



Bioinformatics: scope and challenges in aquaculture research of Bangladesh- a review

K.M.S. Rana^{1*}, K. Ahammad^{1,2} and M.A. Salam¹

Received 31 October 2020, Revised 14 December 2020, Accepted 24 December 2020, Published online 31 December 2020

ABSTRACT

Bioinformatics is one of the ongoing trends of biological research integrating gene based information and computational technology to produce new knowledge. It works to synthesize complex biological information from multiomics data (results of high throughput technologies) by employing a number of bioinformatics tools (software). User convenience and availability are the determining factors of these tools being widely used in bioinformatics research. BLAST, FASTA (FAST-All), EMBOSS, ClustalW, RasMol and Protein Explorer, Cn3D, Swiss PDB viewer, Hex, Vega, Bioeditor etc. are commonly operated bioinformatics software tools in fisheries and aquaculture research. By default, these software tools mine and analyze a vast biological data set using the available databases. However, aquaculture scientists can use bioinformatics for genomic data manipulation, genome annotation and expression profiling, molecular folding, modeling, and design as well as generating biological network and system biology. Therefore, they can contribute in specified fields of aquaculture such as disease diagnosis and aquatic health management, fish nutritional aspects and culture-able strain development. Although having huge prospects, Bangladesh is still in infancy of applying bioinformatics in aquaculture research with limited resources. Research council at national level should be formed to bring all the enthusiastic scientists and skilled manpower under a single umbrella and facilitate to contribute in a collaborative platform. Besides, fully-fledged bioinformatics degree should be launched at University levels to produce knowledgeable and trained work force for future research. This review was attempted to shed light on bioinformatics, as young integrated field of bio-computational research, and its significance in aquaculture research of Bangladesh.

Keywords: Bioinformatics, Aquaculture, Bangladesh.

¹Department of Aquaculture, Bangladesh Agricultural University, Mymensingh, Bangladesh.

²Fisheries Laboratory, University of Tokyo, Hamamatsu, Shizuoka, Japan.

*Corresponding author's email: ranakms@bau.edu.bd (K.M.S. Rana)

Cite this article as: Rana, K.M.S., Ahammad, K. and Salam, M.A. 2020. Bioinformatics: scope and challenges in aquaculture research of Bangladesh- a review. *Int. J. Agril. Res. Innov. Tech.* 10(2): 137-145. <https://doi.org/10.3329/ijarit.v10i2.51587>

Introduction

Bioinformatics, a new era in computational biology, was first introduced to refer the study of information processes in biotic systems (Hesper and Hogeweg, 1970). Since then this branch of science has undergone rigorous development with advanced genetic and networking technology. Today's scientists view bioinformatics as an interdisciplinary field of life sciences for conceptualizing biological systems by integrating biological information with computer based informatics technique. By nature, bioinformatics aims at multifaceted biological problems solving, ranging from molecular biology to physiological processes by the exchange of information and databases through networks (Vinithkumar, 2006). In accordance

with other field of life sciences, bioinformatics has bright prospects in aquaculture (Aquainformatics), especially in skirmishing aquatic diseases and epizootics, domesticating improved brood stock, evolving advanced breeding and culture techniques, cost effective sustainable feed development as well as maintaining congenial aquatic environment. However, aquaculture is the food producing sector that has brought worldwide recognition for Bangladesh occupying top 5th position in global context (FAO, 2018). In order to sustain this sector by reducing production cost, the vast digitally available aquaculture information need to be integrated and synthesized through computational technology to create public

databases of established aquaculture systems and methodologies. Such cloud databases, output of aquainformatics, could save valuable money and time of the aquaculture scientists by guiding towards well defined and specific sectors of interest. In this review, the authors tried to focus on different tools of bioinformatics and its importance in aquaculture of Bangladesh.

Bioinformatics tools and databases for aquaculture research

The Bioinformatics tools are basically the software programs designed for collecting, organizing, retrieving and analyzing large amount of multiomics (genomics, proteomics,

transcriptomics, metabolomics etc.) data produced from high throughput technologies (Fig. 1) to shape the unstructured data into logically structured big frame-work for exploring the general biological phenomena (Khalil and Hill, 2005). Consequently, researchers can connect the complex biological processes by drawing link between *in vitro* systems and *in vivo* animal biology and interpret the results of any deviation (Mohanty *et al.*, 2019). Importantly, the tools must be user friendly and globally available on internet to the mass scientific research community for better implications. However, these tools are broadly classified under four categories (Danish *et al.*, 2017):

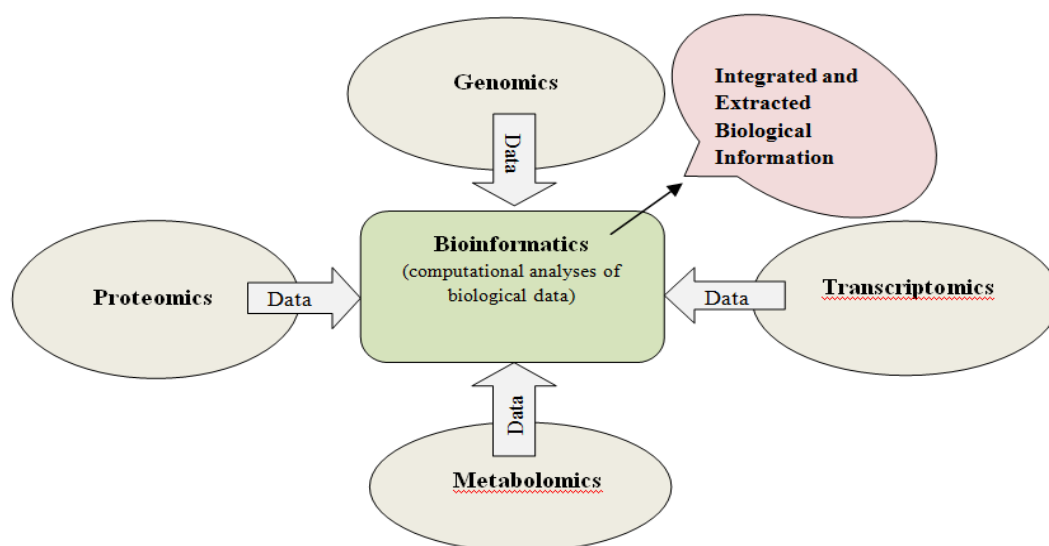


Fig. 1. Integration of biological data through bioinformatics.

a. Homology and similarity tools- identify resemblance of novel query sequences having unidentified structure and function comparing database sequences whose structure and function have been revealed.

b. Protein function analysis- compare query protein sequence with the secondary (or derived) protein databases containing information on motifs, signatures and protein domains to determine biochemical function.

c. Structural analysis- compare structures with the known structure databases.

d. Sequence analysis- conduct detailed analyses (such as evolutionary analysis, identification of mutations, hydropath regions, CpG islands, compositional biases etc.) to ascertain the specific function of the query sequence.

Some commonly used bioinformatics tools in fisheries and aquaculture described by Meena *et al.* (2020):

i. BLAST (The Basic Local Alignment Search Tool)- primarily compares query gene and protein sequences against available references in public databases. Currently available in several forms including PSI-BLAST, PHI-BLAST, BLAST 2 sequences etc.

ii. FASTA (FAST-All)- compare query sequences of nucleotide or peptide with a sequence database based on rapid sequence algorithm.

iii. EMBOSS (The European Molecular Biology Open Software Suite)- a new, open access software analysis bestowed with around 100 programs (applications) for sequence alignment and database searching with sequence patterns along with protein motif identification and domain analysis, nucleotide sequence pattern analysis, codon usage analysis for small genomes and so on.

iv. ClustalW- a common purpose multiple sequence alignment program for DNA or proteins to yield biologically meaningful multiple sequence alignments of divergent sequences, determines the best match for the selected sequences and visualizes the identities, similarities and differences by lining them up.

v. RasMol and Protein Explorer (derivative of RasMol) - widely used for structural display of DNA, proteins, and smaller molecules.

Moreover, Cn3D, Swiss PDB viewer, Hex, Vega, Bioeditor, Bioviewer, Chime etc. are also good names of bioinformatics software tools used in fisheries and aquaculture research.

A large volume of biological data are mined and analyzed by software tools using the available databases to harvest invaluable biological information. Biological databases work as the libraries of life-sciences-information in bioinformatics, which are developed from scientific researches, high-throughput technology, available literature and computational analysis (Attwood *et al.*, 2011). GenBank is a renowned publicly available inclusive database containing nucleotide sequences of more than 240000 named organisms, which is produced and monitored by the National Center for Biotechnology Information (NCBI) and being enriched continuously by individual laboratories and batch

submissions from comprehensive sequencing projects (Benson *et al.*, 2013). The three organizations namely the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI work altogether on constant data sharing basis as part of the International Nucleotide Sequence Database Collaboration (INSDC) (Benson *et al.*, 2013). The Protein Data Bank (PDB), in particular, contains 3D structural data on biological macromolecules (i.e. protein sequences and structures) for more than 30,000 global repositories including many marine organisms (Vinithkumar, 2006). Thereafter, for strengthening aquatic bioinformatics through biographical and ecological analyses, the Micro-Mar database of marine microbial genome has recently been developed with a view to collecting DNA diversity information from Marine prokaryotes (Vinithkumar, 2006). Besides the primary nucleotide database there are also some latest databases for being used in bioinformatics such as protein sequence databases, proteomic databases, protein structure, carbohydrate structure, protein model, protein-protein and other molecular interactions, RNA databases, signal transduction pathway databases, metabolic pathway and protein function, gene expression (mostly microarray data) databases (Kamble and Khairkar, 2016).

Table 1. Bioinformatics tools used in “Omics” database management (Mohanty *et al.*, 2019).

Frequently used bioinformatics tools	Databases
<ul style="list-style-type: none"> ✓ Basic Local Alignment Search Tool (BLAST)- compares nucleotide or protein sequences against databases that contain many archived sequences. ✓ Primer-BLAST- uses Primer 3 to design PCR primers to a sequence template. ✓ Spleign- used for computing cDNA-to-Genomic sequence alignments. ✓ GenBank: BankIt/ Sequin-both are web-based sequence submission tool to the GenBank database. ✓ Sequence Read Archive (SRA) Submission-stores sequencing data from the next generation of sequencing platforms including Roche 454 GS System, Illumina, Life Technologies AB SOLiD System. 	Genomics/ transcriptomics
<ul style="list-style-type: none"> ✓ Translate- Translates a nucleotide sequence to a protein sequence. ✓ FindMod- Predict potential protein post-translational modifications and potential single amino acid substitutions in peptides. ✓ Mascot -Sequence query and MS/MS ion search from Matrix Science Ltd., London. ✓ ProtParam- Physico-chemical parameters of a protein sequence (amino-acid and atomic compositions, isoelectric point, extinction coefficient, etc.). ✓ ScanProsite-Scans target sequence against PROSITE or a pattern against the UniProt Knowledgebase (Swiss-Prot and TrEMBL). ✓ Protein Data Bank (PDB)- archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies. ✓ MS-Fit- correlating Mass Spectrometry data (parent masses only, not fragment masses) with a protein in a sequence database which best fits the data. 	Proteomics
<ul style="list-style-type: none"> ✓ Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway- a collection of databases, utilized for data analysis in genomics, metagenomics, metabolomics and other omics studies, modeling and simulation in systems biology, and translational research. ✓ Biocarta. 	High throughput analysis

However, biological databases, based on their sources of information, can be grouped into primary (information of the sequence or structure alone; e.g. Gen bank and DDBJ for

genome sequence), secondary (derived information from primary and composite databases; e.g. SCOP developed at Cambridge university, CATH at university college of London,

eMOTIF at standford etc.) and composite (information from variety of primary database sources in one platform, e.g. The NCBI nucleotide and protein databases provides free access) databases (Kamble and Khairkar, 2016). Table 1, enumerates the widely used bioinformatics tools for efficient and standard management of “omics” data and databases.

Practical application of Bioinformatics in aquaculture research

Bioinformatics work to organize large biological data, develop appropriate tools and resources for analysis, interpret the results in biological manner and make databases available for global analytical research. Thus, bioinformatics enable researchers to browse existing biological information and submit their experimental findings to global databases (e.g. DNA Data Bank of Japan). Consequently, these approaches pave the way to uncover common principles of biological systems and highlight novel features. Table 2 shows bioinformatics data sources and their potential application. Therefore, Bioinformatics can be employed by aquaculture researchers for numerous applications such as:

i. Genomic data manipulation, genome annotation and expression profiling:

Bioinformatics tools are indispensable part of genome sequencing. For example, currently available BLAST/sequence alignments, in addition to handling, analyzing, comparing, relating, and visualizing DNA sequences, provide

itself a crucial aid in the sequencing process. Furthermore, the rapid and cost effective manners of next generation sequencing (NGS) platforms have proved its efficiency in decoding whole genome sequences of various organisms ranging from human to microscopic viruses including fish genomes with complex polyploidy levels (Oliver *et al.*, 2015; Krampis and Wultsch, 2015). The development of rapid and reliable DNA sequencing techniques (Sanger and Coulson, 1975; Maxam and Gilbert, 1977) have produced large scale sequencing data that are need to be analyzed by efficient bioinformatics methodologies equipped in powerful computers with sufficient memory. For instance, the shotgun sequencing techniques applied in Phage Φ -X174 and *Haemophilus influenza* genome sequencing has generated sequences of tens of thousands of small DNA fragments, ranging from 35 to 900 nucleotide bases and required the assembly of a complete bacterial genome (Sanger *et al.*, 1977; Fleischmann *et al.*, 1995). The terminals of these sequenced shotgun clones overlap and can be assembled using computerized similarity search algorithms (a critical area of bioinformatics research) into the complete genome (Mount, 2004; Abdurakhmonov, 2016). Additionally, manual genome annotation and prediction are literally impossible, therefore require computerized bioinformatics tools. The first genome annotation computer program was designed by Owen White in 1995, to aid analysis and annotation of *H. influenza* genome (Fleischmann *et al.*, 1995).

Table 2. Sources and utilization of bioinformatics data in different subject areas (Kamble and Khairkar, 2016).

Data Sources	Research Areas
Genomes	1) Phylogenetic analysis 2) Linkage analysis relating specific genes to diseases 3) Characterization of protein content metabolic pathways 4) Characterization of repeats 5) Structural assignments to genes
Raw DNA sequence	1) Identification of introns and exons 2) Separating coding and non-coding regions 3) Forensic analysis 4) Gene product prediction
Protein sequence	1) Multiple sequence alignments algorithms 2) Sequence comparison algorithms 3) Identification of conserved sequence motifs
Gene expression	1) Mapping expression data to sequence, structural and biochemical data 2) Correlating expression patterns
Macromolecular structure	1) Protein geometry measurements 2) Secondary, tertiary structure prediction 3) 3D structural alignment algorithms 4) Surface and volume shape calculations 5) Intermolecular interactions

However, all presently available gene annotation and prediction software have been developed based on White's principles (Abdurakhmonov, 2016). In the process of gene finding and annotation (e.g. searching protein-coding genes,

RNA transcripts, other functional sequences etc.), bioinformatics tools recognize the start-stop regions, introns, exons, motifs, repeats, and other regulatory, sensory or signaling regions within a genome which may vary between genes and

among organisms. Bioinformatics tools are also essential in gene and protein expression profiling. The popular gene expression techniques, viz., serial analysis of gene expression (SAGE), expressed sequences tags (ESTs), massively parallel signature sequencing (MPSS), transcriptome profiling, or RNA-Seq, and other microarray profile data often produce biased biological measurement because of their extreme noise-prone nature. Bioinformatics tools therefore are required to separate signal from noise in high-throughput gene expression researches (Abdurakhmonov, 2016). Bioinformatics is also required in protein identification (detecting complex sequence similarity) by protein microarrays and high-throughput mass spectrometry using protein sequence databases (Loo *et al.*, 1999; Tom *et al.*, 2013).

ii. Molecular folding, modeling, and design:

Aquaculture scientists can apply structural bioinformatics in identification of three dimensional (3D) structures of complex protein from aquatic organisms. They can also use molecular modeling, and folding to anticipate the possible function and model behavior of any molecular structures including proteins. They can also fold the molecule to its native functional 3D structure, and design therapeutic drugs for many complex fish diseases by applying structural bioinformatics. It also helps in *denovo* designing of complex bimolecular (e.g. protein, enzyme) structures and their possible interactions (Selkoe, 2013; Abdurakhmonov, 2016). By virtue of, the function of protein is strongly correlated to its primary structure, resulting from the coding DNA sequences. Homology modeling, based on homology patterns in primary protein structure, is used to predict important structural formations and interaction sites of a query protein comparing with other homologous proteins of known structure. Determination of secondary, tertiary, quaternary as well as 3D structures is very crucial for identifying exact function of query protein, whereas failure to fold into native structure may result in toxic or inactive proteins (Sloan *et al.*, 2016). In this regards, bioinformatics through energy landscape and modeling of protein folding approaches can mitigate the problem (Selkoe, 2013; Abdurakhmonov, 2016). Take for example, the “Iterative Threading Assembly Refinement” (I-TASSER) is an open access popular web server/stand-alone software tools widely used for structural and functional characterization of proteins in a comparative scale (Yang and Zhang, 2015). Moreover, bioinformatics through molecular modeling (quantum chemistry approaches) shed light on the molecular behavior

and derived information on structure, dynamics, surface properties, and thermodynamics of complex biological systems. However, the bioinformatics tools for modeling and designing are highly diversified (Abdurakhmonov, 2016).

iii. Generating biological network and system biology

Bioinformatics can aid fisheries and aquaculture researchers in developing aquatic biology network. Bioinformatics approaches including molecular sequence analysis, prediction, annotation and molecular modeling, which comprise the nucleus for building, organizing, and systematizing biological networks of molecules (e.g., metabolic, protein-protein interactions etc.) (Abdurakhmonov, 2016). Such networks for cellular processes, (which are the integration of various forms of genome data such as DNA-RNA-proteins sequences, secondary metabolites, gene expression data etc.) are both physically or functionally connected and therefore useful to understand complex relations among cellular processes and other biological networks. However, sustaining such biological networks for understanding system biology, a simulative, iterative, and model-oriented bioinformatics approaches are required (Ma’ayan, 2011). Examples of such networks are the Kyoto Encyclopedia of Genes and Genomes (KEGG), BioCyc database collection, BRAunschweig ENzyme DATabase (BRENDA), Reactome, Comparative Toxicogenomics Database etc (Abdurakhmonov, 2016).

iv. Database development:

A database is a planned and structured organization of relevant data. A user, seeking information, gets access to a database by an integrated set of computer software, which is termed as “database management system (DBMS)”. The DBMS works to assure consumer access to all the contained data, at the same time execute the entry of defined data, subsequent storage, revision, supervision, and retrieval. The DBMS requires modeling (hierarchical and network models), clustering, query languages and query optimization as well as visualization algorithms for managing large datasets to facilitate users with extraction of information (Mount, 2004; Abdurakhmonov, 2016). Therefore, application of bioinformatics is a prerequisite for successful database development and convenient management. A large extent of bioinformatics tools have been employed to develop an exclusive number of databases, which are different in their data definition, application, format, and access types (Table 1).

Aquaculture fields for bioinformatics devoted research

a. Disease diagnosis and aquatic health management

Molecular and DNA based diagnostic tools incorporating bioinformatics can provide early detection (prior to signs and symptoms) of fish disease. This is done either by direct sequencing of pathogen DNA, amplifying specific sequences or detecting DNA and gene expression using nucleic acid microarrays as bioinformatics tools (Altinok and Kurt, 2003). Although many more are to come, bioinformatics based DNA detection tools are available for infectious hematopoietic necrosis virus (IHNV), viral hemorrhagic septicemia virus (VHSV), viral nervous necrosis virus (VNNV) and *Renibacterium salmoninarum* (Mohanty *et al.*, 2019). Moreover, bioinformatics can be applied in aquatic disease monitoring by assessing environmental variables, genetic factors of the host fish as well as virulence of the pathogenic agents by interpreting genomics, proteomics and next-generation sequencing data on bioinformatics platforms such as “linux”. Correspondingly, vaccination with nucleic acid (enhanced immunity in host), detecting stress resistant protein for improved fish wellbeing, specific drug development etc. are some practical implications of bioinformatics in aquatic health management (Viarengo *et al.*, 2007; Banerjee *et al.*, 2017).

b. Fish nutritional aspects

Bioinformatics, nutrigenomics data in particular, produce information on diet mediated gene expression, deviation from normal metabolism and associated health status (Ngoh *et al.*, 2015; Ganguly *et al.*, 2018). Therefore, worth to conclude on dietary impact of different test diets and cost effectively select the best one for aquaculture. Bioinformatics also aid in assessing flesh quality of fish by metabolic and mineral profiling of edible muscle, hence, can anticipate presence of fish allergen (IgE mediated food hypersensitivity) and associated public health hazard (Godiksen *et al.*, 2009; Mohanty *et al.*, 2019).

c. Suitable strain development for aquaculture

Bioinformatics in combination of biotechnology and informatics can contribute in brood stock improvement and domestication. Besides the traditional genomic biotechnology (e.g. selective breeding, polyploidy, gynogenesis, androgenesis, sex reversal and gene transfer), and DNA marker based genome mapping technology, genome sequencing along with transcriptomic analyses has open a new era of fish breeding. Traditional

biotechnology was primarily based on phenotypic superiority selection and chromosomal manipulation. This trend has largely been replaced by gene sequencing and annotation. The second (next-generation sequencing) and third-generation sequencing technologies can produce terabits of data which require annotation using bioinformatics platform (such as Illumina sequencing) to produce information on genes underlining the performance and production traits. By analyzing the results aquaculture scientists can contribute to new strain development of aquaculture species. However, the majority of aquaculture species genomes have been sequenced using the Illumina technology and few have been supplemented with third-generation sequencing technologies such as PacBio sequencing (FAO, 2017). Whole genome information are available now for more than 30 fish species viz. puffer fish (*Takifugu rubripes* and *Tetraodon nigroviridis*), medaka (*Oryzias latipes*), zebrafish (*Danio rerio*), Atlantic salmon (*Salmo salar*), Atlantic cod (*Gadus morhua*), three spine sticklebacks (*Gasterosteus aculeatus*), rainbow trout (*Oncorhynchus mykiss*), common carp (*Cyprinus carpio*), hilsa shad (*Tenualosa ilisha*) etc (Mohanty *et al.*, 2019; Mollah *et al.*, 2019). Genome annotation on bioinformatics platform of these species could answer hundreds of their evolutionary and behavioral phenomena that are crucial from aquaculture perspectives.

Status of bioinformatics associated aquaculture research in Bangladesh

Biological research including aquaculture has undergone modernization through the incorporation of bioinformatics, a multidisciplinary approach relating statistics, computer technology and molecular biology. Bioinformatics, though a global lucrative tools for biological research, is yet not that flourished in aquaculture research of Bangladesh. Still no agricultural university in Bangladesh is offering bachelor degree in bioinformatics, although there are some semester courses relating bioinformatics under different majors in some universities of Bangladesh. It is imperative that Bangladesh Agricultural University has launched bachelor degree “Bioinformatics Engineering” offered from the Faculty of Agricultural Engineering and Technology in 2020, realizing the need for well developed multidisciplinary bioinformatics course and to avail quality graduate as well as experts in this field. Nevertheless, with few exceptions, many researchers/research-students/research groups are conducting need based bioinformatics research in the country but without collaborative and integrative approach. Therefore, Bangladesh Bioinformatics and Computational

Biology Association (BBCBA) have been formed and working since 2016, to bring the nation-wide bioinformatics venture under a single umbrella with better collaboration.

Compared to the developed countries, where bioinformatics tools have been used exclusively in fisheries and aquaculture research, Bangladesh is still a novice. Even our neighboring country India has gained enviable progress in this field. India has developed a good number of databases, for instance, Biotechnology Information System Network (BTIS net), Environmental Information System (ENVIS), Agricultural Research Information Network (ARISNET), The Indian Ocean Biodiversity Information System (IndOBIS) etc with exclusive provision for marine bioinformatics (Chavan *et al.*, 2004; Vinithkumar, 2006). However, application of bioinformatics has drastically reduced the operation cost of high throughput omics technology and has good potential in developing countries like Bangladesh. This is because bioinformatics projects do not require traditional large funding for world class laboratory development rather high speed internet connectivity with some powerful computers and enthusiastic manpower is merely sufficient (Sikder, 2008; Islam, 2013). Furthermore, the vision and ongoing development of Bangladesh government in information and communication technology (ICT) along with open access to the scientific publications, databases and digital depositories have ease the way of incorporating bioinformatics in aquaculture research. A good example of Bangladeshi scientists' contribution using bioinformatics (the Illumina HiSeqX Platform) is the first draft genome assembly of hilsha (*Tenualosa ilisha*), which has paved the way of identifying genes responsible for physiological, ecological and behavioral adaptations of the high valued fish species (Mollah *et al.*, 2019). Correspondingly, such genomic researches associated with bioinformatics on different aquaculture species can answer to hundreds of questions on aquatic drug design (aquatic pharmacology), effective feed development, stock management, resistant bred development as well as evolutionary biology of commercial fish species. It is encouraging that a good number of people (students/researchers) from Bangladesh are obtaining higher academic degrees by doing research on bioinformatics every year from abroad. The Bangladesh government as well as native scientists working in international platforms can utilize this large folk of young, innovative scientists to develop world class bioinformatics institute as well as globally

accredited curricula in Bangladesh. Such initiative will obviously inspire learner-based discoveries with improved analytical skills.

Challenges of bioinformatics associated aquaculture research

Genome based biotechnology, the core of modern biological invention and discoveries, fundamentally relies on bioinformatics to analyze large data sets. Efficient data mining, analysis, re-analysis, and sharing by bioinformatics are the keys to successful genome research (FAO, 2017). However, integration of bioinformatics in biological research (such as aquaculture) is still limited by the inadequate informatics background of the biological scientists (comprising the end users). Again, informatics experts often have insufficient biological knowledge therefore availing "training-scientists" with combined expertise is very difficult. This driving force has challenged biology students to balance between biological experiments and data analysis using supercomputers. The super computers or high-performance computer clusters (HPC) are the basic equipment to conduct bioinformatics analyses. Such computational dependency has created two bottlenecks; one is associated with the complex and less user-friendly software platforms, and the second one is extravagant purchasing and maintenance cost (FAO, 2017). Notably, users need to have basic knowledge of command lines to use bioinformatics platforms (e.g., Unix or Linux). Furthermore, purchasing and maintaining HPC with regular updates is next to impossible for the developing countries like Bangladesh as often costing over US\$1 million. Although, cloud computing could be an option in this regard but requires a certain level of infrastructure and IT resources to undertake bioinformatics analyses.

Conclusion

Bioinformatics could create major changes in aquaculture research of Bangladesh by developing common knowledge base of aquaculture within a single hub. Such sharing of information would save valuable time and effort of aquaculture scientists and at the same time would reduce haphazard funding for similar research. Therefore, bioinformatics based research is indispensable for making aquaculture more sustainable and cost effective in the country. However, economic and logistic support from both the government and private sectors are crucial to flourish aquainformatics in Bangladesh.

References

- Abdurakhmonov, I.Y. 2016. Bioinformatics: Basics, Development, and Future. In: Abdurakhmonov I.Y. (ed.). Bioinformatics: Updated features and application. IntechOpen Limited, London UK. pp. 3-27. <https://doi.org/10.5772/63817>
- Altinok, I. and Kurt, I. 2003. Molecular diagnosis of fish diseases: a review. *Turk. J. Fish. Aquat. Sci.* 3: 131-138.
- Attwood, T.K., Gisel, A., Eriksson, N.E. and Bongcam-Rudloff, E. 2011. Concepts, historical milestones and the central place of bioinformatics in modern biology: a European perspective. In: Bioinformatics-Trends and Methodologies. IntechOpen Limited, London UK. pp. 1-38. <https://doi.org/10.5772/23535>
- Banerjee, S., Mahanty, A., Mohanty, S., Mazumder, D.G., Cash, P. and Mohanty, B.P. 2017. Identification of potential biomarkers of hepatotoxicity by plasma proteome analysis of arsenic-exposed carp, *Labeo rohita*. *J. Haz. Mat.* 336: 71-80. <https://doi.org/10.1016/j.jhazmat.2017.04.054>
- Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. and Sayers, E.W. 2013. GenBank. *Nucleic Acids Res.* 41 (Database issue): 36-42. <https://doi.org/10.1093/nar/gks1195>
- Chavan, V., Watve, A.V., Londhe, M.S., Rane, N.S., Pandit, A.T. and Krishnan, S. 2004. Cataloguing Indian biota: the electronic catalogue of known Indian fauna. *Curr. Sci.* 87(6): 749-763.
- Danish, M., Trivedi, R.N., Kanyal, P., Khati, A. and Agrawal, R. 2017. Importance of biotechnology in fish farming system: an overview. *Prog. Res.* 12(1): 7-14.
- FAO. 2017. Forthcoming. Genome-based biotechnologies in aquaculture. Food and Agricultural Organization of United Nations (FAO), Rome, Italy. 88p. <http://www.fao.org/3/a-bt490e.pdf>
- FAO. 2018. The State of World Fisheries and Aquaculture, Contributing to food security and nutrition for all. Food and Agricultural Organization of United Nations (FAO), Rome, Italy. 227p.
- Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A. and Merrick, J.M. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Sci.* 269: 496-512. <https://doi.org/10.1126/science.7542800>
- Ganguly, S., Mahanty, A., Mitra, T., Mohanty, S., Das, B.K. and Mohanty, B.P. 2018. Nutrigenomic studies on hilsa to evaluate flesh quality attributes and genes associated with fatty acid metabolism from the rivers Hooghly and Padma. *Food Res. Int.* 103: 21-29. <https://doi.org/10.1016/j.foodres.2017.10.017>
- Godiksen, H., Morzel, M., Hyldig, G. and Jessen, F. 2009. Contribution of cathepsins B, L and D to muscle protein profiles correlated with texture in rainbow trout (*Oncorhynchus mykiss*). *Food Chem.* 113(4): 889-896. <https://doi.org/10.1016/j.foodchem.2008.08.012>
- Hesper, B. and Hogeweg, P. 1970. Bioinformatica: een werkconcept. *Kameleon.* 1: 28-29 (in Dutch).
- Islam, M.M. 2013. Role of bioinformatics in developing country: Bangladesh. *Curr. Trend. Tech. Sci.* 2(1): 160-165.
- Kamble, A. and Khairkar, R. 2016. Basics of bioinformatics in biological research. *Int. J. Appl. Sci. Biotech.* 4(4): 425-429. <https://doi.org/10.3126/ijasbt.v4i4.16252>
- Khalil, I.G. and Hill, C. 2005. Systems biology for cancer. *Curr. Opin. Oncol.* 1: 44-48. <https://doi.org/10.1097/01.cco.0000150951.38222.16>
- Krampis, K. and Wultsch, C. 2015. A review of cloud computing bioinformatics solutions for next-gen sequencing data analysis and research. *Meth. Next Gen. Seq.* 2: 23-34. <https://doi.org/10.1515/mngs-2015-0003>
- Loo, J.A., Brown, J., Critchley, G., Mitchell, C., Andrews, P.C. and Ogorzalek-Loo, R.R. 1999. High sensitivity mass spectrometric methods for obtaining intact molecular weights from gel-separated proteins. *Electrophoresis.* 20(4-5): 743-748. [https://doi.org/10.1002/\(SICI\)1522-2683\(19990101\)20:4/5<743::AID-ELPS743>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1522-2683(19990101)20:4/5<743::AID-ELPS743>3.0.CO;2-I)
- Ma'ayan, A. 2011. Introduction to network analysis in systems biology. *Sci. Signal.* 4(190): tr5. <https://doi.org/10.1126/scisignal.2001965>
- Maxam, A.M. and Gilbert, W. 1977. A new method for sequencing DNA. In: Proceeding of National Academy Sciences of USA. Vol. 74, pp. 560-564. <https://doi.org/10.1073/pnas.74.2.560>
- Meena, D.K., Sahoo, A.K., Panda, D. and Bahera, B.K. 2020. Bioinformatics: its application and perspective in fisheries and aquaculture. Aquafind: Aquatic Fish Database. <http://aquafind.com/articles/Bioinformatics-In-Fisheries.php>

- Mohanty, B.P., Mohanty, S., Mitra, T., Mahanty, A., Ganguly, S. and Singh, S. 2019. Omics technology in fisheries and aquaculture. *Advances in Fish Research*, Vol.-VII. Narendra Publishing House, Delhi, India. pp. 1–30.
- Mollah, M.B.R., Khan, M.G.Q., Islam, M.S. and Alam, M.S. 2019. First draft genome assembly and identification of SNPs from hilsa shad (*Tenualosa ilisha*) of the Bay of Bengal. *F1000 Res.* 8(320): 1-13. <https://doi.org/10.12688/f1000research.18325.1>
- Mount, W.D. 2004. Bioinformatics: sequence and genome analysis. 2nd ed. New York: Cold Spring Harbor Laboratory Press. 692p. <https://doi.org/10.1086/431054>
- Ngoh, S.Y., Tan, D., Shen, X., Kathiresan, P., Jiang, J. and Liew, W.C. 2015. Nutrigenomic and nutritional analyses reveal the effects of pelleted feeds on Asian seabass (*Lates calcarifer*). *PLoS. ONE.* 10(12): e0145456. <https://doi.org/10.1371/journal.pone.0145456>
- Oliver, G.R., Hart, S.N. and Klee, E.W. 2015. Bioinformatics for clinical next generation sequencing. *Clin. Chem.* 61: 124–135. <https://doi.org/10.1373/clinchem.2014.224360>
- Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, C.A., Hutchison, C.A., Slocombe, P.M. and Smith, M. 1977. Nucleotide sequence of bacteriophage phi X174 DNA. *Nature.* 265: 687–695. <https://doi.org/10.1038/265687a0>
- Sanger, F. and Coulson, A.R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* 94: 441–448. <https://doi.org/10.1258/jrsm.99.2.62>
- Selkoe, D.J. 2013. Folding proteins in fatal ways. *Nature.* 426: 900–904. <https://doi.org/10.1038/nature02264>
- Sikder, A.R. 2008. Bioinformatics in Bangladesh: an optimistic view. *Plant Tissue Cult. Biotech.* 18(1): 85-88.
- Sloan, C.A., Chan, E.T., Davidson, J.M., Malladi, V.S. and Strattan, J.S. 2016. ENCODE data at the ENCODE portal. *Nucleic Acids Res.* 44: 726–732. <https://doi.org/10.1093/nar/gkv1160>
- Tom, I., Lewin-Koh, N., Ramani, S.R. and Gonzalez, L.C. 2013. Protein microarrays for identification of novel extracellular protein-protein interactions. *Current Protocol in Protein Sciences.* Chapter 27: Unit 27.3. <https://doi.org/10.1002/0471140864.ps2703s72>
- Viarengo, A., Lowe, D., Bolognesi, C., Fabbri, E. and Koehler, A. 2007. The use of biomarkers in biomonitoring: a 2-tier approach assessing the level of pollutant-induced stress syndrome in sentinel organisms. *Comp. Biochem. Physiol. C. Toxicol. Pharmacol.* 146(3): 281–300. <https://doi.org/10.1016/j.cbpc.2007.04.011>
- Vinithkumar, N.V. 2006. Marine Bioinformatics. In: Advanced Biotech. Monthly Magazine India. Samanthi Publications Pvt. Ltd. pp. 1-12.
- Yang, J. and Zhang, Y. 2015. Protein structure and function prediction using I-TASSER. *Curr. Protoc. Bioinf.* 52: 5.8.1–5.8.15. <https://doi.org/10.1002/0471250953.bio508s52>