

Accomplishments Termination Report

Project No. and Title: NRSP008 National Animal Genome Research Program
Period Covered: 10/2012 to 9/2013
Date of Report: 1/17/14
Annual Meeting Dates: 1/11/2014 to 1/12/2014

Participants

The NRSP8 Aquaculture Genome activities are led by Coordinator John Liu (Auburn University), Co-Coordinator Caird Rexroad (USDA/ARS) and Species Leaders Sylvie Quiniou (catfish, USDA/ARS), Craig Sullivan (striped bass, Aquagyn), Dina Proestou (oyster, USDA/ARS) and Yniv Palti (salmonids, USDA/ARS).

The Aquaculture Genome Workshop was held on January 11, 2014 in conjunction with PAG XXII. Steven Roberts (University of Washington) served as Workshop Chair assisted by Roger Vallejo (USDA/ARS) who served as Chair – Elect. At least 50 participants representing 15 countries (US, Canada, Mexico, Korea, New Zealand, Japan, Chile, Singapore, Thailand, Malaysia, France, Germany, Brazil, Belgium, UK) were in attendance. The workshop had the theme of *Genome Mapping, Tagging and Characterization* and included three invited and sixteen contributed presentations.

Brief Summary of Minutes of Annual Meeting

The meeting was called to order by Caird Rexroad, Aquaculture Genome Co-coordinator. Jim Reecy (Iowa State University) provided an update on Bioinformatics Coordinator activities of interest to researchers of aquatic species. LuAnn Glaser (Affymetrix) provided an update on the availability of SNP arrays for rainbow trout, Atlantic salmon, and catfish. Jeffrey Silverstein (USDA/ARS National Program Leader for Aquaculture) spoke on aquaculture funding opportunities from NOAA, NSF, and NIFA. Ying Sun (BGI, Director of the Marine Biobank) informed participants of the Fish 1000 Transcriptomes Project funded by the Chinese National Genebank, including opportunities to develop collaborations.

Caird Rexroad reviewed funded research support activities for 2013, this effort will be repeated in 2014. New Objectives for the new NRSP8 Project were reviewed.

Mohamed Salem (Middle Tennessee State University) was nominated and elected to serve as the Chair Elect and Chair for the 2015 and 2016 aquaculture genome workshops, respectively.

Acacia Alcivar-Warren announced a special session on the topic of shrimp epigenetics at the Aquaculture America Meeting in Seattle, February 2014.

Accomplishments

Progress towards the Objectives are as follows:

Objective 1: Create shared genomic tools and reagents and sequence information to enhance the understanding and discovery of genetic mechanisms affecting traits of interest.

Catfish

- The channel catfish genome assembly has been improved by changing the assembly algorithm from ABySS to MaSuRCA. The current version of the catfish genome assembly (v1.1) has 95% of the channel catfish genome sequences spanning 780.7Mb in 46,936 contigs and 8,597 scaffolds. We are continuing to work with the developers of Celera Assembler to optimize assembly of long PacBio and Illumina reads.
- The catfish genome was annotated using transcriptome sequencing. Through transcriptome analysis of various tissues, a total of over 23,000 complete cDNAs have been assembled and annotated. Gene families and gene duplications were analyzed.
- A draft genetic linkage map from 3-generation families has been produced and is currently analyzed and will be used to assist genome assembly.
- A 250K SNP array based on Affymetrix Axiom technology has been constructed.

Oyster

- The transcriptome of an adult Eastern oyster (*Crassostrea virginica*) was sequenced with short Illumina reads and assembled into 66,229 contigs. The de novo assembly covers 90% of published ESTs and a set of ~40K contigs have been annotated using public databases. 657 genes related to innate immunity have been identified.
- RNA sequencing of *C. virginica* samples collected before and after the Deep Water Horizon oil spill resulted in a de novo transcriptome assembly where 9,469 transcripts were homologous to Pacific oyster transcripts. RNA seq data are being used to identify potential effects of oiled water and sediments on the Eastern oyster.
- A *Crassostrea gigas* fosmid library was constructed that contains 459,936 clones representing 22.34-fold haploid genome equivalents. End sequencing revealed over 6000 sequences with open reading frames \geq 300 bp, 1 million SNPs, and 3200 SSRs.
- Fifty-six SNPs were identified in *C. gigas* sequences mined from the EST database. Forty-two SNPs conform to Hardy-Weinberg Equilibrium and 28 are polymorphic in a full-sib family, suggesting these SNPs will be useful for pedigree analysis, association studies and marker assisted selection.

Salmonids

- To identify genes and gene products that are essential in the regulation of embryonic development in rainbow trout, RNA-Seq analysis was performed on eight RNA samples isolated from developing embryos. There are 2,020 transcripts that are only expressed in

embryos before cell division, and 34 genes that start to express in 3d embryos, the onset of maternal zygotic transition in rainbow trout. In addition, a total of 50,351 novel transcripts were identified from the dataset, and 3,329 to 17,312 splice variants were observed at different stages of embryonic development.

- The first rainbow trout high density 57K SNP chip was developed and characterized. Approximately 50K of the SNPs were validated in a panel of 18 rainbow trout populations at the standard 97% call rate of the Affymetrix SNP polisher software.

Striped Bass

- Genomic DNA (30.52 Gb) sequenced from 4 domesticated striped bass was assembled into ~517 Mb comprised of 71,500 contigs averaging 7 Kb, with several over 80Mb and one >100 Mb. Contig coverage is generally 30X, with fewer than 10 contigs >400X, suggesting the genome is near 600 Mb, similar to the confamilial European sea bass. Over 200 million unique sequences of small RNAs from ovarian tissues were obtained, with most representing piRNAs expressed in early oogenesis, including ~400 miRNAs known to regulate transcript translation and degradation.
- Striped bass and white bass are the parental species of the hybrid striped bass (white bass, *Morone chrysops* X striped bass, *M. saxatilis*). Major tissues and organs (brain, liver, spleen, kidney, ovary, testes, etc) from 10 individuals from each species (5 male and 5 female) were harvested and RNA sequenced in a lane of Illumina HiSeq2000. A total of 262×10^6 high quality reads were obtained with 135×10^6 reads from striped bass and 127×10^6 reads from white bass. Reads were assembled into 203,587 striped bass contigs and 185,531 white bass contigs. Annotation was carried out by BLAST against the UniProt and nr databases for both species. Again, similar results were obtained from both species, with 18,630 UniProt and 23,605 nr annotated unigenes in striped bass and 18,584 UniProt and 22,354 nr annotated unigenes in white bass.

Objective 2: Facilitate the development and sharing of animal populations and the collection and analysis of new, unique and interesting phenotypes.

Catfish

- Bulk segregant RNA-seq (BSR-Seq) was used to analyze differentially expressed genes and associated SNPs with disease resistance against enteric septicemia of catfish (ESC). A total of 1,255 differentially expressed genes were found between resistant and susceptible fish. In addition, 56,419 SNPs were identified as significant SNPs between susceptible and resistant fish located on 4,304 unique genes. Detailed analysis of these significant SNPs allowed differentiation of significant SNPs caused by genetic segregation and those caused by allele-specific expression. Mapping of the significant SNPs, along with analysis of differentially expressed genes, allowed identification of candidate genes underlying disease resistance against ESC.
- Genotyping-by-sequencing was conducted on individuals from populations of wild and aquacultured blue catfish. Markers were validated and extended using multiplex Sequenom MassArray assays and should be useful in follow up studies of the diversity of cultured blue catfish populations and in parentage studies.
- In-depth transcriptome sequencing of channel catfish resistant and susceptible to *Flavobacterium columnare* as well as microbiome sequencing of channel, blue, and hybrid catfish mucosal tissues.

Oyster

- Sequence polymorphisms and differential gene expression patterns were identified that can distinguish among two *C. gigas* lines exhibiting either high or low survival with respect to summer mortality.

Salmonids

- Several studies were completed to investigate genetic variation of multiple salmonid species including Chinook salmon, steelhead/rainbow trout, and cutthroat trout. Studies included investigation into the genetic basis for traits such as thermal adaptation and migration.
- QTL mapping families for stress response and bacterial cold water resistance in rainbow trout that were previously genotyped with microsatellites, were re-genotyped with ~5,000 restriction-site associated DNA (RAD) SNPs. The major microsatellite QTL were validated by the new RAD SNPs linkage maps. Sequence information from the RAD SNPs is useful for aligning the QTL with sequence contigs from the rainbow trout draft genome assembly in an effort to identify positional candidate genes.

Striped Bass

- Novel supervised machine learning analyses identified networks of expressed ovarian genes and proteins that collectively function to determine a complex phenotype, egg quality. Artificial neural networks (ANNs) were used to reveal a powerful relationship ($R^2 > 90\%$) between profiles of maternal ovary gene expression and subsequent egg

fertility in wild and domestic striped bass. RNAseq data from the same fish is being mined for single nucleotide polymorphisms (SNPs) to determine if there is a genetic basis for egg quality. K-means clustering and support vector machines (SVMs) were applied to quantitative tandem mass spectrometry data to reveal a strong relationship ($R^2 > 83\%$) between ovarian stage and protein profiles during the annual reproductive cycle.

Objective 3: Develop, integrate and implement bioinformatics resources to support the discovery of genetic mechanisms that underlie traits of interest.

The aquaculture community works with the Bioinformatics Coordinator to develop species-specific resources, such as those included in the Animal QTLdb. Large sequence databases are also publicly available at www.animalgenome.org/aquaculture/database/.

Oyster

- SQLShare has been used to store, distribute, and query large genomic datasets from the Pacific oyster. Details of this project, including tutorials for the freely available resource are available at: github.com/sr320/qdod/wiki.

Salmonids

- The working draft of the rainbow trout genome assembly reported in 2012 was placed on the animalgenome.org web site hosted by the NRSP-8 Bioinformatics Coordinators. It is now available for downloading by the general public. In addition, an excel file with the genome location of the 145K RAD SNPs dataset reported in 2012 is available from the same web site and as an appendix file from the journal of Molecular Ecology Resources.

Striped Bass

- A new high performance computing cluster dedicated to NGS analysis in studies of striped bass. This new cyberinfrastructure, built around the open source scientific computing platform – Galaxy, was used successfully to assemble an ovarian transcriptome from RNA-seq data.

Impacts

NRSP-8 coordinators' funds were used to enhance student and postdoctoral fellow participation in national meetings (i.e. PAG, Aquaculture America) through Aquaculture Genomics Travel Awards. Also, the sharing of technology and information from aquaculture community and animal genome community was enhanced through invitation of three speakers to the annual Aquaculture Genome Workshop along with PAG. In addition, through open competition, the NRSP8 Aquaculture Coordinators provided research support for four projects listed below; progress is documented under corresponding objectives.

- 1) Improving the Rainbow Trout Genome Assembly using Moleculo Technology (P.I. Michael Miller, University of California Davis).
- 2) A flexible Platform for Querying Disparate Oyster Datasets (qDOD) (P.I. Steve Roberts, University of Washington).
- 3) NRSP-8 Research Support Funds for Moronid Reference Transcriptomes (P.I. Adam Fuller, USDA ARS).
- 4) Assessing Demographic and Evolutionary Relationships between Shrimp Populations by Sequencing a Shrimp Diversity Panel (P.I. Zhiqiang Du and Max Rothschild, Iowa State University).

Publications

Oyster

Zhang, L., Li, L., Zhu, Y., Zhang, G., Guo, X., 2013. Transcriptome Analysis Reveals a Rich Gene Set Related to Innate Immunity in the Eastern Oyster (*Crassostrea virginica*). Marine Biotechnology, DOI 10.1007/s10126-013-9526-z.

Raghavachari, N., Showmaker, K., Liu, P., Jafari, N., Barker, N., Willett, K.L., Corrales, J., Patterson, H.K., Carmichael, R.H., Baldwin, D., 2013. Genomics Research Group (GRG): Elucidating the Effects of the Deepwater Horizon Oil Spill on the Atlantic Oyster Using Global Transcriptome Analysis. J Biomol Tech. 24(Suppl), S68–S69.

Zhang, L., Li, L., Xu, F., Qi, H., Zhang, G., 2013. Fosmid library construction and end sequences analysis of the Pacific oyster, *Crassostrea gigas*. Molluscan Research 33, 65–73.

Zhong, X., Li, Q., Yu, H., Kong, L., 2013. Development and Validation of Single-nucleotide Polymorphism Markers in the Pacific Oyster, *Crassostrea gigas*, Using High-resolution Melting Analysis. Journal of the World Aquaculture Society 44, 455–465.

Schmitt, P., Santini, A., Vergnes, A., Degremont, L., de Lorgeril, J., 2013. Sequence Polymorphism and Expression Variability of *Crassostrea gigas* Immune Related Genes Discriminate Two Oyster Lines Contrasted in Term of Resistance to Summer Mortalities. PLoS One 8, DOI 10.1371/journal.pone.0075900.

Gavery, M.R., Roberts, S.B., 2013. Predominant Intragenic Methylation is Associated with Gene Expression Characteristics in a Bivalve Mollusc. *PeerJ* 1:e215, doi:10.7717/peerj.215. <http://dx.doi.org/10.7717/peerj.215>

Timmins-Schiffman EB Nunn BL, Goodlett DR and Roberts SB. (2013) Shotgun proteomics as a viable approach for biological discovery in the Pacific oyster. Conservation Physiology. doi:10.1093/conphys/cot009

Catfish

Waldbieser, G.C. and Bosworth, B.G. A standardized microsatellite marker panel for parentage and kinship analyses in channel catfish, *Ictalurus punctatus*. Animal Genetics 44:476-479. 2013.

Bosworth, BG, Waldbieser, GC. 2014. General and specific combining ability of male blue catfish *Ictalurus furcatus* and female channel catfish *I. punctatus* for growth and carcass yield of their F1 hybrid progeny. Aquaculture 420-421: 147-153.

Quiniou S.M.A., Boudinot P. and Bengten E. Comprehensive survey and genomic characterization of Toll-like receptors (TLRs) in channel catfish, *Ictalurus punctatus*: Identification of novel fish TLRs. Immunogenetics 65 (7) 511-530. 2013.

Vollota-Herdoiza D., Pila E.A., Quiniou S., Waldbieser G. and Magor B. Transcriptional regulators of teleost aicda genes. Pt 1 suppressors of promiscuous promoters. Fish and Shellfish Immunology 35(6) 1981-1987, 2013.

Peatman E, Li C, Peterson B, Farmer B, Straus D, Beck B. (2013) Basal polarization of the mucosal compartment in *Flavobacterium columnare* sensitive and resistant channel catfish (*Ictalurus punctatus*). *Molecular Immunology*, 56(4):317-327

Liu L, Li C, Su B, Beck B, Peatman E (2013) Short-term feed deprivation alters immune status of surface mucosa in channel catfish (*Ictalurus punctatus*). *PLoS ONE* 8(9): e74581

Li C, Beck B, Su B, Terhune J, Peatman E (2013) Early Mucosal Responses in Blue Catfish (*Ictalurus furcatus*) Skin to *Aeromonas hydrophila* Infection. *Fish Shellfish Immunology* 34(3):920-8

Li C, Wang R, Su B, Luo Y, Terhune J, Beck B, Peatman E (2013) Evasion of Mucosal Defenses During *Aeromonas hydrophila* Infection of Channel Catfish (*Ictalurus punctatus*) Skin. *Dev Comp Immunol* 39 (2013) 447–455

Wang R, Sun L, Bao L, Zhang J, Jiang Y, Yao J, Song L, Feng J, Liu S, Liu ZJ. 2013. Bulk segregant RNA-seq reveals expression and positional candidate genes and allele-specific expression for disease resistance against enteric septicemia of catfish. *BMC Genomics* BMC Genomics 14: 929.

Feng J, Liu S, Wang X, Kaltenboeck L, Kucuktas H, Li J, Liu ZJ 2014. Hemoglobin genes are differentially regulated under heat stress conditions between sensitive and tolerant catfish in different tissues. *Comparative Biochemistry and Physiology, Part D, Genomics and Proteomics* 9:11-22.

Dunham R, Taylor JF, Rise M, Liu ZJ, 2013. Development of strategies for integrated breeding, genetics and applied genomics for genetic improvement of aquatic organisms. *Aquaculture*, in press.

Jiang Y, Gao X, Liu S, Zhang Y, Liu H, Sun F, Bao L, Waldbieser G, Liu ZJ. 2013. Whole genome comparative analysis of channel catfish (*Ictalurus punctatus*) with four model fish species. *BMC Genomics* 14:780. DOI: 10.1186/1471-2164-14-780.

Sookruksawong S, Sun F, Liu ZJ, Tassanakajon A. 2013. RNA-Seq analysis reveals genes associated with resistance to Taura syndrome virus (TSV) in the Pacific white shrimp *Litopenaeus vannamei*. *Developmental and Comparative Immunology* 41:523-533.

Sun F, Liu S, Gao X, Jiang Y, Perera D, Wang X, Li C, Sun L, Zhang J, Kaltenboeck L, Dunham R, Liu ZJ. 2013. Male-biased genes in catfish as revealed by RNA-Seq analysis of the testis transcriptome. *PLoS One* 8(7):e68452.

Sahu D, Panda SP, Panda S, Das P, Meher PK, Hazra RK, Peatman E, Liu ZJ, Eknath AE, and Nandi S. 2013. Identification of reproduction-related genes and SSR-markers through expressed sequence tags analysis of a monsoon breeding carp rohu, *Labeo rohita* (Hamilton). *Gene* 524:1-14.

Wang R, Feng J, Li C, Liu S, Zhang Y, Liu ZJ. 2013. Four lysozymes (one c-type and three g-type) in catfish are drastically but differentially induced after bacterial infection. *Fish and Shellfish Immunology* 35:136-145.

Liu S, Wang X, Sun F, Zhang J, Feng J, Liu H, Rajendran KV, Sun L, Zhang Y, Jiang Y, Kaltenboeck L., Kucuktas H, and Liu ZJ. 2013. RNA-Seq reveals expression signatures of genes involved in oxygen transport, protein synthesis, folding and degradation in response to heat stress in catfish. *Physiological Genomics* 45:462-476.

Niu D, Wang L, Sun F, Feng B, Jin K, Liu ZJ, Li J. 2013. Development of molecular resources for an intertidal clam, *Sinonovacula constricta* using 454 transcriptome sequencing. *PLoS One* 8: e67456.

Zhang J, Liu S, Rajendran KV, Sun L, Zhang Y, Sun F, Kucuktas H, Liu H, Liu ZJ. 2013. Pathogen recognition receptors in channel catfish: III. phylogeny and expression analysis of Toll-like receptors. *Developmental and Comparative Immunology* 40: 185-194.

Sun F, Ninwichian P, Liu SK, Liu, ZJ. 2013. Identification and application of sex-specific markers in catfish. In: Sex determination in fishes, ed. Songlin Chen, pp. 131-139, Academic Press, Beijing.

Zhang Y, Liu S, Lu J, Jiang Y, Gao X, Ninwichian P, Li C, Waldbieser G, Liu ZJ. 2013. Comparative genomic analysis of catfish linkage group 8 reveals two homologous chromosomes in zebrafish with extensive inter-chromosomal rearrangements. *BMC Genomics*, *BMC Genomics* 14:387.

Liu S, Li Q, Liu ZJ. 2013. Genome-wide identification, characterization and phylogenetic analysis of 50 catfish ATP-binding cassette (ABC) transporter genes. *PLoS ONE* 8(5): e63895.

Jiang Y, Ninwichian P, Liu S, Zhang J, Liu H, Kucuktas H, Liu ZJ. Generation of physical contig-specific sequences useful for whole genome sequence assembly and scaffolding. *PLoS One* 8 (10): e78872.

Preziosa E, Liu S, Gao X, Liu H, Kucuktas H, Terova G, and Liu ZJ. 2013. Effect of Nutrient Restriction and Re-feeding on Calpain Family Genes in Skeletal Muscle of Channel Catfish (*Ictalurus punctatus*). *PLoS One* 8: e59404.

Salmonids

Hecht, B.C., N.R. Campbell, D.E. Holecek, and S.R. Narum. 2013. Genome-wide association reveals genetic basis for the propensity to migrate in wild populations of rainbow and steelhead trout. *Molecular Ecology* 22:3061-3076.

Moran, P., D.J. Teel, M.A. Banks, T.D. Beacham, M.R. Bellinger, S.M. Blankenship, J.R. Candy, J.C. Garza, J.E. Hess, S.R. Narum, L.W. Seeb, W.D. Templin, C.G. Wallace, and C.T. Smith. 2013. Divergent life-history races do not represent Chinook salmon coast wide: the

importance of scale in Quaternary biogeography. *Canadian Journal of Fisheries and Aquatic Sciences* 70:415-435.

Narum, S.R., C.A. Buerkle, J.W. Davey, M.R. Miller, P.A. Hohenlohe. 2013. Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology* 22:2841-2847.

Narum, S.R., N.R. Campbell, K.A. Meyer, M.R. Miller, and R.W. Hardy. 2013. Thermal adaptation and acclimation of ectotherms from differing aquatic climates. *Molecular Ecology* 22:3090-9097.

Pritchard, V.L., N.R. Campbell, S.R. Narum, M.M. Peacock, and J.C. Garza. 2013. Discovery and characterization of novel genetic markers for use in the management of Lahontan cutthroat trout (*Oncorhynchus clarkia henshawi*). *Molecular Ecology Resources* 13:276-288.

Rexroad III, C., Vallejo, R., Liu, S., Palti, Y. & Weber, G. (2013). Quantitative Trait Loci Affecting Response to Crowding Stress in an F2 Generation of Rainbow Trout Produced Through Phenotypic Selection. *Marine Biotechnology*, 15: 613-627.

Salem M, Manor ML, Aussanasuwannakul A, Kenney PB, Gregory Weber GM and Yao J. 2013. Effect of sexual maturation on muscle gene expression of rainbow trout: RNA-Seq approach. *Physiological Reports* 1 (5), e00120, doi: 10.1002/phy2.120.

Steele, C.A, E. Anderson, M.W. Ackerman, M.A. Hess, N.R. Campbell, S.R. Narum, M. Campbell. Validation of parentage-based tagging for hatchery steelhead in the Snake River basin. 2013. *Canadian Journal of Fisheries and Aquatic Sciences* 70:1046-1054.

Wang L, Tripurani SK, Wanna W, Rexroad III CE and Yao J. 2013. Cloning and characterization of a novel oocyte-specific gene encoding an F-Box protein in rainbow trout (*Oncorhynchus mykiss*) *Reproductive Biology and Endocrinology* 11:86. doi:10.1186/1477-7827-11-86.

Wiens, G.D., Vallejo, R.L., Leeds, T.D., Palti, Y., Hadidi, S., Liu, S., Evenhuis, J.P., Welch, T.J. & Rexroad, C.E., Iii (2013). Assessment of Genetic Correlation between Bacterial Cold Water Disease Resistance and Spleen Index in a Domesticated Population of Rainbow Trout: Identification of QTL on Chromosome Omy19. *PLoS ONE*, 8: e75749

Wiens, G.D., Lapatra, S.E., Welch, T.J., Evenhuis, J.P., Rexroad III, C.E. & Leeds, T.D. (2013). On-farm performance of rainbow trout (*Oncorhynchus mykiss*) selectively bred for resistance to bacterial cold water disease: Effect of rearing environment on survival phenotype. *Aquaculture*, 388-391: 128-136

Wolters, W.R., Burr, G.S., Palti, Y. & Vallejo, R.L. (2013). Phenotypic and Genetic Variation in Two North American Arctic Charr, *Salvelinus alpinus*, Stocks Cultured in a Recirculating Aquaculture System. *Journal of the World Aquaculture Society*, 44: 473-485

Hale MC, Thrower FP, Berntson EA, Miller MR, Nichols K. (2013) Evaluating adaptive divergence between migratory and non-migratory ecotypes of a salmonid fish, *Oncorhynchus mykiss*. *G3: Genes, Genomes, Genetics*. 2(9):1113-1127.

Hohenlohe PA, Day MD, Amish SJ, Miller MR, Kamps-Hughes N, Boyer MC, Muhlfeld CC, Allendorf FW, Johnson EA, Luikart G. (2013) Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Molecular Ecology*. 22(11):3002-3013.

Striped bass

Reading, B.J., Williams, V.N., Chapman, R.W., Williams, T.I., and Sullivan, C.V. 2013. Dynamics of the striped bass (*Morone saxatilis*) ovary proteome reveal a complex network of the translasome. *Journal of Proteome Research* **12**:1691-1699.