Animal
Genetic Engineering Summit focused on the use of gene editing to produce large animal models
of human disease (e.g., [131]) rather than modified food animals, possibly in response to an
ongoing climate of regulatory uncertainty, although there was more discussion of gene-edited
animals at the 2018 conference. It also is unclear if consumers will readily accept the
widespread introduction of gene-edited animals in the food chain. Policymakers and regulators
are being encouraged to exercise oversight based on the product rather than the process used
to generate that product [132], but the Court of Justice of the European Union implicitly
rejected this approach when ruling recently that gene-edited crops are subject to the same
regulations as conventional genetically modified organisms [133].

**Consumer acceptance of gene editing**

Many challenges are associated with both genetically engineered and gene-edited animals,
some technical and others related to consumer attitudes towards the technology [134,135].
While the tools available for making changes to animals’ genomes have increased in capability
in recent years, the general public remains concerned about changes made to the genomes of
food crops and livestock. The term “genetically modified organism” often is used in discussions
of consumer and regulatory affairs, language which unfortunately conflates very different
technologies. That term will be used in the following discussion for consistency with the
literature discussed, and should be understood to refer to a broad array of technologies that
includes both genetic engineering and gene editing.

A recent meta-analysis of the literature on consumer preferences suggests that U.S.
respondents have a more favorable view of biotechnologically modified food products than
those from Europe, but most consumers are concerned about genetically modified animals
[136]. Consumers that are generally opposed to the marketing of genetically modified
organisms may moderate those opinions in the presence of another benefit (e.g., increased
levels of omega-3 fatty acids in farmed salmon) [137]. Changing consumer attitudes towards
technologies may be possible, but the discussion should focus on the benefits rather than the
technology [138]. It is difficult to predict how consumers will respond to the idea of dozens or
hundreds of simultaneous edits being made to an individual’s genome, particularly when
current knowledge of interactions among loci is very limited.

Consumers may be more accepting of gene editing in food animals if the technology
focus is on animal health and welfare rather than on productivity [139], and there is less
objection to the promotion of naturally occurring genetic variants [140,141]. For example, the
process of dehorning is traumatic to calves, unpleasant for farmers, and distasteful to
consumers (e.g., [142]). Previous studies [143,144] have shown that increasing the frequency of
polled animals in the Holstein population is difficult because the frequency of the dominant
allele is very low. Carlson et al. [45] have successfully produced polled clones of horned animals
using gene editing with no detectable off-target effects, which showed that the technology can
be used to rapidly propagate desirable genotypes. Gene editing also has been used to produce
animals with increased resistance to disease [145], including porcine reproductive and
respiratory syndrome [146,147] and bovine tuberculosis [54]. Other candidates for gene editing
include casein variants that may have beneficial effects on human health [148], the slick locus
that is involved in adaptation to hot environments [65], and the DGAT1 gene which has
favorable effects on milk composition [149].

Conclusion/Summary

The rapid development of tools for the precision editing of livestock genomes provides an
exciting view of a future in which selection objectives can be rapidly achieved using a
combination of advanced reproductive technologies, genomic selection, and genome
engineering with low risk of accumulation of harmful genetic defects in the population.
However, this future depends on a large body of knowledge that has not yet been generated.
We will probably never learn the exact function of every gene in the bovine genome, or the
precise genetic mechanism that underlies every genetic defect, not because the problem is
insoluble but because it requires human resources and financial capital that are not available to
us. The ultimate goal of this work is to use new technology to feed a growing population with
fewer inputs, which will depend on gaining the consent of consumers who include animal
protein in their diets. It is not enough to assert the safety of these tools, it must be proven with
rigorous studies that are openly discussed if regulatory agencies are to be satisfied that the
animals produced using these tools are safe for human consumption. As the population in the
global south increases there may be a divide in adoption of the technology, with genome
engineered food common in some parts of the world and prohibited in others.

Acknowledgements

This work was supported by appropriated project 1265-31000-096-00, "Improving Genetic
Predictions in Dairy Animals Using Phenotypic and Genomic Information", of the Agricultural
Research Service of the United States Department of Agriculture. Mention of trade names or
commercial products in this article is solely for the purpose of providing specific information
and does not imply recommendation or endorsement by the US Department of Agriculture. The
USDA is an equal opportunity provider and employer.

Maci L. Mueller (University of California - Davis) offered helpful feedback on the
manuscript, and provided an outline for the description of gene editing methodology in the
section titled “Genome Engineering Through Precision Gene Editing”. Two anonymous
reviewers provided thoughtful feedback that substantially increased the quality of the
manuscript.

2. Jonas E, de Koning D-J. Genomic selection needs to be carefully assessed to meet specific requirements in livestock breeding programs. Front Genet. 2015;6. doi:10.3389/fgene.2015.00049


87. Falconer DS, MacKay FC. Introduction to quantitative genetics. Longman; 1996.


128. Ledford H. Salmon approval heralds rethink of transgenic animals. Nat News. 2015;527:417. doi:10.1038/527417a


