



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



True Fish Farming Story in Lake Victoria Basin (TRUEFISH)

RESULT AREA 3.3 IMPROVED PROTECTION OF
BIODIVERSITY: GENETIC SCREENING RESEARCH REPORT

ASILATU H SHECHONGE, CASSIUS ARUHO, KEVIN OBIERO, PAPIUS
DIAS TIBIHIKA, JOHN A.H. BENZIE, ROSE K. BASIITA, BENDULA
WISMEN, ADAM G CIEZAREK, WILFRIED HAERTY

MAY 2024



This project is funded by
the European Union



True Fish Farming Story in Lake Victoria Basin (TRUEFISH)

Result Area 3.3 Improved Protection of Biodiversity: Genetic Screening Research Report

May 2024

Authors:

Asilatu H. Shechonge¹, Cassius Aruho², Kevin Obiero³, Papius Dias Tibihika⁴, John A.H. Benzie⁵, Rose K. Basiita⁴, Bendula Wismen⁵, Adam G. Ciezarek⁶, Wilfried Haerty⁶

¹ Tanzania Fisheries Research Institute (TAFIRI), Tanzania

² National Fisheries Resources Research Institute (NaFIRRI), Uganda

³ Kenya Marine Fisheries Research Institute (KMFRI), Kenya

⁴ WorldFish East and Southern Africa, Zambia

⁵ WorldFish (HQ) Penang, Malaysia

⁶ Earlham Institute, United Kingdom



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Contents

Table of Figures.....	iii
List of Abbreviations.....	iv
1.0 TRUEFISH Project.....	1
1.1 Result Area 3.3: Improved Protection of Biodiversity.....	2
1.2 Research Area: Lake Victoria Basin	3
2.0 Fisheries and Aquaculture Sector in Lake Victoria Basin	4
2.1 Kenya.....	5
2.2 Tanzania	6
2.3 Uganda	6
2.4 Aquaculture in Lake Victoria	7
2.5 Approaches and Methods for Fish Identification	9
2.5.1 Application of Morphometric Approaches	9
2.5.2 Application of Molecular Genetic Approaches	9
3.0 Genetic Screening Research Methodology	11
3.1 Field Sampling.....	11
3.2 Sample Preparation	13
3.3 Genomic Analyses.....	13
3.3.1 Quality control and variant calling	13
3.3.2 Analysis of population genetic structure and hybridization	14
3.3.3 Determination of population genetic diversity	14
3.3.4 Development of a 96-SNP chip panel for species and hybrid assignment.....	15
4.0 Results and Discussion	16
4.1 Sample Collection, Sequencing, Mapping and SNP Calling	16
4.2 4.2 Genetic population structure.....	17
4.3 4.3 Genetic identification of hybrids	23
4.4 4.4 Population genomic statistics	26
4.5 4.5 Morphological versus genetic identification	28
4.5.1 96 SNP diagnostic panel of hybrid identification	29
5.0 Interpretation of the Genetic Screening Results.....	30
6.0 Data Availability	33
7.0 Acknowledgements.....	33
8.0 References.....	34
9.0 Annex 1: Detailed Research Plan.....	36



This project is funded by
the European Union



Table of Figures

Figure 1 Lake Victoria Basin and its catchment area across the five EAC countries. (Source: Nyamweya et al, 2023)	3
Figure 2 Freshwater aquaculture production in Kenya, Tanzania and Uganda. (Source: FAO Statistics, 2023)	4
Figure 3 Tilapia aquaculture production in Kenya, Tanzania and Uganda. (Source: FAO Statistics 2023)	8
Figure 4 Sampling locations across the Lake Victoria Basin, and specifically in the countries of Kenya, Tanzania, and Uganda	11
Figure 5 Extraction of fin clip sample.....	13
Figure 6 PCA of all individuals sampled from Lake Victoria Basin.	17
Figure 7 PCA of all individuals excluding Coptodon species.	18
Figure 8 PCA of <i>O. niloticus</i> individuals sampled.	19
Figure 9 ADMIXTURE analysis of all <i>O. niloticus</i> individuals, with coloured bars corresponding to different ancestry groups.	20
Figure 10 PCA analyses of all non-hybrid <i>O. niloticus</i> from a)Kenya, b) Uganda, and c)Tanzania. Black circles in (b) represent samples from Lake Albert.....	21
Figure 11 PCA analysis of all non-hybrid <i>O. leucostictus</i> samples.	22
Figure 12 PCA analysis of all non-hybrid <i>O. esculentus</i> samples.....	22
Figure 13 Map of all sampling sites coloured according to species combinations of hybrids sampled from LVB.	24
Figure 14 Violin plot of the inbreeding co-efficient, F , per individual in each sampling site with at least five pure <i>O. niloticus</i> individuals. Sites are clustered into countries: Kenya (red), Tanzania (blue) and Uganda (green). Sites indicated with an asterisk are farms.	26
Figure 15 PCA of all individuals (pure and genetic hybrids) for the a)96 SNP panel, and b) genome-wide <i>O. niloticus</i> mapping. Note that PC1 and PC2 are reversed in (b) for easy comparison.	29



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



List of Abbreviations

CV	Cross Validation Score
DNA	Deoxyribonucleic acid
EAC	East African Community
EI	Earlham Institute
GPS	Global positioning system
KMFRI	Kenya Marine Fisheries Research Institute
LVB	Lake Victoria Basin
LVFO	Lake Victoria Fisheries Organization
NGOs	non-governmental organizations
PCA	Principal Component Analysis
SNPs	single nucleotide polymorphism
UK	United Kingdom
VCF	Virtual Contact File



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



1.0 TRUEFISH Project

The True Fish Farming Story in Lake Victoria Basin (TRUEFISH) is a project financed by the European Development Fund, for the benefit of the East African Community (EAC), particularly the Lake Victoria Basin (LVB) in Kenya, Tanzania, and Uganda. This project is aimed at contributing to the development of a competitive, gender-equitable, and sustainable commercial aquaculture sector to support economic development and sustainable management of natural resources in the LVB. The TRUEFISH project was developed with three specific objectives, i) Improve access to commercial networks for aquaculture-related businesses, ii) Increase the availability of skilled workers, and iii) Improve sustainability by mitigating risks related to the aquatic environment, aquatic animal health conditions, and biodiversity. The specific objectives are developed in line with the 2009 Lake Victoria Fisheries Organization's Regional Aquaculture Strategy and Investment Plan and the EAC Fisheries and Aquaculture Policy (2018).

To achieve the specific objectives of the TRUEFISH project, the following components and result areas were identified.

Component 1. BUSINESS (access to commercial networks):

- Result 1.1: Strengthened commercial networks for competitive aquaculture-related businesses.

Component 2. SKILLS (availability and quality of local skilled workers in aquaculture-related businesses):

- Result 2.1: Upgraded practical training delivery by formal training institutions.
- Result 2.2: Strengthened linkages between training institutions, and with aquaculture business operators.

Component 3. SUSTAINABILITY (sustainable and bio-secure regional aquaculture production systems):

- Result 3.1: Strengthened aquatic animal health conditions.
- Result 3.2: Lake Victoria zoning: support an orderly and sustainable cage culture development.
- Result 3.3: Improved protection of biodiversity.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



1.1 Result Area 3.3: Improved Protection of Biodiversity

Activities designed and implemented under Result Area 3.3 were targeted to improve the protection of biodiversity resources, particularly the tilapias, in the LVB of East Africa region, leading towards sustainable and improved commercial aquaculture production and productivity. The key aim of Result Area 3.3 is to address the key risks to the long-term development and sustainability of the aquaculture sector, by minimising the peril of accidental or intentional introduction of genetic material which can threaten native species.

To achieve this broader objective, Result Area 3.3 is focused on building capacity and establishing a system for evaluating and monitoring the biodiversity of key species used for aquaculture, and undertaking an initial assessment, thus providing the foundations and direction for future sustainable aquaculture development.

The approach to Result Area 3.3 will link regional and international research organisations, with regional management bodies, governments, private sector, fish farmer organisations and non-governmental organizations (NGOs) in Kenya, Tanzania, and Uganda.

The genetic screening program, developed as part of Results Area 3.3, envisages the establishment of new or improved techniques that are pivotal to better meet the conservation needs of the diversity of tilapia species in the LVB, by:

- i) A report showing the spatial distribution of different species and genotypes of tilapias in the LVB.
- ii) Identification of fish populations and sites exposed to a high risk of loss of biodiversity potentially due to aquaculture.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



1.2 Research Area: Lake Victoria Basin

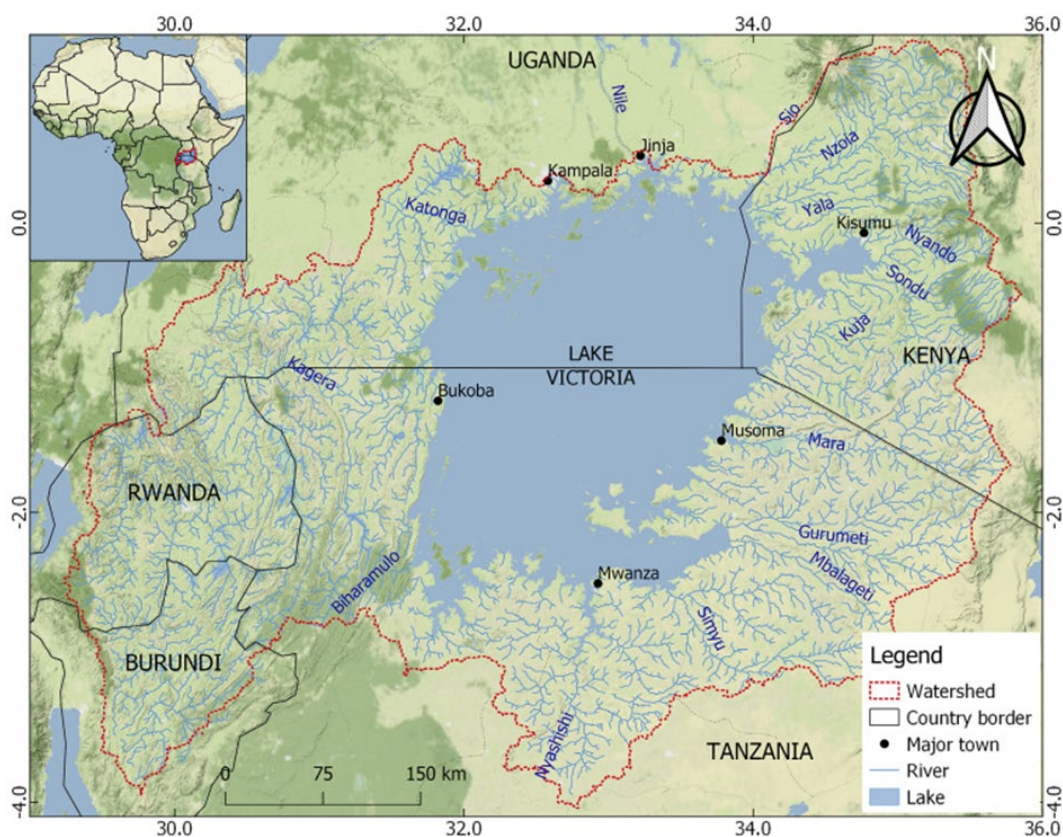


Figure 1 Lake Victoria Basin and its catchment area across the five EAC countries. (Source: Nyamweya et al, 2023)

Lake Victoria Basin (Figure 1) is in the upper reaches of the Nile River and comprises of the world's greatest complexes of lakes, wetlands, and rivers. The basin has a catchment area of approximately 194,200km² traversing through five East African Countries i.e. Tanzania (44%); Kenya (22%); Uganda (16%); Rwanda (11%) and Burundi (7%).

The Lake Victoria ecosystem covers an area of 68,800km² spanning 400km North-South and 240km East-West. It has a shoreline of 3,460km and is a relatively shallow water body with an average depth of 40m while the deepest point is 80m. This trans-boundary waterbody is shared by Kenya (6%), Tanzania (51%), and Uganda (43%). Over 80% of water in the Lake Victoria comes directly from rainfall while the rest comes from tributary runoff from Kagera, Katonga, Sio, Mara, Simiyu, Gurumeti, Yala, Nyando, Migori, and Sondu-Miriu rivers.

Lake Victoria is a vital resource, creating employment for over 200,000 fishers and supporting an estimated 35 million people across Kenya, Tanzania, and Uganda. In 2021, the lake's fisheries provided over 90% of the 1.08 million tons of fish harvested from inland waters in these countries (FAO, 2022).



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



2.0 Fisheries and Aquaculture Sector in Lake Victoria Basin

The current per capita fish consumption in the EAC region is remarkably low (average of 5.4 kg/capita/year) compared to the rest of Africa (10.1 kg/capita/year) and the 2021 global average of 20.5 kg/person/year (FAO, 2022) (Obiero, et al., 2019). Fish consumption habits vary across Africa, with 11 countries consuming more fish than the global average. The low consumption rate of fish and other aquatic foods in Africa can be attributed to several reasons, including population growth, decline in capture fisheries, the relatively small aquaculture sector, and fish supply deficits, among others. The consumption of fish and aquatic foods in Africa is projected to decline in the next 10 years, representing a serious threat in terms of food and nutrition security in the region (FAO, 2022). The increasing population growth and income levels and changes in lifestyles and preferences because of rapid urbanization and globalization imply that the EAC region will need over 5 million tonnes (1.7 MT in Uganda, 1.0 MT in Kenya and 2.5 MT in Tanzania) of fish to fill the demand–supply gaps (Ministry of Agriculture, Animal Industry and Fisheries, 2019). It is anticipated that sustained and targeted growth in the aquaculture sector will contribute to achieving food and nutrition security in the region.

Fish production from Lake Victoria was estimated at 2.23 million tonnes, worth US\$ 6.69 in 2021. Contributions to the national gross domestic product (GDP) of these countries are significant, with 0.7% in Kenya, 1.8% in Tanzania and 2.2% in Uganda.

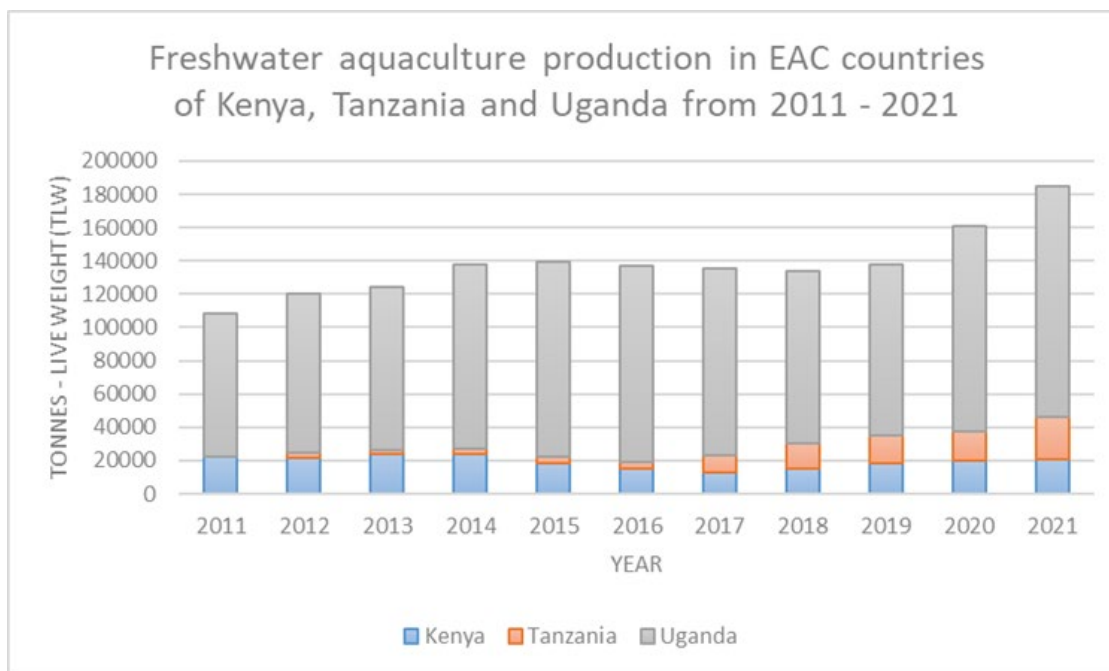


Figure 2 Freshwater aquaculture production in Kenya, Tanzania and Uganda. (Source: FAO Statistics, 2023)



This project is funded by
the European Union



Two non-native species, Nile perch (*Lates niloticus*) and Nile tilapia (*Oreochromis niloticus*), and one native cyprinid, the Lake Victoria sardine (*Rastrineobola argentea*) are commercially important fish species found in Lake Victoria. Despite the significant contribution of the fisheries sector to the economic development in the region, it faces challenges in sustainable exploitation and overall fish production. For instance, the Nile tilapia production from Lake Victoria fluctuated from an annual harvest of 71,500 metric tonnes in 2005 to 41,500 metric tonnes in 2021 (Figure 2). This is largely attributed to increased anthropogenic activities and differing priorities in the lake and has become a basis for the EAC to focus on increasing aquaculture production and productivity in the region. The governments of Kenya, Tanzania, and Uganda have enacted policies to support the expansion of the aquaculture sector in Lake Victoria. Increased investments from public and private sectors, and multilateral development organizations and international financial institutions, were aimed at intensifying culture systems and commercializing the aquaculture supply chain in the region (Obiero, et al., 2019).

2.1 Kenya

The fisheries and aquaculture sector in Kenya supports about 1.2 million people directly and indirectly working as fishers, traders, processors, suppliers and merchants of fishing accessories and their dependents. Recognizing the importance of this economic sector, the Government of Kenya has developed several policies and legal frameworks that support expansion of aquaculture and fisheries. These include the National Oceans and Fisheries Policy (2018), and the Fisheries Management and Development Act (2016) which established the Kenya Fisheries Advisory Council (“Council”), the Kenya Fisheries Service (“Service”), the Fish Marketing Authority (“Authority”), the Fisheries Research and Development Fund and the Fish Levy Trust Fund. The Act also implements obligations under international law concerning fisheries.

The latest policy, the Kenya Fisheries Policy 2023, was approved by the Cabinet on 18th April 2023. The policy seeks to provide a framework for sustainable management and development of fisheries and aquaculture to enhance food and nutrition security, employment creation and revenue generation. The policy identifies fisheries and aquaculture as potential sectors for socio-economic development of the country and therefore strives to build on the phenomenal development of the sector over the last two decades, where the value of national annual fish production increased from an estimated value of Kenya Shillings 9 billion in 2008 to 37.5 billion in 2022 (Kenya National Bureau of Statistics, Economic Survey, 2023).

A technical team was appointed in July 2023 to review the National Aquaculture Policy of 2011 to align with Kenya’s Vision 2030 targets, existing legal frameworks and sectoral policy goals



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



and regional and international commitments and taking into consideration the aquaculture status review report 2021. The current framework in place supports direct investment in the sector at all scales. Although aquaculture development is devolved to the County levels, the current framework provides for coordination between the Central and County Government structures of issues linked to national policy, such as environmental and biodiversity management (Munguti, et al., 2023).

2.2 Tanzania

Fisheries and aquaculture sector has been growing with an average annual rate of 1.5% and contributing towards improving food and nutrition security and social-economic wellbeing in Tanzania. The sector provides direct employment to 197,763 fisherfolk and 34,057 aquafarmers, and approximately 4.5 million Tanzanians are engaged in various activities associated with aquaculture supply chains. About 95% of fish production is from artisanal fishers and subsistence aquafarmers.

The Tanzanian National Five-Year Development Plan 2016/17 – 2020/21 recognizes the importance of fisheries and marine resources sector. Therefore, prioritized activities include inland fisheries, sea and deep-sea fishing, as well as aquaculture. Interventions for the acceleration of the sub-sector include, among others, improvement of research and training; increasing availability and utilization of fisheries inputs and implements, extension services, empowering fishers through subsidies to facilitate fishing activities and strengthening monitoring, control, and surveillance of fisheries resources and trade of fisheries products in marine and inland waters with better management of the respective environment. Specifically, for fisheries and aquaculture, the Development Plan has set targets for the sector over the Plan period and beyond. These include (i) Average growth rate (%) of 2.72, 5.67, and 6.0; (ii) Share of GDP (current prices) (%) of 2.9, 3.3, and 2.8; (iii) Share of total export earnings (%) of 3.7, 4.1, and 4.0; and (iv) Share of total employment (%) of 0.8, 1.5, and 2.0 in 2015, 2020 and 2025, respectively. As of January 2022, Tanzania had implemented legal and policy framework to support aquaculture research and development.

2.3 Uganda

Uganda's National Development Plan NDP III (2020/21 – 2024/25), has set out clear priorities for opportunities in the agriculture sector. Investment in these opportunities follows the entire supply chains, which have guided the identification of priority projects and interventions in these areas. Based on the NDP III, fish has been identified as one of the key 11 agricultural enterprises (others are cotton, coffee, tea, maize, cocoa, cassava, vegetable oil, beef, dairy,



This project is funded by
the European Union



and horticulture). Within these key agricultural sectors, the focus will be on strengthening agricultural research; extension system; technology adaptation at the farm level; increasing access to and effective use of critical farm inputs; increasing access to finance with specific options for women farmers; and strengthening agricultural institutions for effective coordination and service delivery.

While capture fisheries showed stagnating production over the last 10 years, aquaculture has improved in the same period and grew at a rate of 6% per annum from 5,500 metric tonnes in 2005, 120,000 metric tonnes in 2018 and to the current estimate of 124,000 metric tonnes (FAO, 2022). There is potential for the aquaculture sector to grow beyond the current levels against increased fish demand from domestic, regional, and international markets.

Uganda's Fisheries and Aquaculture Policy 2018 (Ministry of Agriculture, Animal Industry and Fisheries, 2019) states that, a total of 1,700,000MT will be required to meet the increasing demand for fish in the local and international markets. Approximately 700,000 metric tonnes are expected to come from inland fisheries while 1,000,000 metric tonnes will come from the aquaculture sector. The aquaculture industry will require close to 2.5 billion quality fingerlings (60% tilapia) for distribution to farmers to achieve the planned target. This will be followed by an array of strategies, including sustained preservation of wild sources from which the broodstocks will be sourced for selective breeding activities, to ensure the industry has a consistent supply of sufficient and quality fingerlings required to increase the production and the productivity. The FAP 2018 has also identified prioritised research areas including fish breeding, fish feed formulation, nutrition, and disease management.

2.4 Aquaculture in Lake Victoria

Fisheries output in EAC region has been dominated by capture fisheries, however, contributions of aquaculture to the total amount of fish produced in the region has grown at a steady pace over the past decade. Aquaculture production in the Lake Victoria Basin of the EAC region is dominated by Nile tilapia (*O. niloticus*) and African Catfish (*Clarias gariepinus*). In 2021, Nile tilapia aquaculture contributed approximately 140,000 metric tonnes Kenya, Tanzania, and Uganda, indicating, an increase of approximately 45,000 metric tonnes compared to production values recorded in 2011 (Figure 3).



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE

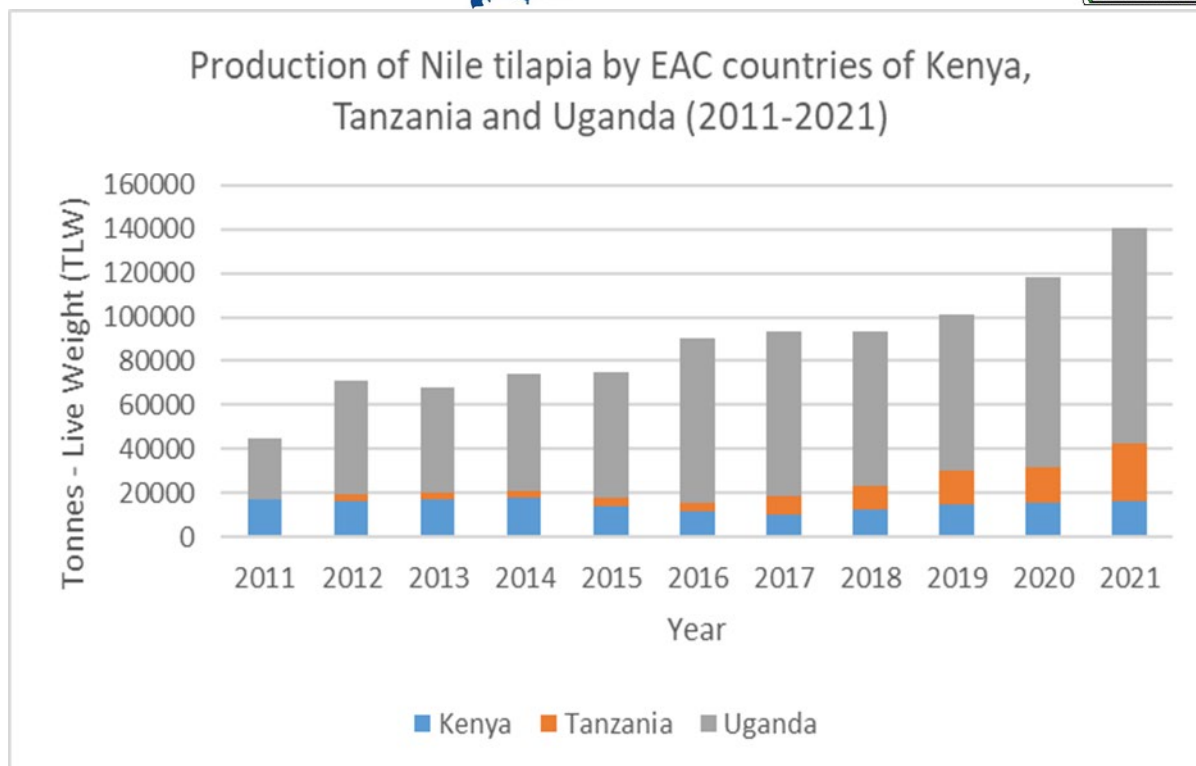


Figure 3 Tilapia aquaculture production in Kenya, Tanzania and Uganda. (Source: FAO Statistics 2023)

Despite the proliferation of Nile tilapia in aquaculture production systems in EAC, there are other wild tilapia species and strains that are potential candidate species for fish farming, with or without genetic improvement. The diverse tilapia species available in the LVB may contain genes contributing to disease resistance, accelerated growth, or tolerance to adverse environmental factors such as temperature or salinity. Currently, knowledge of the status of the wild tilapia strains, their level of distribution and stocking of alien strains and their potential impact on biodiversity is still limited. Furthermore, uncontrolled movement and introduction of genetic materials between populations of these groups of species may potentially result in loss of genetic and species diversity in the region.

Evidently, aquaculture operations can potentially benefit from selective breeding programs and campaigns in areas where it does not have potential negative impacts on sustainability. Therefore, there have been efforts to understand and improve genetic material under local climatic conditions, increase the diversity of cultured fish species, promote hatchery operations through production of fry and fingerling, and increased research on tilapia feed.

In the last three decades, the EAC countries have been developing management strategies for sustainable use of fisheries resources through the co-management approaches (Nunan, 2010). However, it should be noted that conservation and management of the existing fish species might not be achieved if the genetic structure and diversity of the species is not well



This project is funded by
the European Union



understood. A thorough characterization of the populations based on robust molecular tools is critical. There is limited available molecular genetic data on tilapia in the LVB, and these studies have been constrained by the scope, the number of tilapia species included, as well as relatively smaller number of markers studied (Kajungiro, et al., 2019) (Tibihika, et al., 2020). Most importantly, there has been limited studies investigating the genetic diversity and status of the species collectively across the whole LVB, temporally and spatially. The small number of samples collected in isolation cannot provide accurate estimates of the extent of gene flow nor the geographical potential stocks in the LVB (Ciezarek, et al., 2022).

2.5 Approaches and Methods for Fish Identification

Tilapias like other fishes, show a complexity of shapes, sizes, and colours. The characterisation and identification of fish species is of enormous interest to taxonomists and systematists, but also a vital requirement for depicting ecology, fishery management, tracking dispersal patterns, estimation of recruitment and spawn grounds, as well as for breeding/captivity purposes. Two approaches are generally employed in fish characterization and identification: the use of morphological traits, and molecular genetic tools for fish species identification.

2.5.1 Application of Morphometric Approaches

The traditional identification of tilapia species has long relied on the use of traits (phenotypic characteristics) specific to each species (skin, eye, and fin coloration; marks like spots or stripes; fin characteristics; body and head shape). However, there are many limitations associated with the use of only these traits.

Firstly, there are many tilapia species found within the LVB and the propensity of these species to hybridise increases the potential risk of misidentification. Furthermore, due to large morphological variation and plasticity within a species also with respect to the environment, the use of phenotypic traits may not be reliable in fish species identification (Zhang & Hanner, 2012). Additionally morphological species identification relies on specific traits found in sexually mature males, preventing the application of such approach to female tilapia or juveniles. Species identification through morphological traits requires extensive expertise and skills that are not easily acquired, but also the technique may be costly, requiring more staff and time consuming.

2.5.2 Application of Molecular Genetic Approaches

A more accurate approach to species characterisation and identification, hybrid classification, and introgression quantification is the use of molecular genetic markers. For instance, the use



This project is funded by
the European Union



of DNA-based fish identification techniques has proved robust and analytically powerful. Previous work performed in the region demonstrated the increased power of this genetic approach that involves genome sequencing for species and hybrid identification over the use of the traditional morphological traits (Ciezarek, et al., 2022).

Sequencing the full genome across populations identifies millions of single nucleotide polymorphisms (SNPs) that can be used to classify species, separate populations, and quantify genetic diversity, and genetic status (well-defined species, hybrids). The number of markers used enables the identification of ancient or recent introgression events. These markers are essential in the assessment of the stocks in nature and in aquaculture conditions.

The sequencing of the full genomes also enables the identification of genetic markers specific to each tilapia species. These markers ultimately provide a cost-effective and efficient tool for accurate species and population identification. To identify those markers, it is necessary to sample multiple populations per species to ensure a fair representation of the existing genetic diversity and therefore ensure the specificity of the defined markers (Ciezarek, et al., 2022).

The major limitations of the genetic approach focusing on whole genome sequencing have usually been attributed to costs, access to high-performance computing facilities, requisite training, and expertise. However, the output and accuracy outweigh the cost of misidentification. With the genesis of improved sequencing platforms e.g., NGS/Illumina, the genotyping costs could be further brought down by the identification of a small number of genetic markers for the accurate identification of species and hybrids.

The current work based on tilapia SNPs genotyping provides the necessary foundational data, tools, and resources to contribute to the sustainable management of the fisheries resources in the LVB. The tools encompass the genomic data, the specific markers, and the capacity-building programs for the scientists in the region.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



3.0 Genetic Screening Research Methodology

A full description of the methodologies applied in this study as part of the Detailed Research Plan are available in Annex 1.

Prior to the sampling exercise, necessary authorization, permissions, and permits were obtained by the third-party research partners from the EAC fisheries research institutions, in compliance with the local, national authorities and Nagoya Protocols. Samples were then transported with import and export permits obtained by the respective research partners.

3.1 Field Sampling

Field sampling was conducted in close collaboration with local stakeholders (farmers, fisherfolk) using non-destructive fishing methods (experimental gill and seine nets) or bought from local fisherman at the land sites depending on the suitability of the site. Sampling was conducted between April and August 2022 across farms (hatcheries, cages, and ponds) and natural environments in the LVB. During the field sampling, special consideration particularly for the threatened/endangered fish species was observed by using non-lethal gear and the sampled specimen were later returned to the water body. Furthermore, non-targeted or excess fish caught were returned to the environment where they were collected.

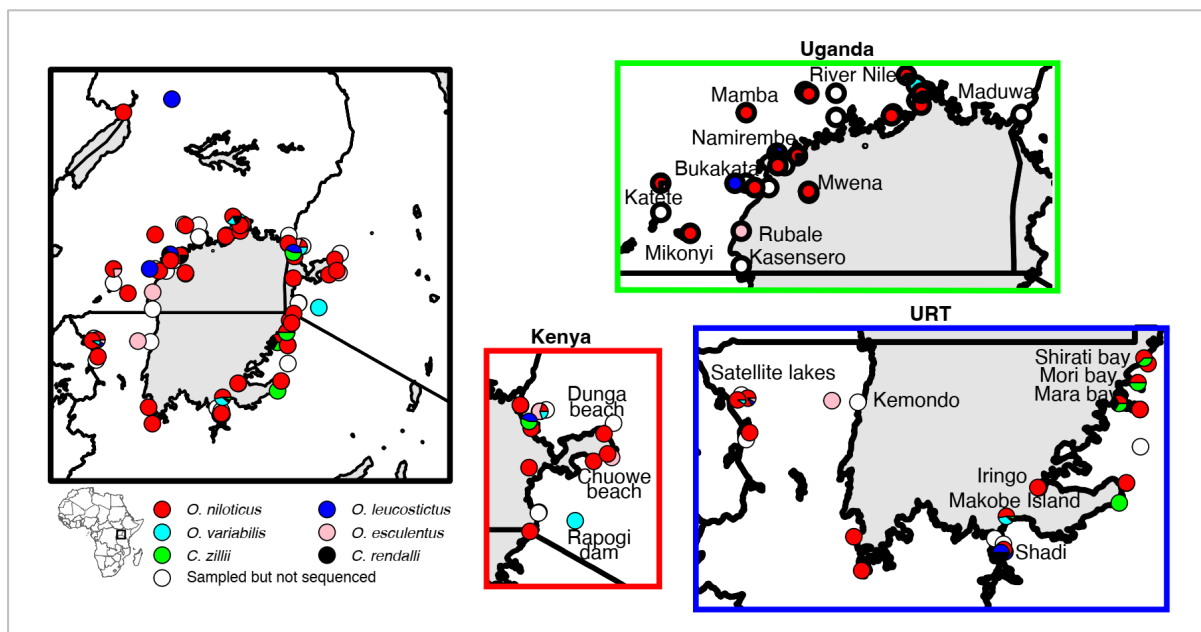


Figure 4 Sampling locations across the Lake Victoria Basin, and specifically in the countries of Kenya, Tanzania, and Uganda



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Principally, the targeted fish species and farmed fish (ponds and cages), were collected, sedated with anaesthesia (clove oil) and identified from morphological traits. Then, a fin clip from the caudal fin was extracted and stored in 99% ethanol in 1.5ml Eppendorf tubes for subsequent genetic analysis. The fish was released back into the environment after extracting the fin clips. Digital images for the individual fish specimens were taken consistently using a camera with an analogous focal length. Exact global positioning system (GPS) coordinates were recorded for all the sampling locations. Species taxonomic treatments were carried out following established and available standards.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



3.2 Sample Preparation

Fin clip samples were shipped to the Earlham Institute (EI) in the UK between March and April 2023 for laboratory and genomic analysis.



Figure 5 Extraction of fin clip sample.

A total of 396 samples (including 12 reference samples from Lake Albert) were selected for genetic analyses focusing on diversity, and population structure, of tilapia species, and assessed hybridisation between species in the LVB. Out of the 396 samples, only 380 were successfully sequenced with sufficient coverage.

DNA was extracted using an automated extraction instrument, the KingFisher Flex (ThermoFisher Scientific). High-throughput genetic sequencing was then carried out (using Illumina NovaSeq X, with LITE library preparation) at EI, with each sample having enough data to cover the genome at least ten times. The current data was supplemented with existing data from Lake Albert (Uganda), representing a reference population of potential non-hybrid or non-admixed individuals (six samples of *O. niloticus* and six samples of *O. leucostictus*).

3.3 Genomic Analyses

3.3.1 Quality control and variant calling

The raw reads obtained from the sequencing were assessed prior to downstream analyses. Here, the low-quality reads and adapter sequences were trimmed using fastp program (Chen, Zhou, Chen, & Gu, 2018), and FASTQ data pre-processing tool. The trimmed reads were then mapped against an existing high-quality reference genome for Zebra fish (*Maylandia zebra*)



This project is funded by
the European Union



(Conte, Gammerdinger, Bartie, Penman, & Kocher, 2017) using the Burrows-Wheeler Aligner (BWA) software (Li & Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, 2009). Samples with a low coverage (less than 1x) were removed from all analyses. Genetic variants were then identified using the bcftools software (Li, et al., 2009). Several subsets of these variants (SNPs) were then extracted, with each filtered for linkage disequilibrium ($r^2 > 0.6$ over 20kb windows), and to exclude sites with substantial missing data (>25%). SNP sets for each species in the *Oreochromis* genus, and *O. niloticus* in each of the three countries, were extracted and bioinformatics analysis performed on each separately. For the single species datasets, only individuals confidently classified (ADMIXTURE ancestry component > 0.99) were included.

3.3.2 Analysis of population genetic structure and hybridization

For each of the datasets, population structure was determined using Principal Component Analysis (PCA) in PLINK (Purcell, et al., 2007), and hierarchical clustering analysis in ADMIXTURE was used to further identify distinct populations within each species, and interspecific hybrids in the overall *Oreochromis* dataset.

To assess the genetic composition of the different populations within each species, an ADMIXTURE analysis was performed using the STRUCTURE program (Hubisz, Falush, Stephens, & Pritchard, 2009). This analysis enabled the identification of genetic clusters for each individual within each species. We also added representatives of each species (*O. leucostictus* and *O. niloticus*) from Lake Albert which were used as references, since the species are native to the waterbody.

To assess the population structure of *O. niloticus* within Lake Victoria and across the sites of LVB, a PCA was carried out on all individuals confidently identified as *O. niloticus* in the ADMIXTURE analysis (ancestry component > 0.99). A subset of 3,300,000 SNPs from these 210 individuals was extracted for these analyses. An ADMIXTURE analysis was carried out from K=1-K=5 (K is the number of subpopulations that make up the total population), and cross-validation scores (CV) calculated to assess which had the best model fit.

3.3.3 Determination of population genetic diversity

To assess the genetic diversity across sampling locations, population genomic statistics were calculated within each sampling site where more than five pure *O. niloticus* were identified. The overall level of genetic diversity per site (π , the number of substitutions divided by callable genome size) was calculated using pixy (v1.2.7. beta1). The inbreeding coefficient for each individual and Tajima's D for each population was calculated using vcfTools (v0.1.16)



This project is funded by
the European Union



(Danecek, et al., 2011). Positive values of Tajima's D can reflect either a recent selective sweep or population contraction after a recent bottleneck, whereas negative values reflect balancing selection or a population expansion. A value of zero indicates no evidence of selection or population size change. The heterozygosity for each individual fish was calculated from their respective VCF file, by dividing the number of heterozygous sites by the callable genome size. Individual heterozygosity and inbreeding coefficients for each individual were averaged for each sampling site.

3.3.4 Development of a 96-SNP chip panel for species and hybrid assignment

To make it easier and cheaper to conduct genetic identification of species or hybrids of *Oreochromis*, a smaller panel of 96 SNPs was identified and selected. These markers were chosen based on their positions on the reference *O. niloticus* genome. To find these markers, they compared genetic data from trimmed DNA sequences to the *O. niloticus* genome instead of a different species used in earlier studies.

Diagnostic SNPs, those which are fixed and specific to either *O. niloticus*, *O. leucostictus*, *O. variabilis* or *O. esculentus*, were identified across the genome using bcftools. These were then filtered to include 24 diagnostic SNPs for each species; one for each species across each of the 21 linkage groups (the repetitive LG3 was excluded). The markers were ensured to be independent, with at least 1 million base pairs apart.

To test the effectiveness of this 96 SNP panel, both PCA and ADMIXTURE analyses were carried out, with results compared to the full-genome data. The comparison was obtained by extracting all SNPs from the 21 linkage groups in the *O. niloticus* genome and removed markers that were too closely linked together (max r^2 0.6 over 20kb windows).



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



4.0 Results and Discussion

4.1 Sample Collection, Sequencing, Mapping and SNP Calling

From the set of species of interest in the LVB, as identified in the Detailed Research Plan, all species, except for *O. spilurus* and *O. jipe* were sampled (Table 1). These species had initially been thought to have been relocated to some fish farms around the LVB. *Oreochromis jipe* and *O. spilurus* are native to lake Jipe and in the coastal rivers of Kenya respectively. However, this study did not show their presence in the main lake and around the LVB.

Table 1: Tilapia samples collected across Lake Victoria Basin

Fish Species (identified based on morphological traits)	Total number of samples collected by country		
	Kenya	Tanzania	Uganda
<i>Oreochromis niloticus</i>	339	577	230
<i>Oreochromis leucostictus</i>	21	349	144
<i>Oreochromis variabilis</i>	62	21	44
<i>Oreochromis esculentus</i>	55	95	69
<i>Coptodon zilli</i>	0	167	5
<i>Coptodon rendalii</i>	30	0	57
Total	507	1209	549

Out of the 396 samples collected (Table 2), 380 samples were successfully sequenced with sufficient coverage of between 3.2-46.7x, with a mapping percentage of 90-97%. The excluded samples were *O. niloticus* (10), *O. esculentus* (2), *O. variabilis* (2) and *C. zilli* (2). The 380 samples successfully sequenced were *O. niloticus* (288), *O. leucostictus* (20), *O. esculentus* (20), *O. variabilis* (20), *C. zilli* (16) and *C. rendalii* (16). Variant and invariant sites were called and filtered separately for each individual fish. These were then merged and filtered to include only SNPs with a minor allele frequency of at least 3 and less than 50% missing data, giving a dataset of 32,125,845 SNPS from 371 individuals (including the 12 Lake Albert reference individuals).



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Table 2: Samples of fish sequenced by country.

Country	No. of fish samples collected	No. of sites sampled	No. of fish from wild sites sequenced	No. of fish from farm/hatcheries sequenced	Total number of fish sequenced
Kenya	507	18	77	16	93
Tanzania	1209	22	131	56	187
Uganda	549	28	105	11	116
Total	2265	68	313	83	396

4.2 Genetic population structure

The filtered all-individuals SNP dataset gave a subset of 4,744,363 SNPs, which were used to determine the PCA. Overall, the analysis showed a clear separation of the different species, with individuals in the *Coptodon* genus clustering together and away from the *Oreochromis* individuals (Figure 6). A clear clustering of the *Oreochromis* individuals by species were also apparent from the PCA.

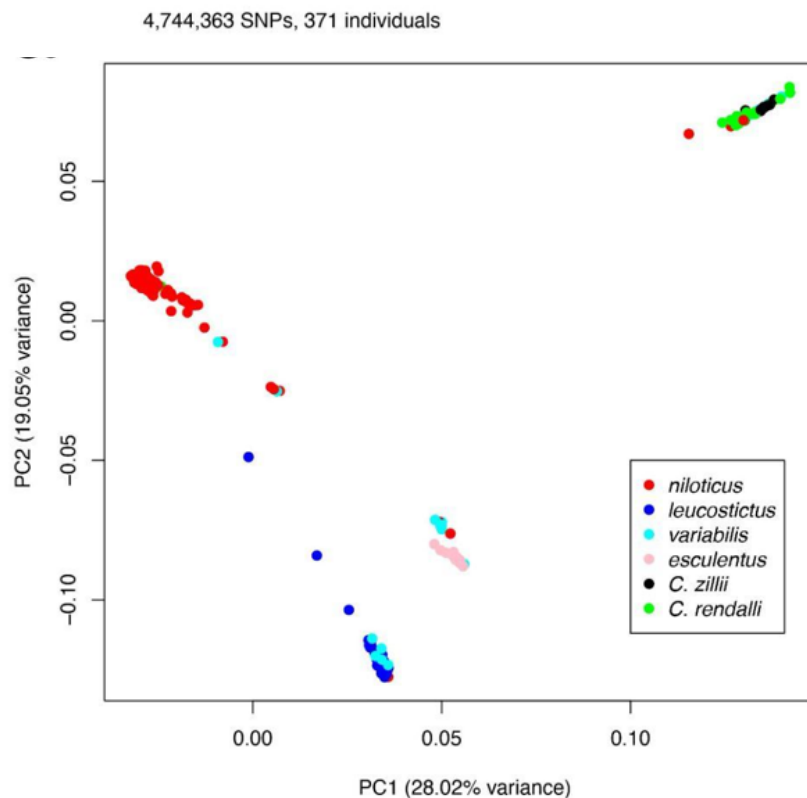


Figure 6 PCA of all individuals sampled from Lake Victoria Basin.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Eight *Oreochromis* individuals (5x *O. variabilis* and 3x *O. niloticus*) that were found to cluster with the *Coptodon* genus were excluded from downstream analyses as they likely resulted from mislabelling. Following the removal of these individuals, a structure of three groups was observed (Figure 7). These individuals are clustered into i) *O. niloticus* populations, ii) *O. esculentus* and *O. variabilis* populations, and iii) *O. leucostictus* and *O. variabilis* populations. However, we noted that some *O. niloticus* individuals were also found in the *O. variabilis* and *O. esculentus* populations, and these are also likely the result of misidentification from morphological traits.

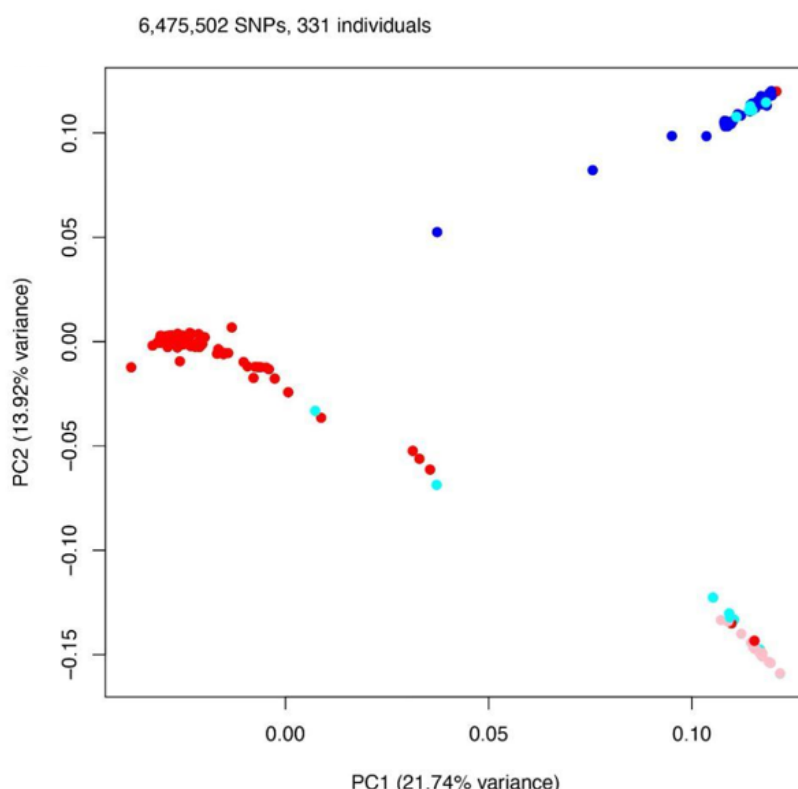


Figure 7 PCA of all individuals excluding *Coptodon* species.

The results from PCA regarding *O. niloticus* analysed for each country separately did not reveal any structure of the populations. However, it does show that the individuals from two sites, Lake Kacheera (Katete) and Lake Kijanebalola (Mitonyi), located west of Lake Victoria in Uganda are genetically distant from the rest of the other individuals originating from the LVB (Figure 8). However, this slight isolation is not significant as samples obtained from LVB generally indicated weak subpopulation structure (F_{st} value = 0.02). Within Lake Victoria, there is a weak pattern of differentiation from North to South. There was no apparent differentiation between individuals sampled from the wild or farms. This could be the result of



This project is funded by
the European Union



the use of local populations as progenitors for the farms, and the relatively recent establishment of the breeding populations, preventing differentiation with the wild stocks.

3.3 million SNPs, 210 individuals (all confidently non-hybrid *O. niloticus*)

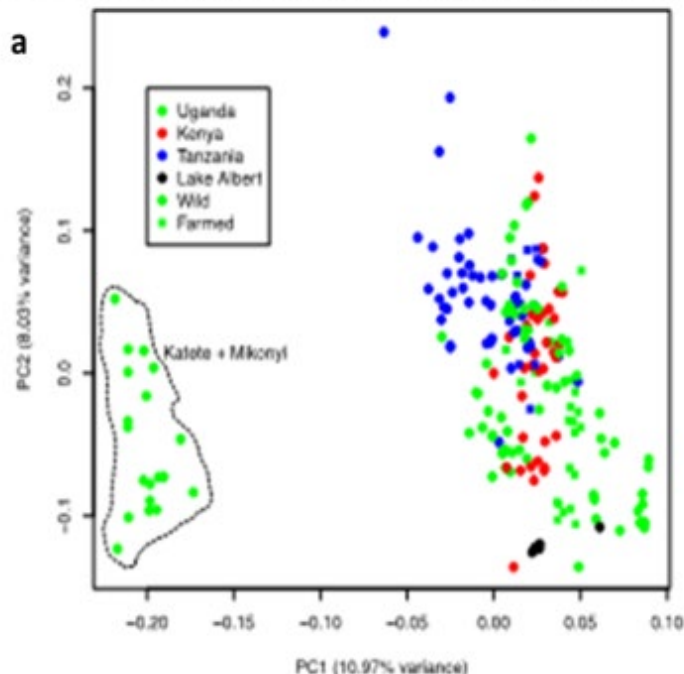


Figure 8 PCA of *O. niloticus* individuals sampled.

The ADMIXTURE analysis suggested that $K=1$ had the lowest CV score (0.227), indicating that *O. niloticus* could be considered as one population within the region (Figure 9). For each value of K , CV gives the 'Cross-Validation Score', with a lower score indicating a better model fit to the data. The lowest CV value (0.227) is for $K=1$, but plots at sub-optimal K values can still be informative about ancestry. $K=2$ suggests that the samples from Lake Albert are of one 'ancestry' group, and those from Lakes Kacheera, and Kijanebalola form another. The samples from Lake Victoria form a mix of those two components, with individuals are more likely to have their ancestors originating from Lake Albert. $K=3$ - $K=5$ show no clear differentiation patterns within Lake Victoria but informs us that introduction of *O. niloticus* into Lake Victoria was not a one-off occurrence, but a series of introductions that occurred over time, with stocks originating from several different ancestors.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE

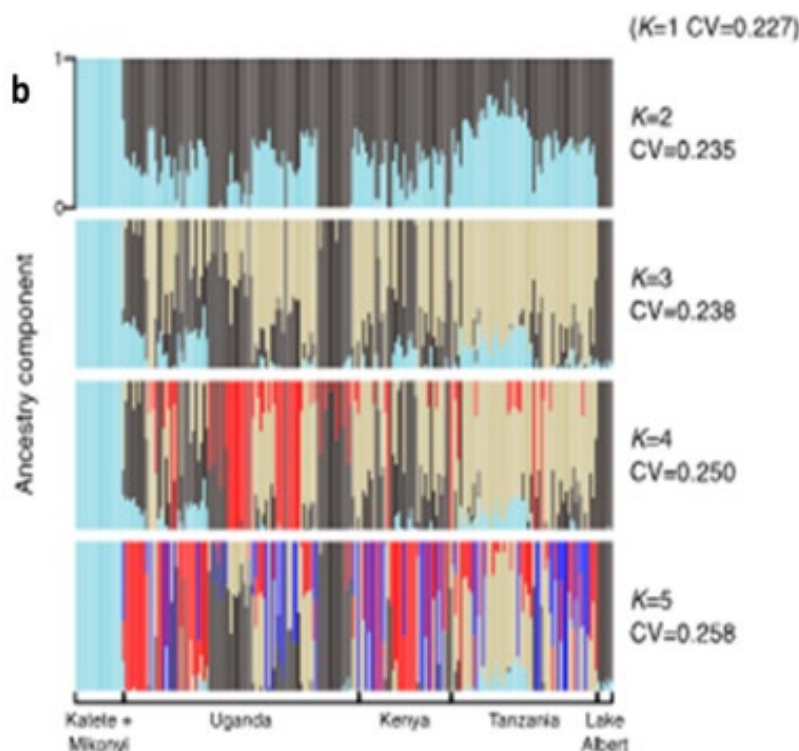


Figure 9 ADMIXTURE analysis of all *O. niloticus* individuals, with coloured bars corresponding to different ancestry groups.

Overall, this could be interpreted as a weak population structure within *O. niloticus*, with a small amount of differentiation segregating the Lakes Kacheera and Kijanebalola from other remaining samples. Differentiation across the lake may be an ongoing process (suggested by the overlapping North-South pattern in the PCA, and slight segregation of Lake Albert samples indicated by the PCA and ADMIXTURE). The results from PCA regarding *O. niloticus* analysed for each country separately did not reveal any structure of the populations (Figure 10). However, *O. niloticus* populations from Uganda indicated the presence of structure (Figure 10b) reflecting the genetic difference between populations of Lakes Kacheera and Kijanebalola, and Lake Victoria



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE

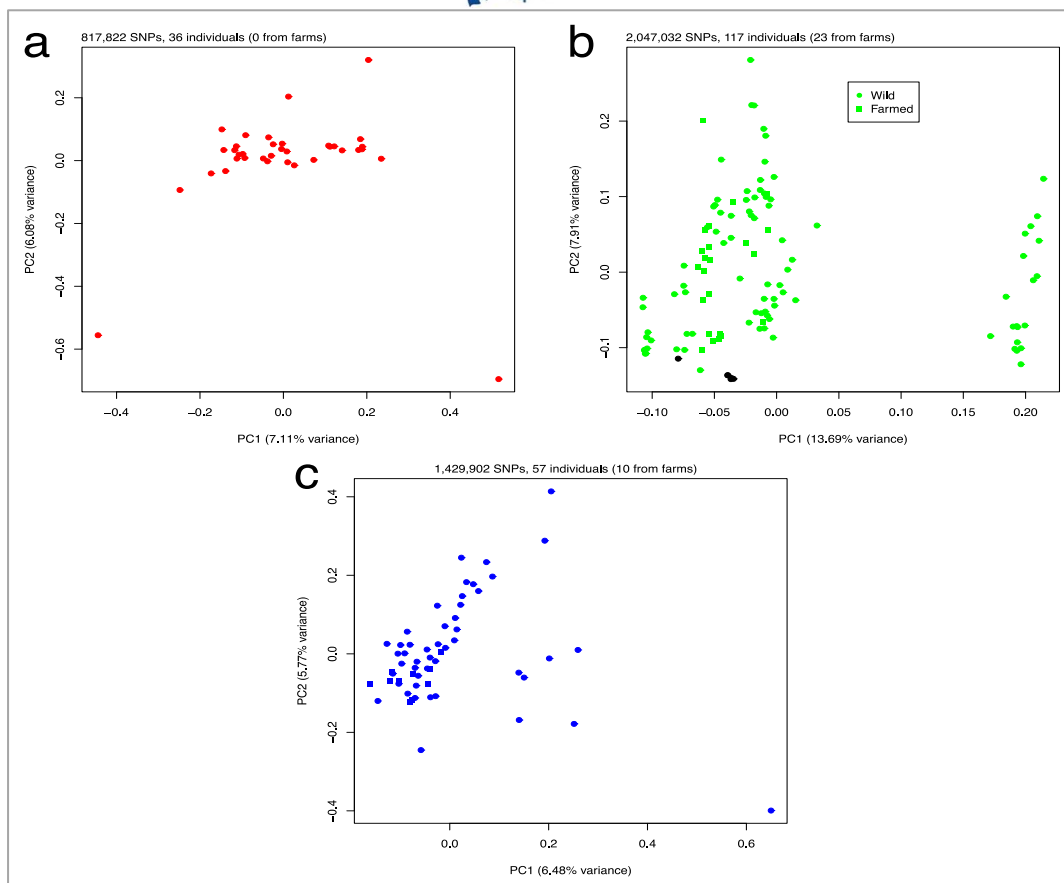


Figure 10 PCA analyses of all non-hybrid *O. niloticus* from a) Kenya, b) Uganda, and c) Tanzania. Black circles in (b) represent samples from Lake Albert.

Similarly, within Lake Victoria, differentiation is not apparent for the non-native *O. leucostictus*, with samples from Lake Albert differentiating along PC1, and samples from the Kazinga farm differentiating along PC2 (Figure 11). By contrast, the native *O. esculentus* showed clear population structure between the three countries (Figure 12). It is possible that there has not been sufficient time for population structure within the non-native species to develop, in contrast to the native species, even if fish from different areas were isolated by distance. Continuous monitoring will be necessary to establish whether differentiation occurs with time.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE

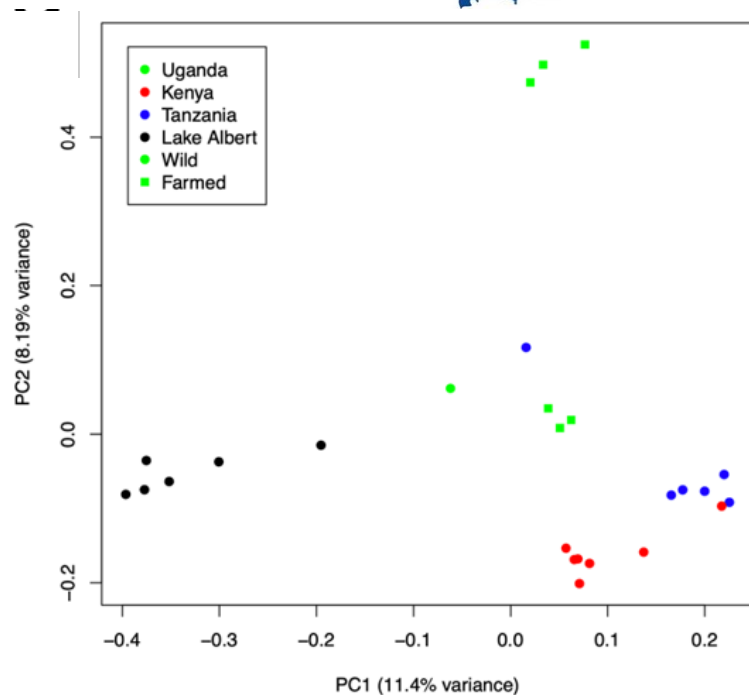


Figure 11 PCA analysis of all non-hybrid *O. leucostictus* samples.

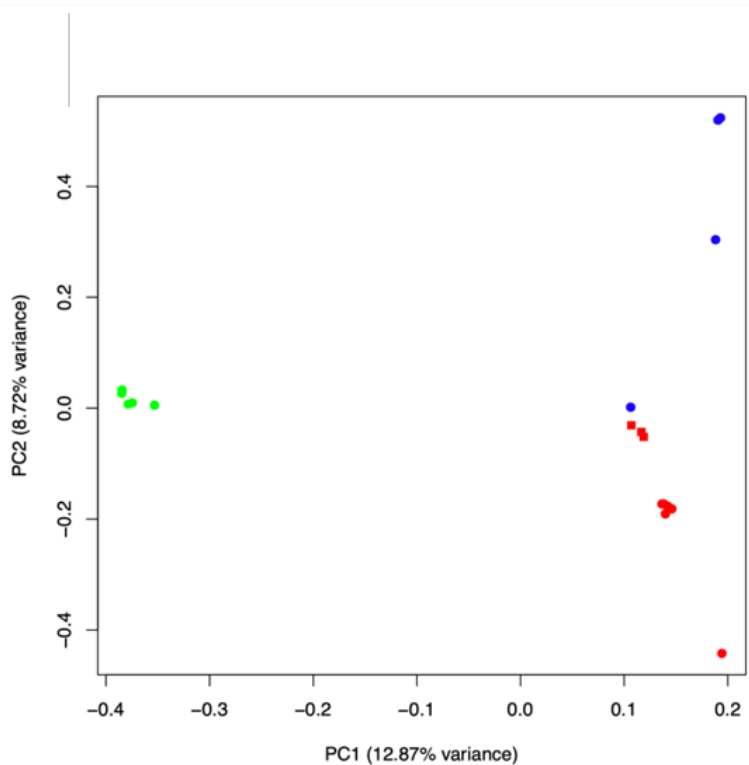


Figure 12 PCA analysis of all non-hybrid *O. esculentus* samples.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



4.3 4.3 Genetic identification of hybrids

Based on ADMIXTURE analysis of all 289 *Oreochromis* individuals, 33 individuals (11.4%) were genetically classified as hybrids, all of which involved *O. niloticus*. Nineteen of these were hybrids with *O. esculentus* (from seven sites), 11 with *O. variabilis* (from six sites) and three with *O. leucostictus* (from one site) (Figure 13). Interestingly, previous genomic works using the same samples, particularly from Lake Victoria, have not identified hybrids between *O. niloticus* and *O. leucostictus* (Ciezarek, et al., 2022) (Ciezarek, et al., 2023). The presence of *O. niloticus* x *O. leucostictus* in one of the sampled farms, and not in the wild, may indicate changes in reproductive behavioural or strategies on individual species resulting from confinement in the small ponds for quite some time; and therefore, leading to hybridisation. Furthermore, this farm (with two 300m² ponds) indicated that fingerlings had been obtained from neighbourhood local farms (not hatcheries) and the fish had taken more than one and a half years in the pond by the time the sampling was done.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE

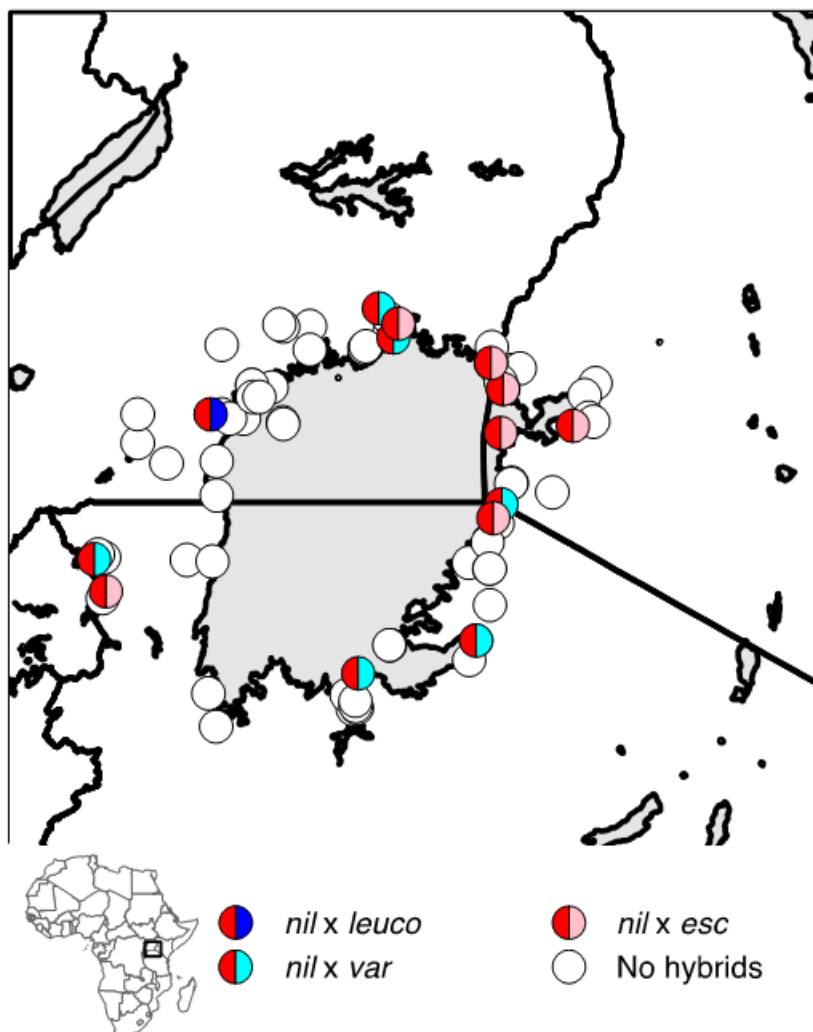


Figure 13 Map of all sampling sites coloured according to species combinations of hybrids sampled from LVB.

This affirms the need for continuously monitoring and developing comprehensive strategies for appropriately identifying and breeding intending but quality species for culturing by the fish farmers. This new data from this study suggests that more hybrids may be possible in farms. More studies with wide scope sampling could be useful to provide a comprehensive picture of the extent of the hybridisation. However, it is also important to note that the study results showed that hybridisation was largely limited to *O. niloticus* breeding with other *Oreochromis* species, and not among other *Oreochromis* species. This may be attributed to the aggressive nature and adaptive reproductive behavioural strategies it exhibits in varying environments compared to other species.

A study by Champneys *et al* 2021 indicated that *O. niloticus* may threaten native tilapia species through dominance in interference competition (Champsey, Genner, & Ioannou, 2021) This can elucidate the invasiveness nature of *O. niloticus* across the region. No site had more than



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



one type of hybrid. These hybrids seem to be particularly clustered in the eastern region of the lake, with *O. variabilis* x *O. niloticus* crosses particularly in the Tanzanian region and *O. esculentus* x *O. niloticus* in the Kenyan region. Both were also found in Lake Mujunju in Tanzania.

Of the wild hybrids, those from Uganda and Kenya were all in Lake Victoria. In Tanzania, hybrids were found in satellite lakes (*O. variabilis* x *O. niloticus* in Lake Kiura and *O. esculentus* x *O. niloticus* in Lake Rwakajunju) as well as in Rubana River (*O. variabilis* x *O. niloticus*). This trend may be associated with the fact that *O. variabilis* and *O. esculentus* are rarely found by the Tanzanian side of Lake Victoria, with remnants only available in the satellite water bodies.

Of the 14 sites where hybrids were recorded, six were fish farms. The *O. leucostictus* x *O. niloticus* hybrids were only found in the Kazinga Mixed Farm in Uganda, with none in the wild. *O. esculentus* x *O. niloticus* hybrids were found in the Masese cage farm in Uganda, Anyanga cages and Lake View and Muga Fish farms in Kenya. A single *O. variabilis* x *O. niloticus* hybrid was found in the Yalelo Fish Farm in Uganda. It is important to note that the relative greater abundance of hybrids detected in farms relative to wild sampling sites is directly influenced by the significant differences in the management of the sites, with much lower number of breeding individuals involved in the formation of the farmed stocks. Thus, many individuals sampled on a farm may have come from one mating event.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



4.4 4.4 Population genomic statistics

Population-genomic statistics of non-hybrid individuals of *O. niloticus* from 24 sites (21 from the wild and 3 farms) with at least 5 such individuals were calculated (Table 3). Heterozygosity per individual was similar across all sites, with all sites in Lake Victoria having an average of 0.0020-0.0028. Lake Albert was slightly lower at 0.0012. Similarly, the level of diversity in each population (π) was very consistent, between 0.0014-0.0017, but again slightly lower in Lake Albert (0.0009).

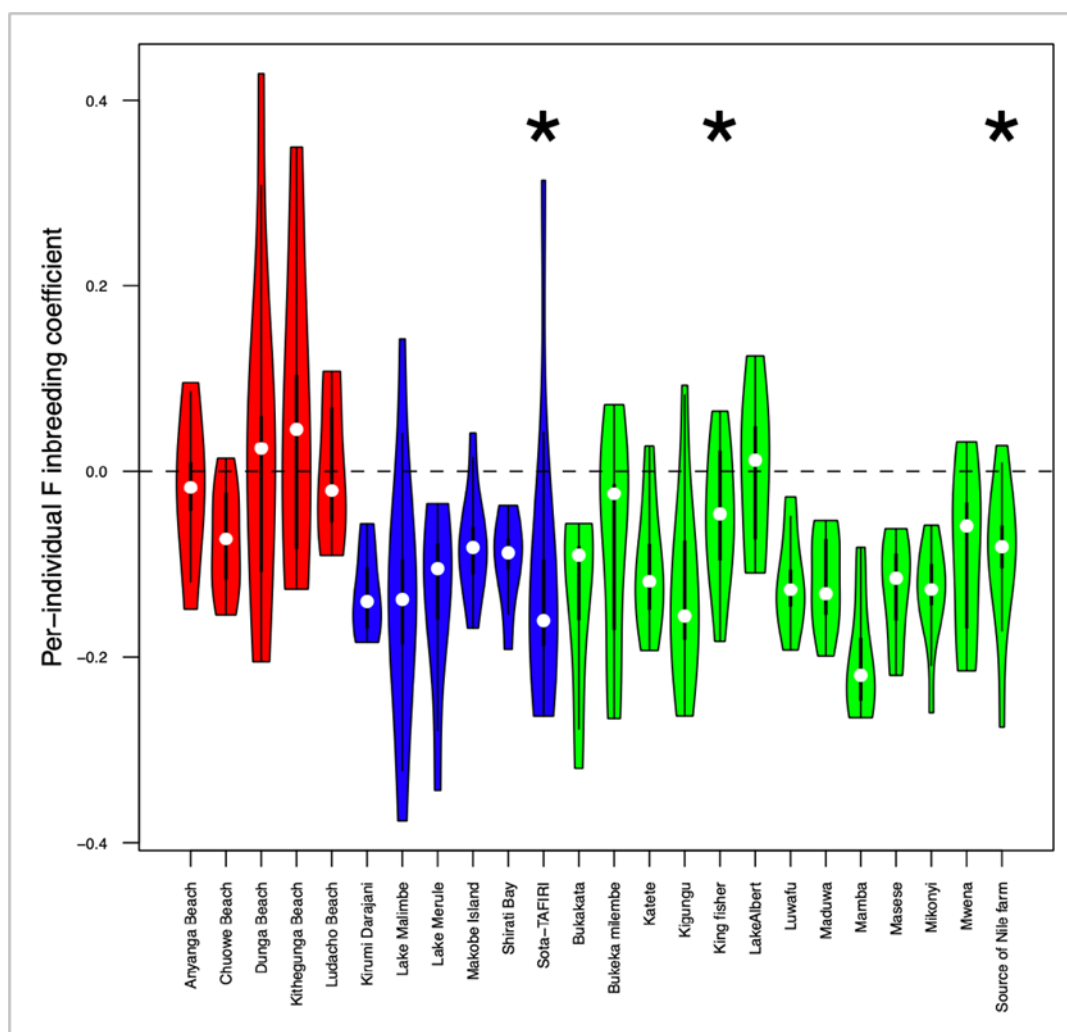


Figure 14 Violin plot of the inbreeding co-efficient, F , per individual in each sampling site with at least five pure *O. niloticus* individuals. Sites are clustered into countries: Kenya (red), Tanzania (blue) and Uganda (green). Sites indicated with an asterisk are farms.

There was no widespread inbreeding in the dataset. Most sites had a slightly negative value (indicating outcrossing), but the average over all sites was near 0 (Figure 14). There were a small number of inbred individuals, including an individual of 0.31 within the TAFIRI farm in Tanzania, an individual of 0.43 in Dunga Beach in Kenya, and another of 0.35 from Kithegunga



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Beach in Kenya. Genome-wide values of Tajimas D varied relatively more between sites, with six sites (one wild and no farmed) having negative values and 18 having positive ones. None of the population genomic statistics showed differences between farmed and wild sites. These analyses did, however, only consider definitively non-hybrid specimens of *O. niloticus*.

Table 3: Population genomic samples for each sampling site with at least five pure O.niloticus samples.

Site	Source	Mean per individual heterozygosity (range)	π	Mean F-statistic (range)	Tajima's D (SD)
Kenya					
Anyanga Beach	Wild	0.0024 (0.0022 – 0.0028)	0.0017	-0.021 (-0.15 – 0.096)	0.24 (0.064)
Chuowe Beach	Wild	0.0023 (0.002 – 0.0025)	0.0015	- 0.075 (-0.15 – 0.014)	- 0.17 (0.074)
Dunga Beach	Wild	0.0021 (0.0014 – 0.0027)	0.0016	0.023 (-0.21 – 0.43)	0.078 (0.017)
Kithegunga Beach	Wild	0.002 (0.0016 – 0.0024)	0.0016	0.058 (-0.13 – 0.35)	0.48 (0.14)
Ludacho Beach	Wild	0.002 (0.0018 – 0.0023)	0.0015	0.0012 (-0.091 – 0.11)	0.24 (0.066)
Tanzania					
Kirumi Darajani	Wild	0.0025 (0.0024 – 0.0027)	0.0016	- 0.13 (- 0.18 – -0.056)	0.049 (0.017)
Lake Malimbe	Wild	0.0026 (0.0018 – 0.0035)	0.0016	-0.14 (-0.34 – 0.035)	-0.0027 (0.025)
Lake Merule	Wild	0.0028 (0.0024 – 0.0035)	0.0018	-0.14 (-0.34 – 0.035)	0.049 (0.017)
Makobe Island	Wild	0.0022 (0.0018 – 0.0024)	0.0015	-0.079 (-0.17 – 0.041)	0.026 (0.019)
Shirati Bay	Wild	0.0024 (0.0022 – 0.0027)	0.0016	-0.097 (-0.19 – 0.037)	0.15 (0.035)
Sota -TAFIRI	Farm	0.0025 (0.0013 – 0.003)	0.0016	-0.11 (-0.26 – 0.31)	-0.091 (0.055)
Uganda					



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Bukakata	Wild	0.0024 (0.0022 – 0.0031)	0.0016	-0.14 (-0.32 – -0.056)	0.15 (0.034)
Bukeka milembe	Wild	0.0024 (0.002 – 0.003)	0.0016	-0.075 (-0.27 – 0.072)	-0.071 (0.042)
Katete	Wild	0.0022 (0.0019 – 0.0025)	0.0015	-0.1 (-0.19 – 0.027)	0.081 (0.018)
Kigungu	Wild	0.0025 (0.002 – 0.003)	0.0016	-0.13 (-0.26 – 0.093)	-0.092 (0.05)
Kingfisher	Wild	0.0021 (0.0019 – 0.0025)	0.0015	-0.048 (-0.18 – 0.065)	0.0072 (0.02)
Lake Albert	Wild	0.0012 (0.001 – 0.0014)	0.0009	-0.0004 (-0.11 – 0.12)	0.62 (0.19)
Luwafu	Wild	0.0022 (0.002 – 0.0024)	0.0015	-0.12 (-0.19 – 0.028)	0.03 (0.023)
Maduwa	Wild	0.0026 (0.0024 – 0.0029)	0.0017	-0.12 (-0.2 – -0.053)	-0.094 (0.051)
Mamba	Wild	0.0024 (0.002 – 0.0026)	0.0014	-0.2 (-0.26 – -0.08)	0.29 (0.077)
Masese	Wild	0.0026 (0.0024 – 0.003)	0.0017	-0.13 (-0.22 – -0.062)	0.026 (0.018)
Mikonyi	Wild	0.0023 (0.0021 – 0.0026)	0.0015	-0.13 (-0.26 – -0.058)	0.022 (0.022)
Mwena	Wild	0.0022 (0.0019 – 0.0025)	0.0015	-0.088 (-0.21 – 0.032)	0.087 (0.02)
Source of the Nile	Farm	0.0022 (0.0019 – 0.0028)	0.0014	-0.09 (-0.28 – 0.027)	0.034 (0.016)

4.5 4.5 Morphological versus genetic identification

In addition to the eight *Oreochromis* that clustered with *Coptodon*, 12 out of the 289 individuals confidently identified as *Oreochromis* morphologically were different to the genetic classification. Two samples morphologically identified as *O. niloticus* were genetically *O. variabilis*, with another three genetically identified as *O. leucostictus*. Five samples morphologically identified as *O. variabilis* were genetically identified as *O. leucostictus*, with one identified as *O. niloticus* and another as *O. esculentus*. Twenty six individuals



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



morphologically identified as *O. niloticus* were genetically identified as such, but with lower confidence (ancestry component 0.9-0.98).

Of the 33 genetically identified hybrids, at least one component was correctly morphologically identified (for example an *O. esculentus* x *O. niloticus* hybrid was identified as *O. niloticus*). The three *O. leucostictus* x *O. niloticus* hybrids were morphologically identified as *O. leucostictus*, the 19 *O. esculentus* x *O. niloticus* hybrids were all morphologically identified as *O. niloticus* and the 11 *O. niloticus* x *O. variabilis* hybrids were morphologically identified as either *O. niloticus* (9) or *O. variabilis* (2).

4.5.1 96 SNP diagnostic panel of hybrid identification

The 96 SNP panel successfully identified all 259 pure species samples consistently with the genomic data. PCA of the 96 SNP panel gave similar results to the full-genome dataset (Figure 15). *O. variabilis* and *O. esculentus* appear more distinct in the 96 SNP panel dataset, owing to the selection of species diagnostic SNPs.

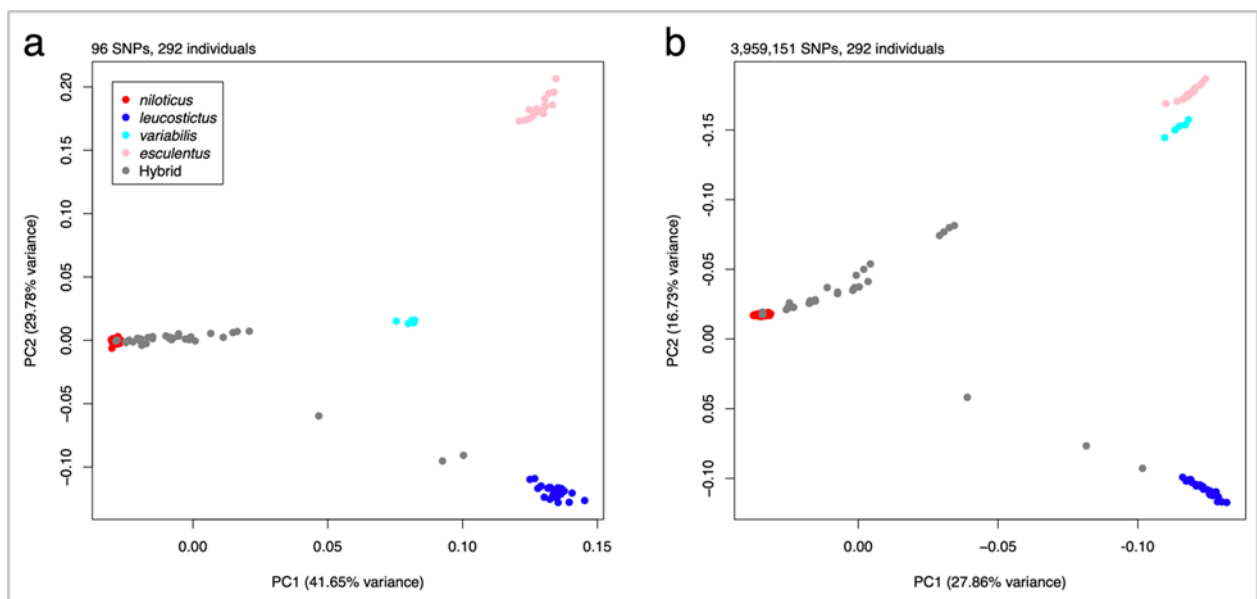


Figure 15 PCA of all individuals (pure and genetic hybrids) for the a) 96 SNP panel, and b) genome-wide *O. niloticus* mapping. Note that PC1 and PC2 are reversed in (b) for easy comparison.

Of the 33 genetically identified hybrids, 17 had ancestry components consistent with being an F1, or first-generation hybrid (minor ancestry component ≥ 0.25). Of these, 16 were successfully identified as hybrids in the 96 SNP panel. Of the remaining 16 hybrids consistent with being a second or third-generation backcross (minor ancestry component 0.1-0.25), only three were identified as hybrid in the 96 SNP panel.



This project is funded by
the European Union



We therefore consider the panel to be accurate in identifying pure species and F1 or first-generation backcross hybrids, but more distant backcrosses may be falsely inferred as pure species.

5.0 Interpretation of the Genetic Screening Results

This part of the work from the TRUEFISH initiative is the first of its kind that aimed to assess the genetic diversity of the tilapia populations across the whole of the LVB using consistent sampling methodologies, molecular approaches, and analyses across selected capture fisheries and aquaculture sites. The data generated provides a unique snapshot of the current state of the tilapia populations across the LVB region and provides the needed reference to assess the evolution of the stocks in the future.

Our analyses revealed high genetic diversity within the lake and a clear separation of the different species analysed. Most importantly, this study revealed very weak population structure for *O. niloticus* across the LVB. Two sites appear to differentiate from the other sampling locations: Lake Kacheera and Lake Kijanebalola in Western Uganda, although the differentiation is slight. The observed genetic difference likely originates from the different origins of the populations used for stocking the lakes or from stochastic differences in the genotypes sampled from the population(s) used to stock the lakes. The populations of Lake Victoria are also genetically similar in part to those from Lake Albert which historical records show as being a source of *O. niloticus* introductions. However, there were likely several introductions from various sources as evidenced from STRUCTURE analyses. Overall, a measure of between population differences (F_{st}) was only 0.02 for Nile tilapia. These data show that nearly all the *O. niloticus* populations we sampled were relatively similar, but most importantly appearing as one large, admixed population in Lake Victoria.

The lack of strong population structure in our analyses compared to previous studies, stems from our focus on the LVB only. Kajungiro et al. 2019, applied double digest restriction site-associated DNA (ddRAD) on seven populations sampled across Tanzania, including three sub-populations from Lake Victoria (Kajungiro, et al., 2019). The authors reported genetic differences between the populations across Tanzania. However, even in their studies, two out of the three populations sampled from the Lake were indistinguishable (FETA and “Lake Victoria”). Likewise, Tibihika et al. 2020, despite reporting significant population structure across African countries for *O. niloticus*, found no difference between Lake Victoria populations except for individuals from Sango Bay (Tibihika, et al., 2020). Our findings are also consistent with recent genetic surveys of natural and farmed populations in Uganda by



This project is funded by
the European Union



Robeldo et al. (2024) which showed high genetic diversity within populations, but with relatively little differences between wild Lake populations ($F_{st} < 0.05$) although contributions from Lakes Albert, Kyoga and Victoria to farm populations could be distinguished (Robeldo, et al., 2024).

Farmed populations are likely to differ in gene frequencies from the native populations from which they are derived simply because the founder fish they represent a small sample from the wild population. As domestication selection occurs, they are likely to diverge further over time. Some degree of inbreeding and associated diversity loss with time is to be expected although both can be reduced by observing the best practices in stock genetic management. However, in the present study, the estimators of diversity indicate an absence of inbreeding within farmed populations. These results, alongside the lack of differentiation with Lake Victoria population, suggest that those farmed populations were likely initiated from Lake Victoria individuals, and that these populations were established recently and or with regular input material from lake populations. Robeldo et al. (2024) noted similarly high genetic diversity and a general absence of inbreeding within Ugandan farm populations that they sampled (Robeldo, et al., 2024), but in their case, the data indicated that the farms had been stocked mainly from Lakes Albert or Kyoga. Previous work in Tanzania identified significant risk for indigenous species from the introduction of species such as the Nile tilapia due to hybridization (Blackwell, et al., 2020) (Ciezarek, et al., 2022) (Ciezarek, et al., 2023). In the present study of the LVB, a total of 33 individuals out of 369 (8.94%) were identified as hybrids. However, the proportion of hybrids within fish farms is much greater than that observed in the wild (18.07% vs 5.75%) and included crosses not seen in the wild (*O. niloticus* x *O. leucostictus*) both in this study and in previous work. Thus, the present study has confirmed that hybridization occurs, but it is not widespread and frequent, but it is sufficiently common that management procedures to reduce the risk of hybridization on farms must be considered.

There was no indication that fish farmers were deliberately attempting to produce hybrids and a single hybrid mating on the farm could produce several hundred fish to be grown in production ponds. Thus, a greater number of hybrid fish being sampled on a farm may not reflect the number of independent hybrids that may have occurred. It is likely that the hybrid mating on farms occur accidentally from the proximity of different fish species and the likely supply of fish at early juvenile stages when different species are very difficult to distinguish. Key to reducing the risk of hybridization, through ensuring strain quality control in hatcheries are tools to accurately identify species and their hybrids so that mixed mating incidences are reduced or eliminated. Further work will be needed to elucidate the occurrence of hybrids in



This project is funded by
the European Union



the wild populations. There were no obvious correlations of wild sites with hybrids in the present sampling.

To facilitate species and hybrids identification in a cost-efficient manner, we designed a panel comprised of 96 SNPs. The panel demonstrates good power to recapitulate whole genome sequencing results, discriminating species, and enabling the identification of first-generation hybrids and first-generation backcrosses.

The genetic screening program has therefore met the two major goals of the study by 1) successfully developing information on the spatial distribution of different species and genotypes of tilapias in the LVB and thus, 2) enabling the identification of risks to biodiversity potentially due to aquaculture. The genetic information has not revealed any populations with unique genetic characteristics that might require specific conservation measures.

However, it appears that the Nile tilapia genotypes from some satellite lakes in EAC, for instance in Lakes Kacheera and Kijanebalola (Uganda), are genetically diverging from both the putative ancestral habitats (e.g., Lake Albert) and the other non-native in Lake Victoria. Therefore, some conservation and sustainable management options may be considered for *O. niloticus* in the satellite water bodies. The current study has identified some hybridization between *O. niloticus* and the other *Oreochromis* species, but the pattern is not extremely common and widespread, with some 8% of individuals sampled being of hybrid origin and more commonly in the eastern part of Lake Victoria. This work has demonstrated the difficulty of identifying species using morphology alone, particularly for females and juveniles. The hybridization observed on farms appears to relate to misidentification of broodstock (perhaps first introduced to the farm as young fish) rather than from deliberate attempts to produce hybrid stock. This finding highlights the need for capacity development to identify the specific species more accurately for culture and for better quality control in hatcheries and permit effective regulation of germplasm transfer in the region to manage issues potentially arising from hybridisation.

It was envisaged that the genetic studies could provide tools to assist aquatic genetic resource management and the work has confirmed that morphological identification of species is inaccurate, particularly when applied to females and juveniles. The genomic screening, involving millions of markers, has demonstrated the capability of genetics to identify species and their hybrids. The present study has also demonstrated that a subset of genetic markers (96 SNPs) can achieve sound species and hybrid identification with high accuracy, with reduced cost of genotyping.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



The present work was limited in resources and by the COVID outbreak. The data gathered have been sufficient, however, to provide a benchmark for the current status of tilapia genetic resources in LVB and to provide a basis for addressing the key questions of managing the risk to biodiversity while developing a viable aquaculture industry in the region. This work can be extended in future studies analysing the samples collected in the present study that could not be analysed and extending studies spatially and temporally utilizing the tools developed here.

6.0 Data Availability

All raw sequencing data are made publicly available on the EBI ENA repository (PRJEB75650). Genotypes for all sequenced individuals, as well as genotypes and positions for the 96 SNPs panel can be accessed from <https://doi.org/10.5281/zenodo.11193182>.

7.0 Acknowledgements

This work was supported with funding from the European Union (FED/2018/039-248). Field collection of samples was made possible through the participation of EAC research institutes, namely, Kenya Marine and Fisheries Research Institute (KMFRRI), Tanzania Fisheries Research Institute (TAFIRI) and the Ugandan National Fisheries Resources Research Institute (NaFIRRI) and their representatives, Dr. Kevin Obiero, Dr. Asilatu Hamisi Shechonge and Dr. Cassius Aruho respectively. Genotyping and sequencing of the samples was made possible by the Earlham Institute, UK, implemented by Dr. Wilfried Haerty and Dr. Adam Ciezarak. The project was also funded in part by the CGIAR Research Initiative on Resilient Aquatic Food Systems for Healthy People and Planet, led by WorldFish and supported by contributors to the CGIAR Trust Fund.

The first draft of this report was initiated by JAHB, BW and PDT of WorldFish, with inputs from WH and AC of Earlham Institute. Subsequent reviews were implemented by the WorldFish (JAHB, RKB, BW and PDT), EAC research partners (CA, AS, KO), and Earlham Institute (WH, AZ).



This project is funded by
the European Union



8.0 References

- Blackwell, T., Ford, A. G., Ciezarek, A. G., Bradbeer, S. J., Gracida, C. A., Smith, A. M., . . . Genner, M. J. (2020). Newly discovered cichlid fish biodiversity threatened by hybridization with non-native species. *Molecular Ecology*, 895-911.
- Champsey, T., Genner, M., & Ioannou, C. (2021). Invasive Nile tilapia dominates a threatened indigenous tilapia in competition over shelter. *Hydrobiologia*, 3747-3762.
- Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, 34(17), 884-890.
- Ciezarek, A. G., Mehta, T. K., Man, A., Ford, A. G., Kavembe, G. D., Kasozi, N., . . . Haerty, W. (2023). Ancient and ongoing hybridization in the *Oreochromis* cichlid fishes. *BioRxiv*.
- Ciezarek, A., Ford, A., Etherington, G., Kasozi, N., Malinsky, M., Mehta, T., . . . George, T. (2022). Whole genome resequencing data enables a targeted SNP panel for conservation and aquaculture of *Oreochromis* cichlid fishes. *Aquaculture*, 548.
- Conte, M. A., Gammerdinger, W. J., Bartie, K. L., Penman, D. J., & Kocher, T. D. (2017). A high quality assembly of the Nile Tilapia (*Oreochromis niloticus*) genome reveals the structure of two sex determination regions. *BMC Genomics*, 1-19.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., . . . Durbin, R. (2011). The variant call format and VCF tools. *Bioinformatics*, 2156-2158.
- FAO. (2022). *The State of World Fisheries and Aquaculture (SOFIA) (2022 ed.)*. Rome: FAO. Retrieved from <https://openknowledge.fao.org/handle/20.500.14283/cc0461en>
- Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 1322-1332.
- Kajungiro, R. A., Palaikostas, C., Lopes Pinto, F. A., Mmochi, A. J., Mtolera, M., Houston, R. D., & de Koning, D. J. (2019). Population Structure and Genetic Diversity of Nile Tilapia (*Oreochromis niloticus*) Strains Cultured in Tanzania. *Frontiers Genetic*, 10. doi:10.3389/fgene.2019.01269
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 1754-1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 2078-2079.
- Ministry of Agriculture, Animal Industry and Fisheries. (November, 2019). *National Fisheries and Aquaculture Policy of Uganda*. Uganda.



This project is funded by
the European Union



- Munguti, J. M., Obiero, K. O., Iteba, J. O., Kiriimi, J. G., Kyule, D. N., Orina, P. S., . . . Tanga, C. M. (2023). Role of multilateral development organizations, public and private investments in aquaculture subsector in Kenya. *Frontiers in Sustainable Food Systems*, 7(1208918).
- Nunan, F. (2010). Governance and Fisheries Co-management on Lake Victoria: Challenges to the Adaptive Governance Approach. *Maritime Studies*, 103 - 125.
- Obiero, K., Meulenbroek, P., Drexler, S.-S., Dagne, A., Akoll, P., Odong, R., . . . Waidbacher, H. (2019). The Contribution of Fish to Food and Nutrition Security in Eastern Africa: Emerging Trends and Future Outlooks. *Sustainability* 11(6):1636.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., . . . Sham, P. C. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American Journal of Human Genetics*, 559-575.
- Robeldo, D., Ogwang, J., Byakora, E., Nascimento-Schulze, J. C., Benda, K. K., Frasin, C., . . . Mukiibi, R. (2024). Genetic diversity and population structure of farmed and wild Nile tilapia (*Oreochromis niloticus*) in Uganda: The potential for aquaculture selection and breeding programs. *Genomics*.
- Tibihika, P. D., Curto, M., Alemayehu, E., Waidbacher, H., Masembe, C., Akoll, P., & Meimberg, H. (2020). Molecular genetic diversity and differentiation of Nile tilapia (*Oreochromis niloticus*, L. 1758) in East African natural and stocked populations. *BMC Evolution Biology*.
- Zhang, J., & Hanner, R. (2012). Molecular Approach to the Identification of Fish in the South China Sea. *PLoS ONE*.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



9.0 Annex 1: Detailed Research Plan

The Detailed Research Plan for Molecular Genetic Screening of Tilapiines in the East African Community (EAC) Lake Victoria Basin; Work Package 3.3.3 Genetic Screening Research.

Implementing Institutions: Earlham Institute (EI), Kenya Marine and Fisheries Research Institute (KMFRI), National Fisheries Resources Research Institute (NaFIRRI), and Tanzania Fisheries Research Institute (TAFIRI).



This project is funded by
the European Union



Table of contents

List of Abbreviations	3
Background and context	4
Materials and Methods	9
Laboratory analysis of samples	11
Research equipment required per institute	11
EI	11
Estimated budgets per institution	11
Readiness of research institutions (EAC third party research partners)	15
KMFR I	15
NaFIRRI	18
TAFIRI	19
Research schedule	22
Year 2021	22
Year 2022	23
Year 2023	23
References	24

List of Abbreviations

EI	Earlham Institute
GDP	Gross Domestic Product
KMFR I	Kenya Marine and Fisheries Research Institute
LVB	Lake Victoria Basin
LVFO	Lake Victoria Fisheries Organization
NaFIRRI	National Fisheries Resources Research Institute
TAFIRI	Tanzania Fisheries Research Institute



This project is funded by
the European Union



Background and context

The East African Great Lakes Region in Uganda, Tanzania and Kenya contribute an estimated 4% of the regional gross domestic product (GDP) annually. The sector produces over 1 million tonnes of fish, worth over US \$600 million, from capture fisheries (LVFO, 2014) and an estimated 160,000 tonnes derived from aquaculture. The sector further contributes to the economic development of the region through the employment of approximately 5 million people.

The annual per capita fish consumption in the EAC region is remarkably low (average of 5.3 kg/person/year) as compared to the rest of Africa (10.1 kg/person/year) and global level of 20 kg/person/year (Obiero et al., 2019). Commercially important fish species in the Lake Victoria Basin (LVB) includes two non-native species, Nile perch (*Lates niloticus*) and Nile tilapia (*Oreochromis niloticus*), and the native cyprinid *Rastrineobola argentea* (Kenya, *Omena*; Uganda, *Mukene*; and Tanzania, *Dagaa*). Despite the significant contribution of fisheries sector to the economic development in the region, the sector is facing significant challenges in transitioning towards sustainable fish production and practices.

Anthropogenic threats to fisheries, mainly through overfishing in the LVB, was evidenced as early as the 1920s (Graham 1929; Welcomme 1966; Worthington 1932). Overexploitation of the native tilapiine cichlids (*O. variabilis* and *O. esculentus*) in the LVB escalated in the 1950s when controversial regulatory policies recommended the use of smaller-sized gillnets (<12.7 cm) (Balirwa 1992; Ogutu-Ohwayo 1990). The decline of the native tilapiine species might have further triggered the introduction of non-native fish species (*O. niloticus*, *L. niloticus*, among others) into the LVB in the 1950s (Balirwa et al. 2003; Ogutu-Ohwayo 1990). Consequently, the reduction of the native species, mainly *O. variabilis* and *O. esculentus*, lead to increased colonization by Nile tilapia and Nile perch resulting in almost the disappearance of the indigenous ichthyofauna (Ogutu-Ohwayo 1990).

The introduction of the non-native tilapiines in the LVB subsequently resulted in introgressive hybridization or segregation in the ecosystem hence contributing to the undesirable ecological changes (Tibihika et al. 2020). In addition, the current expansion of fish farming (aquaculture) activities may further exacerbate introgressive hybridization through the introduction of foreign species/strains into the lake systems. The outcomes of these interactions might be detrimental to the future of the native tilapiines and may lead to the loss of stable populations that are vital for maintaining ecological stability and economic benefits to the people in the region.



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Despite the numerous research efforts by the riparian countries in the LVB, approaches to sustainably manage capture fisheries have failed and the stock of *L. niloticus* and cichlid tilapias have consistently declined. Similarly, the aquaculture sub-sector that would ease pressure on the dwindled capture fisheries has also not fully developed its potential and accounts for only 7-8% of the regional fish consumption. As the EAC's population growth and urbanization continue to increase, the overall demand for fish consumption in the region is projected to grow substantially in the near future. Therefore, interventions focused on developing sustainable aquaculture production systems are crucial and paramount. The development and growth of a competitive aquaculture sector in the EAC are curtailed by various constraints including limited technical skills/training, limited access to quality seed and feed, fish diseases, and biodiversity threats associated with fish translocations of unknown genetic structure and diversity in the course of restocking and aquaculture. Besides, the increase of unregulated aquaculture establishments uncontrolled spatial distribution, and overuse of natural resources profoundly continue to threaten the production of fish sustainably in the EAC countries.

In the last three decades, the EAC countries have been developing management strategies for the sustainable use of fisheries resources through co-management approaches (Nunan et al. 2015). However, it should be noted that conservation and management of the already admixed fish species might not be achieved if the genetic structure of the species in question is not well understood (Allendorf and Luikart 2009). Genetic studies on variation among and within species in LVB were performed mainly on haplochromine cichlids, which indicate little differentiation based on the analysis of mitochondrial (mt) DNA (Meyer et al., 1990; Nagl et al., 2000, Mzighan et al., 2010), nuclear-encoded genes (De Magalhaes et al., 2009; Seehausen et al., 2003), and short interspersed elements (SINEs) (Terai et al., 2004). Tilapiine species in the LVB have undergone various turbulences due to anthropogenic activities, thus, a thorough characterization of the populations based on robust molecular tools is needed. There are available molecular genetic data on tilapiines in the LVB e.g., Tibihika et al. (2020), Kariuki et al. (2021), these studies might have been constrained by the scope, the number of congeneric tilapiine species included, inter alia. Additionally, the fact that these studies were in confined areas and were based on a small number of simple tandem repeats/ short sequence repeats/ microsatellite genetic markers, protein-coding genes, or mitochondrial genes. The interpretation of microsatellite results might have been hampered by the size homoplasy constraints (Narum et al. 2013; Tibihika et al. 2019; Vartia et al. 2016), which may fundamentally render the data less informative. The small number of samples in



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



the studies cannot provide accurate estimates of the extent of gene flow or the geographical potential stocks (Ciezarek et al 2021).

In the efforts aimed at contributing remedies to the fisheries constraints, LVFO Secretariat in partnership with Landell Mills, FAO, and WorldFish, is implementing a project, the “EU-EAC True-Fish farming story in LVB (True-Fish)”. The project received a grant under the 11th European Development Fund (EDF11), to benefit the EAC particularly Kenya, Tanzania, and Uganda. The overall objective of the True-Fish project is “to contribute to the development of a competitive, gender-equitable and sustainable commercial aquaculture sector to support the economic development and sustainable management of natural resources in the LVB”. The project focuses on three Specific Objectives, namely: 1) Enhance capability in access at national and regional levels to commercial networks for aquaculture-related businesses; 2) Develop national capacity to use robust genetic screening consequently increasing the availability of skilled personnel thereby addressing two of the most important limiting factors for aquaculture development, and 3) Improve sustainability by mitigating risks related to the aquatic environment, aquatic animal health conditions, and biodiversity. This project is aligned with the 2009 LVFO Regional Aquaculture Strategy and Investment Plan as well as the EAC Fisheries and Aquaculture Policy (2018).

Under the True-Fish project, WorldFish implements part of Specific Objective 3, with the overall focus on Result Area 3.3: Improved protection of biodiversity. The key aim of Result Area 3.3 is to address the key risks to the long-term development and sustainability of the aquaculture sector, by minimizing the risk of accidental or intentional introduction of genetic material which can threaten native species. In addition, an LVFO Regional Working Group with responsibility for Fish Breeding and Genetics will be established to strengthen national and regional capabilities. A genetic screening and mapping program will be implemented and support will be provided to inform the development of national and regional policies, regulations, and guidelines for aquatic genetic resource management, broodstock management, and fingerling production in the LVB regions of Kenya, Tanzania, and Uganda. Training of fisheries and aquaculture officers in species identification, breeding and administration of the regulatory framework will also be provided.

WorldFish will execute Result Area 3.3 through collaborative research spearheaded by Earlham Institute (EI) UK, Kenya Marine Fisheries Resources Institute (KMFRI) Kenya, National Fisheries Resources Research Institute (NaFIRRI) Uganda and Tanzania Fisheries Resources Research Institute (TAFIRI) Tanzania. In this study, we shall employ full genome



This project is funded by
the European Union



sequencing as well as novel single nucleotide polymorphisms (SNPs) markers to holistically investigate the genetic integrity of tilapia species ranging from captive to natural populations. The molecular tools are based on the protocols for full genome sequencing and population analyses developed at EI and the Nile tilapia SNP chips developed by WorldFish. These approaches are robust, cost efficient and allow investigation of the diversity and genetic structure of tilapia populations including inbreeding, introgressive hybridization, and gene flow. The data collected will allow to inform policy and other regulatory frameworks for sustainable management of the economically important tilapiine biodiversity in LVB.

The current functional and structural frameworks of research in EAC member states do not fully elucidate the regional needs of biodiversity protection and biosecurity related to aquaculture development. This suggests that the delivery of science-based guidelines for biodiversity sustainability to the EAC member states is insufficient. Therefore, Result Area 3.3 of True-Fish will facilitate the establishment of new or improved techniques that are pivotal to better meet the biodiversity needs in EAC, with the following projected outputs:

- i) LVFO Regional Working Group established (i.e. with agreed ToR, composition and working arrangements) to deliver advice on genetics and biodiversity in aquaculture. The aim is to promote relevant, timely and effective sharing of information, and joint research programs on a sustainable basis for both national research and academic institutions, and including the private sector aimed at the needs of existing operators in the aquaculture value chain.
- ii) Aquaculture and wild locations mapped showing distribution of native and exotic genotypes. A genetic survey will be undertaken and using appropriate genotype identification tools and recommendations for a future screening program developed. This will lead to significantly increased information on the potential biodiversity impacts of genetic improvement of farmed species. It will produce a publicly available database including inventory or directory of native aquatic genetic resources (AqGR) and its distribution. The information generated will be used to assess the main risks to biodiversity arising from aquaculture development, the measures which could be adopted to manage the impacts, and to propose future monitoring programme on impact of introduced species on biodiversity and environment. The information will be disseminated to all relevant and interested parties.
- iii) Science-based policy recommendations and model national policies, regulations and guidelines for biodiversity presented for adoption at regional level. This will entail developing Regional Policy and strategic framework and model Regulations and Guidelines for promoting harmonised regulatory approach to aquatic genetic resource management, reflecting the provisions of the Nagoya Protocol (and AU Strategic Guidelines for Coordinated Implementation of the Nagoya Protocol) on Access and Benefit Sharing (ABS) and in line with the EAC Common Market commitments on free movement of goods. New regionally harmonised national regulations will be drafted and presented for adoption, defining the national institutions responsible, and providing the necessary tools and legal powers to limit the impacts of aquaculture.



This project is funded by
the European Union



- iv) Number of fisheries, aquaculture and environmental officers and private hatchery owners having received awareness raising and training on methods and technologies for hatchery and biodiversity management. This will strengthen the capacity in national research institutions to design and conduct collaborative research, and to translate findings into policy recommendations. Officers responsible for environmental risk management in relation to aquatic biodiversity and for enforcement of associated regulatory measures will be trained. An information campaign will ensure that private sector stakeholders in the aquaculture and fishery sector are adequately informed.

Through Component 3.3 of the True-Fish project, WorldFish aims to establish a system and build capacity for evaluating and monitoring the biodiversity of key species used for aquaculture, and undertake an initial assessment, thus providing the foundations and direction for future sustainable aquaculture development. Thus, Work Package 3.3.3: Genetics Screening Research was developed to map the distribution of native and exotic strains of tilapias and their hybrids in the nature and in aquaculture facilities. This Work Package will be implemented with research collaborators, EI, KMFRI, NaFIRRI, and TAFIRI, and the following outputs will be anticipated:

- i) Report showing the spatial distribution of different species and genotypes of tilapias in the LVB;
- ii) Identification of fish populations and sites exposed to a high risk of loss of biodiversity potentially due to aquaculture establishment;
- iii) Recommendations for science-based policy measures for the optimal management and conservation of biodiversity in the region;
- iv) At least one publication in a refereed journal, one LVFO publication and four presentations at regional and or international conferences.

Materials and Methods

This study will target Nile tilapia and its congeneric species in the LVB of the EAC member states; Uganda, Kenya, and Tanzania. Tilapiine specimens will be harvested by experimental gill nets from the water bodies including main Lake (Lake Victoria) and small/satellite lakes and rivers following the acquisition of research permits from the pertinent EAC authorities. Pond seine and scoop nets will be used for sampling ponds and cages respectively.

Fish sampling will be carried out using monofilament gillnets of stretched mesh sizes 1.5 to 3 inches. An effort will be made to sample all identified habitat types. Sampling points will be marked using a GARMIN Global Positioning System (GPS). After hauling, fish will be sorted



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



and identified to species using the taxonomic guide by Eccles (1992). Between 1 and 30 individuals of each species, depending on the catch will be photographed alive (left side of the individual fish) immediately after capture, using a digital camera for further identification if necessary. After fish are anaesthetized with clove oil, the right pectoral or caudal fin of each photographed fish will be clipped and preserved in pure ethanol labelled vials (2 - 5 ml) for genetic analysis. The preserved fin-clips will be shipped to EI for genetic analysis. Care will be observed to avoid cross-contamination of the samples.

From each sampling site, depending on abundance, at least 30 individuals per species (*O. niloticus*, *O. variabilis*, *O. leucostictus*, *O. esculentus*, *C. zillii* etc) will be sampled. For larger water bodies e.g., Lake Victoria, several sampling sites will be explored to have a broader picture of the genetic diversity as well as the likely presence of genetic substructures in the populations. We will also sample *O. niloticus* in Lake Albert, Edwards, George as those populations have previously been used as reference for this species (Ciezarek et al. 2020). The vicinity of fish farms within the LVB will be explored to ascertain the potential prevalence of feral populations, inter alia. The sampling sites will be selected following the recommendations of the EAC partner states. In Uganda, lakes such as Victoria, Wamala, Mburo, Kacheera, Kijanebalola, Nabugabo, Albert, Edward, George, Kyoga as well as rivers such as the, Katonga, Kagera and Victoria Nile will be considered (Table 1; Figure 1). Likewise, in Kenya, lakes Victoria, Baringo, Turkana, Bogoria, Kanyaboli, Naivasha, Nakuru, Elementaita, Jipe, Chala, Ol Bolossat, Kanyaboli and rivers Sio, Yala Sondu, Nyando, Nzoia, Mara, and Migori will be of interest (Table 2). In Tanzania sampling sites will be selected among lakes, Victoria, Eyasi, Manyara, Kitangiri, Burigi, and Ikimba, rivers Mara, Grumeti, and Kagera (Table 3). From each sampling site, GPS coordinates will be taken for GIS analyses.

Table 1: Sampling stations in Uganda

Water body	Stations
Lake Victoria	6
Lake Nabugabo	1
Lake kyoga	3
Lake Kacheera	1
Lake Mbulo	1
Lake Kayugi	1
Lake Wamala	1
Lake Kanyanja	1
Lake Kijjanibarola	1
River Nile	1
River Katonga	1



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



River Kagera	1
Aquaculture facilities	10
<i>Outside the Basin</i>	
Lake Edward	2
Lake George	1
Lake Albert	3

Table 2: Sampling stations in Kenya

Water body	Stations
Lake Victoria	6
Lake Nakuru	1
Lake Elementaita	1
Lake Baringo	1
Lake Naivasha	1
Lake Turkana	1
Lake Jipe	1
Lake Chala	1
Lake Kanyaboli	1
Lake Ol Bolossat	1
River Nzoia	1
River Yala	1
River Nyando	1
Aquaculture facilities	10

Table 3: Sampling in Tanzania

Water body	Stations
Lake Victoria	6
Nyumba ya Mungu	1
Lake Burigi	2
Lake Ikimba	1
Lake Merule	1
Lake Eyasi	1
Lake Kitangiri	1
Lake Jipe	1
Lake Manyara	1
River Kagera	1
River Mara	1
River Grumet	1



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Aquaculture facilities

10

Figure 1. Potential sampling sites in Lake Victoria, depicted by red dots. Other prospective sampling locations from the satellite/minor lakes and aquaculture facilities in the Lake Victoria basin are not indicated.

Laboratory analysis of samples

Assessment of the spatial distribution of genetic material will be determined at the molecular genetic level through DNA extraction, library preparation, and sequencing the genomes of native and stocked forms. DNA will be extracted from up to a total of 360 specimens and Illumina sequencing libraries prepared at the Earlham Institute. Low coverage (5x) Illumina genome sequencing will be performed at the Earlham Institute to enable unbiased assessment of genetic diversity of the different populations and their sustainability

Research equipment required per institute

EI

All research equipment is in place and does not need to be purchased. The protocols for DNA extraction (PureLink Genomic DNA extraction, Thermo Fisher Scientific) from fin clips have been tested. Libraries preparation and sequencing will be performed by the Genomics Pipelines at EI. We aim to sequence the full genome of each sample at a depth of 10x coverage on average.

Data QC and analysis pipelines for variant calling and population analyses are already in place and tested (Ciezarek et al. 2021). Sequencing data will be made publicly available through deposition to ENA. Pictures and other non-sequencing data will be made publicly available on Zenodo, generating DOI for the data sets to facilitate their reuse. We will make all the scripts available to the community through CyVerse UK.

Estimated budgets per institution

Table 4: Proposed inputs per institution

	Unit	No. Required
Earlham Institute:		
Lead scientist fees	days	24
International travel	flight	2
Per diem	night	14
Regional Travel	round trip	2

45



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Researcher	days	330
Laboratory fees: DNA sequencing	sample	360
Laboratory fees: SNP tests	sample	120
Visiting scientist bench fees	days	180
Dissemination activities	lump sum	1
For each of TAFIRI, NaFIRRI and KMFRI:		
Scientist fees: Research planning	days	5
Scientist fees: Sampling	days	60
Scientist fees: Reporting and dissemination activities	days	10
Purchase of equipment	lump sum	1
Transport hire costs	day	60
Fuel Costs	km	10,000
Daily Subsistence Allowance	per diem	60
International travel - Work/study at Earlham	flight	1
Per diems - Work/study at Earlham	per diem	30
Dissemination activities	lump sum	1

Table 5: Detailed costings for EI

Requirements	Unit	No.	Cost €
Salary PI - Wilfried Haerty - 24 days (€837 per day - salary/overheads)			20,088
Salary - Researcher/PDRA - 330 days (€662 per day - salary/overheads)		1	218460
Conference fees	days		
International travel	flight	2	2000
Per diem	night	4	2100
Regional Travel	round trip	2	800
Visiting scientist bench fees	days		
Laboratory fees: DNA extraction and sequencing (consumables)	sample		46800
Laboratory fees: SNP tests (consumables)	sample		4200
Visiting EI lead scientist to EA (meeting with research institutions and short training course)	days		
Dissemination activities	lump sum		900
Total estimated costs			295,348



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Table 6: Detailed costings for KMFI

Items	Unit	No.	Cost (€)
Scientist fees: Research planning	Days	5	1000
Scientist fees: Sampling	Days	60	12000
Sampling fees: Reporting and dissemination activities	Days	10	2000
Purchase of equipment	Lump sum	1	2660
Transport hire costs	Days	60	3000
Fuel costs	KM	10,000	2000
Daily subsistence allowance	Per diem	60	9000
International travel – work/study at EI	Flight	1	1500
Per diem – work/study at EI	Per diem	30	12000
Dissemination activities	Lump sum	1	5585
Research/sampling permits	Lump sum	1	134
Import/export permits	Lump sum	1	300
Administration cost	Lump sum	1	2606
Total	Lump sum	1	53785

Table 7: Detailed costings for NaFIRRI

Items	Units	No. Required	Cost (EUR)
Scientist fees: Research planning	Days	5	900
Scientist fees: Sampling	Days	60	10800
Scientist fees: Reporting and dissemination activities	Days	10	1800
Purchase of equipment	Lump sum	1	2951
Transport hire costs	Days	60	3740
Fuel costs	KM	10,000	1500
Daily subsistence allowance	Per diem	60	5100
International travel – work/study at EI	Flight	1	1628
Per diem – work/study at EI	Per diem	30	9546
Dissemination activities	Lump sum	1	8975
Research/sampling permits	Lump sum	1	233
Import/export permits	Lump sum	1	n/a
Administration cost	Lump sum	1	6612
Total	Lump sum	1	53785



This project is funded by
the European Union



Table 8: Detailed costings for TAFIRI

Requirements	Unit	No	Cost €	Total Cost €
Fee for research planning	Days	5	44	220
Fees for sampling	Days	60	44	2,640
Fee for reporting and dissemination activities	Days	14	44	616
Purchasing of equipment	Lump sum	1	1500	1,500
Transport hire costs	Day	60	120	7,200
Fuel	Km	10,000	1.2	12,000
Per diem (Researcher)	Days	60	44	2,640
Per diem (Researcher)	Days	60	44	2,640
Per diem (Technician)	Days	60	37	2,220
Per diem (Driver)	Days	60	29	1,740
International travel (at Earlham)	flight	1	2000	2,000
Per diem (at Earlham)	per diem	30	80	2,400
Dissemination activities	Lump sum			1,500
Sub-Total				39,316
GPS	Piece	1	250	250
Digital camera	Piece	1	350	350
Wading boots	Pair	2	100	200
2 mm vials	Pkt	1	94	94
Scalpel blades	Pkt	1	25	25
Gloves	Cartons	1	15	15



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



Gillnets	Lump sum			600
Dipnets	Lump sum			450
Seine nets	Lump sum			450
Cast net	Lump sum			450
Electrofisher	Pc	1	3500	3,500
Analytical ethanol	litres	3	45	135
Ethanol GPR	litres	50	54	2,700
Sub-total				9,219
Ethical research clearance	Proposal	1	75	75
Research permit	Pers on	1	285	285
Sub-total				360
TOTAL				48,895
10% Administration cost				4,890
TOTAL				53,785

Readiness of research institutions (EAC third party research partners)

KMFRI

a) Lead Scientists from Partner Institutions

- Dr Kevin Obiero, Senior research Scientist, Kenya Marine and Fisheries Research Institute (Sangoro) and Centre Director in charge of c coordinating research, administrative and financial functions to ensure research activities are aligned to the Institute's strategic plan, vision and mandate.
- Jacob Abwao, Research Scientist at Kenya Marine and Fisheries Research Institute (KMFRI), Sagana Centre. Currently a doctorate student at the University of Nairobi undertaking studies in Animal (fish) genetics and breeding.



This project is funded by
the European Union



- Elijah Kembenya, Research scientist KMFRI, Sangoro Aquaculture Station and a Doctoral student in Kisii University, Kenya

b) Permit requirements for import/ export of live Fish

- The Fisheries Management and Development Act, 2016, provided for prohibition or controlling the importation, exportation and introduction into Kenya of live fish of any kind or species. This is significant because, while the imported improved strains increase production output thus contributing to increased income from aquaculture enterprises, the underlying risks on the natural genetic populations are eminent.
- Any importation will Require fish movement permit

c) Procedure for obtaining live fish importation permit.

- Applying for certification from Kenya Fisheries Services, (KeFS) and Directorate of Veterinary Services (DVS) through KENTRADE agency.
 - KeFS charges: Kshs 5000/= while the
 - DVS charges 25% of the cost of goods

- Apply for Import declaration form (IDF) from KRA;

Requirements:

- ✓ Health certificate from the country of origin
- ✓ Letter of conformity from the country of origin

d) Research Permit from NACOSTI: All persons intending to undertake research in the country must obtain research permit.

- Genetic Resources and Derivatives - Applications to conduct research on genetic resources and derivatives must submit Prior Informed Consent (PIC), Mutually Agreed Terms (MAT); Material Transfer Agreement (MTA) where applicable.
- The charges for the license is 20,000/= for a public institution like KMFRI
- There is need for submission of an approved proposal

e) Identity of the geographical region for sampling

- We shall do mapping of sampling sites in Lakes Victoria - Kenyan side, Baringo, Turkana, Bogoria, Kanyaboli, Naivasha, Nakuru, Elementaita, Jipe, Chala. and Rivers Sio, Yala Sondu, Nyando, Nzoia, and Kuja/Migori and Small Water Bodies/Dams in Kenya.



This project is funded by
the European Union



f) Identity of the species that will be sampled

- *O. niloticus* sub species including, *baringoensis* (Lake Baringo), *Vulcani* (Lake Turkana), and *niloticus* from Lake Victoria.
- *O. esculents* from the dams
- *O. variabilis*
- *O. leucostictus*
- *O. spirulus*
- *O. jipe*
- *Coptodon zillii*

g) Identity of the local experts engaged in the project.

- i) Dr. James Barasa, Lecturer, University of Eldoret, Kenya
- ii) Dr Judith Makwali, Lecturere, University of Eldoret, kenya
- iii) Dr. Simon Omasaki, Lecturer, Kisii University, Kenya
- iv) Dr Paul Orina, KMFRI Kegati Centre Director, Kisii, Kenya
- v) Dr Domitila Kyule, KMFRI Sagana Centre Director, Sagana, Kenya
- vi) Dr Erick Ogello, Lecturer, Maseno University, Kenya
- vii) Dr Zcahry O. Kinaro, Lecturer, Jaramogi Oginga Odinga University of Science and Technology (JOUST), Bondo, Kenya
- viii) Prof Julius Manyala, Jaramogi Oginga Odinga University of Science and Technology (JOUST), Bondo, Kenya
- ix) Dr Dorcas Lusweti, Lecturer, Moi University
- x) Dr Kevin Mbogo, Lecturer, Jomo Kenyatta University of Science and Technology
- xi) Prof Romulus Abila, Ecologist, Maasai Mara University
- xii) Dr Dorothy Wanja Nyingi, National Museums of Kenya, Ichthyology Section
- xiii) Dr Titus Chemandwa Ndiwa, University of Nairobi, Kenya
- xiv) Dr Paul O Angienda, Maseno University, Zoology Department, Maseno, Kenya

NaFIRRI

a) Identify lead scientists from partner institutions

- Dr. Aruho Cassius

b) Identify the permit requirements (sampling, import and export)

- Permission to collect samples



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



- and to transport/ship samples to UK
- At the time of transporting the samples, EI will send a letter to NaFIRRI NARO allowing importation of samples from Uganda

c) Identify the geographical region for sampling

- LVB:
- Lake Lakes Victoria, Nabugabo, Kakyera, Mbulo, Wamala, Kijanebarola, Nakivali, Wamala; and rivers Ruizi, Katonga ad Kagera

d) Identify the species that will be sampled

- Species of Tilapia;
- We shall initially collect and work on these species

- i. *O. niloticus*
- ii. *O. esculentus*
- iii. *O. variabilis*
- iv. *Coptodon zillii*

e) Identify the local experts that we should be engaging with for the project.

- Assoc. Prof. Charles Masembe (Makerere University Kampala)

TAFIRI

a) Lead Scientist from Partner Institutions

- Dr. Asilatu Shechonge: Is a Research Scientist at Tanzania Fisheries Research Institute (TAFIRI) based at the Dar es Salaam Centre. Dr. Shechonge is currently pursuing post-doctoral studies at the University of Bangol (mostly working in Tanzania) in fish genetics.

b) Permit requirements for import/ export of fish samples: (sampling, import and export)

- **Export permit:** The export permit for fish samples (ethanol preserved fin clips) will be requested from Fisheries Division under the Ministry of Livestock and Fisheries. This is to adhere with Fisheries Regulations (2009) and its amendments.
- **Nagoya permit:** Since the samples will be exported for genetic analysis, permit under Nagoya protocol will be required. TAFIRI will apply for Nagoya permit indicating a type of collaboration with WorldFish and give an assurance that the samples will be solely for research purposes. In Tanzania, Nagoya issues including the permit is under the Vice President Office (V.P.O). Submission of an approved proposal may be needed. However, it should be noted that Tanzania is in process of putting in place legal and institutional framework on Access and Benefit Sharing



This project is funded by
the European Union



(ABS). In the meantime, research permits are vested with other relevant research institutions in coordinating research-based activities that falls under the provisions of Nagoya protocol.

- **Research clearance: Research Clearance/approval will be done by TAFIRI**
- **Research permit: If foreign researchers will be visiting Tanzania for field work, TAFIRI will apply research permit for them. This is based on the Tanzania Fisheries Research Institute (Fisheries and Aquaculture Research) Regulations, 2020 and relevant guidelines and requirements by the Tanzania Commission for Science and Technology (COSTECH).**
- **Special sampling permit (from Ministry of natural resources and Tourism):**
Some of the sampling sites will fall under the National conservation or protection areas jurisdiction, which need special clearance to access the sites for field sampling. In that case, TAFIRI will apply for research permit from Tanzania Wildlife Research Institute (TAWIRI) to conduct research activity in those restricted areas

c) Identify the geographical region for sampling

- We shall do mapping of sampling sites in Lakes Victoria Basin/catchment - Tanzanian side, Lake Nyasa, lakes Manyara, Kitangiri, Rukwa, Jipe, Chala, Mtera Dam, Nyumba ya Mungu Dam, rivers Mara, Pangani, Rufiji, Pangani and Kagera and small water bodies/satellite lakes in Tanzania

Table 9: Species that will be sampled

S/No.	Species	Sampling site
1	Nile tilapia (<i>Oreochromis niloticus</i>)	Lake Victoria catchment, Lake Tanganyika catchment, Nyumba ya Mungu
2	Jipe tilapia (<i>Oreochromis jipe</i>)	Pangani catchment
3	Tanganyika tilapia (<i>Oreochromis tanganyicae</i>)	Lake Tanganyika catchment
4	Wami tilapia (<i>Oreochromis urolepis</i>)	Wami, Ruvu, Rufiji catchments
5	Rukwa tilapia (<i>Oreochromis rukwaensis</i>)	Lake Rukwa, Mtera Dam, Upper Ruaha
6	Shire tilapia (<i>Oreochromis shiranus</i>)	Lake Nyasa catchment
7	Victoria tilapia (<i>Oreochromis variabilis</i>)	Lake Victoria catchment
8	<i>Oreochromis esculentus</i>	Lake Victoria catchment, Lake Kitanguri, Lake Rukwa
9	<i>Oreochromis pangani</i>	River Pangani
10	<i>Oreochromis shiranus</i>	Lake Nyasa
11	<i>Oreochromis leucostictus</i>	Lake Victoria Basin



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



--	--	--

d) Identity the local experts engaged in the project.

- i) Dr. Ngatunga – Tuna Alliance (NGO)
- ii) Mr. James Lusana – University of Dar es Salaam
- iii) Mr. Hilary Mrosso, TAFIRI, Mwanza, Tanzania
- iv) Dr. Gabagambi, TAFIRI, Ag. Centre of Director, Kyela, Mbeya
- v) Mr. Huruma Mgana, TAFIRI, Ag. Centre Director, Kigoma
- vi) Dr. Kimirei, TAFIRI, Director General, Dar es Salaam
- vii) Dr. Mary Kishe, TAFIRI, Ag. Director of research

Research schedule

Key activity details	Quarter1			Quarter2			Quarter3			Quarter4		
	J a n	F e b	Mar ch	A pri l	M a y	Ju ne	J ul y	A u g	S ep t	O c t	N o v	D e c
Virtual meeting with the research partners for updating the research progress as well as introducing the WorldFish Biodiversity Advisor (BA)												
Meeting between EI Lead Researcher and regional research institutes												
Assemble research equipment and plan for the sampling												
Research permit clearance and other regulations												
Genetic material sample collection												



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



from wild and farmed tilapiines in the LVB, taking digital images etc												
Preserving, assembling/packaging and air transportation of samples to the UK												

Year 2021

Key activity details	Quarter1			Quarter2			Quarter3			Quarter4		
	J a n	F e b	M a r c h	A p r i l	M a y	J u n e	J u l y	A u g	S e p t	O c t	N o v	D e c
Virtual meeting with the research partners for updating the research progress												
Genetic material sample collection from wild and farmed tilapiines in the LVB, taking digital images etc												
Preserving, assembling/packaging and air transportation of samples to the UK												
Laboratory analysis for DNA extraction, purification and quantification, PCR primer testing and subsequent genomic DNA												



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



amplification, library preparations, Illumina sequencing, bioinformatics and statistics, data interpretation, and reports												
Bioinformatic training – one member from each partner (virtual if travel impossible)												

Year 2022

Year 2023

Key activity details	Quarter1			Quarter2			Quarter3			Quarter4		
	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
Bioinformatics and statistics, data interpretation, and reports												
Report at manuscript preparation												

References

- Allendorf, F. W. & G. Luikart, 2009. Conservation and the genetics of populations. John Wiley & Sons.
- Balirwa, J. S., 1992 The evolution of the fishery of *Oreochromis niloticus* (Pisces: Cichlidae) in Lake Victoria. *Hydrobiologia* 232(1):85-89 doi:<https://doi.org/10.1007/BF00014616>.
- Balirwa, J. S., C. A. Chapman, L. J. Chapman, I. G. Cowx, K. Geheb, L. Kaufman, R. H. Lowe-McConnell, O. Seehausen, J. H. Wanink & R. L. Welcomme, 2003. Biodiversity and fishery sustainability in the Lake Victoria basin: an unexpected marriage? *BioScience* 53(8):703-715 doi:[https://doi.org/10.1641/0006-3568\(2003\)053\[0703:BAFSIT\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2003)053[0703:BAFSIT]2.0.CO;2).



This project is funded by
the European Union



- De Magalhães, J.P., Curado, J. and Church, G.M., 2009. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics*, 25(7), pp.875-881.
- Ciezarek, A., Gardner, L., Savolainen, V. and Block, B., 2020. Skeletal muscle and cardiac transcriptomics of a regionally endothermic fish, the Pacific bluefin tuna, *Thunnus orientalis*. *BMC genomics*, 21(1), pp.1-16.
- Ciezarek, A., Ford, A., Etherington, G., Nasser, K., Malinsky, M., Mehta, T., Penso-Dolfin, L., Ngatunga, B., Shechonge, A., Tamatamah, R. and Haerty, W., 2021. Whole genome resequencing data enables a targeted SNP panel for conservation and aquaculture of *Oreochromis cichlid* fishes. *bioRxiv*.
- Kariuki, J., P. D. Tibihika, M. Curto, E. Alemayehu, G. Winkler & H. Meimberg, 2021. Application of microsatellite genotyping by amplicon sequencing for delimitation of African tilapiine species relevant for aquaculture. *Aquaculture* 537:736501 doi:<https://doi.org/10.1016/j.aquaculture.2021.736501>.
- Nagl, S., Tichy, H., Mayer, W. E., Takezaki, N., Takahata, N. & Klein, J. (2000). The origin and age of haplochromine fishes in Lake Victoria. *Proceedings of the Royal Society B* 267, 1049–1061
- Mzighani, S. I., Nikaido, M., Takeda, M., Seehausen, O., Budeba, Y. L., Ngatunga, B. P., Katunzi, E. F. B., Aibara, M., Mizoiri, S., Sato, T., Tachida, H. & Okada, N. (2010). Genetic variation and demographic history of the *Haplochromis laparogramma* group of Lake Victoria – an analysis based on SINEs and mitochondrial DNA. *Gene* 450, 39–47.
- Narum, S. R., C. A. Buerkle, J. W. Davey, M. R. Miller & P. A. Hohenlohe, 2013. Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology* 22(11):2841-2847 doi:<https://doi.org/10.1111/mec.12350>.
- Nunan, F., M. Hara & P. Onyango, 2015. Institutions and co-management in east African inland and Malawi fisheries: A critical perspective. *World Development* 70:203-214 doi:<https://doi.org/10.1016/j.worlddev.2015.01.009>.
- Ogutu-Ohwayo, R., 1990. The reduction in fish species diversity in Lakes Victoria and Kyoga (East Africa) following human exploitation and the introduction of non-native fishes. *Journal of Fish Biology* 37:207-208 doi:<https://doi.org/10.1111/j.1095-8649.1990.tb05040.x>.
- Seehausen, O., Koetsier, E., Schneider, M.V., Chapman, L.J., Chapman, C.A., Knight, M.E., Turner, G.F., Alphen, J.J.V. and Bills, R., 2003. Nuclear markers reveal unexpected genetic variation and a Congolese-Nilotic origin of the Lake Victoria cichlid species flock. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1511), pp.129-137.
- Tibihika, P. D., M. Curto, E. Alemayehu, H. Waidbacher, C. Masembe, P. Akoll & H. Meimberg, 2020. Molecular genetic diversity and differentiation of Nile tilapia (*Oreochromis niloticus*, L. 1758) in East African natural and stocked populations. *BMC Evolutionary Biology* 20(1):16 doi:<https://doi.org/10.1186/s12862-020-1583-0>.
- Tibihika, P. D., M. Curto, E. Dornstaedter-Schrammel, S. Winter, E. Alemayehu, H. Waidbacher & H. Meimberg, 2019. Application of microsatellite genotyping by sequencing (SSR-GBS) to measure genetic diversity of the East African *Oreochromis niloticus*. *Conservation Genetics*:1-16 doi:<https://doi.org/10.1007/s10592-018-1136-x>.
- Vartia, S., J. L. Villanueva-Cañas, J. Finarelli, E. D. Farrell, P. C. Collins, G. M. Hughes, J. E. Carlsson, D. T. Gauthier, P. McGinnity & T. F. Cross, 2016. A novel method of



This project is funded by
the European Union



TRUEFISH
ADVANCING AQUACULTURE



microsatellite genotyping-by-sequencing using individual combinatorial barcoding.
Royal Society Open Science 3(1):150565 doi:<https://doi.org/10.1098/rsos.150565>.
Welcomme, R., 1966. Recent changes in the stocks of Tilapia in Lake Victoria. Nature
212(5057):52 doi:<https://doi.org/10.1038/212052a0>.
Worthington, E. B., 1932. A report on the fisheries of Uganda investigated by the Cambridge
Expedition to the East African Lakes, 1930-31.