Journal of Fish Biology (2008) **72,** 2377–2390
doi:10.1111/j.1095-8649.2008.01898.x, available online at http://www.blackwell-synergy.com

Utilization of microarray technology for functional genomics in ictalurid catfish

Z. Liu*, R. W. Li† and G. C. Waldbieser‡§

**The Fish Molecular Genetics and Biotechnology Laboratory, Department of Fisheries and Allied Aquacultures and Program of Cell and Molecular Biosciences, Aquatic Genomics Unit, Auburn University, Auburn, AL 36849, U.S.A. †USDA-ARS, Bovine Functional Genomics Laboratory, Beltsville, MD 20705, U.S.A. and ‡USDA-ARS, Catfish Genetics Research Unit, 141 Experiment Station Road, Stoneville, MS 38776, U.S.A.*

The channel catfish, *Ictalurus punctatus*, and the closely related blue catfish, *Ictalurus furcatus*, are important species in aquaculture and serve as biological models for immunology, neurobiology and environmental monitoring. Directed and high-throughput sequencing technologies have produced 44 767 channel catfish expressed sequence tags (ESTs) and 10 764 blue catfish ESTs to date. Experiments incorporating large-scale gene expression analyses have utilized a 660 cDNA macroarray, a 1900 cDNA spotted microarray and *in situ* synthesized oligonucleotide microarrays covering 19 000 and 28 000 genes. These experiments have revealed numerous transcripts that are differentially expressed in response to environmental stressors, immunostimulation and exposure to viral and bacterial pathogens. A current collaboration with the U.S. Department of Energy's Community Sequencing Programme will soon provide a much larger resource of ESTs for both catfish species. This will enhance the coverage of the catfish transcriptome and lead to the production of more comprehensive microarrays for future investigations. Functional characterization of the catfish genome will improve the understanding of genetic variation in catfish populations and their responses to environmental variation and enhance the role of the channel catfish as a model species and source of dietary protein.

Key words: catfish; gene expression; genome; microarray.

CATFISH IN RESEARCH: MODEL SPECIES AND AQUACULTURE PRODUCTION

Channel catfish *Ictalurus punctatus* (Rafinesque, 1818) is a hardy species that tolerates a wide temperature range, moderately turbid water and fresh to brackish water. It is native to the central drainages of North America from Northern Mexico to Southern Canada and also native to areas north and east of the Appalachian Mountains. However, natural populations can be found in most of the continental U.S.A. because of introductions since the 1800s (Jackson, 2004). Channel catfish have also been exported to Central and South America,

§Author to whom correspondence should be addressed. Tel.: +1 662 686 3593; fax: +1 662 686 3567; email: geoff.waldbieser@ars.usda.gov

Europe and Asia for recreational fishing and aquaculture (FAO, 2002). Ease of laboratory culture, tolerance to a variety of stressors and availability of varied size class fish (from 20 g juveniles up to 6 kg adults and larger) has made channel catfish a model species amenable to a variety of experimental studies with the capacity to provide investigators with adequate amounts of specific tissues for molecular assays.

This review is focused on the development of assays to measure differential catfish gene expression in response to environmental and experimental perturbations. While catfish functional genomic research to date has focused on aquaculture-related research, the tools developed are useful for studies in other areas for which catfish serve as models. The channel catfish immune system is well characterized, and it is the only fish species where clonal functionally distinct lymphocyte lines can be readily established (Miller *et al.*, 1998; Shen *et al.*, 2001). These cell lines provide a unique resource for investigating the immune system of this ectothermic vertebrate using approaches not available in any other fish species. A unique quality of some catfish immunocytes is the ability to spontaneously self-immortalize in culture (Lin *et al.*, 1992; Miller *et al.*, 1994), a process that is poorly understood and extremely rare in mammalian cells. Channel catfish are also effective as physiological models for neurobiological research, chemosensation and mechanisms of electroreception, osmoregulation, feed intake and appetite control, hormonal control of reproduction and synchronous oocyte development. Channel catfish in natural systems are often used as sentinels to detect and monitor levels of environmental pollutants and as models to deduce the molecular mechanisms of chemical carcinogenesis.

Channel catfish is important to human health as a safe, sustainable source of dietary protein. In 2006, catfish was the sixth most consumed seafood in the U.S.A. (NFI, 2007). Commercial production of channel catfish as dietary protein is the largest sector of U.S. aquaculture, and over 254 million metric tons of catfish were processed for human consumption in 2006 (USDA, 2007). The processed product is inspected for quality by the U.S. Department of Commerce and available year-round to consumers. Catfish production in the U.S.A. is environmentally sustainable with minimal effluent discharge from earthen ponds, low impact on wild broodstock (Simmons *et al.*, 2006) and low use of fishmeal. Long-term viability of commercial catfish production will be strengthened by genetic improvement of the species for traits such as efficient lean growth, increased feed conversion efficiency, reproductive success, improved filet quality and quantity, and reduced susceptibility to pathogens. Efficient selection of superior broodfish depends on identification of genetic polymorphism linked to genes controlling these traits. Elucidation of genetic pathways controlling these traits, and identification of genetic variation that leads to improved performance, will permit correlation of levels of gene expression and diagnostic proteins with phenotypes.

Genomic tools have been developed to aid in the efficient identification and selection for chromosomal regions that contain DNA sequence variation correlated with superior production phenotypes. Understanding of the structural framework of the catfish genome is supported by genetic linkage and physical framework maps (Waldbieser *et al.*, 2001; Liu *et al.*, 2003; Quiniou *et al.*, 2007; Xu *et al.*, 2007) as well as directed and high-throughput DNA sequencing of

cDNA and large insert clones. Catfish cDNA libraries have been produced from numerous embryonic, juvenile and adult tissues of catfish under varying physiological conditions such as tissues from pathogen-infected fish (Ju *et al.*, 2000; Cao *et al.*, 2001; Karsi *et al.*, 2002; Kocabas *et al.*, 2002; Nonneman & Waldbieser, 2005; Li & Waldbieser, 2006; Li *et al.*, 2007). A combined strategy of both normalization and subtraction coupled with negative-hybridization selection for novel clones (Li *et al.*, 2007) has also been used to maximize the detection of unique and rare transcripts. These efforts have led to the deposition, to date, of 1550 core nucleotide, 44 767 expressed sequence tags (ESTs) and 20 367 genome survey DNA sequences in GenBank specific to channel catfish, as well as 10 764 EST and 24 core nucleotide sequences from blue catfish, *Ictalurus furcatus* (Valenciennes, 1840). Blue catfish is closely related to channel catfish, and interspecific mating produces a viable F1 hybrid that contains favourable traits for commercial production (Dunham *et al.*, 1990; Bosworth *et al.*, 1998; Dunham & Argue, 1998; Li *et al.*, 2004). This report reviews progress in the development of tools for gene expression profiling in catfish (Table I).

CATFISH FUNCTIONAL GENOMICS

INITIAL EXPERIMENTS

Catfish functional genomics experimentation has focused primarily on the role of gene expression associated with catfish production phenotypes. While there have been many experiments examining expression of one or few genes using techniques such as Northern blot hybridization, semi-quantitative polymerase chain reaction (PCR) or quantitative real-time polymerase chain reaction (RT-PCR), this review focuses on experiments investigating large numbers of genes. The first large-scale analysis of gene expression in catfish utilized a low-density macroarray of 660 channel catfish cDNAs spotted onto a nylon membrane. This array was used to identify channel catfish brain genes responsive to cold acclimation at four time points (0, 2, 24 and 48 h) after a shift from 24 to 12° C (Ju *et al.*, 2002). The results of this study (*e.g.* transient induction of chaperone and signal transduction pathway genes) provided insight into the genes and molecular pathways altered in the catfish brain during exposure to a type of environmental stress that is relevant to the aquaculture industry (decreased ambient temperature). While most catfish aquaculture occurs in the warmer south-eastern U.S.A., a comprehensive understanding of the genes involved in cold acclimation could reveal suitable targets (*e.g.* expression biomarkers or single nucleotide polymorphisms) for marker assisted selection of catfish broodstock suitable for colder climates.

This array was also used to examine differentially expressed genes in the ictalurid *Ameiurus catus* (L., 1758) (white catfish), after cooling (Kocabas *et al.*, 2004). In colder waters, channel catfish feeding activity, therefore growth, is significantly reduced, whereas white catfish feed actively during winter months. Brain RNA from white catfish subjected to different temperatures was hybridized with the array. Seven genes (ependymin, ribosomal protein L41, ribosomal protein S27, 16S mitochondrial rRNA and three unknowns)

TABLE I. Evolution of catfish gene expression arrays

Array	Elements	Experiments	Citations
Brain macroarray on nylon membrane	660 channel catfish brain cDNA	Cold acclimation in channel catfish brain	Ju *et al.* (2002)
		Cold acclimation in white catfish brain	Kocabas *et al.* (2004)
19 k oligo microarray	18 989 channel catfish ESTs from multiple tissues, 10 perfectly matched and 10 mismatched 24-mers	LPS-stimulated expression in spleen	Li & Waldbieser (2006)
28 k oligo microarray	21 359 channel + 7159 blue catfish ESTs from multiple tissues, six perfectly matched and six mismatched 24-mers	Expression in channel catfish liver after *Edwardsiella ictaluri* infection	Peatman *et al.* (2007)
		Expression in blue catfish liver after *E. ictaluri* infection	Peatman *et al.* (2008)
2 k cDNA microarray	1497 channel catfish cDNAs from whole fry and leucocytes	Focused on cellular responses to pathogens	Unpubl. data

ESTs, expressed sequence tags.

were up-regulated when white catfish were cultured at low temperature. Twenty-four genes (11 known and 13 unknown) were down-regulated in the brain of white catfish cultured under cold acclimation condition with an initial temperature of 24° C and a final temperature of 12 ± 1° C. Further analysis of these sequences and the pathways in which their products act may provide clues for improving channel catfish feeding activity during colder months.

As the number of sequenced catfish ESTs increased, larger-scale investigations were made possible. The channel catfish sequences contained in GenBank were normalized *in silico* with Paracel Transcript Assembler software to produce the Catfish Gene Index, a web-based resource originally maintained by The Institute for Genomic Research and currently at the Dana Farber Cancer Institute. The first index, in February 2002, contained only 3078 putatively unique sequences derived from 5706 total sequences. The current release of the channel catfish gene index, produced in January 2005 and available online (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=catfish), contains a total of 23 262 unique sequences from adult and juvenile brain, pituitary, liver, stomach, intestine, anterior and posterior kidney, olfactory tissue, ovary, testis, gill, macrophage, mixed leucocytes and whole fry at day 5, 9 and 14 post-hatch. These are subdivided into 4402 tentative consensus sequences, 18 705 singleton ESTs and 155 singleton expressed transcripts. The ESTs were almost entirely 5′ end reads (98%).

FIRST GENERATION 19 K OLIGONUCLEOTIDE MICROARRAY

A pilot study was carried out to measure global expression catfish gene expression on a larger scale (Li & Waldbieser, 2006). Two duplicated sets of 10 24-mer oligonucleotides (features) were designed using proprietary software (NimbleGen Systems Inc., Madison, WI, U.S.A.) to represent each EST; one set of oligos was perfectly matched and positioned throughout the sequence, and the second set of duplicates contained two mismatched bases at positions #13 and #19 from the 5′ end. Oligonucleotides (24-mer) were synthesized on the microarray surface (NimbleGen) through an *in situ* maskless array synthesis technology (Singh-Gasson *et al.*, 1999; Nuwaysir *et al.*, 2002). The feature size for the microarray was 16 × 16 μm, and there were 382 409 features within the 17 × 13 mm array area. Of these, 379 652 were catfish specific and the remaining 2757 were used for quality control of oligonucleotide synthesis and hybridization and for signal normalization. This resulted in the platform containing 18 989 catfish expressed sequences of which only 38% (7182) of the repeat-masked sequences could be putatively annotated based on significant basic local alignment search tool (BLAST)X similarity with sequences in the National Center for Biotechnology Information (NCBI) GenPept database (E value cut-off = 0·001 and minimum bit score = 100).

In order to examine the response to bacterial cell-wall components in catfish immune tissue, the experiment consisted of intraperitoneal injection of lipopolysaccharide followed by tissue sampling at various times post-injection. Biotinylated complementary RNA was produced by *in vitro* transcription of total RNA from spleen only and used to probe the microarray. After

hybridization, the microarrays were scanned with an Axon GenePix 4000B scanner (Molecular Devices Corp., Union City, CA, U.S.A.) at 5 μM resolution, and data were extracted from the raw images with NimbleScan software (NimbleGen, Madison, WI, U.S.A.). A total of 10 microarrays were used for this project: two replicates for the control pool and one microarray for each of the eight lipopolysaccharide (LPS)-stimulated individuals. Relative signal intensity ($\log_2$ transformed) was generated for each feature using the robust multi-array average (RMA) algorithm (Irizarry *et al.*, 2003). The signal intensity was background corrected based on the quantile normalization process [Bolstad *et al.* (2003); Bioconductor open source software for Bioinformatics (http://www.bioconductor.org)] for all microarrays from the entire experiment (under uniform conditions) using data from only perfectly matched oligos. This experiment utilized a microarray platform produced by *in situ* oligonucleotide synthesis of sequences from the Catfish Gene Index. This experiment was placed in the NCBI gene expression omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/index.cgi) as accession number GSE3261.

The data were subjected to stringent criteria to maximize the probability of detecting real differences in gene expression. The first requirement was a ratio of signal to global average background $\geq 2{\cdot}0$. Second, the perfectly matched signal must have been 150% of the mismatched signal. Third, the correlation between biological replicates must have been $>0{\cdot}9$. The normalized data were analysed using Significance Analysis of Microarrays (SAM; Tusher *et al.*, 2001) from the TIGR Multiexperiment Viewer analysis software package [TIGR Microarray software (http://www.tigr.org/software/microarray.shtml)] with two-class unpaired design for each time point. The RMA value from perfectly matched oligos was used to calculate fold differences in expression (average signal intensity of treated animals divided by the average signal intensity of control animals). The final criterion was that genes must have changed at least two-fold with an estimated global false discovery rate $<10\%$. Analysis of the data revealed 138 sequences that were differentially expressed after LPS treatment.

Sixty-four genes were up-regulated by LPS exposure, including cytokines and chemokines such as interleukin (IL)-1β, chemokine (C-C motif) ligand 4 (CCL4), a small CXC/IL 8-like chemokine, and chondromodulin II. Transcriptional factors such as NF-κβ p100 subunit and NF-κβ inhibitor alpha-like proteins A and B, interferon regulatory factor I and adaptor-related protein complex 1 (AP-1) were also up-regulated at least two-fold. Expression of toll-like receptor 5 (TLR5) was up-regulated at 2–4 h post-exposure in spleen. Within the experimental timeframe, almost all induced genes had been induced by 4 h post-exposure. Seventy-four genes were significantly down-regulated in response to LPS including immunoglobulin light and heavy chains, major histocompatibility complex (MHC) class II antigens, invariant chain-like protein 2, an NK lysin-like protein, complement C3, haemoglobin alpha and a CXC chemokine receptor.

Nine expressed sequences representing varying levels of expression were analysed using RT-PCR to verify the microarray results. Linear regression of log-transformed expression data demonstrated a strong positive correlation ($r^2 = 0{\cdot}87$) between the two technologies. While a two-fold difference in expression was an arbitrary value, there were, undoubtedly, significant physiological and immunological consequences to differences in gene expression less than

two-fold. With more replication, this experiment might have been able to discern smaller changes in gene expression with statistical confidence.

SECOND GENERATION 28 K OLIGONUCLEOTIDE MICROARRAY

As more expressed sequences were added to the NCBI EST database, primarily from blue catfish tissue and also from directed sequencing of catfish immunoregulatory genes, the sequences were reclustered using the ContigExpress programme of the Vector NTI software suite (Invitrogen, Carlsbad, CA, U.S.A.) and assembled and maintained in an online database (http://titan.biotec.uiuc.edu/cgi-bin/ESTWebsite/estima_start?seqSet=catfish). Singletons and representative clones from the 4257 contigs were selected and reassembled in ContigExpress to ensure a unique gene set. Channel catfish and blue catfish repeat-masked sequences were annotated by BLASTX comparison against the NCBI non-redundant protein database (E value cutoff = 0·00001). BLASTN searches were also added to the analysis in order to find matches of shorter ESTs from a single transcript that may not have clustered. Inclusion on the microarray of separate sequences from the same gene would not have prevented further characterization but would have initially inflated the results until additional characterization resolved the redundancy. Using the combination of BLASTX and BLASTN, 51% of the sequences were annotated with a putative identity (Li *et al.*, 2007).

These sequences were used to produce a second generation catfish microarray on the NimbleGen platform. One of the advantages of this platform was the relative ease in redesigning the microarray by simply changing the sequence database. For each EST represented on the microarray, at least 12 24-mer oligonucleotides were designed in the same way as for the 19K array. Six were perfect match oligos selected along the length of the sequence, while the other six were duplicates of the first but with two mismatched bases at the #6 and #12 positions from the 5′ end (Peatman *et al.*, 2007, 2008). Reducing the level of sequence redundancy permitted synthesis of oligonucleotide corresponding to 28 518 genes on the array. As with the 19K microarray, sensitivity and reproducibility spike-in controls were included in the 28K oligo array platform. Two experiments investigating catfish hepatic gene expression in response to a bacterial pathogen challenge have utilized this array (GEO accessions numbers GSE6105 and GSE6350).

Channel and blue catfishes were challenged with *Edwardsiella ictaluri*, a pathogenic enterobacterium that causes significant losses in catfish production, and hepatic gene expression was examined 3 days post-exposure (Peatman *et al.*, 2007, 2008). In channel catfish, gene expression levels were compared with the levels observed in non-infected control catfish. After data normalization and calculation of robust multiple average (RMA) values, the resulting expression intensity values were analysed using the significance analysis of microarrays (SAM) software. Again, the criteria of a two-fold or greater change in expression and a global false discovery rate of 10% were chosen to identify up- or down-regulated genes in the infected replicates. Using these criteria, 301 transcripts were significantly up-regulated, of which 207 were believed to

represent unique genes. Only six transcripts (five unique) were significantly down-regulated. The redundant transcripts resulted either from blue and channel putative orthologues of the same gene or multiple transcripts from non-overlapping regions of a large cDNA being included on the microarray. The data showed a wide range of transcriptional activity following *E. ictaluri* exposure. Fourteen genes were up-regulated from 10- to 85-fold following infection; 16 genes were up-regulated from five- to 10-fold; 27 genes were up-regulated from three- to five-fold; and 150 genes were up-regulated from two- to three-fold.

An acute phase response was evident in channel catfish after infection. At least 35 of the 127 unique, annotated transcripts represented acute phase proteins (APP; Bayne *et al.*, 2001), including coagulation factors, proteinase inhibitors, transport proteins and complement components. Many of the APPs were up-regulated greater than five-fold, but several APP included on the microarray were not up-regulated at the studied time point and these included mannose-binding lectin 2, serum amyloid P and heparin cofactor II. Two subgroups of APP, iron transport–homeostasis proteins and complement components, were represented by particularly high numbers of up-regulated transcripts. Transcripts representing at least 15 unique complement components or inhibitors were up-regulated two-fold or greater following infection. These included a short transcript likely representing C1q (CV996365) up-regulated 15·3-fold increase; ficolin-like genes up-regulated as much as 32-fold (BM438750); complement C2/Bf; several C3 isoforms; complement components C4 and C5; complement components C7, C8 and C9 active in the membrane attack complex and several complement regulatory proteins including MAC inhibitor CD59, C1 inhibitor and Factor H. The most highly up-regulated group of functionally related catfish genes was composed of genes involved in iron homeostasis. These included intelectin, the most highly up-regulated gene observed at >85-fold, haptoglobin (>34-fold), haemopexin (>25-fold), ceruloplasmin (8·5-fold), transferrin (greater than seven-fold) and ferritin (greater than two-fold) (Peatman *et al.*, 2007). Trends in expression measured by the microarray were confirmed by quantitative RT-PCR analysis.

Because blue catfish is less susceptible to *E. ictaluri* infection than are channel catfish, the microarray was used to identify differential gene expression between the two species after infection. Both species shared a wide spectrum of similarities in gene expression profiles after infection including an acute phase response and strong induction of complement components and iron regulatory genes at day 3 after infection. However, significant differences were also observed between the expression profiles of the two species. A total of 58 genes were differentially expressed in blue catfish liver but not in channel catfish liver at day 3 after infection (Peatman *et al.*, 2008), including immune-relevant genes such as CC chemokine SCYA106, MHC class I alpha chain and matrix metalloproteinase 13 (MMP13). The microarray results prompted further quantitative (Q)-PCR analyses revealing that MMP13 was 21·8-fold up-regulated in channel catfish liver 24 h post-exposure and only 1·6-fold up-regulated in blue catfish liver at this early time point in the controlled pathogen exposure. Differential regulation of the MMP13 gene demonstrated its potential as a candidate for development of polymorphic DNA markers linked either to MMP13 or to genes regulating MMP13 expression. These markers would be

used to identify haplotypes correlated with disease resistance and select broodstock containing these haplotypes.

FIRST GENERATION 2 K CDNA MICROARRAY

Concurrent with the production of the oligonucleotide microarrays was the production of a spotted cDNA microarray that could be developed and used at less expense than the oligo array platforms. This array was based on channel catfish sequences from the Catfish Gene Index and targeted primarily for immunological research using sequences identified in whole fry, macrophages and mixed leucocyte cell culture (A. Arnizaut, L. Hanson, M. Wilson, E. Bengten, pers. comm.). A total of 1910 clones representing 1497 unique sequences (60% annotated) were spotted onto glass slides (Microarrays Inc., Nashville, TN, U.S.A.). Research that will soon be published indicates that this array is valuable in the evaluation of gene expression patterns in whole catfish and catfish cell lines during responses to pathogens (L. Hanson, G. Chinchar, pers. comm.).

COMPARISONS BETWEEN MICROARRAY EXPERIMENTS

Three published microarray experiments were complementary in providing a better understanding of the catfish immune response. There are several reasons why the suite of regulated genes differed between the two channel catfish microarray experiments (Table I) . First, although LPS is a cell wall component of *E. ictaluri* (Baldwin *et al.*, 1997; Moore & Thune, 1999), LPS injection and live bacterial exposure could induce a different response although basic innate immune response profiles may be similar. Second, the results could also reflect the differential response in spleen *v.* liver. Third, the LPS injection experiment analysed several time points up to 1 day post-injection, whereas the live exposure experiment used the microarray to measure expression at 3 day post-exposure. This could be one reason fewer genes were measured as down-regulated in the live exposure compared with the LPS injection experiment. The research also exposed difficulties inherent in performing microarray analyses from tissues *in vivo*. Differences observed between individuals within each treatment, because of asynchronized sampling and, or genetic variation led to averaged effects that could have masked significant responses in some individuals.

One notable similarity between the two channel catfish experiments was the up-regulation of toll-like receptor (TLR)5, and these results confirmed a previous experiment demonstrating up-regulation of this gene in liver, spleen and anterior kidney of channel catfish exposed to *E. ictaluri* (Bilodeau & Waldbieser, 2005). In mammals, TLR5 recognizes flagellin, whereas TLR4 is the extracellular LPS receptor (Poltorak *et al.*, 1998; Takeuchi *et al.*, 1999). Expression of TLR5 was also shown to increase in zebrafish *Danio rerio* (Hamilton, 1822) exposed to the *Mycobacterium marinum* (Meijer *et al.*, 2004) and in the liver of Atlantic salmon *Salmo salar* L., 1758 exposed to the bacterium *Aeromonas salmonicida* (Ewart *et al.*, 2005). The present data supported the

hypothesis that, in fish, microbial products may induce other TLRs than those specific for their own recognition (Meijer *et al.*, 2004). Thus, the microarray was a useful platform for identification of candidate genes, development and testing of new hypotheses, and verification of prior experimental evidence.

CURRENT STATUS AND FUTURE DIRECTION

Microarray analysis is a powerful tool for the identification of gene expression patterns in response to environmental perturbation, the identification of candidate genes controlling traits of interest and the identification of functional genetic variation within catfish populations. These initial catfish microarray studies have identified numerous candidate genes with potential roles in the response of catfish to *E. ictaluri* and differential sensitivity of channel and blue catfish to this pathogen. While the authors wished to have a microarray that could be useful for both species, they also wanted to differentiate transcription from the blue or channel genomes in channel × blue F1 hybrid catfish in investigations of hybrid catfish physiology and immunology. The primary microarray research has provided a solid foundation for future functional characterization of the molecular basis of physiological processes controlling immunity, growth and development, reproduction and carcass composition. Functional characterization of the genome will improve the understanding of catfish genetic variation and response to environmental variation and enhance the role of the catfish as model species and sources of dietary protein.

Analyses are currently hindered by the relatively low level of annotation in catfish sequence databases. This is particularly a problem with genes related to immune function because of the high levels of sequence divergence between species. A consortium of researchers has recently participated in the Community Sequencing Programme administered by the U.S. Department of Energy – Joint Genome Institute to produce a large number of catfish ESTs derived from normalized and subtracted cDNA libraries from multiple tissues. To date, over 400 000 channel catfish EST sequences have been produced and are in the process of submission to GenBank dbEST, and these sequences should capture a high proportion of the catfish transcriptome. Increased coverage of the transcriptome, increased annotation of other fish species for comparative analyses and functional experimentation will improve catfish EST annotation. Production of a draft catfish genome sequence will also aid identification and annotation of coding sequences.

To date, published research has utilized microarrays upon which each expressed sequence is represented by a collection of short 24-mer oligos designed throughout the sequence. The authors chose 24-mer oligos because they tend to be more specific, and this was especially important for the multispecies 28 k array platform because it made possible the discrimination of EST sequences from these two closely related species. The signal intensities were calculated using the RMA method from the perfect matched oligos only and the mismatched oligos only served the purpose of quality control. Ten matched and 10 mismatched oligos were utilized in the 19 k array, whereas only six each of matched and mismatched oligos were used on the 28 k array. Even with

optimization of oligo selection algorithms, the difference in signal intensities of various oligos along a transcript can vary greatly so it is important to have multiple oligos representing a single EST. Analysis of differential hybridization within species-specific oligos that represent an EST can provide clues to aid identification of closely related members of a gene family and identification of potential alternative splicing events. One drawback of these early experiments was that the available catfish EST sequences were almost entirely from the 5′ end of transcript, potentially more conserved between members of a gene family, and less useful than 3′ end sequence to delineate members of gene families. Improved genome and EST annotation can provide information for the production of new sets of catfish oligos in order to improve hybridization specificity and coverage.

The flexibility of the *in situ* synthesis array technology permits frequent and periodic update when additional EST sequences become available, which also facilitates availability of updated arrays to the scientific community. This technology obviates the logistical burden of maintaining multiwell plates containing cDNA clones and removes the opportunity for sequence contamination during clone re-arraying, sampling and maintenance. However, *in situ* synthesis of arrays is more expensive than spotting amplified cDNA. Future research in defined areas may be more efficiently performed using targeted, smaller spotted cDNA arrays that provide higher levels of hybridization specificity for individual probes than short oligo arrays.

In contrast to microarrays, new platforms based on ultra high-throughput DNA sequencing are rapidly developing (Emrich *et al.*, 2007; Torres *et al.*, 2008). These platforms can now provide large quantities of sequencing reads from cDNA templates without the need for an intermediate bacterial cloning step. For example, the Genome Sequencer FLX System (Roche Applied Science, Penzberg, Germany) can produce, in one run, an average 240 bp reads from 400 000 templates. In contrast, digital gene expression technology (Illumina Inc., San Diego, CA, U.S.A.) can provide up to 4 million short sequence tags (*c.* 35 bp) per sample. The resulting sequence data can be clustered and the number of transcripts quantified to provide a quantification of gene expression, and furthermore, this approach is 'open ended' in that one is not restricted to only those transcripts represented on the microarray, and it can detect and quantify novel and rare transcripts. While prior knowledge of catfish mRNA sequence would not be required, it would be useful for identification of novel transcripts. Utilization of these new open-ended, but expensive, platforms will depend on the requirements and expected outcomes of the experiment. For example, a microarray may be more than sufficient for identification of a gene expression profile characteristic of a physiological response or exposure but not sufficient for a full characterization of the pathways involved.

References

Baldwin, T. J., Collins, L. A. & Newton, J. C. (1997). Antigens of *Edwardsiella ictaluri* recognised by serum antibodies from naturally infected channel catfish. *Fish and Shellfish Immunology* **7,** 261–271.

Bayne, C. J., Gerwick, L., Fujiki, K., Nakao, M. & Yano, T. (2001). Immune-relevant (including acute phase) genes identified in the livers of rainbow trout, *Oncorhynchus mykiss*, by means of suppression subtractive hybridization. *Developmental and Comparative Immunology* **25,** 205–217.

Bilodeau, A. L. & Waldbieser, G. C. (2005). Activation of TLR3 and TLR5 in channel catfish exposed to virulent *Edwardsiella ictaluri*. *Developmental and Comparative Immunology* **29,** 713–721.

Bolstad, B., Irizarry, R., Astrand, M. & Speed, T. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19,** 185–193.

Bosworth, B. G., Wolters, W. R., Wise, D. J. & Li, M. H. (1998). Growth, feed conversion, fillet proximate composition and resistance to *Edwardsiella ictaluri* of channel catfish, *Ictalurus punctatus* (Rafinesque), blue catfish, *Ictalurus furcatus* (Lesueur), and their reciprocal F1 hybrids fed 25% and 45% protein diets. *Aquaculture Research* **29,** 251–257.

Cao, D., Kocabas, A., Ju, Z., Karsi, A., Li, P., Patterson, A. & Liu, Z. (2001). Transcriptome of channel catfish (*Ictalurus punctatus*): initial analysis of genes and expression profiles of the head kidney. *Animal Genetics* **32,** 169–188.

Dunham, R. A. & Argue, B. J. (1998). Seinability of channel catfish, blue catfish, and their F1, F2, F3 and backcross hybrids in earthen ponds. *Progressive Fish Culturist* **60,** 214–220.

Dunham, R. A., Brummett, R. E., Ella, M. O. & Smitherman, R. O. (1990). Genotype-environment interactions for growth of blue, channel, and hybrid catfish in ponds and cages at varying densities. *Aquaculture* **85,** 143–151.

Emrich, S. J., Barbazuk, W. B., Li, L. & Schnable, P. S. (2007). Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Research* **17,** 69–73.

Ewart, K. V., Belanger, J. C., Williams, J., Karakach, T., Penny, S., Tsoi, S. C., Richards, R. C. & Douglas, S. E. (2005). Identification of genes differentially expressed in Atlantic salmon (*Salmo salar*) in response to infection by *Aeromonas salmonicida* using cDNA microarray technology. *Developmental and Comparative Immunology* **29,** 333–347.

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U. & Speed, T. P. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4,** 249–264.

Jackson, D. C. (2004). Natural history and fisheries. In *Biology and Culture of Channel Catfish* (Tucker, C. S. & Hargreaves, J. A., eds), pp. 15–35. Amsterdam: Elsevier.

Ju, Z., Karsi, A., Kocabas, A., Patterson, A., Li, P., Cao, D., Dunham, R. & Liu, Z. (2000). Transcriptome analysis of channel catfish (*Ictalurus punctatus*): genes and expression profile from the brain. *Gene* **261,** 373–382.

Ju, Z., Dunham, R. & Liu, Z. (2002). Differential gene expression in the brain of channel catfish (*Ictalurus punctatus*) in response to cold acclimation. *Molecular Genetics and Genomics* **268,** 87–95.

Karsi, A., Cao, D., Li, P., Patterson, A., Kocabas, A., Feng, J., Ju, Z., Mickett, K. D. & Liu, Z. (2002). Transcriptome analysis of channel catfish (*Ictalurus punctatus*): initial analysis of gene expression and microsatellite-containing cDNAs in the skin. *Gene* **285,** 157–168.

Kocabas, A. M., Li, P., Cao, D., Karsi, A., He, C., Patterson, A., Ju, Z., Dunham, R. A. & Liu, Z. (2002). Expression profile of the channel catfish spleen: analysis of genes involved in immune functions. *Marine Biotechnology* **4,** 526–536.

Kocabas, A. M., Dunham, R. A. & Liu, Z. (2004). Alterations in gene expression in the brain of white catfish (*Ameiurus catus*) in response to cold acclimation. *Marine Biotechnology* **6,** S431–S438.

Li, R. W. & Waldbieser, G. C. (2006). Production and utilization of a high-density oligonucleotide microarray in channel catfish, *Ictalurus punctatus*. *BMC Genomics* **7,** 134.

Li, M. H., Robinson, E. H., Manning, B. B., Yant, D. R., Chatakondi, N. G., Bosworth, B. G. & Wolters, W. R. (2004). Comparison of the channel catfish, *Ictalurus punctatus* (NWAC103 Strain) and the channel X blue catfish, *I. punctatus* × *I. furcatus*, F1 hybrid for growth, feed efficiency, processing yield, and body composition. *Journal of Applied Aquaculture* **15,** 63–71.

Li, P., Peatman, E., Wang, S., Feng, J., He, C., Baoprasertkul, P., Xu, P., Kucuktas, H., Nandi, S., Somridhivej, B., Serapion, J., Simmons, M., Turan, C., Liu, L., Muir, W., Dunham, R., Brady, Y., Grizzle, J. & Liu, Z. (2007). Towards the ictalurid catfish transcriptome: generation and analysis of 31,215 catfish ESTs. *BMC Genomics* **8,** 177.

Lin, G. L., Ellsaesser, C. F., Clem, L. W. & Miller, N. W. (1992). Phorbol ester/calcium ionophore activate fish leukocytes and induce long-term cultures. *Developmental and Comparative Immunology* **16,** 153–163.

Liu, Z., Karsi, A., Li, P., Cao, D. & Dunham, R. (2003). An AFLP-based genetic linkage map of channel catfish (*Ictalurus punctatus*) constructed by using an interspecific hybrid resource family. *Genetics* **165,** 687–694.

Meijer, A. H., Krens, S. F., Medina Rodriguez, I. A., He, S., Bitter, W., Snaar-Jagalska, B. E. & Spaink, H. P. (2004). Expression analysis of the toll-like receptor and TIR domain adaptor families of zebrafish. *Molecular Immunology* **40,** 773–783.

Miller, N. W., Rycyzyn, M. A., Wilson, M. R., Warr, G. W., Naftel, J. P. & Clem, L. W. (1994). Development and characterization of channel catfish long term B cell lines. *Journal of Immunology* **152,** 2180–2189.

Miller, N., Wilson, M., Bengtén, E., Stuge, T., Warr, G. & Clem, L. W. (1998). Functional and molecular characterization of teleost leukocytes. *Immunological Reviews* **166,** 187–197.

Moore, M. M. & Thune, R. L. (1999). Evaluation of antigenic and nonantigenic proteins of *Edwardsiella ictaluri* at two culture temperatures and in enriched and minimal media using two-dimensional polyacrylamide gel electrophoresis. *Journal of Aquatic Animal Health* **11,** 262–274.

Nonneman, D. & Waldbieser, G. C. (2005). Isolation and enrichment of abundant microsatellites from a channel catfish (*Ictalurus punctatus*) brain cDNA library. *Animal Biotechnology* **16,** 103–116.

Nuwaysir, E. F., Huang, W., Albert, T. J., Singh, J., Nuwaysir, K., Pitas, A., Richmond, T., Gorski, T., Berg, J. P., Ballin, J., McCormick, M., Norton, J., Pollock, T., Sumwalt, T., Butcher, L., Porter, D., Molla, M., Hall, C., Blattner, F., Sussman, M. R., Wallace, R. L., Cerrina, F. & Green, R. D. (2002). Gene expression analysis using oligonucleotide arrays produced by maskless photolithography. *Genome Research* **12,** 1749–1755.

Peatman, E., Baoprasertkul, P., Terhune, J., Xu, P., Nandi, S., Kucuktas, H., Li, P., Wang, S., Somridhivej, B., Dunham, R. & Liu, Z. (2007). Expression analysis of the acute phase response in channel catfish (*Ictalurus punctatus*) after infection with a Gram-negative bacterium. *Developmental and Comparative Immunology* **31,** 1183–1196.

Peatman, E., Terhune, J., Baoprasertkul, P., Xu, P., Nandi, S., Wang, S., Somridhivej, B., Kucuktas, H., Li, P., Dunham, R. & Liu, Z. (2008). Microarray analysis of gene expression in the blue catfish liver reveals early activation of the MHC class I pathway after infection with *Edwardsiella ictaluri*. *Molecular Immunology* **45,** 553–566.

Poltorak, A., He, X., Smirnova, I., Liu, M. Y., Van Huffel, C., Du, X., Birdwell, D., Alejos, E., Silva, M., Galanos, C., Freudenberg, M., Ricciardi-Castagnoli, P., Layton, B. & Beutler, B. (1998). Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282,** 2085–2088.

Quiniou, S. M. A., Waldbieser, G. C. & Duke, M. V. (2007). A first generation BAC-based physical map of the channel catfish genome. *BMC Genomics* **8,** 40.

Shen, L., Stuge, T. B., Zhou, H., Khayat, M., Barker, K. S., Quiniou, S. M. A., Wilson, M., Bengtén, E., Chinchar, V. G., Clem, L. W. & Miller, N. W. (2001). Channel

catfish cytotoxic cells: a mini-review. *Developmental and Comparative Immunology* **26,** 141–149.
Simmons, M., Mickett, K., Kucuktas, H., Li, P., Dunham, R. & Liu, Z. J. (2006). Comparison of domestic and wild catfish populations provide no evidence for genetic impact. *Aquaculture* **252,** 133–146.
Singh-Gasson, S., Green, R. D., Yue, Y., Nelson, C., Blattner, F., Sussman, M. R. & Cerrina, F. (1999). Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nature Biotechnology* **17,** 974–978.
Takeuchi, O., Hoshino, K., Kawai, T., Sanjo, H., Takada, H., Ogawa, T., Takeda, K. & Akira, S. (1999). Differential roles of TLR2 and TLR4 in recognition of gram-negative and gram-positive bacterial cell wall components. *Immunity* **11,** 443–451.
Torres, T. T., Metta, M., Ottenwälder, B. & Schlötterer, C. (2008). Gene expression profiling by massively parallel sequencing. *Genome Research* **18,** 172–177.
Tusher, V. G., Tibshirani, R. & Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Science of the United States of America* **98,** 5116–5121.
Waldbieser, G. C., Bosworth, B. G., Nonneman, D. J. & Wolters, W. R. (2001). A microsatellite-based genetic linkage map of the channel catfish, *Ictalurus punctatus*, genome. *Genetics* **158,** 727–734.
Xu, P., Wang, S., Liu, L., Thorsen, J., Kucuktas, H. & Liu, Z. (2007). A BAC-based physical map of the catfish genome. *Genomics* **90,** 380–388.

Electronic References

FAO (Food and Agriculture Organization) (2002). *The State of World Fisheries and Aquaculture*. Rome: Fishery Resources Division, Fisheries Department, United Nations Food and Agriculture Organization. Available at http://www.fao.org/DOCREP/003/X8002E/X8002E00.HTM
NFI (National Fisheries Institute) (2007). *Top 10 U.S. Consumption of Seafood by Species Chart*. McLean, Virginia: National Fisheries Institute. Available at http://www.aboutseafood.com/media/top_10.cfm
USDA (U.S. Department of Agriculture) (2007). *Catfish Processing*. Washington, DC: National Agricultural Statistics Service, U.S. Department of Agriculture. Available at http://usda.mannlib.cornell.edu/usda/current/CatfProc/CatfProc-09-21-2007.txt