ANNUAL PROGRESS REPORT
NATIONAL RESEARCH SUPPORT PROJECT - NRSP-8
Year Ending 2004
Preliminary Information-Not for Publication

Washington Agriculture Experiment Station
Pullman, WA 99164-6351

I. PROJECT TITLE:

National Animal Genome Research Program (NAGRP)

II. COOPERATING AGENCIES AND PRINCIPAL LEADERS:

A. Agencies and Departments Cooperating: **Washington Agriculture Experiment Station and Animal Sciences Department, Washington State University**

B. Leader of the Project: **Zhihua Jiang**

C. Cooperating Investigators: Raymond W. Wright Jr., Xiao-Lin Wu and Jennifer J. Michal, Washington State University; Bradley A. Freking, Gary A. Rohrer and Thomas H. Wise, USDA, MARC, Clay Center; and Peter Dovc and Tanja Kunej, University of Ljubljana, Slovenia.

III. NATURE OF WORK AND PRINCIPAL RESULTS OF YEAR:

Objective: 1. Enhance and integrate genetic and physical maps of agriculturally important animals for cross species comparisons and sequence annotation.

Experiment #1: **Comparative interval mapping of SSC3 to human, dog, mouse and rat genomes.**
A comprehensive RH map of SSC3 was constructed with a total of 116 genes/markers, including 16 that have been placed on the cytogenetic map and 21 on the linkage map. The retention frequency of these 116 genes/markers varied from 8.5 to 53.3% with an average of 27.62%. Overall, SSC3 had a total of 11 conserved segments paired with human, 13 with dog, 17 with rat and 22 with mouse, respectively. Alignment of ~192 Mb of orthologous regions in these five species led to the identification of provisional conserved ancestral blocks (CABs) and the characterization of breakpoint regions, which provides an alternative for further determination of the evolutionary makeup of mammalian genomes.

Experiment #2: **HAPPY mapping of the porcine genome on microarrays.**
Common drawbacks for the current mapping approaches, such as linkage mapping, RH mapping and HAPPY mapping are their cost and time consumption. Microarray technology has allowed researchers to develop cost-effective and revolutionary types of molecular analysis for the detection or quantitation of multiple genes in a single sample. Evidence has shown that comparative genomic hybridization (CGH) is a powerful tool to provide high-resolution views of the physical genome in different species. About 100 pig HAPPY lines were probed on a 19K human cDNA array to detect the presence and absence of genes in each HAPPY line. Microarray data are being collected and analysis is under way.

Objective: 2. Facilitate integration of genomic, transcriptional, proteomic and metabolic approaches toward better understanding of biological mechanisms underlying economically important traits.

Experiment #3: **Selection-based mapping of quantitative trait loci for ovulation rate and uterine capacity in pigs.**

Long-term selection, either within breed or within-line, has served as a main vehicle for genetic improvement of livestock, which aims for improving a breed or line as a source of superior germplasm for commercial production. Research conducted at the USDA Meat Animal Research Center, Clay Center has demonstrated that eleven generations of selection for superior ovulation rate and uterine capacity effectively improved these two traits in pigs. In comparison to the control animals, ovulation rate increased by more than three eggs and uterine capacity by more than one piglet at generation 11. AFLP analysis was conducted on 12-pooled DNA samples to screen QTL linked markers for these two traits. So far, two significant markers for ovulation rate and one for uterine capacity have been identified. Isolation and characterization of these AFLP products are under way.

Experiment 4: **Identification of candidate genes involved in energy metabolism in pigs.**

Energy metabolism is a complex process determined by the action of several endogenous and exogenous factors. Difference between energy intake and energy dissipation is reflected in changes of body composition. In eukaryotes the mitochondrial function and copy-number can be increased in response to external stimuli. The important player in this process is PPARgamma coactivator 1 (*PPARGC1*), which stimulates mitochondrial biogenesis and respiration in muscle cells. *PPARGC1* promotes transcription and co-activates transcription of mitochondrial transcription factor A (*TFAM*). The full length cDNA sequence of the pig TFAM gene was cloned and sequenced. A SNP was developed and genotyped on ~250 animals from 12 pig breeds, including 7 Chinese and 5 European/American pig breeds. Gene expression and mapping are under way.

Objective: 3. Facilitate and implement bioinformatic tools to extract, analyze, store and disseminate information. (See Attachment 1 for more details on objectives.)

Experiment #5: **Genome transcriptomes and tissue/organ transcriptomes in livestock species.**

The number of protein coding genes in the finished human genome sequence has decreased from ~100,000 estimated a decade ago to only 20,000 – 25,000. We have developed a comparative gene-based approach to census orthologous gene sequences in domestic animals by annotation of the ESTs (expressed sequence tags) deposited in the public databases. Like in humans, the census estimated the number of protein-coding genes to be 22,817 in swine, 20,898 in cattle and 20,005 in chicken. By fitting a quadratic equation of the trend between the number of genes and the number of ESTs derived from different cDNA libraries, we estimated that a single tissue might express 9,200 genes in mammals, while 12,000 in birds. The new estimate on the number of genes in the finished human genome sequences and our estimates in different domestic animals may lead to a conclusion in the number of protein-coding genes that contribute to make-up of mammals and birds. However, the number of genes expressed in a single tissue remains open for further verification and discussion.

Experiment #6: **Collection and generation of full-length cDNA sequences of orthologous genes in livestock species.**

Animal genomics research has been undergoing a rapid development during the past decade. To date, an enormous amount of genomics information has been produced and continues to be produced as a result of extensive animal genome community efforts. We have decided to take three steps to collect and generate full-length cDNA sequences of orthologous genes in livestock species, i.e., puzzle sorting, puzzle retrieving and puzzle making for a final puzzle solving. We have developed a bioinformatics tool, **ELF-Walking** (electronic flanking walking) to facilitate large-scale *in silico* cloning of full-length cDNA sequences by mining the sequence databases. Using 21,775 human coding genes as references, we were able to generate unique full-length cDNA sequences of 3,881 genes and partial cDNA sequences of 10,358 genes in pigs, and unique full-length cDNA sequences of 4,308 genes and partial cDNA sequences of 10,785 genes in cattle. All sequence data and annotation information can be downloaded from our Bioinformatics website at http://www.ansci.wsu.edu/programs/bioinformatics/. All of these bioinformatics tools and reagents will contribute to the development of a Livestock Orthologous Gene (LOG) database.

IV. APPLICATION OF FINDINGS:

Our work has been focused on developing a program in "Comparative Genome Biology" by targeting orthologous gene sequences, map locations, expressions and functions with a database development. This program would provide the community with cutting edge tools and sophisticated reagents to speed up genome mapping and QTL mapping in farm animals. Our program would also promise to greatly benefit the livestock industry by providing knowledge and technologies that can help optimize production, quality, nutritional value and resistance to diseases.

V. WORK PLANNED FOR NEXT YEAR:

1. Characterization of QTL lined AFLP markers for ovulation and uterine capacity in pigs
2. HAPPY mapping of the porcine genome on microarrays
3. Generation of full-length cDNA sequences in animals
4. Development of Livestock Orthologous Gene (LOG) database

VI. PUBLICATIONS:

Jiang Z, Wu X-L, Garcia MD, Griffin KB, Michal JJ, Ott TL, Charley T. Gaskins CT, Raymond W. Wright Jr. 2004. Comparative Gene-based *In Silico* Transcriptome Analysis of Different Tissues/Organs in Cattle. Genome 47:1164-1172.

Cao H, Robinson JA, **Jiang Z**, Melville JS, Golovan SP, Jones MW and Verrinder Gibbins AM. 2004. A high-resolution radiation hybrid map of porcine chromosome 6. Anim. Genet. 35:367-378.

Wu X-L, Griffin KB, Garcia MD, Michal JJ, Xiao Q-J, Wright Jr. RW, **Jiang Z**. 2004. Census of orthologous genes and self-organizing maps (SOM) of biologically relevant transcriptional patterns in chickens (Gallus gallus). Gene 240:213-225.

Kunej T, Wu X-L, Milosevic Berlic T, Michal JJ, **Jiang Z**, Dovc P. 2005. Frequency distribution of a Cys430Ser polymorphism in peroxisome proliferator-activated receptor-gamma coactivator-1 (PPARGC1) gene between Chinese and Western pig breeds. J. Anim. Breed. Genet. (**in press)**.