Hydrobiologia
https://doi.org/10.1007/s10750-018-3570-7

ADVANCES IN CICHLID RESEARCH III

Species integrity and origin of *Oreochromis hunteri* (Pisces: Cichlidae), endemic to crater Lake Chala (Kenya–Tanzania)

Jorunn Dieleman · Moritz Muschick · Wanja Dorothy Nyingi · Dirk Verschuren

Received: 14 October 2017 / Revised: 24 February 2018 / Accepted: 27 February 2018

Abstract Extensive transfer of tilapia between lakes throughout East Africa has often led to hybridisation with indigenous fish populations. The endemic *Oreochromis hunteri* of Lake Chala, an isolated crater lake near Mount Kilimanjaro, is potentially susceptible to introgression from a species formerly identified as *Oreochromis korogwe*, introduced ~ 30 years ago. We combined whole-body geometric morphometry on 104 specimens of both taxa with molecular phylogenetic analysis of mitochondrial loci from 15 *O. hunteri* and 9 *O.* cf. *korogwe* specimens to assess whether hybridisation has occurred. Using fishes from Lake Jipe and Nyumba ya Mungu reservoir, we expanded our analysis to all four *Oreochromis* species currently inhabiting the Upper Pangani River system to determine the closest relative of *O. hunteri*, and hence the possible source population of the ancestral species that colonised Lake Chala. Our results indicate no interbreeding occurs between *O. hunteri* and *O.* cf. *korogwe*, and suggest *O. jipe* to be the closest living relative of *O. hunteri*. The introduced *O.* cf. *korogwe* is a phenotypically uniform but genetically variable population, the identity of which remains unknown. The high haplotype diversity of *O. hunteri* is consistent with fossil evidence indicating that its ancestor colonised Lake Chala at least 25,000 years ago.

Keywords Introgression · Colonisation · Crater lake · Cichlids · Geometric morphometrics

Guest editors: S. Koblmüller, R. C. Albertson, M. J. Genner, K. M. Sefc & T. Takahashi / Advances in Cichlid Research III: Behavior, Ecology and Evolutionary Biology

Jorunn Dieleman and Moritz Muschick have contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10750-018-3570-7) contains supplementary material, which is available to authorized users.

J. Dieleman (✉) · D. Verschuren
Limnology Unit, Department of Biology, Ghent University, K. L. Ledeganckstraat 35, 9000 Ghent, Belgium
e-mail: jorunn.dieleman@ugent.be

M. Muschick
Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

M. Muschick
Aquatic Ecology & Evolution, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland

M. Muschick
Department of Fish Ecology & Evolution, EAWAG, Swiss Federal Institute for Aquatic Science and Technology, 6047 Kastanienbaum, Switzerland

Published online: 04 April 2018

Introduction

Tilapia is one of the most productive food fishes in Africa. Tilapia-based fisheries provide an often indispensable protein source for local food security, and especially *Oreochromis* species have been utilised for this purpose in global aquaculture (Eknath & Hulata, 2009). This has led to extensive transfer from their natural ranges into other regions and countries (Eknath & Hulata, 2009). The invasive nature of some *Oreochromis* species and their propensity for hybridisation rightfully has raised concerns when they are being introduced to regions with indigenous tilapiine communities (e.g. Agnèse et al., 1997; D'Amato et al., 2007; Nyingi & Agnèse, 2007; Angienda et al., 2011; Deines et al., 2014; Ndiwa et al., 2014). It has been suggested that the introgression of alien genes into local species may also contribute to rapid speciation in cichlids (Salzburger et al., 2002; Meier et al., 2017). However, most often it simply induces a loss of genetic diversity, through the homogenisation of gene pools (Nyingi & Agnèse, 2007; Crispo et al., 2011; Firmat et al., 2013).

Lake Chala (locally 'Challa', after a nearby village) is a crater lake bridging the border between Kenya and Tanzania, immediately to the southeast of Mt Kilimanjaro in East Africa. It harbours the only natural population of *Oreochromis hunteri* Günther, 1889 (Seegers et al., 2003), the type species of the genus *Oreochromis* (Günther, 1889; Trewavas 1983). Until the early 1980s it also seems to have been the only fish species inhabiting Lake Chala, as surveys carried out in 1889, 1902, 1946, 1952 and 1980 did not reveal other species (Günther, 1889; Dadzie et al., 1988). However, sometime in the late twentieth century two other tilapiine species were introduced, namely *Coptodon rendalli* (Boulenger, 1896) and *Oreochromis korogwe* (Lowe, 1955) (Dadzie et al., 1988; Seegers et al., 2003), as well as a small cichlid identified as *Haplochromis* spec. "Chala" (Seegers et al., 2003). The two *Oreochromis* species have the potential to hybridise, as the deep open-water environment of the lake encircled by near-vertical rocky crater walls appears to offer limited potential for reproductive or niche segregation.

W. D. Nyingi
Ichthyology Section, Zoology Department, National Museums of Kenya, P.O. Box 40658-00100, Nairobi, Kenya

Recent research on the long-term evolutionary and ecological dynamics of *O. hunteri* in Lake Chala is based on the analysis of its fossil teeth recovered from the sediment record (Dieleman et al., 2015). This study used the extant *O. hunteri* population as principal reference for variation in fossil tooth morphology, and assumed it to be genetically pure (Dieleman et al., 2015). However, introgression of genetic material due to interbreeding with an introduced species could potentially impact important morphological features (Parnell et al., 2012; Holzman & Hulsey, 2017). Such a recently compromised species integrity would complicate the comparison of modern phenotypes with the fossil record. Quantitative examination of general body morphology and the shape of oral teeth by Dieleman et al. (2015) found the two local *Oreochromis* species to be clearly distinct, arguing against recent and/or ongoing hybridisation. However, demonstration of the presence or absence of shared haplotypes with molecular genetic methods would provide a more sensitive test of hybridisation. Such genetic assessment is particularly relevant for the endemic *O. hunteri* population in Lake Chala, where detection of recent hybridisation would have implications for both taxonomy and conservation. Also, comparing genetic data of *O. hunteri* with that of *Oreochromis* populations from nearby waters might help reveal its phylogenetic associations, by providing information on the possible source population(s) from where the isolated crater lake was colonised, and on the approximate timing of this colonisation (Barluenga et al., 2006; Elmer et al., 2012; Genner & Turner, 2014).

In this study, we aim to (i) validate with genetic evidence the quantitative morphological differences and apparent lack of interbreeding between the endemic and introduced *Oreochromis* in Lake Chala; (ii) identify the closest relative of *Oreochromis hunteri* by molecular phylogenetic analysis of the *Oreochromis* fauna inhabiting Lake Chala and the only two nearby lakes, namely Lake Jipe and the Nyumba ya Mungu reservoir; and (iii) provide an estimate of the timing of *O. hunteri*'s arrival in Lake Chala.

Study area

Lake Chala (3°19′S, 37°42′E) is a freshwater lake with a surface area of 4.5 km^2 and a maximum depth that

has fluctuated between 92 and 98 m since 1999 (Wolff et al., 2014). The lake fills a volcanic caldera basin at ~ 880 m above sea level immediately east of Mount Kilimanjaro, and is hydrologically mainly influenced by subsurface in- and outflow (Moernaut et al., 2010). Despite its isolated location, biogeographically it is considered part of the Upper Pangani River basin (Seegers et al., 2003), which also contains Lake Jipe (Kenya/Tanzania) and the man-made reservoir Nyumba ya Mungu (Tanzania; Fig. 1). The former is a medium-sized (30 km^2) but very shallow (< 3 m) muddy lake located immediately east of the Pare Mountains. It is fed by the Lumi River and drains into the Ruvu River, both at the northern end of the lake, and its shoreline is fringed with swamps. Nyumba ya Mungu was created in 1965, has a surface area of 110–180 km^2 depending on rainfall and drawdown, and gradually increases in depth from 4 in the north to 41 m at the dam (Denny, 1978). It lies in the north–south trending valley between the Lelatema and Pare Mountains, at the former confluence of the Ruvu and Kikuletwa rivers. The reservoir drains into the Pangani River, which flows into the Indian Ocean 500 km to the southeast.

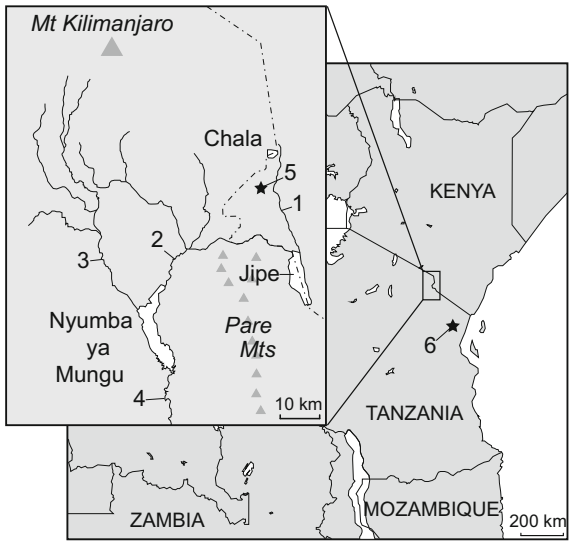

Fig. 1 Skeleton maps of East Africa and the Upper Pangani River basin (inset) in northern Tanzania, with indication of the sampled surface waters Chala, Jipe and Nyumba ya Mungu. River systems (1–4) and towns (5–6) mentioned in the text are indicated by numbers. *1* Lumi, *2* Ruvu, *3* Kikuletwa, *4* Pangani, *5* Taveta, *6* Korogwe

Regional ichthyofauna of the Upper Pangani system

The indigenous *Oreochromis* fauna of the Upper Pangani region is rather modest. Apart from *O. hunteri*, endemic to Lake Chala, *Oreochromis jipe* (Lowe, 1955) is considered indigenous to both Lake Jipe and the Pangani River itself (Lowe, 1955; Trewavas, 1983; Seegers et al., 2003). Trewavas (1983) also reports the species *Oreochromis pangani* (Lowe, 1955), with subspecies *O. pangani pangani* in the Pangani River and *O. pangani girigan* in Lake Jipe. However, following Seegers et al. (2003) we regard these as junior synonyms of *O. jipe*. Morphologically, *O. jipe* is considered the closest relative of *O. hunteri*, sharing high numbers of vertebrae, scales and dorsal fin rays (Trewavas, 1983), but to our knowledge this has never been confirmed by genetic data. Transfer of tilapiine fishes to improve local fisheries started influencing the ichthyofauna in this region from the 1970s onwards. For most of these transfers, no written records are available, and nearly all dates mentioned below are based on observations made during field surveys, rather than actual accounts.

A survey in 1980 found only the endemic *O. hunteri* in Lake Chala, but in 1985 catches also included *Coptodon rendalli* and one specimen that the collectors identified as *O. pangani* (Dadzie et al., 1988). Since the latter species does not occur in any recent catches, it either rapidly disappeared after its introduction, or the specimen was in fact *O. hunteri* misidentified as *O. pangani*. More recently a second *Oreochromis* species has been found in significant numbers, and has been thought to be *O. korogwe* (Seegers et al., 2003). Morphologically, however, Lake Chala specimens do not fully correspond to the *O. korogwe* holotype, which naturally inhabits the Lower Pangani River. Seegers et al. (2003) thus recommended to confirm this association, but no such study has been carried out so far.

The first written account of catches of *Oreochromis esculentus* (Graham, 1928) in Lake Jipe (which naturally occurs only in Lake Victoria) dates from 1983, followed by *C. rendalli* (naturally distributed throughout the Congo River basin, lakes Tanganyika and Malawi, and southern Africa) in 1985 (Dadzie et al., 1988). However, Trewavas (1983) suggests that at least *O. esculentus* must have been introduced there

already in the 1950s, but that it went unnoticed in previous surveys. In Nyumba ya Mungu, *O. esculentus* and *C. rendalli* were caught for the first time in respectively 1973 and 1974 (Bailey, et al., 1978). It is unclear whether *O. jipe* colonised this reservoir in a natural way, or was introduced (Trewavas, 1983).

This general lack of written records complicates determining the exact source populations of introduced species, but some inferences can be made. Probably, most of the region's lakes were stocked with fish from nearby fish ponds, because transporting living fry had to be logistically feasible. The fish ponds still present today around the town of Taveta and near Lake Jipe are good candidates, as they date back to the late 1940s when the British colonel Ewart Scott Grogan settled in the region and established a sisal farm that included such ponds (Dadzie et al., 1988). Tilapiines were also reared in governmental fish ponds at the town of Korogwe, located downstream along the Pangani River, for stocking reservoirs and ponds throughout the Tanganyika region (Lowe, 1955; Bailey et al., 1978). Among those species were *O. jipe* and *O. korogwe*, but also *O. esculentus* and *Oreochromis variabilis* (Boulenger, 1906) from Lake Victoria, and *C. rendalli* and *Oreochromis macrochir* (Boulenger, 1912) that had previously been raised in ponds in the D.R. Congo province of Katanga (Lowe, 1955; Dadzie et al., 1988). It seems thus most plausible that species introduced in the Upper Pangani Region, including Lake Chala, derive from populations that were reared in the Korogwe or Taveta ponds.

Methods

Specimen collection

For this study, we obtained 104 specimens from local fishermen servicing nets on the Kenya (southeast) side of Lake Chala, in September 2012, January 2014 and September 2015. Similarly, 15 specimens were obtained from fishermen along the southeast shore of Lake Jipe and 10 from the north shore of Nyumba ya Mungu (hereafter, NyM) during the same periods. Initially the local names given to the diverse species were recorded for each specimen; the corresponding nominal species names were assigned afterwards. From Lake Chala, two *Oreochromis* species were distinguished, being *O. hunteri* ('Chala') and *O.* cf. *korogwe* ('Bandia'). From Lake Jipe, we collected *O. jipe* ('Asilia') and *O. esculentus* ('Polana'). At NyM we collected *O. jipe* ('Asilia') and *O. esculentus* ('Polana') (Supplementary Table 1).

Geometric morphometric analysis

Whole-body photographs of fresh specimens were taken in the field prior to preservation, from a perpendicular angle. This was carried out by positioning specimens on graph paper with their left side facing up and fins spread out.

Overall body morphology was analysed as in Dieleman et al. (2015), using 16 traditional landmarks (Fig. 2) digitised in tpsDig2 version 2.17 (Rohlf, 2010). The present study includes data on 95 Lake Chala specimens previously presented by Dieleman et al. (2015). Size was calibrated using the graph paper visible on the photograph. The digitised landmark dataset was aligned via Procrustes superimposition in the program CoordGen6 h of the Integrated Morphometrics Package software (Sheets, 2008). Procrustes coordinates and rescaled centroid size were saved as data matrix file in IMP format. All further analyses were performed on these Procrustes coordinates.

Overall variation in body shape was analysed using principal component analysis (PCA) in R (R

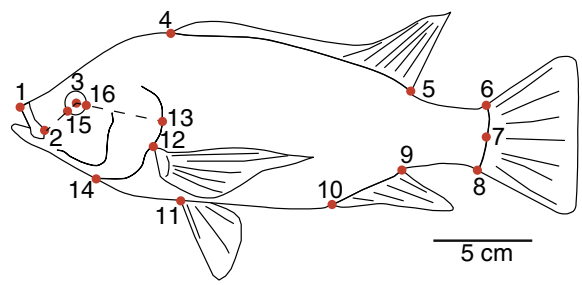

Fig. 2 Outline drawing of a female *Oreochromis hunteri* from Lake Chala with indication of the digitised landmarks. *1* rostral tip of the upper jaw, *2* caudodorsal tip of maxillary bone, *3* centre of the eye, *4* rostral insertion of the dorsal fin, *5* caudal insertion of the dorsal fin, *6* base of the dorsal caudal fin ray, *7* intersection between lateral line and insertion of the caudal fin, *8* base of the ventral caudal fin ray, *9* caudal insertion of the anal fin, *10* rostral insertion of the anal fin, *11* rostral insertion of the pelvic fin, *12* base of the dorsal pectoral fin ray, *13* most caudal point of the operculum, *14* ventral intersection between the branchiostegal membrane and body outline, *15* intersection between the line connecting landmarks 2 and 3 and the eye outline, *16* intersection between the line connecting landmarks 3 and 13 and the eye outline. After Dieleman et al. (2015)

Development Core Team, 2016). Multivariate Analysis of Variance (MANOVA) with four constraints was used to determine whether the fish taxa as distinguished by their local names indeed reflect significantly distinct shape morphs, and Canonical Variate Analysis (CVA) was used to determine the axes of maximal group separation. Both analyses were conducted on minimal Mahalanobis distances in PAST v. 2.17 (Hammer et al., 2001), and results were cross-validated by comparing leave-one-out (jackknifed) classifier tables to the original confusion matrix. Substantial differences between these two classifier tables would reflect the important influence of one or few specimens on the observed outcome, and hence reveal any unreliable results.

MtDNA genotyping and analysis

We isolated DNA from tissue sample collected and stored in the field in 99.5% absolute ethanol using the Blood and Tissue DNA isolation kit (Qiagen) following the manufacturer's specifications. Two mitochondrial loci, NADH dehydrogenase subunit 2 (ND2) and part of the control region (CR), were amplified via polymerase chain reaction (PCR), using the published primers MET and TRP for ND2 (Kocher et al., 1995), and L-Pro-F (Meyer et al., 1994) and TDK-D (Lee et al., 1995) for CR. PCR was performed in 25 µl reaction volumes, containing 1 µl DNA extract, 2.5 µl PCR buffer II (Applied Biosystems), 0.5 µl of either primer [10 µM], 1 µl $MgCl_2$ solution [25 mM], 0.4 µl dNTP solution [10 mM], 0.1 µl AmpliTaq Gold DNA polymerase [5U/µl] (Applied Biosystems), and 19 µl water. Thermocycling was performed with an initial denaturation for 3 min at 95°C, then 30 cycles with 30 s at 95°C, 30 s at 55°C, and 1 min at 72°C, followed by final elongation for 7 min at 72°C. Reaction products were cleaned up using ExoSAP-IT PCR Product Cleanup Reagent (Affymetrix), following the manufacturer's instructions. Purified PCR products were used for cycle sequencing reactions using the BigDye Terminator Mix v3.1 (Applied Biosystems). Cycle sequencing was performed for each sample and primer combination using 4 µl of BigDye Terminator 3.1 Ready Reaction Mix, 1.5 µl of primer [10 µM], 2 µl cleaned PCR product, and 2.5 µl water. Thermocycling was performed with 1 min at 96°C, then 25 cycles with 10 s at 94°C, 5 s at 50°C, and 4 min at 60°C. Cycle sequencing reactions were cleaned using ethanol precipitation. Sanger sequencing was performed on an ABI 3730 48-well capillary DNA Analyser (Applied Biosystems, Foster City, CA, USA). Electropherograms and their automatic translation were checked by eye and trimmed. Overlapping sequence reads from either direction were merged for each sample and locus. For each locus, 38 sequences were generated for this study (Supplementary Table 1).

In total, four sequence alignments were generated. Lengths differed between alignments due to the occurrence of gaps with more distantly related species, and as a result of trimming positions with excess missing data from the alignment ends. *O. hunteri* CR sequences ($N = 15$) were aligned (430 positions) and used for demographic analyses. All new *Oreochromis* sp. sequences ($N = 38$) were aligned for each locus separately (1051 positions for ND2, 435 for CR), and concatenated to generate a haplotype network. Additional, published sequences of each locus were downloaded from Genbank (https://www.ncbi.nlm.nih.gov, Supplementary Tables 2 and 3) and analysed jointly with those new to this study. The datasets were pruned and trimmed, resulting in two alignments with 1040 and 437 positions, and 71 and 90 sequences, for ND2 and CR, respectively. These alignments were used for locus-wise phylogenetic analysis in RAxML (version 8.2.4; Stamatakis, 2014). To find the best-scoring maximum likelihood tree in each case, we performed rapid bootstrap analyses using the GTRGAMMAI model of sequence evolution and 100 alternative runs from distinct starting trees. The resulting trees were visualised in, and figures created with, FigTree version 1.4.3 (Rambaut, 2009) and fitchi (Matschiner, 2016), for the locus-wise molecular phylogenies and the haplotype network, respectively.

To test for a genetic signature of past population expansion in *O. hunteri's* CR sequences, we performed a haplotype mismatch distribution analysis in Arlequin 3.5 (Excoffier et al., 2005) and a coalescent Bayesian skyline plot analysis (BSP; Drummond et al., 2005) in BEAST2, v2.4.5 (Bouckaert et al., 2014). Populations that have undergone a period of sudden or exponential growth in the past exhibit a characteristically unimodal, wave-like pattern in the distribution of haplotype mismatches. If present, the mode of this distribution together with estimates of generation time and mutation rate can be used to infer the approximate timing of population expansion (e.g. Barluenga et al.,

2006). We performed the BSP analysis using a strict molecular clock model with the base substitution rate estimate of 0.0324 changes per site per million years (SE 0.0139) of Genner et al. (2010). Our analysis employed a non-coding site model, and the Hasegawa–Kishino–Yano (HKY) substitution model with empirical base frequencies (Hasegawa et al., 1985), identified as best choice for these data, using jModelTest v. 2.1.10 (Darriba et al., 2012). Chain length was 25,000,000 steps, and the first 10% was discarded as burn-in. This analysis also estimated the timing of the deepest coalescence event of the *O. hunteri* CR sequences.

Results

Morphological characterisation of regional tilapiine taxa

The first two axes of the PCA (Fig. 3a) together explain 47.14% of the observed variation in general body morphology. Principal component 1 (PC1) mainly reflects variation in body elongation and depth, with long and slender bodies on the positive side and shorter, deeper bodies on the negative side of the axis (Fig. 3c). PC1 distinguishes *O. hunteri* from the other *Oreochromis* species. PC2 reflects differences within body depth: specimens with positive PC2 values are ventrally flattened, whereas negative PC2 values correspond to ventrally extended specimens (Fig. 3d).

Although assumptions for CVA and MANOVA were not met, jackknifed confusion matrices did not differ substantially from the original classification table. Therefore, we consider these results as reliable,

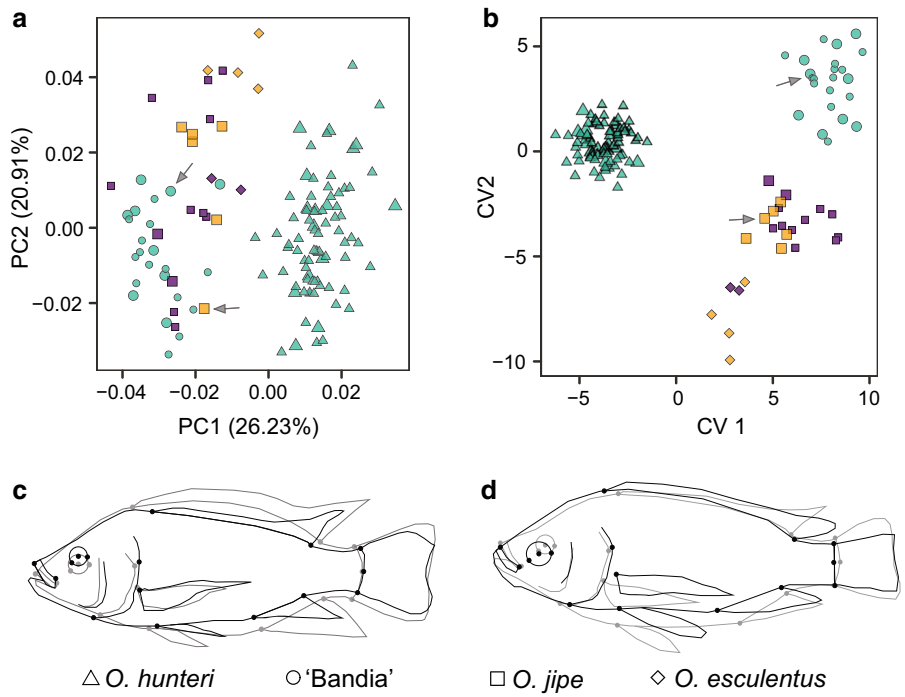

Fig. 3 **a** PCA ordination plot synthesising overall variation in body morphology among the four *Oreochromis* species currently inhabiting the Upper Pangani River basin. Green, purple and yellow symbols represent specimens from respectively Lake Chala, Lake Jipe and NyM reservoir, with large symbols of each type and colour indicating sequenced specimens. The two arrows point to the specimens of *O. jipe* and 'Bandia' with genetically distinct positions in the phylogenetic trees. **b** CVA scatter plot showing maximal phenotypic separation of the four groups of specimens attributed to each of the four taxa. The outline drawings represent **c** the positive (black) and negative (grey) extremes of PC1, and **d** the positive (black) and negative (grey) extremes of PC2

except for *O. esculentus*, where the small sample sizes (2 specimens from Jipe, 4 from NyM) prevented comparison with the other species except for *O. hunteri*. Uncorrected and Bonferroni-corrected pairwise comparison results indicate that *O. hunteri*, *O.* cf. *korogwe* and *O. jipe* specimens as identified by their local names differ significantly in overall body morphology at the (overall) 5% confidence level (Table 1; Wilks' $\lambda = 0.003$, $P < 0.001$); the CVA scatter plot (Fig. 3b) shows all four taxa to be clearly distinct from one another. *O. hunteri* is separated from the other species along the first axis (CV1), whereas the other species are separated along the second axis (CV2).

Genetic characterisation of regional *Oreochromis* taxa

Phylogenetic trees

Figure 4 depicts the phylogenetic trees of a selection of relevant African *Oreochromis* taxa based on sequence data from the mitochondrial control region (CR) and NADH dehydrogenase subunit 2 (ND2) genes, respectively. Specimens of the four *Oreochromis* taxa as identified by local fishermen, and as validated by geometric morphometric analysis, largely cluster into distinct mitochondrial clades. Further, CR and ND2 sequences suggest very similar phylogenetic affiliations, as could be expected from a non-recombining pair of markers. Only few of the studied specimens appear genetically distinct from these clusters. Specimens representing 'Chala' (*O. hunteri*) from Lake Chala and 'Asilia' (*O. jipe*) from both Lake Jipe and NyM cluster together in one clade, but with a distinct, albeit nested, split between the two species. Comparison with relevant GenBank sequences confirms the identity of 'Asilia' as *O. jipe* and suggests a close relationship with *Oreochromis amphimelas* (Hilgendorf, 1905) specimens from an unknown locality. The latter species occurs naturally in lakes of the Eastern Rift Valley in Tanzania (Manyara, Eyasi, Kitangiri, Singida), situated to the west of the Upper Pangani region. However, CR sequences of *O. amphimelas* from Lake Manyara occur in very different parts of the tree. Nevertheless, the latter clade is separated from *O. hunteri* and *O. jipe* by nodes with very low support values, so that their close relationship with *O. amphimelas* cannot be ruled out.

One 'Asilia' specimen from NyM (ASILIA_nym011) clusters with *Oreochromis niloticus* (Linnaeus, 1758) genotypes, although morphologically it groups with the other 'Asilia' examined in this study (Fig. 3a). The Lake Chala fishes identified by local fishermen as 'Bandia' (*O.* cf. *korogwe*), although morphologically uniform (Fig. 3a), display a fairly distinct split within their main clade, with CR and ND2 sequences separating the same two sub-groups of specimens. This differentiation appears greater than the variation within *O. hunteri*, or even the difference between *O. hunteri* and *O. jipe* (Fig. 4). Most of our 'Bandia' specimens cluster phylogenetically with *Oreochromis urolepis* (Norman, 1922). One available *O. urolepis* CR sequence, representing a specimen from the Wami river, was even identical to a subset of 'Bandia' sequences. This river belongs to the natural range of *O. urolepis*, and is located just south of the Pangani River basin. One particular Lake Chala specimen (BANDIA_cha053) is phylogenetically even further removed from the main 'Bandia' clade. Our phylogenetic reconstruction based on the CR gene suggests close affinity with an unidentified *Oreochromis* specimen (HT-1639) collected from Pangani River, and dubbed *Oreochromis* 'Korogwe' in the original publication (Nagl et al., 2001). Its ND2 sequence reveals that a specimen of *Oreochromis mweruensis,* (Trewavas, 1983) collected in Lake Mweru Wantipa (Zambia), may be a close relative (Klett & Meyer, 2002).

Table 1 Uncorrected (above the diagonal) and Bonferroni-corrected (below the diagonal) post hoc results, showing that all taxa as identified by their local names differ significantly in overall body morphology. The limited sample size of Polana caused pairwise comparisons to be non-applicable (N.A.) in some cases

	'Chala'	'Asilia'	'Bandia'	'Polana'
O. hunteri				
'Chala'		< 0.001	< 0.001	< 0.001
O. jipe				
'Asilia'	< 0.001		< 0.001	N.A.
O. cf. *korogwe*				
'Bandia'	< 0.001	0.003		N.A.
O. esculentus				
'Polana'	< 0.001	N.A.	N.A.	

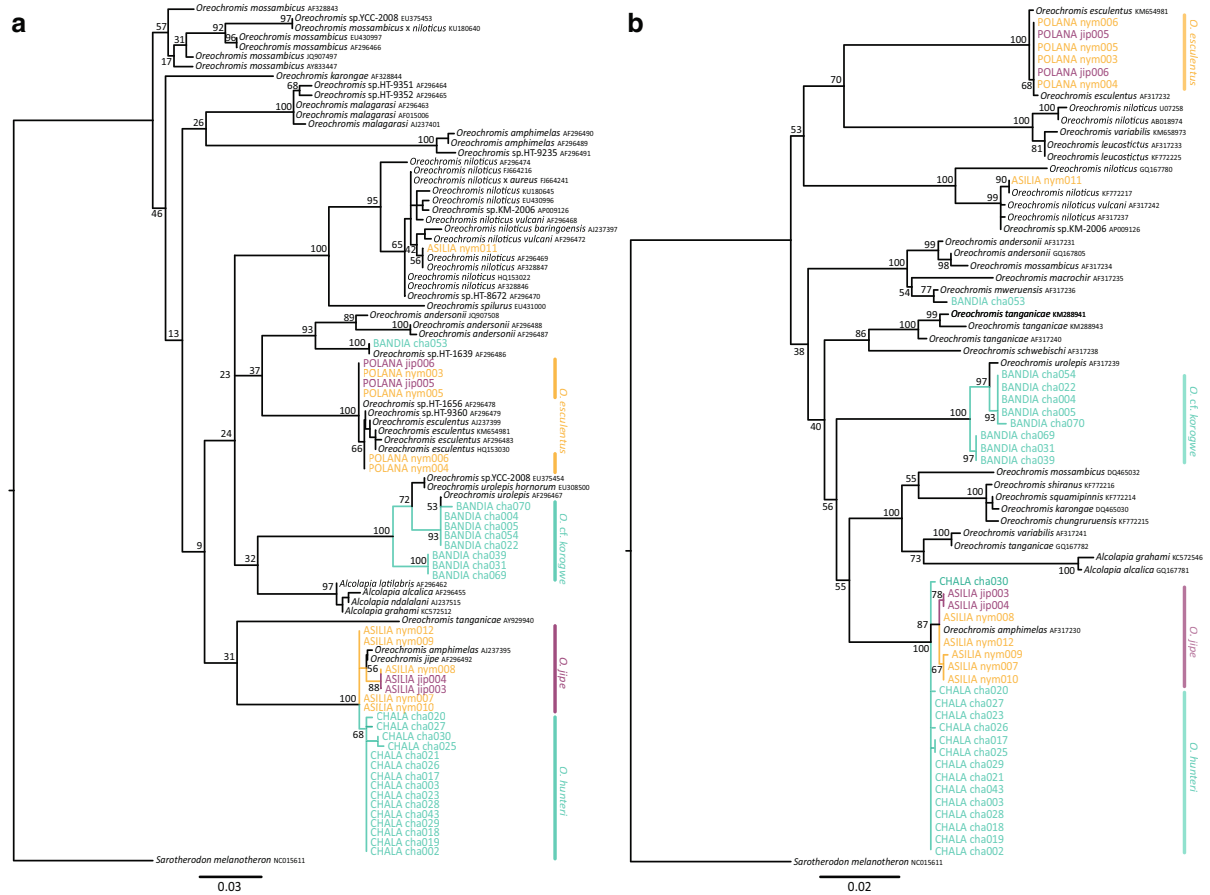

Fig. 4 Maximum likelihood phylogenetic trees of relevant African *Oreochromis* taxa based on sequences of the **a** control region (CR) and **b** NADH dehydrogenase subunit 2 (ND2) gene in mitochondrial DNA. *Sarotherodon melanotheron* was used as outgroup. The specimens sequenced in this study are colour-coded per site, as in Fig. 3: Chala (green), Jipe (purple) and NyM (yellow). Node support values are given in percent and are based on 100 bootstrap replicates. Note the low support for deeper nodes in both trees, but high support for those nodes on which identification of the closest relatives of the target species relies

Haplotype network

An unrooted haplotype network of all genotyped *Oreochromis* specimens based on concatenated CR and ND2 sequences reveals 20 different haplotypes (Fig. 5). Again, the four taxa locally recognised as such (and distinguished by morphology, Fig. 3a) also group into distinct mitochondrial clusters. Among our 15 *O. hunteri* specimens, haplotype 20 is the most abundant (9 specimens), with six less common haplotypes (one specimen each) differing from haplotype 20 by up to five mutation steps. These seven *O. hunteri* haplotypes share a hypothetical common ancestor (haplotype 22) with *O. jipe*. Our seven *O. jipe* specimens comprise five haplotypes, of which one is found in Lake Jipe (2 specimens) and four in NyM, three of which are more closely related to *O. hunteri* (Fig. 5).

The deep split in the main group of Lake Chala 'Bandia' (*O.* cf. *korogwe*) specimens separates haplotype 16 (3 specimens) from haplotypes 14, 15 and 17 (together 5 specimens), and reiterates the observations in the phylogenetic trees. Haplotypes of the aberrant 'Bandia' specimen from Lake Chala (12; near *O. mweruensis*) and 'Asilia' specimen from NyM (10; near *O. niloticus*) are also here strongly isolated from their respective, morphology-based, clusters.

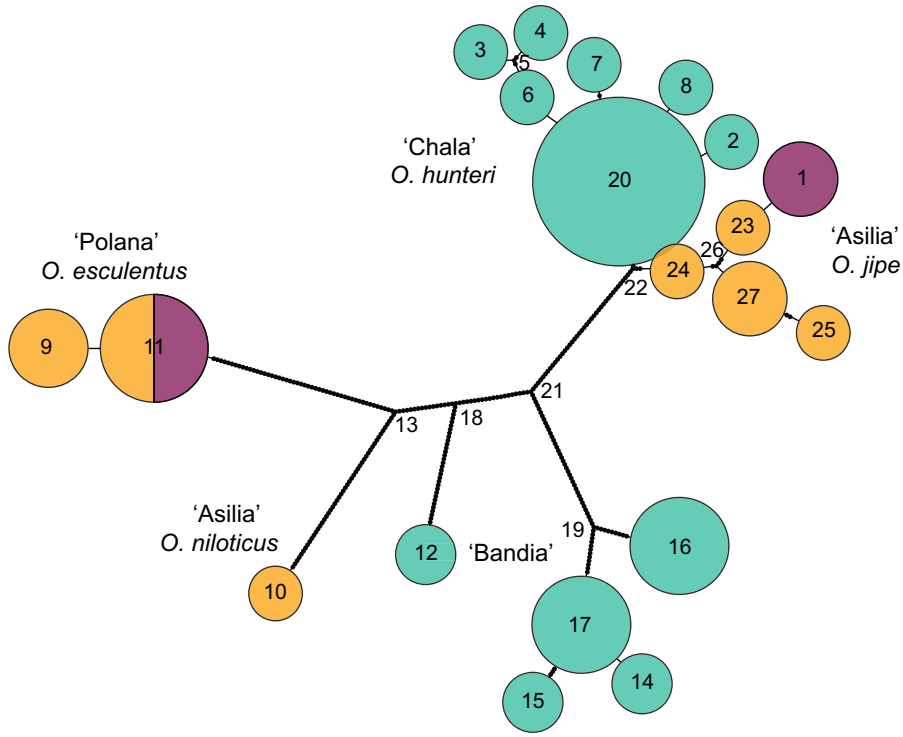

Fig. 5 Unrooted haplotype network of all 38 *Oreochromis* specimens genotyped in this study, based on mitochondrial CR and ND2 sequences. Each circle represents one haplotype, numbered 1–27 including the seven hypothetical haplotypes at network nodes. The diameter of the circles is proportional to the number of individuals with a certain haplotype (1–9); the black dots represent undocumented mutation steps between the haplotypes. Colour codes are by water body, as in Figs. 3 and 4: Chala (green), Jipe (purple) and NyM (yellow)

Mismatch distribution and skyline plot

The observed haplotype mismatch distribution of the 15 genotyped *O. hunteri* specimens is not unimodal (Fig. 6a), hence no distinct event of past population expansion can be defined. Likewise, the Bayesian skyline plot (Fig. 6b) does not provide evidence of a sudden large change in population size that might indicate a post-colonisation expansion. Instead, both results suggest that the size of the Lake Chala population has been relatively stable, or only slightly and continuously increasing, over an extended period of time. We estimate the oldest coalescence event for the CR sequences to have occurred approximately 100,000 years ago (100 ka), with a median of 92.4 ka and a 95% highest posterior density (HPD) interval ranging from 230 ka to 19 ka.

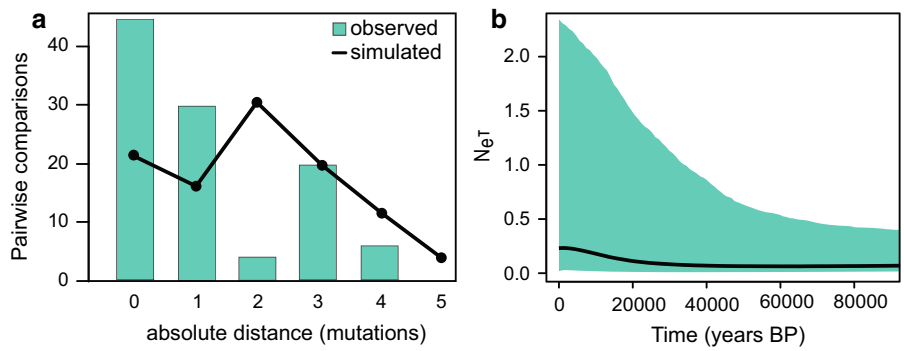

Fig. 6 **a** Empirical distribution of haplotype mismatches among the CR sequences of 15 *O. hunteri* individuals (green bars) and their expected distribution from simulations (black line). **b** Bayesian Skyline Plot visualising the modelled median effective female population size (N_eT, scaled by generation time τ) versus time in years before present (BP)

Discussion

Oreochromis in Lake Chala

To allow treatment of the extant *O. hunteri* as a modern-day reference for the fossil record of Lake Chala, hybridisation with species recently introduced to Lake Chala should ideally be ruled out. Although clear morphological distinction between *O. hunteri* and *O.* cf. *korogwe* was demonstrated earlier (Dieleman et al., 2015) and is confirmed by this study (Fig. 3), a complementary molecular genetic approach enabled direct assessment of the likelihood of past hybridisation events. The results of our phylogenetic analyses (Figs. 4, 5) show that Lake Chala specimens assigned to these two *Oreochromis* species form distinct, well-supported genetic clades and do not share mitochondrial haplotypes, suggesting that hybridisation in Lake Chala is absent or at least very rare. Hybridisation can be detected more reliably with a combination of mitochondrial and nuclear markers, as introgression does not necessarily always affect both genomes at the same time (Nyingi & Agnèse, 2007; Angienda et al., 2011). However, mitochondrial introgression appears to happen more readily than introgression of nuclear loci, due to, for example, interspecific matings being more likely to occur between females of the (initially) rare invading species and males of the normally more abundant native species (Wirtz, 1999). Mitochondrial alleles might also be comparatively neutral in a new genetic background, as opposed to alleles of nuclear genes, or of loci those genes are linked to (Martinsen et al., 2001). In *Oreochromis*, cases of mitochondrial introgression have been demonstrated to occur without apparent nuclear introgression (Rognon & Guyomard, 2003; Nyingi & Agnèse, 2007), or with parallel introgression of only a few nuclear loci (Ndiwa et al., 2014). While this discrepancy might in part reflect a bias towards traditionally studied mitochondrial loci, these studies at least demonstrate that mitochondrial introgression in *Oreochromis* appears to occur readily and frequently. As *O. hunteri* and *O.* cf. *korogwe* from Lake Chala can easily be distinguished phenotypically and are not found to share mitochondrial haplotypes despite three decades of syntopy, we consider the species integrity of *O. hunteri* to be currently intact, and regard its morphology as a reliable modern-day reference frame for interpretation of its fossil record.

The uniform and spatially contiguous habitat of Lake Chala does not provide much opportunity for species segregation, but the observed lack of mitochondrial exchange between the native and the introduced *Oreochromis* does suggest some form of reproductive isolation. Postzygotic barriers are not very prominent in closely related cichlid species (Stelkens et al., 2010), but various forms of premating isolation may prevent the two species from interbreeding. Visual identification of species-specific coloration patterns, olfactory cues and sound recognition are important segregation mechanisms in cichlid species (Fryer & Iles, 1972). In fact, Stelkens & Seehausen (2009) found that phenotypic divergence predicts assortative mating better than does genetic distance between species. Visual cues may be less important in tilapias, but other mechanisms of isolation have been suggested in this group, such as separation in spawning time and distinctions in microhabitat preference (Pullin & Lowe-McConnell, 1982; Lowe-McConnell, 1987; Beveridge & McAndrew, 2000).

Although distinction between species within tilapiine genera is often notoriously difficult (Nagl et al., 2001), the phenotypic/genotypic clusters of the four species analysed in this study coincide largely with usage of their local names, indicating that fishermen in each lake most often differentiate accurately and consistently between these fish taxa. However, one 'Asilia' specimen (in principle *O. jipe*) has a mitochondrial genotype clustering with *O. niloticus*¸ despite its phenotypical clustering with *O. jipe*. Our morphometric dataset does not contain *O. niloticus* specimens, and although this species has never been encountered in past Lake Jipe surveys (Dadzie et al., 1988; Seegers et al., 2003), the possibility exists that it has been introduced in recent years, and that this specimen is an actual *O. niloticus* or an individual carrying an introgressed mitochondrial haplotype. Importantly, the phenotypically uniform 'Bandia' (*O.* cf. *korogwe*) in Lake Chala also contains one individual with a very distinct genotype and a pronounced subdivision of the other specimens into two clades. A possible explanation for the occurrence of phenotype-genotype mismatches is that the aberrant mitochondrial genotype introgressed via an interbreeding event, either ancient or recent (Rognon & Guyomard, 2003; Ndiwa et al., 2014). This process does not explain the genotypic division within the

main 'Bandia' cluster, however. Historical collections suggest that *O.* cf. *korogwe* has been introduced to Lake Chala only during the 1980s (Dadzie et al., 1988). Therefore, it is unlikely that this deep split, more pronounced than even the split between *O. hunteri* and *O. jipe*, has arisen locally in such a short time, while the morphological uniformity of 'Bandia' implies that the population indeed consists of one species. We propose that multiple stocking events, or one single event containing a mix of genotypes, are a likely cause for these distinct genotypes currently coexisting in Lake Chala.

There also remains doubt about the true identity of 'Bandia', which Seegers et al. (2003) attributed to *O. korogwe* with some reticence. Whereas all nine specimens genotyped in this study are morphologically similar (Fig. 3a), most of their CR and ND2 sequences cluster with *O. urolepis*. The one exception has a CR sequence most similar to that of a specimen identified by Nagl et al. (2001) as *O.* 'Korogwe', but this only refers to the eastern Tanzanian village where it was collected (in Genbank this specimen is listed as *Oreochromis* sp.). The ND2 sequence of the same specimen suggests close relationship with a completely different species, *O. mweruensis.* No other *O. korogwe* sequences are currently available, and at this point we cannot rule out that our sequences would align with other *O. korogwe* specimens. Our data nevertheless suggest that the Lake Chala 'Bandia' (*O.* cf. *korogwe* in this paper) were stocked from at least two fish ponds, each containing a distinct *O. urolepis* or *O. korogwe* population which itself may already have undergone prior interbreeding with other *Oreochromis* taxa. Future genetic studies using nuclear markers may confirm this, and may elucidate the exact identity of 'Bandia'.

The likely ancestor of *O. hunteri*

Trewavas (1983) first suggested that *O. jipe* and *O. hunteri* might be closely related, on the basis of their similar number of vertebrae, which is generally higher (31–34) than in other *Oreochromis* species such as *O. esculentus* (30–31). One would therefore expect the Upper Pangani tilapiines to have relatively elongated bodies, and to cluster together in a PCA where body elongation is an important character separating phenotypes along the principal axis of variation (PC1). Our geometric morphometric data do not support this suggestion. Based on morphological data alone (Fig. 3), *O. hunteri* is distinct from the other three species, whereas *O. jipe* clusters with *O.* cf. *korogwe* and *O. esculentus*.

Our molecular phylogenetic analyses, in contrast, do reveal *O. jipe* and *O. hunteri* to be each other's closest relative (Fig. 4). The two species form a polytomy in both phylogenetic trees, but based on CR sequences *O. hunteri* is nested within *O. jipe*, whereas the opposite is true for ND2. This indicates that the employed markers may not be optimal to resolve the exact relationship of the two species, and further interpretation would be prudent. Nevertheless, genetic diversity of the *O. jipe* indigenous to Lake Jipe seems to be nested in the greater genetic diversity of modern-day *O. jipe* from NyM for both markers, despite the fact that the lacustrine habitat of NyM reservoir is only 50 years old. We cautiously suggest that this high diversity reflects standing variation retained through time in riverine populations of *O. jipe*, which seeded the newly formed NyM as well as the natural but climate-sensitive Lake Jipe, after a (relatively recent) environmental perturbation had eradicated its lacustrine population. In this context, we follow Seegers et al. (2003) in considering the riverine *O. pangani* as conspecific with the lacustrine *O. jipe*, notwithstanding some morphological differences in oral and pharyngeal teeth that had led Bailey et al. (1978) and Trewavas (1983) to describe them as two distinct species. Although no *O. pangani* specimens were available for us to address this issue with genetic analyses, we surmise that *O. pangani* may well be the riverine representative of *O. jipe*, which ensured the species' survival in the Upper Pangani River basin through past episodes of climatic drought when the region's shallow lakes fell dry. Although *O. jipe* clusters with an *O. amphimelas* specimen in both trees, most well-described specimens of the latter occur in other parts of the tree, and that particular specimen may be misidentified or may have been subject of mitochondrial introgression. Therefore, also the apparent relationship between *O. amphimelas* and *O. jipe* should be treated with caution.

Timing and mode of the colonisation of Lake Chala

The colonisation of isolated crater lakes by fish is still an enigmatic process (Barluenga & Meyer, 2010;

Elmer et al., 2012). The main mechanisms considered are human introduction, a hypothetical former aquatic connection and natural introduction by air (Elmer et al., 2012). Although undocumented, late 20th century human introduction is almost certainly how *O.* cf. *korogwe* and *C. rendalli* arrived in Lake Chala. Yet fossil fish teeth and bones occurring throughout the presently recovered part of the sediment record reveal presence of *O. hunteri* in Lake Chala since at least 25,000 years ago (Dieleman et al., 2015). Ancient, or at least pre-colonial, stocking of fishless lakes in this region of East Africa, if it did occur, was most likely restricted to the period after ca. 1000 AD, when Bantu farmers first settled in the Mt. Kilimanjaro region (Håkansson, 2008). Therefore, the introduction of fish by air, such as the transfer of fertilised eggs by birds, arguably remains the only plausible explanation of how the ancestor of *O. hunteri* arrived in Lake Chala, given that a direct hydrographic connection of the high-rimmed Chala crater to Upper Pangani tributary streams can be ruled out.

Assuming that the niche space available to the colonising *O. hunteri* ancestor was not filled by other (now locally extinct) fish species, the ancestral *O. hunteri* population probably expanded rapidly after this initial colonisation. If so, the genetic signature of this ancient population expansion in today's *O. hunteri* population should provide an estimate of the time passed since then. However, as neither the haplotype mismatch distribution nor the Bayesian skyline plot of the 15 *O. hunteri* specimens we sequenced (Fig. 6) reveal an unambiguous signal of rapid population expansion, we cannot conclude with certainty that such rapid population expansion has actually occurred. Although sample sizes similar to ours have allowed the detection of past population expansions in some studies (e.g. Genner & Turner, 2014), some authors suggest that sample sizes must be on the order of 20–40 (Drummond & Bouckaert, 2015) or even 50 (Grant, 2015) for this purpose. Estimating the approximate timing of the putative population expansion which followed the colonisation of Lake Chala by the ancestor of *O. hunteri,* using a coalescence approach, must hence await the sequencing of additional specimens.

The structure of the modern-day haplotype network of *O. hunteri*, in which 15 specimens yield seven haplotypes with up to five mutations between them, does suggest that this endemic population is relatively ancient, i.e. in line with the fossil evidence. The age of Lake Chala itself is estimated at approximately 250,000 years, based on the total depth of its sedimentary record as revealed by seismic reflection stratigraphy (Moernaut et al., 2010) relative to the radiocarbon-dated upper portion of this record (Verschuren et al., 2009; Blaauw et al., 2011). The HPD interval derived from our genetic data suggests the age of the oldest coalescence event within *O. hunteri* to range between 230,000 and 19,000 years. Although this is a rather wide bracket of time, its upper (older) end is consistent with the current best estimate of the age of Lake Chala, whereas its lower (younger) end is only a slight underestimation of the minimum age of the population based on fossil evidence. Given the modest number of sequences currently available, and therefore the potential for more distant haplotypes to remain undocumented at this time, we consider an early rather than late colonisation most plausible.

Acknowledgements This study was carried out under Memorandum of Understanding A14/TT/0923 between Ghent University and the National Museums of Kenya (NMK), and institutional affiliation of DV with NMK. We thank Caxton Oluseno and the fishermen of lakes Chala and Jipe for assistance in acquiring fish specimens for this study, and the staff of the NERC Biomolecular Analysis Facility and the Nosil lab, both at the University of Sheffield, for support. We further thank three anonymous reviewers for their suggestions to improve this manuscript. This research was sponsored by the Ghent University Special Research Fund through Collaborative Research Activity 'DeepCHALLA', including PhD support to JD. MM received support from the Swiss National Science Foundation, the University of Sheffield, and the International Continental Scientific Drilling Program. The Euler HPC cluster at ETH Zürich was used for phylogenetic analyses. We thank Michael Matschiner for advice on the program Fitchi. The mitochondrial CR and ND2 sequences produced by this study are accessible through GenBank, and vouchers of all sequenced fish specimens will be archived at NMK.

References

Agnèse, J., B. Adépo-Gourène & L. Pouyaud, 1997. Natural hybridization in tilapias. In Agnèse, J. F. (ed.), Genetics and Aquaculture in Africa. ORSTOM, Paris: 95–103.

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

Bailey, R. G., S. Churchfield, T. Petr & R. Pimm, 1978. The ecology of the fishes in Nyumba ya Mungu reservoir,

Tanzania. Biological Journal of the Linnean Society 10: 109–137.

Barluenga, M. & A. Meyer, 2010. Phylogeography, colonization and population history of the Midas cichlid species complex (*Amphilophus* spp.) in the Nicaraguan crater lakes. BMC Evolutionary Biology 10: 326.

Barluenga, M., K. N. Stölting, W. Salzburger, M. Muschick & A. Meyer, 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439: 719–723.

Beveridge, M. C. & B. McAndrew, 2000. Tilapias: Biology and Exploitation. Springer Science & Business Media, Dordrecht.

Blaauw, M., B. van Geel, I. Kristen, B. Plessen, A. Lyaruu, D. R. Engstrom, J. van der Plicht & D. Verschuren, 2011. High-resolution 14C dating of a 25,000-year lake-sediment record from equatorial East Africa. Quaternary Science Reviews 30: 3043–3059.

Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C. H. Wu, D. Xie, M. A. Suchard, A. Rambaut & A. J. Drummond, 2014. BEAST 2: a software platform for bayesian evolutionary analysis. PLoS Computational Biology 10: e1003537.

Crispo, E., J. S. Moore, J. A. Lee-Yaw, S. M. Gray & B. C. Haller, 2011. Broken barriers: human-induced changes to gene flow and introgression in animals. BioEssays 33: 508–518.

D'Amato, M. E., M. M. Esterhuyse, B. C. W. Van Der Waal, D. Brink & F. A. M. Volckaert, 2007. Hybridization and phylogeography of the Mozambique tilapia *Oreochromis mossambicus* in southern Africa evidenced by mitochondrial and microsatellite DNA genotyping. Conservation Genetics 8: 475–488.

Dadzie, S., R. D. Haller & E. Trewavas, 1988. A note on the fishes of Lake Jipe and Lake Chale on the Kenya–Tanzania border. Journal of East African Natural History 192: 46–51.

Darriba, D., G. L. Taboada, R. Doallo & D. Posada, 2012. JModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9: 772.

Deines, A. M., I. Bbole, C. Katongo, J. L. Feder & D. M. Lodge, 2014. Hybridisation between native *Oreochromis* species and introduced Nile tilapia *O. niloticus* in the Kafue River, Zambia. African Journal of Aquatic Science 39: 23–34.

Denny, P., 1978. Nyumba ya Mungu reservoir, Tanzania, the general features. Biological Journal of the Linnean Society 10: 5–28.

Dieleman, J., B. Van Bocxlaer, C. Manntschke, D. W. Nyingi, D. Adriaens & D. Verschuren, 2015. Tracing functional adaptation in African cichlid fishes through morphometric analysis of fossil teeth: exploring the methods. Hydrobiologia 755: 73–88.

Drummond, A. J. & R. R. Bouckaert, 2015. Bayesian Evolutionary Analysis with BEAST 2. Cambridge University Press, Cambridge.

Drummond, A. J., A. Rambaut, B. Shapiro & O. G. Pybus, 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Molecular Biology and Evolution 22: 1185–1192.

Eknath, A. E. & G. Hulata, 2009. Use and exchange of genetic resources of Nile tilapia (*Oreochromis niloticus*). Reviews in Aquaculture 1: 197–213.

Elmer, K. R., T. K. Lehtonen, S. Fan & A. Meyer, 2012. Crater lake colonization by neotropical cichlid fishes. Evolution 67: 281–288.

Excoffier, L., G. Laval & S. Schneider, 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1: 47–50.

Firmat, C., P. Alibert, M. Losseau, J. F. Baroiller & U. K. Schliewen, 2013. Successive invasion-mediated interspecific hybridizations and population structure in the endangered cichlid *Oreochromis mossambicus*. PLoS ONE 8: e63880.

Fryer, G. & T. D. Iles, 1972. The Cichlid Fishes of the Great Lakes of Africa, their Biology and Distribution. Oliver and Boyd, Edinburgh.

Genner, M. J. & G. F. Turner, 2014. Timing of population expansions within the Lake Malawi haplochromine cichlid fish radiation. Hydrobiologia 748: 121–132.

Genner, M. J., M. E. Knight, M. P. Haesler & G. F. Turner, 2010. Establishment and expansion of Lake Malawi rock fish populations after a dramatic Late Pleistocene lake level rise. Molecular Ecology 19: 170–182.

Grant, W. S., 2015. Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography. Journal of Heredity 106: 333–346.

Günther, A., 1889. On some fishes from the Kilima-Njaro district. Proceedings of the Scientific Meetings of the Zoological Society of London for the Year 1889: 70–72.

Håkansson, N. T., 2008. The decentralized landscape: regional wealth and the expansion of production in northern Tanzania before the eve of colonialism. In Cliggett, L. & C. A. Pool (eds.), Economies and the Transformation of Landscape. AltaMira Press, Creek: 239–266.

Hammer, Ø., D. A. T. Harper & P. D. Ryan, 2001. PAST: paleontological statistics software package for education and data analysis. Palaeontologia Electronica 4(1): 1–9.

Hasegawa, M., H. Kishino & T. A. Yano, 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22: 160–174.

Holzman, R. & C. D. Hulsey, 2017. Mechanical transgressive segregation and the rapid origin of trophic novelty. Scientific Reports 7: 40306.

Klett, V. & A. Meyer, 2002. What, if anything, is a Tilapia? Mitochondrial ND2 phylogeny of tilapiines and the evolution of parental care systems in the African cichlid fishes. Molecular Biology and Evolution 19: 865–883.

Kocher, T. D., J. A. Conroy, K. R. McKaye, J. R. Stauffer & S. F. Lockwood, 1995. Evolution of NADH dehydrogenase subunit 2 in east African cichlid fish. Molecular Phylogenetics and Evolution 4(4): 420–432.

Lee, W. J., J. Conroy, W. H. Howell & T. D. Kocher, 1995. Structure and evolution of teleost mitochondrial control regions. Journal of Molecular Evolution 41: 54–66.

Lowe, R. H., 1955. New species of Tilapia (Pisces, Cichlidae) from Lake Jipe and the Pangani River, East Africa. Bulletin of the British Museum (Natural History) Zoology 2(12): 349–368.

Lowe-McConnell, R. H., 1987. Ecological Studies in Tropical Fish Communities. Cambridge University Press, Cambridge.

Martinsen, G. D., T. G. Whitham, R. J. Turek & P. Keim, 2001. Hybrid populations selectively filter gene introgression between species. Evolution 55: 1325–1335.

Matschiner, M., 2016. Fitchi: haplotype genealogy graphs based on the Fitch algorithm. Bioinformatics 32: 1250–1252.

Meier, J. I., D. A. Marques, S. Mwaiko, C. E. Wagner, L. Excoffier & O. Seehausen, 2017. Ancient hybridization fuels rapid cichlid fish adaptive radiations. Nature Communications 8: 14363.

Meyer, A., J. M. Morrissey & M. Schartl, 1994. Recurrent origin of a sexually selected trait in *Xiphophorus* fishes inferred from a molecular phylogeny. Nature 368: 539–542.

Moernaut, J., D. Verschuren, F. Charlet, I. Kristen, M. Fagot & M. De Batist, 2010. The seismic-stratigraphic record of lake-level fluctuations in Lake Challa: hydrological stability and change in equatorial East Africa over the last 140 kyr. Earth and Planetary Science Letters 290: 214–223.

Nagl, S., H. Tichy, W. E. Mayer, I. E. Samonte, B. J. McAndrew & J. Klein, 2001. Classification and phylogenetic relationships of African tilapiine fishes inferred from mitochondrial DNA sequences. Molecular phylogenetics and evolution 20: 361–374.

Ndiwa, T. C., D. W. Nyingi & J. F. Agnese, 2014. An important natural genetic resource of *Oreochromis niloticus* (Linnaeus, 1758) threatened by aquaculture activities in Loboi Drainage, Kenya. PLoS ONE 9: e106972.

Nyingi, D. W. & J. F. Agnèse, 2007. Recent introgressive hybridization revealed by exclusive mtDNA transfer from *Oreochromis leucostictus* (Trewavas, 1933) to *Oreochromis niloticus* (Linnaeus, 1758) in Lake Baringo, Kenya. Journal of Fish Biology 70: 148–154.

Parnell, N. F., C. D. Hulsey & J. T. Streelman, 2012. The genetic basis of a complex functional system. Evolution 66: 3352–3366.

Pullin, R.S & R.H. Lowe-McConnell, 1982. The biology and culture of tilapias. Proceedings of the International Conference on the Biology and Culture of Tilapias, 2–5 September 1980 at the Study and Conference Center of the Rockefeller Foundation, Bellagio, (Vol. 7). WorldFish, 1982: 426–432.

R Development Core Team, 2016. R: A Language and Environment for Statistical Computing. Foundation for Statistical Computing, Vienna.

Rambaut, A., 2009. FigTree v1. 3.1: Tree Figure Drawing Tool. Website: http://tree.bio.ed.ac.uk/software/figtree.

Rognon, X. & R. Guyomard, 2003. Large extent of mitochondrial DNA transfer from *Oreochromis aureus* to *O. niloticus* in West Africa. Molecular Ecology 12: 435–445.

Rohlf, F. J., 2010. TpsDig2: Thin Plate Spline Digitise (Version 2.16). Stony Brook University, New York.

Salzburger, W., S. Baric & C. Sturmbauer, 2002. Speciation via introgressive hybridization in East African cichlids? Molecular Ecology 11: 619–625.

Seegers, L., L. De Vos & D. O. Okeyo, 2003. Annotated checklist of the freshwater fishes of Kenya (excluding the lacustrine haplochromines from Lake Victoria). Journal of East African Natural History 92: 11–47.

Sheets, H. D., 2008. IMP: Integrated Morphometrics Package. Department of Physics, Canisius College, Buffalo.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312–1313.

Stelkens, R. B. & O. Seehausen, 2009. Phenotypic divergence but not genetic distance predicts assortative mating among species of a cichlid fish radiation. Journal of Evolutionary Biology 22: 1679–1694.

Stelkens, R. B., K. A. Young & O. Seehausen, 2010. The accumulation of reproductive incompatibilities in African cichlid fish. Evolution 64: 617–633.

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

Verschuren, D., J. S. Sinninghe Damsté, J. Moernaut, I. Kristen, M. Blaauw, M. Fagot & G. H. Haug, 2009. Half-precessional dynamics of monsoon rainfall near the East African Equator. Nature 462: 637–641.

Wirtz, P., 1999. Mother species-father species: unidirectional hybridization in animals with female choice. Animal Behaviour 58: 1–12.

Wolff, C., I. Kristen-Jenny, G. Schettler, B. Plessen, H. Meyer, P. Dulski, R. Naumann, A. Brauer, D. Verschuren & G. H. Haug, 2014. Modern seasonality in Lake Challa (Kenya/Tanzania) and its sedimentary documentation in recent lake sediments. Limnology and Oceanography 59: 1621–1636.