# Population genomics of introduced Nile tilapia Oreochromis niloticus (Linnaeus, 1758) in the Democratic Republic of the Congo: Repeated introductions since colonial times with multiple sources

Peer-reviewed author version

GERAERTS, Mare; Vangestel, Carl; ARTOIS, Tom; Fernandes, Jorge Manuel de Oliveira; JORISSEN, Michiel; Manda, Auguste Chocha; Mizani, Celestin Danadu; SMEETS, Karen; Snoeks, Jos; Sonet, Gontran; Yang Tingbao; VAN STEENBERGE, Maarten; Vreven, Emmanuel; Wamuini, Soleil Lunkayilakio; VANHOVE, Maarten & Huyse, Tine (2022) Population genomics of introduced Nile tilapia Oreochromis niloticus (Linnaeus, 1758) in the Democratic Republic of the Congo: Repeated introductions since colonial times with multiple sources. In: MOLECULAR ECOLOGY, 31 (12) ,p. 3304-3322.

DOI: 10.1111/mec.16479
Handle: http://hdl.handle.net/1942/37438

# Population genomics of introduced Nile tilapia (*Oreochromis niloticus* (Linnaeus, 1758)) in the Democratic Republic of the Congo: repeated introductions since colonial times with multiple sources

## Introduced Nile tilapia in the Congo Basin (running title)

Mare Geraerts[1†], Carl Vangestel[2,3], Tom Artois[1], Jorge Manuel de Oliveira Fernandes[4], Michiel W. P. Jorissen[1], Auguste Chocha Manda[5], Célestin Danadu Mizani[6], Karen Smeets[1], Jos Snoeks[7,8], Gontran Sonet[2], Yang Tingbao[9], Maarten Van Steenberge[2,8], Emmanuel Vreven[7,8], Soleil Lunkayilakio Wamuini[10,11], Maarten P. M. Vanhove[1,8,12,13], and Tine Huyse[7]

[1]Research Group Zoology: Biodiversity and Toxicology, Centre for Environmental Sciences, Hasselt University, Diepenbeek, Belgium

[2]Royal Belgian Institute of Natural Sciences, Brussels, Belgium

[3]Terrestrial Ecology Unit, Ghent University, Ghent, Belgium

[4]Faculty of Biosciences and Aquaculture, Nord University, Bodø, Norway

[5]Unité de recherche en Biodiversité et Exploitation durable des Zones Humides (BEZHU), Faculté des Sciences Agronomiques, Université de Lubumbashi, Lubumbashi, République Démocratique du Congo

[6]Département d'Ecologie et Biodiversité des Ressources Aquatique, Centre de Surveillance de la Biodiversité (CSB), Université de Kisangani, Kisangani, République Démocratique du Congo

[7]Department of Biology, Royal Museum for Central Africa, Tervuren, Belgium

[8]Laboratory of Biodiversity and Evolutionary Genomics, KU Leuven, Leuven, Belgium[9]Institute of Aquatic Economic Animals and Key Laboratory forImproved Variety Reproduction of Aquatic Economic Animals, Zhongshan University, Ghangzhou, China
[10]Département de Biologie, I. S. P. Mbanza-Ngungu, Mbanza-Ngungu, République Démocratique du Congo
[11]Functional and Evolutionary Morphology Laboratory, University of Liège, Liège, Belgium
[12]Zoology Unit, Finnish Museum of Natural History, University of Helsinki, Helsinki, Finland
[13]Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic

[†]Corresponding author: mare.geraerts@uhasselt.be

## Abstract

During colonial times, Nile tilapia *Oreochromis niloticus* (Linnaeus, 1758) was introduced into non-native parts of the Congo Basin (Democratic Republic of the Congo, DRC) for the first time. Currently, it is the most farmed cichlid in the DRC, and is present throughout the Congo Basin. Although Nile tilapia has been reported as an invasive species, documentation of historical introductions into this basin and its consequences are scant. Here, we study the genetic consequences of these introductions by genotyping 213 Nile tilapia from native and introduced regions, focussing on the Congo Basin. Additionally, 48 specimens from 16 other tilapia species were included to test for hybridisation. Using RAD sequencing (27,611 SNPs), we discovered genetic admixture with other tilapia species in several morphologically identified Nile tilapia from the Congo Basin, reflects their ability to interbreed and the potential threat they pose to the genetic integrity of native tilapias. Nile tilapia populations from the Upper Congo and those from the Middle-Lower Congo are strongly differentiated. The former show genetic similarity with Nile tilapia from the White Nile, while specimens from the Benue Basin and Lake Kariba are similar to Nile tilapia from the Middle-Lower Congo, suggesting independent introductions using different sources. We conclude that the presence of Nile tilapia in the Congo Basin results from independent introductions, reflecting the dynamic aquaculture history, and that their introduction probably leads to genetic interactions with native tilapias, which could lower their fitness. We therefore urge avoiding further introductions of Nile tilapia in non-native regions and to use native tilapias in future aquaculture efforts.

## Keywords

Invasive species, cichlid, RAD sequencing, genetic integrity, genetic structure, independent introductions

## Introduction

Aquaculture production is one of the fastest-growing food-producing sectors in the world (FAO, 2020). Together with fisheries, it plays a significant role in reducing hunger, promoting health, and reducing poverty by providing jobs and livelihood to millions of people (Dugan et al., 2010; FAO, 2020). Many people in Africa, especially those living near major rivers (Congo, Nile, and Niger rivers) and the Great Lakes (lakes Tanganyika, Victoria and Malawi), depend primarily on fish as a source of animal protein (Brummett et al., 2008; FAO, 2016; Satia, 2017).

Tilapias are, after carps, the world's most important group of freshwater species used in aquaculture (Eknath & Hulata, 2009), and they also have been introduced for capture fisheries and sportfishing (Trewavas, 1983; Welcomme, 1988). In this study, we use 'tilapia' to refer to a paraphyletic species assemblage, composed of several tribes (Dunz & Schliewen, 2013), belonging to the so-called haplotilapiine lineage within the cichlids (Teleostei: Cichliformes: Cichlidae). The most commonly farmed tilapia species is Nile tilapia, *Oreochromis niloticus* (Linnaeus, 1758), which belongs to the tribe Oreochromini (Dunz & Schliewen, 2013; FAO, 2016; Lind et al., 2012) and comprises eight subspecies that are recognised based on morphological characteristics, biogeography, behaviour, development, feeding, and analysis of partial mitochondrial DNA sequences (Seyoum & Kornfield, 1992; Trewavas, 1983) (Table 1; Figure 1). Nile tilapia is native to 22 countries, and its natural distribution roughly comprises the Nile Basin, several river basins in West Africa (Senegal, Gambia, Volta, Niger, Benue and Chad), various waterbodies of the East African Rift Valley (lakes Albert, Edward, Kivu, Baringo, Turkana, some shallow parts of Lake Tanganyika, and the Omo and Suguta Basins), Lake Tana in Ethiopia, and the Yarkon Basin in Israel (Trewavas, 1983) (Table 1; Figure 1). However, because of worldwide introductions, both deliberate through stocking and

99 unintentional through aquaculture escapees (Welcomme, 1988), its presence is now reported
100 in 105 countries (Froese & Pauly, 2021).
101 The popularity of Nile tilapia in aquaculture stems from its fast growth and reproductive rate,
102 and its ability to feed at a range of trophic levels and being tolerant to a range of environmental
103 conditions (Canonico et al., 2005; Philippart & Ruwet, 1982; Zengeya et al., 2012). However,
104 these same characteristics predispose it to be a successful invasive species (Canonico et al.,
105 2005; Trewavas, 1983; Welcomme, 1988; Zengeya et al., 2012). Farmed fish can escape from
106 aquaculture systems, establish themselves in local waterbodies and form feral populations
107 (Lind et al., 2012). Here, they can predate on eggs and small fish, compete with native fishes
108 for food and habitat resources, and introduce aquatic pathogens and parasites (Canonico et al.,
109 2005; Deines et al., 2016; Jorissen et al., 2020; Lind et al., 2012; Naylor et al., 2001;
110 Welcomme, 1988). These processes can cause a decline in the population size of native fish
111 species (including native tilapias), which indirectly results in the loss of genetic diversity. The
112 introduction of Nile tilapia can also have a direct genetic impact on native tilapia populations
113 through hybridisation, a process that is often exploited for aquaculture purposes (Bezault et al.,
114 2012; Brummett et al., 2004; Brummett & Ponzoni, 2009; Wohlfarth & Hulata, 1981).
115 Unintentional hybridisation between escaped Nile tilapia and native tilapia species is a major
116 concern for the genetic integrity of the latter and can cause a reduction of their overall degree
117 of adaptation or fitness (Brummett & Ponzoni, 2009; Lind et al., 2012; Shechonge et al., 2018).
118 Several cases of hybridisation have been recorded in the wild. The introduction of *O. niloticus*
119 has been linked to the decline of native tilapias through hybridisation in Lake Victoria
120 (Balirwa, 1992; Goudswaard et al., 2002), the Limpopo River system (D'Amato et al., 2007;
121 Firmat et al., 2013; Moralee et al., 2000) , and the Kafue River (Deines et al., 2014).
122 Aquaculture in sub-Saharan Africa is a relatively new activity and is characterised by
123 fluctuations caused by political instabilities and civil wars (Brummett et al., 2008; Toguyeni,

124 2004). Tilapia aquaculture probably originated during the Second World War in the region of
125 Lubumbashi, in the province Haut-Katanga in the then Belgian Congo (now the Democratic
126 Republic of the Congo (DRC)) (Charpy, 1954; Micha, 2013; Robert, 1976; Toguyeni, 2004),
127 producing mainly the native species *O. macrochir* and *Coptodon rendalli* (Boulenger, 1897)
128 (Huet, 1957, 1959; Micha, 2013; Thys van den Audenaerde, 1964; Toguyeni, 2004). These
129 species were also imported from Haut-Katanga into the Republic of the Congo and into the
130 Central African Republic (then part of French Equatorial Africa), where they were used in
131 aquaculture in the Middle Congo (Charpy, 1954; Lemasson, 1958). After WWII, aquaculture
132 production in the Upper Congo increased by the creation of several fry production centres. At
133 this point, the main cultured species were native *O. macrochir* and *C. rendalli*, and introduced
134 *O. niloticus* of unknown origin, with the latter outperforming the former two by the end of the
135 1950s (Micha, 2013; Thys van den Audenaerde, 1988). However, after the independence of
136 the country in 1960, aquaculture activities encountered numerous negative impacts due to the
137 hasty departure of the Belgian supervisory staff, lack of trained personnel, and remaining
138 political unrest (Brummett et al., 2008; Toguyeni, 2004). However, since 1996, fish production
139 restarted and has been growing since (Toguyeni, 2004). Currently, tilapia production in the
140 DRC is the highest among the Central African countries (Satia, 2017). Its annual tilapia
141 production increased from 2000 tons/year to 3000 tons/year between 2000 and 2010, with Nile
142 tilapia being the most farmed, followed by *O. macrochir* and *C. rendalli* (El-Sayed, 2013).
143 The Congo River Basin covers almost the entire area of the DRC and parts of its neighbouring
144 countries (Runge, 2007; Snoeks et al., 2011), and is divided into three main sections: the Upper
145 Congo running from its source until the Boyoma Falls, upstream of Kisangani; the Middle
146 Congo running from these falls until Pool Malebo near Kinshasa; and the Lower Congo running
147 from the outlet of Pool Malebo until its estuary in the Atlantic Ocean (Brummett et al., 2011;
148 Roberts & Stewart, 1976; Runge, 2007) (Figure 2). Nile tilapia is naturally present only in a

149 small part of the Congo Basin (in some shallow parts of Lake Tanganyika and Lake Kivu)
150 (Thys van den Audenaerde, 1964) (Figure 1). However, due to its extensive (historical)
151 introduction for aquaculture purposes, and the possible secondary unintentional dispersal of
152 aquaculture escapees, it has established itself throughout the entire basin (Decru et al., 2017a;
153 Kisekelwa et al., 2020; Lunkayilakio et al., 2010).
154 In view of the well-documented negative effects that the introduction of Nile tilapia can have
155 upon native species, it is paramount to identify and trace introductions. Moreover, regarding
156 the current efforts being made to boost Nile tilapia aquaculture (Micha, 2013), it is important
157 to understand the distribution of genetic diversity and structure of introduced Nile tilapia in the
158 Congo Basin as genetic diversity is a critical indicator for the evolutionary potential of
159 populations, an attribute that could be of great value for the management of aquaculture stocks
160 (Lind et al., 2012). We aim to gain insight into the historical introduction of Nile tilapia in the
161 Congo Basin and to assess possible genetic consequences on native tilapias and on introduced
162 Nile tilapia itself. We use a RAD sequencing approach to study the genetic structure of Nile
163 tilapia populations from the Upper, Middle, and Lower Congo Basin, including farmed as well
164 as feral populations. We hypothesise that: (i) a certain degree of genetic admixture exists in
165 introduced Nile tilapia due to their ability to interbreed with other tilapia species, (ii) feral
166 populations have higher genetic variation than farmed populations as a result of mixing of
167 escapees from different sources of farmed populations in combination with inbreeding and
168 artificial selection under farmed conditions, (iii) several (independent) introductions took place
169 using populations with different genetic backgrounds due to a turbulent aquaculture history
170 characterised by political instabilities and civil wars, and (iv) Nile tilapia populations from the
171 Upper, Middle, and Lower Congo Basin are genetically similar, i.e. that there is no genetic
172 differentiation between populations from the different sections, since aquaculture was first
173 developed in the Upper Congo (Lubumbashi), following transfer of specimens from the Upper

174 to the Middle Congo, as already reported for *O. macrochir* and *C. rendalli* (Charpy, 1954;
175 Lemasson, 1958).
176

## 177 Materials and Methods

### 178 Sample areas

179 In the present study, we focus on the part of the Congo Basin in the DRC, excluding Lakes
180 Tanganyika and Kivu (Figure 2), and will refer to this area as the 'Congo River Basin' (CRB).
181 When referring to the 'Upper' Congo, we intend the sections of the Congo Basin that fall within
182 the CRB. A total of 272 samples, consisting of fins, (dorsal) muscles, spleen, and gills stored
183 in 99% ethanol (v/v), were selected. Of these, 96 museum specimens were morphologically
184 identified as *O. niloticus*, and originated from different locations in the CRB: 33 from the
185 Upper, 29 from the Middle, and 34 specimens from the Lower Congo. These include specimens
186 from fish farms and feral specimens from rivers and lakes (Figure 2). Additionally, 74
187 specimens of *O. niloticus* from its native range were included (Nile River, Senegal River, the
188 Albertine Rift Valley (lakes Albert, Edward, George, Tanganyika, and Kivu, and the Ruzizi
189 River), Lake Tana, Lake Hashenge, and the Benue Basin), and 43 specimens from regions
190 where it has been introduced (China, Jordan, Madagascar, Uganda (Lake Victoria), Benin,
191 Togo, and Zimbabwe) to infer the origins of introduced specimens (Table S2; Table S3)..
192 Further, 48 specimens of other tilapia species, present in the collection of the Royal Museum
193 of Central Africa (RMCA), were included to study possible hybridisation between introduced
194 Nile tilapia and other (native) tilapia species: *Oreochromis aureus* (Steindachner, 1864), *O.*
195 *macrochir*, *O. andersonii*, *O. upembae* (Thys van den Audenaerde 1964), *O. leucostictus*
196 (Trewavas, 1933), *O. salinicola* (Poll, 1948), *O. schewebischi* (Sauvage, 1884), *Coptodon zillii*
197 (Gervais, 1848), *C. rendalli*, *C. congicus* (Poll & Thys van den Audenaerde, 1960), *Congolapia*
198 *bilineata* (Pellegrin, 1900), *Tilapia sparrmanii* Smith, 1840, *T. ruweti* (Poll & Thys van den

199 Audenaerde, 1965), *Sarotherodon melanotheron* Rüppell, 1852, *S. galilaeus* (Linnaeus, 1758),
200 and *Pelmatochromis ocellifer* Boulenger, 1899. Additionally, five morphologically identified
201 hybrid specimens obtained through crossing of *O. aureus* and *O. niloticus* were included,
202 caught in a natural ecosystem in Israel (escapees, or their descendants, from aquaculture
203 facilities), and six morphologically identified hybrid specimens from the Upper Congo River
204 (five between *O. niloticus* and *O. macrochir*, and one between *O. niloticus*, *O. macrochir*
205 and/or *C. rendalli*) (Table 1; Table S4).
206
207 **DNA extraction**
208 Total genomic DNA was extracted from the samples using the DNeasy® Blood & Tissue Kit
209 (Qiagen) following the manufacturer's instructions. The concentration of DNA extracted from
210 each individual was quantified with a Qubit® 2.0 Fluometer (Life Technologies, Paisley
211 (UK)).
212
213 **RAD library preparation**
214 Seventeen RAD libraries, each including 16 individuals, were prepared according to the
215 protocol described in Baird et al. (2008) (Baird et al., 2008) and Etter et al. (2011) (Etter et al.,
216 2011). First, the DNA of each individual was enzymatically digested with *SbfI-HF*® (NEB,
217 cut site 5'-CCTGCA^GG-3'). A first adaptor, containing forward amplification and Illumina
218 sequencing primer sites, was ligated to each digested DNA fragment. The uniquely barcoded
219 samples were then pooled into multiplexed libraries, followed by random mechanical shearing
220 with the Covaris® S220 Focused-ultrasonicator. Subsequently, fragments between 250 and
221 700 bp were selected using a BluePippin™ device (Sage Science, Beverly, MA, USA). Next,
222 the DNA was ligated to a second adaptor with a unique barcode ensuring PCR amplification

223 and the identification of different libraries. RAD libraries were 101 bp paired-end sequenced
224 on an Illumina platform with a HiSeq 4000 system at Macrogen Korea (Seoul, South Korea).
225
226 **SNP discovery and genotyping**
227 The overall quality of the reads in each library was checked with the software FastQC version
228 0.11.7 (Andrews et al., 2011). Raw reads were processed using Stacks v2.3b (Catchen et al.,
229 2011; Catchen et al., 2013). The intactness of the RAD cut site was checked and reads were
230 demultiplexed using the *process_radtags* module. Reads with a dubious RAD cut site or a low
231 quality score were discarded with the filtering options '-r', '-c', and '-q'. Next, PCR clones
232 were identified and discarded with the *clone_filter* module. Using the *kmer_filter* module, reads
233 were filtered according to the number of abundant *k*-mers they contained with the filtering
234 option '--abundant'. Reads were then mapped against the reference genome of *O. niloticus*
235 (O_niloticus_UMD_NMBU, accession number MKQE02000000 (Conte et al., 2017)) using
236 the BWA-MEM algorithm of the software BWA version 0.7.17 (Li & Durbin, 2009). Next, a
237 sequence dictionary was made with the same command line tools, and the reference sequence
238 and BAM files were indexed with SAMtools version 1.7. For the actual SNP discovery and
239 genotyping, the software GATK version 4.0.0.0 (McKenna et al., 2010) was used. With this
240 software, local realignment was performed with the 'RealignmentTargetCreator' and
241 'IndelRealigner' option so that the number of mismatching bases was minimised across all the
242 reads. Finally, SNPs were called with the 'UnifiedGenotyper' option using a Bayesian
243 genotype likelihood model. The resulting VCF file was filtered using VCFtools version 0.1.13
244 (Danecek et al., 2011) to include only high-quality SNPs with the following parameters: only
245 bi-allelic SNPs with a quality score above 30, and only SNPs that were successfully genotyped
246 in 80% of the individuals. Only one SNP per RAD tag was kept to minimise linkage

247 disequilibrium. To remove possible paralogues, sites characterised by heterozygosity excess
248 ($q$-value <0.05) were discarded. The final dataset included 27,611 SNPs.
249
250 **Genetic structure**
251 To investigate genetic population structure, the software STRUCTURE version 2.3.4
252 (Pritchard et al., 2000) was used. For each value of $K$ (number of clusters) ranging from one to
253 ten, ten iterations were run using the admixture model (generations = 20 000; burn-in = 10
254 000). The optimal number of clusters $K$ was inferred in Structure Harvester version 0.6.94 (Earl
255 & vonHoldt, 2012) by the Ln$P(K)$ and the derived delta $K$ calculated by the method of Evanno
256 et al. (2005) (Evanno et al., 2005). Because independent iterations resulted in different
257 outcomes, the optimal alignment of the three iterations with the highest estimated log
258 probability was determined using CLUMPP version 1.1.2 (Jakobsson & Rosenberg, 2007). For
259 each cluster, the individual's membership coefficient values ($Q$-values) were estimated. Plots
260 were visualised in DISTRUCT version 1.1 (Rosenberg, 2004). In case of a bimodal support for
261 different $K$ values, plots with both $K$ values were visualized. Genetic structure was further
262 assessed by performing a non-scaled, non-centred Principal Coordinate Analysis (PCoA) using
263 the R package 'adegenet' version 2.1.3 (Jombart, 2008) in R version 4.1.0 (R Core Team,
264 2021). For these analyses, the Albertine Rift Valley was split into the northern Nilotic (lakes
265 Albert, George, and Edward) and southern Congolese part (lakes Tanganyika and Kivu, and
266 Ruzizi River). Populations from Benin and Togo were considered together as both were
267 sampled in the Mono Basin. Populations from the Nile River (Egypt and Sudan) were
268 considered separately because of the large geographical distance between them.
269
270 **Discovery of admixed Nile tilapias**

271 Species identification based on morphology alone can be challenging (Blackwell et al., 2020;
272 Bradbeer et al., 2019; Ciezarek et al., 2021; Rhymer & Simberloff, 1996). In order to detect
273 the possible source of Nile tilapia populations from the Congo Basin, and to not overestimate
274 the genetic diversity and structure in these populations, we performed some additional analyses
275 to exclude possible hybrids or misidentified specimens from further analyses. We use the term
276 'admixture' to refer to genetic introgression resulting from interspecific crossings. First, an
277 exploratory STRUCTURE analysis was performed, revealing some aberrant individuals with
278 a high membership coefficient value to clusters Q3 and Q4 (when $K = 4$) (see Results). Based
279 on these results, individuals were selected as 'potential hybrids or misidentified specimens'
280 when their membership coefficient values (*Q*-values) to the minor clusters (Q3 and Q4) were
281 above 5%. Next, SNPs with large allele frequency differences between the population of
282 purebred *O. niloticus* (specimens of *O. niloticus* excluding the ones selected in the previous
283 step) and each of the other tilapia species (Table 1) were identified using PLINK version 1.9
284 (Purcell et al., 2007). Subsequently, all specimens were assigned to hybrid classes based on the
285 selected SNPs using the R package 'Hybrid index estimation' or 'HIest' version 2.0
286 (Fitzpatrick, 2012). Given the large number of markers included in this study (27,611 SNPs),
287 the assignment of an individual to a certain hybrid class was considered reliable when the log-
288 likelihood of the best-fit class was over two units greater than the log-likelihood of the second
289 best-fit class and within two units of the maximum log-likelihood (Fitzpatrick, 2012). Class '1'
290 is purebred Nile tilapia, class '2' is the purebred other tilapia species, class '3' is a F1 hybrid,
291 class '4' is a F2 hybrid, class '5' is a backcross to Nile tilapia, and class '6' is a backcross to
292 the other tilapia species. In total, 17 HIest analyses were performed, each time with *O. niloticus*
293 and another tilapia species from Table 1 as parental species. Individuals that were significantly
294 assigned to one of the hybrid classes (class 3, 4, 5, or 6) or purebred other tilapia species (class
295 2) were considered to be admixed. The same applies to those that were not assigned to the class

296 for purebred Nile tilapia (class 1) in any of the tests. We checked the performance of the HIest
297 analysis on six specimens from the Upper Congo that were morphologically identified as
298 hybrids.
299
300 **Genetic diversity and differentiation**
301 Pairwise $F_{st}$ values were calculated between specimens from each of the geographical regions
302 included in this study (Table S2) with Arlequin version 3.5 (Excoffier & Lischer, 2010), using
303 1,000 permutations. To account for multiple testing, FDR adjusted *p*-values were calculated
304 using the Benjamini and Hochberg procedure (Benjamini & Hochberg, 1995) with the R
305 package 'BiocManager' version 1.30.16 (Morgan, 2021) and a significance level of 0.05.
306 To explore whether the genetic differentiation between specimens within the Congo Basin
307 increases with geographical distance between them ('isolation by distance'), a Mantel test,
308 implemented in GenAlEx version 6.5 (Peakall & Smouse, 2006, 2012) with 999 permutations
309 was performed between the matrices of Euclidean genetic distances and hydrological distances.
310 The shortest hydrological distance was measured between each locality with QGIS version
311 3.18.1 by mapping the locations on a river network, splitting the network into segments, and
312 measuring the length of the segments between each pair of specimens. For the Mantel test, a
313 subset of 8,098 SNPs was randomly sampled with the '--thin 80 000' option in VCFtools, and
314 farmed specimens were excluded as they are not free to move.
315 Additionally, basic population genetic parameters (mean number of individuals typed per locus
316 per population ($N$), mean observed heterozygosity per locus ($H_o$), and mean expected
317 heterozygosity per locus ($H_e$)) were calculated using the R package 'diveRsity' version 1.9.90
318 (Keenan et al., 2013) for all locations, and for farmed and feral specimens of the different
319 sections of the Congo Basin. Mean allelic richness per locus ($A$) and mean private allelic

320 richness per locus ($A_{pr}$) were estimated using the rarefaction algorithm implemented in HP-
321 Rare version 1.1 (Kalinowski, 2005).
322

323 ## Results

324 **Detection and exclusion of admixed specimens of Nile tilapia**

325 The exploratory STRUCTURE analysis, including only specimens morphologically identified
326 as Nile tilapia ($n$ = 213), showed a bimodal $K$ value: $K$ = 2 (highest delta $K$ value), and $K$= 4
327 (highest mean Ln$P(K)$) (Table S5) (Figure 3). The following results are based on the optimal
328 number of four clusters ($K$ = 4). Most individuals had high membership coefficients to clusters
329 Q1 and Q2. Overall, the membership coefficient to cluster Q1 was higher for individuals from
330 the Middle and Lower Congo Basin than for individuals from the Upper Congo Basin, which
331 had a high membership coefficient to cluster Q2. A total of 86 individuals were identified as
332 'potentially admixed' based on the estimated membership coefficient to Q3 and Q4 (Table S6).
333 The HIest test resulted in the selection of 39 specimens of admixed Nile tilapia, each of which
334 had also been identified as admixed in the STRUCTURE analysis: six from the Upper Congo
335 (all farmed specimens), seven from the Middle Congo (all feral specimens), five from the
336 Lower Congo (two farmed and three feral specimens), four from the northern Albertine Rift
337 Valley, eight from the southern Albertine Rift Valley, two from Madagascar, two from the
338 Senegal Basin (Senegal), one from the Nile River (Sudan), one from Lake Victoria (Uganda),
339 and three from Lake Kariba (Zimbabwe) (Table S7). In the HIest test, five specimens were
340 significantly classified as a purebred 'other' tilapia species (class 2): one from the Middle
341 Congo (MC_MS_36_1) was classified as *O. macrochir* and another (MC_ULI_39_1) as *C.*
342 *bilineata*; one from the northern Albertine Rift Valley (UG_GRG_32_5) as *O. upembae*; one
343 from the southern Albertine Rift Valley (DRC_NYA_35_6) as *S. melanotheron*; and one from
344 Lake Kariba (ZIM_KAR_25_2) as *C. rendalli*. One specimen from the Lower (LC_INK_51_1)

345 and one from the Middle Congo (MC_BOO_43_1) were classified as an F2 hybrid (class 4)
346 between *O. niloticus* and *O. upembae*. Some specimens (e.g. UC_KAT_DEP_1_1) were
347 significantly assigned to class 1 in different HIest analyses, using a different tilapia species as
348 second parental species. This outcome means that these specimens were significantly assigned
349 to the class of purebred Nile tilapia in these analyses. The HIest test with morphologically
350 identified hybrids (Table S8) assigned one of the morphologically identified hybrids to the
351 class of purebred Nile tilapia (class 1). Also, specimens of other tilapia species were not always
352 assigned to the class of purebred other tilapia species (class 2) (Table S8).
353
354 **Genetic structure of purebred Nile tilapia**
355 PCoA and STRUCTURE analyses were performed without the admixed specimens of Nile
356 tilapia identified with the HIest analysis. The STRUCTURE analysis including only purebred
357 specimens from the CRB had a bimodal optimal number of clusters: $K = 2$ (highest delta $K$),
358 and $K = 6$ (highest mean Ln$P(K)$) (Table S5). In the plot (Figure 4a), individuals from the
359 Upper Congo had an overall high membership fraction to cluster Q2 (when $K = 2$) and Q3
360 (when $K = 6$). Individuals from the Middle and Lower Congo Basin had a high membership
361 fraction to Q1 (when $K = 2$) and Q6 (when $K = 6$). In the PCoA results, a clear geographical
362 genetic clustering was visible (Figure 4b; Figure S1). The first three principal components
363 explained respectively 13.9%, 6.6%, and 3.6% of the variation. Individuals from the Upper
364 Congo Basin were separated from individuals from the Middle and Lower Congo Basin by the
365 first principal component PCo1. Individuals from the Middle and Lower Congo Basin were
366 separated mainly by the third principal component PCo3. Individuals on the positive side of
367 PCo1 space (Figure 4b) had a higher membership fraction to Q2 than to Q1 in the
368 STRUCTURE analysis (when $K = 2$) (Figure 4a), while individuals on the negative side of
369 PCo1 space had a higher membership fraction to Q1 than to Q2 (when $K = 2$) (Figure 4a).

370 The STRUCTURE analysis including purebred Nile tilapia from all geographical regions also
371 had a bimodal optimal number of clusters: $K = 3$ (highest delta $K$), and $K = 5$ (highest mean
372 Ln$P(K)$) (Table S5). In the PCoA including purebred Nile tilapia from all geographical regions,
373 the first three principal components explained respectively 11.4%, 6.0%, and 4.7% of the
374 variation. In both STRUCTURE (Figure 5a), as well as the PCoA (Figure 5b; Figure S2)
375 analyses, most individuals from the Upper Congo Basin clustered together with most native
376 Nile tilapia from the Nile River (Sudan) and introduced Nile tilapia from feral populations from
377 Madagascar and China. Most individuals from the Middle and Lower Congo Basin clustered
378 with native Nile tilapia from the Nile River (Egypt), the Benue Basin (Cameroon), and the
379 southern Albertine Rift Valley (Ruzizi River: populations 37, 39, and 41 in Figure 5a), and
380 with introduced Nile tilapia from feral populations in the Mono Basin (Benin), and Lake Kariba
381 (Zimbabwe). Individuals from Lake Tana (Ethiopia) formed a small cluster separate from all
382 other locations (Figure 5b) with a high membership fraction to cluster Q5 (when $K = 5$) in the
383 STRUCTURE analysis (Figure 5a). The population from Lake Tana was genetically distinct
384 from all other Nile tilapia populations based on the PCoA (Figure 5b) and the STRUCTURE
385 analysis with $K = 5$ (Figure 5a).
386
387 **Genetic differentiation**
388 Pairwise $F_{st}$ analysis of purebred Nile tilapia from the CRB indicated that individuals from the
389 respective sections of the Congo Basin were significantly differentiated from each other.
390 Individuals from the Middle Congo were genetically most similar to those from the Lower
391 Congo ($F_{st} = 0.049$), and individuals from the Upper Congo were most differentiated from
392 individuals from the Lower Congo ($F_{st} = 0.161$) (Table 2 and S9).
393 When considering other introduced and native Nile tilapias (Table S9), pairwise $F_{st}$ values
394 indicated that individuals from the Upper Congo Basin were not significantly differentiated

395 from individuals from Sudan ($F_{st}$ = 0.010), and that the genetic differentiation with individuals
396 from Lake Hashenge ($F_{st}$ = 0.044), the Benue Basin ($F_{st}$ = 0.040), Lake Victoria ($F_{st}$ = 0.045),
397 and China ($F_{st}$ = 0.056) was significant, but relatively low (Table 2). Individuals from the
398 Middle Congo and Lower Congo Basin were not significantly differentiated from individuals
399 from the Benue Basin ($F_{st}$ = -0,009 and $F_{st}$ = 0.0100, respectively), and individuals from the
400 Lower Congo were not significantly differentiated from individuals from Lake Kariba ($F_{st}$ = -
401 0.010) (Table 2). The genetic differentiation between individuals from the Middle Congo and
402 the Lower Congo ($F_{st}$ = 0.04923), Nile Basin in Egypt ($F_{st}$ = 0.017), Mono Basin ($F_{st}$ = 0.016),
403 and Lake Kariba ($F_{st}$ = 0.041) was significant although relatively low (Table 2). Finally, the
404 genetic differentiation between individuals from the Lower Congo and the Nile Basin in Egypt
405 ($F_{st}$ = 0.081) and Mono Basin ($F_{st}$ = 0.048) was significant, though, relatively low (Table 2).
406 The Mantel test including purebred feral Nile tilapia from the CRB demonstrated no significant
407 correlation between the Euclidean genetic distances and the hydrological distances ($R$ = 0.157,
408 $p$-value = 0.090) (Figure S3). When considering fish from each section of the basin separately,
409 again no significant correlation was found in the Middle Congo ($R$ = 0.247, $p$-value = 0.230)
410 and Lower Congo ($R$ = 0.075, $p$-value = 0.400) (Figure S3). A Mantel test for feral individuals
411 from the Upper Congo was not performed, as there was a negligible hydrological distance
412 between the individuals.
413
414 **Genetic diversity**
415 When considering only purebred Nile tilapia from the CRB, no significant differences were
416 found between the respective sections of the basin in terms of genetic diversity (Table 3). Also,
417 within each section, no significant difference was found in the genetic diversity between
418 farmed and feral populations. (Table 3; Figure S4). When considering all other introduced and

419 native Nile tilapia, no statistically significant differences were found between geographic
420 regions (Table 4; Figure S5).
421

## 422 Discussion

423 Traditional morphometric and molecular markers used in previous studies (Agnèse et al., 1997;
424 Bezault et al., 2011; Seyoum & Kornfield, 1992; Trewavas, 1983; Vreven et al., 1998) have
425 low resolving power to unveil genetic differentiation within and between populations. Because
426 of the high number of SNPs that can be identified, the rise of NGS techniques provides an
427 efficient approach to increase resolution in population genomic studies and has already proven
428 its value in the assessment of population structure and diversity in cultured and feral
429 populations of Nile tilapia in Tanzania (Kajungiro et al., 2019). In the present study, 27,611
430 SNPs were derived from RAD-seq data to investigate: (i) whether introduced Nile tilapia
431 suffered from genetic contamination from other tilapia species due to their ability to interbreed,
432 (ii) whether feral populations have higher genetic variation than farmed populations as a result
433 of mixing of escapees from different sources of farmed populations in combination with
434 inbreeding and artificial selection in farmed conditions, (iii) whether one or rather several
435 (independent) introductions took place in the Congo Basin using different genetic backgrounds,
436 and (iv) the possible source(s) of historical introductions in this river basin.
437

### 438 Genetic contamination of introduced Nile tilapia

439 The classification of Nile tilapia into eight subspecies was contradicted by studies using
440 morphometric, allozyme, restriction fragment length polymorphism data, and microsatellite
441 data (Agnèse et al., 1997; Bezault et al., 2011; Rognon & Guyomard, 2003; Tibihika et al.,
442 2020; Vreven et al., 1998). Still, these studies based on traditional markers gave some

443 inconsistent results, which suggests that these genetic markers have insufficient resolving
444 power to characterise variation and/or mixing between phyletic lineages.
445 In addition to the debated taxonomy of Nile tilapia subspecies, morphological identification of
446 tilapias is challenging because divergence of phenotypic traits can be influenced by
447 environmental factors (Hornsby et al., 2013; Tibihika et al., 2018; Wohlfarth & Hulata, 1981).
448 In addition, morphological divergence can be induced by anthropogenic activities, e.g. the
449 introduction of populations with different genetic backgrounds, followed by intraspecific
450 admixture (Tibihika et al., 2018). Consequently, misidentifications in the field are inevitable.
451 Additionally, several taxonomic issues are known. *Coptodon zillii*, for example, is sometimes
452 used when referring to *C. rendalli* and *vice versa* (Wohlfarth & Hulata, 1981). Also, *O. aureus*
453 in Israel has been misidentified in the past as *O. niloticus* (Wohlfarth & Hulata, 1981).
454 Interspecific and intergeneric hybridisation between different species of tilapia make
455 identification based on morphology alone even more complicated, particularly after several
456 generations of backcrossing (Bezault et al., 2012; Brummett et al., 2004; Brummett & Ponzoni,
457 2009; Rhymer & Simberloff, 1996; Wohlfarth & Hulata, 1981). Moreover, introgression
458 resulting from hybridisation is not always reflected in morphology or in traits that can be easily
459 measured (Blackwell et al., 2020; Bradbeer et al., 2019; Ciezarek et al., 2021; Rhymer &
460 Simberloff, 1996; Shechonge et al., 2018).
461 By performing a HIest test on our RAD-seq data, we classified about 20% of the
462 morphologically identified Nile tilapia as admixed. For the CRB, this was the case for six
463 individuals from the Upper (all farmed), seven from the Middle (all feral), and five from the
464 Lower Congo (two farmed specimens and three feral) (Table S7). The results from the HIest
465 analysis suggest that most of the admixed specimens probably were backcrosses, implying the
466 viability of these hybrids and ongoing introgression. The presence of this hybrid swarm can

467 potentially have a negative effect on the native species through genetic swamping (Facon et
468 al., 2005; Gibson et al., 2019; Hohenlohe et al., 2013; Todesco et al., 2016).
469 Our findings should, however, be interpreted with caution due to some methodological
470 limitations that could have influenced our results. First, we selected a specimen as potentially
471 admixed when the membership coefficient values (*Q*-values) to the minor clusters (Q3 and Q4)
472 in the STRUCTURE analysis were above 5%. This selection could have influenced the
473 selection of SNPs that are divergent between the two parental species, and, subsequently, the
474 assignment of individuals to hybrid classes based on these SNPs. Secondly, results of the HIest
475 test that included specimens that were phenotypically identified as hybrids, assigned one
476 specimen to the class of purebred Nile tilapia (Table S8). This may reflect a high intraspecific
477 divergence of phenotypic traits. It may also indicate introgression that has been masked by
478 several generations of backcrossing, resulting in a low membership coefficient fraction for the
479 introgressed specimen and, consequently, resulting in the wrong hybrid class assignment in the
480 HIest analysis (Ciezarek et al., 2021). Additionally, classifying individuals in a limited set of
481 hybrid classes is not suitable after many generations of hybridisation and backcrossing
482 (Fitzpatrick, 2012). Also, the outcome of the HIest test is considered credible only if the log-
483 likelihood of the best-fit class was over two units greater than the log-likelihood of the second
484 best-fit class and within two units of the maximum log-likelihood (Fitzpatrick, 2012). This is
485 an arbitrary cut-off, which might influence which specimens we then consider to be admixed.
486 Unfortunately, we were not able to identify the exact hybrid status and respective parental
487 species of the admixed specimens. Not all candidate parental species present in the basin (e.g.
488 *Oreochromis lepidurus* (Boulenger 1899), *O. mortimeri* (Trewavas 1966), *O. mweruensis*
489 Trewavas 1983, *O. spilurus* (Günther 1894) (Froese & Pauly, 2021)) and aquaculture strains
490 (and their parental species) were included in our analyses. Therefore, the presence of admixed
491 specimens might have been underestimated. Also, some specimens of the parental species

492 might not themselves be pure species or even be misidentified, given that they were not all
493 significantly assigned to 'class 2' in the HIest analysis (Table S8).
494 Another factor underestimating the presence of hybrids is the fact that we focused on
495 morphologically identified Nile tilapia, ignoring those with a deviating morphology. In order
496 to understand the real impact of Nile tilapia introduction on native tilapias in the Congo Basin,
497 future research is required to more accurately identify the parental species of hybrids, and the
498 direction and extent of introgression. To reach this goal, more specimens phenotypically
499 resembling the respective native species and specimens with deviating morphology should be
500 included, as well as more pure native species and commonly used aquaculture strains.
501
502 **Genetic diversity of populations in the Congo Basin**
503 We hypothesised that Nile tilapia from the Upper Congo were the donor population for
504 aquaculture in the Middle and Lower Congo Basin, as aquaculture in the Congo Basin was first
505 developed in the Upper Congo, and because transfer of other tilapia species from the Upper
506 Congo to the rest of the Congo Basin has already been reported in the past (Charpy, 1954;
507 Lemasson, 1958). If this was the case, we would expect a higher genetic diversity in the Upper
508 Congo and a lower diversity in populations from the Middle-Lower Congo Basin due to
509 founder effects and similar genotypes of Nile tilapia populations from the different sections of
510 the basin.
511 Contrary to our expectations, Nile tilapia populations from the three sections of the CRB, as
512 well as Nile tilapia from the other sampled regions, are not significantly different in terms of
513 genetic diversity, suggesting multiple introductions into the Middle and Lower Congo (Table
514 3; Table 4; Figure S4: Figure S5). We also hypothesised that feral populations would have
515 higher genetic variation than farmed populations because of interbreeding of escapees from
516 different sources of farmed populations in combination with inbreeding and artificial selection

517 under farmed conditions. However, such an outcome was not apparent in our results (Table 3;
518 Figure S4). Similar results were found, for example, for freshwater bream *Abramis brama*
519 (Linnaeus 1758) (Hosseinnia et al., 2014) and Eurasian perch *Perca fluviatilis* Linnaeus 1758
520 (Khadher et al., 2016), where the relatively high genetic diversity in farmed populations was
521 ascribed to the swapping of broodstock between different farms and the regular introduction of
522 wild individuals (Khadher et al., 2016). In case of Nile tilapia, the relatively high diversity of
523 farmed populations could have resulted from multiple introductions from different sources, as
524 reported, for example, in Lake Victoria (Balirwa, 1992).
525 Interestingly, the overall levels of observed and expected heterozygosities in our study are
526 considerably lower than in prior genetic studies on Nile tilapia (Angienda et al., 2011; Dias et
527 al., 2016; Hassanien & Gilbey, 2005; Kajungiro et al., 2019; Lind et al., 2019; Mireku et al.,
528 2017; Moses et al., 2020; Romana-Eguia et al., 2005; Rutten et al., 2004; Sukmanomon et al.,
529 2012; Tibihika et al., 2019) (Table S10). As most aforementioned studies used microsatellites
530 as genetic markers, a direct comparison with these studies is not appropriate, as SNPs are bi-
531 allelic. However, when comparing our results with previous studies using SNPs,
532 heterozygosity values in the present study were still remarkably low (Kajungiro et al., 2019;
533 Lind et al., 2019; Moses et al., 2020). This difference could possibly be caused by the exclusion
534 of admixed individuals in the present study and by recent genetic bottlenecks of feral and
535 inbreeding of cultured populations.
536
537 **The use of several source populations**
538 No significant correlation was found between genetic and hydrological distances within the
539 CRB (Figure S3). Results from the PCoA (Figure 4b, 7, S1, and S2), STRUCTURE analyses
540 (Figure 4a and 5a), and pairwise $F_{st}$ comparisons (Table 2) suggested a clear genetic split
541 between populations from the Upper and Middle-Lower Congo, and a high genetic similarity

542 between populations of the Middle and Lower Congo. The presence of waterfalls between the
543 Middle and Upper Congo Basin (Runge, 2007) could preclude upstream migration of Nile
544 tilapia. Also, the well-developed social behaviour (i.e. non-random mating) and substrate
545 affinity (i.e., male territorial guarding and female parental care) makes Nile tilapia a rather
546 sedentary species, influencing population differentiation on a small geographical and temporal
547 scale (Bezault et al., 2011). This behaviour could cause a genetic divergence between
548 populations living at large geographical distances from each other. However, given the
549 relatively short history of modern aquaculture in the Congo Basin, it is implausible that this
550 mechanism has caused the observed population differentiation in the Congo Basin. Combining
551 the results from the PCoA, the STRUCTURE analyses, the pairwise $F_{st}$ comparisons, and the
552 fact that there is no significant difference between the genetic diversity of the different sections
553 of the Congo Basin, we suggest that Nile tilapia from the Upper Congo was not the main donor
554 for aquaculture in the Middle-Lower Congo. The current genetic structure of Nile tilapia in the
555 CRB can be explained by human-mediated gene flow in the form of independent introductions,
556 using different sources in the Upper and in the Middle-Lower Congo.
557 The genetic differentiation of the population from Lake Tana from all other populations was
558 also found by Tibihika et al. (2020) (Tibihika et al., 2020), and supports its status as a separate
559 subspecies, as suggested by Seyoum and Kornfield (1992) (Seyoum & Kornfield, 1992).
560

561 **Possible source(s) of Nile tilapia populations in the Congo Basin**

562 Documentation about historical Nile tilapia introductions in the Congo Basin is scant: only a
563 few introductions have been reported from Sudan to Brazzaville (Lower Congo, Republic of
564 the Congo) (Froese & Pauly, 2021) and from the Bouaké station (Ivory Coast) to Brazzaville
565 (Lower Congo, Republic of the Congo) and to Bangui (Middle Congo, Central African
566 Republic) (Thys van den Audenaerde, 1988). Introduction from the Lake Edward/George

567 system into the Middle Congo Basin was proposed by Decru et al. (2017a, 2017b) (Decru et
568 al., 2017a; Decru et al., 2017b),though, this introduction was not formally registered. In recent
569 years, aquaculture in the Upper Congo has been influenced by aquaculture activities in
570 Southern Africa, introducing several aquaculture strains believed to include improved strains,
571 such as the 'GIFT' (Genetically Improved Farmed Tilapia) and 'Chitralada' strain. These
572 introductions are probably facilitated by the border position of the Lubumbashi area and the
573 less-restrictive Congolese aquaculture policy (pers. obs., A. Chocha Manda, University of
574 Lubumbashi). But, here too, the exact origin of Nile tilapia introductions is undocumented.
575 The results from the PCoA (Figure 5b and S2), STRUCTURE analyses (Figure 5a), and
576 pairwise $F_{st}$ analysis (Table 2 and S9) restricted to purebred Nile tilapia from native and
577 introduced populations suggest some possible source populations for aquaculture in the Congo
578 Basin (Figures 5, and S2; Table 2 and S9). Native Nile tilapias from the Nile Basin in Sudan
579 are genetically similar to introduced Nile tilapia from the Upper Congo. Native Nile tilapia
580 from the Benue Basin (Cameroon) are genetically similar to introduced Nile tilapia from the
581 Middle and Lower Congo Basin. In addition, introduced Nile tilapias from Lake Kariba
582 (Zimbabwe) are genetically similar to introduced Nile tilapia from the Lower Congo. Possible
583 introductions from Sudan or the Ivory Coast to the Lower Congo Basin, from the Ivory Coast
584 to the Middle Congo Basin, from the Lake Edward/George system to the Middle Congo Basin,
585 or from Southern Africa to the Upper Congo Basin, could not be validated in the present study
586 because of our limited dataset and limited documentation of transfers of aquaculture stock.
587 We also found relatively low (though significant) genetic differentiation between specimens
588 from the Upper Congo and native specimens from Lake Hashenge (Ethiopia) and the Benue
589 Basin (Cameroon), and between specimens from the Upper Congo and introduced specimens
590 from Lake Victoria (Uganda) and the Songtao and Gaozhou Reservoirs (China) (Table 2). In
591 Lake Victoria, Nile tilapia was introduced from Lake Edward, from fish ponds in Kajjansi

592 (Uganda) and Lake Turkana (Balirwa, 1992; Fuerst et al., 2000; Pullin & Capili, 1988). In
593 China, only one introduction of Nile tilapia from Sudan has been documented (Pullin & Capili,
594 1988).
595 Nile tilapia from the Middle and Lower Congo are genetically relatively similar to each other
596 and to native individuals from the Nile (Egypt) and Benue basins (Cameroon), and to
597 introduced individuals from the Mono Basin (Benin and Togo) and Lake Kariba (Zimbabwe)
598 (Table 2). In the Mono Basin, introductions have been documented from stations in Ivory Coast
599 and Burkina Faso (Lazard, 1990; Lederoun et al., 2018; Montcho et al., 2015). In Lake Kariba
600 (Zimbabwe), one introduction was documented from Nakambala Estate Farm in Zambia
601 (Marshall, 1988) and recent research demonstrated the use of several strains of Nile tilapia used
602 in aquaculture in Lake Kariba (Makeche et al., 2020).
603 Clearly, Nile tilapia in most countries originate from multiple introductions using different
604 populations, of which the native source is often unknown. Even within a section of the Congo
605 Basin, e.g. the Upper Congo, multiple strains are currently used (e.g. 'Kipopo' and 'Israel'
606 strain at farm Kipopo). Furthermore, little introduction events are documented. Therefore, from
607 the data we have now, we cannot reach reliable conclusions upon the exact source of the strains
608 being cultured in the Congo Basin.
609 Besides the poorly documented introductions of Nile tilapia, our study was limited by the lack
610 of well-defined genetically improved aquaculture strains that are popular in Nile tilapia
611 aquaculture and the parental species used to produce them (such as 'GIFT', 'Chitralada',
612 'Ghana', etc.). The inclusion of these strains could help clarify our results. Especially the
613 inclusion of Nile tilapia strains from the Bouaké station (Ivory Coast) would be interesting, as
614 introductions from this station have been reported in the Republic of the Congo and the Central
615 African Republic (Thys van den Audenaerde, 1988). Also, the inclusion of specimens from
616 aquaculture facilities in southern Africa could highlight their current role in aquaculture in the

area of Lubumbashi. Additionally, we have considered Nile tilapia populations coming from the native regions to be native. However, given the worldwide transportation of genetically improved strains, and the high chance of escapees through pond flooding or floating cage breakages (Lind et al., 2012), we cannot exclude the possibility that individuals from the native region are already products of admixture with other strains of Nile tilapia or other tilapia species. To avoid this problem, historical, pre-aquaculture samples should be included in the analyses.

To conclude, our genetic results reflect the complex history of frequent and rather careless introduction and translocation events of Nile tilapia throughout the Congo Basin, without considering the genetic consequences that now emerge. Whilst introduced Nile tilapia dominates tilapia culture in the DRC (Toguyeni, 2004), several native tilapia species, such as *O. macrochir*, *S. galilaeus*, and *C. rendalli*, have a proven aquaculture potential (Lind et al., 2012). To reconcile conservation with the growing demand for fish, future initiatives should promote the use of native tilapias that are most suitable to local conditions and use these species as a genetic resource for potential breeding programs (Lind et al., 2012).

## Acknowledgements

We would like to acknowledge the financial support of the Belgian Science Policy (BELSPO) through JEMU (Joint Experimental Molecular Unit), Belgian Federal Science Policy Office (BRAIN-be Pioneer Project BR/132/PI/TILAPIA), EMBRC Belgium – FWO project GOH3817N, European Research Council project (no. 683210), Research Council of Norway (no. 250548/F20), VLIR-UOS South Initiative ZRDC2014MP084, the OCA type II project S1_RDC_TILAPIA and Mbisa Congo project (2013-2018), the latter two being framework agreement projects of RMCA with the Belgian Development Cooperation, and the Czech Science Foundation (P505/12/G112 - European Centre of Ichtyoparasitology ECIP). M.G. was

supported by the Bijzonder Onderzoeksfonds (BOF-project 7NI02) of Hasselt University for the analysis, interpretation of data and writing of the manuscript. M.W.P.J. is supported by a BOF Reserve Fellowship from Hasselt University. M.P.M.V. received travel grant K220314N from the Research Foundation – Flanders (FWO-Vlaanderen) and support from the Belgian Directorate-General for Development Cooperation and Humanitarian Aid (CEBioS programme at the Royal Belgian Institute of Natural Sciences), and is financed by the Special Research Fund of Hasselt University (BOF20TT06). Additionally, we appreciate the help of the people involved in field work and sampling. In particular, we would like to thank Eva Řehulková, Maria Lujza Červenka Kičinja, and Andrea Šimková (Masaryk University, Czech Republic), Shevy Rothman (Steinhardt Museum of Natural History, Tel Aviv University, Israel), Stephan Koblmüller (University of Graz, Austria), Antoine Pariselle (Institute de Recherche pour le Développement, University of Montpellier, France; Université Mohammed V de Rabat, Morocco), Jeanne Rasamy Razanabolana, Jean Robertin Rasoloariniaina, Roger Daniel Randrianiana, Sarah Rakotomamonjy and Natacha Rasozolaka (University d'Antananarivo, Madagascar), and Maxwell Barson (University of Zimbabwe, Zimbabwe) for providing valuable samples for our dataset. Fieldwork was carried out under permission 06/AR.ED./15 from the General Directorate for Fishery Resources and Fisheries, Ministry of Fisheries Resources and Fisheries of Madagascar, and mission statements 863/2014 (Faculté des Sciences Agronomiques, Université de Lubumbashi), C/075/2015/I.S.P./MBNGU/AUT.AC and AC/076/2015/I.S.P./MBNGU/AUT.AC. We acknowledge the help of Massimiliano Virgilio in the analyses and interpretation of our results, and of Dirk F.E. Thys van den Audenaerde (Royal Museum for Central Africa / KU Leuven) for sharing his insights on tilapia. Finally, we would like to thank Ria Vanderspikken (Hasselt University) and Natascha Steffanie (Hasselt University) for their technical support in the laboratory.

## Data Accessibility and Benefit-Sharing

Genotype data are available on DataDryad (doi:10.5061/dryad.sxksn035k).

## Author Contributions

T.H., M.P.M.V., and T.A. supervised the study. T.H., M.P.M.V., T.A., M.G., C.V., G.S., M.W.P.J., and K.S., helped in the setup of the sampling design. M.P.M.V., A.C.M., C.D.M., S.L.W., J.M.O.F., Y.T., M.V.S., E.V., and J.S. contributed to the collection and morphological identification of fish, and provided scientific background information. M.G. conducted the lab work and genetical analyses. C.V., G.S. and T.H. helped in the interpretation and discussion of the results. The first draft of the manuscript was written by M.G. All authors critically revised the draft and approved the final manuscript.

## Tables and Figures

**Table 1**
Species and number of tilapia specimens other than Nile tilapia and its hybrids included in this study with their introduction state (native or introduced) and sampling location

| Tilapia species or hybrid | Number of individuals | Sampling location | Native/Feral |
|---|---|---|---|
| *Congolapia bilineata* (Pellegrin, 1900) | 3 | Lefini River (Middle Congo), DRC | Native |
| *Coptodon congicus* (Poll & Thys van den Audenaerde, 1960) | 5 | Lindi, Lefini and Sangha River (Middle Congo), and Inkisi River (Lower Congo), DRC | Native |
| *Coptodon rendalli* (Boulenger 1897) | 4 | Lindi and Kasaï River (Middle Congo), Lake Kipopo (Upper Congo), DRC | Native |
| *Coptodon zillii* (Gervais, 1848) | 3 | Lake Edward, Uganda and Epulu River (Middle Congo), DRC | Feral in Lake Edward, native in Epulu River |

| | | | |
|---|---|---|---|
| *Oreochromis andersonii* (Castelnau 1861) | 3 | Kabwe and Kapabi Swamp (Upper Congo), Zambia | Feral |
| *Oreochromis aureus* (Steindachner, 1864) | 3 | Oued Draa River, Morocco | Native |
| *Oreochromis leucostictus* (Trewavas, 1933) | 3 | Ruzizi River and Lake George, Uganda | Native |
| *Oreochromis macrochir* (Boulenger 1912) | 5 | Lufira, Kiswishi, and Kimbeimbe River, and Lake Kipopo (Upper Congo), DRC | Native |
| *Oreochromis salinicola* (Poll, 1948) | 3 | Kalombe and Kabunda River (Upper Congo), DRC | Native |
| *Oreochromis schwebischi* (Sauvage, 1884) | 1 | Nyanga River, Republic of the Congo | Native |
| *Oreochromis upembae* (Thys van den Audenaerde 1964) | 3 | Lake Kabwe, Lake Kabele, and Fungwe River (Upper Congo), DRC | Native |
| *Pelmatochromis ocellifer* Boulenger, 1899 | 1 | Congo River (Middle Congo), DRC | Native |
| *Sarotherodon galilaeus* (Linnaeus, 1758) | 3 | Inkisi Basin (Lower Congo), DRC | Native |
| *Sarotherodon melanotheron* Rüppell, 1852 | 3 | Mono Basin, Benin | Native |
| *Tilapia ruweti* (Poll & Thys van den Audenaerde, 1965) | 2 | Kasaï River (Middle Congo), DRC | Native |
| *Tilapia sparrmanii* Smith, 1840 | 3 | Lufira and Fungwe River, Mulenda Lake (Upper Congo), DRC | Native |
| *O. niloticus* x *O. aureus* | 5 | Coastal Levant, Israel | Feral |
| *O. niloticus* x *O. macrochir* | 5 | Lake Kipopo and Bumaki farm(Upper Congo), DRC | Feral |
| *O. niloticus* x *O. macrochir* x *C. rendalli* | 1 | Lake Kipopo (Upper Congo), DRC | Feral |

**Table 2**

Matrix of population differentiation based on pairwise $F_{st}$ estimators between purebred Nile tilapia populations of the different sections of the CRB and purebred Nile tilapia from the other geographical regions. The number of individuals per regions is given between parentheses.

| | **Upper Congo, CRB (DRC)** | **Middle Congo, CRB (DRC)** | **Lower Congo, CRB (DRC)** |
|---|---|---|---|
| **Upper Congo, CRB (DRC)† (n = 27)** | / | 0.12523*** | 0.16182*** |
| **Middle Congo, CRB (DRC)† (n = 22)** | 0.12523*** | / | 0.04923*** |
| **Lower Congo, CRB (DRC)† (n = 29)** | 0.16182*** | 0.04923*** | / |
| **Senegal Basin (Senegal) (n = 3)** | 0.11235*** | 0.10721** | 0.20813** |

| | | | |
|---|---|---|---|
| **Nile Basin (Egypt) (n = 6)** | 0.07574*** | 0.01650* | 0.08123** |
| **Nile Basin (Sudan) (n = 2)** | 0.01017 | 0.13771** | 0.19954** |
| **Northern Rift Valley (Uganda) (n = 11)** | 0.14693*** | 0.20364*** | 0.28431*** |
| **Southern Rift Valley (DRC, Burundi) (n = 20)** | 0.11821*** | 0.08655*** | 0.15415*** |
| **Lake Tana (Ethiopia) (n = 9)** | 0.34724*** | 0.42356*** | 0.45884*** |
| **Lake Hashenge (Ethiopia) (n = 4)** | 0.04440* | 0.08569** | 0.14422** |
| **Jordan Basin (Jordan)† (n = 3)** | 0.07516*** | 0.13976** | 0.20413** |
| **Mono Basin (Benin, Togo)† (n = 8)** | 0.09486*** | 0.01626* | 0.04755* |
| **Benue Basin (Cameroon) (n = 4)** | 0.03951** | -0.00918 | 0.00990 |
| **Betsiboka, Rianila, Sofia Basin (Madagascar)† (n = 10)** | 0.06324*** | 0.14395*** | 0.18445*** |
| **Lake Victoria (Uganda)† (n = 4)** | 0.04500** | 0.07374** | 0.12459** |
| **Lake Kariba (Zimbabwe)† (n = 2)** | 0.17749** | 0.04051* | -0.00996 |
| **Songtao and Gaozhou Reservoir (China)† (n = 10)** | 0.05552*** | 0.15872*** | 0.19895*** |

Geographical regions with a '†' indicate the regions where Nile tilapia has been introduced. The five lowest values are underlined for each section of the CRB. Significancy levels (FDR adjusted *p*-values<0.05) are indicated by an asterisk (*<0.05, **<0.01, and ***<0.001)

## Table 3

Summary of genetic diversity among all purebred Nile tilapia in the Upper, Middle and Lower Congo (CRB) as well as among farmed and feral populations for these three sections of the CRB. Values are given as the mean per locus (and standard deviation) for the mean number of individuals typed per population (N), mean allelic richness (A), private allelic richness ($A_{pr}$), observed heterozygosity ($H_o$), and expected heterozygosity ($H_e$)

| Geographical region | Farmed/Feral | N | A | $A_{pr}$ | $H_o$ | $H_e$ |
|---|---|---|---|---|---|---|
| Lower Congo (CRB) | All | 25,6981 (1,8238) | 1,0239 (0,0761) | 0,0042 (0,0165) | 0,0164 (0,0607) | 0,0235 (0,0746) |
| | Feral | 6,9134 (1,5385) | 1,0337 (0,0999) | 0,0178 (0,0604) | 0,0165 (0,0655) | 0,0310 (0,0920) |
| | Farmed | 18,7847 (0,8098) | 1,0198 (0,0750) | 0,0064 (0,0266) | 0,0165 (0,0676) | 0,0193 (0,0731) |
| Middle Congo (CRB) | All | 19,8472 (1,6870) | 1,0269 (0,0808) | 0,0051 (0,0185) | 0,0223 (0,0704) | 0,0262 (0,0789) |
| | Feral | 14,8812 (1,6474) | 1,0281 (0,0839) | 0,0116 (0,0335) | 0,0220 (0,0696) | 0,0271 (0,0809) |
| | Farmed | 4,9660 (0,2229) | 1,0222 (0,0881) | 0,0082 (0,0411) | 0,0223 (0,0967) | 0,0199 (0,0792) |
| Upper Congo (CRB) | All | 26,7629 (0,7865) | 1,0259 (0,0803) | 0,0046 (0,0188) | 0,0243 (0,0790) | 0,0256 (0,0789) |
| | Feral | 4,9549 (0,2385) | 1,0267 (0,0957) | 0,0116 (0,0492) | 0,0260 (0,1011) | 0,0240 (0,0861) |
| | Farmed | 21,8080 (0,7927) | 1,0256 (0,0805) | 0,0107 (0,0342) | 0,0240 (0,0787) | 0,0250 (0,0789) |

## Table 4

Summary of genetic diversity of all purebred Nile tilapia from all geographical regions. Values are given as the mean per locus (and standard deviation) for the mean number of individuals typed per population (N), mean allelic richness (A), private allelic richness ($A_{pr}$), observed heterozygosity ($H_o$), and expected heterozygosity ($H_e$)

| Geographical region | Introduced/ Native | N | A | $A_{pr}$ | $H_o$ | $H_e$ |
|---|---|---|---|---|---|---|
| Upper Congo, CRB (DRC) | Introduced | 26.7629 (0.7865) | 1.0259 (0.0803) | 0.0046 (0.0188) | 0.0243 (0.0790) | 0.0256 (0.0789) |
| Middle Congo, CRB (DRC) | Introduced | 19.8472 (1.6870) | 1.0269 (0.0808) | 0.0051 (0.0185) | 0.0223 (0.0704) | 0.0262 (0.0789) |

| | | | | | | |
|---|---|---|---|---|---|---|
| Lower Congo, CRB (DRC) | Introduced | 25.6981 (1.8238) | 1.0239 (0.0761) | 0.0042 (0.0165) | 0.0164 (0.0607) | 0.0235 (0.0746) |
| Senegal Basin (Senegal) | Native | 2.8719 (0.3890) | 1.0299 (0.1135) | 0.0075 (0.0491) | 0.0232 (0.1021) | 0.0256 (0.0982) |
| Nile River (Egypt) | Native | 5.2802 (0.8465) | 1.0290 (0.0976) | 0.0056 (0.0317) | 0.0227 (0.0833) | 0.0261 (0.0879) |
| Nile River (Sudan) | Native | 1.9708 (0.1738) | 1.0207 (0.1061) | 0.0034 (0.0381) | 0.0189 (0.1071) | 0.0164 (0.0843) |
| Northern Rift Valley (Uganda) | Native | 10.4329 (0.8615) | 1.0264 (0.0815) | 0.0064 (0.0260) | 0.0232 (0.0773) | 0.0254 (0.0779) |
| Southern Rift Valley (DRC, Burundi) | Native | 18.3266 (1.5671) | 1.0277 (0.0830) | 0.0051 (0.0190) | 0.0218 (0.0663) | 0.0269 (0.0808) |
| Lake Tana (Ethiopia) | Native | 8.5739 (1.0865) | 1.0083 (0.0514) | 0.0035 (0.0378) | 0.0085 (0.0611) | 0.0103 (0.0694) |
| Lake Hashenge (Ethiopia) | Native | 1.6927 (0.8171) | 1.0192 (0.1111) | 0.0085 (0.0755) | 0.0079 (0.0753) | 0.0306 (0.1488) |
| Jordan Basin (Jordan) | Introduced | 2.9201 (0.2976) | 1.0280 (0.1096) | 0.0066 (0.0461) | 0.0261 (0.1130) | 0.0234 (0.0917) |
| Mono Basin (Benin, Togo) | Introduced | 6.7902 (1.1501) | 1.0279 (0.0920) | 0.0055 (0.0281) | 0.0188 (0.0686) | 0.0257 (0.0848) |
| Benue Basin (Cameroon) | Native | 1.6270 (1.1250) | 1.0165 (0.0961) | 0.0060 (0.0559) | 0.0099 (0.0747) | 0.1963 (0.3897) |
| Betsiboka, Rianila, Sofia Basin (Madagascar) | Introduced | 9.9021 (0.4381) | 1.0252 (0.0854) | 0.0048 (0.0262) | 0.0227 (0.0834) | 0.0242 (0.0814) |
| Lake Victoria (Uganda) | Introduced | 3.9550 (0.2241) | 1.0253 (0.0968) | 0.0034 (0.0259) | 0.0203 (0.0865) | 0.0222 (0.0849) |
| Lake Kariba (Zimbabwe) | Introduced | 1.9821 (0.1423) | 1.0172 (0.0969) | 0.0025 (0.0296) | 0.0160 (0.1005) | 0.0142 (0.0810) |
| Songtao and Gaozhou Reservoir (China) | Introduced | 9.9351 (0.3563) | 1.0234 (0.0815) | 0.0043 (0.0232) | 0.0229 (0.0846) | 0.0225 (0.0777) |

**Figure 1**

Map of Africa with rivers and lakes in black, regions with natural occurence of Nile tilapia shaded in grey (based on Trewavas (1983) (Trewavas, 1983) and Bezault et al. (2011) (Bezault et al., 2011)), and sampling locations depicted as red dots (sampling locations in China not shown): **a.** Genetic clusters as identified by Bezault et al. (2011) (see Discussion section), framed region is expanded in **b.**, **b.** Geographical distribution of subspecies of *O. niloticus* following Trewavas (1983) and Seyoum and Kornfield (1992)

**Figure 2**

Map of Africa (top left) with the framed region expanded. Outline of the Congo Basin in orange and the part of the basin that we focus on in this study shaded in grey. Sampling locations of introduced Nile tilapia within this area are indicated as red dots. Boyoma Falls and Pool Malebo define the transition from the Upper Congo to the Middle Congo, and the Middle Congo to the Lower Congo, respectively. Numbers refer to the population identifiers in Table S2. Kinshasa, Kisangani and Lubumbashi indicated by green stars. Main rivers and lakes in black (shapefiles downloaded from Figure.landscapeportal.org, maps created using QGis 3.18.1 software)

**Figure 3**

Population structure plot resulting from individual-based clustering using STRUCTURE with the two optimal $K$ values ($K$ = 2, $K$ = 4). All native and introduced specimens that were morphologically identified as Nile tilapia were considered in the analyses. Geographical regions are shown at the top. Each bar represents one individual, which is partitioned into as many as K coloured segments. The length of a coloured bar represents the estimated membership coefficient fraction ($Q$-values) in each of the K inferred clusters. Numbers at the bottom of the STRUCTURE plot represent the population identifiers as in Table S2

**Figure 4**

Visualisation of population structure including only purebred Nile tilapia from the CRB. **a.** Individual-based clustering using STRUCTURE with the two optimal *K* values (*K* = 2, *K* = 6). Geographical regions are indicated at the top. Each bar represents one individual, and is partitioned into as many as *K* coloured segments. The length of a coloured bar represents the estimated membership coefficient fraction (*Q*-values) in each of the K inferred clusters. Numbers at the bottom of the STRUCTURE plot represent the population identifiers as in Table S2. **b.** Genetic scatter plot of PCo1 versus PCo3 resulting from the PCoA. Each dot represents one individual. Colours represent different geographical regions. Ellipses are drawn at a confidence level of 0.95

**Figure 5**

Visualisation of population structure plot including purebred Nile tilapia from all native and introduced populations. **a.** Individual-based clustering using STRUCTURE with the two optimal *K* values (*K* = 3, *K* = 5). Geographical regions are indicated at the top. Each bar represents one individual, which is partitioned into as many as *K* coloured segments. The length of a coloured bar represents the estimated membership coefficient fraction (*Q*-values) in each of the K inferred clusters. Numbers at the bottom of the STRUCTURE plot represent the population identifiers as in Table S2. **b.** Genetic scatter plot of PCo1 versus PCo3 of the PCoA. Each dot represents one individual. Colours represent different geographical regions. Ellipses are drawn at a confidence level of 0.95

## Additional files

Additional supporting information may be found online in the Supporting Information section.

## References


Agnèse, J. F., Adépo-Gourène, B., Abban, E. K., & Fermon, Y. (1997). Genetic differentiation among natural populations of the Nile tilapia *Oreochromis niloticus* (Teleostei, Cichlidae). *Heredity*, *79*(1), 88-96. https://doi.org/10.1038/hdy.1997.126

Andrews, S., Lindenbaum, P., Howard, B., & Ewels, P. (2011). *FastQC high throughput sequence QC report*. In www.bioinformatics.babraham.ac.uk/projects/

Angienda, P. O., Lee, H. J., Elmer, K. R., Abila, R., Waindi, E. N., & Meyer, A. (2011). Genetic structure and gene flow in an endangered native tilapia fish (*Oreochromis esculentes*) compared to invasive Nile tilapia (*Oreochromis niloticus*) in Yala swamp, East Africa. *Conservation Genetics*, *12*, 243-255.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One*, *3*(10). https://doi.org/10.1371/journal.pone.0003376

Balirwa, J. S. (1992). The evolution of the fishery of *Oreochromis niloticus* (Pisces: Cichlidae) in Lake Victoria. *Hydrobiologia*, *232*, 85-89. https://doi.org/10.1007/BF00014616

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*(1), 289-300.

Bezault, E., Balaresque, P., Toguyeni, A., Fermon, Y., Araki, H., Baroiller, J.-F., & Rognon, X. (2011). Spatial and temporal variation in population genetic structure of wild Nile tilapia (*Oreochromis niloticus*) across Africa. *BMC Genetics*, *12*(1). https://doi.org/10.1186/1471-2156-12-102

Bezault, E., Rognon, X., Clota, F., Gharbi, K., Baroiller, J.-F., & Chevassus, B. (2012). Analysis of the meiotic segregation in intergeneric hybrids of tilapias. *International Journal of Evolutionary Biology*, *2012*, 10, Article 817562. https://doi.org/10.1155/2012/817562

Blackwell, T., Ford, A. G. P., Ciezarek, A. G., Bradbeer, S. J., Gracida Juarez, C. A., Smith, A. M., Ngatunga, B. P., Shechonge, A., Tamatamah, R., Etherington, G., Haerty, W., Di Palma, F., Turner, G. F., & Genner, M. J. (2020). Newly discovered cichlid fish biodiversity threatened by hybridization with non-native species. *Molecular Ecology*, *30*(4), 895-911. https://doi.org/10.1111/mec.15638

Bradbeer, S. J., Harrington, J., Watson, H., Warraich, A., Shechonge, A., Smith, A., Tamatamah, R., Ngatunga, B. P., Turner, G. F., & Genner, M. J. (2019). Limited hybridization between introduced and critically endangered indigenous tilapia fishes in northern Tanzania. *Hydrobiologia*, *832*, 257-268. https://doi.org/10.1007/s10750-018-3572-5

Brummett, R., Stiassny, M., & Harrison, I. (2011). Chapter 1. Background. In E. G. E. Brooks, D. J. Allen, & W. R. T. Darwall (Eds.), *The Status and Distribution of Freshwater Biodiversity in Central Africa*. IUCN.

Brummett, R. E., Angoni, D. E., & Pouomogne, V. (2004). On-farm and on-station comparison of wild and domesticated Cameroonian populations of *Oreochromis niloticus*. *Aquaculture*, *242*(1-4), 157-164.

Brummett, R. E., Lazard, J., & Moehl, J. (2008). African aquaculture: Realizing the potential. *Food Policy*, *33*(5), 371-385.

Brummett, R. E., & Ponzoni, R. W. (2009). Concepts, alternatives, and environmental considerations in the development and use of improved strains of tilapia in African aquaculture. *Reviews in Fisheries Science*, *17*(1), 70-77.

Canonico, G. C., Arthington, A., McCrary, J. K., & Thieme, M. L. (2005). The effects of introduced tilapias on native biodiversity. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *15*, 463-483. https://doi.org/10.1002/aqc.699

Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3*, *1*, 171-182.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124-3140.

Charpy, B. (1954). La pisciculture du tilapia en Afrique Equatoriale Française. *Revue Bois et Forêts des Tropiques*, *38*, 27-30.

Ciezarek, A., Ford, A. G. P., Etherington, G. J., Kasozi, N., Malinsky, M., Mehta, T., Penso-Dolfin, L., Ngatunga, B. P., Shechonge, A., Tamatamah, R., Haerty, W., Di Palma, F., Genner, M. J., & Turner, G. F. (2021). Whole genome resequencing data enables a targeted SNP panel for conservation and aquaculture of *Oreochromis* cichlid fishes. *bioRxiv*. https://doi.org/10.1101/2021.03.24.436760

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*, *18*(314). https://doi.org/10.1186/s12864-017-3723-5

D'Amato, M. E., Esterhuyse, M. M., van der Waal, B. C. W., Brink, D., & Volckaert, F. A. M. (2007). Hybridization and phylogeography of the Mozambuque tilapia *Oreochromis mossambicus* in southern Africa evidenced by mitochrondrial and microsatellite DNA genotyping. *Conservation Genetics*, *8*, 475-488. https://doi.org/10.1007/s10592-006-9186-x

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & Group, G. P. A. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156-2158.

Decru, E., Vreven, E., Danadu, C., Walanga, A., Mambo, T., & Snoeks, J. (2017a). Ichthyofauna of the Itimbiri, Aruwimi, and Lindi/Tshopo rivers (Congo Basin): Diversity and distribution patterns. *Acta Ichthyologica et Piscatoria*, *47*(3), 225-247.

Decru, E., Vreven, E., & Snoeks, J. (2017b). The occurences of an Eastern African haplochromine cichlid in the Ituri River (Aruwimi, Congo basin): Adaptive divergence in an introduced species? *Hydrobiologia*, *791*(1), 209-220.

Deines, A. M., Bbole, I., Katongo, C., Feder, J. L., & Lodge, D. M. (2014). Hybridisation between native *Oreochromis* species and introduced Nile tilapia *O. niloticus* in the Kafue River, Zambia. *African Journal of Aquatic Science*, *39*(1), 23-34. https://doi.org/10.2989/16085914.2013.864965

Deines, A. M., Wittmann, M. E., Deines, J. M., & Lodge, D. M. (2016). Tradeoffs among ecosystem services associated with global tilapia introductions. *Rev Fish Sci Aquac*, *24*(2), 178-191.

Dias, M. A. D., de Freitas, R. T. F., Arranz, S. E., Villanova, G. V., & Hilsdorf, A. W. S. (2016). Evaluation of the genetic diversity of microsatellite markers among four strains of *Oreochromis niloticus*. *Animal Genetics*, *47*, 345-353.

Dugan, P., Delaporte, A., Andrew, N., O'Keefe, M., & Welcomme, R. (2010). *Blue Harvest: Inland Fisheries as an Ecosystem Service*. WorldFish Center.

Dunz, A. R., & Schliewen, U. K. (2013). Molecular phylogeny and revised classification of the haplotilapiine cichlid fishes formerly referred to as "*Tilapia*". *Molecular Phylogenetics and Evolution*, *68*(1), 64-80. https://doi.org/10.1016/j.ympev.2013.03.015

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program visualizing STRUCTURE output and implementing the Evanno method. *Conservation of Genetics Resources*, *4*(2), 359-361. https://doi.org/10.1007/s12686-011-9548-7

Eknath, A. E., & Hulata, G. (2009). Use and exchange of genetic resources of Nile tilapia (*Oreochromis niloticus*). *Rev Aquacult*, *1*(3-4), 197-213. https://doi.org/10.1111/j.1753-5131.2009.01017.x

El-Sayed, A.-F. M. (2013). Tilapia feed management practices in sub-Saharian Africa. In M. R. Hasan & M. B. New (Eds.), *On-Farm Feeding and Feed Management in Aquaculture* (Vol. 583, pp. 377-405). FAO.

Etter, P. D., Bassham, S., Hohenlohe, P. A., Johnson, E. A., & Cresko, W. A. (2011). SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods Mol Biol*, *772*, 157-178.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*(8), 2611-2620.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, *10*, 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x

Facon, B., Jarne, P., Pointier, J. P., & David, P. (2005). Hybridization and invasiveness in the freshwater snail *Melanoides tuberculata*: Hybrid vigour is more important than increase in genetic variance. *Journal of Evolutionary Biology*, *18*(3), 524-535.

FAO. (2016). *The State of World Fisheries and Aquaculture 2016*. http://www.fao.org/3/a-i5555e.pdf

FAO. (2020). *The state of world fisheries and aquaculture 2020. Sustainablilty in Action*. https://doi.org/10.4060/ca9229en

Firmat, C., Alibert, P., Losseau, M., Baroiller, J.-F., & Schliewen, U. K. (2013). Succesive invasion-mediated interspecific hybridizations and populations structure in the endangered Cichlid *Oreochromis mossambicus*. *PLoS One*, *8*(5). https://doi.org/10.1371/journal.pone.0063880

Fitzpatrick, B. M. (2012). Estimating ancestry and heterozygotsity of hybrids using molecular markers. *BMC Evolutionary Biology*, *12*. https://doi.org/10.1186/1471-2148-12-131

Froese, R., & Pauly, D. (2021, 08/2021). *Fishbase. World Wide Web electronic publication. www.fishbase.org, version (08/2021).* www.fishbase.org

Fuerst, P. A., Mwanja, W. W., & Kaufman, L. (2000). The genetic history of the introduced Nile tilapia in Lake Victoria (Uganda - E. Africa): The population structure of *Oreochromis niloticus* (Pisces: Cichlidae) revealed by DNA microsatellite markers. In k. Fitzsimmons & J. C. Filho (Eds.), *Tilapia Aquaculture in the 21st Century: Proceedings from the Fifth International Symposium on Tilapia Aquaculture* (pp. 30-40). Panorama da Aquicultura Magazine.

Gaskin, J. F., & Schaal, B. A. (2002). Hybrid *Tamarix* widespread in U.S. invasion and undetected in native Asian range. *Proceedings of the National Academy of Sciences*, *99*(17), 11256-11259.

Gibson, I., Welsh, A. B., Welsh, S. A., & Cincotta, D. A. (2019). Genetic swamping and possible species collapse: tracking introgression between the native Candy Darter and introduced Variegate Darter. *Conservation Genetics*, *20*, 287-298. https://doi.org/10.1007/s10592-018-1131-2

Goudswaard, P. C., Witte, F., & Katunzi, E. F. B. (2002). The tilapiine fish stock of Lake Victoria before and after the Nile perch upsurge. *Journal of Fish Biology*, *60*, 838-856. https://doi.org/10.1006/jfbi.2002.1888

Hassanien, H. A., & Gilbey, J. (2005). Genetic diversity and differentiation of Nile tilapia (*Oreochromis niloticus*) revealed by DNA microsatellites. *Aquaculture Research*, *36*, 1450-1457.

Hohenlohe, P. A., Day, M. D., Amish, S. J., Miller, M. R., Kamps-Hughes, N., Boyer, M. C., Muhlfeld, C. C., Allendorf, F. W., Johnson, E. A., & Luikart, G. (2013). Genomic patterns of introgression in rainbow and westslope cutthroat trout illuminated by overlapping paired-end RAD sequencing. *Molecular Ecology*, *22*(11), 3002-3013. https://doi.org/10.1111/mec.12239

Hornsby, M. A., Sabbah, S., Robertson, R. M., & Hawryshyn, C. W. (2013). Modulation of environmental light alters reception and production of visual signals in Nile tilapia. *Journal of Experimental Biology*, *216*, 3110-3122.

Hosseinnia, Z., Shabany, A., & Kolangi-Miandare, H. (2014). Comparison of genetic variation of wild and farmed bream (*Abramis brama orientalis; Berg, 1905*) using microsatellite markers. *Molecular Biology Research Communications*, *3*(3), 187-195.

Huet, M. (1957). *Dix années de pisciculture au Congo belge et au Ruanda-Urundi: compte rendu de mission piscicole*. Direction de l'Agriculture, des Forêts et de l'Elevage.

Huet, M. (1959). Aperçu de la pisciculture dans les régions tropicales en Extrême-Orient et en Afrique Centrale. *Bulletin Français de Pisciculture*, *193*, 129-144.

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with switching and multimodality in analysis of population structure. *Bioinformatics*, *23*(14), 1801-1806.

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403-1405.

Jorissen, M. W. P., Huyse, T., Pariselle, A., Lunkayilakio, S. W., Bukinga, F. M., Manda, A. C., Kasembele, G. K., Vreven, E. J., Decru, E., Artois, T., & Vanhove, M. P. M. (2020). Historical museum collections help detect parasite species jumps after tilapia introductions in the Congo Basin. *Biological Invasions*, *22*, 2825-2844. https://doi.org/10.1007/s10530-020-02288-4

Kajungiro, R. A., Palaiokostas, 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 in Genetics*, *10*. https://doi.org/10.3389/fgene.2019.01269

Kalinowski, S. T. (2005). HP-Rare 1.0: A computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes*, *5*, 187-189.

Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, *4*, 782-788.

Khadher, S. B., Fontaine, P., Milla, S., & Agnèse, J.-F. (2016). Genetic characterization and relatedness of wild and farmed Eurasian perch (*Perca fluviatilis*): possible implications for aquaculture practices. *Aquaculture Reports*, *3*, 136-146.

Kisekelwa, T., Snoeks, J., & Vreven, E. (2020). An annotated checklist of the fish fauna of the river systems draining the Kahuzi-Biega National Park (Upper Congo: Eastern DR Congo). *Fish Biology*, *96*, 700-721.

Lazard, J. (1990). Transferts de poissons et développement de la production piscicole: exemple de trois pays d'Afrique subsaharienne. *Revue d'Hydrobiologie Tropicale*, *23*(3), 251-265.

Lederoun, D., Snoeks, J., Lalèyè, P., Vandewalle, P., & Vreven, E. (2018). An updated checklist of the ichthyofauna of the Mono River basin (Benin and Togo: West Africa). *Ichthyological Exploration of Freshwaters*, *28*(2), 135-156.

Lemasson, J. (1958). Chronique piscicole. La pisciculture en Afrique équatoriale française et au Cameroon. *Revue Bois et Forêts des Tropiques*, *57*, 57-61.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754-1760.

Lind, C. E., Agyakwah, S. K., Attipoe, F. Y., Nugent, C., Crooijmans, R. P. M. A., & Toguyeni, A. (2019). Genetic diversity of Nile tilapia (*Oreochromis niloticus*) throughout West Africa. *Scientific Reports*, *9*. https://doi.org/10.1038/s41598-019-53295-y

Lind, C. E., Brummett, R. E., & Ponzoni, R. W. (2012). Exploitation and conservation of fish genetic resources in Africa: Issues and priorities for aquaculture development and research. *Reviews in Aquaculture*, *4*(3), 125-141. https://doi.org/10.1111/j.1753-5131.2012.01068.x

Lunkayilakio, S. W., Vreven, E., Vandewalle, P., Mutambue, S., & Snoeks, J. (2010). Contribution à la connaissance de l'ichtyofaune de l'Inkisi au Bas-Congo (RD du Congo). *Cybium*, *34*(1), 83-91.

Makeche, M. C., Muleya, W., & Nhiwatiwa, T. (2020). Characterization of *Oreochromis niloticus* strains of Lake Kariba culture fisheries using morphological and meristic methods. *American Scientific Research Journal for Engineering, Technology, and Sciences*, *74*(1), 31-40.

Marshall, B. (1988). The status of wild and cultured tilapia genetic resources in various countries: Africa: Zimbabwe. In R. S. V. Pullin (Ed.), *Tilapia Genetic Resources for Aquaculture. ICLARM Conference Proceedings 16*. International Center for Living Aquatic Resources Management.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*, *20*(9), 1297-1303.

Micha, J.-C. (2013). Fish farming in the Congo Basin: past, present and future. In *Nutrition and Food Production in the Congo Basin* (pp. 147-171). Royal Academies for Science and the Arts of Belgium.

Mireku, K. K., Kassam, D., Changadeya, W., Attipoe, F. Y. K., & Adinortey, C. A. (2017). Assessment of genetic variations of Nile tilapia (*Oreochromis niloticus* L.) in the Volta Lake of Ghana using microsatellite markers. *African Journal of Biotechnology*, *16*(7), 312-321.

Montcho, S. A., Agadjihouèdé, H., Montchowui, E., Lalèyè, P. A., & Moreau, J. (2015). Population parameters of *Oreochromis niloticus* (Cichlidae) recently introduced in Lake Toho (Benin, West Africa). *International Journal of Fisheries and Aquatic Studies*, *2*(3), 141-145.

Moralee, R. D., van der Bank, F. H., & van der Waal, B. C. W. (2000). Biochemical genetic markers to identify hybrids between the endemic *Oreochromis mossambicus* and the alien species, *O. niloticus* (Pisces: Cichlidae). *Water SA*, *26*(2), 263-268.

Morgan, M. (2021). BiocManager: access the bioconductor project package repository. https://CRAN.R-project.org/package=BiocManager

Moses, M., Mtolera, M. S. P., Chauka, L., J, Lopes, F. A., de Koning, D. J., Houston, R. D., & Palaiokostas, C. (2020). Characterizing the genetic structure of introduced Nile tilapia (*Oreochromis niloticus*) strains in Tanzania using double digest RAD sequencing. *Aquaculture International*, *28*, 477-492. https://doi.org/10.1007/s10499-019-00472-5

Naylor, R. L., Williams, S. L., & Strong, D. R. (2001). Aquaculture - A gateway for exotic species. *Science*, *294*, 1655-1656.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: genetic analyses in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*, 288-295.

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. *Bioinformatics*, *28*, 2537-2539.

Philippart, J., & Ruwet, J. (1982). Ecology and distribution of tilapias. In R. S. V. Pullin & R. H. Lowe-McConnell (Eds.), *The Biology and Culture of Tilapias. ICLARM Conference Proceedings 7* (pp. 15-59). International Center for Living Aquatic Resources Management.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945-959.

Pullin, R. S. V., & Capili, J. B. (1988). Genetic improvement of tilapias: problems and prospects. In R. S. V. Pullin, T. Bhukasawan, K. Tonguthai, & J. L. Maclean (Eds.), *The Second International Symposium on Tilapia in Aquaculture. ICLARM Conference Proceedings 15*. Department of Fisheries, Bankok, Thailand, and International Center for Living Resources Management.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool

set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559-575. https://doi.org/10.1086/519795

R Core Team. (2021). *R: A language and environment for statistical computing*. In R Foundation for Statistical Computing. https://www.R-project.org/

Rhymer, J. M., & Simberloff, D. (1996). Extinction by hybridization and introgression. *Annual Review of Ecology and Systematics*, *27*, 83-109.

Robert, R. (1976). La pisciculture du tilapia a la Réunion. *Revue Forestière Française*, 378-387.

Roberts, T. R., & Stewart, D. J. (1976). An ecological and systematic survey of fishes in the rapids of the Lower Zaïre or Congo River. *Bulletin of the Museum of Comparative Zoology*, *147*(6), 239-317.

Rognon, X., & Guyomard, R. (2003). Large extent of mitochondrial DNA transfer from *Oreochromis aureus* to *O. niloticus* in West Africa. *Molecular Ecology*, *12*(2), 435-445. https://doi.org/10.1046/j.1365-294x.2003.01739.x

Romana-Eguia, M. R. R., Ikeda, M., Basiao, Z. U., & Taniguchi, N. (2005). Genetic changes during mass selection for growth in Nile tilapia, *Oreochromis niloticus*), assessed by microsatellites. *Aquaculture Research*, *36*, 69-78.

Rosenberg, N. A. (2004). Distruct: A program for the geographical display of population structure. *Molecular Ecology Notes*, *4*, 137-138.

Runge, J. (2007). The Congo River, Central Africa. In A. Gupta (Ed.), *Large Rivers: Geomorpohology and Management* (pp. 293-309). John Wiley & Sons.

Rutten, M. J. M., Komen, H., Deerenberg, R. M., Siwek, M., & Bovenhuis, H. (2004). Genetic characterization of four strains of Nile tilapia (*Oreochromis niloticus* L.) using mircosatellite markers. *International Society for Animal Genetics*, *35*, 93-97.

Satia, B. P. (2017). Regional review on status and trends in aquaculture development in sub-Saharan Africa - 2017. *FAO Fisheries and Aquaculture Circular*, *1135/2*. https://www.fao.org/3/i6873e/i6873e.pdf

Seyoum, S., & Kornfield, I. (1992). Taxonomic notes of the *Oreochromis niloticus* subspecies-complex (Pisces: Cichlidae), with a description of a new subspecies. *Canadian Journal of Zoology*, *70*(11), 2161-2165. https://doi.org/10.1139/z92-291

Shechonge, A., Ngatunga, B. P., Tamatamah, R., Bradbeer, S. J., Harrington, J., Ford, A. G. P., Turner, G. F., & Genner, M. J. (2018). Losing cichlid fish biodiversity: Genetic and morphological homogenization of tilapia following colonization by introduced species. *Conservation Genetics*, *19*, 1199-1209. https://doi.org/10.1007/s10592-018-1088-1

Snoeks, J., Harrison, I. J., & Stiassny, M. L. J. (2011). The status and distribution of freshwater fishes. In W. R. T. Darwall, K. Smith, D. Allen, R. Holland, I. Harrison, & E. Brooks (Eds.), *The Diversity of Life in African Freshwaters: Underwater, Under Threat. An Analysis of the Status and Distribution of Freshwater Species throughout mainland Africa*. IUCN.

Sukmanomon, S., Senanan, W., Kapuscinski, A. R., & Na-Nakorn, U. (2012). Genetic diversity of feral populations of Nile tilapia (*Oreochromis niloticus*) in Thailand and evidence of genetic introgression. *Aquaculture and Natural Resources*, *46*(2), 200-216.

Thys van den Audenaerde, D. F. E. (1964). Revision systematique des especes congolaises du genre *Tilapia* (Pisces, Cichlidae). *Annalen Koninklijk Museum voor Midden-Afrika Zoologische Wetenschappen*, *124*, 1-155.

Thys van den Audenaerde, D. F. E. (1988). Resource papers on tilapias in Africa: Natural distribution of tilapias and its consequences for the possible protection of genetic resources. In R. S. V. Pullin (Ed.), *Tilapia Genetic Resources for Aquaculture.*

*ICLARM Conference Proceedings 16* (pp. 108). International Center for Living Aquatic Resources Management.

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 Evolutionary Biology*, *20*. https://doi.org/10.1186/s12862-020-1583-0

Tibihika, P. D., Curto, M., Dornstauder-Schrammel, E., Winter, S., Alemayehu, E., Waidbacher, H., & Meimberg, H. (2019). Application of microsatellite genotyping by sequencing (SSR-GBS) to measure genetic diversity of the East African *Oreochromis niloticus*. *Conservation Genetics*, *20*, 357-372.

Tibihika, P. D., Waidbacher, H., Masembe, C., Curto, M., Sabatino, S., Alemayehu, E., Meulenbroek, P., Akoll, P., & Meimberg, H. (2018). Anthropogenic impacts on the contextual morphological diversification and adaptation of Nile tilapia (*Oreochromis niloticus*, L. 1758) in East Africa. *Environmental Biology of Fishes*, *101*, 363-381.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., Heredia, S. M., Hahn, M. A., Caseys, C., Bock, D. G., & Rieseberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications*, *9*(7), 892-908. https://doi.org/10.1111/eva.12367

Toguyeni, A. (2004). Tilapia production and its global impacts in Central African countries. In R. Bolivar, G. Mair, & K. Fitzsimmons (Eds.), *6th International Symposium on Tilapia in Aquaculture*. ICLARM.

Trewavas, E. (1983). *Tilapiine Fishes of the Genera Sarotherodon, Oreochromis and Danakilia*. British Museum (Natural History).

Vreven, E. J., Adépo-Gourène, B., Agnèse, J.-F., & Teugels, G. G. (1998). Morphometric and allozyme variation in natural populations and cultured strains of the Nile tilapia *Oreochromis niloticus* (Teleostei, Cichlidae). *Belgian Journal of Zoology*, *128*(1), 23-34.

Wasonga, A. G., Daniel, W. A., & Brian, O. (2017). Interspecific hybridization of tilapiines in Lake Victoria, Kenya. *Journal of Fisheries and Livestock Production*, *5*(2). https://doi.org/10.4172/2332-2608.1000235

Welcomme, R. L. (1988). International introductions of inland aquatic species. *FAO Fisheries Technical Paper*, *294*, 318.

Wohlfarth, G. W., & Hulata, G. I. (1981). *Applied Genetics of Tilapias* (Vol. 6). international Center for Living Aquatic Resources Management.

Zengeya, T. A., Robertson, M. P., Booth, A. J., & Chimimba, C. T. (2012). Ecological niche modeling of the invasive potential of Nile tilapia *Oreochromis niloticus* in African river systems: Concerns and implications for the conservation of indigenous congenerics. *Biological Invasions*, *15*(7), 1507-1521. https://doi.org/10.1007/s10530-012-0386-7