Proc. Assoc. Advmt. Anim. Breed. Genet. 22:191-196

CUTTING AND PASTING: THE FUTURE OF GENETIC IMPROVEMENT FOR FOOD ANIMAL GENOMES

T.S. Sonstegard[1], D. Carlson[1], P.B. Hackett[2], A. Watson[1] and S.C. Fahrenkrug[1]

[1]Recombinetics, Inc. St. Paul, MN, USA
[2]College of Biological Sciences, University of Minnesota, Minneapolis, MN, USA

SUMMARY

Gene editing technologies based on site-directed nucleases continue to improve at a rapid pace and have evolved to a point where they can be useful for direct introgression of high effect alleles into naïve populations of food animals. Herein, we review basic mechanics of site-directed nuclease action and how this activity is deployed to produce precision bred alleles into animal genomes. We also discuss the variety of traits being deployed, and differences between introducing alleles already found in nature versus gene knock out and other rationale design approaches for genetic improvement. Finally, perspectives for regulatory approval and commercialization are summarized to highlight some of the obstacles, which may hinder the widespread adoption of gene editing technology as one of the primary tools of animal breeding.

BACKGROUND

Since the first livestock domestication events approximately 10,000 years ago, the efficiency of animal production in the developed world has continued to improve through selection for desirable traits related to protein yield. The long tradition of selective breeding has relied on superior production traits emerging from the natural genetic flux. Many other advances in animal husbandry, like advanced reproductive techniques (ART), new animal medicines, and feed additives have supported production increases derived from selective breeding outcomes. Although modern genetic techniques like genome selection are increasing the accuracy with which we can find and select for these valuable alleles, genetic improvement is still limited by the availability and frequency of beneficial alleles in our current breed populations and slowed by linkage disequilibrium (LD) and long generation intervals. Furthermore, antagonistic effects are problematic in breeding practice due to tight linkages of alleles with opposite (pleiotropic) effects, e.g., the antagonistic effects of dairy fertility and/or disease resistance with milk production.

The importance of precision in animal breeding is further underscored by challenges related to global food security (FAO 2017). A burgeoning middle class of consumers estimated to be growing from 1.8 to 4.9 billion by 2030 will significantly contribute to the increasing human demand for animal protein. Furthermore, the geographical regions where this growth in consumption is highest underscore the inefficiencies of local adapted livestock varieties. These unrefined varieties are not capable of sustainably meeting demands in rapidly expanding markets, unless rapid improvements are made in average production output per animal. Crossbreeding provides an alternative to rapidly improve production in these adapted varieties, but this strategy has historically provided only short term bursts of increased performance due to heterosis and the introgression of beneficial alleles. The long term downside of such admixture is that locally adapted or purpose-bred genetics can be diluted, requiring additional, lengthy backcrossing to reach the breeding objectives of a more productive, adapted animal. Therefore, new technologies that augment current selection methods for genetic improvement must be used as part of the solution across production systems. There are thousands of yet to be discovered, desirable traits in animals that allow them to survive well in their current environments. Ultimately, solutions based on animal breeding are hampered by generation interval times, and economic feasibility and practicality for low input production systems.

The opportunity provided introgression of alleles initiated by gene editing. Advanced breeding techniques based on genome editing offer an alternative method for rapid acceleration of genetic improvement in a single generation that is directed and sustainable. The results can be disseminated in subsequent generations through traditional breeding methods. This technology allows for adaptive breeding of elite production lines, rapid production improvement of adapted lines, and conservation of diversity by introduction of heirloom alleles lost through intensive selection for a single production trait. Thus, precision genome editing as an animal breeding tools has the potential to be a "Game Changer". The possible applications and precision of site-directed nucleases (SDN)s could not, and were not predicted or made possible until very recently. By breaking a DNA molecule at a specific site, we can induce a designed genetic change by instructing the cell's repair mechanisms. The various molecular scissors provide the power previously used by plant breeders to introduce new traits by double stranded DNA breaks (DSB) made through treatment with radiation or mutagenic compounds. SDNs, like TAL effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9, have improved upon the accuracy of gene targeting by 10,000,000,000-fold, so there is no need to wade through all the random undesirable mutant outcomes of random mutational breeding. Therefore, the precision of genome editing presents a unique, asymmetric opportunity, with negligible risk and the potential for relatively very high genetic gain.

The process of editing and its deployment. Foreign DNA is often touted by the anti-biotech advocates as a primary reason for safety concerns, even though the average human consumes 10^{14} unknown genes per day. From a mechanistic perspective, the molecular components used to trigger the editing process (post-DSB) do not actually introduce or transfer any foreign material into target genome (Jasin and Haber 2016). Rather, a competition for the selected repair mechanism takes place that is dependent on the availability of cellular factors to initiate non-homologous end joining (NHEJ) or homology directed repair (HDR). HDR is facilitated but still not always favoured when a DNA template is provided to direct allelic information for gene conversion, whereby specified nucleotides can be copied in reverse-complement into the DSB site (Bozas et al. 2009; Jasin and Haber, 2016). There is firm experimental evidence suggesting template information can be provided as either ssDNA or dsDNA with potentially no size limitation just differences in efficiency for the gene conversion (Paix et al. 2016). Conversion by HDR takes place by one of two components: invasion-mediated synthesis-dependent strand annealing (SDSA) pathway or by single-strand annealing (SSA). This underscores that editing based on DSB followed by HDR is a completely natural process, because there is no transferred or introduced recombinant DNA constructs or synthetic DNA placed into the target genome.

There are currently only two proven ways to deploy gene editing technology in food animals. These are through transfection of fibroblasts destined for nuclear transfer cloning (primordial germ cells in poultry) or by microinjection of mammalian one celled embryos (Tizard et al. 2016). The transfection/injection of the "molecular scissors" into animal germplasm (fibroblasts/one-cell embryos) only increases the frequency of DNA breaks at the target locus billions fold over the natural processes of mutation. The DSBs stimulate gene conversion, which is not something novel or unnatural. In the case of sister chromatid repair, the template is copied, not physically acquired. This is analogous to gene conversion processes during meiosis, which is supported by evidence in yeast and *Drosophila* that somatic repair goes by the SDSA mechanism and involves invasion, copying, then withdrawal of the extended strand and re-pairing with the other end of the break (Bozas et al. 2009). There is also evidence in *C. elegans* that indicates that the whole of a large insert is copied *de novo* (Paix et al. 2016). This type of recombineering blurs the distinctions laid

out between categorizing gene edits as something mechanistically different when considering differences in sequence length post gene conversion (e.g.,. a 10 bp versus a 1000 bp increase in allele size). The only way to separate an introduction of natural alleles from trans or cis-genesis events is to define these events as "could not be obtained by conventional breeding" or not, respectively. The opportunities now exist to introduce any sequence without the need for a recombinant DNA construct.

The deployment of editing tools for precision breeding has a spectrum of efficiencies depending on the input parameters of the editing tools, such that in most cases, many of the resultant animals from microinjection of IVF embryos may have no edits or be mosaic for an edit (Wei et al. 2015). This means some animals produced from IVF embryo injection may need to go through a Mendelian transmission test to confirm commercial viability as a germ plasm product. For regulatory considerations and commercial viability of the technology, it would be beneficial if any animals produced by injection treatment of IVF embryos, which retain no edits or cannot transmit an edit by sexual reproduction, would be treated as conventionally bred animals. In essence, the mutagenic treatment failed as if there were no treatment applied. Furthermore, any recipient animals carrying edited clones or IVF embryos should have no restrictions under conventional animal quality measures relative to entering the food chain for human consumption.

Some concern has also been raised relative to the stability of an edit and other unintended edits at off-target sites in the genome. This concern is raised based on past observations of some transgenes being lost over time from modified genomes. However, the terms used by regulators regarding "genetic stability" & "genome integrity" are meaningless phrases in the context of genome editing. For example, deep sequencing reveals that in a typical human genome there is an average of 1 mutation every 1000 bp ($6X10^6$ total), >50 Loss-Of-Function mutations in disease-related genes, hundreds sequence (gene) duplications and translocations, 10^{15} active transposons in a single human (>100/cell), and 60-100 new *de novo* mutations not from the parents ($1/10^8$ bp) with mutations varying from cell-to-cell in a single person. The rates of natural mutagenesis have been shown to be similar in cattle (Kadri et al. 2016), and negative outcomes relative to phenotype can occur from such "normal" mutagenic events even though the animals are predicted to be of superior genetic merit (Schutz et al. 2016).

Current traits and demonstration animals. Besides the need to integrate with existing systems of selection, we believe commercial deployment of genome editing should initially be focused on those traits that are beneficial along the entire value chain from animal to consumer. Such a strategy promotes animal welfare, sustainability, and consumer acceptance of the technology. For example, one of the first traits deployed in cattle was introgression of the *celtic* polled allele into horned dairy animals (Carlson et al. 2016), where changes were made using natural occurring alleles; consistent with conventional breeding principles. The resultant animals demonstrated that genome editing could benefit animal welfare by eliminating stressful management practices (dehorning), while achieving acceptance from animal advocacy groups that influence consumer food product decisions. The only remaining challenge for commercial deployment of *polled* by gene editing is regulatory approval.

In addition to consumer acceptance and regulatory approval, another limiting factor for commercial deployment is the availability of known sequence variants for target traits that add substantial value to offset the costs of trait deployment by ART. The current number of known variants that have both major effects on production, health and/or welfare and exist in low allele frequency in the most popular breeds is severely limited, especially with respect to poultry and

swine. Some of this can be attributed to the fact that most livestock traits are "complex", i.e., variation in phenotype is due to effects from numerous loci with additive gene action, usually in the range of 10-100, sometimes thousands of genes. Whole genome selection in a breeding program iteratively enriches for desirable production alleles at all these loci, and the molecular markers used for guiding selection decisions rarely correspond to causal polymorphisms. Identification of causal alleles is not necessary for substantial genetic progress by genome selection, but is necessary for moving alleles to new genetic backgrounds.

The historical emergence and selection for major effect alleles indicates that the right mutation in the right gene can have a dramatic effect on a complex trait. Indeed, our appreciation of genetic potential is likely limited by epistasis, pleiotropy, and small effective population sizes. Most variants of this type as targets for editing have been reported in cattle. These variants have major effects on traits like thermotolerance (*SLICK*), muscling and tenderness (*double muscling*), coat colour, milk components (*DGAT1*), and fertility (multiple recessive lethals like HH1 and JH1). To date, these variants have only been introduced or corrected in bovine fibroblast cell lines using TALENs and HDR templates, and only a single Nelore bull was made with a *myostatin* knock out by IVF injection of TALENs into single cell embryos (Carlson et al. unpublished). Eventually, these bovine traits will be deployed for commercialization, but they may not represent the best opportunities for the use of precision breeding in food animals.

A strong argument can be made that genome editing for disease resistance traits represent the best commercial opportunity, because these traits improve animal well-being while changing industry inputs through reduction of reliance on antibiotics, vaccination, and other physical biosecurity protocols and surveillance. Again, there are practically no known variants that contribute to a substantial proportion of the phenotypic variance for resistance to a specific pathogen. Discovery of such variants may improve with revised reference genomes that contain more accurate assemblies of immune complex gene clusters. Better SNP tools and sequence alignments can then be applied to improve past and future variant discovery efforts. However, until then, we must rely on our limited knowledge of host:pathogen interactions for editing by rationale design of candidate genes. Other methods for informing rationale design and testing hypothesized causal alleles, especially for disease resistance, are needed. Recent methods by Yueng et al. (2017) demonstrate the power of using CRISPR/Cas9 to edit candidate immune genes in stem cells that can be differentiated into macrophages for pathogen challenge testing to compare how variants change response to infection and disease. Also, new gene targets for resistance can be identified using CRISPR library screening methods to interrogate gene function across an animal genome in a systematic and comprehensive manner (Zhou et al. 2014).

Although bovine traits based on "rationale design" have received publicity for conferring resistance to TB (Gao et al. 2017) and bovine respiratory disease, neither case has demonstrated resistance through a pathogen challenge of the edited animals. In pigs, Carlson et al. (unpublished), have potentially developed pigs with natural variant edits that are resistant to foot and mouth disease virus (FMDV). FMDV is a member of the picornaviruses, which replicate after infection by taking over the host cell's protein synthesis machinery. The viral proteses expressed early in the viral life cycle target and disable cap dependent translation of mRNA to shift protein synthesis to its own uncapped mRNA. The edits copied into these potentially FMD resistant pig were based on a two amino acid change to *EIF4G1* at a predicted viral cleavage site targeted by FMDV proteases (L^{pro}). L^{pro} cleavage of EIF4G1 disables the mRNA bridging function that helps allow cap dependent proteins synthesis. The amino acid variation in edited sequence was derived from *EIF4G2*; and thus, represents a natural variant found in swine. We showed these edits in *EIF4G1* gave pig

embryonic fibroblasts protection from cell death caused by active L^{pro}, and the edited fibroblast cells with altered *EIF4G1* genotypes were also resistance to L^{pro} cleavage. This result provides support for further investigation to test mutant *EIF4G1* cells and animals with live FMDV challenges in a biosecure facility.

To date, probably the best case of gene editing to make an animal resistance to a pathogen was reported by Whitworth and colleagues (2015), where a knockout of swine *CD163* resulted in protection from infection by porcine reproductive and respiratory syndrome virus (PRRSV). However, subsequent efforts by Wells et al. (2016) to swap domains with a human paralog of *CD163* and by Burkard et al. (2017) to only delete exon 7 of *CD163* suggest that the other functions of this gene must be maintained to make a commercially viable, healthy animal. Thus, the continued focus to only alter a specific portion of *CD163* that facilitates infection of PRRSV into the host macrophage.

Regulation and impediments to commercialization. Gene editing has emerged as a powerful research tool that can be used to systematically test hypothesized genotype/phenotype associations, particularly for major effect alleles. Furthermore, editing can be a powerful tool to study epistatic and pleiotropic effects, enabling the comparison of phenotypes presented when a polymorphism resides in the context of original versus a comparator genome. It seems very likely that gene editing may serve not only as a research tool, but also a way to achieve non-meiotic introgression of high value polymorphisms into commercial populations, breeds, or elite individuals where they don't already exist or are present at frequencies too low for effective enrichment by selection.

So what is impeding widespread activity of such research towards commercialization of gene edited animals? The answer is probably the uncertainty of the regulatory approval systems, which inhibits investment and negatively affects innovation and commercialization. Even though molecular genetics is a highly precise science (an investment of over $300 Billion dollars in knowledge about DNA since the 1980s), our regulators in the US insist that SDN mutagenesis of DNA, where the cell's DNA repair mechanisms act naturally to produce the genetic change, is not a breeding technique but rather a drug treatment. Furthermore, the parameters laid out for approval of gene edited animals mirror those for testing efficacy of chemical compounds that may have residues in food products. Editing has no residues. Unintended mutations are possible, but requiring a measurement of a "Durable Genome" and the purity of the enzyme used to induce the DSB is scientifically a nonsensical concept. Genome editing is very predictable (more so than sex), and the technological innovation based on SDN activity is exponential. It is inevitable that over-regulation will always be based on outdated concepts, which opens up opportunities to amplify public fears by the big business of anti-biotech advocates (e.g., Greenpeace). The consequences of lengthy, expensive regulation have the potential to greatly diminish the widespread use of editing in livestock. If regulation is expensive, in a commercial endeavour with narrow profit margins for the genetics provider; then only a few "blockbuster" products will make it through regulatory approval. Expensive regulation leaves out countries that need gene edited animals the most, and supports a few multinational companies (e.g., Bayer) that can eventually overcome cost barriers to profit. However, if costs go down, then even farmers in small ecosystems can benefit from optimized agricultural animals.

We have the power to improve the world –will we use it? Today the conversation remains at a 1980s level, due to wilful ignorance by well-meaning, intelligent people. Challenges for acceptance are public "concerns" and resistance to the technology is more apparent than real. So even though the science is straight forward; regulatory approval is not. In the current schema and based on

previous approval of a single GMO animal for consumption, one could predict that regulation will be 99% of financial costs of precision breeding by genome editing in the US. This will extinguish the use of gene editing for animal improvement. The fact-based argument, that ever-increasing changes in micro-climates and emerging disease threats around the world supports the application of single generation re-tooling of animals to accommodate global demand for protein in future generations, seems to have no resonance with the anti-biotechnology advocates. Their position blocks the intuitive need for genetic retooling of agricultural products, to better suit production in new environments world-wide to accommodate population increases and environmental constraints on agriculture. That can only happen if new, climate-adapted strains of animals are available at reasonable cost in poor countries. Ultimately, most rationale well intended people want the same outcomes from the use of new technology - healthy people, animals and environment – worldwide. Genome editing offers risk-free solutions – worldwide.

REFERENCES

Bozas A., Beumer K.J., Trautman J.K., Carroll D. (2009) **182**:641-51.

Burkard C., Lillico S.G., Reid E., Jackson B., Mileham A.J., Ait-Ali T., Whitelaw C.B., Archibald A.L. (2017) *PLoS Pathog.* **13:** e1006206.

Carlson D.F., Lancto C.A., Zang B., Kim E.S., Walton M., Oldeschulte D., Seabury C., Sonstegard T.S., Fahrenkrug S.C. (2016) *Nat Biotechnol.* **34:** 479-81.

FAO. (2017) The future of food and agriculture – Trends and challenges. Rome.

Gao Y., Wu H., Wang Y., Liu X., Chen L., Li Q., Cui C., Liu X., Zhang J., Zhang Y. (2017) *Genome Biol.* **18:** 13.

Jasin M., Haber J.E. (2016) *DNA Repair* **44:** 6-16.

Kadri N.K., Harland C., Faux P., Cambisano N., Karim L., Coppieters W., Fritz S., Mullaart E., Baurain D., Boichard D., Spelman R., Charlier C., Georges M., Druet T. (2016) *Genome Res.* **26:** 1323-1332.

Schütz E., Wehrhahn C., Wanjek M., Bortfeld R., Wemheuer W.E., Beck J., Brenig B. (2016) *PLoS One.* **11:**e0154602.

Tizard M., Hallerman E., Fahrenkrug S., Newell-McGloughlin M., Gibson J,. de Loos F., Wagner S., Laible G., Han J.Y., D'Occhio M., Kelly L., Lowenthal J., Gobius K., Silva P., Cooper C., Doran T. (2016) *Transgenic Res.* **25**: 575-95.

Wei J., Wagner S., Lu D., Maclean P., Carlson D.F., Fahrenkrug S.C., Laible G. (2015) *Sci Rep.* **5:**11735.

Wells K.D., Bardot R., Whitworth K.M., Trible B.R., Fang Y., Mileham A., Kerrigan M.A., Samuel M.S., Prather R.S., Rowland R.R. (2017) *J Virol.* **91:** e01521-16.

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. (2015) *Nat Biotechnol.* **34:** 20-2.

Yeung A.T.Y., Hale C., Lee A.H., Gill E.E., Bushell W., Parry-Smith D., Goulding D., Pickard D., Roumeliotis T., Choudhary J., Thomson N., Skarnes W.C., Dougan G., Hancock R.E.W. (2017) *Nat Commun.* **8:** 15013.

Zhou Y., Zhu S., Cai C., Yuan P., Li C., Huang Y., Wei W. (2014) *Nature.* **509:** 487-91.