

RESEARCH

Open Access



Pangenomics of the cichlid species (*Oreochromis niloticus*) reveals genetic admixture ancestry with potential for aquaculture improvement in Kenya

John G. Mwaura^{1,2}, Clabe Wekesa^{1,4}, Kiprotich Kelvin¹, Ang'ienda Paul³, Philip A. Ogutu¹ and Patrick Okoth^{1*} 

Abstract

Background Nile tilapia has a variety of phenotypes suitable for aquaculture farming, yet its entire gene pool with potential for breeding climate ready strains in resource-limited settings remains scanty and poorly documented. SNP calling has become increasingly popular in molecular genetic studies due to their abundance and precision in estimating and identifying an organism's genetic makeup. SNPs are versatile in trait-specific breeding, which, in contrast to pedigree breeding, is affordable and speeds up genetic advancement by allowing animals to be used as parents sooner.

Results Clustering analysis revealed a strong correlation between the experimental sample, *Oreochromis niloticus*, *Oreochromis aureus* and *Betta splendens*. Three other species indicated evolutionary independence. Comparative genomics identified similarities between fishes with common genetic and evolutionary ancestry, allowing for better adaptation to local environmental conditions. Some of the selected genes exhibiting substantial effect on immunity include: *Prxs*, *MMR1 like*, *ZMYM4-like partial*; stress-reactive genes including: *PALLD-like* gene, *LPLBAG6-like* and growth-related genes including: *NF1x like*, *PEDF* and *CL like*. Experimental sample, *O. niloticus*, *O. aureus* and *Danio rerio*, can hybridize in their natural environments bringing about genetic admixture ancestry that hybridizes new genes which confer beneficial phenotypes.

Conclusion Breeding for specific traits could be a useful addition to aquaculture to allow expansion of the ecological niche and genetic base for tilapia. Some of the beneficial genes that can be hybridized include *Slc25a24* and *Slc12member 10*, *tandem duplicate 1*, for salinity tolerance and *Abca1*, *bcl2a* and *mylk13* for hypoxia tolerance. Breeders should introduce beneficial traits into fish breeds to ensure they are climate ready and able to weather climate shocks. This will allow aquaculture to contribute to food and nutrition security in line with SDG2 and improve the economic status of fish-farming communities in the Global South countries.

Keywords Pangenome, Genetic admixture ancestry, Introgression, Stress, Nile tilapia, Pedigree breeding

*Correspondence:

Patrick Okoth
okothpatrick@mmust.ac.ke

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

The demand for fish and fisheries products has been on an upward trajectory due to an ever-increasing world population. Capture fisheries are overstretched and can no longer produce enough fish to cope with the ever-rising demand. Therefore, aquaculture production has become the alternative to fill the gap. With the limited land and water resources, intensification practices have been adopted to enhance aquaculture production. This, however, has far-reaching ramification on fish (Mangang & Pandey, 2021). This is due to the accumulation of fish waste and left-over feeds in the culture systems that eventually lead to an increase in the levels of ammonia in the water and high stocking densities with calamitous consequences. Ammonia exists in two forms, the ionized (NH_4^+ and the unionized NH_3) with the unionized being more toxic to fish due to its ability to cross the epithelial linings into the blood of fish. Interconversion between these two forms of ammonia is influenced by temperatures. High temperatures increase ammonia NH_3 concentrations (Zhong et al., 2022; Song et al., 2022). Climate change has currently resulted into increase in water temperatures and with cumulative increase in ammonia toxicity in the culture systems resulting into increased fish mortalities. Kenyan aquaculture is predominantly freshwater. Due to the limited freshwater resources and the ensuing competition wrought by agriculture, human and animal use, aquaculture expansion is curtailed. However, Kenya has enormous untapped resources in marine and brackish water which could be utilized for aquaculture (El-Sayed, 2006). There are limited fish species of aquaculture importance that can survive in marine and brackish water environments. For fish to survive in harsh environments, they have to possess genes that confer advantageous phenotypes. Tilapias are well-known for their high yields, rapid growth rates, and tolerance to a wide range of environmental conditions including the effects of high stocking densities (Aketch et al., 2014), making it the second most farmed fish in the world (Munguti et al., 2022; Mwaura et al., 2023). *Sarotherodon*, *Oreochromis*, Tilapia, *Tristromella*, and *Danakilia* are the five genera of tilapia. *Sarotherodon*, *Oreochromis*, and Tilapia form the largest species in the wild and in aquaculture. Tilapia is native to Africa and has been introduced to over 140 countries, to boost fishing output and enable the growth of aquaculture, with Nile tilapia (*O. niloticus*) featuring more prominently (Prabu et al., 2019). This distribution beyond their native range is majorly dependent on their genetic flexibility to adapt to a varying range of environmental conditions (Lind et al., 2019). Tilapia can be grown in diverse farming systems and is omnivorous, requiring minimal fish meal in its feed. It has a naturally high tolerance to variable water quality and can grow

in both freshwater and marine environments. Tilapias are hardy and have good disease resistance. With these genetic advantages, tilapia may modify their interactions with other species causing remarkable changes in the fish community structures of local waters (Chapman, 2019). Non-native fish species introduced into freshwater systems can reduce biodiversity and change local community dynamics (Aloo et al., 2017). Exotic species have been recognized as the third largest cause of vertebrate species extinction in aquatic habitats. Exotic species introductions typically jeopardize ecological stability, leading to extinction through long-term predation and competition, as well as the replacement of native species by alien ones (Havel et al., 2015; Twongo, 1995). Humans' unintentional or intentional spread of fish species endangers the biodiversity and character of freshwater environments. Invasive fish studies have thus focussed on understanding the biological characteristics of prospective invasive fish, forecasting invasion outcomes at local and large-scale levels, and analysing the consequences of invasions on ecosystems.

Many tilapia species are extremely adaptable, making them particularly prone to becoming dangerous invasive species. Tilapia can adapt to brackish conditions, and some may even survive at salinities as low as 35 parts per thousand (ppt), which is typical of sea water (Yan et al., 2013). Tilapia can withstand significant fluctuations of the environmental conditions including salinity, dissolved oxygen, and temperature. Its tolerance to environmental instability, as well as their high reproduction, rapid growth rates, and omnivorous food habits, all contribute to their successful invasions (Martin et al., 2010).

Genetic changes in aquatic ecosystems and populations due to such invasions must be considered for successful long-term management and conservation of the aquatic ecosystems. To successfully understand the aquatic biodiversity, molecular genetic data on phylogeny, population genetics, and genomics, as well as quantitative features, are required (Hohenlohe et al., 2021). For instance, studies by Angienda et al. (2011) provided empirical evidence of genetic hybridization and introgression of alleles between the introduced Nile tilapia and the native Singida tilapia within Lake Victoria Basin based on analysis of mitochondria and microsatellite DNA markers. The results indicated gene flow between the two Cichlid species, suggesting genetic admixture ancestry and a compromise of the genetic integrity of the species. Various methods have been used to determine genetic sequence with the single reference analysis method being the most prominently used. However, the single reference analysis method has suffered shortcomings in the identification of genetic variations among populations. These shortcomings, however, can be effectively overcome by

constructing a pangenome. A pangenome aims to capture the complete genetic diversity within a species and reduce bias in genetic analysis inherent in using the single reference genome (Hurgobin & Edwards, 2017). The complete set of genes in a particular species is known as the pangenome. It is made up of the essential genes, which are found in every member of the species and the shell genes which are found in some but not all members of the species. Using pangenomes for phylogenetic analysis has the benefit of allowing SNPs found in Presence/Absence Variat (PAV)-displaying areas to be utilized to infer more precise connections between accessions. The phylogenomic tree may be used to plot the numbers that represent which genes are specifically present and absent in each accession using variable genes. The amount of nucleotide changes per site determines the length of each branch in the tree, and the evolutionary patterns shown in the tree may be connected to the biological characteristics of each accession. Pangenome analysis is mostly useful in the identification of Single nucleotide Polymorphism Markers (SNPs) which aid in the identification of the genetic variance within a species. SNPs are high-resolution molecular markers, used to analyse the neutral and adaptive genetic diversity of populations with large numbers (Wenne, 2023). SNPs calling has been used to identify species and hybrids in natural environments, as well in examining the genetic implications of restocking as a conservation effort and the deleterious consequences on wild populations of fish accidentally escaping from culture systems (Williams et al., 2010). SNPs are extremely beneficial for identifying genomic regions associated with phenotypic polymorphisms that are important for aquatic biodiversity conservation and management. Traditionally, species evolution has been understood as a long-term process, lasting up to hundreds or thousands of millions of years (Wiens, 2004). Nonetheless, there is rising evidence that recent speciation occurs in natural environments resulting in morphological divergence (Via, 2009). Recent improvements in sequencing techniques have led in the collection of massive data sets of molecular markers useful in identifying genetic diversity in populations and genomic areas influenced by natural selection (Bansal & Boucher, 2019). SNP loci can be found in coding or non-coding genome regions and are useful markers in defining populations and species that are closely related, characterization of their polymorphism and how it changes over time and helps in the understanding of the adaptations and genetic polymorphism, of a population, at specific trait loci (Garg et al., 1999). The current study aims to identify the relationships between natural populations of cichlids using SNPs markers. This will be beneficial in the identification and breeding of Nile tilapia strains that are high yielding

and resistant to environmental stress. Over the past few decades, aquaculture has expanded significantly, in part due to ongoing innovations. The creation of Genetically Improved Farmed Tilapia (GIFT) strain in 1988 and the subsequent start of genetic improvement are two factors that have contributed to the success of fish species today (Li et al., 2014). Currently, other cutting-edge genetic techniques, such as the use of mutational breeding and CRISPR-Cas9 genome editing, are being used and are revolutionizing the aquaculture industry.

Methods

Sample preparation and sequencing

Female Nile tilapia was collected from a local pond in Ilala Fish farm in Kakamega East Sub-County, Kakamega County, Western Kenya. Genomic DNA from muscle tissue was extracted using Qiagen GenomicTip100 (Qiagen, Germantown, MD, USA). The Masinde Muliro University of Science and Technology Institutional Ethics and Review Committee (MMUST-IERC) [REF: MMU/COR: 403012 Vol 5 (01)] approved the animal procedures, and all tests were carried out in accordance with the regulations.

DNA extraction and sequencing

DNA was extracted from muscle tissue using the protocol as described in Mayjonade et al. (2016). Tissue samples about 50 mg were ground to a fine powder in liquid nitrogen. The powder was then placed in 1.5-mL microtubes containing 0.7 mL 2% CTAB extraction buffer [20 mM EDTA, 0.1 M Tris-HCl pH 8.0, 1.4 M NaCl, 2% CTAB, plus 0.4% β -mercaptoethanol added just before use]. The solution was incubated at 65 °C for 45 min, gently mixing by inversion every 15 min; 500 μ L of chloroform-isoamyl alcohol (24:1) was added to the tubes and gently mixed for 1 min. Samples were centrifuged for 10 min at 12,000 rpm; 0.6 mL of the supernatant was then transferred to a fresh tube following the addition of 500 μ L chloroform-isoamyl alcohol (24:1); this procedure was repeated twice. 500 μ L of the supernatant was then transferred to a fresh tube with 0.7 mL of cold isopropanol (−20 °C); samples were gently mixed by inversion and centrifuged at 12,000 rpm for 10 min, and so it was possible to visualize the DNA adhered to the bottom of the tube. The liquid solution was then released and the DNA pellet washed with 1 mL of 70% ethanol to eliminate salt residues adhered to the DNA, and set to dry for approximately 12 h, or until the next day, with the tubes inverted over a filter paper, at room temperature. The pellet was then re-suspended in 100 μ L TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0) plus 5 μ L Ribonuclease (RNase 10 mg mL^{−1}) in each tube; this solution was incubated at 37 °C for 1 h and after stored

at $-20\text{ }^{\circ}\text{C}$. The isolated genomic DNA was subsequently applied to construct DNA sequencing libraries using Illumina MiseQ. We further trimmed 5 bases in both ends of the raw reads, discarded those duplicated reads, and removed reads with 10 or more Ns and low-quality bases to improve the quality of sequenced reads.

Genome assembly and sequence annotation

Genome was assembled de novo by Abbyss v2.8.5 and gaps filled with ABYss sealer (v1.12-r6, default parameters) to construct contigs and original scaffolds by using clean reads. The genome completeness was checked by BUSCO v4.1.2, and contigs less than 500 were removed by in-house bash script. Raw reads and the genome assembly have been deposited in the NCBI under the project ID PRJNA848236. Repeat sequences in the experimental sample assembly were predicted by an integration of three routine approaches, including de novo, tandem repeat predictions and homology annotation (Mitra et al., 2021). For the de novo prediction, RepeatMasker v4.0.9_p2 was used to identify and mask the repeats (Tarailo-Graovac & Chen, 2009). Ab initio gene prediction was performed using Augustus v3.5.0 and GeneMark-ES. Augustus was run on a predefine *Danio rerio* genome training set. The tandem repeats were subsequently predicted using Tandem Repeat Finder (Benson, 1999) (version 4.04). These repeat data from above three approaches were integrated to generate a non-redundant repeat set. For the homology annotation, protein sequences of *D. rerio*, *O. niloticus*, *O. aureus*, *B. splendens*, *L. chalumnae*, *S. chuatsi*, and *P. altivelis* were downloaded from the Ensembl database (release 75). These sequences were aligned onto the F1 strain assembly to generate alignments using Python script with an e-value $< 1.0 \times 10^{-5}$. Python script was used to select consensus genes predicted by both Augustus v3.5.0 and GeneMark-ES.

Comparative genomics

Orthologous analysis was performed with orthoMCL and COGtriangles programmes, and perl script from GET_HOLOGUES programme was used to generate pan genome matrix from intersection between orthologues generated by both programmes.

Construction of the phylogenomic and divergence time trees

Phylogenomic analysis was produced by GET_PHYLO-MARKERS programme which implements the FigTree programme. Protein sequences of each single-copy gene family were aligned to each other. The protein alignments were then converted to their corresponding coding sequences using an in-house Perl script. These nucleotide

sequences were linked into a continuous sequence for each species. Non-degenerate sites, obtained from the continuous sequence of each species, were then joined into a new sequence of each species to build a phylogenomic tree using MrBayes (Huelsenbeck & Ronquist, 2001) (Version 3.2, with the GTR+gamma model). The tree was then visualized using the FigTree programme which revealed the experimental sample as closely related to *O. niloticus*, *O. aureus*, and *Betta splendens*. To determine the number of genes shared or unique between them, a Venn diagram was constructed using the Venn programme.

Results

A whole-genome sequence script was carried out for the experimental fish, and the raw reads together with the genome assembly have been deposited in the NCBI database [project ID PRJNA848236]. Gene families were identified by OrthoMCL (v1.4) and COGtriangle. First, nucleotide and protein data of seven species representative of different teleost families (*O. niloticus*, *B. splendens*, *O. aureus*, *D. rerio*, *L. chalumnae*, *S. chuatsi*, and *P. altivelis*) were downloaded from Ensembl (Release 70) and National Center for Biotechnology Information (NCBI) to co-analyse with the experimental sample genome assembly. The genomes of the isolates contained 95,275 and 93,972 orthologous groups, respectively, identified by the orthoMCL and COGtriangle algorithms, of which 60,045 ortho groups represented the intersection point of the two programmes. The orthologous groups of both orthoMCL and COGtriangle intersected to produce the pangenome matrix below (Fig. 1).

A panel of SNPs markers was developed for the identification of six fish species of interest (*O. niloticus*, *Betta*



Fig. 1 Venn diagram representing the pangenome matrix generated from the intersection between orthologues generated from the orthologous analysis performed with orthoMCL and COGtriangles programmes and a perl script from GET_HOLOGUES programme

splendens, experimental sample, *O.aureus*, *D. rerio*, *Latimeria chalumnae*, and for comparison purposes) (Fig. 2). A total of 27,577 markers were obtained. The pangenome matrix was used to generate maximum likelihood pangenome phylogenomic tree with IQ-TREE v1.6.12 after determining the best fit model by ModelFinder (Kalyanamoorthy et al., 2017) in GET_PHYLOMARKERS (Vinuesa et al., 2018) and visualized by FigTree v1.4.4 (Rambaut, 2012).

We determined the divergence of species and found that the experimental sample was closely related to *O. aureus* and had diverged together from common maternal ancestry. The phylogenomic analysis revealed a true relative of experimental sample as *O. niloticus* (the Nile tilapia) as they clustered into one clade with similar Operational Taxonomic Units (OTUs) (Fig. 2). *L. chalumnae* and *S. chuatsi* evolved from a common ancestor and they both seem to have shared similar evolutionary relationships with *B. splendens* and *D. rerio* because they clustered together in a common clade. The species *P. altivelis* clustered to its own clade indicating a distinctive relationship with other species of study (Fig. 2).

A total of 27,577 expressed sequence tags were collected. From these tags a total of 11,020 genomes had a hit with *B. splendens* (39.96%), 16,333 with the experimental sample (59.23%), 18,272 with oreochromis

niloticus (66.3%), and 11,852 with *Oreochromis aureus* (42.98%). The sequence tags expressed were designed from the region having the best homology with the reference genome (Fig. 3).

We identified the number of genotypes and the cumulative number of genotypes (%) called at increasing levels of sequencing depth as shown in Fig. 4.

The total number of transitions (Ts) and transversions (Tv) were investigated. The total number of transversions (TV) (A/G, C/T, G/A, and T/C) were significantly higher for all the pairs. The frequencies of A/C and T/G, A/T and T/A, C/A and G/T, and C/G and G/C were at similar levels (Fig. 5).

About 19.42% of SNPs from the validation panel were A ↔ T and G ↔ C transitions representing 10.64% and 8.78% of the total SNPs, respectively. On the other hand, a total of 80.58% of SNPs from the validation panel were A ↔ C, A ↔ G, C ↔ T, and G ↔ T transversions representing 9.38%, 30.89%, 30.91% ,and 9.40%, respectively (Table 1).

A total number of 2,787,593 SNPs sites were detected with a transition/transversion ratio of 1.62 (Table 2). The rate of mutation was not high. Fatal mutations occur when transition/transversion ratio is more than 2. From the total single nucleotide polymorphism counts the number of genes were related to the *B. splendens*.

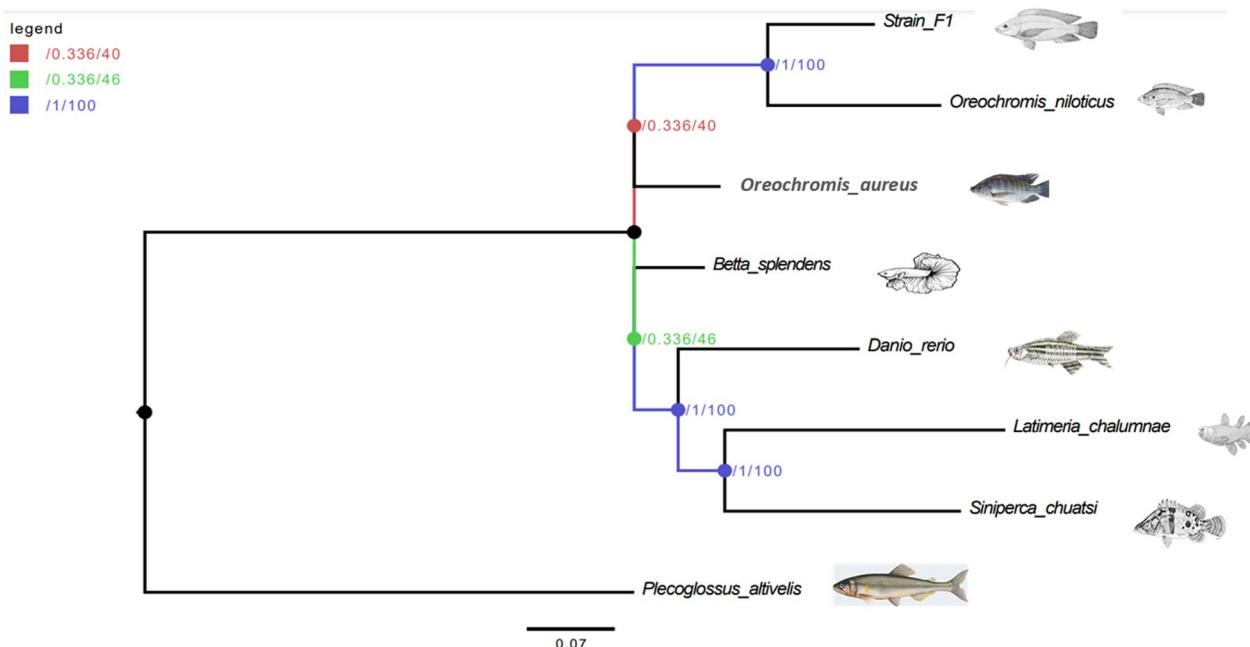


Fig. 2 A pangenome tree showing phylogenomic relation between strain F1 and the seven strains of interest (*O. niloticus*, *O. aureus*, *B. splendens*, *D. rerio*, *L. chalumnae*, *S. chuatsi*, and *P. altivelis*). The nodes are coloured per the legend in which the first value corresponds to approximate Bayes branch support values, and the second is the UFBoot support values. The node with less than 95% support (red node) was collapsed

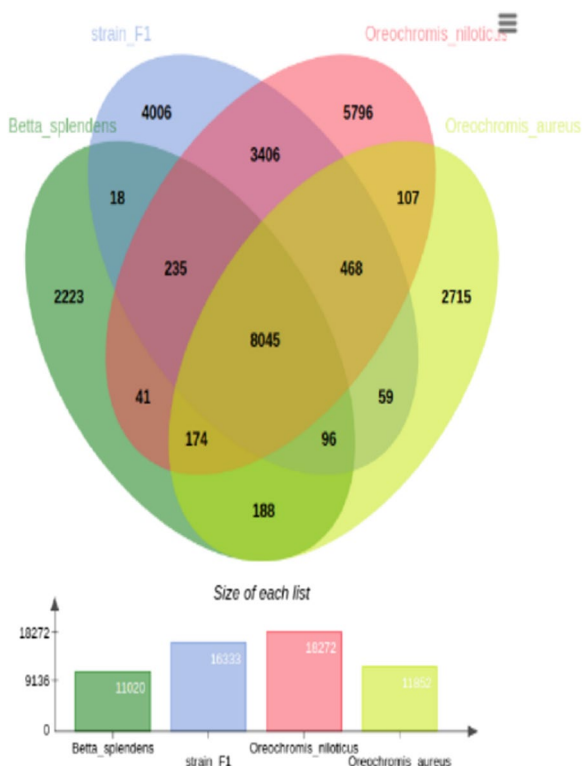


Fig. 3 Venn diagram representing the distribution of genes shared by our sample and *O. niloticus*, *Betta splendens*, experimental sample, and *O. aureus*. Each ellipse represents a model species, and each intersection depicts the number of genes shared by two species or more. For each model species, the number of genes and the percentage of 27,577 genes are presented. *O. niloticus* had highest number of shared genes with the other species (18,272) amounting to 66.3%, while experimental sample followed with 16,333 (59.23%) shared genes. *O. aureus* shared 11,852 (42.98%) genes, while *B. splendens* shared 11,020 (39.96%) genes with other species. *O. niloticus* had 5796 unique genes. Experimental sample had 4006 unique genes, *O. aureus* 2715, and *B. splendens* with 2223 unique genes

Discussion

The estimation of a pangenome infers that either the species has an open or closed pangenome. It is open when the number of genes in a pangenome increases with the addition of further genomes or a closed pangenome, when the additional sequenced genomes do not add new genes into the existing pangenome. Species colonizing multiple environments can easily exchange genetic material and tend to have an open pangenome while, species living in an isolated habitat have less possibility to exchange genetic material and have closed pangenome. Tilapia has an open pangenome that allows for the introgression of beneficial genes resulting in the development of strains that are more superior to their parents. The Genetically Improved Farmed

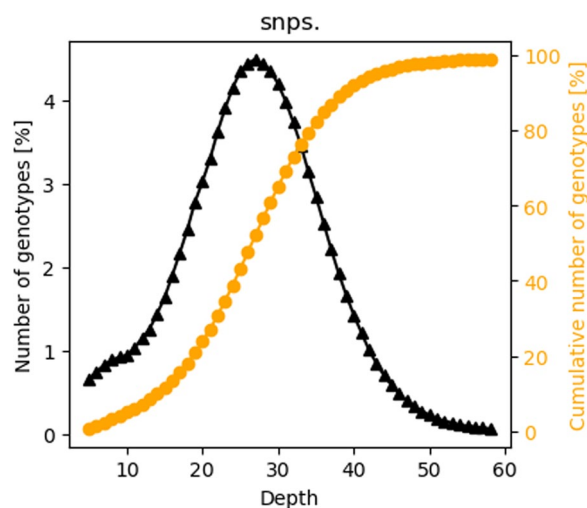


Fig. 4 A combined graph representing the SNPs and genotype counts from genotyping by sequencing (GBS) data. The black curve displays the number of genotypes called at increasing levels of SNPs sequencing depths, while the orange curve represents the cumulative number of genotypes called at increasing levels of SNPs sequencing depths

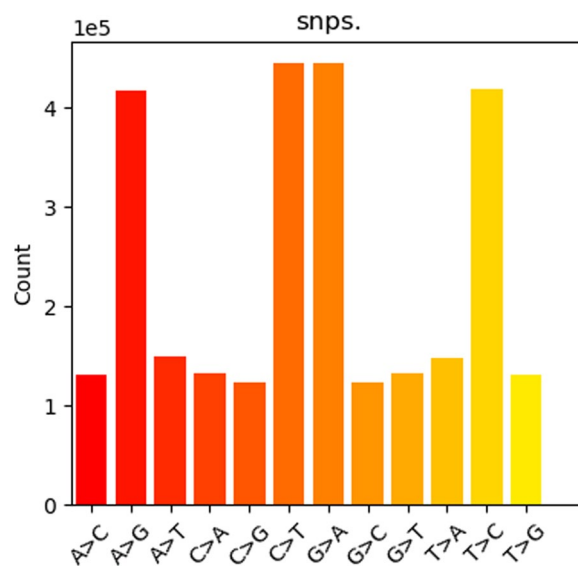


Fig. 5 A substitution graph showing distribution of detected SNPs types (transversions and transitions) with each bar corresponding to the total counts. A > G, C > T, G > A, and T > C are the SNPs with the highest counts. (transversions are more common compared to the transitions)

Tilapia has been subjected to breeding programmes over a long period of time, and this acts as evidence that the tilapia genome is open for expansion through the introgression of genes conferring beneficial phenotypes. The chromosomal-level assembly of the GIFT strain exposes its ancestry and potential functions for

Table 1 Summary table representing the transition/transversions values of the isolated SNPs markers

1	A>T	148,581	Transition
2	C>G	122,599	Transition
3	G>C	122,247	Transition
4	T>A	147,913	Transition
5	A>C	130,005	Transversion
6	A>G	417,100	Transversion
7	C>A	131,462	Transversion
8	C>T	444,180	Transversion
9	G>A	444,045	Transversion
10	G>T	131,825	Transversion
11	T>C	417,563	Transversion
12	T>G	130,073	Transversion

introgressed areas in the development of certain phenotypes (Etherington et al., 2022). Beneficial genes can be found and introgressed into Nile tilapia through breeding programmes in order to develop strains that are high yielding and resistant to the specific environmental stressor. This can be done by genetically analysing fish species that are endemic to various environmental conditions characteristic of the Kenyan waters. Thus pangenome analysis serves as a framework to determine and understand genomic diversity. In the current study, the phylogenomic tree developed from the shared SNP markers showed significant separation between the fish species. The tree showed *P. altivelis* as the most distant relative of the experimental fish. *L. chalumnae* and *S. chuatsi* diverged in time together with *D. rerio*, followed by *B. splendens*. Our experimental sample was in the same clade with *O. niloticus* (Fig. 2). The result of this study suggests a possibility of exchange of genetic material during their evolutionary development; since *O. niloticus* is a widely spread species with the ability to colonize multiple environments, it has increased chances of exchanging its genetic material with other species indicating genetic hybridization and possible genetic admixture ancestry.

A pangenome analysis was carried out and a total of 8045 core genes and 19,532 variable genes were identified among four species that showed a close genetic relationship (Fig. 3). Experimental sample and *O. aureus* had 59 common unique genes. Since the experimental sample is a strain of *O. niloticus*, the 59 genes probably have been introgressed from *O. aureus*. On the other hand, 18 genes were common between *Betta splendens* and the experimental sample probably as a consequence of hybridization between *O. niloticus* and *O. aureus*. Among the 18 genes we identified four introgressed genes that conferred beneficial traits to our experimental sample including the S100 calcium-binding protein B (S100B) that reduces *il6* production in malignant melanoma via inhibition of RSK cellular signalling, the triple functional domain protein-like that is involved in coordinating actin remodelling, which is necessary for cell migration and growth, peroxiredoxin-like 2A, an adipocyte-derived PAMM that may suppress macrophage activation by inhibiting MAPK signalling pathway and the coxsackievirus and adenovirus receptor homolog that is thought to regulate the cytoskeleton through interactions with actin and microtubules in fish. These genes play a major role in the enhanced growth performance, and their introgression has led to the development of superior growing fish strains that have the ability to withstand high stocking densities and high ammonia concentrations. A total of 96 genes were common to *B. splendens*, *O. aureus*, and the experimental sample but were absent in *O. niloticus*. Comparison to *O. niloticus* lost a total of 6118 genes, while it gained a total of 4179 genes. We hypothesize that the hybridized genes are the basis of the beneficial phenotypes exhibited by experimental sample in terms of resilience to stress, immunity, and faster growth rate in comparison with *O. niloticus* (Fig. 3). Among the unique genes found in experimental sample in this study and absent in *O. niloticus* in this study were classified according to the functions they augmented either immunity, stress resistance, or growth-making experimental sample better adapted to the local environmental conditions.

Table 2 Summary table representing the SNPs numbers representing the SNPs count and the transition/transversion (ts/tv) ratio

Callset	SNPs			Indels		MNPs	Others
	n	ts/tv	(1st ALT)	n	frm*		
snps	2,787,593	1.62	1.62	0	–	0	0
Callset	Singletons (AC = 1)			Indels		Multiallelic	
	SNPs	ts/tv		Indels		Sites	SNPS
snps	100.0%	1.62		0.0%		0	0

*Frameshift ratio: out/(out + in)

Some of the selected genes exhibiting substantial effect on immunity include: peroxiredoxin, macrophage mannose receptor 1-like, and zinc finger MYM-type protein 4-like partial, while stress-reactive genes included Palladin like gene, large proline-rich protein BAG6 like, and growth-related genes such as nuclear factor 1-x-type like, pigment epithelium-derived factor, and cathepsin L like.

O'Leary et al. (2014) demonstrated that peroxiredoxin-1-like has the ability to directly interact with other proteins, which may have an impact on other cellular processes including apoptosis, iron metabolism, proliferation, and the growth and operation of tissues, organs and systems, in pathogen infection, as well as in defence against cell death, tissue healing after injury, and tumour growth; they may function as inflammatory modulators. The current study has demonstrated the introgression of these genes into the genome of the experimental fish. This may confer better immune response of the experimental fish compared to its nilotic parent. According to Krata et al. (2022) peroxiredoxins modulate oxidative stress and can also be used as indicators of oxidative stress in humans. Peroxiredoxins are crucial regulators of oxidative stress because they effectively reduce peroxides levels, thereby regulating the peroxidases signalling. Peroxiredoxins account for around 90% of the decline of peroxidases activity in the body (Perkins et al., 2015). Peroxiredoxin has also been demonstrated to act as a chaperon for apurinic endonuclease (APE 1) which is essential for the activity of interleukin-8 (il-8) and nuclear factor kappa B (nf- κ B) O'Leary et al. (2014). Il-8 and nf- κ B play an important role in ammonia and heat stress resistance in fish (Esam et al., 2022). Therefore, peroxiredoxin is an important component in stress management as it is directly involved in oxidative, ammonia, and thermal stress management. The introgression of peroxiredoxin genes into the genome of *O. niloticus* and related cichlids could be beneficial in enhancing their resistance to ammonia stress. Kenyan aquaculture, which is mainly pond based, suffers elevated ammonia levels which lowers fish productivity. Breeding for ammonia tolerance would be a useful addition to the battery of tools available for increasing fish productivity.

Palladin-like protein that attracts VASP to these sites to aid in the formation of dorsal stress fibres (Price & Brindle, 2000). When stress fibres started to form in Rcho-1 cells of adult mice, palladin expression was seen to rise. Palladin is widely distributed in developing tissues, which raises the possibility that this protein plays a unique role in the organization of the actin cytoskeleton in cells that are differentiating structurally and functionally. Large proline-rich protein BAG6-like endoplasmic reticulum stress-induced pre-emptive quality control. BAG6 is also involved in endoplasmic reticulum stress-induced

pre-emptive quality control, a process that reroutes freshly produced proteins to the cytosol for proteasomal destruction while selectively attenuating their translocation into the endoplasmic reticulum. Misfolded proteins are broken down when they are moved from the endoplasmic reticulum and directed to the cytoplasmic proteasomes for destruction. BAG6 alters the dynamics of stress granules brought on by environmental challenges such as high temperatures, oxidative stress, osmolarity, and viral infections (Mediani et al., 2020). This study revealed our experimental fish had the BAG6 gene introgressed into its genome implying a superior stress management as compared to its nilotic parents. We hypothesize that the introgression of the BAG6 gene into tilapia and related cichlids, would improve their stress resistance and thus improve productivity.

Cathepsin L, a cysteine protease belonging to the papain superfamily, is essential for carrying out typical cellular processes such as general protein turnover, antigen processing, and bone remodelling essential to the architecture and function of the heart plays a crucial part in the morphogenesis and cycling of hair follicles as well as the differentiation of the epidermis. Endopeptidase activity formed as a pro-enzyme in lysosomes. A papain-like lysosomal enzyme called cathepsin L (CTS-L) breaks down endocytosed proteins to produce immunogenic antigens for adaptive immunity (Zhu et al., 2023a). According to Chen et al. (2020), Cathepsin has also been reported to be linked with the presentation and processing of antigens, as well as the control of immune responses in turbot fish. The endosomal/lysosomal system's non-specific bulk proteolysis, which breaks down both intracellular and extracellular proteins, is predominantly carried out by cathepsins, which are primarily intracellular enzymes. However, because of the restricted proteolysis processing, cathepsins play a role in the production of immune modulators. As a result, cathepsin L is a crucial gene that affects how the innate immune system responds.

MMR1-like enables macrophages to phagocytose and endocytose glycoproteins binds polysaccharide chains that are both sulphated and unsulphated. MMR1-like interacts with the glycoproteins and glycolipids that are present on the surface of bacteria, fungi, and viruses that cause disease. Studies on zebra fish show that the MMR exhibits expression in every tissue investigated and shares highly conserved structures with MMRs from other species (Zheng et al., 2015). The fact that MMR is expressed in the kidney and spleen indicates that it is involved in the immunological responses to infection. Nuclear factor 1 x-type like (*NFix*) recognizes and binds the palindromic sequence and has been shown to play an important role during development especially in stem

cell differentiation, maturation, and self-renewal (Harris et al., 2015).

Yu et al. (2022) demonstrated a mixed ancestry for the Sukumandi strain which had genes from Nile tilapia (*O. niloticus*) with approximately 0.36% of the genome having been derived from the blue tilapia (*O. aureus*). These shared genes may have arisen from the interbreeding between *O. niloticus* and *O. aureus*. About 0.11% of the genome was derived from *B. splendens*. Hybridization of female *O. niloticus* and male *O. aureus* has been used to obtain predominantly male off springs that are desirable for aquaculture. This aspect of hybridization between *O. niloticus* and *O. aureus* also happened in the wild environments. The experimental sample has accumulated genes from other fish species, making it more adaptive to environmental stressors as compared to *O. niloticus*.

Data provided from this study indicate the existence of genetic admixture ancestry in experimental sample. From the results of the current study, 155 genes are shared between experimental sample, *B. splendens* and *O. aureus*, which are absent in *O. niloticus*. This clearly demonstrates that the alleles have been introgressed into *O. niloticus* from *O. aureus* and *B. splendens*. In the wild *O. niloticus*, *B. splendens* and *O. aureus* live in the same environment. There is tangible evidence that these fish can interbreed to produce an offspring of mixed ancestry and in the process acquire and transmit genes that are not traditionally found in them. The resultant effect of these introgression is the production of a more resilient fish that is able to not only withstand environmental stressors but also yield highly. This study lays a foundation for the selection of genes conferring beneficial traits that would enable fish to survive in different regions of Kenya. With this background it is plausible to design a breeding programme that target genes conferring these advantageous traits to the particular fish. Desirable traits such as fast growth, resistance to disease and parasite, tolerance to high salinity, high stocking densities, and high levels of ammonia tolerance are important in the improvement of the aquaculture industry. There are established methods of delivering selected genes such as Tol2 system and CRISPR-Cas9 (Li et al., 2016). Tol2 transgenesis system is based on Tol2 transposon, a mobile genetic element that has been found to function in zebra fish and other fish species (Keng et al., 2009). The CRISPR-Cas9 on the other hand is an RNA-directed endonuclease that can generate double-stranded breaks in the genome. It can also be used for whole-genome genetic screens. CRISPR-Cas9 can therefore be a useful tool in the quick introgression of desirable genes into the genome of fish. Once a desirable gene has been identified and extracted from a related fish species. CRISPR-Cas9 can be employed to precisely transfer these genes into specific locations on

the genome of the receiving fish. CRISPR-Cas9 has been successfully used in disruption of myostatin gene in common carp resulting in significantly phenotypes (Zhong et al., 2016). Similarly Khalil et al. (2017) demonstrated the same effect in channel cat fish using the CRISPR-Cas9 tools. In Nile tilapia CRISPR-Cas9 has successfully been used to deliver mutation to *foxl2* and *dmrt1* genes, and these mutations were shown to be transmitted through the germ line (Li et al., 2014). CRISPR-Cas9 can therefore be used to accurately edit specific genes in specific locations to yield a desired result within a short period of time and in a cost-effective manner. Besides being used in gene “knock outs”, CRISPR-Cas system can also be used for gene “knock ins” utilizing the specificity of CRISPR-Cas to introduce a double-stranded break at a specific location of the DNA and the precision of the homology-directed repair (HDR) mechanism (Banan, 2020). In this case the HDR pathway is activated in the presence of a donor template carrying the genetic material to be inserted. Polymerization is initiated by the 3' end of the broken strand invading the intact donor template and repairing the damaged strand (Albadri et al., 2017). A major shortcoming of this method is that HDR works only in dividing cell specifically S and G2 phases of cell cycle.

Salinity tolerance

B. splendens and *D. rerio* species have a better salinity tolerance compared to *O. niloticus* and *O. aureus*. *B. splendens* has a solute carrier family 25-member 24 (*Slc25a24*) gene, an ATP Mg/Pi carrier, which is involved in the control of energy metabolism which enable it to cope with the salinity stress (Yu et al., 2022). This gene is absent in *O. aureus* and *O. niloticus* but is present in *D. rerio* species. On the other hand, *D. rerio* also has solute carrier family 12 member 10, tandem duplicate 1 gene which plays a key role in sodium, potassium, and chloride ions homeostasis. The protein expressed from this gene enables sodium potassium symporter activity. It is the authors considered opinion that the introgression of these two genes into the genome of *O. niloticus* could improve its salinity tolerance and allow it to be adaptable to saline water environments. In result, this will expand the aquaculture scope from not only fresh water bodies but also to marine and brackish water environments.

Hypoxia tolerance

D. rerio is more tolerant to hypoxic conditions relative to *O. niloticus*, *O. aureus*, and *B. splendens*. This can be associated with the presence of genes such as ATP-binding cassette transporter a1 (*Abca1*), B-cell lymphoma 2a (*bcl2a*), and (*mylk3*). *Abca1* is also present in *O. aureus* and has been shown to increase hypoxia tolerance by

increasing cholesterol efflux from peripheral cells, particularly foam cells in atherosclerotic plaques (Bogomolova et al., 2019). *Abca1* helps in clearing cholesterol in the atherosclerotic plaques and loading it to the *ApoA-I* which transports it to the liver. This clearance reduces the thickness of the arterial walls allowing for better oxygen diffusion (Linton et al 2019). *Bcl2a* prevents cell death by blocking oxidative stress-induced apoptosis (Susnow et al., 2009). On the other hand, *mylk3* gene has been shown to be a key gene in the regulation of cardiac myocyte contraction. This enables the heart to contract faster and more intensely to deliver more blood and thus oxygen to the body cells (Tobita et al., 2017). The evidence of the beneficial traits conferred by these genes and their absence in *O. aureus* and *O. niloticus* provide a basis to aver that their introgression would improve hypoxia tolerance of *O. niloticus*. This will help in overcoming the challenge of hypoxia stress which is a limiting factor in tilapia aquaculture intensification. Taken together, hybridization of salinity and hypoxia tolerance genes into *O. niloticus* would benefit the aquaculture industry by providing strains that would not only survive in these stressful conditions but also yield highly. Breeding for salinity tolerance and hypoxia tolerance would be a suitable area of interest that will enable aquaculturists to generate parental generations with these genes and thus can pass them down naturally to subsequent generations. Breeding for a specific trait would be a faster and more accurate method compared to pedigree-based methods (Vallejo et al., 2017; Yoshida et al., 2018). Pedigree-based breeding techniques are time-consuming, costly, and inaccurate in the transmission of a particular characteristic. Pedigree techniques convey other genes in addition to the beneficial features, which might dilute the benefits of earlier breeding efforts on a particular strain. On the other hand, breeding for a particular characteristic offers the chance to expand on current efforts with little chance of losing other advantageous traits currently present in the organism. While pedigree breeding has been used for a long time, it is time-consuming, laborious, and costly especially for the requirement of maintaining proper pedigree records. Where large numbers of progenies are involved, extra care is required and its success hinges on the skill of the breeder (Zenger et al., 2019). On the other hand trait-specific breeding enables early selection, which is economical and speeds up genetic progress by allowing animals to be utilized as parents sooner (García-Balasteros et al., 2021; Zenger et al., 2019). Whereas there are methods for the introgression of new alleles in to fish to improve productivity and stress tolerance, there are controversies regarding the safety of consumption of genetically modified organisms, health, and environmental impact. The attendant effect is however not investigated

conclusively though they are thought to have far-reaching ramifications such as the rise in cancer cases in human beings (Sitinjak et al., 2023). Introduction of genetically modified fish in the natural ecosystem has been found to alter the population structure of the wild fishes (Devlin et al., 2004). This is associated with the genetically modified fish having advantageous phenotypes that enable them thrive better than the wild species. This is usually accompanied by shrinking of the gene pool which may be a threat to the species in question.

Conclusions

The experimental sample is predominantly of *O. niloticus* origin. Introgressions from blue tilapia and *B. splendens* have given the experimental sample advantage over Nile tilapia. The experimental sample has displayed superior phenotypes in growth performance, immunity, and stress tolerance. Therefore, in light of a growing global population and a constantly changing climate, the development of variable genes and their associated SNPs can significantly aid in the improvement of fish breeds. These discoveries establish the groundwork for the creation of breeding programmes that would enhance the Nile tilapia and other related species' capacity for growth, immunity, and stress tolerance. With the ever-changing weather patterns, it is prudent for breeders to come up with fish breeds that are climate ready and are able not only to weather the shocks of changing climate but also produce highly. This will allow aquaculture to contribute its rightful share in the attainment of food and nutrition security in line with SDG 2 as well as improve the economic status of the fish-farming communities to address SDG 1 which are also highlighted in items 57 and 60 of the Agenda 64 of the United Nations.

Abbreviations

Abca1	ATP-binding cassata transporter a1
ATP	Adenosine tri-phosphate
BCL2a	B-cell lymphoma 2a
CRISPR-Cas9	Clustered regulatory interspaced short palindromic repeats and CRISPR: associated protein 9.
CTS-L	Cathepsin L like
DNA	Deoxyribonucleic acid
IL-6	Interleukin-6
IL-8	Interleukin-8
LPLBAG6	Large proline-rich protein BAG6 like
MAPK	Mitogen-activated protein kinase
MMR1	Macrophage mannose receptor 1
Mylk3	Myosin light chain kinase 3
NF1X	Nuclear factor 1-x-type
NF-kb	Nuclear factor kappa B
PALLD	Palladin like
PAMM	Peroxiredoxin activated in M-CSF-stimulated monocytes
PEDF	Pigment epithelium-derived factor
PPT	Parts per thousands
PRXS	Peroxiredoxin
RSK	Ribosomal s6 kinase
SDG	Sustainable development goals

Slc12	Solute carrier family 12 member 10
Slc25a24	Solute carrier family 25 member 24
SNPs	Single nucleotide polymorphism
VASP	Vasodilator-stimulated phosphoprotein
ZMYM4	Zinc finger MYM-type containing 4
HDR	Homology-directed repair

Acknowledgements

The author wishes to thank Kenya Climate Smart Agricultural Project (KCSAP) for the financial support during the study. The authors also wish to express their gratitude to the Department of Fisheries Kakamega County for hosting the research setup. Finally, the authors would like to appreciate the Department of Biological Sciences, Masinde Muliro University of Science and Technology, for providing valuable technical assistance and laboratory facilities throughout the research period.

Author contributions

JGM conceptualized the study, collected data, analysed the data, and wrote the original manuscript. CSW gave technical support provided expert advice and critical review. PAO conceptualized the study, curated the data, provided expert advice, critical review, and supervised the work. PO conceptualized the study, curated the data, provided expert technical advice, critical review, and supervised the work. AP curated the data, provided expert advice, and critical review. KK curated the data, provided technical advice, and critical review. All the authors read and approved the manuscript.

Funding

This study was funded by the Kenya Climate Smart Agricultural Project (KCSAP).

Availability of data and materials

The data set analysed during the current study are available from the corresponding author upon request.

Declarations

Ethics approval and consent to participate

The Masinde Muliro University of Science and Technology Institutional Ethics and Review Committee (MMUST-IERC) (REF: MMU/COR: 403012 Vol 5 (01)) approved the animal procedures, and all tests were carried out in accordance with the regulations.

Consent for publication

Not applicable.

Competing interests

The authors declare no conflicts of interest.

Author details

¹Department of Biological Sciences, School of Natural and Applied Sciences, Masinde Muliro University of Science and Technology, P.O. Box 190, Kakamega 50100, Kenya. ²Department of Fisheries, County Government of Kakamega, P.O. Box 190, Kakamega 50100, Kenya. ³Department of Zoology, School of Biological and Physical Sciences, Maseno University, P.O. Box, Private Bag, Maseno, Kenya. ⁴Max Planck Institute for Chemical Ecology, Hans-Knöll-Straße 8, 07745 Jena, Germany.

Received: 8 May 2023 Accepted: 12 September 2023

Published online: 08 October 2023

References

- Aketch, B. O., Angienda, P. O., Radull, J. O., & Waindi, E. N. (2014). Effect of stocking density on the expression of glucose transporter protein 1 and other physiological factors in the Lake Victoria Nile tilapia, *Oreochromis niloticus* (L.). *International Aquatic Research*, 6(2), 1–8. <https://doi.org/10.1007/s40071-014-0069-3>
- Albadri, S., Del Bene, F., & Revenu, C. (2017). Genome editing using CRISPR/Cas9-based knock-in approaches in zebrafish. *Methods*, 121–122, 77–85. <https://doi.org/10.1016/j.ymeth.2017.03.005>
- Aloo, P. A., Njiru, J., Balirwa, J. S., & Nyamweya, C. S. (2017). Impacts of Nile Perch, *Lates niloticus*, introduction on the ecology, economy and conservation of Lake Victoria, East Africa. *Lakes and Reservoirs*, 22, 320–333.
- Angienda, P. O., Lee, H. J., Elmer, K. R., Abila, R., Waindi, E. N., & Meyer, A. (2011). Genetic structure and gene flow in an endangered native tilapia fish (*Oreochromis esculentus*) compared to invasive Nile tilapia (*Oreochromis niloticus*) in Yala swamp, East Africa. *Conservation Genetics*, 12, 243–255.
- Banan, M. (2020). Recent advances in CRISPR/Cas9-mediated knock-ins in mammalian cells. *Journal of Biotechnology*, 308, 1–9. <https://doi.org/10.1016/j.jbiotec.2019.11.010>
- Bansal, V., & Boucher, C. (2019). Sequencing technologies and analyses: Where have we been and where are we going? *iScience*, 18, 37–41. <https://doi.org/10.1016/j.isci.2019.06.035>
- Benson, G. (1999). Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research*, 27(2), 573–580. <https://doi.org/10.1093/nar/27.2.573>
- Bogomolova, A. M., Shavva, V. S., Nikitin, A. A., Nekrasova, E. V., Dizhe, E. B., Larionova, E. E., Kudriavtsev, I. V., & Orlov, S. V. (2019). Hypoxia as a factor involved in the regulation of the apoA-1, ABCA1, and complement C3 gene expression in human macrophages. *Biochemistry*, 84(5), 529–539. <https://doi.org/10.1134/S0006297919050079>
- Chen, J., Zhang, L., Yang, N., Cao, M., Tian, M., Fu, Q., Su, B., & Li, C. (2020). Characterization of the immune roles of cathepsin L in turbot (*Scophthalmus maximus* L.) mucosal immunity. *Fish Shellfish Immunology*, 97, 322–335. <https://doi.org/10.1016/j.fsi.2019.12.005>. Epub 2019 Dec 2 PMID: 31805413.
- Devlin, R. H., D'Andrade, M., Uh, M., & Biagi, C. A. (2004). Population effects of growth hormone transgenic coho salmon depend on food availability and genotype by environment interactions. *Proceedings of the National Academy of Sciences*, 101(25), 9303–9308. <https://doi.org/10.1073/pnas.0400023101>
- El-Sayed, A. F. M. (2006). *Tilapia Culture*. CABI.
- Esam, F., Khalafalla, M. M., Gewaily, M. S., Abdo, S., Hassan, A. M., & Dawood, M. A. (2022). Acute ammonia exposure combined with heat stress impaired the histological features of gills and liver tissues and the expression responses of immune and antioxidative related genes in Nile tilapia. *Ecotoxicology and Environmental Safety*, 231, 113187. <https://doi.org/10.1016/j.ecoenv.2022.113187>
- Etherington, G. J., Nash, W., Ciezarek, A., Mehta, T. K., Barria, A., Peñaloza, C., Khan, M. G. Q., Durrant, A., Forrester, N., Fraser, F., Irish, N., Kaithakottil, G. G., Lipscombe, J., Trong, T., Watkins, C., Swarbreck, D., Angiolini, E., Cnaani, A., Gharbi, K., ... Haerty, W. (2022). Chromosome-level genome sequence of the Genetically Improved Farmed Tilapia (GIFT, *Oreochromis niloticus*) highlights regions of introgression with *O. mossambicus*. *BMC Genomics*, 23(1), 832. <https://doi.org/10.1186/s12864-022-09065-8>
- García-Ballesteros, S., Fernández, J., Toro, M. N., & Villanueva, B. (2021). Benefits of genomic evaluation in aquaculture breeding programs with separate rearing of families. *Aquaculture*, 543, 737004. <https://doi.org/10.1016/j.aquaculture.2021.737004>
- Garg, K., Green, P., & Nickerson, D. A. (1999). Identification of candidate coding region single nucleotide polymorphisms in 165 human genes using assembled expressed sequence tags. *Genome Research*, 9(11), 1087–1092. <https://doi.org/10.1101/gr.9.11.1087>
- Harris, L., Genovesi, L. A., Gronostajski, R. M., Wainwright, B. J., & Piper, M. (2015). Nuclear factor one transcription factors: Divergent functions in developmental versus adult stem cell populations. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, 244(3), 227–238. <https://doi.org/10.1002/dvdy.24182>
- Havel, J. E., Kovalenko, K. E., Thomaz, S. M., Amalfitano, S., & Kats, L. B. (2015). Aquatic invasive species: Challenges for the future. *Hydrobiologia*, 750(1), 147–170. <https://doi.org/10.1007/s10750-014-2166-0>
- Hohenlohe, P. A., Funk, W. C., & Rajora, O. P. (2021). Population genomics for wildlife conservation and management. *Molecular Ecology*, 30(1), 62–82. <https://doi.org/10.1111/mec.15720>
- Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics (oxford, England)*, 17(8), 754–755. <https://doi.org/10.1093/bioinformatics/17.8.754>

- Hurgobin, B., & Edwards, D. (2017). SNP discovery using a pangenome: Has the single reference approach become obsolete? *Biology*, 6(1), 21. <https://doi.org/10.3390/biology6010021>
- Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587–589. <https://doi.org/10.1038/nmeth.4285>
- Keng, V. W., et al. (2009). Efficient transposition of Tol2 in the mouse germline. *Genetics*, 183(4), 1565–1573.
- Khalil, K., Elayat, M., Khalifa, E., Daghash, S., Elswad, A., Miller, M., Abdelrahman, H., Ye, Z., Odin, R., Drescher, D., Vo, K., Gosh, K., Bugg, W., Robinson, D., & Dunham, R. (2017). Generation of Myostatin Gene- Edited Channel Catfish (*Ictalurus punctatus*) via Zygote Injection of CRISPR/Cas9 System. *Scientific Reports*. <https://doi.org/10.1038/s41598-017-07223-7>
- Krata, N., Foroniewicz, B., Zagodzón, R., Moszczuk, B., Zielenkiewicz, M., Pączek, L., & Mucha, K. (2021). Peroxiredoxins as markers of oxidative stress in iga nephropathy, membranous nephropathy and lupus nephritis. *Archivum Immunologiae Et Therapiae Experimentalis*, 70(1), 3. <https://doi.org/10.1007/s00005-021-00638-1>
- Li, M., Yang, H., Zhao, J., Fang, L., Shi, H., Li, M., Sun, Y., Zhang, X., Jiang, D., Zhou, L., & Wang, D. (2014). Efficient and heritable gene targeting in tilapia by CRISPR/Cas9. *Genetics*, 197(2), 591–599. <https://doi.org/10.1534/genetics.114.163667>
- Li, M., Zhao, L., Page-McCaw, P. S., & Chen, W. (2016). Zebrafish genome engineering using the CRISPR-Cas9 system. *Trends in Genetics: TIG*, 32(12), 815–827. <https://doi.org/10.1016/j.tig.2016.10.005>
- Lind, C. E., Agyakwah, S. K., Attipoe, F. Y., Nugent, C., Crooijmans, R. P. M. A., & Toguyeni, A. (2019). Genetic diversity of Nile tilapia (*Oreochromis niloticus*) throughout West Africa. *Science and Reports*, 9(1), 16767. <https://doi.org/10.1038/s41598-019-53295-y>. PMID: 31727970; PMCID: PMC6856548.
- Linton, M. F., Yancey, P. G., Davies, S. S., Jerome, W. G., Linton, E. F., Song, W. L., Doran, A. C., & Vickers, K. C. (2019). The role of lipids and lipoproteins in atherosclerosis. In K. R. Feingold, B. Anawalt, M. R. Blackman, A. Boyce, G. Chrousos, E. Corpas, W. W. de Herder, K. Dhatariya, K. Dungan, J. Hofland, S. Kalra, G. Kaltsas, N. Kapoor, C. Koch, P. Kopp, M. Korbonits, C. S. Kovacs, W. Kuohung, B. Laferrère, ... D. P. Wilson (Eds.), *Endotext*. South Dartmouth: MDText.com, Inc.
- Mangang, Y. A., & Pandey, P. K. (2021). Hemato-biochemical responses and histopathological alterations in the gill and kidney tissues of *Osteobrama belangeri* (Valenciennes, 1844) exposed to different sub-lethal ionized ammonia. *Aquaculture*, 542, 736887. <https://doi.org/10.1016/j.aquaculture.2021.736887>
- Martin, C. W., Valentine, M. M., & Valentine, J. F. (2010). Competitive interactions between invasive Nile tilapia and native fish: The potential for altered trophic exchange and modification of food webs. *PLoS ONE*, 5(12), e1395. <https://doi.org/10.1371/journal.pone.0014395>
- Mayjonade, B., Gouzy, J., Donnadiou, C., Pouilly, N., Marande, W., Callot, C., Langlade, N., & Muñoz, S. (2016). Extraction of high-molecular-weight genomic DNA for long-read sequencing of single molecules. *BioTechniques*, 61(4), 203–205. <https://doi.org/10.2144/000114460>
- Mediani, L., Galli, V., Carrà, A. D., Bigi, I., Vinet, J., Ganassi, M., Antoniani, F., Tiago, T., Cimino, M., Mateju, D., Cereda, C., Pansarasa, O., Alberti, S., Mandrioli, J., & Carra, S. (2020). BAG3 and BAG6 differentially affect the dynamics of stress granules by targeting distinct subsets of defective polypeptides released from ribosomes. *Cell Stress & Chaperones*, 25(6), 1045–1058. <https://doi.org/10.1007/s12192-020-01141-w>
- Mitra, I., Huang, B., Mousavi, N., Ma, N., Lamkin, M., Yanicky, R., Shleizer-Burko, S., Lohmueller, K. E., & Gymrek, M. (2021). Patterns of de novo tandem repeat mutations and their role in autism. *Nature*, 589(7841), 246–250. <https://doi.org/10.1038/s41586-020-03078-7>
- Munguti, J. M., Nairuti, R., Iteba, J. O., Obiero, K. O., Kyule, D. D., Opiyo, M. A., Abwao, J., Kirimi, J. G., Outa, N., Muthoka, M., Githukia, C. M., & Ogello, E. O. (2022). Nile tilapia (*Oreochromis niloticus* Linnaeus, 1758) culture in Kenya: Emerging production technologies and socio-economic impacts on local livelihoods. *Aquaculture, Fish and Fisheries*, 2(4), 265–276. <https://doi.org/10.1002/aff2.58>
- Mwaura, J. G., Wekesa, C., Ogotu, P. A., & Okoth, P. (2023). Whole transcriptome analysis of differentially expressed genes in cultured Nile tilapia (*O. niloticus*) subjected to chronic stress reveals signaling pathways associated with depressed growth. *Genes*, 14, 795. <https://doi.org/10.3390/genes14040795>
- O'Leary, P. C., Terrile, M., Bajor, M., Gaj, P., Hennessy, B. T., Mills, G. B., Zagodzón, A., O'Connor, D. P., Brennan, D. J., Connor, K., Li, J., Gonzalez-Angulo, A. M., Sun, H. D., Pu, J. X., Pontén, F., Uhlén, M., Jirström, K., Nowis, D. A., Crown, J. P., ... Gallagher, W. M. (2014). Peroxiredoxin-1 protects estrogen receptor α from oxidative stress-induced suppression and is a protein biomarker of favorable prognosis in breast cancer. *Breast Cancer Research: BCR*, 16(4), R79. <https://doi.org/10.1186/bcr3691>
- Perkins, A., Parsonage, D., Nelson, K., Ogba, O., Cheong, P. Y., Poole, L., & Karplus, P. (2016). Peroxiredoxin Catalysis at Atomic Resolution. *Structure*, 24(10), 1668–1678. <https://doi.org/10.1016/j.str.2016.07.012>
- Prabu, E., Rajagopalsamy, C. B. T., Ahilan, B., Jeevagan, I. J. M. A., & Renuhadevi, M. (2019). Tilapia—An excellent candidate species for world aquaculture: A review. *Annual Research & Review in Biology*. <https://doi.org/10.9734/arrb/2019/v3i1330052>
- Price, C. J., & Brindle, N. P. (2000). Vasodilator-stimulated phosphoprotein is involved in stress-fiber and membrane ruffle formation in endothelial cells. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 20(9), 2051–2056. <https://doi.org/10.1161/01.atv.20.9.2051>. PMID: 10978248.
- Song, P., Li, M., Zhang, M., Jiang, H., Shao, J., Wang, R., Qian, Y., & Feng, D. (2022). Inhibition of argininosuccinate synthase (ASS) affected ammonia excretion in yellow catfish *Pelteobagrus fulvidraco* during acute ammonia poisoning. *Aquaculture Reports*, 22, 100931. <https://doi.org/10.1016/j.aqrep.2021.100931>
- Sitinkaj, B. D. P., Murdaya, N., Rachman, T. A., Zakiyah, N., & Barliana, M. I. (2023). The potential of single nucleotide polymorphisms (SNPs) as biomarkers and their association with the increased risk of coronary heart disease: A systematic review. *Vascular Health and Risk Management*, 19, 289–301. <https://doi.org/10.2147/vhrm.s405039>
- Susnow, N., Zeng, L., Margineantu, D., & Hockenbery, D. M. (2009). Bcl-2 family proteins as regulators of oxidative stress. *Seminars in Cancer Biology*, 19(1), 42–49. <https://doi.org/10.1016/j.semcancer.2008.12.002>
- Tarailo-Graovac, M., & Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. *Current Protocols in Bioinformatics*. <https://doi.org/10.1002/0471250953.bi0410s25>
- Tobita, T., Nomura, S., Morita, H., Ko, T., Fujita, T., Toko, H., Uto, K., Hagiwara, N., Aburatani, H., & Komuro, I. (2017). Identification of MYLK3 mutations in familial dilated cardiomyopathy. *Science and Reports*, 7(1), 17495. <https://doi.org/10.1038/s41598-017-17769-1>. PMID: 29235529; PMCID: PMC5727479.
- Twongo, T. (1995). Impact of fish species introductions on the tilapias of Lakes Victoria and Kyoga. In T. J. Pitcher & P. J. B. Hart (Eds.), *The impact of species changes in African lakes* (pp. 47–57). Chapman and Hall.
- Vallejo, R. L., Leeds, T. D., Gao, G., Parsons, J. E., Martin, K. E., et al. (2017). Genomic selection models double the accuracy of predicted breeding values for bacterial cold water disease resistance compared to a traditional pedigree-based model in rainbow trout aquaculture. *Genetics, Selection, Evolution*, 49, 17.
- Via, S. (2009). Natural selection in action during speciation. *Proceedings of the National Academy of Sciences*, 106(Suppl 1), 9939–9946. <https://doi.org/10.1073/pnas.0901397106>. Epub 2009 Jun 15. PMID: 19528641; PMCID: PMC2702801.
- Vinuesa, P., Ochoa-Sánchez, L. E., & Contreras-Moreira, B. (2018). *GET_PHYLO-MARKERS, a Software Package to Select Optimal Orthologous Clusters for Phylogenomics and Inferring Pan-Genome Phylogenies*. Used for a Critical Geno-Taxonomic Revision of the Genus *Stenotrophomonas*: *Frontiers in Microbiology*. <https://doi.org/10.3389/fmicb.2018.00771>
- Wenne, R. (2023). Single nucleotide polymorphism markers with applications in conservation and exploitation of aquatic natural populations. *Animals*, 13(6), 1089. <https://doi.org/10.3390/ani13061089>
- Wiens, J. J. (2004). What is speciation and how should we study it? *The American Naturalist*, 163(6), 914–923. <https://doi.org/10.1086/386552>
- Williams, L. M., Ma, X., Boyko, A. R., Bustamante, C. D., & Oleksiak, M. F. (2010). SNP identification, verification, and utility for population genetics in a non-model genus. *BMC Genetics*, 11, 32. <https://doi.org/10.1186/1471-2156-11-32>
- Yan, B., Wang, Z. H., & Zhao, J. L. (2013). Mechanism of osmoregulatory adaptation in tilapia. *Molecular Biology Reports*, 40(2), 925–931. <https://doi.org/10.1007/s11033-012-2133-7>
- Yoshida, G. M., Bangerla, R., Carvalheiro, R., Correa, K., Figueroa, R., Lhorente, J. P., & Yáñez, J. M. (2018). Genomic prediction accuracy for resistance against

- Piscirickettsia salmonis* in farmed rainbow trout. *G3 (Bethesda, Md.)*, 8(2), 719–726. <https://doi.org/10.1534/g3.117.300499>
- Yu, X., Setyawan, P., Bastiaansen, J. W., Liu, L., Imron, I., Groenen, M. A., Komen, H., & Megens, H. J. (2022). Genomic analysis of a Nile tilapia strain selected for salinity tolerance shows signatures of selection and hybridization with blue tilapia (*Oreochromis aureus*). *Aquaculture*, 560, 738527. <https://doi.org/10.1016/j.aquaculture.2022.738527>
- Zenger, K. R., Khatkar, M. S., Jones, D. B., Khalilisamani, N., Jerry, D. R., & Raadsma, H. W. (2019). Genomic selection in aquaculture: Application, limitations and opportunities with special reference to marine shrimp and pearl oysters. *Frontiers in Genetics*, 9, 693. <https://doi.org/10.3389/fgene.2018.00693>
- Zhang, M., Wang, S., Sun, L., Gan, L., Lin, Y., Shao, J., Jiang, H., & Li, M. (2022). Ammonia induces changes in carbamoyl phosphate synthetase I and its regulation of glutamine synthesis and urea cycle in yellow catfish *Pelteobagrus fulvidraco*. *Fish & Shellfish Immunology*, 120, 242–251. <https://doi.org/10.1016/j.fsi.2021.11.023>
- Zheng, F., Asim, M., Lan, J., Zhao, L., Wei, S., Chen, N., Liu, X., Zhou, Y., & Lin, L. (2015). Molecular cloning and functional characterization of mannose receptor in zebra fish (*Danio rerio*) during infection with *Aeromonas sobria*. *International Journal of Molecular Sciences*, 16(5), 10997–11012. <https://doi.org/10.3390/ijms160510997>
- Zhong, Z., Niu, P., Wang, M., Huang, G., Xu, S., Sun, Y., Xu, X., Hou, Y., Sun, X., Yan, Y., & Wang, H. (2016). Targeted disruption of sp7 and myostatin with CRISPR-Cas9 results in severe bone defects and more muscular cells in common carp. *Scientific Reports*. <https://doi.org/10.1038/srep22953>
- Zhu, C. S., et al. (2023a). Identification of procathepsin L (pCTS-L)-neutralizing monoclonal antibodies to treat potentially lethal sepsis. *Science Advances*, 9, eadf4313. <https://doi.org/10.1126/sciadv.adf4313>
- Zhu, L., Du, W., Li, Y., Li, D., Wei, W., Zhao, J., & Wang, X. (2023b). Chiral SPINOL-Based Pt(II) Metallacycles For Immunogenic Cell Death. *Inorganic Chemistry*. <https://doi.org/10.1021/acs.inorgchem.3c01635>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
