ANNUAL PROGRESS REPORT
NATIONAL RESEARCH SUPPORT PROJECT – NRSP-8
Year Ending 2016
Preliminary Information-Not for Publication

Submitted by Christopher K. Tuggle and Max F. Rothschild
Iowa State University
January 14, 2017

I. Project: NRSP-8: Swine Genome Committee

II. Cooperating Agencies and Principal Investigators

A. Agencies and Departments Cooperating: Iowa Agriculture Experiment Station, Department of Animal Science, Iowa State University

B. Co-Leaders of the Project: Christopher K. Tuggle and Max F. Rothschild

C. Cooperating Investigators: Jack Dekkers, Elizabeth Huff-Lonergan, Steven Lonergan, Zhiliang Hu, Ken Stalder, Nick Gabbler, John Patience, James Reecy, Francesca Bertolini, Iowa State University; Many International Swine Genome Sequence committee collaborators.

III. Objectives

Objective 1: Advance the status of reference genomes for all species, including basic annotation of worldwide genetic variation, by broad sequencing among different lines and breeds of animals.

Objective 2: Develop strategies to identify and exploit genes and allelic variation that contribute to economically relevant phenotypes and traits, in part through improving functional annotation of the genomes of our species.

Objective 3: Facilitate analysis, curation, storage, distribution and application of the enormous datasets now being generated by next-generation sequencing and related "omics" technologies with regard to animal species of agricultural interest.

IV. General Project Plan

A. Use SNP chip to identify and better understand genetic control of traits (Objective 1).

B. Participate in on-going pig genome functional annotation efforts (Objective 2).

C. Development and analysis of populations measured for response to PRRSV (Objective 2)

D. Development and analysis of populations selected for increased feed efficiency (Objective 2)

E. Participate in further database and other bioinformatic resource development (Objective 3)

V. Work Progress

A. Use SNP chip and share results to better understand genetic control of traits (Objective 1).

Porcine Periweaning Failure to Thrive Syndrome (PFTS) affects young pigs around 2-3 weeks post-weaning and it is characterized by anorexia, lethargy and progressive debilitation that frequently leads to death. A genetic basis of this syndrome has been recently hypothesized but until now only one GWAS study have been reported, with a moderate suggestive association with the disease. In this experiment, 107 pigs were collected in North America within 2-3 weeks after weaning from commercial farms. A total of 70 met the PFTS case definition and were euthanized (cases) while 37 were aged-matched penmates (control). Cases and controls were balanced across the farms using a 2:1 ratio. All animals were genotyped with the 80K porcine SNP chip and 53,810 filtered autosomal SNPs were considered for the analyses. Population structure analysis was performed, showing that the samples have a similar genetic background even between cases and controls. Then, three analysis were performed comparing cases and controls. These included Fst analysis, considering single SNPs and the average Fst calculated in 500Kb overlapping windows and logistic regression analysis. Only moderate suggestive associations were found. A total of four regions not identified previously on chromosome 1, 3 and 11 were concordant for at least two types of analyses and contain some genes that are involved in energy homeostasis, caloric intake and growth and mature size. With this limited analysis we could not confirm major regions of difference for PFTS but we identified the presence of novel genomic regions that may be moderately associated with this syndrome.

B. Participate in on-going pig genome functional annotation efforts (Objective 2).

Work with colleagues to continue sequencing regions of interest and examine them for traits of interest in the pig, as well as annotate genes of interest, especially those involved in the immune system. Initiate collaborative work on functional annotation of the pig genome, especially related to ENCODE-type projects.

We have completed a project to assemble a deeply comprehensive peripheral blood transcriptome, so that all or nearly all isoforms of gene transcripts are catalogued. Using both de novo assembly and genome-guided approaches, over six billion Illumina RNAseq reads from five independent datasets were assembled into putative transcripts (PT). After filtering for low-quality assemblies, a final merged set of 132,928 PTs resulted. The exon-intron junctions of these PTs were validated using novel PacBio Isoseq data from other tissues, in collaboration with Tim Smith (USDA-MARC). Completeness of the 5' termini of 37,569 PTs was validated by public CAGE data.

We have initiated a project to annotate genes expressed in circulating neutrophils, as a complementary project to the Fr-AgENCODE project in France, which is investigating T lymphocytes. Our goal was to test whether human neutrophil-specific expression data could be used to identify porcine neutrophil-specifically expressed genes, as there is a lack in the literature for such expression patterns in porcine immune cells. We also wanted to test whether genes whose expression level was correlated with neutrophil numbers (as measured by standard Complete Blood Count assays, were specifically expressed in such cells. Quantitative PCR experiments using the Fluidigm system verified many of these genes to be 2-10 fold

higher in expression in isolated neutrophils compared to peripheral blood mononuclear (PBMC) preparations.

C. Development and analysis of populations measured for response to PRRSV (Objective 2)

Work with ISU and international collaborators to further analyze data resulting from PRRS population and trait measurement. As part of PRRSv Host Genetics Consortium, cohorts of 200 healthy crossbred pigs were procured from commercial farms and they were challenged with PRRS. Over 3000 have been challenged with PRRS. A number of traits were measured and the animals were sacrificed. All these animals were genotyped with the Porcine SNP60 SNPchip. Growth traits were also recorded in these animals and all the data have been entered into the PHGC relational database. New work has included analysis of blood RNA expression data from PRRSv-infected pigs to find QTN and gene pathways controlling variation in response to PRRSv. The latter work revealed pathway differences between pigs with alternate genotypes for a major host response QTL, and these results have now been published (Schroyen et al. 2016).

D. Development and analysis of populations selected for increased feed efficiency (Objective 2)

A Yorkshire population selected for RFI was maintained at ISU and is managed by Jack Dekkers. Both genotyping as well as substantial numbers of phenotypes (physiological and molecular) has been collected and is now being analyzed.

Further analysis of a novel mutation causing Severe Combined Immune Deficiency (SCID) in the RFI lines has identified that two mutations exist in the same gene (DCLRE1C) that cause this phenotype. In addition, the NK cells which are the only remaining lymphocytes in the affected pigs are apparently normal, as they can kill human cells in vitro with appropriate cytokine stimulation. This observation has now been extended to proteins expressed in these cells such as perforin, further documenting the normal intrinsic function of NK cells in SCID pigs. This work has now been published (Powell, et al. 2016).

E. Participate in further database and other bioinformatic resource development (Objective 3)

In conjunction with the Bioinformatics coordination effort we are developing a Pig Genome Database (PGD). The PGD has integrated the functions of the Pig QTLdb, GBrowse, Biomart, ANEXdb, VCmap, and SNPlotz to provide a research database tool for the community. As the pig genome research is quickly evolving, the purpose of the PGD is to make it a supporting platform with continued updates of data and tools. Our emphasis is on the integration of data from not only current but also historical studies to make it a central hub, and add tools for interactive information mining (URL: http://www.animalgenome.org/Pig/Genome/DB/).

We completed migration of an open-source database devoted to storage, analysis and dissemination of porcine transcriptional profiling data (http://www.ANEXdb.org) and hosted at AnimalGenome.ORG. Expansion of this database and website to all livestock species continues.

Development and use of a PRRS database (PHGCdb) is continuing in order to find genes and pathways linked to phenotypes recorded in the database.

We initiated hosting of a website (www.faang.org) and WIKI for use by members of FAANG (Functional Annotation of ANimal Genomes). The intent of this website is to popularize this new collaborative project to improve the annotation of all livestock genomes, and to serve as a collection space for protocols, member information, and communication to the public. For data specific to pig, this will be linked to the Pig Genome Database to facilitate analysis and dissemination of additional porcine genome data.

We are continuing to develop a database to contain all data and communicate via Web pages the results of the USDA-funded project entitled "Enhancing sustainability and competitiveness of the US pork industry by improving nutrient utilization and feed efficiency through innovative scientific and extension approaches " in 2011. A summary front page of this database is at: http://rfidb.swinefeedefficiency.com and contains summary statistics of this project; i.e., there are 18, 012 animals in the selection project, and over 1,900 of these pigs have SNP60 genotypes. The website that covers this project as well as allied projects in swine feed efficiency is titled: National Program of Swine Feed Efficiency, and is at www.swinefeedefficiency.com.

VI. Additions to the Project

None.

VII. Applications of Findings

A. Additional sequence information and annotation of this sequence in the pig will help inform agricultural and biomedical researchers for future research.

B. A large number of SNPs associated with traits of interest have been identified and mapped by ISU researchers. An emphasis has been made to confirm these results.

C. Several new genes that may be important QTL have been or are being mapped. These include genes associated with growth, feed efficiency, IMF and other performance traits. Of high importance is the documentation of gene variant(s) that appear to be genetically controlling a significant component of the variation in response to PRRSV, and the new diseases PEDV and now PFTS.

D. Several gene clusters were associated with weight gain or viral load in pigs extreme for these PRRSV phenotypes. These gene groups explain phenotypic differences between these extremes, and may be useful clues as to genes genetically controlling these traits.

E. Manual annotation of immune related genes in the existing genome build has identified many new transcripts as well as artifactual duplications. Transcriptomic analysis of gene expression patterns during infections has provided functional data extending known immune genes in the pig.

F. The creation of a comprehensive peripheral blood transcriptome will provide excellent annotation tools for researchers interested in studying blood as a source of biomarkers for quantitative traits.

G. The results on exploring neutrophil gene expression supports human cell-type specific expression data combined with correlation of gene expression to cell-type proportion can be a tool to predict similar expression patterns in swine.

VIII. Future Project Plans

Objective 1: Advance the status of reference genomes for all species, including basic annotation of worldwide genetic variation, by broad sequencing among different lines and breeds of animals.

Experiment A. Possibly collect additional samples from PED virus "susceptible" and "resistant" animals and use SNP chip to determine genomic regions and gene responsible for response to PEDV. Share results.

Experiment B. Possibly collect additional samples from PFTS "susceptible" and "resistant" animals and use SNP chip to determine genomic regions and gene responsible for response to PFTS. Share results.

Objective 2: Develop strategies to identify and exploit genes and allelic variation that contribute to economically relevant phenotypes and traits, in part through improving functional annotation of the genomes of our species

Experiment C. Add to growing effort to create and analyze functional annotation of the pig genome.

Experiment D. Validate blood RNA biomarkers associated with selection for RFI. Complete work on differences in RNA and protein responses to inflammatory and other stressors in RFI lines for blood and peripheral tissues.

Experiment E. Complete study on neutrophil gene annotation through RNA-seq analysis of isolated cell samples.

Objective 3: Facilitate analysis, curation, storage, distribution and application of the enormous datasets now being generated by next-generation sequencing and related "omics" technologies with regard to animal species of agricultural interest.

Experiment F. Continue database development and improvement of PGD, FAANG and other general applicability databases. Continue development and use of PRRS resistance and Feed Efficiency databases.

VIII. Publications

A. Publications during the year

Abell, C.E., R.L. Fernando, T.V. Serenius, M.F. Rothschild, K.A. Gray, and K.J. Stalder. 2016. Genetic relationship between purebred and crossbred sow longevity. J. Animal Sci and Biotech. **7**:51. DOI: 10.1186/s40104-016-0112-x

Bertolini F, J. Harding, B. Mote, G.S. Plastow and Rothschild MF. 2016. Genomic Differences between Preweaning Survival and Mortality of Piglets Following PEDV Outbreaks. Animal Industry Report

Bertolini F., J. Harding, B. Mote, G.S. Plastow, and M.F. Rothschild. 2016. Genomic differences between pre-weaning survival and mortality of piglets following PEDV outbreaks. Plant & Animal Genome XXIV, San Diego, California.

Groenen, M.A.M., A. Archibald, and C.K. Tuggle. Communications as an important Component of the Functional Annotation of Animal Genomes. Proceedings of Plant and Animal Genome XXIV. San Diego, CA, USA. Abstract # P0097.

Grubbs, J., J.C.M. Dekkers, E. Huff-Lonergan, C. K. Tuggle, S. Lonergan. 2016. Identification of serum biomarkers to predict feed efficiency in young pigs. J. Animal Sci. 94:1482-92.

Kern, C., P.J. Ross, P. Saelao, Y. Wang, M.M. Halstead, J. L. Chitwood, I. Korf, M. Delany, J.F. Medrano, H. Cheng, A. Van Eenennaam, C.K. Tuggle, C.W. Ernst, and H. Zhou. Identification of Tissue-Specific Long Non-Coding RNAs in Three Livestock Species. Proceedings of Plant and Animal Genome XXIV. San Diego, CA, USA. Abstract #P0675

Liu, H., N. T. Yet, D. Nettleton, J. C. M. Dekkers, and C. K. Tuggle. Post-Weaning Blood Transcriptomic Differences Between Yorkshire Pigs Divergently Selected for Residual Feed Intake. Proceedings of Plant and Animal Genome XXIV. San Diego, CA, USA. Abstract #W915.

Liu, H., N. Yet, D. Nettleton, J.C.M. Dekkers, C.K. Tuggle. Post-weaning blood transcriptomic differences between Yorkshire pigs with divergent residual feed intake phenotypes. 2016. BMC Genomics, 17:73.

Lunney, J.K., I. Choi, H. Bao, A. Kommadath, L.L. Guan, G. S. Plastow, R. R. R. Rowland, S. M. Abrams, J. M. Reecy, E. Fritz-Waters, C.K. Tuggle, J.C.M. Dekkers, and P. Stothard. Probing Porcine Reproductive and Respiratory Syndrome Virus (PRRSV) Infection Control Mechanisms using Differential Gene Expression. Proceedings of Plant and Animal Genome XXIV. San Diego, CA, USA. Abstract #W908.

Outhouse, A. C., J. Grubbs, C. Tuggle, N. Gabler, A. Rakhshandeh, and S. M. Lonergan. 2016. Immune system stimulation by repeated lipopolysaccharide injections alters longissimus dorsi sarcoplasmic protein profile in pigs. Meat Science 112.178. doi:10.1016/j.meatsci.2015.08.173

Powell, E.J., J. Cunnick, S. Knetter, E. Waide, J.C.M. Dekkers, C.K. Tuggle. 2016. NK cells are intrinsically functional in pigs with Severe Combined Immune Deficiency (SCID) caused by natural mutations in the Artemis gene. Veterinary Immunology and Immunopathology 175:1-6.

Rothschild M. F. and G. Plastow. 2016. Applications of Technology for Small Holders´ Livestock to Meet Global Food Production. Ceiba 54: 3-13.

Rothschild M.F. 2016. Applications of genomics to address adaptation of livestock to stressful environments to prevent food insecurity in the developing world. Proceedings of the 35th International Society for Animal Genetics Conference, 23-27 July 2016, Salt Lake City, Utah, USA.

Rothschild M.F. 2016. This little piggy went to market: applications of genomics to improved health for sustainable pig production. Proceedings of the EAAP meetings in Belfast Ireland.

Schroyen, M., C. Eisley, J.E. Koltes, E. Fritz-Waters, I. Choi, G.S. Plastow, L. Guan, P. Stothard, H. Bao, A. Kommadath, J. M. Reecy, P. Liu, J. K. Lunney, R.R.R. Rowland, J.C.M. Dekkers and C. K. Tuggle. 2016. Bioinformatic analyses in early host response to Porcine Reproductive and Respiratory Syndrome virus (PRRSV) reveals pathway differences between pigs with alternate genotypes for a major host response QTL. BMC Genomics 17:196.

Serão, E.D. Mauch, M.F. Rothschild and J.C.M. Dekkers. 2016. Genome-wide association study (GWAS) for residual feed intake (RFI) and component traits of feed efficiency in pigs divergently selected for RFI. Am. Soc of Anim Sci Midwest Meetings.

Stothard, H. Bao, A. Kommadath, J. M. Reecy, P. Liu, J. K. Lunney, R.R.R. Rowland, J.C.M. Dekkers and C. K. Tuggle. 2016. Bioinformatic analyses in early host response to Porcine Reproductive and Respiratory Syndrome virus (PRRSV) reveals pathway differences between pigs with alternate genotypes for a major host response QTL. BMC Genomics 17:196.

Smith, T.P.L. S. Koren, A. Phillippy, D.M. Bickhart, B. D. Rosen, K. C. Worley, S. G. Schroeder, B. L. Sayre, I. Liachko, S. T. Sullivan, J. N. Burton, A. R. Hastie, T. S. Sonstegard, C. M. Kelley, A. L. Archibald, M. Watson, R. Green, J. Chin, D. Nonneman, G. A. Rohrer, C. K. Tuggle, H. Liu, 140. D.C. Ciobanu, J.F. Medrano, A. Zimin, S. J. Schultheiss, D. E. Hagen, C. G. Elsik, B. Dalrymple, J. W. Kijas, N. Cockett, and M. P. Heaton. Approaches Taken, Progress Made, and Enhanced Utility of Long Read-based Goat, Swine, Cattle and Sheep Reference Genomes Proceedings of Plant and Animal Genome XXIV. San Diego, CA, USA. Abstract #W635.

Thekkoot D.M., J.M. Young, M.F. Rothschild, and J.C.M. Dekkers. 2016. Genome wide association analysis of sow lactation performance traits in lines of Yorkshire pigs divergently selected for residual feed intake during grow-finish. J. Anim Sci 94:2317–2331. doi:10.2527/jas2015-0258.

Thekkoot D.M., R.A. Kemp, M.F. Rothschild, G.S. Plastow, and J.C.M. Dekkers. 2016 Estimation of genetic parameters for traits associated with reproduction, lactation and efficiency in sows. J. Anim Sci. 94:4516–4529. doi:10.2527/jas2015-0255.

Tuggle, C.K., E. Giuffra, E., S.N. White, L. Clarke, H. Zhou, P.J. Ross, H. Acloque, J.M. Reecy, A. Archibald, R.R. Bellone, M. Boichard, A. Chamberlain, H. Cheng, R.P.M.A. Crooijmans, M.E. Delany, C.J. Finno, M.A.M. Groenen, B. Hayes, J.K. Lunney, J.L. Petersen, G.S. Plastow, C.J. Schmidt, J. Song, M. Watson. 2016. GO-FAANG: a Gathering On Functional Annotation of ANimal Genomes. Animal Genetics 47:528-33.

Walugembe, M., P.M. Swantek, M.S. Honeyman, J.W. Mabry, K.J. Stalder and M.F. Rothschild. 2016. Evaluation of growth, deposition of backfat, and loin muscle for purebred Berkshire pigs housed in bedded hoop buildings. J. Anim. Sci. 94:1-5

Zhou, H., Kern, C., P.J. Ross, C. Kern, P. Saelao, Y. Wang, M.M. Halstead, J. L. Chitwood, I. Korf, T.H. Kim, M. Delany, H. Cheng, J.F. Medrano, A. Van Eenennaam, C.K. Tuggle, and C.W. Ernst. Genome-wide Functional Annotation of Regulatory Elements in Livestock Species. Proceedings of Plant and Animal Genome XXIV. San Diego, CA, USA. Abstract #P0676.

B. Publications Planned

Bertolini F., T. Yang, Y. Huang, J.C.S. Harding, M.F. Rothschild and Plastow G.S. 2017. Failure to Thrive Syndrome (PFTS): is there a genetic component? PAG, 25^{h} Plant and Animal Genomics Conference, 14-18 January, San Diego, CA, US.

Bertolini F., J.C.S. Harding, B. Mote, A. Ladinig, G.S. Plastow and M.F. Rothschild. 2017. Genomic investigation of piglet resilience following porcine epidemic diarrhea outbreaks. Animal Genetics. (In press).

Huang, J., M. Schroyen, Y.Nguyen, N. Gabler, D. Nettleton, J.C.M. Dekkers, C.K. Tuggle. 2017. Identifying tissue specific gene expression using RNAseq data from multiple porcine tissues. 25^{h} Plant and Animal Genomics Conference, 14-18 January, San Diego, CA, US.

Liu, H., N. Manchanda 1, D. Nonneman, T.P.L. Smith, C.K. Tuggle 2017. Cataloguing multi-tissue transcriptomes by PacBio IsoSeq and Illumina RNA-seq, and its application in annotating new-generation swine reference genome assemblies: Lessons learned from and recommendations given. 25^{h} Plant and Animal Genomics Conference, 14-18 January, San Diego, CA, US.

Liu, H., T.P.L. Smith, D.J. Nonneman, J.C.M. Dekkers, C.K. Tuggle et al., A high-quality annotated transcriptome of swine peripheral blood. (in review).

Rajao, D.S., C.L. Loving, P.C. Gauger, E. H. Waide, J.C.M. Dekkers, C.K. Tuggle, and Amy L. Vincent. 2017. Inability of Severe Combined Immune Deficient (SCID) pigs to control IAV replication despite innate immune activation. J. Innate Immunity, in press.

Waide, E., C.K. Tuggle, N.V.L. Serão, M. Schroyen, A. Hess, R.R.R. Rowland, J.K. Lunney, G. Plastow, and J.C.M. Dekkers. 2017. Genome-wide Association of Piglet Response to one of two isolates of the Porcine Reproductive and Respiratory Syndrome Virus. J. Animal Science, in press.

Zurbrigg K., T. van Dreumel, M. F. Rothschild, D. Alves, R. Friendship, T. O'Sullivan. 2017. Pig-level risk factors for in-transit losses in swine: a review. Can J. Ani. Sci (in review).