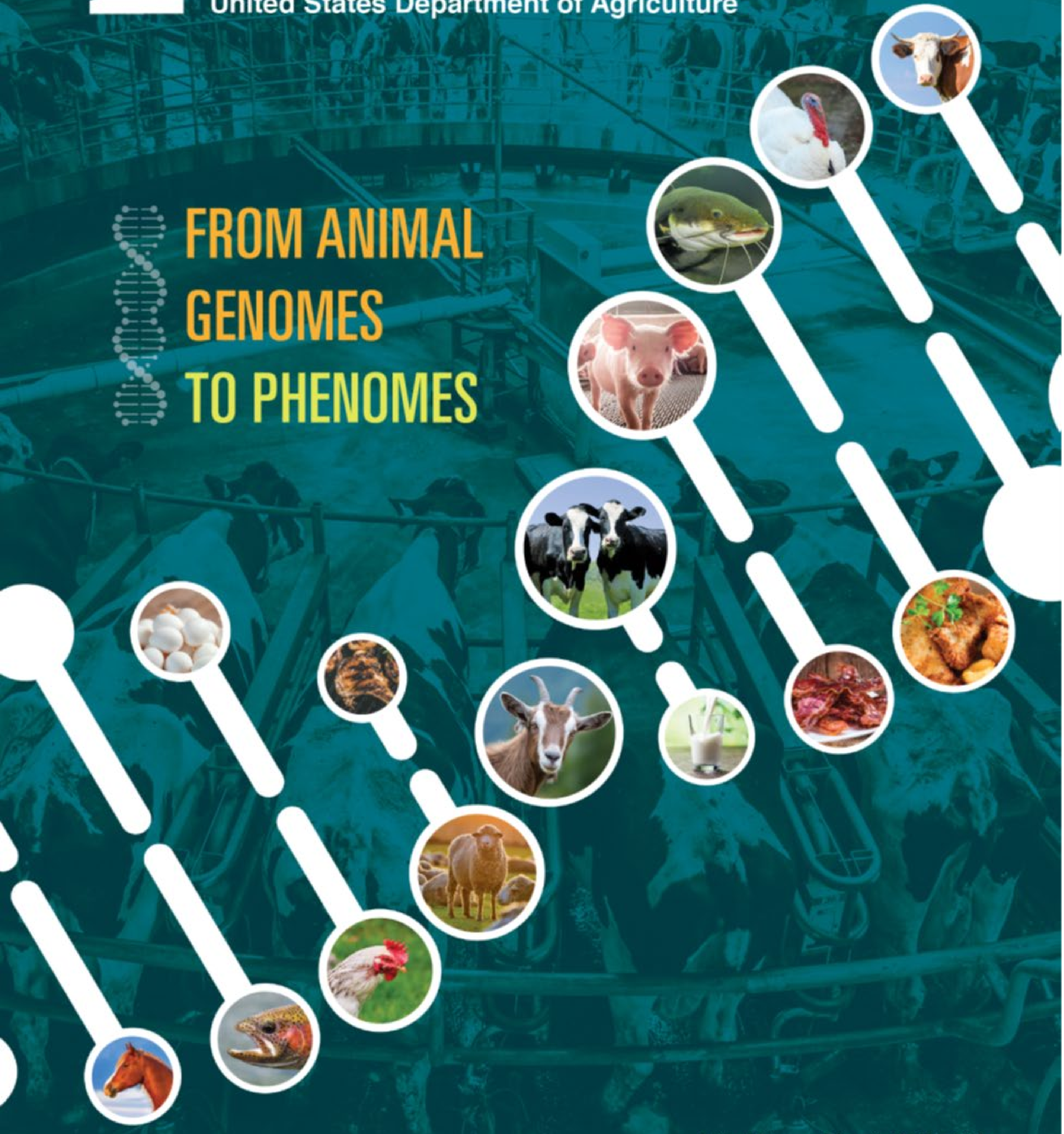
United States Department of Agriculture

FROM ANIMAL GENOMES TO PHENOMES



Animal Genome Blueprint 2018 – 2027

An Executive Summary of "Genome to Phenome: Improving Animal Health, Production, and Well-Being – A New USDA Blueprint for Animal Genome Research 2018–2027."[1]

Caird Rexroad[1], Angelica Van Goor[2], Lakshmi Kumar Matukumalli[2], and Jeffrey Vallet[1]

[1]USDA Agricultural Research Service
[2]USDA National Institute of Food and Agriculture

Summary

In 2008, a consortium led by the Agricultural Research Service (ARS) and the National Institute for Food and Agriculture (NIFA) published the "Blueprint for USDA Efforts in Agricultural Animal Genomics 2008–2017", which served as a guiding document for research and funding in animal genomics. In the decade that followed, many of the goals set forth in the blueprint were accomplished. However, several other goals require further research. In addition, new topics not covered in the original blueprint, which are the result of emerging technologies, require exploration. To develop a new, updated blueprint, ARS and NIFA, along with scientists in the animal genomics field, convened a workshop titled "Genome to Phenome: A USDA Blueprint for Improving Animal Production" in November 2017, and these discussions were used to develop new goals for the next decade. This document summarizes the strategy outlined in the new USDA Animal Genome Blueprint that will significantly increase our Nation's capacity to meet global consumer demands for animal products over the next decade.

---

[1] Frontiers in Genetics May 2019 Volume 10:327 doi: 10.3389/fgene.2019.00327

Introduction

By the year 2050, the global population is projected to grow from 7.6 billion to 10 billion, along with a general rise in the standard of living. Both these increases are expected to generate a growing demand for animal products, which currently provide 18 percent of total calories and 39 percent of the protein in human diets worldwide. To meet future demands, animal agriculture, which contributed $441 billion to U.S. economic output in 2014 alone[2], will need to boost production—and do so while maintaining or reducing its requirements for land, water, and feed, and while adjusting to fluctuating production factors associated with climate change, consumer production shifts, and disease challenges.

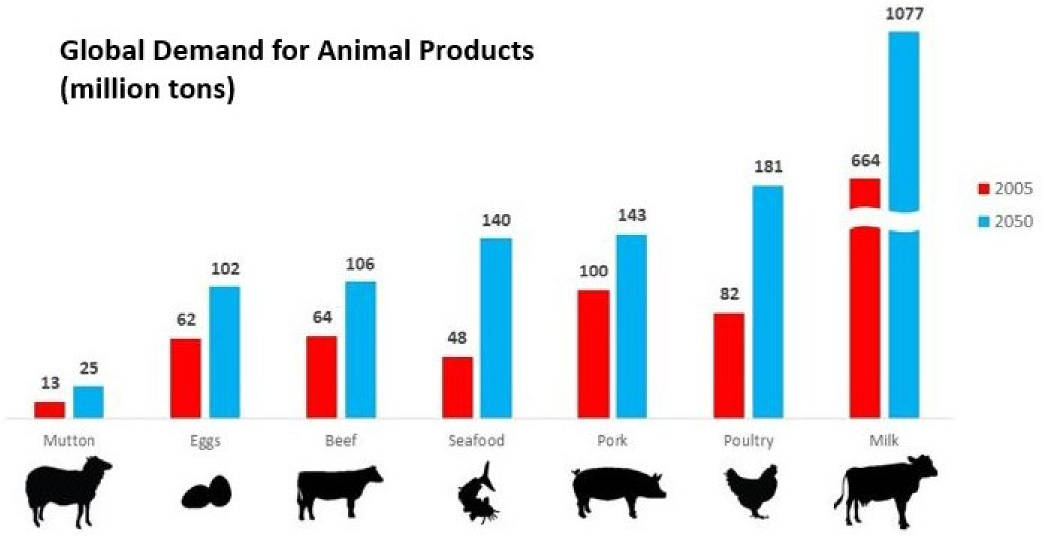


Figure 1. The global demand for animal products is expected to dramatically increase by 2050[3].

Animal production research has dramatically improved individual animal productivity and its environmental sustainability. In 2005, 2-month-old chickens weighed four times more than their 1957 counterparts, while dairy cows now produce three times as much milk as they did in 1960. The genetic selection of animals for improved performance and genomic technologies have greatly accelerated and contributed to these improvements in livestock production efficiency, and these improvements must continue to meet future global demands for animal products.

In 2008, the USDA Agricultural Research Service (ARS) and the National Institute of Food and Agriculture (NIFA) collaborated with scientists conducting animal genomics research at other institutions and released a Blueprint for Animal Genomics Research. This document was developed to provide a framework for animal genomics research for the following decade. The Blueprint was divided into three broad sections encompassing the work needed to fully realize the potential of genomic technologies for improving animal production efficiencies. **Science to Practice** focused on research intended to implement genomic selection in commercial animal

---

[2] Economic Analysis of Animal Agriculture 2004 – 2014 (https://unitedsoybean.org/wp-content/uploads/Economic-Analysis-of-Animal-Agriculture-FINAL-Low-Resolution.pdf)

[3] Modified from Future of Food 2013 (https://www.gatesnotes.com/About-Bill-Gates/Future-of-Food)

production. **Discovery Science** described the basic knowledge of animal genomes that would be needed to realize genome enhanced animal selection. **Infrastructure** described the computing, data analysis, and enhanced training that would be needed to advance the field.

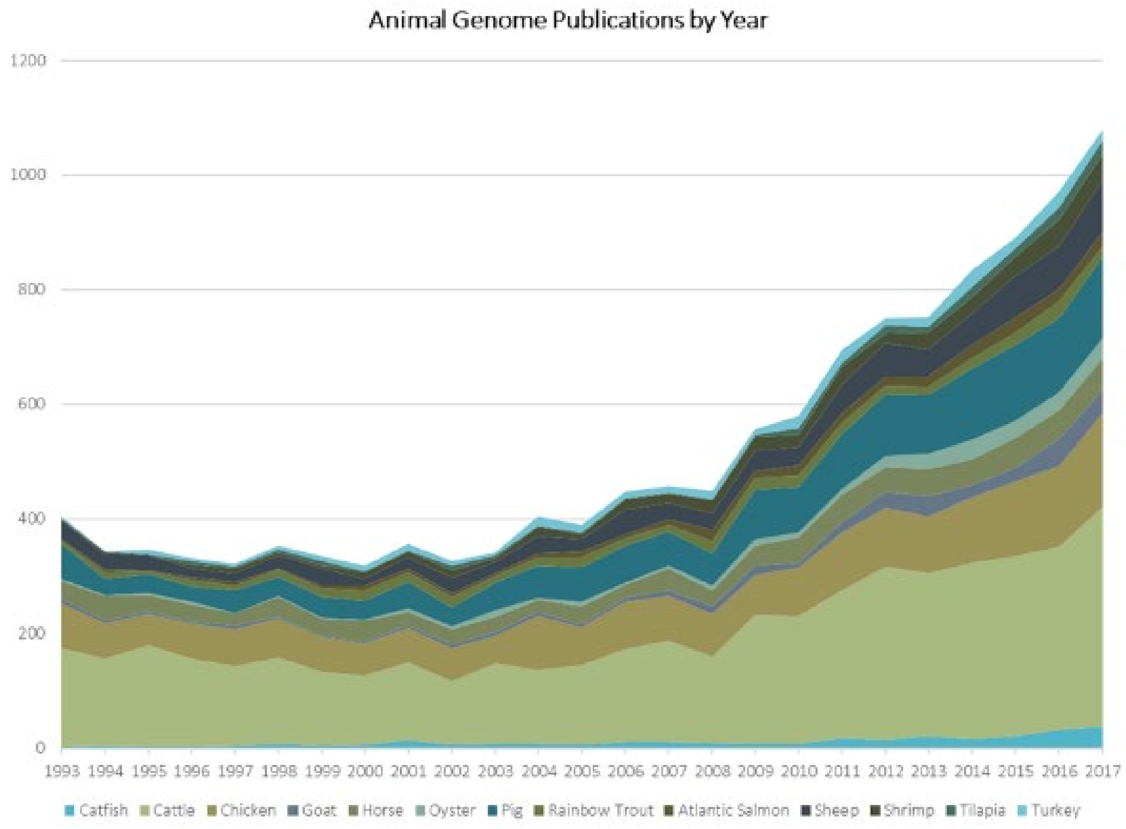


Figure 2. Genomics publications per year have nearly tripled in the last two decades.[4]

The research that resulted from the Animal Genomics Blueprint has been extraordinarily successful. Studies proposed in the first Blueprint for defining the genomes of livestock species have resulted in the production of complete genomes for cattle, swine, sheep, goats, chickens and catfish. Technologies and analysis methods have been developed for the comprehensive genotyping of individual animals at a variety of scales, including complete genome sequencing. The availability of livestock genomes has supported the development of techniques for analyzing every transcribed gene, as well as identifying and measuring all the proteins produced in various animal tissues. Genomic selection has been implemented in livestock species that are key components in animal agriculture, and the groundwork has been produced for future studies. The first Blueprint was written to guide animal genomics research from 2008 to 2017 and findings from these studies have generated new research goals and targets. A new Animal Genomics Blueprint is now needed to guide work in the next decade to address these new goals and targets.

---

[4] From Appendix 5, A New USDA Blueprint for Animal Genome Research 2018 – 2027, Front. Genet. 10:327. doi: 10.3389/fgene.2019.00327.

To generate this new Blueprint, ARS and NIFA scientists worked with their colleagues in the animal genomics research community to update unmet goals from the first Blueprint and develop new goals. The new Blueprint retains the three broad categories of Science to Practice, Discovery Science, and Infrastructure. As in the previous Blueprint, these broad topics are divided into specific research areas, including a vision and exploration of the current state and future goals for each research area. The new Blueprint contributes to the four main goals for animal production: 1) Providing Nutritious Food for a Growing Human Population, 2) Improving Sustainability of Animal Agriculture, 3) Increasing Animal Fitness and Improving Animal Welfare, and 4) Meeting Consumer Needs and Choices.

SCIENCE TO PRACTICE

Genomic Selection in U.S. Animal Agriculture: Commercial Implementation of Genomic Technology

A primary goal of animal genomics research is to use genomic information for improving how animals respond to selection. Over the last decade, genome-enabled technologies became integral components of commercial animal breeding for many species as advances in DNA sequencing and genotyping dramatically increased the ability to obtain genome information. More dramatic gains through breeding have been achieved using low-heritable traits because the accuracy of predicted genetic values increases proportionally for low-heritable traits than for more highly heritable traits. The dairy industry implemented genome-enabled genetic predictions and led the agri-genomic revolution, and similar progress has been achieved in beef cattle, swine, and poultry breeding as well. However, more opportunities are needed for using genomic information for breeding in other agriculturally relevant species, while improved public-private partnerships are needed to facilitate technology transfer.

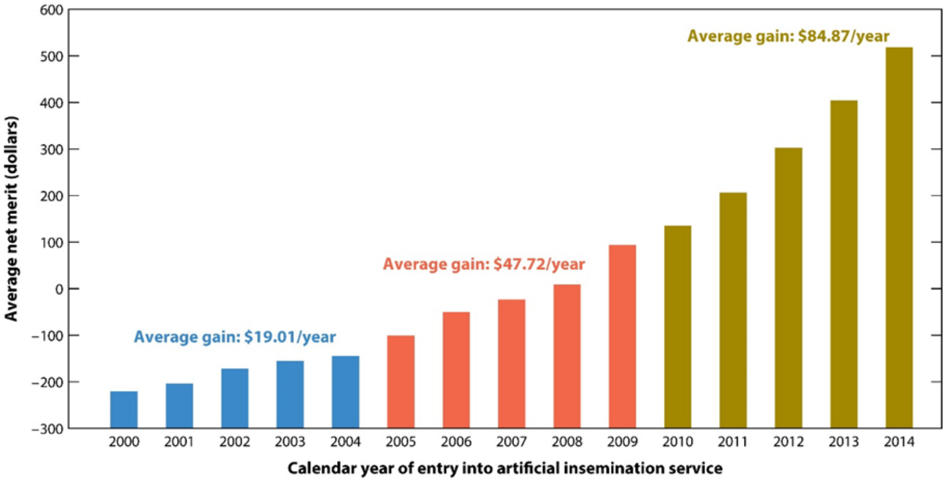


Figure 3. The dairy industry has led the agri-genomic revolution with the implementation of genome-enabled genetic predictions that have resulted in rapid growth in the economic value of animals.[5]

---

[5] Reproduced from Wiggans *et al*. 2017 (Anim Rev Animal Bioscience 5:309-327) which is not subject to copyright protection.

Implementing Genome Science into Animal Production

Whole genome-enabled animal selection, a goal of the previous Blueprint, has been successfully implemented in dairy and beef cattle, swine, poultry and some aquaculture species such as rainbow trout. Genomic selection has doubled the rate of genetic progress in dairy cattle, largely through enhancing breeding structures and reducing generation intervals. While implementing genomic selection in pigs and poultry has increased genetic progress rates by 30 to 50 percent. Reducing generation intervals via genomic selections in livestock will dramatically improve the genetic advancement of animal populations, and future progress will require defining functional variations in genome sequences. In addition, more useful animal phenotypes are needed to expand genomic animal research, and automated processes are needed to collect this data. Furthermore, it is possible that in animals in which heterosis is appropriately used, the resulting production improvement may equal improvements achieved with genomic selection based on additive inheritance. Finally, improvements in routine genotyping technologies are needed to increase their cost-effectiveness. These improvements will also expand use of existing technologies such as parentage verification, which is already routine in the beef and dairy industries.

Figure 4. Genomic selection in pigs has increased the rate of genetic progress by 30 to 50 percent for economically important traits such as litter size.[6]

Optimizing Animal Production through Precision Breeding and Management

New genomic selection technologies are needed for many food animal species, particularly aquaculture species. For some livestock species, some genomic information is available, but more is needed about specific DNA variations that influence gene function, particularly gene expression. For all livestock species, current genomic selection technologies focus on selection averaged over numerous environmental factors such as temperature and humidity and those influenced by management, such as feeding, stocking densities, and housing environments. Genomic technologies will be developed that incorporate environmental variation into models and result in targeted environmental selections or production values. Environment-specific

---

[6] From the Genus Annual Report 2015, with permission courtesy of Ernst Van Orsouw (https://www.genusplc.com/investors/results-reports-and-presentations).

production values will use genomic factors to develop precision management strategies for livestock according to their genotype, which will enhance production gains. Finally, while current genomic selection is based on additive inheritance, new genomic technologies will be developed that enhance nonadditive inheritance effects like heterosis and epigenetic mechanisms.

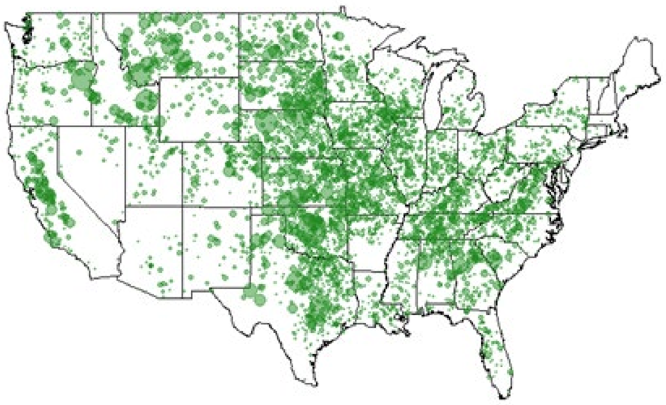


Figure 5. Although genomic information is widespread, as shown on this map of 521,645 genotyped Angus, opportunities exist to incorporate variation in the environment into genomic predictions. .[7]

DISCOVERY SCIENCE

Understanding Genome Biology to Accelerate Genetic Improvement of Economically Important traits

Much of our understanding of livestock genome biology comes from work in humans and rodents. When similar genes are present in livestock species their gene function may be inferred from available information about human and/or rodent gene function, but this inference may be partially or completely incorrect and needs to be validated through additional research. In addition, there are several livestock genes that do not have counterparts in model species. Variation among livestock species also occurs via post-translational modifications in the form of alternative splicing or other RNA-RNA interactions. Finally, although protein sequences can be predicted from gene sequences, very little information is available on regulatory elements influencing gene expression in livestock. Using model species information to study the function of noncoding sequences like promotors and enhancers and predict their expression in livestock is often not very useful. More work is needed on linking genes to functions, cataloging gene expression (including differential RNA modification), and cataloging DNA sequence motifs and other modifications that alter gene expression. The Functional Annotation of Animal Genomes

---

[7] Figure kindly provided by Dan Moser, American Angus Association and Front. Genet. 10:327. doi: 10.3389/fgene.2019.00327.

(FAANG) is collaborative effort within the animal genomics community with the purpose of addressing these research needs.

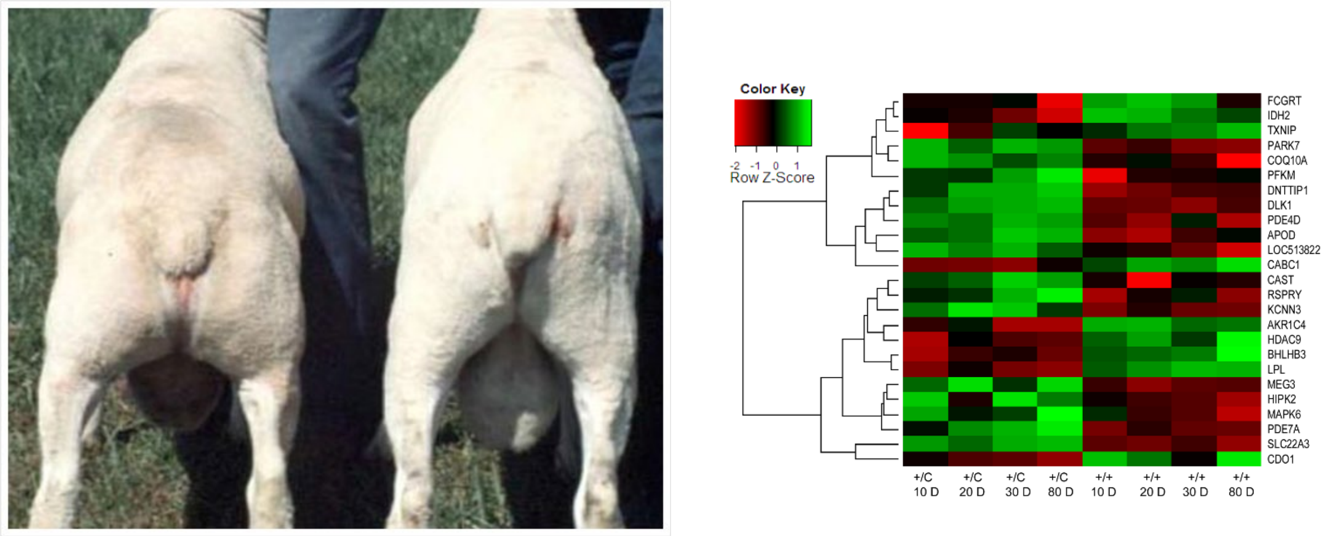


Figure 6. A genetic mutation in sheep that changes gene expression results in the callipyge, Greek for "beautiful buttocks", phenotype.[8]

Reducing the Effects of Animal Diseases

Genome-enhanced selection can be used to improve livestock disease resistance/tolerance, but livestock disease resistance/tolerance traits can also be some of the most difficult to approach with current technologies. Resistance is only displayed when the animal is exposed to the disease; even during an outbreak, exposure can occur in a very sporadic manner and these conditions are costly to recreate in a laboratory setting. To make progress in genomic selection for disease resistance, comprehensive animal disease surveys and databases are needed to fully employ naturally-occurring disease processes within the animal industries and identify potentially resistant/tolerant individuals. In addition, strategies are needed to identify measurable immune activity and other physiological indicators in healthy individuals before and during disease challenges to assess resistance/tolerance in healthy individuals. More information is also needed about host-pathogen-environment interactions and how these factors affect disease resistance. Work associated with these efforts will include continuing the development of livestock -specific biological reagents that can be used to obtain the necessary data and measures.

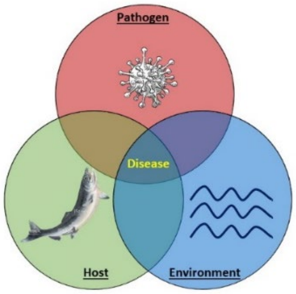


Figure 7. The host-pathogen-environment interactions must be examined to identify disease resistance/tolerance.

---

[8] From Fleming-Waddell JN, Olbricht GR, Taxis TM, White JD, Vuocolo T, Craig BA, Tellam RL, Neary MK, Cockett NE, Bidwell CA. Effect of DLK1 and RTL1 but not MEG3 or MEG8 on muscle gene expression in Callipyge lambs. PloS one. 2009 Oct 9;4(10):e7399.

Applying Precision Agriculture Technologies to Animal Phenotyping

Successful research conducted in alignment with the first Blueprint has resulted in accurate and cost-effective processes for genotyping many livestock species. Now work is shifting from genotyping itself to work associated with phenotyping, including the collection and measurement of economically important traits in individual animals. Some routine measures are collected during commercial production, but in most cases, there are not central locations where these measures can be accessed and used to further genomics research and inform selection. Opportunities exist to develop electronic sensors combined with improved on-farm data infrastructure to automate the routine collection of economically relevant measures. However, further research is needed to define these measures and develop the needed sensor technologies. Routine, automated measures are also needed for collecting data in production environments to inform environment-specific and management-specific genomic selection techniques. The simultaneous development of a data infrastructure is needed to support the full use of the collected data and to ensure the data are available to scientists and breeders to use for effective genomic selection.

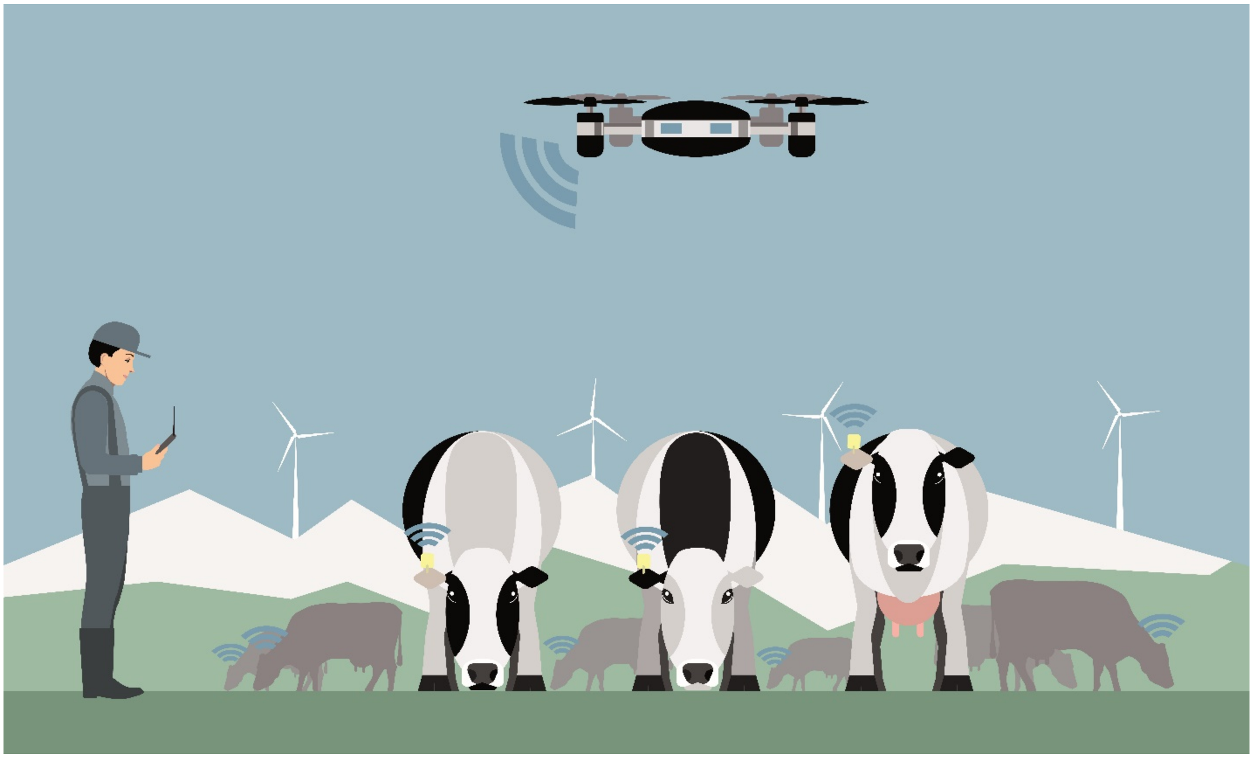


Figure 8. There is an opportunity to develop electronic sensors combined with improved on farm data infrastructure to automate the routine collection of economically relevant measures.[9]

---

[9] Shutterstock image753946003.

Harnessing the Microbiome to Improve the Efficiency and Sustainability of Animal Production

All individual animals are a collection of organisms consisting of the host animal and its microbiome—the symbiotic, commensal, and often parasitic or pathogenic organisms that reside either in or on the animal. Because of vital links between an animal and its individual microbiome, animal performance is influenced by the presence or absence of these organisms and their genomics and biochemistry. Investigations on the genomics of the microbiome (metagenomics) are in their infancy. To begin to address the role of the microbiome in livestock production and health, adequate reference genomes are needed for the many organisms that make up the microbiome, which often provides biochemical capabilities such as vitamin synthesis that are important to host function. These microbial capabilities can be specific to the genus, species and/or strain of each organism, and new technologies and assessment analyses are needed to comprehensively study and elucidate the varied and essential roles of microbiomes in livestock health and physiology. Efforts are needed to determine the role of microbiome during different animal states such as illness and different animal production facilities and environments. The goal of this research is to acquire the information and develop the techniques needed to beneficially manipulate the microbiome, either through host genomics, dietary strategies, or other management interventions, to optimize the function of the microbiome for livestock health and production efficiency.

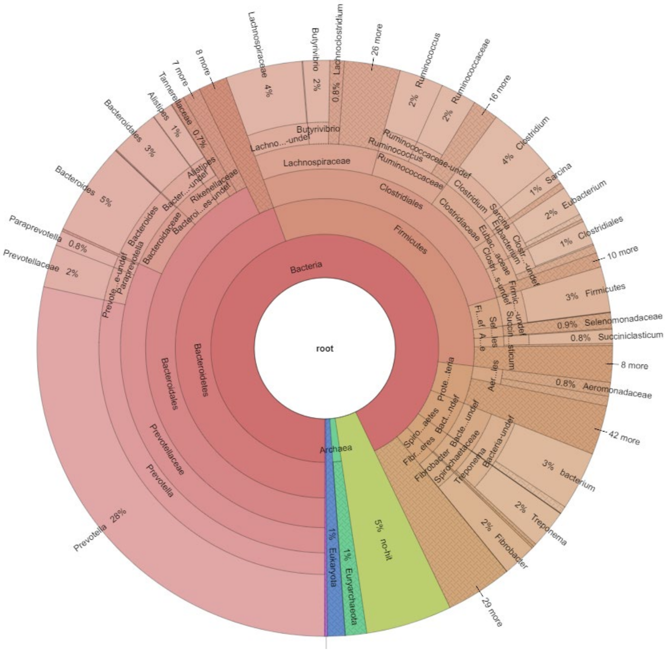


Figure 9. Microbial diversity in the cattle rumen captured through long read (PacBio) metagenome assembly. Image credits Derek Bickhart.

Figure 10. Every individual animal is a collection of organisms, made up of the host animal, and its microbiome, the symbiotic, commensal and often parasitic or pathogenic organisms that reside either in or on the animal that can affect animal productivity. DNA sequence information can indicate the status of an individual animal. [10]

---

[10] Photo of microbiome used with permission from Derek Bickhart. Gut microbiome from iStock photo gm586183710-100613233.

INFRASTRUCTURE

Training the Next Generation of Animal Scientists

Incorporating genomic information into selection is complex and requires sophisticated multidisciplinary training in quantitative genetic techniques and expertise in molecular biology. Training programs will be needed to ensure farm production staff possess the computational skills required for the automated, comprehensive collection of individual animal data. Precision management strategies for livestock developed using environmental and management variables associated with specific genotypes will also require expert on-farm data collection. The development and implementation of university curricula and farm staff training programs to support these needs will allow scientists and livestock production personnel to interact seamlessly, exchange animal data for analysis and genomic predictions, and adjust activities as needed to optimize management. At all levels, curricula that integrate computational instruction with production experience, such as animal handling and exposure to commercial production practices, are likely to be the most efficient way to develop these seamless interactions.

Figure 11. A multidisciplinary approach that includes computation, molecular biology, and animal production experience is required to train the next generation of animal scientists.[11]

Developing Advanced Genomic Tools, Technologies, and Resources for Agricultural Animals

Single-consensus genomes for many livestock species have been developed, but in many cases these genomes do not represent any animal that has ever or will ever exist. To fully understand how the genome functions and to use that information, genomes must be annotated to include information on sequence segments that control protein amino acid sequences, gene transcription, and higher-order genomic architecture. The FAANG consortium has started to provide this information for single-consensus genomes and genomes that are currently available. Additional work is now needed to generate the same information about the genomes of many individuals

---

[11] iStock photos 184602304 and 683253578

within each livestock species—a group of genomes called the pan-genomes. Fully annotated pan-genomes are needed to completely understand how sequence variations among individuals translate into changes in genome function in various tissues. To help with this effort, improved technologies and analysis techniques for global analysis of gene expression (transcriptomics), protein production (proteomics), and metabolite production (metabolomics) are needed to conduct comprehensive analyses of links between livestock genomes and economically important traits.

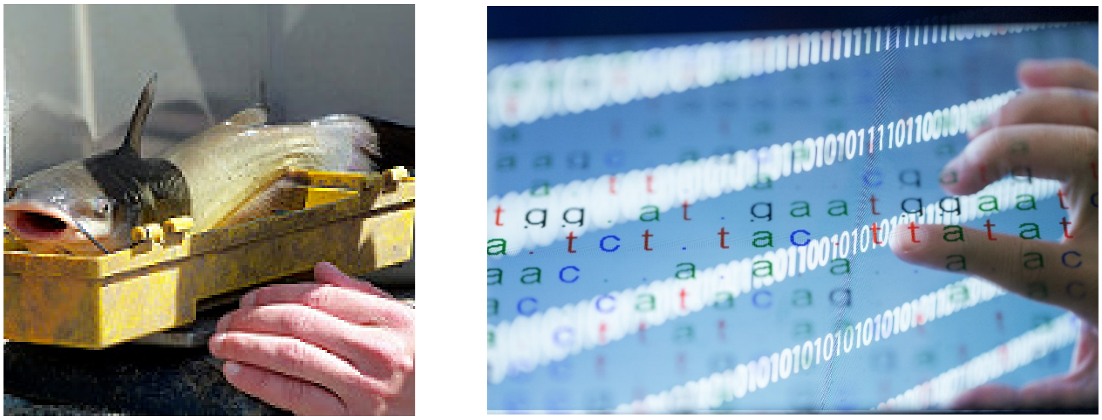


Figure 12. To be fully useful, genomes must be annotated to include information on the portions of sequence that control protein amino acid sequences, gene transcription and higher order genomic architecture.[12]

Creating Big Data Tools and Infrastructure for Animal Production

Current computer and data handling capabilities are sufficient to perform routine genomic selection analyses on animal measures that have already been collected. But the computer and data infrastructure needed to process, store, and use of all the genomic data, animal and environmental measures, animal health data, and microbiome measures included in this Blueprint are orders of magnitude beyond current capabilities. It will be necessary to significantly increase current computational capacities, including faster computers, greater storage, and faster and more interconnected networks, to achieve the goals included in this Blueprint. Better and faster software algorithms will be needed to effectively analyze these data, and data standards and defined interconnectivity among data will be needed. This will allow data to be effectively used for analysis and reused as analysis methods improve and systems approaches integrating many types of data are developed. Maintaining data accessibility and integrity is expensive, and support for permanent data storage will be needed to make maximum use of the animal genomics efforts described in this Blueprint.

---

[12] Catfish photo flickr 7463089188_38cd90f7ed_k; Genetic code photo iStock 695176924.

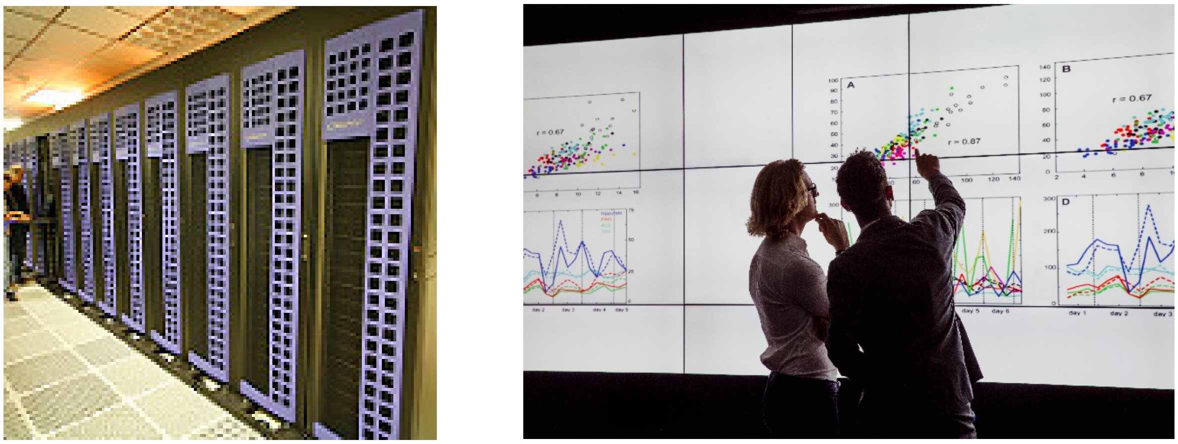


Figure 13. Hardware, tools, and data standards are required to support big data.[13]

Advancing Biotechnology to Improve the Sustainability and Efficiency of Animal Production

It is now possible using site directed nucleases (e.g., CRISPR/Cas9) to make very specific changes to livestock DNA. As more animal genomics information becomes available, it will be able more accurately predict how changes to DNA can enhance economically important traits in livestock. To fully use gene-editing technologies, advanced reproductive technologies will be needed to extend desirable modifications made in a single individual to much larger production populations, especially in livestock with long generation intervals. This will require improvements in *in vitro* fertilization of elite females, development of embryo genomic selection methods, and cloning technologies to expand the population size of embryos with elite desirable genotypes. To increase the use of these editing approaches, research is needed on using gene-editing in species where it has not been employed to date. In addition, to effectively modify complex multiple-gene traits, methods must be improved to increase the number of possible modifications in a single individual.

---

[13] Servers flickr phto 14650699941_da30665d1_k; iStock photo 872019580.

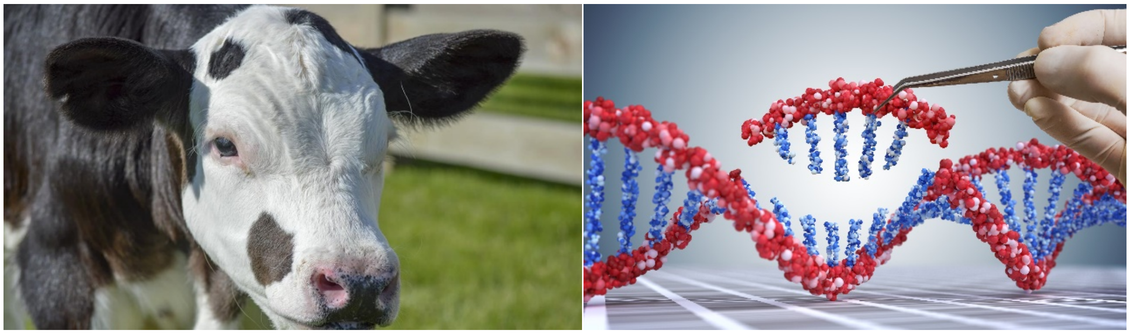


Figure 14. Gene editing by CRISPR/Cas9 can be used to improve animal welfare: one example is the development of hornless cows.[14]

Characterizing and Preserving Genetic Diversity for the Future of Animal Production

Figure 15. Dr. Sue Lamont holds an inbred chicken from a line used frequently for disease resistance/tolerance studies at Iowa

Throughout this new Blueprint, research has been proposed to improve livestock selection through various genomic approaches. This research will advance breeding activities for the development of livestock with genetic adaptations that support and improve health, increase disease resistance, and enhance resilience in the face of weather extremes and other challenges. However, obstacles to livestock production are always evolving. As research and development continues to result in the modification of animal genomes, it is essential to preserve current genetic resources to ensure their availability for addressing challenges in the future. The National Animal Germplasm Preservation facility in Fort Collins, Colorado, and other germplasm storage facilities in other countries, protect and preserve viable livestock germplasm in case it is needed. However, effective preservation methods are not yet available for all livestock species, so further work in preservation methods is needed for those species. For species where preservation methods are available, some of the available genetic diversity in underutilized livestock "heritage" breeds needs additional sampling and improved storage.

---

[14] Gene edited bull calf, photo credit: Hannah Smith Walker, Cornell Alliance for Science. DNA editing photo iSTock 104030074.

CONCLUSION

Current U.S. and worldwide animal production is the result of dramatic improvements in animal health and production efficiency achieved using genetic selection. Many of the goals outlined in the previous Animal Genome Blueprint have been achieved, including doubling the genetic progress in dairy cattle and increasing the rate of progress in other livestock species by 30 to 50 percent. However, some goals in the first Blueprint still need to be realized, while the research achievements that were realized have generated new goals. Further improvements in animal genomics will require additional support for genome discovery science, computing and data infrastructure and training, and practical solutions for incorporating environmental factors into genomic selection. With appropriate support, past advances will accelerate improvements in animal health and production and reduce the environmental footprints of current livestock production systems.