

GENOMIC SELECTION IN CATFISH: A REVIEW

*Umar, B. D., Sa'adu, H. and Danjuma, M.

Department of Biology, Federal College of Education, Katsina, Nigeria

*Corresponding Author: bashirumardayi0123@gmail.com; +2348168004506

ABSTRACT

Aquaculture genomics identifies the genetic basis of performance and production traits and utilizes such information for selective breeding programs. Fisheries genomics, an emerging field that applies genomic tools to solve questions related to fisheries management, has been discussed less frequently as compared to aquaculture genomics. Genes specific to important environmental and terrestrial adaptation, locomotion, immune system, anti-microbial properties, thermoregulation, air-breathing, etc. were identified and critically analysed. The analysis clearly indicated that Catfish genome possessed several unique and duplicate genes similar to that of terrestrial or amphibians' counterparts in comparison to other teleostean species. The genome information will be useful in conservation genetics, not only for this species but will also be very helpful in such studies in other fishes.

Keywords: Aquaculture, Genomic selection, next-generation sequencing, whole genome sequencing,

INTRODUCTION

Catfish (order Siluriformes) is one of the largest orders of teleosts containing ~4100 species, representing ~12% of all teleosts and ~6.3% of all vertebrates (Eschmeyer and Fong, 2014). Most catfish have a cylindrical body with a flattened ventral to allow for benthic feeding (Bruton, 1996). Catfish are so-named because of their whisker-like barbels, which are located on the nose, each side of the mouth, and on the chin. Most catfish possess leading spines in their dorsal and pectoral fins. Catfish are scaleless, a characteristic of catfishes distinguishing them from most other teleost fish. However, some catfish, such as plecos, possess bony dermal plates covering their skin (Arce *et al.*, 2013).

Aquaculture genomics identifies the genetic basis of performance and production traits and utilizes such information for selective breeding programs. Fisheries genomics, an emerging field that applies genomic tools to solve questions related to fisheries management, has been discussed less frequently as compared to aquaculture genomics. The use of genomic tools in fisheries can detect adaptive diversity in response to fishing effort, reveal stock identity and structure, and estimate abundance and spawning stock biomass (Casey, *et al.*, 2016). However, when compared to agriculture, aquaculture and fisheries appear to fall behind in applying genomics in selective breeding and resources management (Bernatchez *et al.*, 2017). In this review, we intend to fill this gap by introducing sequencing technologies and genomics and describing genomic findings of major fish in world fisheries and aquaculture concerning genome structure, organization, and function. Aquaculture genomics identifies the genetic basis of performance and production traits and utilizes such information for selective breeding programs. Considering the contribution to world fisheries and aquaculture production, we are determined to focus on 4 major catfish species in aquaculture.

MATERIALS AND METHODS

Search Method

Google and Google Scholar were used to research the grey literature. Looking through the references included in the chosen publications also allowed us to find more research. Our research then included pertinent articles that were connected to the supplied keywords.

Analysis and Presentation the Review

A comprehensive analysis and presentation system is essential in making information realistic and contextually relevant. This article focuses on genomic techniques and selective breeding strategies in aquaculture.

Table 1: Major catfish in aquaculture discussed in this review and their genomic significance.

Common Name	Scientific Name	Order	Role	Aquaculture/fisheries (genomic significance)	significance
Channel catfish	<i>Ictalurus punctatus</i>	Siluriformes	Aquaculture	Over 60% of U.S. aquaculture production (Broad geographic distribution, tolerance of low water quality and resistance against various infectious agents, evolutionary loss of scale)	Over 60% of U.S. aquaculture production (Broad geographic distribution, tolerance of low water quality and resistance against various infectious agents, evolutionary loss of scale)
African catfish	<i>Clarias gariepinus</i>	Siluriformes	Aquaculture	Found primarily throughout Africa, where it was first introduced in aquaculture around the mid-1970s. This omnivorous fish is quite resilient due to its ability to cope with extreme environmental conditions, tolerate various land-based farming practices and a large diet spectrum	
walking catfish	<i>Clarias magur</i>	Siluriformes	Aquaculture	The walking catfish <i>Clarias magur</i> (magur) is an important catfish species inhabiting the Indian subcontinent. It is considered as a highly nutritious food fish and has the capability to walk to some distance, and survive a considerable period without water	
walking catfish	<i>Clarias batrachus</i>	Siluriformes	Aquaculture	Walking and air-breathing adaptations	
bighead catfish	<i>Clarias macrocephalus</i>	Siluriformes	Aquaculture	Bighead catfish, which is native to Southeast Asia, is highly preferred by consumers. However, aquaculture of this species is generally hampered by its slow growth rate and disease susceptibility	

RESULTS AND DISCUSSION

DNA Sequencing Technologies

Next-Generation Sequencing (NGS): NGS techniques were further evolved into second-generation sequencing, third-generation sequencing and newly fourth-generation sequencing techniques (Ku, *et al.*, 2013). Next-Generation Sequencing (NGS) technologies have revolutionized genetic research in aquaculture by providing high-throughput, cost-effective, and accurate sequencing capabilities (Satam, *et al.*, 2023) (Table 1). The emergence of next-generation sequencing (NGS) technology has made it possible to identify and characterize genetic markers, microsatellites, and single nucleotide polymorphisms (SNPs) across the transcriptome and genome of a variety of non-model fish species for which there are currently insufficient or nonexistent genomic resources (Kumar & Kocour, 2017).

For example, using NGS techniques polymorphic microsatellite markers have been developed in large number of marine fishes (Barnes, *et al.*, 2014) and freshwater fishes (Rodriguez-Zarate *et al.*, 2014). Whole Genome Sequencing (WGS) projects have been completed for several key aquaculture species. These reference genomes serve as essential resources for comparative genomics and functional genomics studies (Johnston, 2024). Breeding plans and conservation initiatives have benefited from the understanding of genome organization, gene composition, and evolutionary history that WGS data have brought to light (Hohenlohe, *et al.* 2021). Aquaculture has been greatly influenced by Next-Generation Sequencing (NGS), which has made it possible to conduct in-depth genetic and genomic investigations that guide breeding, disease control, and environmental adaptation tactics. The efficiency, sustainability, and productivity of the aquaculture sector might be further improved with the further development of NGS technology and their incorporation into practices.

Genome assembly, completeness and characterization

Genome sequencing and assembly

Sequencing was conducted de novo using next-generation sequencing technologies such as 454 and Illumina and PacBio (Table 2). Most sequencing projects were advanced based upon previous research on linkage mapping and quantitative traits linkage (QTL), which enabled chromosome-level genome assemblies (Wang *et al.*, 2015; Xu *et al.*, 2014). The estimated genome sizes ranged from 544 Mb in turbot (Figueras *et al.*, 2016), a flatfish to 2.97 Gb in Atlantic salmon (Lien *et al.*, 2016). Among the 4 major species, brown trout, a tetraploid species in Salmonidae, possessed the second-largest genome of 2.18 Gb (Berthelot *et al.*, 2014). The estimated genome sizes ranged from 544 Mb in turbot (Figueras *et al.*, 2016), a flatfish to 2.97 Gb in Atlantic salmon (Lien *et al.*, 2016). Among the 14 major species, brown trout, a tetraploid species in Salmonidae, possessed the second-largest genome of 2.18 Gb (Berthelot *et al.*, 2014). Common carp, which is also a tetraploid species, had a genome of 1.83 Gb (Xu *et al.*, 2014). The flatfish species and northern snakehead had relatively small genomes (Figueras *et al.*, 2016; Xu *et al.*, 2017). The genome size of grass carp was approximately 1.07 Gb for a female and 0.9 Gb for a male (Wang *et al.*, 2015). The scaffold N50, an indicator of the genome completeness, ranged from 137 kb in Pacific Bluefin tuna to 7.7 Mb in channel catfish (Nakamura *et al.*, 2013). The final assembly obtained in this study resulted in high continuity and completeness of the genome as the N50 value was higher than the *C. batrachus* (Table 2). It is worth noting that multiple factors, such as the coverage and assembly software, may affect the estimated scaffold N50 values. The analyses of this genome provide a comprehensive understanding of the evolution of *C. magur* with respect to other fish species and the genes/gene families which were evolved in *C. magur* for environmental/terrestrial adaptation. Additionally, approximately 92% of *C. gariepinus* transcripts were successfully mapped to our assemblies, with over 90% coverage and more than 90% identity, demonstrating high functional completeness.

The GC content in *C. magur* genome (39.83%) is slightly higher than the *C. batrachus* (39.2%). Repetitive sequences constituted 43.94% of the *C. gariepinus* genome, which roughly corresponded to the estimated repeat content of 46% based on k-mer analysis. Unlike that found in other catfish species, including *Clarias magur* (43.72%), *Clarias macrocephalus* (38.28%). However, it is higher than in *Clarias batrachus* (30.30%). The distribution of genes and repeats on chromosomes followed the typical pattern observed in vertebrate genomes, with higher gene densities in GC-rich regions and lower gene densities in repeat-rich distal and pericentromeric regions. The estimated repeats content in *C. magur* was slightly higher than the *C. batrachus* and other teleosts. The variation in repeat coverage as compared to *C. batrachus* indicated that *C. magur* had undergone slightly more active adaptive evolution (Table 2). The variation in repeat content plays an important role in adaptive evolution and genome structure in fishes and other vertebrates due to unequal recombination. Although *C. batrachus* and *C. magur* are closely related but later one contains higher repeat elements. This might be one of the reasons for the higher genome size (1.02 Gb) in *C. magur* compared to *C. batrachus* (900 Mb).

Genome Evolution

Comparative insights of evolution of genes related to specific characteristics of catfish

A total of 23 copies of myoglobin genes were reported in *C. magur*, which is higher than the *C. batrachus* (15 copies), lungfish (7 copies), and most of them other vertebrates (2–3 copies). These genes were arranged on five scaffolds of *C. magur* genome. Out of 23 genes' copies, 19 were arranged as

tandem repeats on Scaffold 320 (14 copies) and on Scaffold 248 (5 copies), which is also reported to be tandemly

duplicated in *C. batrachus*. Myoglobin genes role are crucial for adaptation in hypoxic condition, where they rapidly oxygenate and deoxygenate to maintain oxygen balance during the period of fluctuation in oxygen supply and demand. Ten copies of sult16b gene were significantly expanded in *C. magur*, while 12 copies were reported in *C. batrachus*. Sult16b gene eliminates or neutralizes the deleterious effect of different xenobiotic compounds from aquatic and terrestrial environments and, thereby, may protect the *C. magur* in the hypoxic conditions.

Evolution of genes specific to environmental and terrestrial adaptation in Catfish

Temperature is one of the major environmental factors for catfish. In response to heat stress, genes involved in oxygen transport (Haemoglobin, myoglobin, and cytoglobin genes), protein folding and degradation, and metabolic process were found to be dramatically up regulated, while general protein synthesis was highly repressed (Feng *et al.*, 2014, 2015; Liu *et al.*, 2013). After cold-stress challenge, Ju *et al.* (2002) identified differential gene expression in response to cold acclimation in the brain of channel catfish. Temperature is one of the major environmental factors for catfish. In response to heat stress, genes involved in oxygen transport (Haemoglobin, myoglobin, and cytoglobin genes), protein folding and degradation, and metabolic process were found to be dramatically upregulated, while general protein synthesis was highly repressed (Feng *et al.*, 2014, 2015; Liu *et al.*, 2013). After cold-stress challenge, Ju *et al.* (2002) identified differential gene expression in response to cold acclimation in the brain of channel catfish.

Table 2: Genome assembly and gene annotation of major catfish in aquaculture

Common name	Chromosome number (2n)	Genome size (Mb)	Scaffold N50 (Mb)	Coverage	%GC	Number of genes	No of protein-coding genes	Sequencing platforms	Reference
Channel catfish	58	771.9	7.7	167X	31.5	30,530	26,661	Roche 454, IlluminaHiSeq 2000 and SOLiD	Xu <i>et al.</i> (2019); Xu <i>et al.</i> (2014)
African catfish walking catfish	56	980	33.71	90X	39.0	26,661	25,655	PromethION 2	Juliana <i>et al.</i> (2024)
		941	1.31	94X	39.8	23,748		multiplatform (short, medium and long reads) sequencing strategy	Kushwah <i>et al.</i> (2021)

Immunological adaptation: The adaptive/acquired immune system in vertebrates comprises major histocompatibility complex (MHC) I and II proteins along with their regulator proteins. The MHC I involves in presentation of antigens derived from the intracellular environment, while MHC II present antigens derived from the antigen presenting cells, like macrophages, B cells or dendritic cells. We identified 16 MHC I genes in distributed in lineages, viz. five copies of U lineage, five copies of Z lineage, five copies of L lineage and one copy of S lineage. MHC II genes consist of 12 alpha and 15 beta copies. The variation in MHC I genes present in *C. magur* may provide additional benefits as more diverse range of pathogens are found on the land. The species needs an extra gadget of immune system for land adaptation to deal with the pathogens of both the land and the aquatic habitats. The presence of transcriptional regulators, thymus transcription factor and T cell receptor might also provide strength to the immune system of the *C. magur*. In the *magur* genome, we also identified all the genes and/or components that might be involved in the inflammasome assembly and its activation.

Air-breathing adaptation: Some of the air-breathing fishes have developed lungs or a respiratory swim bladder, while others have modified their gills, branchial cavities, skin, pharynx, pneumatic duct or intestine for aerial respiration during their terrestrial habitat. In *C. magur*, the accessory respiratory

organ comprises supra-branchial chambers which is located dorsally to the gill cavities and has the respiratory membrane lining, the fan or gill plates and the respiratory tree.

Elastin b gene showed contraction, in terms of copy number, in *C. magur*, which is a major component reported for neofunctionalization and acquisition of bulbus arteriosus which is a respiratory component in aquatic teleost. For terrestrial adaptation, *C. magur* might have acquired cardiac muscle for air-breathing rather than the aquatic teleost-specific smooth muscle. Gene is responsible for vascular

development and angiogenic patterning during angiogenesis. The gene, involved in angiogenesis, has undergone strong selection in catfish.

Table 3: Genes associated with functional characterization important in fish aquaculture.

Specie	Functional characterization	Candidate genes associated with phenotypes	References
Channel catfish	Scale formation and regeneration	<i>Eda, edar, fgfr, lef1, tcf7, spp1, odam, Scppfamilies, Gsp37, Fa93e10, Sparcl1</i>	Jiang <i>et al.</i> (2013); Liu <i>et al.</i> (2016)
African catfish	Air breathing catfish		Julien <i>et al.</i> (2024)
walking catfish	ability for locomotion on land, especially during or just after the rainfall, covering a good distance.	HOXC9b and HOXA9 genes	Kushwaha <i>et al.</i> , (2021)

CONCLUSION

The genome provides a comprehensive understanding of evolution of Catfish with respect to other fish species and the genes/gene families which have evolved for environmental and terrestrial adaptations. It is evidenced in present study that the catfish genome possesses large numbers of unique and species-specific genes that have evolved in due course of evolutionary process and their specific functions support Catfish for living in adverse environmental conditions. The review also reveals that the presence of evolved specific genes/gene families may have facilitated the development of additional capabilities for environmental adaptations particularly in the catfishes. The genome information is a valuable genomic resource for its conservation management and would be a very useful model for studying genes responsible and their molecular mechanism in hypoxia/ammonia tolerance, locomotion, vision, hearing, olfaction, respiration, osmoregulation, antimicrobial substances, metabolic depression, pollutant degradation, antioxidant defence system, etc. not only for this species but also will be very helpful in such studies for other teleosts too.

REFERENCES

- Aree, M., Reis, R.E., Geneva, A.J., Pérez, M.H.S., (2013). Molecular phylogeny of thorny catfishes (Siluriformes: Doradidae). *Mol. Phylogenet. Evol.* 67, 560–577.
- Barnes, T.C., Izzo, C., Bertozzi, T., Saint, K.M., Donnellan, S., Hammer, M.P., Gillanders, B.M. (2014) Development of 15 microsatellite loci from mullet, *Argyrosomus japonicus* (Pisces: Sciaenidae) using next generation sequencing and an assessment of their cross amplification in other sciaenids. *Conserv. Genet. Resour.* 6) 345–348.
- Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., ... Withler, R. E. (2017). Harnessing the power of genomics to secure the future of seafood. *Trends in Ecology & Evolution*, 32(9), 665–680. <https://doi.org/10.1016/j.tree.2017.06.010>.
- Berthelot, C., Brunet, F., Chalopin, D., Juanchich, A., Bernard, M., Noël, B., ... Alberti, A. (2014). The rainbow trout genome provides novel insights into evolution after wholegenome duplication in vertebrates. *Nature Communications*, 5, 3657.
- Bruton, M.N., (1996). Alternative life-history strategies of catfishes. *Aquat. Living Res.* 9, 35–41.
- Casey, J., Jardim, E., & Martinsohn, J. T. (2016). The role of genetics in fisheries management under the EU common fisheries policy. *Journal of Fish Biology*, 89(6), 2755–2767.
- Eschmeyer, W., Fong, J., (2014). Available from: <http://research.calacademy.org/research/ichthyology/catalog/SpeciesByFamily.asp>

- Feng, J., Liu, S., Wang, R., Zhang, J., Wang, X., Kaltenboeck, L., Li, J., Liu, Z., 2015. Molecular characterization, phylogenetic analysis and expression profiling of myoglobin and cytoglobin
- Feng, J., Liu, S., Wang, X., Wang, R., Zhang, J., Jiang, Y., Li, C., Kaltenboeck, L., Li, J., Liu, Z., 2014. Channel catfish Haemoglobin genes: Identification, phylogenetic and syntenic analysis, and specific induction in response to heat stress. *Comp. Biochem. Physiol. Part D Genomics Proteomics* 9, 11–22.
- Figueras, A., Robledo, D., Corvelo, A., Hermida, M., Pereiro, P., Rubiolo, J. A., ... Gut, M. (2016). Whole genome sequencing of turbot (*Scophthalmus maximus*; Pleuronectiformes): A fish adapted to demersal life. *DNA Research*, 23(3), 181–192.
- Hohenlohe, Paul A et al. (2021). Population genomics for wildlife conservation and management, *Molecular ecology* vol. 30, 62-82. <https://doi.org/10.1111/mec.15720>
- Jiang, Y., Gao, X., Liu, S., Zhang, Y., Liu, H., Sun, F., Bao, L., Waldbieser, G., Liu, Z., 2013a. Whole genome comparative analysis of channel catfish (*Ictalurus punctatus*) with four model fish species. *BMC Genomics* 14, 780.
- Johnston, I. A., Kent, M. P., Boudinot, P., Looseley, M., Bargelloni, L., Faggion, S., ... & Lien, S. (2024). Advancing fish breeding in aquaculture through genome functional annotation, *Aquaculture*, 740589
- Ju, Z., Dunham, R., Liu, Z., (2002). Differential gene expression in the brain of channel catfish (*Ictalurus punctatus*) in response to cold acclimation. *Mol. Genet. Genomics* 268, 87–95
- Juliena. Nguinkal, Yedomon A. B. Zoclanclounon, Ronald M. Brunner, Yutang Chen & tom Goldammer. (2024). Haplotype-resolved and near-T2T genome assembly of the african catfish (*Clarias gariepinus*). *Scientific Data Open*.
- Ku, C.S.K., Roukos, D.H. (2013) From next-generation sequencing to nanopore sequencing technology: paving the way to personalized genomic medicin, *Expert Rev. Med. Devices* 10 (1) 1–6.
- Kumar, G., & Kocour, (2017). M. Applications of next-generation sequencing in fisheries research: A review, *Fisheries Research*, 186) 11-22.
- Kushwaha Basdeo, Manmohan Pandey, Paramananda Das, Chaitanya G. Joshi, Naresh S. Nagpure, Ravindra Kumar, Dinesh Kumar, Suyash Agarwal, Shreya Srivastava, Mahender Singh, Lakshman Sahoo, Pallipuram Jayasankar, Prem K. Meher, Tejas M. Shah, Ankit T. Hinsu, Namrata Patel, Prakash G. Koringa, Sofia P. Das, Siddhi Patnaik, Amrita Bit, Mir A. Iquebal, Sarika Jaiswal, and Joykrushna Jena, (2021). The genome of walking catfish *Clarias magur* (Hamilton, 1822) unveils the genetic basis that may have facilitated the development of environmental and terrestrial adaptation systems in air-breathing catfishes. *Genomes Explored*. Oxford. *DNA Research*, 0(0), 1–16 doi: 10.1093/dnares/dsaa031
- Lien, S., Koop, B. F., Sandve, S. R., Miller, J. R., Kent, M. P., Nome, T., ... Zimin, A. (2016). The Atlantic salmon genome provides insights into rediploidization. *Nature*, 533(7602), 200.
- Liu, S., Wang, X., Sun, F., Zhang, J., Feng, J., Liu, H., Rajendran, K., Sun, L., Zhang, Y., Jiang, Y., 2013. RNA-Seq reveals expression signatures of genes involved in oxygen transport, protein synthesis, folding, and degradation in response to heat stress in catfish. *Physiol. Genomics* 45, 462–476.
- Liu, Z., Liu, S., Yao, J., Bao, L., Zhang, J., Li, Y., Jiang, C., Sun, L., Wang, R., Zhang, Y., Zhou, T., Zeng, Q., Fu, Q., Gao, S., Li, N., Koren, S., Jiang, Y., Zimin, A., Xu, P., Phillippy, A., Geng, X., Song, L., Sun, F., Li, C., Wang, X., Chen, A., Jin, Y., Yuan, Z., Yang, Y., Tan, S., Peatman, E., Lu, J., Qin, Z., Dunham, R., Li, Z., Sonstegard, T., Feng, J., Danzmann, R., Schroeder, S., Scheffler, B., Duke, M., Ballard, L., Kucuktas, H., Kaltenboeck, L., Liu, H., Armbruster, J., Xie, Y., Kirby, M., Tian, Y., Flanagan, M., Mu, W., Waldbieser, G., 2016. The channel catfish genome sequence provides insights into the evolution of scale formation in teleosts. *Nat. Commun.* 7, 11757.
- Nakamura, Y., Mori, K., Saitoh, K., Oshima, K., Mekuchi, M., Sugaya, T., ... Inouye, K. (2013). Evolutionary changes of multiple visual pigment genes in the complete genome of Pacific bluefin tuna. *Proceedings of the National Academy of Sciences of the US A*, 110(27), 11061–11066. <https://doi.org/10.1073/pnas.1302051110>.
- Rodriguez-Zarate, C.J., Carvalho, D.C., Hammer, M.P., Beheregaray, L.B. (2014) A set of microsatellite markers for the threatened Murray hardyhead, *Craterocephalus fluviatilis* (Pisces: Atherinidae) from the southern Murray-Darling basin, *Conserv. Genet. Resour.* 6 473– 475

- Satam H, Joshi K, Mangrolia U, Waghoo S, Zaidi G, Rawool S, Thakare RP, Banday S, Mishra AK, Das G, Malonia SK. (2023). Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology (Basel)*, 12(7) 997. <https://doi.org/10.3390/biology12070997>
Erratum in: *Biology (Basel)*. 2024 Apr 24;13(5):286. <https://doi.org/10.3390/biology13050286>
- Wang, Y. P., Lu, Y., Zhang, Y., Ning, Z. M., Li, Y., Zhao, Q., ... Zhu, Z. Y. (2015). The draft genome of the grass carp (*Ctenopharyngodon idellus*) provides insights into its evolution and vegetarian adaptation (vol 47, pg 625, 2015). *Nature Genetics*, 47(8), <https://doi.org/10.1038/ng0815-962a962-962>.
- Xu, J., Bian, C., Chen, K., Liu, G., Jiang, Y., Luo, Q., ... Huang, Y. (2017). Draft genome of the Northern snakehead, *Channa argus*. *GigaScience*, 6(4), 1–5.
- Xu, P., Xu, J., Liu, G., Chen, L., Zhou, Z., Peng, W., ... Sun, X. (2019). The allotetraploid origin and asymmetrical genome evolution of the common carp *Cyprinus carpio*. *Nature Communications*, 10(1), 4625. <https://doi.org/10.1038/s41467-019-12644-1>.
- Xu, P., Zhang, X., Wang, X., Li, J., Liu, G., Kuang, Y., ... Sun, X. (2014). Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nature Genetics*, 46(11), 1212–1219. <https://doi.org/10.1038/ng.3098>.