

Repeated trans-watershed hybridization among haplochromine cichlids (Cichlidae) was triggered by Neogene landscape evolution

Julia Schwarzer, Ernst Roelof Swartz, Emmanuel Vreven, Jos Snoeks, Fenton Peter David Cotterill, Bernhard Misof and Ulrich Kurt Schliewen

Proc. R. Soc. B 2012 **279**, 4389-4398 first published online 5 September 2012
doi: 10.1098/rspb.2012.1667

Supplementary data

["Data Supplement"](#)

<http://rsjb.royalsocietypublishing.org/content/suppl/2012/09/04/rspb.2012.1667.DC1.html>

References

[This article cites 52 articles, 16 of which can be accessed free](#)

<http://rsjb.royalsocietypublishing.org/content/279/1746/4389.full.html#ref-list-1>

Subject collections

Articles on similar topics can be found in the following collections

[evolution](#) (1323 articles)

[molecular biology](#) (115 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

Repeated trans-watershed hybridization among haplochromine cichlids (Cichlidae) was triggered by Neogene landscape evolution

Julia Schwarzer^{1,2}, Ernst Roelof Swartz³, Emmanuel Vreven^{4,5},
Jos Snoeks^{4,5}, Fenton Peter David Cotterill⁶, Bernhard Misof²
and Ulrich Kurt Schlieven^{1,*}

¹Bavarian State Collection of Zoology, Münchhausenstrasse 21, 81247 München, Germany

²Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany

³South African Institute for Aquatic Biodiversity, Private Bag 1015, Grahamstown 6140, South Africa

⁴Royal Museum for Central Africa, Leuvensesteenweg 13, 3080 Tervuren, Belgium

⁵Laboratory of Biodiversity and Evolutionary Genomics, Ch. Deberiotstraat 32, 3000 Leuven, Belgium

⁶Africa Earth Observatory Network (AEON), Geocodynamics Research Hub, University of Stellenbosch, Private Bag X1, Matieland 7602, South Africa

The megadiverse haplochromine cichlid radiations of the East African lakes, famous examples of explosive speciation and adaptive radiation, are according to recent studies, introgressed by different riverine lineages. This study is based on the first comprehensive mitochondrial and nuclear DNA dataset from extensive sampling of riverine haplochromine cichlids. It includes species from the lower River Congo and Angolan (River Kwanza) drainages. Reconstruction of phylogenetic hypotheses revealed the paradox of clearly discordant phylogenetic signals. Closely related mtDNA haplotypes are distributed thousands of kilometres apart and across major African watersheds, whereas some neighbouring species carry drastically divergent mtDNA haplotypes. At shallow and deep phylogenetic layers, strong signals of hybridization are attributed to the complex Late Miocene/Early Pliocene palaeohistory of African rivers. Hybridization of multiple lineages across changing watersheds shaped each of the major haplochromine radiations in lakes Tanganyika, Victoria, Malawi and the Kalahari Palaeolakes, as well as a miniature species flock in the Congo basin (River Fwa). On the basis of our results, introgression occurred not only on a spatially restricted scale, but massively over almost the whole range of the haplochromine distribution. This provides an alternative view on the origin and exceptional high diversity of this enigmatic vertebrate group.

Keywords: East African cichlid radiations; haplochromines; hybridization; Neogene landscape evolution

1. INTRODUCTION

Cichlid fishes of the haplochromine lineage gave rise to one of the most spectacular vertebrate radiations on our planet, the megadiverse cichlid species flocks endemic to the East African Great Lakes and southern Africa [1,2]. Reconstructing robust phylogenetic relationships of haplochromine cichlids has proved to be difficult owing to limited taxon sampling and lack of phylogenetically informative characters. Until recently, phylogenetic analyses have delivered support for the monophyly of each of the major haplochromine species flocks, including the Lake Victoria superflock [3], southern African serranochromines and Tropheini of Lake Tanganyika [1,3–5]. The assumed monophyly of Lake Malawi haplochromines was recently falsified after inclusion of riverine haplochromines [6,7]. In general, available comprehensive analyses

of haplochromines including several important riverine lineages have relied exclusively on mitochondrial DNA (mtDNA) [8,9]. Multilocus nuclear datasets are available only for a subset of haplochromine taxa focusing on particular subgroups from the Lake Victoria region [10], Lake Malawi [6,7] or Lake Tanganyika [5]. In these analyses, riverine haplochromines inhabiting different regions of the Congo basin and Angola are represented by very few Congo basin taxa [10], whereas those in the Upper Kwanza basin (Angola) have yet to be represented at all. This is unfortunate because it has been assumed that Lake Tanganyika, on the edge of vast Congo basin, is the origin of haplochromine diversification [9]. Furthermore, several studies have identified selected Congolian haplochromines as sister taxon to ‘modern haplochromine’ sublineages [10], or as sistergroup to members of the southern African serranochromine species flock, which is assumed to have originated in the Kalahari Palaeolakes, previously referred to as the Makgadikgadi flock [1,8,9,11].

Other studies have provided phylogenetic evidence for massive introgression and hybridization among ancient lineages in evolving species flocks [12–15], but phylogenetic

* Author for correspondence (schlieven@zsm.mwn.de).

Ulrich Kurt Schlieven dedicates to Prof. W. Wickler on the occasion of his 80th birthday.

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2012.1667> or via <http://rspb.royalsocietypublishing.org>.

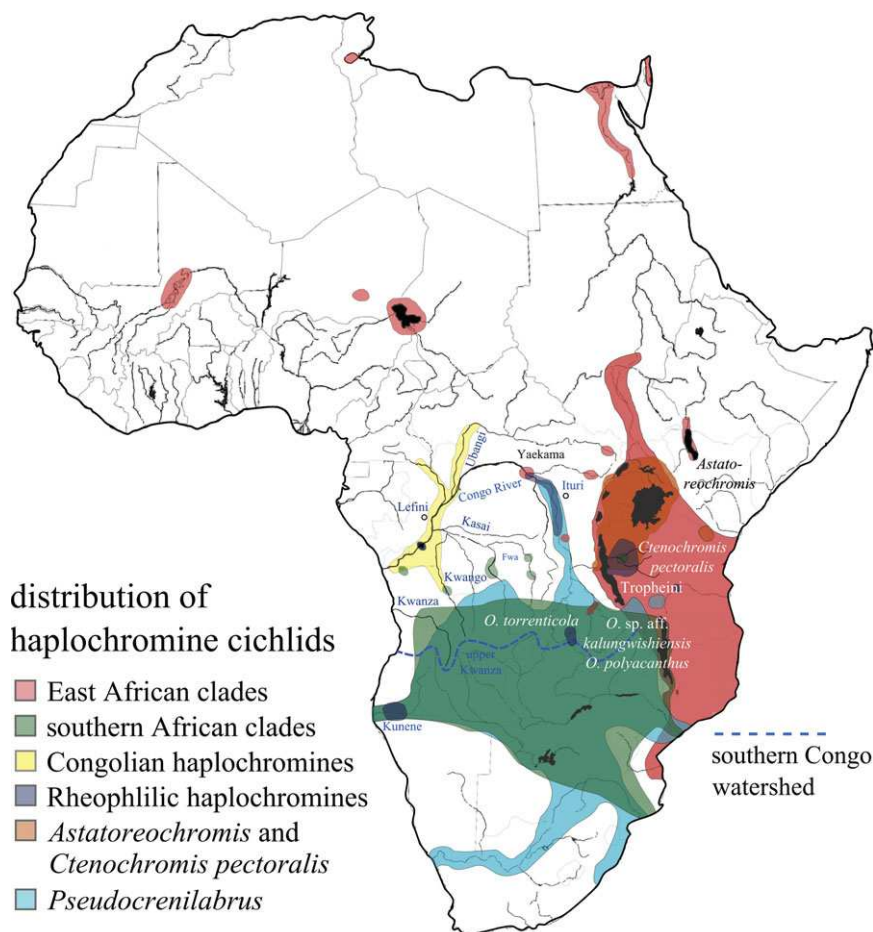


Figure 1. Distributions of the major groups of haplochromine cichlids. Colour codes correspond to distribution ranges. The blue dotted line marks the border of the southern Congo watershed.

tests for a potential contribution of hybridization to the evolution of haplochromine lineages remain scarce ([7], but see [10,16]). Here, our extensive taxon sampling focuses on riverine haplochromines. It includes key Angolan and most haplochromine species from the Congo basin. Phylogenetic hypotheses are based on more than 2000 amplified fragment length polymorphism (AFLP) loci as well as two mitochondrial genes. Critically, following Seehausen [17], an experimental approach enabled us to recover otherwise obscured phylogenetic signals in the reticulate dataset. It allowed us not only to decipher the phylogeographic origin of novel riverine hybrid lineages, which had formed after secondary contact of previously isolated drainages, but also to estimate their contribution to the origin and diversification of megadiverse lacustrine cichlid species flocks.

2. MATERIAL AND METHODS

Sampling focused on an extensive and representative coverage of all major haplochromine lineages (see the electronic supplementary material, table S1) and biogeographic regions (figure 1). *Tilapia bilineata* as well as *Lamprologus* sp., were chosen as outgroups based on results in Schwarzer *et al.* [18]. The mitochondrial ND2 and part of the cytochrome *b* gene (Cytb) were amplified and sequenced for 67 taxa (48 species), using primers MET and TRP [19] and L14724 H15149 [20]. AFLP genotypes were obtained for the same species ($n = 68$, see the electronic supplementary material, table S1). Peaks between 100 and 499 bases were scored unambiguously for presence/absence. The analysis was conducted automatically

using GENEMAPPER v. 4.0. Eight individuals were genotyped twice to test for reproducibility. The error rate per individual was calculated as the ratio between observed number of differences and the total number of scored fragments [21], resulting in a mean error rate of 0.03. Sequence data of mitochondrial Cytb and ND2 genes and the AFLP matrix have been deposited in GenBank under the accession numbers JX156995–JX157126 and in the Dryad data repository (doi:10.5061/dryad.72h4m).

(a) Phylogenetic inference

Sequence alignment was conducted using the CLUSTALW [22] algorithm implemented in BIOEDIT v. 7.0.4.1. [23]. Identification of ambiguous alignment positions was carried out using ALISCOPE v. 0.2 under default settings [24], and identified positions were manually deleted. We used a partition separating first and second codon positions from the third. The GTR + Γ model best-fitted variability in codon positions one and two, whereas the HKY + Γ model was assigned to third codon positions based on results from the Bayes factor test [25]. Bayesian analyses were performed using MRBAYES v. 3.1.2 [26] including two parallel runs each having 10^6 generations starting with random trees and sampling of trees every 1000 generations. To ensure convergence, the first 10^5 generations of each run were treated as burn-in and excluded. The remaining trees from all Bayesian analyses were used to build a 50 per cent majority rule consensus tree. A maximum-likelihood (ML) phylogenetic analysis was conducted using RAXML v. 7.0.3 [27] using the GTR + Γ model and the rapid bootstrap algorithm

with following search for the best-scoring ML tree. Branch support was evaluated based on non-parametric bootstrapping (BS) consisting of 1000 pseudoreplicates. For the AFLP data, a neighbour joining tree was calculated using TREECON v. 1.3 based on Link *et al.*'s [28] algorithm that takes shared fragments into account and ignores shared absent bands [28]. BS values were calculated based on 1000 pseudoreplicates.

(b) Inferring hybrid signal

Following Seehausen [17], we applied a tree-based method to test for homoplasy excess in our dataset. The expectation is that the inclusion of hybrid taxa increases the conflict in the dataset and reduces support values for affected nodes in a phylogenetic tree more strongly than the inclusion of non-hybrid taxa [17]. The exclusion of a hybrid taxon from the dataset should therefore increase support values only for affected nodes. This detection of potential hybrid signal focuses on the AFLP dataset, as hybridization cannot be reliably detected in maternally inherited mitochondrial markers [29]. All clades showing discordant signal between the nuclear (nc) and mitochondrial (mt) trees as well as all monophyletic clades (in the ncDNA tree) were successively removed from the dataset resulting in 86 removal experiments (see the electronic supplementary material, figure S1). Subsequently, distance trees were built for each reduced dataset with 500 bootstrap replicates using TREECON v. 1.3. The resulting trees and BS support values were checked manually for all remaining clades. Results of the homoplasy excess test were visualized in boxplots for major phylogenetic nodes with initially low BS support values. To test for random effects on BS support, additional removal experiments were conducted with a certain number of randomly chosen taxa. The number of excluded taxa depended on the number of individuals that caused an effect on node support and ranged from $n = 1$ to 6. For each n , the random removals were repeated 100 times. Tree generation and BS support evaluation was conducted as described earlier. A heatmap based on BS outliers was generated representing the change in BS support values for all removal experiments over the whole dataset. Outliers were defined following Tukey [30], as data points located outside the 1.5* inter-quartile distance displayed in boxplots as whiskers.

(c) Dating and diversification rates

On the basis of the mtDNA dataset a relaxed-clock, Bayesian approach implemented in BEAST v. 1.6.2 [31] was applied. Uniform priors were set for the split of lamprologines from haplochromines based on the 95% confidence intervals obtained by Schwarzer *et al.* [18]. Constraints were set from 10.6 to 20.4 Ma in run A (including the fossil prior *Oreochromis lorenzoi*† [32]), and from 12.8 to 28.4 Ma in run B (without the fossil prior). The Bayesian tree was used as starting tree, the Yule model selected as tree prior and an uncorrelated lognormal model was applied to estimate rate variation along branches. The analysis was run for 10 million generations, and the effective sample size was checked using TRACER v. 1.4. [33].

3. RESULTS

The final AFLP dataset comprised 68 taxa with 2106 AFLP loci. Of these, 1984 (1889 without outgroups) fragments were polymorphic. The ND2 dataset consisted of

1022 bp and the Cytb dataset of 405 bp (total = 1427 bp) with 399 (ND2) and 130 (Cytb) variable sites and empirical base frequencies of A = 0.26, C = 0.35, G = 0.12, T = 0.27 and A = 0.24, C = 0.30, G = 0.17, T = 0.29, respectively. The mean sequence divergence of the mitochondrial dataset was 0.086 ± 0.029 . '*Haplochromis*' *snoeksi* failed to amplify for one AFLP primer combination (ACT-CAT*), the missing data were coded as undefined character states. To legitimate this approach, the AFLP analysis was also conducted with a reduced dataset of 11 primer-combinations, showing no topological differences (data not shown).

(a) Phylogenetic hypotheses based on AFLP data and mtDNA

Analyses of AFLPs and mitochondrial genes yielded statistically highly supported phylogenetic hypotheses. In the AFLP dataset, 11 major clades are discernable (figure 2) reflecting a mostly congruent biogeographic division into an eastern, a Congolian and a southern group. '*Haplochromis*' cf. *bakongo* and '*H.*' *snoeksi* from lower River Congo tributaries, however, appear closer to the southern clade (figure 2, node E), rendering the Congolian clade paraphyletic. On the basis of the AFLP dataset, the single-included *Pseudocrenilabrus* captures a position as sistergroup to all remaining haplochromines (figure 2, BS = 99). Several, but not all of the rheophilic haplochromines currently assigned to *Orthochromis*, i.e. *O.* cf. *stormsi* 'Kinshasa', *O. stormsi* 'Kisangani', *O. polyacanthus* and *O.* sp. aff. *kalungwishiensis*, form sistergroup to the Tropheini from Lake Tanganyika, the East African clades, the Congolian clade, '*H.*' *snoeksi* and '*H.*' cf. *bakongo* and the southern clades (figure 2, node H, BS = 78). Members of the East African clades (figure 2, BS = 92) appear as sistergroup to the Tropheini (figure 2, BS = 99 and node F, BS = 62). On the basis of the ncDNA dataset, *Astatoreochromis alluaudi* and '*H.*' *burtoni* capture a sistergroup position to the remaining East African clades (BS = 92, BS = 100). On the basis of the mtDNA tree, however, *A. alluaudi* appears together with *Pseudocrenilabrus multicolor* and *O.* sp. aff. *kalungwishiensis* unresolved at the base of the East African clades and Tropheini (BS < 50), and '*H.*' *burtoni* appears within the Lake Victoria superflock (BS/Bayesian posterior probability (BPP) = 51/0.64). Lake Malawi haplochromines form a monophyletic group in both trees (BS = 100, BS/BPP = 81/0.93). Interestingly, the recently discovered '*H.*' sp. 'Yaekama' falls into the Lake Victoria superflock clade based on both trees (figure 2, BS = 100 and 100/1.0) even though it is distributed in the northeastern River Congo drainage (near Kisangani). The East African clades and Tropheini (figure 2, node F, BS = 62) appear as sistergroup to the Congolian clade (BS = 100), '*H.*' *snoeksi* and '*H.*' cf. *bakongo* (BS = 76) and the southern clades (BS = 60) based on the ncDNA dataset. The integrity of the southern clades is only weakly supported, as is the sistergroup relationship of '*H.*' cf. *bakongo*/*H.*' *snoeksi* and the southern African clades (figure 2, node E, BS = 60). Within the Congolian clade, the Pool Malebo and central Congo basin '*H.*' *polli* and '*H.*' *oligacanthus* and the three lower River Congo rapids species '*H.*' *fasciatus*, '*H.*' *demeusii* and '*H.*' sp. 'Sanzikwa' appear monophyletic (BS = 100). Within the southern clades, all included species from rivers Fwa, Kasai and Kwango (all Kasai

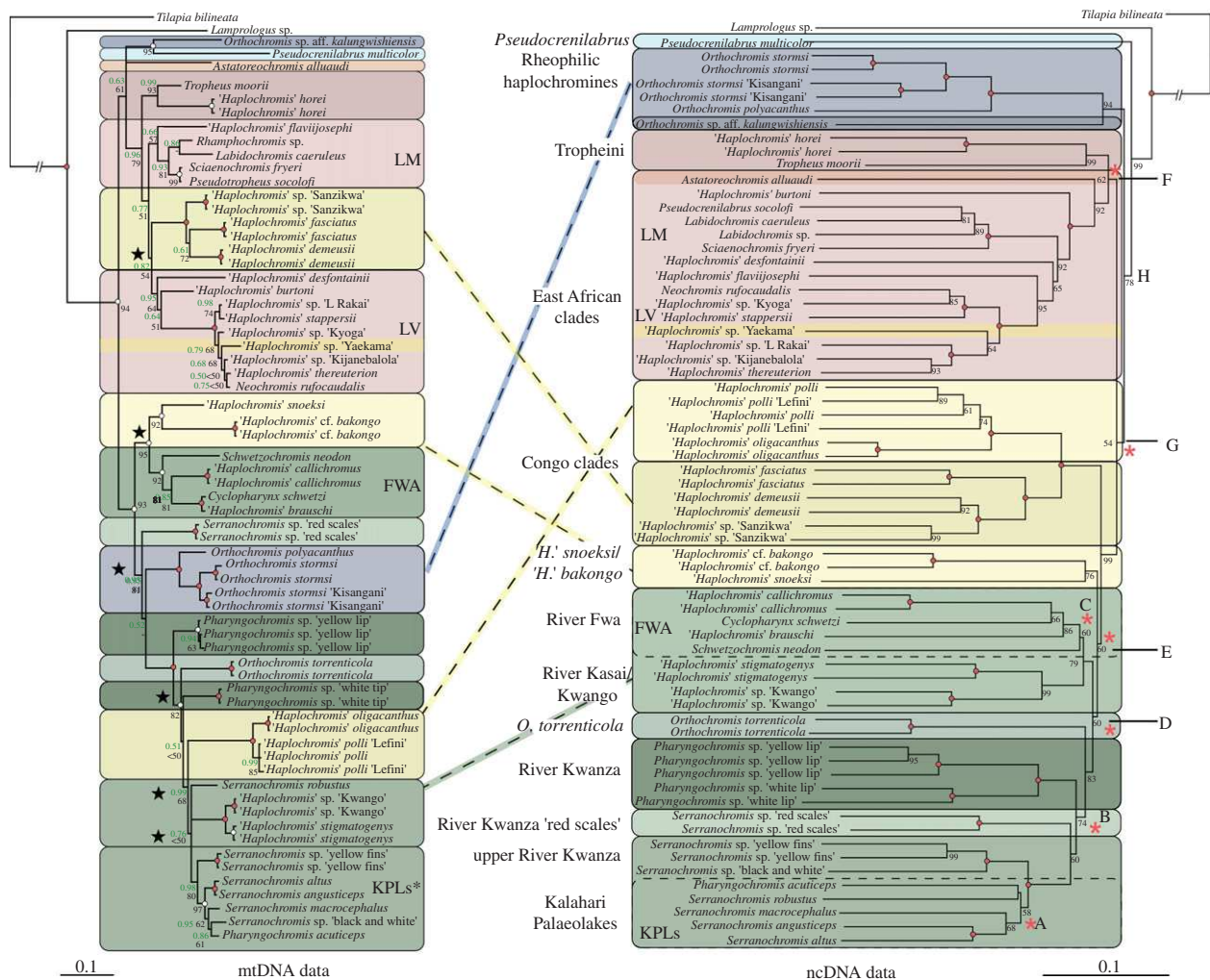


Figure 2. Phylogenies based on mtDNA and AFLP data. The datasets comprise mitochondrial sequences of ND2 and cytochrome *b* and more than 2000 AFLP markers. Black numbers on nodes refer to bootstrap values (BS, 1000 pseudo-replicates) of the ML run (left side) or the neighbour joining tree (right side). Green numbers refer to posterior probabilities from the Bayesian inference (BPP). Filled red circles represent 100% BS (right side) or 100/1.0 BS- and BPP values (left side). Empty circles on the mtDNA tree indicate 1.0 BPP. Major groups within the phylogeny are marked with coloured frames. The colours correspond to those used in figure 1. Clades that are placed differently in AFLP and mtDNA trees are connected by dotted lines. Red asterisks in the AFLP tree indicate nodes with initially low BS values caused by homoplasy excess. Black stars in the mtDNA tree mark nodes with age estimates provided in the electronic supplementary material, table S2. LM, Laka Malawi; LV, Lake Victoria super flock; FWA, River Fwa; KPL, Kalahari Palaeolakes; KPLs* Kalahari Palaeolakes not monophyletic.

drainage, Congo basin) form a monophyletic group (BS = 79) that is sistergroup (BS = 60) to species from the Angolan River Kwanza system and another rheophilic species, *O. torrenticola* (figure 2, BS = 83). Except for the smaller *O. torrenticola* and the *Serranochromis* sp. 'red scales' clades, none of the major clades supported by the ncDNA phylogeny is recovered in the tree based on mtDNA (figure 2).

Two well-supported, but geographically heterogeneous clades are recovered in the mtDNA phylogeny. One is composed of members of the East African clades and '*H. fasciatus*'/'*H. demeusii*' and '*H.* sp. 'Sanzikwa' from the lower River Congo (BS/BPP = 77/0.51) with the Tropheini as sistergroup (BS/BPP = 79/0.96) and the other one of '*H. snoeksi*'/'*H.* cf. *bakongo*' and species from River Fwa (BS/BPP = 95/1.0) forming the sistergroup to the Angolan *Serranochromis* sp. 'red scales', *O. stormsi*/*O.* cf. *stormsi* 'Kisangani' and *O. polyacanthus*, members of the southern clades and the Congolian clade composed of '*H. polli*' and '*H. oligacanthus*' (BS/BPP = 98/1.0). '*Haplochromis*' *polli* and '*H.* *oligacanthus*' appear as

sistergroup to species from southern African rivers Kwango, Kasai, upper River Kwanza, and the serranochromine radiation in the Kalahari Palaeolakes (BS/BPP = 68/0.99, figure 2).

(b) Dating and diversification rates

Owing to extensive reticulate signal in the dataset, age estimates based on mtDNA do not serve as estimator of the actual age of haplochromine radiations or species flocks, but rather represent the age of mitochondrial diversity and the timing of mitochondrial introgression events. Age estimates for the five major introgression events within the haplochromines (indicated by dotted lines in figure 2) ranged between 2.4 and 11.33 Ma, or between 2.8 and 15.2 Ma depending on calibration priors (with or without fossil prior; electronic supplementary material, table S2). The obtained time frames largely matched those from previously published studies, except for the age of the Lake Victoria super flock, which appears older

in our dataset (see the electronic supplementary material, table S2). The lack of calibration points for terminal nodes and the limited taxon sampling for the Lake Victoria superflock probably caused this uncertainty [34].

(c) *Cytonuclear discordance and homoplasy excess test*

Cytonuclear discordances indicating hybridization events are present throughout the whole haplochromine phylogeny. Major discrepancies between the mtDNA and AFLP (ncDNA) phylogenetic hypotheses are as follows. (i) The Congolian species (without '*H.*' *snoeksi* and '*H.*' cf. *bakongo*) appear monophyletic in the tree based on ncDNA (figure 2), but based on the mtDNA data, the two Congolian subclades ('*H.*' *pollii*/*H.*' *oligacanthus* and '*H.*' *fasciatus*/*H.*' *demeusii*/*H.*' sp. 'Sanzikwa') are deeply nested within southern clade species or East African clades, as sistergroup to members of the Lake Victoria superflock. '*Haplochromis*' *snoeksi* and '*H.*' cf. *bakongo* form a sistergroup to the southern clades based on ncDNA data, but in the tree based on mtDNA, they are sistergroup to species from the River Fwa (*Schwetochromis* *neodon*, '*H.*' *callichromus*, '*H.*' *brauschi* and *Cyclopharynx* *schwetzi*; figure 2). Within the River Fwa species, '*H.*' *brauschi* appears as sister to *C. schwetzi* and '*H.*' *callichromus* based on the ncDNA data (BS = 86), but to *C. schwetzi* based on mtDNA data (BS/BPP = 100/1.0). (ii) The rheophilic haplochromines *O. stormsi* 'Kinshasa'/*O. stormsi* 'Kisangani', *O. polyacanthus* and *O. sp. aff. kalungwishiensis* are based on the ncDNA dataset monophyletic (BS = 94) but *O. stormsi* 'Kinshasa'/*O. stormsi* 'Kisangani' and *O. polyacanthus* are nested within a clade of southern species and the '*H.*' *pollii*/*H.*' *oligacanthus* clade (BS/BPP = 81/0.95) and *O. sp. aff. kalungwishiensis* appears as sistergroup to *P. multicolor* based on the mtDNA data (BS/BPP = 95/1.0). (iii) *Astatoreochromis* *alluaudi* forms the sistergroup to the remaining species from the East African clades based on ncDNA data (BS = 92), but remains unresolved at a basal position based on the mtDNA dataset. Further cytonuclear discordances within the East African clades are present concerning the position of '*Haplochromis*' *flavijosephi* and '*H.*' *burtoni*. '*Haplochromis*' *flavijosephi* is sister to members of the Lake Victoria superflock based on ncDNA (BS = 95), but sister to the Lake Malawi clade based on the mtDNA dataset (BS/BPP = 66/0.57). '*Haplochromis*' *burtoni* appears as sister to the remaining East African clades (without *A. alluaudi*) based on the ncDNA (BS = 92), but is nested in a clade composed of the Lake Victoria superflock and '*H.*' *desfontainii* based on the mtDNA dataset (BS/BPP = 64/0.95). (iv) Within the southern clade, several discordances are obvious concerning within-group relationships (figure 2). *Serranochromis* *robustus* clusters within the upper River Kwanza/Kalahari Palaeolakes clade based on ncDNA data (BS = 100), but not so based on mtDNA data (figure 2), where it appears in a clade with rivers Kwango/Kasai and River Kwanza/Kalahari Palaeolakes species (BS/BPP = < 50/0.76).

In 86 removal experiments, effects on different BS values are evident across the whole AFLP haplochromine phylogeny (figures 3 and 4). Strong effects with more than 50 per cent increase or decrease based on the mean BS value exist concerning the support values of

five nodes (9, C, D, 70 and F, figure 3; see the electronic supplementary material, figure S1) and medium effects (25% increase or decrease of BS) are present on six additional nodes (A, 8, B, E, 42 and G, figure 3; see the electronic supplementary material, figure S1). Removals causing an increase in BS values (indicating a decrease of homoplastic signal in the dataset) are caused mainly by members of the following subclades: upper River Kwanza/Kalahari Palaeolakes, *S. sp.* 'red scales', *O. torrenticola*, '*H.*' *snoeksi*/*H.*' cf. *bakongo* and River Congo clades, Tropheini, Lake Victoria superflock, riverine haplochromines and *P. multicolor* (figures 2 and 3). Boxplots were generated for BS values of nodes with an initially low support in the AFLP tree (indicated by letters A–H in figure 2). The node support for the Kalahari Palaeolakes species flock increases when two species from the flock (*Serranochromis* *altus* and *S. angusticeps*) are removed from the dataset (node A, BS = 58–86, figure 4). Support for node B comprising species from the Angolian River Kwanza and Kalahari Palaeolakes is increasing when *O. torrenticola* or *S. sp.* 'red scales' are removed (node B, BS = 74–96 or 97, respectively, figure 4). Effects on node support values of species from River Fwa are present when '*H.*' *brauschi* or the sistergroup to the River Fwa clade from rivers Kasai and Kwango is removed (node C, BS = 60–90 or 95, figure 4). An exclusion of '*H.*' cf. *bakongo* and '*H.*' *snoeksi* entails an increase in BS for node D comprising the southern African clades (BS = 60–89). Exclusion of '*H.*' *pollii* and '*H.*' *oligacanthus* increases the BS of node E, comprising the southern African clades and '*H.*' cf. *bakongo*/*H.*' *snoeksi* (BS = 60–81), and the removal of the Tropheini member *T. moorii* (but not '*H.*' *horei*) and of the rheophilic clade leads to an increase in BS for the East African clades and Tropheini (node F, BS = 62–80 or 90, figure 4). The BS for the node comprising all haplochromines excluding the rheophilic haplochromines and *Pseudocrenilabrus* increases, when all rheophilic haplochromines, *O. sp. aff. kalungwishiensis* or *P. multicolor* are removed (node G, BS = 54–82, 96 or 99 respectively, figure 4).

4. DISCUSSION

Recent molecular studies indicate significant introgression, through hybridization, of riverine haplochromines into the Lake Malawi species flock [6,7]. The inclusion of additional riverine haplochromines covering almost the whole range of their distribution in this study, however, highlights for the first time, to our knowledge, to which large extent hybridization has shaped the evolution of haplochromines. Nuclear data, based on more than 2000 AFLP markers, reflect close relationships of geographically adjacent haplochromine species and clades (figures 1 and 2), whereas the mtDNA data yield clearly conflicting phylogenetic signal. Combining phylogenetic information of both datasets, each reflecting different parts of evolutionary history, with analytical approaches targeted to unravel hybridization has proved necessary to decipher the complex evolution of haplochromine cichlids [5–7]. Theoretically, cytonuclear discordance can also be explained by ancient shared polymorphisms as a result of incomplete lineage sorting [17]. A strong argument against incomplete lineage sorting in the present case is the unequal spatial distribution of well-separated mtDNA haplotypes

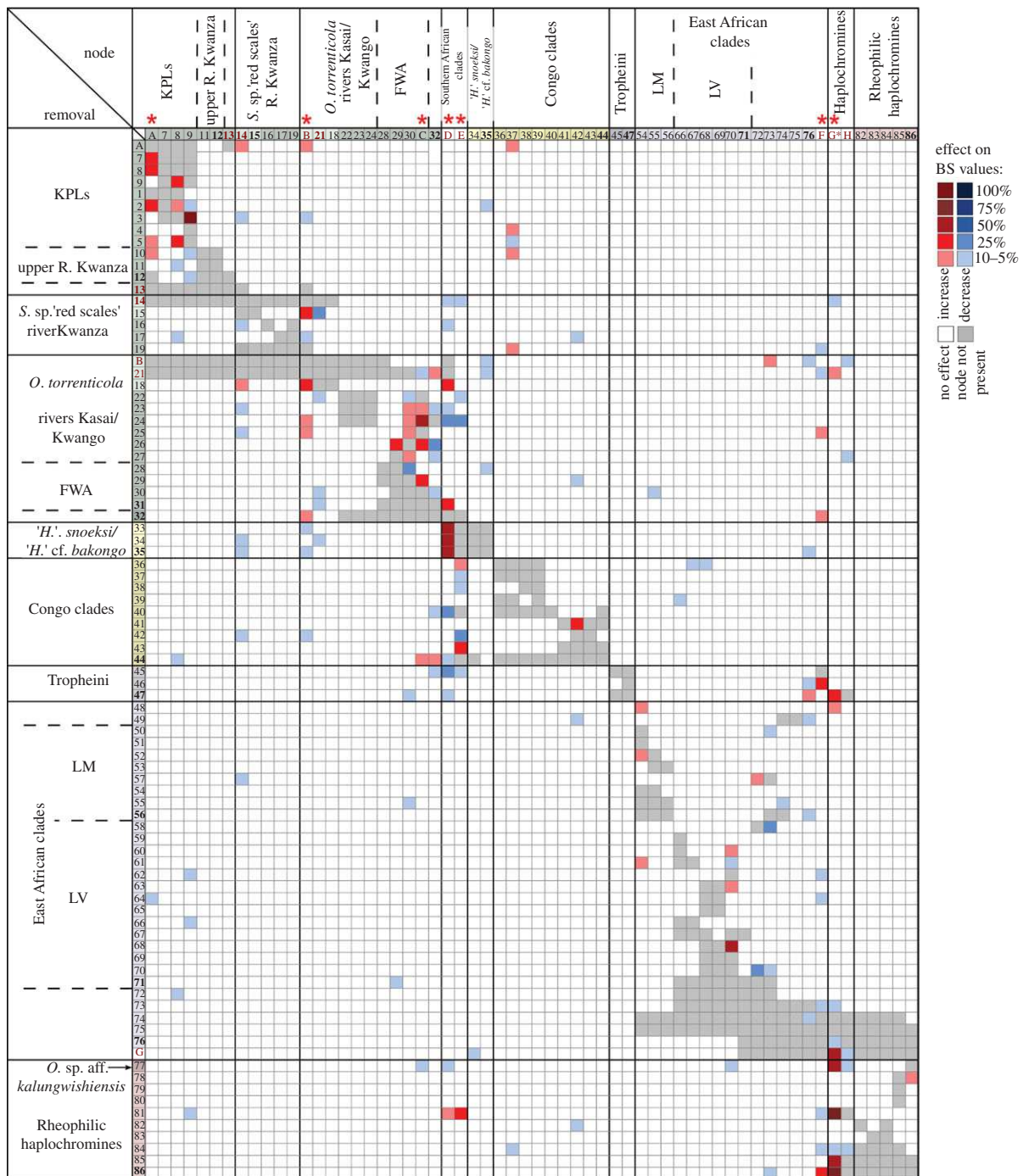


Figure 3. Heatmap of all removal experiments representing changes in BS values. A heatmap based on bootstrap outliers was generated representing the change in bootstrap support values for all removal experiments over the whole dataset. Outliers were defined as data points located outside the $1.5 \times$ inter-quartile distance displayed (in boxplots) as whiskers. Each clade and each single species was removed iteratively from the dataset (resulting in $n = 86$ experiments, see the electronic supplementary material, figure S1). Subsequently neighbour joining trees based on 500 bootstrap replicates were recalculated using TREECON v. 1.3. Stars mark nodes highlighted in figures 2 and 4. Numbers correspond to node names and removal experiments specified in the electronic supplementary material, figure S1.

and evidence from removal experiments based on AFLP data, showing that homoplasy excess is induced by single species or clades and not randomly (figures 3 and 4).

(a) Gene flow patterns can be explained through landscape evolution

The age of haplochromine cichlid diversification is a highly debated topic [8,35,36]. The lack of reliable fossil and geological calibration points and the high amount of reticulate signal within the haplochromines

make it difficult to calculate robust age estimates. Here, we have used constraints lying outside the haplochromine cichlid radiations [18] to estimate the timing of mitochondrial introgression events within the radiation. All major introgression events were constrained in the Late Miocene/Early Pliocene (2.4–11.3 Myr ago or 2.8–15.2 Myr ago, depending on calibration priors; see the electronic supplementary material, table S2). They coincided with pronounced Neogene tectonics that reshaped drainage systems across Central and East

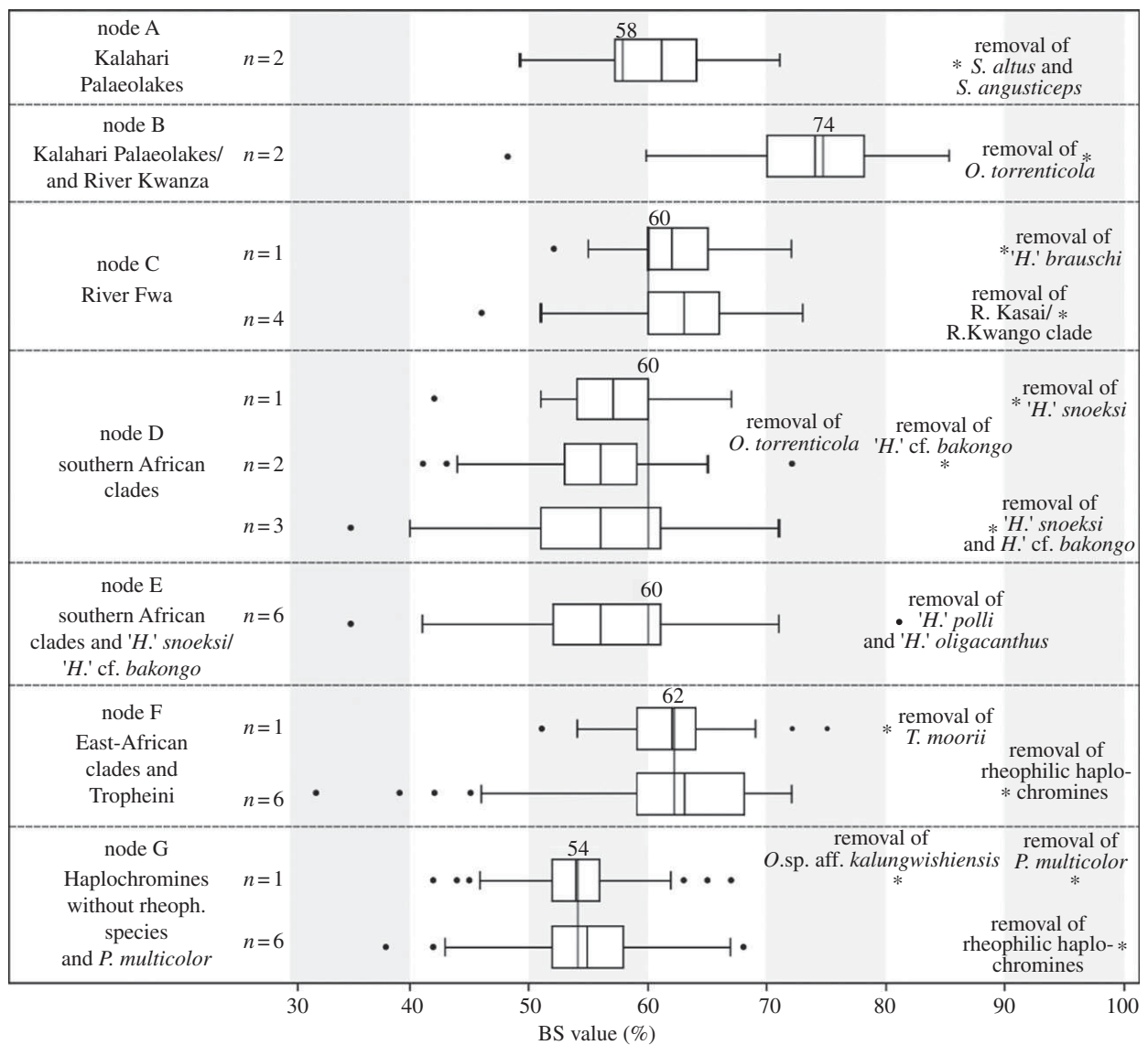


Figure 4. Homoplasmy excess tests of nodes with low BS values in the AFLP tree. The boxplots show the distribution of bootstrap support values (%) for selected basal nodes with initially low BS values (values in red). Letters correspond to nodes indicated in the AFLP phylogenetic tree in figure 2. Zero-distributions of BS values (derived from 100 randomly conducted removals) are shown as empty boxes. The random removals were conducted with a certain number of randomly chosen taxa (from $n = 1$ to 6) depending on the number of individuals that caused an effect on node support. For each n , the random removals were repeated 100 times.

Africa, with focal impacts across the northern Kalahari Plateau, especially the Congo–Zambezi watershed and the Eastern African Rift System. This episode of landscape evolution across the core area of ‘explosive’ haplochromine diversification, including the Albertine Rift [37–39], may be of focal relevance to understand the evolution of the lacustrine species flocks. New links between ancient watersheds connected previously isolated cichlid lineages, as tectonic activity caused river captures across the region, e.g. the Chambeshi/upper Kafue disruption [37]. Our results identify several rivers as foci of drainage evolution, specifically river captures, previously not recorded for the Kalahari Plateau [37–39]. Phylogeographic signatures of ancient hybridization testify to these geomorphic events, and shed new light on the role of landscape evolution as one determinant of haplochromine megadiversity: (i) phylogenetically closely related mtDNA haplotypes are shared across present day separated drainage systems, i.e. across upper Kwanza–Okavango and Congo–Zambezi watersheds (figure 2); (ii) close

phylogenetic relationships based on mtDNA are evident among haplotypes from ‘*Haplochromis*’ species of rivers Fwa, Inkisi and Kwilu; and (iii) putative hybrid taxa linking the southern clade with the Congolian lineage formed at geographically intermediate locations in the Lufira (*O. torrenticola*) and the upper Lucalla River (*Serranochromis* sp. ‘red scales’).

Our results point to landscape dynamics as the prominent control of cichlid evolution. These events that reshaped drainage topology provided multiple opportunities for accumulation of lineage specific alleles leading to phenotypic divergence, and repeated reconnection of discrete lineages in a dynamically changing landscape. It appears these events have led to substantially increased levels of novel phenotypes in haplochromine species flocks inhabiting the southern Congo basin and the northern Kalahari plateau. The molecular dates on nodes in this expanded haplochromine phylogeny have the potential to complement recent geobiological studies [37,39,40]. Generally, node age estimates of the

evolutionary events in aquatic biota set constraints on when formative events reshaped river topologies. Specifically, the here presented geobiological evidence for cichlid evolution sets hitherto unavailable chronological constraints—in the Late Neogene—on when tectonism reshaped the Southern Equatorial divide.

(b) *Long-distance dispersal of East African haplotypes*

A sistergroup relationship of mtDNA haplotypes is evident between East African haplochromines and a small group of narrow endemics, today confined within the lower River Congo rapids (figure 2). Interestingly, ‘*Haplochromis*’ species of the lower River Congo rapids are parapatrically distributed, corresponding to a trisection of the river stretch [40] in upstream (‘*H. polli*’), central (‘*H. demeusii*’) and downstream (‘*H. fasciatus*’) reaches of the lower River Congo. The distribution of ‘*H. polli*’ expands further upstream into the central Congo basin (figure 1). Surprisingly, ‘*H. fasciatus*’/‘*H. demeusii*’ carry haplotypes closely related to those of the East African haplochromines, whereas mt-haplotypes of their directly neighbouring upstream relatives (according to ncDNA data) ‘*H. polli*’ and ‘*H. oligacanthus*’ are rather distinct (figure 2). There is no trace of introgression (indicated by homoplasmy excess) detectable in the ncDNA dataset (figure 3). A complete replacement of mtDNA haplotypes can occur without any evidence of nuclear introgression as shown for salmonids, mountain hare, green pond frogs as well as for cichlids [41–45]. The mtDNA haplotype can become fixed by chance (drift) or by positive selection [46]. As there are no obvious indications for a selective advantage, potential drift through spatial isolation of the downstream lower River Congo species, as indicated by Schwarzer *et al.* [40], can explain the retention of the ancient, eastern mtDNA haplotypes. A connection of the River Congo with eastern African drainage systems (including the Lake Victoria basin and other rift lakes) is indicated by the recently detected occurrence of ‘modern’ haplochromines in the eastern River Congo drainage (J. Schwarzer & U. K. Schliewen 2009, personal observation): ‘*H.*’ sp. ‘Yaekama’, from the upper River Congo near Kisangani, for example, carries ocellated egg-spots typical of East African ‘modern’ haplochromines and clusters with members of the Lake Victoria superflock based on our mtDNA and ncDNA datasets (figure 2). Interestingly, Verheyen *et al.* [3] identified the mitochondrial sistergroup to the remaining Lake Victoria region modern haplochromines in Lake Kivu, a lake that is geographically intermediate between the Lake Victoria region and the central River Congo area around Yaekama. This ichthyogeographical connection is supported by closely related non-cichlid fish species occurring both on the east and on the west side of the northern Albertine Rift, i.e. poeciliids of the genus *Hypsopanchax* [47], and points to an area of connection between the Nile and the Congo basin drainages around the upper Ituri drainage (a northeastern River Congo affluent) and the Nilotic Lake Kivu/Lake Edward region [48].

(c) *Lacustrine origin of haplochromine species diversity?*

Hybridization between distantly related lineages can increase genetic and phenotypic potential, and thus

favour the onset of rapid adaptive radiations [17]. Empirical evidence comes from different studies in animals and plants [49–51], and initial hybridization is hypothesized to have shaped the cichlid radiations of the Kalahari Palaeolakes [1] and Lake Malawi [7]. A close relationship of the geographically proximate upper River Kwanza species and the Kalahari Palaeolakes haplochromines is supported by our data (figure 2). The occurrence of closely related mtDNA haplotypes of upper River Kwanza and Kalahari Palaeolakes haplochromines (figure 2), as well as excess homoplasmy in the Palaeolakes species flock (figure 4), supports the previously hypothesized hybrid origin of this young lake radiation [1]. Our data further suggest that lineages of the upper River Kwanza, upper River Zambezi and River Congo contributed to the hybrid swarm origin of the proposed Kalahari Palaeolakes radiation [1].

River Fwa, represented by four of five existing species in our dataset, appears to harbour a monophyletic species flock (though with low BS support). On the basis of ncDNA, the species of River Fwa are closely related to neighbouring species in rivers Kwango and Kasai (figure 2), but based on mtDNA they appear closely related to ‘*H.*’ cf. *bakongo* and ‘*H.*’ *snoeksi* from the lower River Congo system. Controlling for reticulation by removing species from rivers Kasai, Kwango and Fwa unequivocally increased BS for River Fwa monophyly, especially when removing the River Fwa ‘*H.*’ *brauschi* (figure 4, node C). Such evidence for reticulate (‘homoplastic’) genetic signals reveals how the hybrid origin of the radiation was seeded by distantly related Congolian lineages that thereafter became geographically isolated.

Our results clearly show that hybridization among both ancient (*Orthochromis*, *Pseudocrenilabrus* and *Astatoreochromis*) and recent (lower River Congo ‘*Haplochromis*’) riverine lineages qualifies as a fundamental event in the evolution of haplochromine radiations. This does not refute the assumed lacustrine origin of the East African clades, but underscores a much more complex history for the megadiverse haplochromine radiations. It is important to acknowledge how alternating wet and dry climates over the Late Cenozoic affected the entire East African region, as shown from sediment cores of its lakes [52–55]. These events probably complemented more widespread impacts of the Neogene tectonism across southern and east Africa, where uplift of the Kalahari Plateau [37,38,56] transfigured the drainage, exemplified in the incision of the lower River Congo rapids into Africa’s western margin [40,57]. The critical consequence of these impacts on upper River Congo affluents was to reshape haplochromine distributions, thereby facilitating extensive hybridization between previously separated lineages.

5. CONCLUSION

Hybridization among even geographically distant haplochromine lineages has had a major influence on the evolution of haplochromines. Our results demonstrate trans-watershed dispersal of mtDNA haplotypes present within haplochromines as well as strong, multiple signals of introgression and potential hybrid speciation, thereby questioning simplistic assumptions about the evolution of some of the major haplochromine lineages. Our results ask for more thorough phylogenetic analyses, including

tests of reticulate evolution on the basis of fully representative taxon sampling of both riverine and lacustrine haplochromine lineages.

This work was supported by grants of the Deutsche Forschungsgemeinschaft to B.M. (DFG MI649/8-1) and U.K.S. (SCHL567/4-1) and a graduate student grant of the University of Bonn as well as a travel grant from the supporters of the ZFMK to J.Sch. We gratefully thank E. Schraml, O. Seehausen and L. Rüber for providing tissues. We thank R. Schelly, D. Neumann, J. G. Frommen, P. Alibert, A. Dunz, M. Levy, V. Mamonekene, D. Tweddle, A. Ibala-Zamba, J. Punga, U. Ali-Patho, P. Mongindo, C. Danadu, F. Bapeamoni and T. Kadangé Ngongo for indispensable assistance in the field. Special thanks go to K. Langen, F. Eppler and B. Müller who helped with laboratory routines. The Angolan samples came from surveys that were coordinated by D. Neto, conducted by staff from SAIAB and INIP, and funded by the National Research Foundation (South Africa). We thank two anonymous referees and R. Peters for valuable comments on an earlier version of the manuscript.

REFERENCES

- Joyce, D. A., Lunt, D. H., Bills, R., Turner, G. F., Katongo, C., Duftner, N., Sturmbauer, C. & Seehausen, O. 2005 An extant cichlid fish radiation emerged in an extinct Pleistocene lake. *Nature* **435**, 90–95. (doi:10.1038/nature03489)
- Turner, G. F., Seehausen, O., Knight, M. E., Allender, C. J. & Robinson, R. L. 2001 How many species of cichlid fishes are there in African lakes? *Mol. Ecol.* **10**, 793–806. (doi:10.1046/j.1365-294x.2001.01200.x)
- Verheyen, E., Salzburger, W., Snoeks, J. & Meyer, A. 2003 Origin of the superclade of cichlid fishes from Lake Victoria, East Africa. *Science* **300**, 325–329. (doi:10.1126/science.1080699)
- Albertson, R. C., Markert, J. A., Danley, P. D. & Kocher, T. D. 1999 Phylogeny of a rapidly evolving clade: the cichlid fishes of Lake Malawi, East Africa. *Proc. Natl Acad. Sci. USA* **96**, 5107–5110. (doi:10.1073/pnas.96.9.5107)
- Koblmüller, S., Egger, B., Sturmbauer, C. & Sefc, K. M. 2010 Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. *Mol. Phylogenet. Evol.* **55**, 318–334. (doi:10.1016/j.ympev.2009.09.032)
- Genner, M. J. & Turner, G. F. 2012 Ancient hybridization and phenotypic novelty within Lake Malawi's cichlid fish radiation. *Mol. Biol. Evol.* **29**, 195–206. (doi:10.1093/molbev/msr183)
- Joyce, D. A., Lunt, D. H., Genner, M. J., Turner, G. F., Bills, R. & Seehausen, O. 2011 Repeated colonization and hybridization in Lake Malawi cichlids. *Curr. Biol.* **21**, R108–R109. (doi:10.1016/j.cub.2010.11.029)
- Koblmüller, S., Schlieven, U. K., Duftner, N., Sefc, K. M., Katongo, C. & Sturmbauer, C. 2008 Age and spread of the haplochromine cichlid fishes in Africa. *Mol. Phylogenet. Evol.* **49**, 153–169. (doi:10.1016/j.ympev.2008.05.045)
- Salzburger, W., Mack, T., Verheyen, E. & Meyer, A. 2005 Out of Tanganyika: genesis, explosive speciation, key-innovations and phylogeography of the haplochromine cichlid fishes. *BMC Evol. Biol.* **5**, 17. (doi:10.1186/1471-2148-5-17)
- Seehausen, O., Koetsier, E., Schneider, M. V., Chapman, L. J., Chapman, C. A., Knight, M. E., Turner, G. F., van Alphen, J. J. M. & Bills, R. 2003 Nuclear markers reveal unexpected genetic variation and a Congolese–Nilotic origin of the Lake Victoria cichlid species flock. *Proc. R. Soc. Lond. B* **270**, 129–137. (doi:10.1098/rspb.2002.2153)
- Moore, A. E., Cotterill, F. P. D. & Eckardt, F. D. In press. The evolution and ages of Makgadikgadi palaeo-lakes: consilient evidence from Kalahari drainage evolution. *S. Afr. J. Geol.*
- Herder, F., Nolte, A. W., Pfaender, J., Schwarzer, J., Hadiaty, R. K. & Schlieven, U. K. 2006 Adaptive radiation and hybridization in Wallace's Dreamponds: evidence from sailfin silversides in the Malili Lakes of Sulawesi. *Proc. R. Soc. B* **273**, 2209–2217. (doi:10.1098/rspb.2006.3558)
- Salzburger, W., Baric, S. & Sturmbauer, C. 2002 Speciation via introgressive hybridization in East African cichlids? *Mol. Ecol.* **11**, 619–625. (doi:10.1046/j.0962-1083.2001.01438.x)
- Schelly, R. C., Salzburger, W., Koblmüller, S., Duftner, N. & Sturmbauer, C. 2006 Phylogenetic relationships of the lamplogine cichlid genus *Lepidiodiamprologus* (Teleostei: Perciformes) based on mitochondrial and nuclear sequences, suggesting introgressive hybridization. *Mol. Phylogenet. Evol.* **38**, 426–438. (doi:10.1016/j.ympev.2005.04.023)
- Schlieven, U. K. & Klee, B. 2004 Reticulate sympatric speciation in Cameroonian crater lake cichlids. *Front. Zool.* **1**, 1–12. (doi:10.1186/1742-9994-1-1)
- Bezault, E., Mwaiko, S. & Seehausen, O. 2011 Population genomic tests of models of adaptive radiation in Lake Victoria region cichlid fish. *Evolution* **65**, 3381–3397. (doi:10.1111/j.1558-5646.2011.01417.x)
- Seehausen, O. 2004 Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–207. (doi:10.1016/j.tree.2004.01.003)
- Schwarzer, J., Misof, B., Tautz, D. & Schlieven, U. K. 2009 The root of the East African cichlid radiations. *BMC Evol. Biol.* **9**, 186. (doi:10.1186/1471-2148-9-186)
- Kocher, T. D., Conroy, J. A., Mckaye, K. R., Stauffer, J. R. & Lockwood, S. F. 1995 Evolution of NADH dehydrogenase subunit 2 in East African cichlid fish. *Mol. Phylogenet. Evol.* **4**, 420–432. (doi:10.1006/mpev.1995.1039)
- Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Paabo, S., Villablanca, F. X. & Wilson, A. C. 1989 Dynamics of mitochondrial-DNA evolution in animals: amplification and sequencing with conserved primers. *Proc. Natl Acad. Sci. USA* **86**, 6196–6200. (doi:10.1073/pnas.86.16.6196)
- Pompanon, F., Bonin, A., Bellemain, E. & Taberlet, P. 2005 Genotyping errors: causes, consequences and solutions. *Nat. Rev. Genet.* **6**, 847–859. (doi:10.1038/nrg1707)
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. 1997 The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl. Acids Res.* **25**, 4876–4882. (doi:10.1093/nar/25.24.4876)
- Hall, T. A. 1999 BioEDIT: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acid. Ser.* **41**, 95–98.
- Misof, B. & Misof, K. 2009 A Monte Carlo approach successfully identifies randomness in multiple sequence alignments: a more objective means of data exclusion. *Syst. Biol.* **58**, 21–34. (doi:10.1093/sysbio/syp006)
- Nylander, J. A. A., Ronquist, F., Huelsenbeck, J. P. & Nieves-Aldrey, J. L. 2004 Bayesian phylogenetic analysis of combined data. *Syst. Biol.* **53**, 47–67. (doi:10.1080/10635150490264699)
- Huelsenbeck, J. P. & Ronquist, F. 2001 MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755. (doi:10.1093/bioinformatics/17.8.754)

- 27 Stamatakis, A. 2006 RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690. (doi:10.1093/bioinformatics/btl446)
- 28 Link, W., Dixkens, C., Singh, M., Schwall, M. & Melchinger, A. E. 1995 Genetic diversity in European and Mediterranean Faba bean germ plasm revealed by RAPD markers. *Theor. Appl. Genet.* **90**, 27–32. (doi:10.1007/BF00220992)
- 29 Chan, K. M. A. & Levin, S. A. 2005 Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution* **59**, 720–729.
- 30 Tukey, J. W. 1977 *Exploratory data analysis*. Reading, MA: Addison-Wesley.
- 31 Drummond, A. J. & Rambaut, A. 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214. (doi:10.1186/1471-2148-7-214)
- 32 Carnevale, G., Sorbini, C. & Landini, W. T. 2003 *Oreochromis lorenzoi*, a new species of tilapiine cichlid from the late Miocene of central Italy. *J. Vertebr. Paleontol.* **23**, 508–516. (doi:10.1671/1858)
- 33 Rambaut, A. & Drummond, A. J. 2007 TRACER, v. 1.4. See <http://beast.bio.ed.ac.uk/Tracer>.
- 34 Ho, S. Y. W. & Larson, G. 2006 Molecular clocks: when times are a-changin'. *Trends Genet.* **22**, 79–83. (doi:10.1016/j.tig.2005.11.006)
- 35 Genner, M. J., Seehausen, O., Lunt, D. H., Joyce, D. A., Shaw, P. W., Carvalho, G. R. & Turner, G. F. 2007 Age of cichlids: new dates for ancient lake fish radiations. *Mol. Biol. Evol.* **24**, 1269–1282. (doi:10.1093/molbev/msm050)
- 36 Koblmüller, S., Sefc, K. M. & Sturmbauer, C. 2008 The Lake Tanganyika cichlid species assemblage: recent advances in molecular phylogenetics. *Hydrobiologia* **615**, 5–20. (doi:10.1007/s10750-008-9552-4)
- 37 Cotterill, F. P. D. & de Wit, M. J. 2011 Geoecodynamics and the Kalahari Epeirogeny: linking its genomic record, tree of life and palimpsest into a unified narrative of landscape evolution. *S. Afr. J. Geol.* **114**, 489–514. (doi:10.2113/gssajg.114.3-4.489)
- 38 Moore, A. E., Cotterill, F. P. D. W., Main, M. P. L. & Williams, H. B. 2007 The Zambezi River. In *Large rivers: geomorphology and management* (ed. A. Gupta), pp. 311–332. New York, NY: John Wiley & Sons, Ltd.
- 39 Goodier, S. A. M., Cotterill, F. P. D., O’Ryan, C., Skelton, P. H. & de Wit, M. J. 2011 Cryptic diversity of African tigerfish (genus *Hydrocynus*) reveals palaeogeographic signatures of linked neogene geotectonic events. *PLoS ONE* **6**, e28775. (doi:10.1371/journal.pone.0028775)
- 40 Schwarzer, J., Misof, B., Ifuta, S. N. & Schliewen, U. K. 2011 Time and origin of cichlid colonization of the lower Congo rapids. *PLoS ONE* **6**, e22380. (doi:10.1371/journal.pone.0022380)
- 41 Alves, P. C., Harris, D., Melo-Ferreira, J., Branco, M., Ferrand, N., Suchentrunk, F., Melo-Ferreira, J. & Boursot, P. 2006 Hares on thin ice: introgression of mitochondrial DNA in hares and its implications for recent phylogenetic analyses. *Mol. Phylogenet. Evol.* **40**, 640–641. (doi:10.1016/j.ympev.2006.02.016)
- 42 Liu, K., Wang, F., Chen, W., Tu, L., Min, M. S., Bi, K. & Fu, J. 2010 Rampant historical mitochondrial genome introgression between two species of green pond frogs, *Pelophylax nigromaculatus* and *P. plancyi*. *BMC Evol. Biol.* **10**, 201. (doi:10.1186/1471-2148-10-201)
- 43 Melo-Ferreira, J., Boursot, P., Suchentrunk, F., Ferrand, N. & Alves, P. C. 2005 Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Mol. Ecol.* **14**, 2459–2464. (doi:10.1111/j.1365-294X.2005.02599.x)
- 44 Nevado, B., Koblmüller, S., Sturmbauer, C., Snoeks, J., Usano-Aleman, J. & Verheyen, E. 2009 Complete mitochondrial DNA replacement in a Lake Tanganyika cichlid fish. *Mol. Ecol.* **18**, 4240–4255. (doi:10.1111/j.1365-294X.2009.04348.x)
- 45 Wilson, C. C. & Bernatchez, L. 1998 The ghost of hybrids past: fixation of arctic charr (*Salvelinus alpinus*) mitochondrial DNA in an introgressed population of lake trout (*S. namaycush*). *Mol. Ecol.* **7**, 127–132. (doi:10.1046/j.1365-294x.1998.00302.x)
- 46 Ballard, J. W. O. & Whitlock, M. C. 2004 The incomplete natural history of mitochondria. *Mol. Ecol.* **13**, 729–744. (doi:10.1046/j.1365-294X.2003.02063.x)
- 47 Poll, M. & Lambert, J. G. 1965 Contribution à l’étude systématique et zoogéographique des Procatopodinae de l’Afrique centrale (Pisces, Cyprinodontidae). *Bull. Séanc. Acad. R. Sci. Outre-Mer* **1965**, 615–631.
- 48 Roller, S., Hornung, J., Hinderer, M. & Ssemmanda, I. 2010 Middle Miocene to Pleistocene sedimentary record of rift evolution in the southern Albert Rift (Uganda). *Int. J. Earth Sci.* **99**, 1643–1661. (doi:10.1007/s00531-010-0560-z)
- 49 Hudson, A. G., Vonlanthen, P. & Seehausen, O. 2011 Rapid parallel adaptive radiations from a single hybridogenetic ancestral population. *Proc. R. Soc. B* **278**, 58–66. (doi:10.1098/rspb.2010.0925)
- 50 Rieseberg, L. H., Archer, M. A. & Wayne, R. K. 1999 Transgressive segregation, adaptation, and speciation. *Heredity* **83**, 363–372. (doi:10.1038/sj.hdy.6886170)
- 51 Shaw, K. L. 2002 Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proc. Natl Acad. Sci. USA* **99**, 16 122–16 127. (doi:10.1073/pnas.242585899)
- 52 deMenocal, P. B. 2004 African climate change and faunal evolution during the Pliocene–Pleistocene. *Earth Planet. Sci. Lett.* **220**, 3–24. (doi:10.1016/S0012-821X(04)00003-2)
- 53 Trauth, M. H., Maslin, M. A., Deino, A., Strecker, M. R., Bergner, A. G. N. & Dühnforth, M. 2007 High- and low-latitude forcing of Plio-Pleistocene East African climate and human evolution. *J. Hum. Evol.* **53**, 475–486. (doi:10.1016/j.jhevol.2006.12.009)
- 54 Cohen, A. S., Soreghan, M. J. & Scholz, C. A. 1993 Estimating the age of formation of lakes: an example from, L. Tanganyika, East African Rift system. *Geology* **21**, 511–514. (doi:10.1130/0091-7613(1993)021<0511:ETAOFO>2.3.CO;2)
- 55 Trauth, M. H. *et al.* 2010 Human evolution in a variable environment: the amplifier lakes of Eastern Africa. *Quaternary Sci. Rev.* **29**, 2981–2988. (doi:10.1016/j.quascirev.2010.07.007)
- 56 Partridge, T. C. 2010 Tectonics and geomorphology of Africa during the Phanerozoic. In *Cenozoic mammals of Africa* (eds L. Werdelin & J. C. Sanders), pp. 3–16. Berkeley, CA: University of California Press.
- 57 Lavie, L. L., Steckler, M. S. & Brigaud, F. 2001 Climatic and tectonic control on the Cenozoic evolution of the West African margin. *Mar. Geol.* **178**, 63–80. (doi:10.1016/S0025-3227(01)00175-X)