

SUMMARY OF NRSP-8 ACCOMPLISHMENTS: 2008-2012

Overview of accomplishments for all NRSP-8 technical committees

An important accomplishment of the NRSP-8 project has been the formation of a large community of scientists working world-wide to achieve the goals of this National Research Support Project. Specific accomplishments include sharing of resources, development of open-access multi-species bioinformatic tools, sequencing and assembly of genomes, organization of workshops and conferences, communication of results, support for travel for students and invited speakers, preparation of multi-institutional grant proposals, and formation of large collaborative research groups. Another significant accomplishment of the project has been the communication and sharing of information among the different species committees. The experience of one group has informed and influenced the directions and approaches taken by other groups in the areas of tool development, genome mapping, QTL studies and genome sequencing.

The NRSP-8 participants and their collaborators have far reaching national and international impacts in animal genomics. The community supported by NRSP-8 has developed and adopted new genomics technologies, including high-throughput short read sequencing, high density SNP detection and several types of microarrays to sequence genomes, generate high quality genetic maps, and analyze gene expression of a rapidly increasing number of species. The NRSP-8 members have facilitated the transfer of experience from the earliest species with sequenced genomes to the current efforts to produce higher quality integrated genetic maps and genome sequences for all species supported by the project. The NRSP-8 participants and their collaborators are using genomics technology, bioinformatics resources, and diverse animal models to investigate fundamental mechanisms affecting production efficiency, product quality, animal health, disease resistance and food safety. Genomics technology has been successfully translated for genetic selection programs for some sectors of animal agriculture. The NRSP-8 membership has fostered development of critical resources for animal genomics in the USA and throughout the world. The annual NRSP-8 workshops have become an essential component for development of collaborations, training and dissemination of new information to government, academic, and industry stakeholders in animal agriculture.

A summary of the major accomplishments for each of the technical committees (aquaculture, bioinformatics, cattle, horse, poultry, sheep, and swine) are listed below. No single overall summary is included given that each species committee and their research are at different stages of development.

Websites for each coordination group:

Aquaculture : <http://www.animalgenome.org/aquaculture/>

Bioinformatics : <http://www.animalgenome.org/bioinfo/>

Cattle : <http://www.animalgenome.org/cattle/>

Horse: <http://www.uky.edu/Ag/Horsemap/>

Poultry: <http://poultry.mph.msu.edu/>

Sheep: <http://www.animalgenome.org/sheep/>

Swine: <http://www.animalgenome.org/pigs/>

Highlights of Accomplishments: AQUACULTURE SPECIES

- An integration of the genetic and chromosome maps of rainbow trout and Atlantic salmon was completed.
- A SNP-based linkage map was constructed for Shrimp.
- A medium density linkage map for striped bass (*Morone saxatilis*) based on 498 microsatellite markers was developed.
- Whole genome sequencing for Tilapia was initiated.
- The Oyster Genome Consortium was successful in establishing a whole genome sequencing project for the Pacific oyster (*Crassostrea gigas*), and whole genome sequencing of the Pacific oyster is in progress. (Coordinator's funds were used to supplement the arraying and distribution of BAC libraries.)
- Projects to identify ESTs and characterize the transcriptomes of various tissues were conducted in catfish, rainbow trout, brook trout and striped bass.
- A second generation integrated physical and genetic map for Rainbow trout was produced and is available through G-browser at the Animal Genome website of the NRSP-8 bioinformatics group. (Coordinator's funds were used to help construct BAC libraries.)
- A high density genetic map composed of approximately 5,000 single nucleotide polymorphism markers (SNPs) was produced for rainbow trout.
- A first SNP genetic map for Pacific white shrimp was built with 418 SNP markers mapped onto 45 sex-averaged linkage groups.
- A 1,772 SNP genetic map was produced for sockeye salmon.
- A deep sequencing of restriction-site associated DNA marker (RAD-seq) method was used to find genetic markers involved in disease resistance in Pacific white shrimp.
- For the Pacific oyster, several groups developed SNPs, BAC end sequences, a BAC physical map, and integrating these with genetic and cytogenetic maps.
- Deep sequencing of a doubled-haploid channel catfish transcriptome yielded 25,144 annotated contigs. The channel catfish genome assembly currently contains contigs produced from Illumina paired end libraries, and existing sequences from larger insert libraries (3 kb, 8 kb, 36 kb).
- A low density genetic map was produced for the striped bass based on 289 microsatellite DNA markers, and performance traits in the mapping populations contained ~68 QTL, many with very strong potential for predicting performance of growth and body composition.
- Next-generation sequencing of striped bass genomic DNA produced 14 Gb of sequence data to support initial assembly of the draft genome sequence. In addition, 5.4 Gb of sequence data for microRNAs were obtained for these striped bass.
- High throughput sequencing of mRNA (RNA-seq) of fast- and slow-growing hybrid striped bass revealed 1,076 genes that were differentially expressed in fast- versus slow-growing fish. This analysis also identified 270,000 single nucleotide polymorphisms (SNPs) with large numbers of SNPs being found only in fast growing or slow growing fish.
- Linkage maps for shrimp (SNP-based) and striped bass (microsatellite-based) were constructed.
- A DNA methylation-enriched Pacific oyster DNA library was produced which described functional roles of DNA methylation in oysters.
- Deep sequencing of a doubled-haploid channel catfish transcriptome yielded 25,144 annotated contigs.
- QTL mapping populations were established to study stress tolerance, salinity tolerance in salmonids, spawning times in females and maturation timing in males/females.
- University/commercial cooperation continued to develop inbred oyster lines to crossbreed F1 hybrids for use in the oyster farming industry, and F2 families useful for mapping QTL for survival, growth, and sex determination.
- Species-specific bioinformatics resources were developed to support efforts aimed at identifying genes of interest. Most efforts focused on database development, including development of pipelines for next-generation sequencing data processing, analyses and annotation, and in cooperation with NRSP8-supported bioinformatics capacity. (Coordinator's funds were used to provide partial support for bioinformatics training efforts.)

- The rainbow trout WebFPC BAC physical map was updated with genetic markers and BACs sequence data integrated onto the BAC contigs.
- A rainbow trout QTL database was placed on the NRSP8 Animal Genome website.
- A pipeline was developed for identification, characterization, and selection of oyster SNPs in a mixture of Sanger and next generation cDNA sequences.
- Catfish RNASeq, ESTs, and related SNP information was disseminated through the Catfish Genome Database, cBARBEL, <http://www.catfishgenome.org/cbarbel/>.
- The collection of over 11,000 high-quality, annotated, striped bass transcriptome sequences was deposited in the NCBI Short Read Archive (GenBank: SRX007394) and maintained for public access on the National Animal Genome Project website.
- The vast majority of these research accomplishments were achieved through competitive funding from USDA NRI and recently by grants from NIFA AFRI. The participation of students and postdoctoral fellows in the program was supported through travel grants. Other functions of coordinator's funds included facilitation of resource sharing, and training in bioinformatics.

Highlights of Accomplishments: BIOINFORMATICS

- The www.ANEXdb.org database was migrated into www.AnimalGenome.ORG for long-term maintenance, to expand its storage capabilities to all widely used livestock gene expression profiling tools, and to create comparative annotation of the sequence elements on these profiling tools. (Minimal support provided by coordination funds.)
- Relational database and user tools to store and disseminate phenotypic and genotypic data from large genomic projects in farm animals were developed to facilitate sharing of information and data analysis among members (<http://www.animalgenome.org/lunney/index.php>). The database model and tools can be re-used to support similar projects for research labs to share information among collaborators, and facilitate data analysis. (Mostly industry and federal grant support with modest contribution of coordination funds and server infrastructure.)
- In addition to information available in traditional browsers (UCSC, NCBI, Ensembl), GBrowse is utilized to display custom mapping data for each species. The custom mapping data is also made available through a public data repository created for the community to share data. This allowed members of the community to freely deposit and share various categories of genomic and functional information. (Coordination funds have been used to fund Dr. Zhi-Liang Hu's efforts.)
- The Bioinformatics coordination program expanded the capabilities of the portal www.AnimalGenome.ORG to integrate resources, tools, services, news and updates for each of the 6 species groups (web sites for aquaculture, cattle, pig, sheep/goat; Complimentary web information for horse and poultry). (Coordination funds have been used to fund Dr. Zhi-Liang Hu's efforts.)
- Additional functionalities were implemented to assist with standardization of trait terms across species, including the Vertebrate Trait Ontology (<http://www.animalgenome.org/bioinfo/projects/ato/>) for hierarchy visualization and Virtual Comparative Map (<http://bioneos.com/VCMMap/>), for multiple dataset comparisons. (Primarily funded by USDA-AFRI competitive funding with minimal support from coordinator funds.)
- Traits specific to livestock products were incorporated into a new Livestock Product Trait Ontology. As the first stage outcome, the cattle, pig, chicken, and sheep QTL traits were mapped to Vertebrate Trait Ontology (VT), Product Trait Ontology (PT) and Clinical Measurement Ontology (CMO) to help standardize the trait nomenclature used in the QTLdb. (Primarily funded by USDA-AFRI competitive funding with minimal support from coordinator funds.)
- Minimal standards for publication of QTL and gene association data were developed (<http://miqas.sourceforge.net/>) to facilitate standardized data portal development and meta data analysis. (Primarily accomplished by via international collaboration.)
- About 13,045 new QTL were curated into the Animal QTLdb since 2008 (include ~2,500+ chicken, 4,700+ cattle, 5,140+ pig, 630+ sheep, and 70+ rainbow trout QTL). (Primarily funded by USDA-AFRI competitive funding with modest support from coordinator funds.)
- The Animal QTL database (QTLdb) has been re-developed to accommodate a number of new functions: Inclusion of GWAS data, comparative display of cytogenetic bands, meta-QTL plots, addition of sheep and

rainbow trout in the QTLdb species families, etc. (Primarily funded by USDA-AFRI competitive funding with modest support from coordinator funds.)

- Coupled with QTLdb, GBrowse display of QTL made it possible to align the cattle, pig and chicken QTL against their respective genomes and aligned with the high density SNPs, and oligo arrays along with NCBI/and/or Ensembl annotated gene information. (Coordination funds have been used to fund Dr. Zhi-Liang Hu's efforts.)
- A local copy of Biomart software was installed on the AnimalGenome.ORG server to serve the cattle, chicken, pig, and horse communities. (Coordination funds have been used to fund Dr. Zhi-Liang Hu efforts.)
- The STS Cattle 770K high-density SNP and 4.1M dbSNP data were mapped and made available in both Gbrowse, to align with QTL, and in SNPlotz, for genome analysis. (Coordination funds have been used to fund Dr. Zhi-Liang Hu efforts.)
- A new pig genome database (PGD) is under active development. A number of tools developed has served the pig genome community in several pig genome annotation projects. (Coordination funds have been used to fund Dr. Zhi-Liang Hu efforts.)
- The web site and user forum listserv for improving the CRIMAP software has been maintained and is actively used. (Coordination funds have been used to fund Dr. Zhi-Liang Hu efforts.)
- Continued improvements to ANGENMAP listserv to ensure a trouble-free and effective communication channel. (Coordination funds have been used to fund Dr. Zhi-Liang Hu efforts.)
- NRSP-8 funds were used to support the development of BirdBase resources, such as the Chicken Gene Nomenclature Committee (CGNC) database, which is now linked with the NCBI Entrez Gene chicken gene pages. To date, we have approved nomenclature for 1,214 genes and automatically assigned nomenclature for 18,638 genes. (Primarily funded by other resources with modest contribution of coordination funds.)
- CGNC also provides manual biocuration targeting genes that are specific to chicken. We currently are working on assigning nomenclature for 212 egg genes (193 approved, 8 pending, 11 remaining) and 74 MHC genes (30 approved, 12 pending, 32 remaining). (Primarily funded by competitively funded resources with modest contribution of coordination funds.)
- Via AgBase, we were recruited by the Phenotype Ontology Research Coordination Network (NSF DBI 0956049) to develop an avian anatomy ontology to be integrated with other, existing ontologies to describe phenotypes. (Primarily funded AgBase.)
- The NRSP-8 Bioinformatics Online Tool Box was actively updated. Several major software upgrades were made to SNPlotz, Gene Ontology CateGORizer, BEAP, and the Expeditor. (Coordination funds have been used to fund Dr. Zhi-Liang Hu efforts.)
- The Virtual Comparative Map (VCMaP) tool passed its initial development stage and was transferred to AnimalGenome.ORG for more application development. (Primarily funded by USDA-AFRI competitive funding with modest support from coordinator funds.)
- Genome2Seq, an online tool that rapidly retrieves a FASTA file of sequences based on genome coordinates generated from RNA-Seq data, was developed. (Collaboratively developed with AgBase, coordinator funds have been used to fund Dr. Zhi-Liang Hu's efforts.)
- A bird comparative genome browser was developed via BirdBase that currently includes the chicken, turkey, and zebra finch genomes. (Funded by AgBase.)

Highlights of Accomplishments: CATTLE

- SNP markers in candidate genes associated with milk fat composition, milk and protein in yield were identified in chromosome BTA4. These markers were validated in association studies in large beef cattle populations in Argentina (Brangus) and Uruguay (Hereford), extending this work for graduate student training and research collaborations with the international community.
- 40 genes related to the GH/IGF signaling pathway were resequenced in 52 samples from *Bos taurus* (beef and dairy) and *Bos indicus* origin, allowing for selection of 250 SNPs for genotyping and marker-trait association studies.
- The transcriptome of milk somatic cells in Holstein cows was analyzed at early, peak and late lactation using RNA sequencing (RNAseq). This work was extended to examine the mammary transcriptome in relation to

oligosaccharide composition of milk, to annotate the bovine milk glycome and to study of the expression of key glycosylation genes in Holstein, Jersey and Brown Swiss cows at different stages of lactation. This is the first comprehensive study conducted to characterize the milk glycome in any mammalian species.

- RNASeq analyses of reproductive tissues were combined with GWAS to construct gene regulatory networks indicative of heifer puberty in *Bos indicus*-influenced heifers.
- SNP markers that can be used to predict genetic merit for twinning rate in the Holstein population were identified and validated.
- Two loci were identified that were linked with bovine respiratory disease (BRD) and found to be also associated with persistent infection of the BVD virus.
- Y-chromosome haplotypes (SNPs/indels) associated with fertility were identified that can improve marker assisted selection, eliminate potential genetic defects and reduce maintenance costs prior to breeding.
- A NimbleGen 12plex gene expression array for all known genes included in mtDNA was developed to evaluate the role of mitochondrial function and energy use in cattle with extreme differences in residual feed intake.
- An RH panel was developed, and the genetic map was refined, with focus in the MHC region, for river buffalo in collaboration with E. Amaral and Brazilian agencies.
- Animal resources, genomic, functional and proteomic data were developed for identification of genes and markers for growth, fertility, feed efficiency and meat tenderness in beef cattle.
- A database was developed to house DNA, genotypes, phenotypes and herd data for evaluation and assessment of different DNA-enabled approaches for predicting the genetic merit of herd sires on commercial beef ranches.
- A performance comparison was performed of the new high density SNP genotyping platforms, the Illumina High-Density Bovine BeadChip Array (777,962 SNP) and the Affymetrix Axiom Genome-Wide BOS 1 Array (648,874 SNP). The study demonstrated that both platforms are well designed and provide high quality genotypes and similar coverage of informative SNP, and that the BovineHD platform measured Copy Number Variation more efficiently.
- A collaborative project was initiated to develop a genotyped, phenotyped population to enable the evaluation and/or assessment of different DNA-enabled approaches for predicting the genetic merit of herd sires on commercial beef ranches.
- Efforts were taken to improve the bovine genome assembly included 30x genome coverage of Illumina short and long-insert sequence reads of Dominette. In addition, an optical mapping project was contracted with OpGen Inc (Maryland, USA) to generate a high-resolution, ordered, whole genome restriction maps from Dominette DNA. (Coordinator funds were used to support the sequencing and optical mapping efforts.)
- Phenotypes were collected from the Cycle 1 (F2 Nellore-Angus cows), Cycle 2 (reciprocal F2 steers and heifers) and Cycle 3 (F3 Nellore-Angus steers and heifers) McGregor Genomics populations to determine the genetic basis of variation in immunological response to vaccination for BVDV using steers from Cycle 2 and Cycle 3.
- A resource family was created to map the location of a major gene for ovulation rate.
- A whole genome association and fine mapping studies continue to correlate genotypes with susceptibility to infection by *Mycobacterium avium* subsp. *Paratuberculosis*, bovine viral diarrhea persistent infection and bovine respiratory disease.
- The bovine UMD 3.1 genome assembly was made available on the NRSP8 Bioinformatics site, as were 8.4 million SNP loci, data on the Illumina ~770K HD SNP chip, an updated QTLdb, and access to genomic data through Biomart.
- Development of the BovineSNP50 genotyping assay has been implemented for whole genome selection in dairy cattle and adopted by practically all dairy production and breeding industries around the world
- Several beef cattle breed associations are incorporating molecular breeding values into their multi-trait genetic evaluations. The molecular breeding values are a product of genotypes from BovineSNP50 and HD beadchips.
- A joint effort by NRSP8 investigators to produce new high density genotyping chip resources (600K+) in Illumina and Affimatrix platforms has the potential of truly refining QTL position for fine mapping and expanding the application of targeted genome selection to a large group of cattle breeds.
- High throughput computational resources, including the use of parallelized graphics processing units, were developed to solve advanced computational problems such as sophisticated models used to predict genetic merit from candidate broodstock.
- Nine SNP-chips have now entered the market for use in GWAS and genomic selection. These are the Illumina 3K, 7K, 9K, 50K, and 90K for *Bos taurus* and *Bos indicus*. Both Illumina and Affymetrix market an 800K chip. (All of these tools were developed from the sequence resources supported by NRSP8.)

- The majority of these research accomplishments were achieved through competitive funding from USDA NIFA AFRI grants and with funding from commodity stakeholders and universities. NRSP-8 coordinator funds have been used to develop resources to improve the bovine reference genome sequence and to provide conference travel grants to students and speakers at the Plant and Animal Genome meeting.

Highlights of Accomplishments: EQUINE

- The assembled Horse Whole Genome Sequence was published. (Coordination funds were used to support assembly and workshops in support of the activity.)
- RNA-seq data from 8 tissues were analyzed to improve annotation of equine protein-coding genes. (Industry support of NRSP8 member laboratory based on PAG interactions)
- In order to enable comprehensive studies of equine copy number variation (CNV), a whole genome tiling array was designed based on the EquCab2 assembly. (Resource developed as consequence of point 1.)
- One of the major tools for investigating the horse genome has been the Illumina SNP50 (~50,000) chip. A study using this tool to describe whole genome linkage disequilibrium across many breeds, inbreeding, inter-breed and intra-breed relationships, and proof of principle association analyses was published. (Coordination funds used to develop and share this tool.)
- Illumina discontinued the equine SNP50 chip in 2010 and a 74K SNP chip on the Illumina platform was designed and produced during 2011 based on collaborations between NRSP8 scientists and scientists at Geneseek, a division of Neogen, Inc. (Coordination funds used for workshops on this activity)
- Agilent oligonucleotide microarrays were produced and compared for investigation of gene expression in several laboratories. (Application by individual research laboratories was derivative of point 1.)
- At least 30 horses have been sequenced by independent research laboratories during 2011-2012 using next-generation DNA sequencing technology. Through the auspices of NRSP8, this information is being made publically available. Genome-wide association analyses of economically important traits were conducted using the SNP50 chip. These traits included osteochondrosis, swayback in American Saddlebred horses, dwarfism in Miniature horses, Neuroaxonal Dystrophy (NAD) in the American Quarter Horse, Chronic Progressive Lymphedema (CPL) in draft horses, Foal Immunodeficiency Syndrome (FIS) and Lavendar Foal Syndrome (LFS) in Arabian horses, Polysaccharide Storage Myopathy (PSSM) in Quarter Horses (and other breeds), Recurrent Exertional Rhabdomyolysis and racing performance in Thoroughbreds. Most of these studies identified chromosomal segments to interrogate further and several led to the discovery of the causative gene and development of diagnostic tests. (This application was made by individual laboratories and NRSP8 collaborators, was derivative of point 1 and SNP chip development and presented at PAG NRSP8 workshop)
- A polymorphism of the gene *MSTN* was found to influence racing performance, with one form associated with endurance and another with sprinting ability (Application by individual laboratories was a derivative of point 1 and SNP chip development, presented at PAG NRSP8 workshop.)
- MHC genetic structure, functions, and polymorphisms were refined. (Application by individual laboratories was a derivative of point 1 and presented at PAG NRSP8 workshop.)
- RNA sequencing of 43 diverse horse tissues is being used to elucidate the equine transcriptome and to identify candidate mutations for traits including Congenital Stationary Night blindness in the Appaloosa (Application by individual laboratories was a derivative of point 1 and presented at PAG NRSP8 workshop.)
- Gene expression studies are enabling understanding of chronobiology, exercise-induced stress and complex diseases such as laminitis, among others.
- The equine breed diversity consortium was developed to facilitate large-scale population genetic analyses in the domestic horse using whole genome SNP data. This is an international collaboration of scientists from 22 different institutions and represents genetic data from approximately 40 horse breeds. The main forum for this consortium is the NRSP8 workshop.
- Conferences and workshops facilitated collaborations in which data were shared between laboratories to investigate diverse hereditary conditions, including investigations of developmental bone diseases, respiratory

disease, stable vices, immunology and population analyses. (Financial assistance for these meetings was provided from coordination funds.)

- Data from testing animals with the SNP50 chip or with RNA-seq technologies have been shared for use as control samples or in collaborations through established community databases based on NRSP8 workshops.
- A special issue of the journal *Animal Genetics* (volume 43, supplement 2) was published and dedicated to horse genomics as a consequence of workshop activities and collaborations. (Financial support of the special issue came from industry and from coordination funds.)
- Bioinformatic development is a major limitation to the advance of horse genomics. Work is underway to coordinate with NRSP8 bioinformatics capacity and Agbase.
- A 685K SNP chip is being designed on an Affymetrix platform for use in 2013. SNP selection and design is being made by a subcommittee from NRSP8. (Coordination funds are being used to evaluate 2M SNPs for possible inclusion on the new tool.)

Highlights of Accomplishments: POULTRY

- Chicken linkage mapping transitioned to being almost solely based on high-throughput SNP assays (see below).
- A 60K SNP Illumina iSelect genotyping array was developed and used to enhance the SNP linkage map of the Wageningen reference panel. (Coordination funds were committed to SNP chip development and distribution to member labs, along with research grant support.)
- SNP genotyping was used to catalog the genetic diversity of commercial breeds, worldwide, and to demonstrate the effects of long term selection on the diversity of chickens. Virtual sharing of populations involved the collection of DNA samples from several populations from Federal, state and industry locations and the coordination of a common high density genotyping effort. Relates to bullet 2 above.
- SNP arrays have been used by a large academic-industry consortium to test and implement genome-wide marker-assisted selection (GMAS) in commercial chickens. (Funded by grant and private sources.)
- Very high density SNP mapping (ca. 600,000 SNP) panels recently have been developed and are being employed in genome-wide association studies and GMAS. (Coordination funds this year supplying 600K SNP chip genotyping to several members through a contract with DNA Landmarks.)
- A regularly updated, detailed comparative chicken-turkey BAC contig-based comparative map was made available at <http://poultry.mph.msu.edu/resources/resources.htm#TurkeyBACChicken>. (Mostly grant funded with small contribution from coordination support.)
- A first draft sequence of the turkey genome using next-generation sequencing (NGS) was produced by the Turkey Genome Sequencing Consortium. (Coordination funding made a modest contribution along with academic and industry support. Draft sequence eventually led to \$1M AFRI grant for further development.)
- Coordination funds contributed to transcriptomics studies and supported the development of new Agilent arrays and high throughput RNA sequencing by NGS.
- A new build, Galgal4.0, of the chicken genome sequence was released which combined traditional Sanger sequence with NGS. The Z chromosome sequence was published at near-finished quality. (Mostly NIH funded.)
- A number of additional chicken genomes have been or are being re-sequenced with NGS technology (see below for that part of the effort supported by coordination funds).
- Coordination funds supported a project with DNA Landmarks to sequence 20 different chicken lines of interest. NGS data for genomes from the DF1 and DT40 chicken cell lines have also been obtained and are currently being analyzed and compared to the Galgal4.0 chicken genome assembly.
- A homepage has been maintained for NRSP-8 U.S. Poultry Genome project (<http://poultry.mph.msu.edu>) that provides a variety of genome mapping resources, including our newsletter archive. The Poultry Genome Newsletter is published quarterly and is distributed through our Homepage and on the angenmap email discussion group. (All done by coordination support.)
- Coordinators are currently working on a project to fill gaps and missing microchromosome sequence in Galgal4.0 and to provide improved annotation and a contact point for user group questions and suggestions for improvement. Industry funds have already been committed, along with a modest amount of coordinator funds and hopes for AFRI grant support.

Highlights of Accomplishments: SHEEP and GOAT

- A high resolution ovine RH map was constructed for each ovine chromosome using the USUoRH5000 panel, and then combined into a whole genome RH map comprised of 39,856 SNPs. Markers included sequences that cover autosomal chromosome ends and the regions flanking interspecies chromosomal breakpoints, as well as SNPs from a sheep 1.5K SNP pilot chip. Coordinator funds were used to support part of the technical help needed to generate the RH map, as well as the synthesis of marker primers used to amplify the RH panel.
- Various sheep maps (RH, linkage, cytogenetic) and a virtual sheep genome assembly were integrated and then used to improve the assembly of the ovine whole genome reference sequence. Dr. Jill Maddox, University of Melbourne, took the lead on integration of the ovine maps. (Her work on the ovine genome assembly was supported by coordinator funds.)
- A high-density ovine SNP array (Illumina Ovine SNP50 BeadChip) was developed by the International Sheep Genome Consortium (ISGC). Over 11,000 sheep have been genotyped with the chip to date. (Coordinator funds contributed to the identification of SNPs used on the chip and have also contributed to the genotyping of animals in projects of NRSP-8 committee members who are investigating traits of economic importance.)
- The Ovine SNP50 BeadChip was also typed on the USUo5000RH panel and the International Mapping Flock. (Coordinator funds contributed to the typing of these samples; the generated data were incorporated back into the ovine RH and genetic maps, respectively.)
- Version 3.0 of the ovine whole-genome reference sequence (Oar v3.0) was publicly released in September, 2012 and contains 2.71 Gb of sequence with an N50 of 1.08 Mb. It covers 93.1% of the genome, with 2.57 Gb placed onto chromosomes. (NRSP-8 coordinator funds have facilitated collaboration of the international participants on this project by supporting travel to the NRSP-8 annual meeting and also partially supporting the researchers working on the assembly and annotation of the sequence.)
- The world-wide ovine HapMap project includes 3064 sheep from 74 breeds and strains, seven species of wild sheep and nine outgroup species. These HapMap samples were typed with the Ovine SNP50 BeadChip. Samples from two U.S. breeds (Rambouillet and Gulf Coast Natives) were collected and genotyped using NRSP-8 coordinator funds.
- The genomic resources developed with NRSP-8 have been applied to research projects conducted by NRSP-8 scientists. Examples of projects that used these genomics resources include:
 - The identification of significant QTL for aseasonal reproduction and milk production traits in a resource Dorset population selected for aseasonality and prolificacy (Cornell University).
 - Mapping of genomic regions on OAR9 and OAR19 associated with *Haemonchus contortus* parasites using a sheep population segregating for parasite resistance/susceptibility (Louisiana State University, Utah State University).
 - Gene expression analyses used to study muscle growth in callipyge lambs and to identify pathways involved in the muscle hypertrophy phenotype (Purdue University, Utah State University).
 - Analysis of feed efficiency phenotypes in sheep using the Ovine SNP50 Beadchip.
- A 5,000 rad RH panel (90 clones) was created using DNA from a male Boer goat. This panel was used to develop a low-resolution RH map with XX loci on all XX chromosomes. (Grant funding has supported the construction of the panel and the RH map.)
- A predicted order of goat ESTs with 3190 elements was developed using *in situ* comparisons of the goat genetic map with the sheep genetic maps and the bovine and ovine reference sequences. In this project, genome conservation analysis of ESTs was performed relative to sheep, cow, human, mouse and rat. (Grant funding to NRSP-8 members has supported the project.)
- Two projects have been initiated by the International Goat Genome Consortium that will result in the goat genome reference sequence. The public project was undertaken by the research group at BGI in China as part of the BGI “1000 genomes” project. The initial sequencing has been completed on Illumina GX sequencers, and the sequences are currently in the assembly process. NRSP-8 members have contributed to this project by providing virtual and RH maps to the caprine genome assembly team.

Highlights of Accomplishments: SWINE

- The swine genome sequence assembly was improved using high-resolution radiation hybrid mapping and pig-human comparative mapping, as well as transcriptome sequencing. (Coordination funds were committed to develop bioinformatics tools that used these resources.)
- A QTL database, Pig QTLDB, continues to be expanded as a part of the AnimalQTLDB. An open-source transcript profiling database and website that allows the user to store and submit Affymetrix profiling data to NCBI-GEO (www.anexdb.org) was refined. (Genome coordination funds were used to support development of this resource.)
- The Pig Genome Database (PGD) is under development and integrates the functions of the Pig QTLdb, GBrowse, Biomart, ANEXdb, Vcmap, and SNPloz to provide a research database tool for the community. (Coordination funds were used to support bioinformatics tool development and sharing of information to the pig genome community.)
- A high density SNP genotyping platform has been developed and employed (i.e. 60K Illumina Porcine SNP60 BeadChip; Ramos et al., 2009. PLoS One 4:e6524). The Porcine SNP60 BeadChip has been used for several genome-wide association studies (GWAS) to understand the genetic control of important pig traits, such as reproduction, structural soundness, residual feed intake, disease resistance, and the genetic control of resistance to PRRS virus infection and related growth effects. (Coordination funds were committed to SNP chip development and distribution to member labs, along with research grant support.)
- The PRRS Host Genetics Consortium (PHGC) has been developed to determine the role of host genetics in resistance to PRRS and in effects on pig health and related growth effects. Genome wide association studies (GWAS) using the PorcineSNP60 Genotyping BeadChip have identified chromosomal regions associated with PRRS resistance and/or improved weight gain, with an area on SSC4 correlated with both lower viral load and higher weight gain.
- Transcriptional profiling of differentially expressed genes in longissimus dorsi muscle of Piau and Yorkshire x Landrace pigs was performed at 40 and 70 d of gestation (encompassing the transition from primary to secondary fibers). (Coordination funds were committed to synthesize the oligonucleotides for development of the long oligonucleotide microarray.)
- Scientists estimated linkage disequilibrium in four US breeds of pigs (Duroc, Hampshire, Landrace, and Yorkshire) and subsequently calculated persistence of phase among them using the Porcine SNP60 BeadChip. (Coordination funds were committed to help provide SNP genotyping support.)
- Nomenclature for the swine Major Histocompatibility Complex (MHC), the swine leukocyte antigen (SLA) complex, was updated. (Genome coordination funds were used to support the database work related to this activity.)
- Computational curation tools (Otterlace suite of programs) from the Sanger Institute were used to refine the currently available automated annotation of the pig genome. (Coordination funds were used to support bioinformatics tool development and sharing of information to the pig genome community.)
- The porcine reference genome has been completed (build 10.2 see links below). The sequence is based on Sanger sequencing of BAC clones, complemented with additional NGS WGS data. The draft sequence was published in the November 15, 2012 issue of Nature (with cover and several companion papers have also been or will be published. (Coordination funds were used in part to support the sequencing and research efforts.)
- NGS has allowed for sequencing the genomes of multiple pig breeds. Several large scale sequencing projects on individual pigs and pools have already resulted in individual genome sequences for around 200 individuals from a large variety of breeds. NGS data has resulted in the identification of over 25 million SNPs and thousands of CNVs. (Coordination funds were used in part to support the sequencing and research efforts.)
- Genome sequencing of other Suids (*Phacochoerus Africanus*, *Sus barbatus*, *Sus celebensis*, *Sus cebifrons*, *Sus verrucosus*) has provided insight into genome evolution and domestication of the pig and its close relatives. (Coordination funds supported development of the SNP chip used in this process.)
- Commercial breeding companies have begun to test and employ the Porcine SNP60 BeadChip for genomic selection in their breeding protocols. A new 8K chip for imputation has been developed. Coordination funds have helped to support testing and development of this smaller chip. Microarrays and, more recently, NGS technologies have generated a wealth of transcriptomic data for various porcine tissues (embryo's and adults).

Many of the RNAseq experiments were conducted on tissues derived from clones of TJTabasco (used for the reference sequence). The data is currently being used for the gene annotation efforts by Ensembl. (Coordination funds have supported these efforts.)

- Manual annotation of immune genes has been completed on a large number of genes. (Coordination funds have supported these efforts.)
- The HapMap consortium has characterized over 3000 individual pigs from across the world including commercial breeds, wild boars, feral pigs and museum specimen. (Illumina 60K chip + mtDNA). (Coordination funds have supported these efforts.)