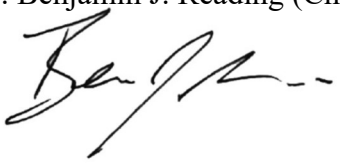
FINAL REPORT: NATIONAL RESEARCH SUPPORT PROGRAM NRSP-8 ACCOMPLISHMENTS

National Animal Genome Research Program (NAGRP)
10/01/2018 to 09/30/2023
Report Prepared by: Benjamin J. Reading (Chair) and Annette M. McCoy (Chair-Elect)



February 29, 2024

Background and Executive Summary: The NRSP-8 National Animal Genome Research Program (NAGRP) is a collaborative initiative involving multiple institutions, researchers, and stakeholders aimed at advancing genomic research in animals. From 1993 to 2023, the NRSP-8 NAGRP has been hugely successful, exceeding all expectations, by helping to deliver complete genome sequences of pig, cattle, sheep, goat, horse, chicken, turkey, and eleven aquaculture species including finfishes and shellfishes. The program provided genetic tools and resources that have revolutionized the animal breeding industry. Genome-enabled technologies co-developed under NRSP-8 NAGRP are now helping to deliver commercial animal breeding and production for many animal species, resulting in a multifold return on investment to United States stakeholders and producers. These successes, along with concomitant advances in genomics-enabled technologies, resulted in the release of the 2018-2027 USDA Animal Genome Blueprint (2018) which outlines key areas for future research and funding. A crucial element identified in the Blueprint report is "Science to Practice", which is the application of genomics-enabled technologies to traits and phenotypes critical to animal industries. The "Science to Practice" goals are undoubtedly attainable, in large part because of the genetic resources and technologies developed through the NRSP-8 NAGRP. Further, a long-term goal since 1993 of the NRSP-8 NAGRP community has been to support the development of tools to link genomics data with important animal traits and applications and to utilize this information within animal industries; this effort led to the successful NRSP-8 renewal project for 2023-2028 "Building Applied Genomic Capacity for Animal Industries". Lastly, the NRSP-8 NAGRP also contributed to the U.S. Department of Homeland Security report (2021) on "Threats to Food and Agricultural Resources" (TFAR), which is related to National Security Memorandum-16 (NSM-16) "Strengthen the Security and Resilience of United States Food and Agriculture" released in 2022. The NRSP-8 NAGRP has a very important mission in this regard as purveyors and stewards of genomic data critical to the safety and resilience of the United States domestic food supply (food security is national security).

These accomplishments reflect the broad scope and impact of NRSP-8 NAGRP animal genomic research and the diverse applications of genomic technologies in addressing key challenges facing animal agriculture, biodiversity conservation, and public health that encourage innovation to ensure continued access to safe, sustainable, and abundant food supplies for future generations.

Business Plan and Management: Organization of the NRSP-8 NAGRP included a national Chair and Chair-Elect and then seven subgroups thereon of bioinformatics services and commodity species including aquaculture, bovine, equine, porcine (swine), ovine (sheep and small ruminants) and poultry, each of which was headed by a national coordinator. Within each of the subgroups the suborganization included workshop coordinators, writing teams, committees, and other delegations that were designed based on the structures defined by each of the national coordinators and that varied slightly between groups. Leadership was designated through election or appointment. A national writing committee was also designated to lead the NRSP-8 renewal proposal efforts for the 2023-2028 "Genomic Capacity: Building Applied Genomic Capacity for Animal Industries" (https://nimss.org/projects/view/mrp/outline/18969). The collective NRSP-8 NAGRP group conducted activities directed towards achieving three main objectives and each year from 2018-2023 the national coordinators reported out to the Chair and Chair-Elect, who then completed the annual reporting on behalf of the group.

Successful efforts were made to develop platforms to facilitate collaborative research for collection and analysis of new information and to develop, integrate, and implement bioinformatic resources to support the discovery of genetic mechanisms underlying agriculturally important traits. The NRSP-8 continues to host mailing lists/websites for various research groups in the NAGRP community (https://www.animalgenome.org/community/). This includes groups like AnGenMap (with about 3,000 subscribers from 67 countries/regions of the world), Functional Annotation of Animal Genomes (FAANG) international consortium working groups (with 8 working group mailing lists and websites), and CRI-MAP users (670+ subscribers), and user bulletin boards to coordinate and facilitate meetings, among other user forums.

Tens of millions of $USD were leveraged from federal, state, private sector/industry, and other extramural sources to support NRSP-8 NAGRP activities during the reporting period and these efforts were facilitated using the modest and sustained support of $2.5 million USD allocated to the species coordinators from 2018-2023. In the previous NRSP-8 NAGRP funding cycle (2013-2017), the group showed a similar impressive return on research investment (ROI) in the form of leveraged research dollars of 40:1, indicating a strong project record of success for the group that continued for a decade through the 2018-2023 cycle.

NRSP Mission and National Scope: The specific objectives of the NRSP-8 NAGRP was in the development of genomics enabling technologies and, in the training, and education that supported its applications across a range of agricultural animal sciences and veterinary medicine disciplines and across invertebrate and vertebrate animal species in the United States. The NRSP-8 NAGRP brought together researchers in animal science, genomics, genetics, and bioinformatics as well as industry representatives to enhance national capacity with support from the State Agricultural Experiment Stations (SAES) and the overall goal of promoting sustainable and resilient food animal agriculture production in the United States.

OBJECTIVES: The NRSP-8 project was renewed for reporting period 10/01/18-09/30/2023, with the following three objectives:

1. Advance the quality of reference genomes for all agri-animal species by providing high contiguity assemblies, deep functional annotations of these assemblies, and comparison across species to understand structure and function of animal genomes;

2. Advance genome-to phenome prediction by implementing strategies and tools to identify and validate genes and allelic variants predictive of biologically and economically important phenotypes and traits; and

3. Advance analysis, curation, storage, application, and reuse of heterogeneous big data to facilitate genome-to-phenome research in animal species of agricultural interest.

Accomplishments and Outcomes: The NRSP-8 NAGRP made tremendous efforts in working toward project objectives and specific accomplishments that are listed below by objective. For detailed information on achievements and research outcomes of the NRSP-8 NAGRP program during the 2018-2023 period, please refer to the individual annual project reports hosted at the NIMSS portal (https://nimss.org/projects/view/mrp/outline/18464) and the International Plant and Animal Genome (PAG) Conference proceedings (PAG 27, 2019; PAG 28, 2020; PAG 29, 2022; PAG 30, 2023; and PAG 31, 2024; PAG in 2021 was cancelled due to COVID-19 pandemic travel restrictions).

Collectively, scores (score = 20) of conference travel awards were provided by the species coordinators to students, post-doctoral researchers, and early career faculty to attend NRSP-8 NAGRP events, workshops, and activities during the reporting period. This is critical for training the next generation of scientists and researchers in advanced animal genomics and to introduce them to the NRSP-8 NAGRP community, which is the best talent in the nation for the topic. The importance of NRSP-8 funding to provide leverageable and matching funds to support grants and activities relevant to the mission objectives across species groups cannot be overstated.

The NRSP-8 NAGRP group averaged about 70 publications generated from community members directly relevant to the objectives each year for a total of over 300 products from 2018-2024. A letter attesting to the request for comments from the OSTP/FDA regarding the modernization of biotechnology (i.e., genetic editing of animals in the United States) was produced as commentary from the NRSP-8 community as well as contributions to national security efforts in the realm of food and agriculture production.

1. Advance the quality of reference genomes for all agri-animal species by providing high contiguity assemblies, deep functional annotations of these assemblies, and comparison across species to understand structure and function of animal genomes;

The NRSP-8 NAGRP group capitalized on advancements including second generation short-read sequencing technologies, optical mapping, Pacific Biosciences sequencing and other emergent

technologies such as third generation sequencing and chromatin mapping. Across all species, improved genome assemblies reached high-quality chromosome levels, eliminated most of the regions with ambiguous sequences, and in some cases provided sequence for previously unknown genome regions. Genomes were completed or augmented for swine, cattle, sheep, goat, horse, chicken, and turkey, including robust telomere-to-telomere sequence assemblies and the sequencing of hundreds or thousands of individual animals to create powerful pangenome resources for several of the species such as chicken, bovine (cattle and dairy), and swine. It is worth noting that this was the first time some of these technologies were applied to these species. Aquaculture members contributed to genomes that were completed for 11 species, including 8 finfish (channel catfish, blue catfish, 2 subspecies of rainbow trout, North American Atlantic salmon, Nile tilapia, striped bass, white bass) and 3 shellfish (eastern oyster, Pacific geoduck, Pacific Olympia oyster) and genome sequencing projects are underway for 7 additional finfish subspecies of rainbow trout and cutthroat trout and 1 shellfish species, North American Pacific oyster. Altogether, these resources support the major terrestrial and aquatic animal agriculture industries in the United States.

Advances in bioinformatics tools facilitated the generation of high-quality genome assemblies and comprehensive gene annotations for these important agriculture animal species. The Animal QTLdb, CorrDB, Bioinformatics Tools, and the NAGRP data repository actively supported research activities for multiple agricultural animal species. Ontology hierarchy display tools were implemented that facilitated our efforts at expanding and exploring the use of Vertebrate Trait (VT) Ontology, Livestock Product Trait (LPT) Ontology, Clinical Measurement Ontology (CMO), and other ontology hierarchies. All curated association data were automatically ported to NCBI, Ensembl, UCSC genome browser, and Reuters Data Citation Index in a timely fashion. The NRSP-8 NAGRP made significant efforts to improve the curator tools to facilitate trait-variant related data curations and continued to strengthen the data quality control procedures to help improve data quality. The Data Repository for the aquaculture, cattle, chicken, horse, pig, and sheep communities have proven to be very useful in facilitating research (https://www.animalgenome.org/repository). Additionally, the AquaMine RefSeq Database for aquatic species was created (https://aquamine.elsiklab.missouri.edu/) which includes information from 21 vertebrate finfish and 4 invertebrate mollusks, and 3 invertebrate crustaceans (28 total species).

Efforts were initiated in 2014 to annotate additional functional elements of the cow, pig, chicken, horse, sheep, rainbow trout, and pacific oyster genomes as part of the Functional Annotation of ANimal Genomes (FAANG) consortium and these efforts continued during 2018-2023. The FAANG consortium was formed to accelerate genome-to phenome discovery, and this includes NRSP-8 NAGRP priority species. Work is ongoing among members of the FAANG and the Agricultural Genome to Phenome Initiative (AG2PI) project to standardize collection techniques, experimental protocols, and data analysis pipelines to maximize the utility and availability of data produced by NRSP-8 NAGRP efforts.

2. Advance genome-to phenome prediction by implementing strategies and tools to identify and validate genes and allelic variants predictive of biologically and economically important phenotypes and traits; and

Investigations of the functional elements of agriculture animal genomes, including coding and non-coding regions, regulatory elements, and epigenetic modifications, were enhanced by NRSP-8 NAGRP community members. Researchers identified genes and genetic variants associated with important production traits such as disease resistance, growth rate and efficiency, reproduction, animal welfare, and adaptation to various changing environmental and climate conditions. Genome resources developed for bovine have been impactful in the beef and dairy industries, especially regarding environmental impacts targeting reduction of carbon footprint and greenhouse gas emissions and genome editing in pig is expanding, including in the field of biomedicine. Functional annotations, single guide RNA (sgRNA) libraries, muscle cell lines, and spatial transcriptomes, among other completed tool refinements and resources are now available to poultry researchers. Across species tools developed under this objective by the NRSP-8 NAGRP community have allowed for identification of alleles responsible for important economic and production traits, including the genetic basis of infectious disease resistance, susceptibility, and resilience in animals. Additionally research into animal genetic sex determination mechanisms and other animal welfare traits have provided an understanding to produce sterile animals for culture as well. Thus, the group also works to deliver scientific solutions to challenges of escapement, sustainability, and wildlife conservation.

Development and application of molecular markers and genomic prediction models for breeding programs in livestock and other animal populations is another important outcome of the resources created by NRSP-8 NAGRP community members. Genome-wide association studies (GWAS) and genomic signatures of selection for identification of genomic regions harboring alleles for traits of interest have been explored, which allow for improved accuracy of predicted breeding values, enabling genomic selection, and permitting estimation of genetic diversity in breeds and populations of interest. Genomic selection allows breeders to make more informed decisions regarding trait improvement, genetic diversity, conservation, and the management of inbreeding depression, leading to accelerated genetic gains and more resilient animal populations and more efficient cultivation. Genomic selection has been initiated in terrestrial agricultural animals, but also aquaculture species including eastern oyster. Other aquatic species have domestication and selective breeding programs that interface with industry for commercial production including striped bass, white bass, rainbow trout, and channel catfish. Priority aquaculture species identified by USDA and NOAA National Marine Aquaculture Initiative (2019) include potential future targets for genome sequencing and genomic tool development (about 14 additional species total). These are species that already have commercial production or are deemed commercially viable for production; however, they completely lack genome resources.

3. Advance analysis, curation, storage, application, and reuse of heterogeneous big data to facilitate genome-to-phenome research in animal species of agricultural interest.

The most important accomplishment of the NRSP-8 NAGRP has been the formation of a large community of scientists working worldwide to advance animal genomics through the sharing of resources, development of open-access multi-species bioinformatic tools, sequencing and assembly of genomes, organization of workshops and conferences, communication of results, travel support for students, early career faculty and invited speakers, preparation of multi-institutional grant proposals, and formation of large collaborative research groups. The communication and sharing of information among the different species technical committees fostered by NRSP-8 NAGRP has led to significant achievements under the objectives outlined for 2018-2023. Across committees, the experience of one group has often informed and advanced the directions and approaches taken by other groups and this shared knowledge has accelerated tool development and discovery for all supported animal species.

Development of genomic databases, bioinformatics pipelines, and computational resources to support data storage, analysis, visualization, and sharing within the research community have strong user support and have empowered research across the United States and the world. Training programs, workshops, and educational initiatives aimed at building capacity in genomic research, bioinformatics analysis, and technology transfer among scientists, students, farmers, policymakers, and other stakeholders in the animal agriculture and veterinary sectors also is an important aspect of the NRSP-8 NAGRP. Open-access genomic data repositories and collaborative platforms that facilitate data integration, meta-analysis, and knowledge dissemination across diverse research domains were utilized and showcased by NRSP-8 NAGRP community members at various local, national, and international conferences. The NRSP-8 NAGRP conducted numerous Species Workshops for animal genomics from 2018-2023 except for 2021 when travel was restricted due to the COVID-19 Pandemic and even during that time the best efforts were made to host virtual events as a replacement for in person events. During each of the years 2018-2020 and 2022-2024, workshops hosted by the NRSP-8 NAGRP species groups at the annual International Plant and Animal Genome (PAG) Conference were attended by an average of 450 persons from at least 2 dozen different countries and were represented with attendees from academic, private sector, and government institutions; many additional workshops and training sessions were conducted outside of this venue as well during the reporting period including the topics of pangenomes, functional genomics, genomic enabled breeding, and machine learning.

Finally, the annual NRSP-8 workshops have become an essential component for the development of collaborations, training, and dissemination of new information to government, academic, and industry stakeholders in animal agriculture. The greatest strength of the NRSP-8 NAGRP is the community that it has built.