International Research Journal of Applied and Basic Sciences. Vol., 2 (8), 303-319, 2011
Available online at http://www. irjabs.com

IRJABS

IN-SILICO STUDIES OF INDIAN CATFISH: AN UPDATE

SURAJIT DEBNATH

Department of Medical Laboratory Technology, Women,s Polytechnic, Govt. of Tripura
Hapania,P.O:Amtali via Sekerkot,Tripura (W).PIN:799130 , India.

Corresponding author: E-mail: Surajit03@yahoo.co.in. Ph/FAX: 91-381-2376814. 0-9402335151 (M).

Abstract: Computational analysis of nucleic acids and proteins from fishes has enriched our understanding of several fundamental problems in classical fish biology and genetics. Mining of data which is analogous to information discovery is the process of automating knowledge discovery. Analysis of data in all possible manners and recognition of a pattern which has a scientific basis using an efficient algorithm is the fundamental principle of Bioinformatics. Ever-increasing number of Bioinformatics tools and computational packages are now a days used in tandem to analyze sequence of a loci and to model a novel gene product. High efficiency algorithms with excellent homology searching functions have lead to this advancement. Cat fishes, a group of diverse bony fishes have been studied with respect to genetic analysis in several species and with several peptides. Tracking of the evolutionary foot prints using genomic libraries and among the expressed sequences has been generating much interest. Recent development in bioinformatics research in catfishes needs a comprehensive review to develop text for the academia as well as the industry for ready reconer to use and maximize potential of already developed tools and software packages. Identifying the sectors where catfish bioinformatics necessitate further research is also needed.

Key words: *In-silico* biology; Applications; Tropical; Catfish; Genetic diversity.

INTRODUCTION

Advances in new technologies to study genetic and molecular hallmarks of living beings has revolutionized the way the scientific community engages in research and also the way it represents the findings from a wet lab. Genomics and proteomics have largely taken over the tools used in basic research and increasing amount of data is regularly generated by molecular techniques throughout the globe. New methodologies as high throughput DNA sequencing for studies on functional , structural and evolutionary genetics and cDNA microarrays for study of expression pattern of these genes are now combined with a third dimension in which the microprocessor is used for presenting the findings in a more application oriented and customer centric format. The *in-silico* studies of laboratory findings have generated much interest and the demarcation between the lab and the industry is getting thinner day by day. A sequence data of a peptide is now reported with a full scale homology and structural analysis using a software program to identify motifs with economical importance. These computer packages has given scientists a wing for imagination where they can predict, model as well analyze functions in-silico of genes and proteins.Research in fisheries and aquaculture has also been blessed with the high efficiency computational methods where genetic and proteomic data has been utilized in analyzing

function of multiple physiological, morphological and behavioral traits of interests in fishes. In an interesting presentation to indicate the efficiency of microprocessor technology for molecular analysis Hauser & Seeb in 2008 reported that in every 2 years there is a doubling of throughput of number of base pairs that can be sequenced for one dollar (US$1) (Fig:1). This shows the economy and feasibility of computation in biological research and automated processing by computers. Study of novel genes and peptides can be extended to analyze phylogenetics of populations and phylogeography as well as to devise protocol for conservation practices. Genome projects of several fish species are proposed which will leave no stone unturned in the understanding life and biology of these species. Among various fish groups the catfishes has also been studied with respect to several loci by fish geneticist but genetic studies in Indian catfish *Clarias batrachus* is fragmentary in literature. This section reviews the current status in computational biological studies in catfish in general and *Clarias batrachus* in particular. A comparatively simple culture characteristic with efficient food conversion (Hecht *et al*., 1996; Hargreaves & Tucker, 2003; Ali & Jauncey, 2005) and excellent nutritional profile of tissue (Rui *et al* 2007, Pruszyński 2003, Debnath 2009) makes *Clarias batrachus* very suitable for commercial intensive culture. According to FAO data there has been a regular growth in the global production and subsequent earning from different catfish varieties. There has been an year wise growth in the production and processing of the catfishes (Fig: 2) and a high demand resulted in to an increase in the earning from these varieties as well.Coordination between government bodies with respect to skill up gradation of the workers, market regulation etc together with the scientific community ensuring timely delivery of better quality seed stock will generate success stories in intensive *Clarias batrachus* culture (Fig:3). In this context the role of bioinformatics is not less, instead search for quantitative trait loci (QTL) and experimental modeling of active ingredients promoting somatic growth should be encouraged for their future benefits that may reap into rich commercial as well as knowledge based gains.

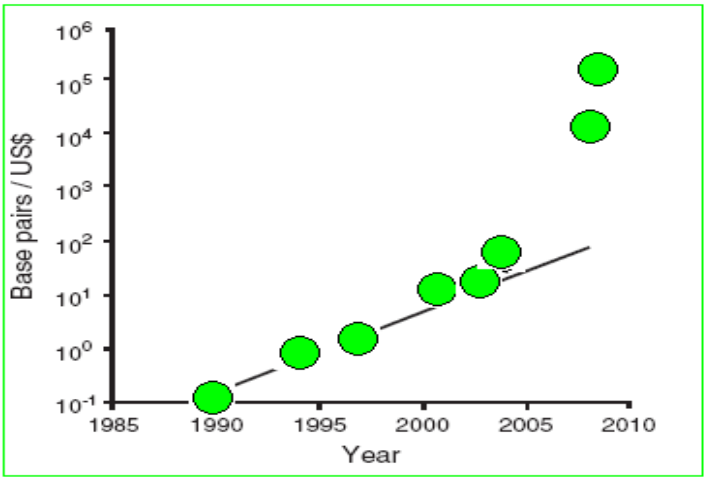


Figure: 1, Economy of gene sequencing. (After Hauser & Seeb, 2008).

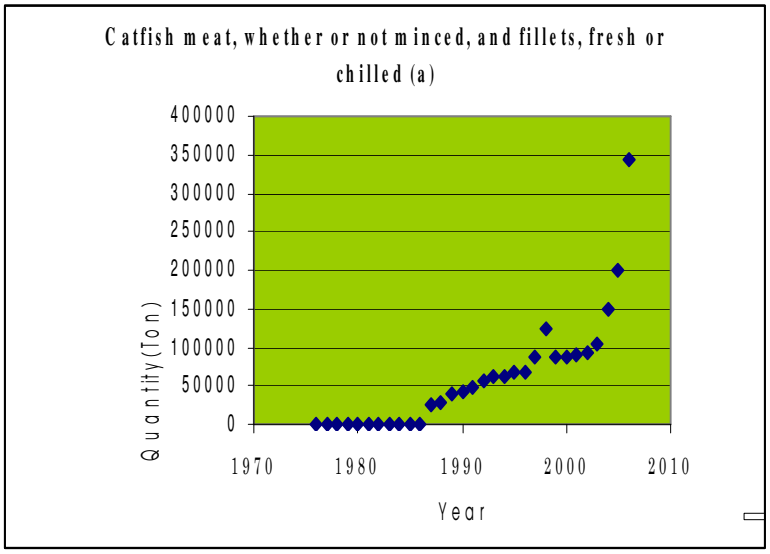


Figure 2: Year wise Global Production in Catfish [Quantity in Ton] (Meat, fillets fresh or chilled).

Figure 3: *Clarias batrachus* culture in a site of Tripura,North-East India

Sequence data and Genome projects in catfishes

Development in technologies to detect molecular variation and automation of several steps in laboratory analyses has led to the production of large amounts of genetic data. The availability of these data has stimulated the development of new statistics and computational programs, which provide an insight of data that was not previously possible (Zhang and Hewitt, 2003). Full genome sequences can be based on automated sequencer, extension of DNA primer by walking or by relatively conventional technique using bacterial artificial chromosomes (BAC). Sequence data generated from whatever protocol are regularly deposited in the global databases as the GenBank.

Among several catfish varieties *Ictalurus punctatus* has been studied in some detail. *Ictalurus punctatus* is the major aquaculture species in the United States. It is one of the six species included in the U.S. National Animal Genome Project NRSP-8. This agitation has resulted in the development of a number of genome resources in catfish including a large number of molecular markers (Kucuktas *et al* 2009), genetic linkage maps (Liu *et al* 2003, Kucuktas *et al* 2009 , Waldbieser *et al* 2001) , BAC libraries (Xu *et al* 2006,Quiniou *et al* 2003,Wang *et al* 2007) and BAC-based physical maps (Somridhivej *et al* 2008, Quiniou *et al* 2007 , Xu *et al* 2007).

Earliest report of BAC library has been established in Channel catfish *Ictalurus punctatus* (Quiniou *et al*. 2003) and has been used to construct a physical map (Quiniou *et al*. 2007). In *Ictalurus punctatus* , Male map length is 1593 , number of linkage groups 44, average distance between markers 3.8 (Liu *et al*. 2003) , in female map length 1958 , number of linkage groups 32, average distance between markers 8.7 (Waldbieser *et al* 2001). In walking catfish *Clarias macrocephalus* Male, the values are 1342, 2037, 31 and 17.1 respectively (Poompuang and Na-Nakorn, 2004).Nakomet al (1999) has isolated and characterized microsatellite loci of *Clarias macrocephlus* to study genetic diversity. Comparative mapping is a powerful tool to transfer genomic information from sequenced genomes to closely related species for which whole genome sequence data are not yet available. Successful mapping using computers is depended on availability of genetic data in a related fish species and high percentage of BLAST hits, an application globally used by bioinformaticians. It is a function of high levels of genome co linearity. Liu *et al* (2009) reported comparative analysis of catfish BAC end sequences with the zebrafish genome. This group used fully sequenced genetic map of zebrafish (*Danio rerio*) [http://www.ensembl.org/index.html] from the order Cypriniformes. Channel Catfish (*Ictalurus punctatus*) BAC Library is available online [http://bacpac.chori.org/library.php?id=103]. A recently developed Catfish Genome Consortium (http://web.uvic.ca/cbr/grasp/) has given scientists engaged in catfish research a platform for mutual sharing of data and tools. In Indian context similar consortium is much anticipated.

Bioinformatics in search of novel molecules in catfish

Molecules with significant influence in fish breeding such as inducing of gamatogenesis, maturation of gamates and larval development etc , has been studied for several species by computational biology. Search for an efficient inducer of reproductive phase related changes in fish has been a major area of interest for fish physiologists. Gonadotropin releasing hormone (GnRH) and its receptor (GnRHR), Insulin like growth factors (IGF,s) etc falls in the above criteria of being in an economical trait locus. *Clarias batrachus* and other indigenous catfishes has yet to be studied in detail with respect to novel and customized gene products.

Computational methods have helped in homology search of related gene segments, enhancers and promoters for successful expression of molecules influencing fish breeding. According to White *et al* (1998) the variants of GnRh in different species are commonly named according to species of origin, but may also be named GnRH-I, -II or -III, based on sequence alignments solely. In this connection Torgersen *et al* (2002) reported that the catfish GnRH (cfGnRH) from *Clarias gariepinus* can be termed as GnRH-I (Sequence-QHWSHGLNPG) which share much homology with Human (*Homo sapiens*) , Gilt head seabream sbGnRH (*Sparus aurata*) ,Chicken cGnRH-I (*Gallus gallus*) and Guinea pig gpGnRH (*Cavia porcellus*) , all of which are classified as GnRH-I, based on extensive sequence homology. Bogerd *et al.* (2002) has cloned GnRH receptor cDNAs for African catfish, *Claria gariepinus* and reported that there is no differences in ligand selectivity, while tissue distribution differed with highest expression in gonads , as was expected from a peptide that influence reproductive physiology. GnRH receptors are grouped in three main subtypes designated GnRHR1, GnRHR2 and GnRHR3, the latter only found in teleost fishes (Millar *et al.* 2004). In search of gene product with novel characters.Blomenro *et al* (2002) reported a computer generated model of Chimaeric gonadotropin-releasing hormone (GnRH) (Figure 4) peptides with improved affinity for the catfish (*Clarias gariepinus*) GnRH receptor. Therefore it is evident that similar peptides for Indian catfishes as *Clarias batrachus* and *Heteropneustes fossilis* must be studied , characterized and classified.

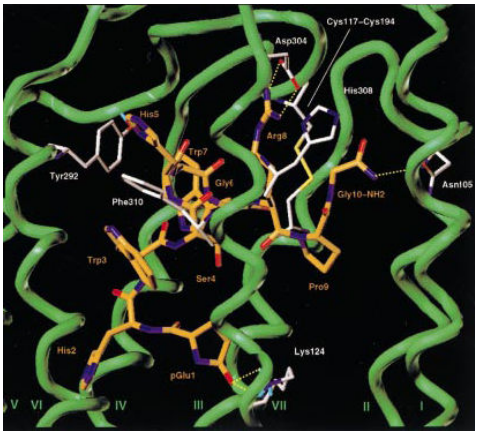


Figure: 4 High affinity GnRH receptor (Catfish) model regions with the program SYBYL 6.4, (Blomenro *et al* 2002).

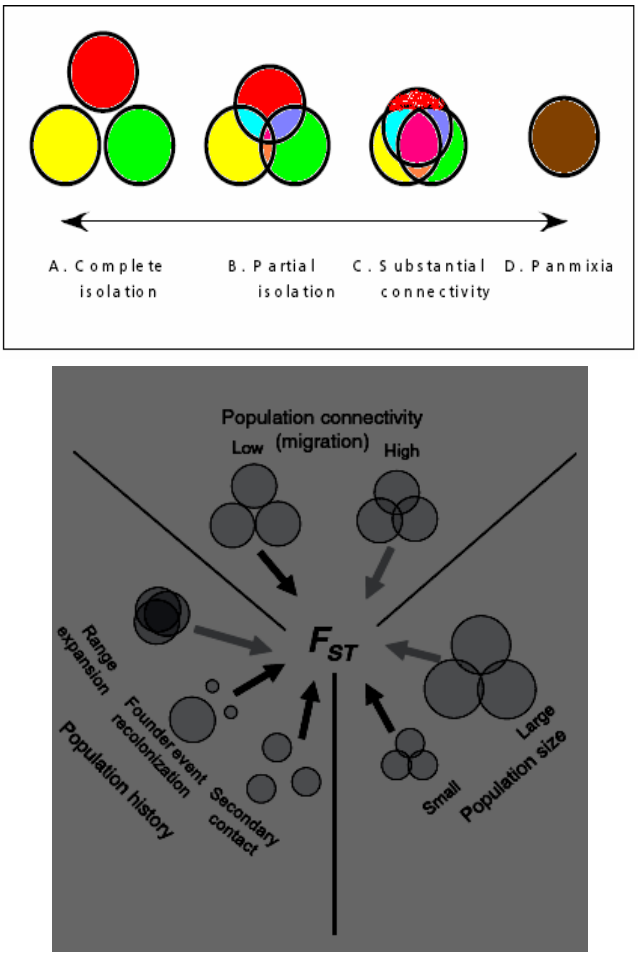


Figure: 5 (a) Effect of gene flow altering the nature of a population (After Waples & Gaggioti 2006) .Figure: 5 (b) Conceptual diagram (After Lorenz Hauser & Gary R Carvalho, 2008) showing the influence of various factors on levels of genetic differentiation (measured by FST).

Phylogenetics of catfishes

Natural population of catfish verities are constantly under pressure and wild stocks are gradually declined because of overexploitation and environmental degradation in one side and supplementation of natural populations with exotic varieties on the other. For the sake of the farmer artificially propagated individuals to sustain the fishery and widespread practice of easy hybridization is a cause of worry for conservation biologists. In this environment of high entropy where wild genes are under threat and populations are either facing dilution of genetic diversity or depression for inbreeding genetic counseling to the catfishes is an immediate need.

Bartley *et al*., (2001) have reported that interspecific fish hybrids as *Clarias gariepinus* x *Clarias macrocephalus* , etc has contributed significantly in aquaculture production.These hybridizations in catfish varieties need to be addressed properly. It is interesting to mention that biological species is conceptualized on the basis of reproductive isolation between taxa (Mayr, 1942, 1963). Wide spread hybridization may however bring about delimitation in various taxa and as a result need detail genetic analysis for species identification and phylogenetics. In morphologically similar sympatric taxons, species are defined by estimating the phylogeny of closely related populations. The phylogenetic species concept makes no reference to reproductive isolation. This again necessitates the need of proper regulation in the market by appointing regulatory authorities.

Waples & Gaggioti (2006) has described the change in population status due to gene flow from an independently, self-sustaining population towards a Panmixia (Fig:5 a) where only a single population exists with individuals (or gametes) .Hauser & Carvalho (2008) considered other factors as connectivity, history and size to propose their effect on a function F_{ST} , defined as the fraction of total genetic variation attributable to differences among populations (Fig:5 b) .

Figure 6: Wildtype Indian *Clarias batrachus* (observe coloration)

Figure 7: Exotic *Clarias* variety available in Tripura (*C. gariepinus*)

In order to study species, populations and varieties of catfishes phylogenetic analysis has been generated with available sequence data. Correlation of evolutionary events with phylogeography is also a well established area of bioinformatics research which has been exploited in some studies with respect to catfishes. Algorithms and programs catering to the need of phylogeneticist are well established and interesting outcomes are been reported.

In India studies on *Clarias batrachus* though less in terms of volume but some reports are indicating towards the global trend og genetic dilution in this species. Global distribution of the species (*Clarias batrachus*) suggests that the fish has dispersed widely and adapted itself well in many areas. In order to find ways for stock improvement and conservation for any fish species information on genetic variation is essential. According to the data released by the Fisheries and Aquaculture department, Food and Agriculture Organization of the United Nations, *Clarias batrachus* has been propagated throughout the Asia from Thailand and Indonesia (Java). The species has been introduced to as far as Europe (United Kingdom), USA and Australia (Papua New Guinea) from various pockets of South and South East . As far as the studies on population genetics, diversity and genetic makeup of *Clarias batrachus* is concerned some of the recent observations worth mentioning are as follows: Khedkar et al in 2009 studied genetic similarity and diversity of catfish *Clarias batrachus* (Linn.1758) populations collected from three regions of Indian riverine system by using randomly amplified polymorphic DNA-polymerase chain reaction (RAPD-PCR).They reported that *Clarias batrachus* population lacks genetic diversity in major riverine system of India . The wild type Indian *Clarias batrachus* genotype (blackish yellow coloration with strong jaw bones) (Fig:6) is constantly threatened by the exotic *Clarias* (gray body coloration with much flexible jawbones) (Fig:7) and the this may be due to rearing of the fish species in the same environmental conditions, migration or by inbreeding during several generations .In nearer future, the lack in genetic diversity can lead to inbreeding depression, that can result in poor growth and disease susceptibility. Islam *et al* (2007) described the genetic structure of different populations of walking catfish (*Clarias batrachus* L.) in Bangladesh and mentioned

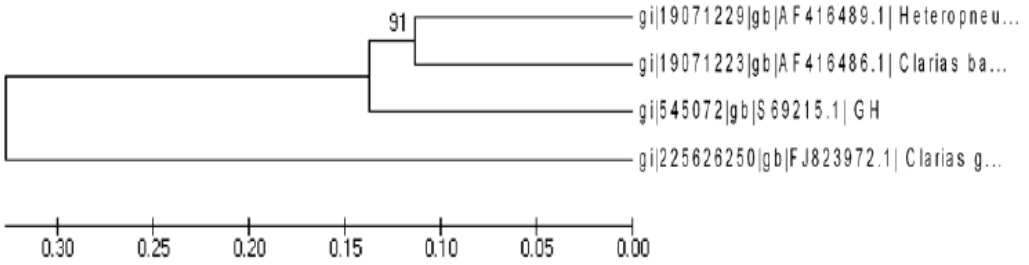


Figure 8 : Phylogenetic tree of a nuclear gene (GH) by UPGMA Method in Catfishes.

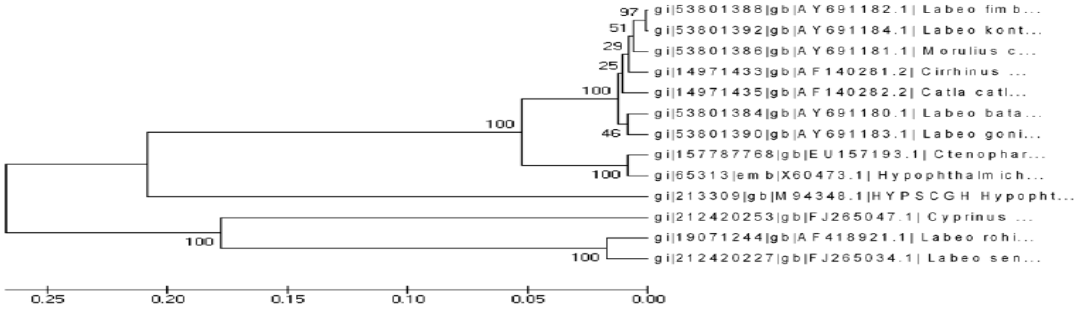


Figure 9: Phylogenetic tree of a nuclear gene (GH) by UPGMA Method in Carps.

The potentialities for improving this species through a selective breeding program. Their study revealed a recent bottleneck in some wild populations of this species which necessitates habitat protection to increase the population size and lower the risk of vulnerability of genetic dilution of this species in the future. In order to evaluate diversity Padhi *et al* (1998) characterized the MboI satellites in *Clarias batrachus*. Phylogenetic inference from the correlation of some microsatellite DNA segments for indirect assessment of genetic diversity in *Clarias batrachus* has been evaluated (Debnath and Gupta , 2009). In a study Ahmad & Hasnain in (2006) reported correlation between biochemical properties and adaptive diversity of skeletal muscle myofibrils and myosin of some air-breathing teleost including *Clarias batrachus*.

Comparative phylogeny is a simple technique by which degree of divergence of a locus can easily be studied. Phylogenetic relationship of a nuclear gene with economical traits (Growth Hormone) in catfishes (Fig 8) and carp fishes (Fig 9) (Debnath 2009) can be mentioned in this context.Nucleic acid data were retrieved from Genbank and aligned by ClustalW inbuilt in MEGA,4. Bootstaps reveal the degree of homologyand represent it in evolutionary timescale. In the dendrogram (UPGMA) the genus *Labeo* showed dispersion throughout the studied gene population although less than 100 bootstrap values indicate at least 9 species of the studied population has much structural similarity. All available sequences for growth hormone in catfishes, ie *Clarias batrachus*,*C gariepinus*, a haplotype and *Heteropneustes fossilis*, are conserved with minimal divergence.

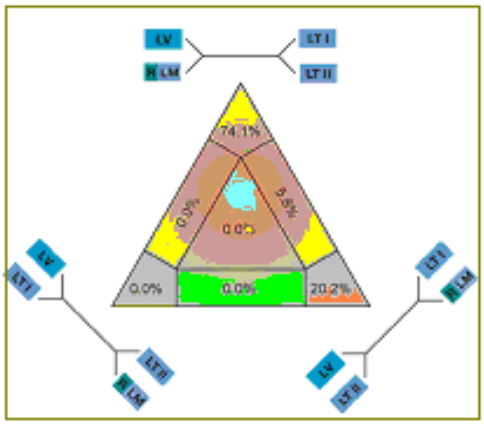


Figure 10: A four-cluster likelihood mapping analysis using the program TREE-PUZZLE 5.1(Applied to test the hypothesis of the monophyletic of the two endemic lake clades of catfishes) (Koblmüller et al, 2006).

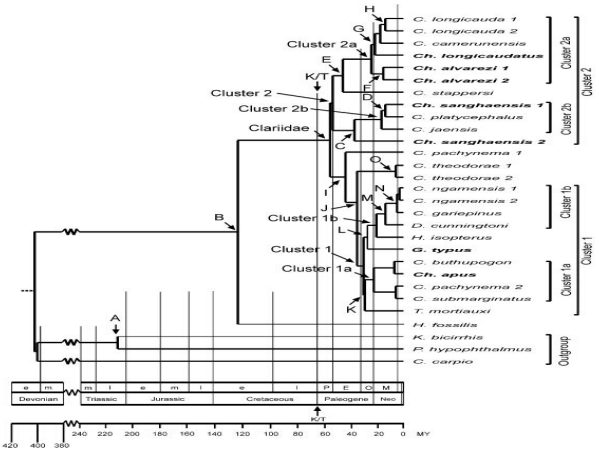


Figure 11: Dated phylogenetic tree of African Clariidae obtained from semiparametric rate smoothing. (Jansen et al, 2006)

The old siluriform *Clarias* since originated in the old African continent as the phylogeography suggests, several studies in this genus has been reported from Africa .Koblmüller *et al* (2006) studied mitochondrial phylogeny and phylogeography of East African squeaker catfishes (Siluriformes: *Synodontis*) and proposed that that the ancestral lineage of today's East African squeaker catfish fauna has colonized the area before the Great East African Lakes have formed. In this study (Fig 10) Koblmüller *et al* used Shimodaira- Hasegawa test to observe significant differences between the topologies obtained by the different tree building algorithms. Jansen *et al* (2006) has studied Phylogenetic relationships and divergence time estimate of African anguilliform catfish (Siluriformes: Clariidae) using ribosomal genes and spacer sequences .The phylogenetic tree established by this team is given in the figure 11.

Pouyaud *et al* (2008) studied the phylogenetic structure of habitat shift and morphological convergence in Asian *Clarias* .They have inferred from the analysis of habitat transitions with Markovian models that Asian *Clarias* evolved from an ancestor probably living in clear waters in Africa.These valuable findings are encouraging and dated phylogenetics and phylogeographical studies in Indian catfishes are therefore need to be carried out.

Expression pattern studies in cat fishes

Microarrays can be utilized for screening for polymorphisms throughout the genome and have virtually countless applications from studying developmental stage related gene expression and even in studing population genetics (Luikart *et al*. 2003). Different expression pattern in microarrays may help to identify samples from different localities and differing in functional traits (Rise *et al*. 2004a).Development of expressed sequence tags, which can be done in high throughput by microarray, can bring about dynamic changes in catfish genetics. According to the review of Liu and Cordes (2004) the value of EST resources is perhaps underestimated currently in the aquaculture genetics community, primarily because of the lack of bioinformatics capabilities. This kind of scientific bottlenecks must be trounced. A greater emphasis on applications of bioinformatics in aquaculture genetics/genomics is inevitable, and it is expected that various EST databases will serve as rich sources of genomic information not only for aquaculture geneticists, but also for aquaculture physiologists, immunologists, biotechnologists, and the like. Microarrays have been developed for catfish (Li and Waldbieser 2006) and applications can be as versatile as studying developmental stage specific gene expression and regulatory networks. In studying host parasite interaction Peatman *et al*. (2007) reported the acute phase response following infection of catfish with *Edwarsiella ictaluri*, causing enteric septicemia .It was studied by high density in situ oligonucleotide microarray. Several microsatellites were associated with resistance and susceptible phenotypes. These markers have been incorporated in the catfish linkage map, that will facilitate finding resistance determining QTL and will help in the development of MAS (Marker Associated Selection) programs. Bao *et al*. (2006) identified 26 chemotactic cytokine genes, sequenced them and studied their expression in catfish. In the catfish *Ictalurus punctatus* several hundred thousands of ESTs [Serapion *et al* (2004) , He *et al* (2003) , Liu, unpublished data , Li *et al* (2007) , Steinke *et al* (2006)], microarray platforms (Li and Waldbieser 2006, Liu *et al* 2008 , Peatman *et al* 2007 , 2008, Sha *et al* 2009) has been generated.

Computational Morphometrics in catfish

Selection pressure moulds an organism in both macro and micolevel.Morphological peculiarities are nothing but a function of fitness acquired in the phylogenetic history of an animal. Computational morphometry has gained substantial popularity due to its easy operation, time tested theories and exciting applications. Agne` se *et al*. (1997) examined morphological and genetic variation in sympatric populations of *Clarias gariepinus* (Burchell) and *Clarias anguillaris* L. using morphometry, as one of the parameters. Taxonomic characterization can be reliably done by morphometric measurements (Moyle & Cech, 2000). Basic features as Length – weight relationship in fishes may also reflect differences in taxon and population. The length-weight relationship can also be used in setting yield equations for estimating the number of fish landed and comparing the population in space and time (Beverton and Holt 1957). Adriaens & Verraes (2002) studied regression of the overall shape changes in the bony neurocranium in a sample of the larval catfish *Clarias gariepinus* (Teleostei: Clariidae), during a limited period of

ontogeny, to size and age. Regression studies using statistical analysis tool box of MATLAB©, 7 in a local exotic *Clarias* species (log W = - 0.1617 + 1.699 log L) (Figure 12) compared to the wild *Clarias batrachus (*log W = - 0.8628 + 2.097 log L.*)* of the same length groups showed distinct features of the two populations. The condition factor indicate higher increment of body mass in exotic variety (Figure 13) (Debnath,2008) in the same rearing condition. Catfish phylogeny and morphometrics have generated valuable insight in on the overall development of the group and how they have evolved. In an interesting report Pouyaud et al (2008) suggested a complex pattern of morphological evolution in Asian *Clarias* by analyzing morphometrics. The variety displays a repeated convergence towards an elongated morphological form in black waters. This is so apparent that morphometrics is a superior contrivance for studying developmental evolutionary peculiarities.

Tools used in catfish Bioinformatics research

Testing of hypothesis, generation of new variables and simulation of situations as hydrophobic interactions during modeling of a peptide etc are possible during research. Most programs are available online whereas some are custom-made. Automation and regular development in the fine tuning of data analysis with a satisfactory solution in a reasonable time (Agard and Kusiak, 2004; Ghosh and Nath, 2004; Li and Wang, 2004) is the goal of Bioinformatics. Some software's which have been used in studying catfish biology is mentioned here.

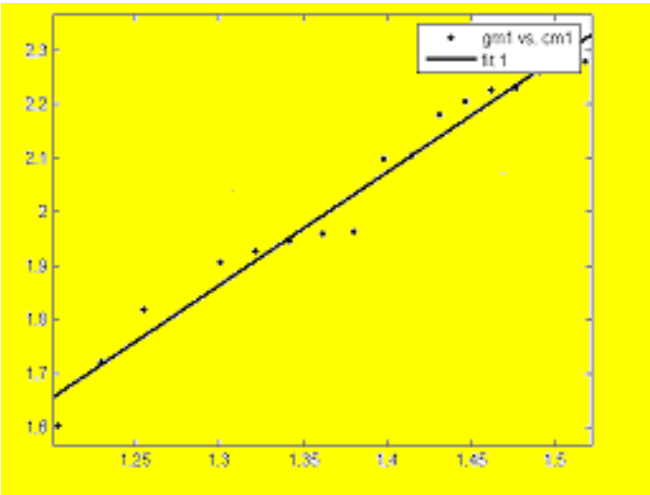


Figure 12: Regression Curve of wild *Clarias batrachus* .

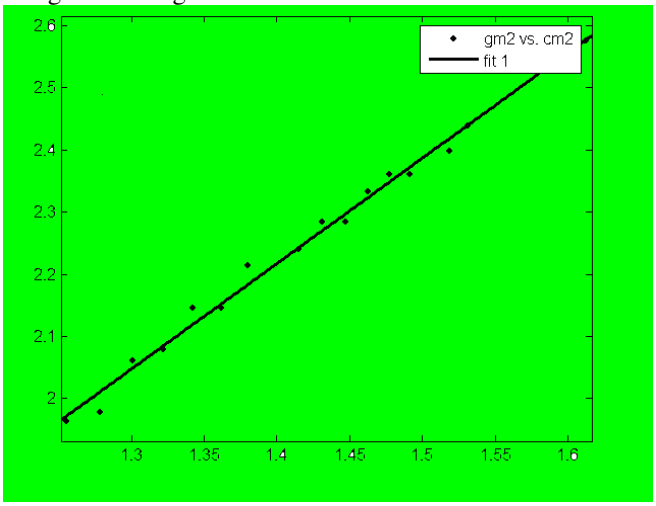
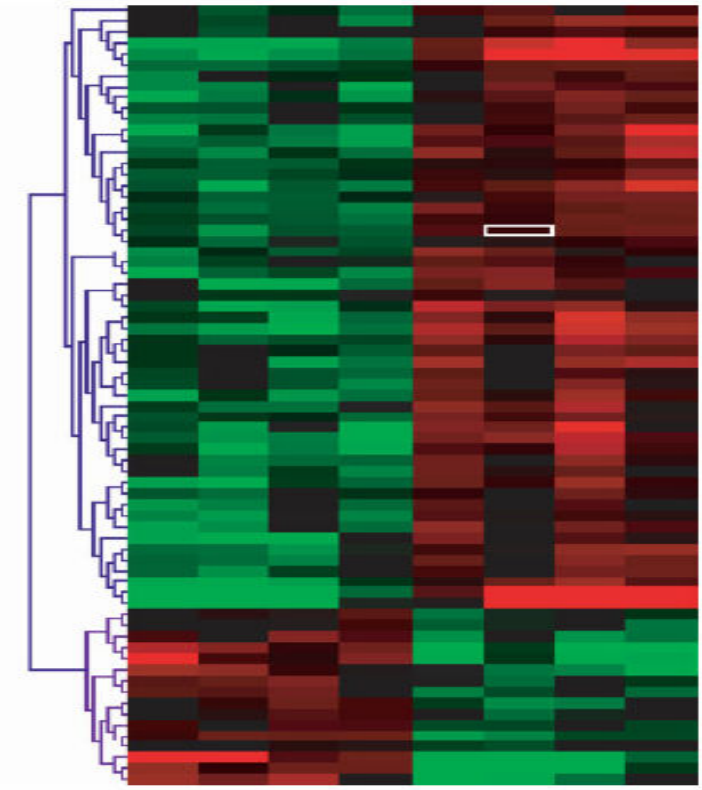


Figure 13: Regression Curve of exotic *Clarias variety*. Figure



14: Microarray data coupled with Phylogenetics of the expressed genes. TIGR Multiple Experiment Viewer software; The Institute for Genomic Research, J. Craig Venter Institute, Rockville, MD, USA) (After Panserat *et al* 2008),

For sequence analysis and alignment several bioinformatics tools have been in use. Sequence homology search made by BLASTN (for nucleotide similarity) and BLASTX .Sequence search for microsatellite repeats can be done with Sputnik 2 computer software (http://espressosoftware.com/pages/sputnik.jsp). Richardson *et al* in (2006) used MBEToolbox (Cai *et al* 2005) of MATLAB (MathWorks) package for sequence analysis which contain an alignment tool command as alignSeqFile. Nucleotide sequences can be aligned using CodonAlign2.0 (Hall, 2004). DNASTAR software (DNASTAR Inc.) has also been used by several workers for sequence edition and alignment.

In order to study population genetics in catfishes several tools have been used. Hoh *et al* (2007) for isolation and Development of DNA Microsatellite Markers in the River Catfish (*Mystus nemurus*) used PCR generated data to analyze population genetics characters using the POPGENE (ver 1.32) computer software (Yeh and Boyle 1997).These markers will eventually be used in the construction of a genetic linkage map, and in testing for associations with Quantitative Trait Loci (QTL) so that markers assisted selection (MAS) may become feasible in breeding programs. Kiss and Cheng (2008) in astudy on molecular diversity and genomic organisation of the α, β and γ eye lens Crystallins proteins of several fishes including *Clarias batrachus* and *Clarias fuscus* used in BioEdit v7.0.5.2 (Hall, 1999) with the inbuilt tools like ClustalW and amino acid identity matrices (IDENTIFY, BLOSUM62, GONNET). Nasren *et al* (2009) in a study to asses genetic variation in random microsattelite loci for stinging catfish *Heteropneustes fossilis* used Arlequin V 3.0 for estimation of allelic variations, observed and expected heterozygocity values, deviation from Hardy-weinberg equilibrium, Fst values for population differentiation and analysis of molecular varience. Nei,s genetic distance and geneflow among the populations were studied by Genalex V 6.1. Ryynänen and Primmer (2006) studied events of natural selection during teleost

growth hormone gene evolution including *Clarias.* SignalP v. 3.0 (Bendtsen *et al*. 2004) tool was used to predict Putative signal peptide cleavage sites for each teleost datasets. SWAPSC (Fares 2004) software is used to infer a statistically optimum codon window size based on simulated data and the software slides it along the alignment for each branch in the phylogeny separately. Posada and Crandall in 1998 developed MODELTEST, a program for testing the model of DNA substitution.

For protein sequence analysis studies several softwares has been used. Panprommin *et al* (2007) in a study to characterize Expressed Sequence Tags (EST) from liver and muscle tissues of walking Catfish, *Clarias macrocephalus* used GENETYX Version 7.0 to delete vector and polylinker sequences and to check for the quality of genetic data. Similarly sequence chromatograms can be edited and assembling of consensus sequences can be done by using the Sequencher software package (Gene Codes, Ann Arbor, MI). In the Panprommin (2007) study , protein sequence similarities between different forms of GnRHR from some species were calculated with GeneDoc (Nicholas *et al.* 1997). The protein motifs were identified using the Prosite database (Falquet *et al.* 2002). The putative transmembrane domains (TMs) were determined with TMHMM 2·0 (Krogh *et al.* 2001). PAUP software (version 4·0, Swofford 2002) was used to generate Parsimonious consensus trees.

For phylogenetic research several Bioingformatics tools have been used. Moncaut *et al* (2005) studied gonadotropin-releasing hormone receptors (GnrhR) in a teleost and identified the phylogeny of the peptide with the similar peptides from other fishes including catfish *Clarias gariepinus.* Multisequence alignments of the full-length proteins were carried out using CLUSTAL X (Thompson *et al.*1997). Among several algorithms and packages for computation Huelsenbeck and Ronquist (2001) developed MRBAYES a program to study Bayesian inference of phylogeny. Computational algorithms for phylogenetic analysis as Maximum Parsimony (MP),Neighbour Joining (NJ), Maximum Likelyhood (ML)and UPGMA etc can be generated by PHYML (Guindon and Gascuel 2003) ,PHYLIP (Felsenstein 1989), ProtTest 1.2 (Abascal et al. 2005) etc. Alternative tree topologies can be compared applying the approximately unbiased test as implemented in the CONSEL package (Shimodaira and Hasegawa 2002), using the sidewise likelihood values estimated by PAML (Yang 1997). The TREE-PUZZLE tool based on a maximum likelihood phylogenetic analysis using quartets and parallel computing was developed by Schmidt *et al* in (2002). Relative-rate tests between groups of sequences on a phylogenetic tree can be studied by RRTree program developed by Rechavi et al (2000). Maddison and Maddison (2005) developed a modular system for evolutionary analysis named Mesquite.Divergence times for different teleost lineages from their most recent common ancestor were estimated using a Bayes MCMC package (Thorne and Kishino 2002). In another approach during construction of phylogenetic trees nucleotide or amino acid insertions and deletions can be coded as binary characters and can be added to the matrix, following the method described by Barriel (1994) and implementation can be done by the software BARCOD (http://wwwabi.snv.jussieu.fr/people/billoud/). Bremer support indices (BSI), a concept similar to bootstrapping for each node of a dendrogram can be calculated by the software TreeRot (Sorenson, 1999). Among the phylogenetic tools mentioned MEGA, version 4 developed by Tamura *et al* (2007) is very much user-friendly and has several inbuilt algorithms for alignment as CLUSTAL W and BLOSUM. All common dendrogram generating methods are also inbuilt.

Proteomics has also generated a plethora of software application for studying gene expression patterns and several post translational modification modeling .Amino acid sequences can be classified by homology using BLASTP (Swiss Institute of Bioinformatics, Basel) conserved-domain analysis and manual inspection. Hansen *et al* (2005) used similar approach to break down peptide sequences into individual Ig superfamilial C domains in a study of B cell development .The group reported a unique Ig heavy-chain isotype (IgT) in rainbow trout. In this study transmembrane regions were identified by using TMPRED (www.chnet.org), and N-linked glycosylation sites were predicted by using the NETNGLYC 1.0 server (www.cbs.dtu.dk/ services/NetNGlyc).Similar studies in catfishes should be done expecting novel discoveries.

To analyze the massive gene expression data generated by microarray technology GeneSight® is an efficient data mining, visualization, and reporting tool. It has

applications like GenePieTM visualization, 2-D scatter plots; interactive ratio histogram plotting, hierarchical, K-means, and neural network clustering, principal components analysis. The confidence analyzer tool can use replicated gene expression data for identifying genes having expression patterns which can differentiate between classes of experimental conditions such as disease states. BioArray Software Environment (BASE) database as well as TIGR Multiple Experiment Viewer software (TMEV; The Institute for Genomic Research, J. Craig Venter Institute, Rockville, MD, USA), (Fig 14) can anlyze micro array data (Panserat *et al* 2008). GoMiner software (Genomics and Bioinformatics Group, National Institutes of Health, Bethesda, MD, USA; http://discover.nci.nih. gov/gominer/) can be used for organization of genes for biological interpretation in the context of gene ontology. For classical genomic experiments as visualizing of stained nucleic acid and protein, several softwares are available. Frequently used methods like Gel documentation and analysis of nucleic acid can be done by the software DNAfrag, version 3.03.

Morphometric data in fishery research is an important tool has also been studied with several computer softwares. These programs were mainly used in normalizing data land to determine several variables as Euclidean distances. Tseng *et al* (2009) used PRIMER 5 (www.primer-e.com) for cluster analysis based on a single linkage mode (Clarke & Gorley, 2001). Morphometric was examined with a principal component analysis (PCA) using the SYN-TAX 2000 (Podani, 2001).

The above discussion noticeably shows that studies on indigenous catfish varieties are scanty and Indian Bioinformaticians has therefore ample prospect to engage in analyzing the genetic wherewithal of Indian Catfishes and develop newer applications to see the sights of interesting correlations.

FUTURE COURSE OF ACTION FOR INDIAN CATFISH RESEARCH IN BIOINFORMATICS: Overcoming the scientific bottlenecks

Catfishes available in African and American subcontinent are most well studied. Therefore our knowledge can be biased by the studies reflecting specific geographical types and species.Genbank database has scanty information on two genus of catfishes available in India , as *Clarias* and *Heteropneustes* (Table 1). There are 29 species for which records are available in *Clarias* while two species of *Heteropneustes* are recorded in Genbank.Many of the *Clarias* species have very little available information, (Table 2) and Genome project of only *Clarias fuscus* has been reported in NCBI and only one species , ie *Clarias gariepinus* has 04 GSS data . Species with less than 5 entries in Protein or Nucleotide heads are neglected but *Heteropneustes microps* has been an exception in this case. A further review in existing literature has pointed out that studies on microarray development for Indian varieties have not been intiated.Modelling of novel peptides and *in-silico* studies on molecular interactions as in hormone-receptor or ligand-receptor in Indian context must be developed. The genepool of the Indian fish stocks must be conserved from dilution. In recent years many steps have been taken by the Indian regulatory authorities against the import of exotic fish varieties and their introduction in the wild.Besides largescale industrial hybridization are also discouraged with much stringency.Genome projects of Indian catfishes should be collaboratively initiated among several laboratories to increase economy of the project. Several national level institutes are engaged in developing indigenous techniques for yield increment and revenue generation, similarly high end research is also going on .India's human resource in the IT industry can be harnessed to develop computational packages with easy user interface and functions relevant for Indian milieu. Goverment funding though promotes new technologies but proper implementation and dedicated taskforce can bring about the much needed paradigm shifts in bioinformatics research in India. By convention bioinformatics tools are written by software developers and therefore the view of the biologists towards a problem may be not visualized by the developer. Greater cooperation of interdisciplinary fields is consequently obligatory. Conserving Indian fish genepool yet maximizing yield increase is the contemporary challenge to which active Bioinformatics research can counter. For this major studies in catfish genetics integrating the wet labs and computation is needed. In this connection the work of Matthew Et al (2004) worth mention in which this group undertook a major study to develop EST Database and cDNA Microarray in Salmonid and applied them to

understand interspecific hybridization characteristics. Canadian government genome initiative called the Genome Research on Atlantic Salmon Project (GRASP) accelerated efforts of DNA marker development and genome mapping, similar initiatives are highly expected in India too.

Table 1: Recent updates of NCBI database on two major catfish genus available in India.

Gen	DNA	Exprsd. Seq. Tags (EST)	Gen.Su r.Seq. Rec. (GSS)	Protein	Uni STS Markr/ Map data	Popln. data (popSet)	Exprsn./mo l.abndnce Pro. (Geo profiles)	Depo. Chem. Sub. Rec.	Gen.P roj.	Exp. Set. Geod ata
Clarias	610	1934	04	218	30	27	01	02	01	00
Heteropneustes	061	059	00	026	00	07	04	01	00	06

Table 2: Species wise data available on Genbank

Species	Taxonomic ID	Nucleotides	Nucleotide EST	Protein	Popsets	UniSTS
Clarias batrachus	59899	70	1923	23	07	27
Clarias fuscus	33541	22	00	11	02	00
Clarias gariepinus	13013	191	00	94	01	02
Clarias liocephalus	391045	06	00	06	01	00
Clarias macrocephalus	35657	153	02	01	02	01
Clarias maurus	222699	10	00	10	00	00
Clarias meladerma	222698	12	00	12	00	00
Heteropneustes fossilis	93621	60	00	26	07	00
Heteropneustes microps	575559	01	00	00	01	00

REFERENCES

Adriaens D, Verraes W (2002) An Empirical Approach to Study the Relation between Ontogeny, Size and Age Using Geometric Morphometrics. Topics in Functional and Ecological Vertebrate Morphology, pp. 293-324.m P. Aerts, K. D'Août, A. Herrel & R. Van Damme, Eds. © Shaker Publishing, ISBN 90-423-0204-6

Agne se JF, Teugels GG, Galbusera P, Guyomard R (1997) Morphometric and genetic characterization of sympatric populations of *Clarias gariepinus* and *C. anguillaris* from Senegal. Journal of Fish Biology 50, 1143–1157.

Bao B, Peatman E, Xu P, Baoprasertkul P, Wang G, Liu Z (2006) Characterization of 23 CC chemokine genes and analysis of their expression in channel catfish (*Ictalurus punctatus*). Dev. Comp. Immunol. 30, 783-796.

Bartley DM, Rana K, Immink AJ (2001) Interspecific hybrids in aquaculture and fisheries. Rev. Fisheries and Fish Biol., 10: 325-337.

Beverton RJH, Holt SJ (1957) On the dynamics of exploited fish populations.Fisheries investigation. London, Series 2: 533.

Blomerno Marion HR, Ter Laak T, KU Ronald H, Beyermen M, Hund E, Bogerd J, Leurs R (2002) Chimaeric gonadotropin-releasing hormone (GnRH) peptides with improved affinity for the catfish (*Clarias gariepinus*) GnRH receptor. Biochem. J. 361, 515-523

Bogerd J, Bas Diepenbroek W, Hund E, van Oosterhaut F, Teves AC, Leurs R, Blomenröhr M (2002) Two gonadotropin-realising hormone receptor in the African catfish: No differences in ligand selectivity, but differences in tissue distribution. Endocrinology 143 4673–4682.

Cai JJ, Smith DK, Xia X, Yuen K (2005) MBEToolbox: a MATLAB toolbox for sequence data analysis in molecular biology and evolution. BMC Bioinformatics, 6, 64.

CHORI-212: Channel Catfish (*Ictalurus punctatus*) BAC Library. http://bacpac.chori.org/library.php?id=103

Debnath S (2008) Growth Economics and Biometric Comparison of Length Weight Relationship in Two Intra Generic

Varieties of *Clarias* Isolated from Dumbur Lake, in a Rural District of Tripura, India., XI Manipur science Congress, Manipur University, March 28-30,

Debnath S (2009) Phylogenetic Correlation among Carpfishes and Catfishes using a Nuclear Protein coding Genetic Data (GH) retrieved from GENEBANK using a freeware Computational Evolutionary Biology package MEGA. International Conference on Biotechnological Solution for Environmental Sustainability, Vellore Institute of Technology, Vellore, Tamil Nadu, October 21-23,

Debnath S (2009) Traditional Consumption of Magur (*Clarias batrachus*) an Air Breathing Catfish among the Population of North Eastern India is Rationalized by Its Blood Lipid Parameters. CHEMFERENCE, Annual Research Symposium, IIT Madras 22 - 23rd August, 2009.

Debnath S, Gupta R (2009) Phylogenetic Inference from the Correlation of Some Microsatellite DNA Segments for Indirect Assessment of Genetic Diversity in the Asian Catfish, Clarias batrachus" .The ICFAI Journal of Computational Mathemetics.,Vol. II, No. 2, 2009.

Devaere S, Jansen G, Adriaens D, Weekers P (2007) Phylogeny of the African representatives of the catfish family Clariidae (Teleostei, Siluriformes) based on a combined analysis: independent evolution towards anguilliformity.Journal compilation, Blackwell Verlag, Berlin

Hansen John D, Landis Eric D, Phillips Ruth B (2005) Discovery of a unique Ig heavy-chain isotype (IgT) in rainbow trout: Implications for a distinctive B cell developmental pathway in teleost fish, PNAS, 102, 19, 6919–6924

He C, Chen L, Simmons M, Li P, Kim S, Liu ZJ (2003) Putative SNP discovery in interspecific hybrids of catfish by comparative EST analysis. Anim Genet.; 34:445–448.

Jansen G, Devaere S, Weekers PHH, Adriaens D (2006) Phylogenetic relationships and divergence time estimate of African anguilliform catfish (Siluriformes: Clariidae) inferred from ribosomal gene and spacer sequences Molecular Phylogenetics and Evolution 38 , 65–78

Koblmüller S, Sturmbauer C, Verheyen E, Meyer A, Salzburger W (2006) Mitochondrial phylogeny and phylogeography of East African squeaker catfishes (Siluriformes: Synodontis),*BMC Evolutionary Biology*, 6:49 doi:10.1186/1471-2148-6-49

Kucuktas H, Wang S, Li P, He C, Xu P, Sha Z, Liu H, Jiang Y, Baoprasertkul P, Somridhivej B, Wang Y, Abernathy J, Guo X, Liu L, Muir W, Liu Z (2009) Construction of Genetic Linkage Maps and Comparative Genome Analysis of Catfish Using Gene-associated Markers. Genetics.; 181:1649–1660.

Kucuktas Serapion J, Feng H, J, Liu Z (2004) Bioinformatic mining of type I microsatellites from expressed sequence tags of channel catfish (*Ictalurus punctatus*) Mar Biotechnol (NY);6:364–377

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.

Li RW,Waldbieser GC (2006) Production and utilization of a highdensity oligonucleotide microarray in channel catfish, *Ictalurus punctatus*. BMC Genomics 7, 134.

Liu Hong, Jiang Yanliang, Wang Shaolin, Ninwichian Parichart, Somridhivej Benjaporn, Xu Peng, Abernathy Jason, Kucuktas Huseyin, and Liu Zhanjiang (2009) Comparative analysis of catfish BAC end sequences with the zebrafish genome. BMC Genomics. ; 10: 592.

Liu Z (2003) A review of catfish genomics: progress and perspectives. Comp Funct Genomics.; 4:259–265.

Liu Z (2008) In: Genome Mapping and Genomics in Fishes and Aquatic Animals. Kocher T, Kole C, editor. Vol. 2. Berlin Heidelberg: Springer Catfish; pp. 85–100. [Kole C (Series Editor): Genome Mapping and Genomics in Animals, vol. 2]

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.

Liu Z, Li RW, Waldbieser GC (2008) Utilization of microarray technology for functional genomics in ictalurid catfish. J Fish Biol.; 72:2377–2390.

Liu ZJ (2007) Expression analysis of the acute phase response in channel catfish (*Ictalurus punctatus*) after infection with a Gram negative bacterium. Dev. Comp. Immunol. [doi.10.1016/j.dci.(2007).03.003].

Lorenz H, Gary RC (2008) Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. Fish and Fisheries, , 9, 333–362

Lorenz H, James ES (2008) Advances in molecular technology and their impact on fisheries genetics , Fish and Fisheries, 9, 473–486.

Millar RP, Lu ZL, Pawson AJ, Flanagan CA, Morgan K, Maudsley SR (2004) Gonadotropin-releasing hormone receptors. *Endocrine Reviews* 25 235–275.

Moncaut N, Somoza G, Power DM, Canário Adelino VM (2005) Five gonadotrophin-releasing hormone receptors in a teleost fish: isolation, tissue distribution and phylogenetic relationships, Journal of Molecular Endocrinology 34, 767–779

Msatfinder: detection and characterisation of microsatellites. http://www.genomics.ceh.ac.uk/msatfinder/

Na-Nakom U, Taniguchi N, Nuguroho E, Seki S, Kamonrat W (1999) Isolation and characterization of microsatellite loci of *Clarias macrocephlus* and their application to genetic diversity study. Fisheries Sci 65:520–526.

Nasren S, Nazrul Islam M, Golam KKM, Shahidul Islam Md, Alam Md (2009)Genetic variation and differentiation in the stinging catfish *Heteropneustes fossilis* (Bloch), populations assessed by heterologous microsattelite DNA markers.Indian J.Biotechnol.

Panprommin D, Poompuang S, Srisapoome P (2007) Characterization of Expressed Sequence Tags from Liver and Muscle Tissues of Walking Catfish, *Clarias macrocephalus,* ScienceAsia 33: 461-468

Panserat S, Kolditz C, Richard N, Plagnes-Juan E, Piumi F, Esquerre´ D, Me´dale F, Corraze G, Kaushik S (2008)Hepatic gene expression profiles in juvenile rainbow trout (*Oncorhynchus mykiss*) fed fishmeal or fish oil-free diets. British Journal of Nutrition, 100, 953–967

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. Dev Comp Immunol.;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. Mol Immunol.;45:553–566.

Poompuang S, Na-Nakorn U (2004) A preliminary genetic map of walking catfish (*Clarias macrocephalus*). Aquaculture, 232, 195- 203.

Quiniou SMA, Katagiri T, Miller NW, Wilson M, Wolters WR, Waldbieser GC (2003) Construction and characterization of a BAC library from a gynogenetic channel catfish, *Ictalurus punctatus*. Genet. Sel. Evol. 35, 1–11.

Quiniou SMA, Waldbieser GC, Duke MV (2007), A first generation BAC-based physical map of the channel catfish genome. BMC Genomics 8, 40.

RepeatMasker Open-3.0. http://www.repeatmasker.org

Richardson Devid, Vanwye Jeffrey D, Exum Amy M, Cowen Robert K, Crawford Doughlus L (2006) High-throughput species identification: from DNA isolation to bioinformatics. Molecular Ecology Notes, 10.1111/j.1471-8286.(2006).01620.x

Schmidt HA, Strimmer K, Vingron M, von Haeseler A (2002) TREE-PUZZLE: a maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics, 18:502-504.

Sha Z, Abernathy WJ, Wang S, Li P, Kucuktas H, Liu H, Peatman E, Liu Z (2009) NOD-like subfamily of the nucleotide-binding domain and leucine-rich repeat containing family receptors and their expression in channel catfish. Dev Comp Immunol.;31:991–999.

Shimodaira H, Hasegawa M (2002) Multiple comparisons of log-likelihoodsn with

applications to phylogenetic inference. *Mol Biol Evol*, 16:1114-1116.

Somridhivej B, Wang S, Sha Z, Liu H, Quilang J, Xu P, Li P, Hu Z, Liu Z (2008) Characterization, polymorphism assessment, and database construction for microsatellites from BAC end sequences of channel catfish (*Ictalurus punctatus*): A resource for integration of linkage and physical maps. Aquaculture.;275:76–80.

Steinke D, Salzburger W, Meyer A (2006) Novel relationships among ten fish model species revealed based on a phylogenomic analysis using ESTs. J Mol Evol.; 62:772–784.

Torgerse J, Nourizadeh-Lillabadi R, Husebye H, Aleström P (2002) *In silico* and *in situ* characterization of the zebrafish (*Danio rerio*) *gnrh3* (sGnRH) gene.*BMC* Genomics (2002), 3:25

Waldbieser GC, Bosworth BG, Nonneman DJ, Wolters WR (2001) A microsatellite-based genetic linkage map for channel catfish, *Ictalurus punctatus*. Genetics. 158:727–734.

Wang S, Xu P, Thorsen J, Zhu B, de Jong PJ, Waldbieser G, Kucuktas H, Liu Z (2007) Characterization of a BAC library from channel catfish *Ictalurus punctatus*: indications of high levels of chromosomal reshuffling among teleost genomes. Mar Biotechnol (NY); 9:701–711.

Waples RS, Gaggioti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Molecular Ecology, 15, 1419–1439.

White RB, Eisen JA, Kasten TL, Fernald RD (1998) Second gene for gonadotropin-releasing hormone in humans. Proc Natl Acad Sci *USA*, 95:305-309

Xu P, Wang S, Liu L, Peatman E, Somridhivej B, Thimmapuram J, Gong G, Liu Z (2006) Channel catfish BAC-end sequences for marker development and assessment of syntenic conservation with other fish species. Anim Genet.; 37:321–326.

Xu P, Wang S, Liu L, Thorsen J, Kucuktas H, Liu Z (2007) A BAC-based physical map of the channel catfish genome. Genomics.; 90:380–388.