Title: Introgression dynamics of sex-linked chromosomal inversions shape the Malawi cichlid adaptive radiation

Authors: L. M. Blumer[1]†, V. Burskaia[2]†, I. Artiushin[2]†, J. Saha[2]†, J. Camacho Garcia[2], J. Elkin[3], B. Fischer[1], C. Zhou[1,4], S. Gresham[2], M. Malinsky[5], T. Linderoth[6], W. Sawasawa[2], I. Bista[7,8], A. Hickey[3], M. Kucka[9]‡, S. Louzada[10]§, R. Zatha[11], F. Yang[12], B. Rusuwa[11], M. E. Santos[3], Y. F. Chan F[13,9], D. A. Joyce DA[14], A. Böhne[15], E. A. Miska[16,1], M. Ngochera[17], G. F. Turner[18], R. Durbin[1,4]*, H. Svardal[2,19]*

[1]Department of Genetics, University of Cambridge, Cambridge, England, CB2 3EH, UK.
[2]Evolutionary Ecology Group, Department of Biology, University of Antwerp, 2020 Antwerp, Belgium.
[3]Department of Zoology, University of Cambridge, Cambridge, CB2 3EJ, UK.
[4]Wellcome Sanger Institute, Tree of Life, Wellcome Genome Campus, Hinxton, CB10 1SA, UK.
[5]Institute of Ecology and Evolution, Department of Biology, University of Bern, Baltzerstrasse 6, 3012 Bern, Switzerland.
[6]W.K. Kellogg Biological Station, Michigan State University, Hickory Corners, MI 49060, United States of America.
[7]Senckenberg Research Institute and Natural History Museum, 60325 Frankfurt am Main, Germany.
[8]LOEWE Centre for Translational Biodiversity Genomics, 60325 Frankfurt am Main, Germany.
[9]Friedrich Miescher Laboratory of the Max Planck Society, Max-Planck-Ring 9, 72076 Tübingen, Germany.
[10]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK.
[11]School of Natural and Applied Sciences, University of Malawi, Zomba, Malawi.
[12]School of Life Sciences and Medicine, Shandong University of Technology, Zibo, China.
[13]Groningen Institute for Evolutionary Life Sciences (GELIFES), University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands.
[14]Evolutionary and Ecological Genomics Group, School of Natural Sciences, University of Hull, Hull, UK.
[15]Leibniz Institute for the Analysis of Biodiversity Change, Museum Koenig Bonn, Adenauerallee 127, 53115 Bonn, Germany.
[16]Department of Biochemistry, University of Cambridge, Cambridge, England, CB2 3EH, UK.
[17]Department of Fisheries Headquarters, P.O. Box 593, Lilongwe, Malawi.
[18]School of Environmental and Natural Sciences, Bangor University, Bangor, Gwynedd, LL57 2TH, Wales, UK.
[19]Naturalis Biodiversity Center, 2333 Leiden, Netherlands.

†These authors contributed equally to this work

‡Present address: Department of Translational Genomics, University of Cologne, 50931 Cologne, Germany

§Present addresses: Laboratory of Cytogenomics and Animal Genomics, Department of Genetics and Biotechnology, University of Trás-os-Montes and Alto Douro, Vila Real, Portugal; BioSystems and Integrative Sciences Institute, Faculty of Sciences, University of Lisbon, Lisbon, Portugal.

*Corresponding author. Email: hannes.svardal@uantwerpen.be, rd109@cam.ac.uk

Abstract

Chromosomal inversions contribute to adaptive speciation by linking co-adapted alleles. Querying 1,375 genomes of the species-rich Malawi cichlid fish radiation, we discovered five large inversions segregating in the benthic subradiation that each suppress recombination over more than half a chromosome. Two inversions were transferred from deepwater pelagic *Diplotaxodon* via admixture, while the others established early in the *deep benthics*. Introgression of haplotypes from lineages inside and outside the Malawi radiation coincided with bursts of species diversification. Inversions show evidence for transient sex linkage and a striking excess of protein changing substitutions pointing towards selection on sensory, behavioural and reproductive genes. We conclude that repeated interplay between depth adaptation and sex-specific selection on large inversions has been central to the evolution of this iconic system.

Main

Understanding how biodiversity evolves is a fundamental question in biology. While some evolutionary lineages remain virtually unchanged over hundreds of millions of years (*1*), others give rise to a great diversity of species over short evolutionary timescales (*2*). Adaptive radiations are particularly remarkable examples of explosive diversification, with many ecologically, morphologically, and behaviourally differentiated species emerging rapidly from a common ancestor. It is still not well understood how evolutionary lineages can produce such bursts of organismal diversity, but recent insights from genome sequencing point to a widespread contribution of "old" genetic variants (*3*), often introduced into populations by hybridisation (*4*), and reused in new combinations that provide adaptation to novel ecological niches (*5*). A conundrum, however, is the role of meiotic recombination in this process. On the one hand, recombination can create beneficial combinations of adaptive alleles (*6*). On the other hand, recombination can break adaptive combinations apart (*7*), especially in the face of gene flow, producing unfit intermediates (*8*), and impeding speciation (*9*).

Chromosomal inversions – stretches of DNA that are flipped in their orientation – provide a mechanism to break the apparent deadlock between the beneficial and detrimental effects of recombination on species diversification, by strongly suppressing recombination between the inverted haplotype and its ancestral configuration (*7*, *10*, *11*). Inversions acting as "supergenes" can link together adaptive alleles that confer a fitness advantage in a specific environmental context or species background (*12*). In recent years, inversions have increasingly been found to contribute to adaptation (*13*, *14*), genetic incompatibilities (*15*), assortative mating (*16*), sexual dimorphism (*17*, *18*), mating systems (*19*), social organisation (*20*), life-history strategies (*21*) and other complex phenotypes (*11*). Inversions are more common between sympatric than allopatric sister species in fruit flies (*22*), rodents (*23*), and passerine birds (*24*), pointing to their involvement in speciation with gene flow. However, despite their evolutionary relevance in other

systems, there is relatively little information on their role in shaping large vertebrate adaptive radiations (*25–27*).

With over 800 known extant species, Lake Malawi cichlids constitute the most species-rich recent vertebrate adaptive radiation (*28*, *29*). The radiation was able to unfold and generate extraordinary morphological and ecological diversity, despite repeated hybridisation (*30*, *31*) and conserved fertility across species (*32*). Intriguingly, previous studies found broad genetic association peaks for a behavioural phenotype important in assortative mating (*33*) and suppressed recombination in species crosses (*34*). This raises the question of whether recombination-suppressing mechanisms such as inversions contributed to the adaptive diversification of Malawi cichlids.

Here we show that five large inversions segregate across and within many species and groups in the Lake Malawi radiation, and systematically investigate their evolutionary histories and functions. By suppressing recombination, large chromosomal inversions can cause affected genomic regions to show evolutionary histories consistently distinct from the rest of the genome (*35*, *36*). To detect regional deviations from the genome wide evolutionary history we obtained whole genome sequencing (WGS) data from 1,375 individuals of 240 Malawi cichlid species (table S1), detected 84 million single nucleotide polymorphisms (SNPs), and first inferred genome-wide relationship patterns as a backbone (Fig. 1) (see materials and methods). Previous work suggested that the Malawi radiation evolved through serial diversification of three subradiations from a riverine-like ancestor (*30*): 1. A pelagic grouping of the mostly mid-water *Rhamphochromis* and mostly deep-living *Diplotaxodon*; 2. an ecologically and morphologically highly diverse *benthic* subradiation consisting of three subgroups – *deep benthics*, *shallow benthics*, and semi open-water *utaka*; and 3. the predominantly rock-dwelling *mbuna*. The generalist-like stem lineage is represented today by the Malawi cichlid species *Astatotilapia calliptera* (which provided the reference genome for the present study) (*30*, *37*). Phylogenetic inference on our larger dataset generally confirms these major groupings and supports the branching order (Fig. 1).

Large inversions suppress recombination

We identified extensive genomic outlier regions consistent with polymorphic inversions on five chromosomes (2, 9, 10, 11 and 13), each spanning more than half a chromosome (between 17 and 23 Mbp), using a clustering approach on our SNP data set (Fig. 2A, top panel) (see materials and methods). A windowed principal component (PC) analysis (*36*, *38*) of genetic variation revealed that these regions showed relationships among species of the diverse *benthic* clades of Malawi cichlids which were dramatically different from the rest of the genome (Fig. 2A, bottom panel, fig. S2). While in the rest of the genome PC1 tended to separate the three *benthic* subclades, in the focal regions three distinct clusters emerged (two on chromosome 9) that separated different groups of individuals and explained a much higher proportion of the genetic variance (fig. S3). Individuals in the intermediate cluster showed strongly increased

heterozygosity as expected for the heterozygous state of two divergent inversion haplotypes (fig. S4). With double crossover events in only two *deep benthic* individuals, patterns were consistent with almost complete recombination suppression as observed in inversions in other systems (fig. S6) (*22*, *36*, *39*).

We confirmed and characterised inversions in the five regions using a combination of cytogenetics with long and linked read sequencing. We first cultured cell lines of *A. calliptera* and the *deep benthic Aulonocara stuartgranti* (Fig. 2B) and performed fluorescence in situ hybridisation (FISH), confirming a differentially fixed state for inversions on chromosomes 9 and 11, with the latter corresponding to a double inversion (Fig. 2C, fig. S5). Next, we obtained long read sequencing, chromosome conformation capture (Hi-C), and/or linked-read haplotagging (*40*) data and generated *de novo* chromosome-level assemblies of all five major clades of Malawi cichlids and pairwise genome alignments (table S4, figs. S7 to S12). This confirmed the presence of inversions on chromosomes 2, 9, 11 and 13 and revealed three additional smaller inversions that went undetected in the SNP analysis: one small inversion that is nested inside the large inversion on chromosome 2 and two adjacent inversions on chromosome 20 (Supplementary text 1, fig. S13).

To confirm the suppression of recombination between inverted and ancestral haplotypes, we performed an interspecific cross between *A. calliptera* and *Au. stuartgranti* and whole genome sequenced 290 individuals up to generation F3 (table S5). The absence of switching between the inversion-state clusters on chromosomes 9 and 11 on genomic PC1 axis in F2 and F3 individuals confirmed that recombination is fully suppressed in inversion regions of heterozygous F1s (Fig. 2E). Segregation ratios in F2s were Mendelian, except for chr11 showing a moderate deficiency of homozygotes for the the *A. calliptera* haplotype (genotype proportions 20:66:43; χ^2 test on Mendelian ratios p=0.016).

Querying our *Diplotaxodon* high quality PacBio HiFi-based genome assemblies for centromeric repeats allowed us to locate and then project centromeric regions to the reference genome, revealing that only the large inversion on chromosome 2 is pericentric (Fig. 2A, fig. S14, table S6). Analysing the large insertion sequences between assemblies for occurrence of transposable elements, we found a slight trend that inversion regions tended to harbour fewer recently inserted and deleted DNA sequences compared to the rest of the genome (fig. S15) but this trend was not consistent across superfamilies (LTR, DNA, LINE or SINE) and inversions (fig. S16).

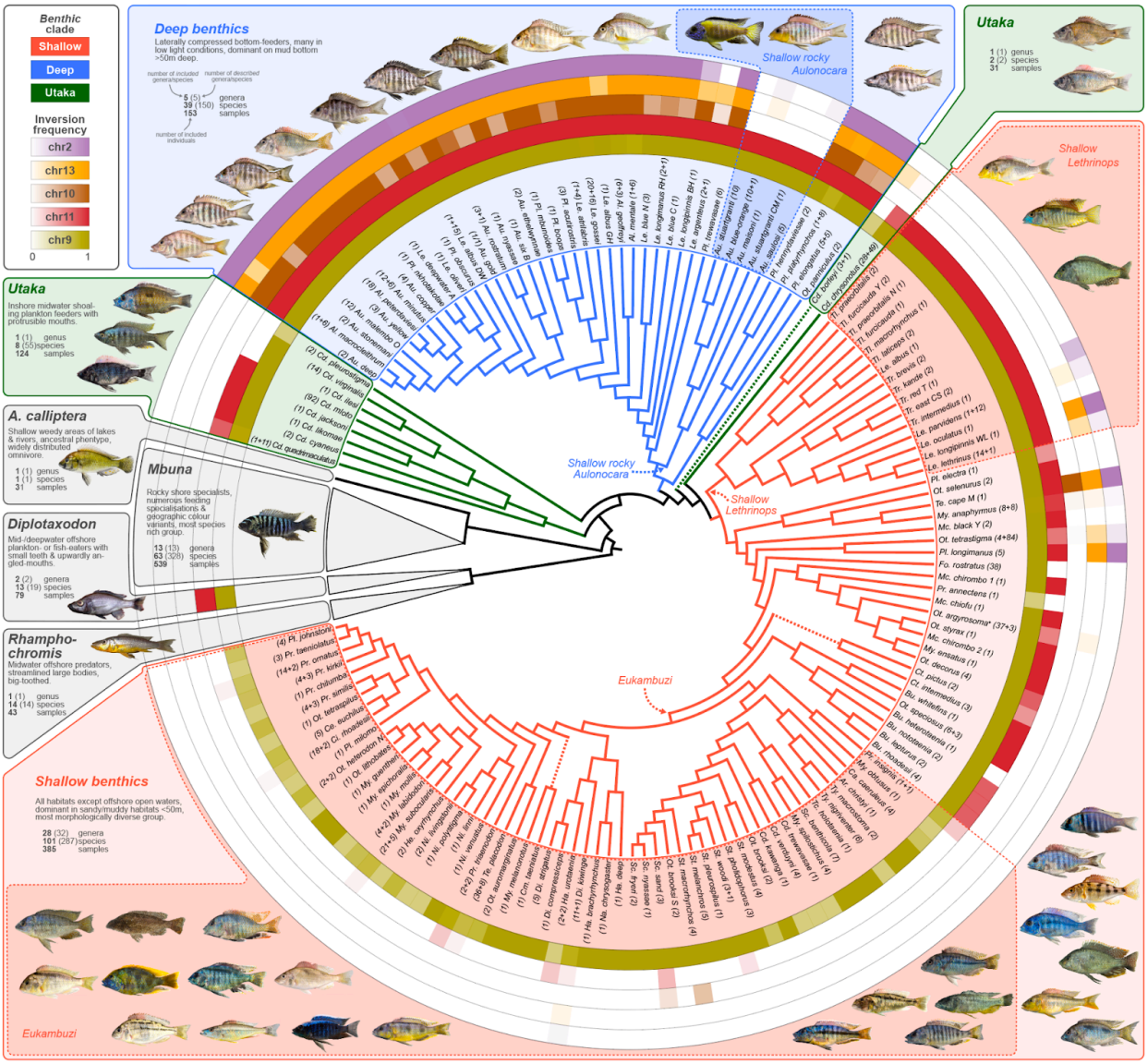

Fig. 1: Study system and prevalence of five large inversions. Consensus phylogeny of Malawi cichlid species used in this study (data S1) with inversion frequency based on WGS and PCR-typing (see materials and methods) shown in rings around the phylogeny (the same colours are used throughout the article). The *benthic* subradiation is expanded here to show the phylogenetic position of each species. Note that *utaka* are not monophyletic in this phylogeny. Non-*benthic* groups of Malawi cichlids, i.e., the *pelagic* subradiations of *Rhamphochromis* and *Diplotaxodon*, the subradiation of predominantly rock-dwelling *mbuna*, and *Astatotilapia calliptera* – a species distributed in rivers and margins around the lake that shares its putatively ancestral characteristics and genus assignment with riverine haplochromines outside of the radiation – are each represented by a single grey triangle approximately reflecting species richness. Dashed lines indicate branches with uncertain placement. See data S1 and S2 for full phylogenies with branch lengths and support values. Annotations next to species/clade names provide the numbers of sequenced/inversion-genotyped samples (additional samples which were inversion-genotyped with PCR are indicated as + *n* in the annotation). Further details are given in table S2. Full species names are given in table S3. Species names for representative photographs are given in fig. S1.

Inversions segregate within *benthic* subradiation

Next, we investigated the distribution of inversion states across the phylogeny using a multi-step PC approach to infer inversion genotypes for all 1,375 sequenced individuals ("WGS-typing") (Fig. 1, tables S 7 and S8; figs. S17 and S18) and denoting as non-inverted or ancestral the orientation of the outgroup species *Oreochromis niloticus* (fig. S20) and *Pundamilia nyererei* (fig. S19). To further increase the number of genotyped individuals, we identified TE insertions in breakpoint regions highly correlated with inversion state and PCR-typed these insertions in an additional 401 individuals (fig. S22, tables S9 to S11). Together, these analyses revealed that all specimens of the *mbuna* and *Rhamphochromis* subradiations and *A. calliptera* were fixed for the non-inverted ancestral state for all five large inversions. All *Diplotaxodon* specimens also lacked the inversions on chromosomes 2, 10 and 13 but showed the same intermediate signal in the typing of the chromosome 9 and 11 inversions. In our *de novo* assembly for *D. limnothrissa*, both inversions are present (fig. S21, fig. S8), suggesting that *Diplotaxodon* are fixed for the inverted chromosome 9 and 11 haplotypes.

Among the *benthic* clades, the five inversions showed strikingly different frequencies across the species: for chromosomes 2, 10, and 13, the inverted state was fixed or at high frequency in most *deep benthic* species, but almost absent among *shallow benthic* species and *utaka*. For chromosomes 9 and 11, the inverted states were fixed in most *benthics*, except for a large monophyletic subclade of *shallow benthics* in which chromosome 11 was mostly fixed for the ancestral non-inverted state and chromosome 9 mostly polymorphic. We will refer to this group as "*eukambuzi*", inspired by the local name Kambuzi for some members of this group (Fig. 1). In summary, our inversion typing is consistent with a scenario in which inversions on chromosomes 2, 10, and 13 rose to high frequency in an ancestor of the *deep benthic* lineage, while the inversions on chromosomes 9 and 11 rose to high frequency in the ancestors of two non-sister groups – pelagic *Diplotaxodon* and *benthics* – but with a monophyletic subgroup of the *benthics* (*eukambuzi*) retaining or re-gaining the non-inverted ancestral state.

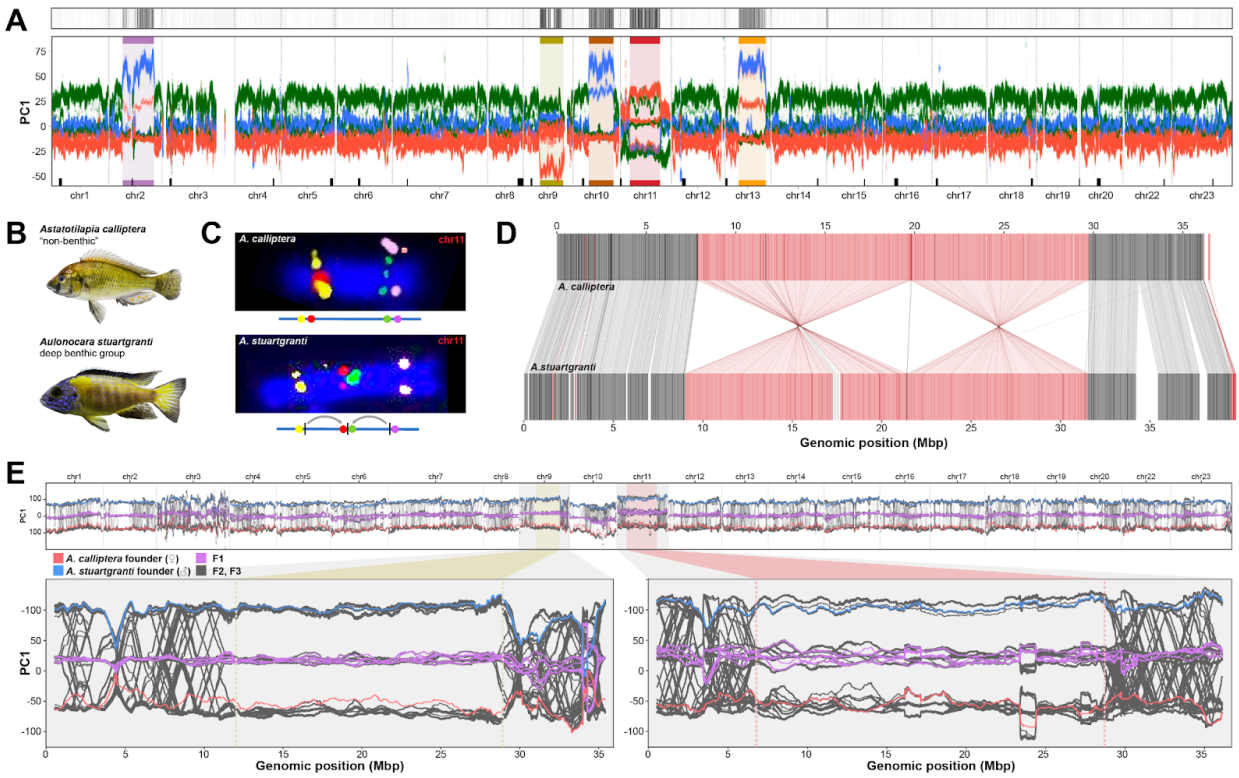

Fig. 2: Characterisation of inversions. (**A**) (Top panel): Identification of genomic regions with clusters of aberrant phylogenetic patterns (see materials and methods). (Bottom panel): First genetic principal component in overlapping 1 Mbp windows along chromosomes, using the same colours for the *benthic* subclades as in Fig. 1. Outlier regions from the top panel are highlighted and colour-labelled. Centromeric satellite regions are indicated as black rectangles on top of the x-axis. (**B**) Representative photographs of the species used in panels C-D: *A. calliptera*, a lineage of the Malawi radiation distinct from *benthics*, from which the reference genome was produced, and *Aulonocara stuartgranti*, a species that genetically belongs to the *deep benthic* group, but lives in shallow rocky habitats. According to WGS-typing the species are expected to show fixed different orientations for the chromosome 9 and 11 inversions. (**C**) Fluorescence in situ hybridisation (FISH) of markers on chromosome 11 left and right of the putative inversion breakpoints show the expected non-inverted orientation (upper panel) in *A. calliptera*. In *Au. stuartgranti* we see a double inversion (lower panel). (see fig. S5 for FISH of chromosome 9) (**D**) Whole genome alignment of an ONT duplex long-read assembly of *Au. stuartgranti* to the *A. calliptera* reference assembly (which was re-scaffolded with chromosome conformation capture (Hi-C) data, see material and methods) confirms the double inversion on chromosome 11. (For other chromosomes see fig. S7.) (**E)** Top: Genomic windowed PC1 values of whole genome sequenced founders and progeny of an interspecific cross. Among 290 F2 and F3 individuals no crossing-over events were observed in the inversion regions of chromosomes 9 (left bottom) and chromosome 11 (right bottom), while recombination is frequent in the flanking regions and on other chromosomes.

Inversions contribute to sex determination

Inversions are a common feature in the evolution of suppressed recombination on sex chromosomes in many organisms (*41*, *42*). To investigate sex-association in Malawi cichlids, we examined inversion genotype frequencies in species with at least one heterozygous sample (table S12). For chromosomes 9, 10 and 11 we observed a relative excess of heterozygous genotypes and a clear lack of homozygous ancestral states (Fig. 3A). Furthermore, while females tended to be homozygous for the inverted state, there was a significant correlation of male sex with the heterozygous state (Fisher's exact test p= $3.7x10^{-11}$, 0.014, and 5.3x10-11, respectively) (Fig. 3A). The associations remained when WGS and PCR-typed samples were considered independently (except for chromosome 10, which was not significant for WGS samples alone).

Species tended to have heterozygotes for only one of these three sex-associated inversions, with only one out of 61 species showing double heterozygotes for two inversions among WGS-typed samples. For chromosome 9, heterozygotes were almost exclusively present among (*shallow benthic*) *eukambuzi* (Fig. 1) and, despite the presence of 77 heterozygous individuals across 12 species, not a single homozygous ancestral state was present among all *benthics*. The existence of an X/Y sex-determining system on this chromosome is further supported by a study analysing a cross between *A. calliptera* and the *eukambuzi Protomelas taeniolatus (43)*. Species showing evidence for a role of the chromosome 10 inversion in sex determination were mainly *deep benthic*, while for chromosome 11, species with heterozygotes present clustered phylogenetically intermediately between the basal *shallow Lethrinops* and *eukambuzi*, except for *utaka* (*Copadichromis*) species (Fig. 1).

However, the association of inversion heterozygosity with male sex was usually not perfect within species, which could be due to additional genetic or environmental effects contributing to sex determination or due to errors in our sex assignment (see materials and methods). An exception to this incomplete penetrance is *C. chrysonotus*, in which all of our 28 samples showed perfect correlation of sex with chromosome 11 inversion state (Fig. 3B, C, Fisher's exact test p-value = $4.7x10^{-8}$).

To further investigate the generality of the role of inversions in sex determination of haplochromine cichlids, we applied our SNP-based inversion detection approach to publicly available sequencing data for the Lake Victoria adaptive radiation (*44*), which confirmed that the sex-linked regions identified by Feller et al. (*43*) on chromosomes 9 and 23 are in fact chromosomal inversions (the one on chromosome 9 being distinct from the one present in Malawi) (figs. S19 and S23).

In summary, our analysis revealed robust, though imperfect associations of male sex with heterozygous inversion states on chromosomes 9, 10, and 11 in Malawi cichlid species which are polymorphic for these inversions (Fig. 3A) and two chromosomes with distinct sex-linked chromosomal inversions in the closely related Lake Victoria radiation, consistent with a rapid turnover of an incompletely penetrant Y-like function of these inversions in different species.

These results suggest a widespread role for chromosomal inversions in the highly dynamic sex determination of haplochromine cichlids (*43*, *45*, *46*).

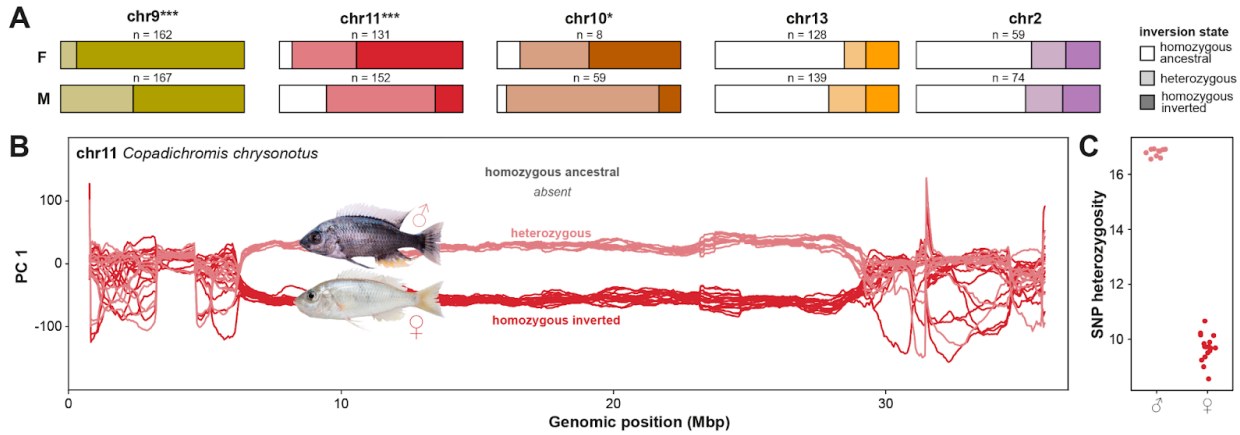

Fig. 3: Sex association of inversions. (**A**) Proportions of homozygous/heterozygous inversion genotypes in males and females in species with heterozygotes present, according to WGS and PCR typing of 809 samples (67 species). Asterisks denote significance levels of Fisher's exact tests of inversion genotype–sex correlation: * p<0.05, *** p<0.001 (**B**) Windowed PC analysis along chromosome 11 demonstrates perfect association of inversion genotype with sex in our sample of 28 *Copadichromis chrysonotus*. (**C**) SNP heterozygosity among our sample of male and female *C. chrysonotus*, measured as number of het SNPs per 10 kbp.

Origins and introgression patterns of the inversions

To better understand evolutionary histories in inversion regions, we estimated genetic divergence times both within *benthics* and from the other major Malawi clades for the five inversion regions (both the inverted and ancestral haplotypes) as well as for the remaining non-inverted regions of the genome (Fig. 4A). We also rephased heterozygous SNPs (see material and methods) in inversion heterozygote individuals using SNP allele frequencies in the different inversion genotypes and built consensus phylogenies for all haplotypes in all inversion regions (figs. S24 to S28). The results imply that the evolutionary histories of all the inversion regions were shaped by multiple introgression events, as detailed below.

Genome-wide phylogenetic trees for Malawi cichlids (e.g. ref. (*30*), Fig. 1, data S2) suggest that *benthics* are a sister clade to *A. calliptera* and *mbuna*. However, in the top row of Fig. 4A we see that both *shallow* and *deep benthics* are closer to *Diplotaxodon* than to *Rhamphochromis*, and closer to *A. calliptera* than to *mbuna*. These results are inconsistent with the genome-wide trees. It was already established that there is evidence for gene flow from *Diplotaxodon* into the *benthics* (*30*). However, here we suggest that the most likely resolution of these new observations is that the *benthics* arose through admixture between the *Diplotaxodon* and *A.*

calliptera lineages after their splits from *Rhamphochromis* and *mbuna* respectively, as shown in Fig. 4B. This hybrid *benthic* origin model changes the relative order of divergence of *benthics* and *mbuna* from the *A. calliptera* stem, which is a natural consequence of the inability of the binary tree phylogeny to account for admixture. In the context of this revised model of the origin of the *benthics*, the sharing of the chromosome 9 and 11 inversions (as well as the chromosome 20 compound inversion, see figs. S7 to S10) between all *Diplotaxodon* and some *benthics* is most parsimoniously explained by them having arisen in an ancestor of *Diplotaxodon* and passing into the *benthics* in their founding admixture event, as indicated by ① in Fig. 4B.

Considering subsequent events, the chromosome 9 region is fixed inverted in most *benthic* species except for the *eukambuzi* and a few others where it is heterozygous (Fig. 1). However, it is striking that the non-inverted state found in these few shallow *benthics* is much more divergent from the rest of the Malawi radiation than any other inversion haplotype and the rest of the *benthic* genome (Fig. 4A, row 2). This suggests that this haplotype in the ancestral, non-inverted orientation re-introgressed into shallow *benthics* from a lineage that is phylogenetically outside the entire Malawi radiation. To search for the origin of this haplotype, we analysed a second SNP callset containing representatives of all major Malawi clades as well as a wide variety of related riverine and lacustrine cichlid species ("haplochromines") (*31*) (fig. S30, table S13, materials and methods). Computing ABBA-BABA tests (table S14) revealed excess allele sharing of the non-inverted *benthic* chromosome 9 haplotype with a number of distantly related outgroup species, with the strongest signal being to *Pseudocrenilabrus philander*, one of the few outgroup species present in the catchment of Lake Malawi (D=0.45, block-jackknifing z-score 6.5; FWER corrected $p = 4x10^{-9}$). We conclude that the chromosome 9 non-inverted *benthic* haplotype is not closely related to other Malawi haplotypes (whether inverted or not), but arrived in the ancestor of *eukambuzi* through admixture with a distantly related *Pseudocrenilabrus*-like cichlid lineage, as indicated by ② in Fig. 4B, and spread to the few other species in which it is found by subsequent introgression.

Turning next to the chromosome 11 inversion, this is fixed inverted in the *deep benthics* and *shallow Lethrinops* and variably present in other shallow species, being mostly non-inverted in the *eukambuzi* (Fig. 1). Unlike in all the other inversion regions, the non-inverted haplotype is closest to *A. calliptera* (Fig. 4A, row 3), suggesting that it re-introgressed into shallow benthics from an *A. calliptera* ancestor or an extinct close relative, as indicated by ③ in Fig. 4B. This haplotype shows lower heterozygosity (and nucleotide diversity) than the inverted one (fig. S31), consistent with a genetic bottleneck upon re-introgression. Signatures of gene flow from *A. calliptera* into the *shallow benthics* can also be seen outside of chromosome 11: (i) outside the inversions, *shallow benthics* are closer to *A. calliptera* than *deep benthics* are (Fig. 4A, row 1), and (ii) the non-inverted haplotypes of the chromosome 2 and 13 inversions are closer to *A. calliptera* than the inverted haplotypes, which are present predominantly in the *deep benthics* and protected from introgression because of their orientation (Fig. 4A). The f-branch analysis in Fig. 3 of our earlier study (29) further supports this genome-wide gene flow, showing significant

excess allele sharing of the ancestral branch of shallow *benthics* with *A. calliptera* relative to *deep benthics*. Finally, some *utaka*, including *C. chrysonotus* where it is sex-determining (see above), also received the *A. calliptera*-like non-inverted haplotype in the chromosome 11 region. Overall, consensus phylogenies and genetic divergence measures suggest that *utaka* are a complex (and variable) mixture of *shallow benthic* and other contributions, which also explains their non-monophyletic clustering in the genome-wide phylogeny (Fig. 1, figs. S32 and S33).

The remaining inversions on chromosomes 2, 10 and 13 are all common among *deep benthics*, and rare or absent among *shallow benthics* and *utaka*, suggesting that they first appeared and rose in frequency early in the deep *benthic* lineage (Fig. 4B). This is also consistent with the inversion region haplotype phylogenies (figs. S24, S26, S28), which suggest that the few cases of *deep benthics* with non-inverted states and *shallow benthics* with inverted states are due to a limited number of later gene flow events (fig. S29). Most of these events transmitted more than one inversion haplotype between *deep* and *shallow benthic* lineages and also genetic material outside the inversion regions, as seen in ABBA-BABA tests of excess allele sharing outside inversion regions (Fig. 4C) (*47*). This underlines that the sharing of inverted haplotypes among *deep* and *shallow benthic* species is generally not a result of random segregation through incomplete lineage sorting but due to gene flow. The most consequential of these transmission events are indicated in Fig. 4B. Specifically, inverted haplotypes introgressed from *deep benthics* into the *shallow Lethrinops*, which show traits similar to deepwater *Lethrinops* consistent with their taxonomic assignment to the same genus despite their distinct molecular phylogeny (④ in Fig. 4B). Furthermore, ancestral chromosome 10 haplotypes introgressed from a shallow source into a broader *deep benthic* group of species (indicated by ⑤ in Fig. 4B), where this haplotype now is male-linked and hence might function as a Y chromosome. Finally, the ancestral haplotypes for chromosomes 2, 10, and 13 introgressed from some *shallow benthic* source into *shallow rocky Aulonocara* (⑥ in Fig. 4B), where they now appear to be fixed. Further events, shown in fig. S29), are qualitatively similar to these in that they mostly correspond to inversion transmissions between *deep* and *shallow benthic* species that were collected at a similar depth (fig. S34). This suggests that ancestral/inverted haplotypes of these inversions carry alleles favourable for a shallow/deep water lifestyle, a possibility which we will further investigate below.

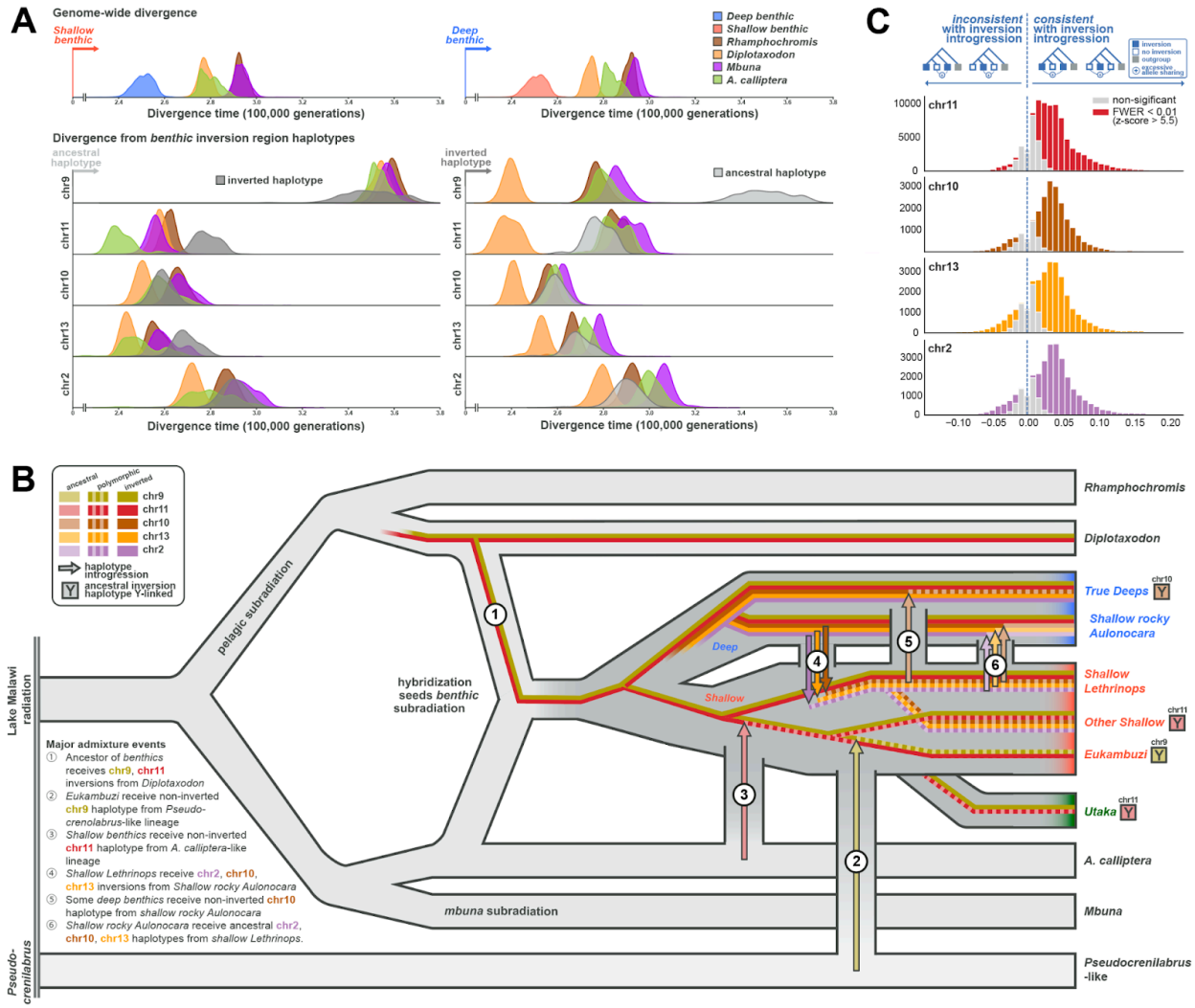

Fig. 4: Evolutionary history of inversion haplotypes. (**A**) Density plots of pairwise sequence divergence translated into divergence (coalescence) times assuming a mutation rate of 3×10^{-9} bp/generation (*30*). The top panels show results for the genome outside the five large inversions, comparing all major clades against *shallow benthics* (left) and *deep benthics* (right). Panels below the top row show divergence in inversion regions for the ancestral (left) and inverted (right) *benthic* haplotypes. (**B**) A simplified model for the evolutionary history of the Malawi cichlid radiation, which includes several inversion haplotype transmission events. Vertical grey connections indicate major gene flow events. Numbered arrows indicate transfer of inversion-region haplotypes. For further events see fig. S29. Some inversion haplotypes are sex-associated with Y-like function of re-introgressed non-inverted haplotypes in extant lineages. (**C**) Evidence for transfer of inversion haplotypes through introgressive hybridisation. Histograms of ABBA-BABA statistics *D(P1, P2, P3, Outgroup)*, calculated outside the inversions for all species trios in which the two more closely related species P1 and P2 differ in their inversion state but one of them shares the inversion state with P3. Significantly positive values are suggestive of gene flow between the species sharing inversion states, while significantly negative values suggest gene flow between species not sharing inversion states. Under the null hypothesis of no inversion introgression the statistic is symmetric around zero.

Pervasive signatures of adaptation on inversion haplotypes

Next, we investigated whether the evolution of inversion haplotypes was driven by adaptive processes and, specifically, whether they could constitute "supergenes" of co-adapted alleles.

The above analyses suggest that the inversion haplotypes rose in frequency relatively rapidly in early lineages of the Malawi radiation. Furthermore, we found that these regions show increased relative divergence, reduced cross-coalescence rates (a measure of genetic exchange) and, for chromosomes 2, 10 and 13, also disproportionally high effect sizes in a population-structure-corrected genome-wide association study between *deep* and *shallow benthics* (materials and methods, Fig. 4A, figs. S32, S35 to S37). All this suggests that inversion regions disproportionally contribute to divergence among these lineages. To identify variants potentially under selection, we computed correlation coefficients and significance scores ($-\log_{10}$ p-value) between SNP and inversion genotypes (figs. S38 and S39). We expected the most highly inversion correlated SNPs ("ICS") to correspond to variants present on the early inversion haplotypes that might have driven the early evolution of these haplotypes.

ICS were much more likely to fall into protein coding regions compared to other SNPs on all five chromosomes (Fig. 5A), suggesting that relaxed purifying and/or positive selection contributed to inversion haplotype divergence. Furthermore, the neutrality index (NI) and the direction of selection (DoS) statistic both indicated an excess of non-synonymous (amino acid changing) divergence between inversion haplotypes on all five chromosomes (fig. S40), a pattern often attributed to positive selection. To confirm that this observation was not exclusively driven by relaxed purifying selection (*48*), we calculated the normalised ratio of non-synonymous to synonymous mutations (dN/dS). While dN/dS ratios smaller than 1.0 can reflect a mixture of evolutionary forces, values larger than 1.0 are only consistent with adaptive evolution. We found that dN/dS ratios increased with increasing inversion correlation, with highly positive ICS showing dN/dS ≥ 1 for all five chromosomes (Fig. 5C). Therefore, we conclude that positive selection contributed to diversification between ancestral and inverted haplotypes on all five chromosomes (Supplementary text 2, table S15).

Inversions contribute to sensory and physiological adaptation

Since all five inverted haplotypes are found at higher frequencies in deepwater-living species (fig. S34, table S16), we hypothesised a contribution to niche divergence along a depth gradient with relevant sensory and physiological adaptations related to changes in light, oxygen, and hydrostatic pressure, as observed in many organisms (*49–51*) including cichlids (*30*, *52*, *53*). Querying ICS for functional associations, we found strong support for this hypothesis (Supplementary text 3). Gene ontology (GO) enrichment analyses near ICS (see materials and methods) showed vascular system-related categories functionally linked to responses to hypoxia stress to be enriched in three inversions and sensory system-related categories in all five of them

(table S17). Furthermore, querying expression of the 330 genes with non-synonymous ICS (nsICS) in a multi-tissue single-cell gene expression atlas of the zebrafish model (Daniocell database (*54*), tables S18, S19, S20) revealed that homologs of those genes were significantly more likely to be highly expressed in the eye compared to randomly selected genes (FDR-corrected binomial test $p=6\times10^{-6}$, fig. S41) and generally often highly expressed in sensory and vascular tissues (Fig. 5C, D; fig. S41). Finally, among the 83 genes with two or more nsICS, we found strong candidate genes for depth adaptation, including master regulators of vasoconstriction (*agt*) and urate degradation (*urad*), genes involved in signal transduction in photoreceptor cells (*arr3a, gucy2d, pou6f2*), otolith tethering (*tecta*), or essential for sound perception in mammals (*myo7ab*) (Fig. 5D, Supplementary text 3, table S22). Interestingly, some genes showed a close relationship between the amino acid sequence coded by the inverted haplotype and that of *Diplotaxodon*, even for the inversions not present in *Diplotaxodon* (e.g., *arr3a* on chromosome 10, see also Malinsky et al. (*30*)) (fig. S42), which is consistent with the hybrid ancestry of *benthics* being differentially selected between inversion region haplotypes. Together, our results suggest that widespread functional divergence in genes related to vascular and sensory systems shaped the early evolution of inversion haplotypes. This makes sense, because sensory systems mediate sound perception, mechanoreception and vision, essential for navigation and feeding in fishes (*55*), making them important targets of ecological adaptation to differing underwater environments.

Adaptations related to behaviour, sex and reproduction

In addition to their importance in ecological adaptation, sensory systems are further relevant because of their ability to promote speciation through sensory drive (*56*) and their role in assortative mating (*57*), which was specifically linked to differences in the underwater light environment (*58*). Among *benthic* Malawi cichlids, assortative mating involves complex behavioural phenotypes such as the construction of mating platforms ("bowers") by males. Bower building is a highly species-variable, sexually-selected male extended phenotype (*59*), which has previously been conjectured to be associated with the existence of supergenes (*60*). Intriguingly, a previous study had identified conspicuous associations between bower architecture and genetic divergence peaks inside our chromosome 2 and 11 inversion regions (*33*).

Although we found no significant correlation between bower type ('pit' vs 'castle') and the presence of the five inversions when testing in a Bayesian framework that accounts for phylogeny (materials and methods; fig. S43, table S21; but see Supplementary text 1), investigating nsICS for functional association yielded robust evidence for pervasive behavioural adaptation on inversion haplotypes (Fig. 5D, Supplementary text 3). First, among the twelve zebrafish tissues examined for consistently high expression of nsICS gene orthologs, beyond the eye (discussed above), nervous tissue more generally was significantly over-represented (FDR-corrected binomial test $p=5\times10^{-3}$, fig. S19). Second, we found nsICS to be significantly overrepresented among neuroreceptor genes that have been previously associated with social

(glutamate) and affiliative (oxytocin and arginine vasopressin/vasotocin, opioid receptors, dopamine, serotonin) behaviour (*60*) (6 out of the 46 candidate genes, Fisher's exact test, p= 0.0011) (table S23). These are three glutamate receptors (*gria3a* and *gria4b* on chromosome 10, *grik3*, on chromosome 11), one opioid receptor (*oprd1b* on chromosome 11), one dopamine receptor (*drd2l* on chromosome 11), and one serotonin receptor (*htr7a* on chromosome 13) (fig. S42). Remarkably, genes and GO categories associated with neurotransmitter regulation are not only associated with fish social behaviour in general, but were specifically connected to bower building behaviour in Malawi cichlids (*61*).

An additional cue linking the inversions to reproduction is their role in sex determination as illustrated in Fig. 3. Assortative mating and the evolution of sex linked regions are both forms of sex-specific selection (*62*). Although the interplay between these forms of selection is not well understood, it might give rise to synergistic evolutionary dynamics, potentially mediated by sexually antagonistic selection (*63*, *64*). Consistent with this, we found many of the strongest candidate genes for adaptive evolution on inversion haplotypes to be related to sex and reproduction (Fig. 4D; Supplementary text 3). Among these, *smchd1* – a chromatin repression regulator responsible for ovarian differentiation in mammals (*65*) – is among the eight genes with two or more nsICS on chromosome 9. On chromosome 11, the only gene with five or more nsICS is *plecb* (fig. S42), whose ortholog mediates sex differentiation in mouse gonads (*66*). On chromosome 10, the only gene with four nsICS is *pof1b* (fig. S42) linked to premature ovarian failure in mammals (*67*). Additionally, the only two genes on chromosome 2 with five or more nsICS are *ahnak*, involved in sex determination in the cichlid *Oreochromis mossambicus*, and *atrx* – X-linked and sex-reversing in humans (*68*).

Given that the evolution of sex-determining regions often involves changes in gene expression between male and female haplotypes, we obtained transcriptomic data of five tissues (muscle, liver, brain, gills, gonads) for 11 males of *C. chrysonotus*, the *utaka* species in which all males were heterozygous for the chromosome 11 inversion while females were fixed for the inverted state, and investigated allele specific expression (ASE). Further, lacking access to appropriate female samples, we obtained equivalent data for eleven male *C. mloto*, a congeneric species fixed for the derived chromosome 11 inversion state, to perform differential gene expression (DE) analysis between the two inversion states. Expression of most genes was similar across haplotypes in male *C. chrysonotus*, with a moderate bias towards lower expression of the Y-like non-inverted haplotype among genes with significant ASE (Supplementary text 3), a pattern seen in many organisms (*69*). Several of the significant ASE and DE genes (FDR < 0.05) were implicated in sex determination, sex specific expression or gonad function in other (fish) species (table S22) and ICS were significantly overrepresented among genes with significant allele specific expression (Fisher's exact test $p=1.6x10^{-7}$).

In summary, our analyses support the presence of adaptive alleles on all inversion haplotypes, contributing to physiology, behaviour, reproduction, sex determination, or sex-specific function (Fig. 5D).

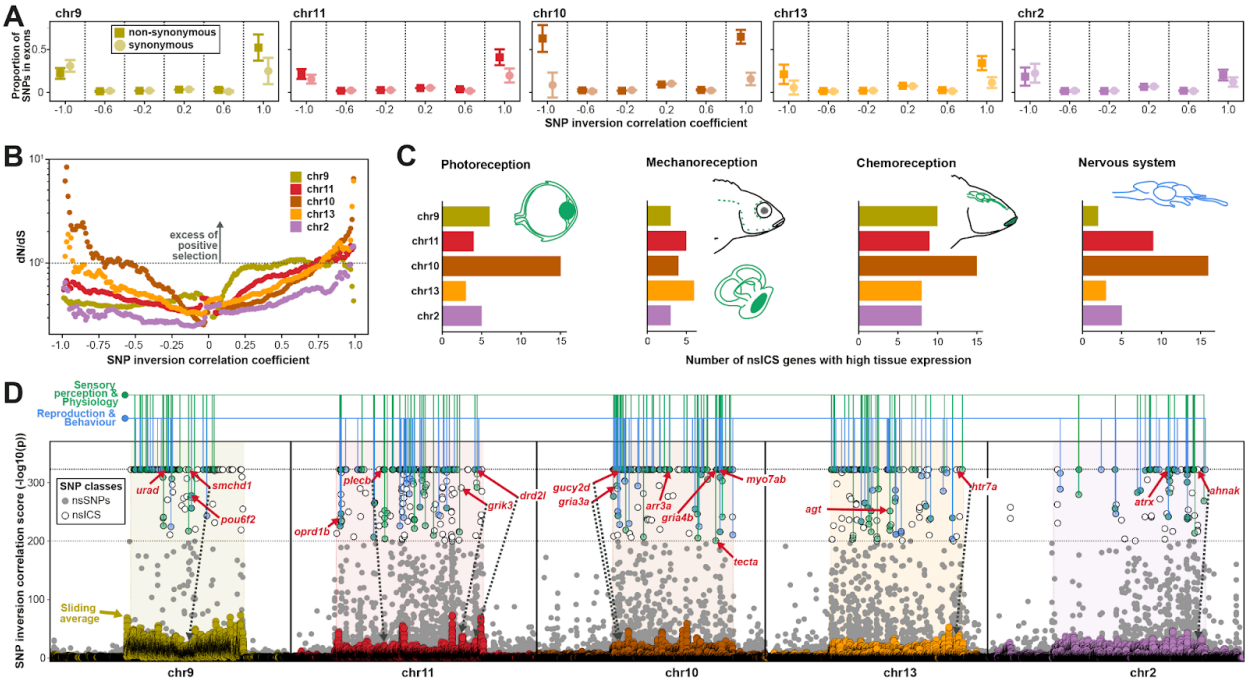

Fig. 5: Adaptive evolution of inversion haplotypes. (**A**) Proportion of coding SNPs among all SNPs in inversion regions measured for SNPs in intervals of inversion genotype correlation. A positive correlation coefficient corresponds to the derived SNP allele being more common on the inverted haplotype, while a negative coefficient corresponds to the derived SNP allele being more common on the ancestral haplotype. (**B**) dN/dS measured for SNPs as a function of inversion genotype correlation (see materials and methods). (**C**) Number of nsICS genes per inversion, which are highly expressed in the main sensory (green) and nervous system (blue) tissues in developing zebrafish, based on the single-cell expression atlas Daniocell (*54*). The tissues are grouped into the functional categories: vision (eye), mechanoreception (lateral line, ear), chemoreception (taste, olfaction), and nervous system (neural). (**D**) Non-synonymous SNPs (empty dots; if ICS: grey dots) and averaged ICS scores in 100 SNP rolling windows (markers colour-coded according to inversion) on the five inversion chromosomes. We annotated nsICS if located in genes with high expression in zebrafish tissue groups related to reproduction, vascular function and sensory perception or had neural function (coloured vertical connector lines, markers are also filled accordingly). nsICS in candidate genes discussed in the main text are annotated with red arrows; an additional grey dotted arrow signifies if the same gene overlaps with the top 0.5 percentile of sliding average ICS windows (table S22).

Discussion

In this article we identify large chromosomal inversions present in Lake Malawi cichlids and present evidence that the interplay between introgression and selection on inversions helped shape the Malawi radiation. Chromosomal inversions have been implicated in adaptation, sex

determination, and speciation in many systems, especially in the context of adaptive divergence with gene flow (*7*, *11*, *15–22*). Gene flow is common in young adaptive radiations (*5*, *70*, *71*), with hybridisation and genetic introgression having fuelled the large cichlid radiations in lakes Tanganyika (*72*, *73*), Victoria (*4*, *74*) and Malawi (*31*, *71*), but so far with little evidence for chromosomal inversion locking together adaptive alleles. Covering more than half of each of chromosomes 2, 9, 10, 11, and 13 (11% of the genome), the five large inversions studied here appeared early in the Malawi radiation and contributed to sex determination and ecological adaptation. Given the additional sex-linked inversions we found in Lake Victoria cichlids (fig. S23), potentially also involved in adaptive introgression (see ref (*44*), Fig. 2C), and large scale differences in male/female DNA sequence in Lake Tanganyika cichlids (*45*), we suggest that such rearrangements may be a common feature of adaptive radiation in cichlids.

The inversions we study show signs of repeated transfer between lineages since their origin, in one case involving a relatively divergent lineage from outside the Malawi radiation. We suggest that these introgression events were driven by selection on the supergene of coadapted alleles, with a primary driver being water depth adaptation which is a frequent axis of differentiation in fishes (*52*, *75*). We also see that the most recently introgressed haplotypes act as Y chromosomes in some extant species, with signatures of selection on genes related to sex and reproduction on all inversion haplotypes. Consistent with the highly dynamic nature of sex determination in many fishes (*76*) and specifically cichlids (*43*, *45*, *46*), our results point to a relatively easy recruitment of sex determination loci (SDLs), possibly as a direct consequence of introgression of relatively divergent haplotypes affecting a sex determination threshold, or due to heterozygote advantage of introgressed inversions selecting for the recruitment of SDLs (*77*).

Sexual selection has been identified as a major predictor of successful radiation in cichlids (*78*), and assortative mating is a main driver of cichlid reproductive isolation (*57*). Both of these processes rely heavily on the same sensory systems (e.g., vision (*57*), olfaction (*79*), and hearing (*80*)) that are also relevant for adaptation to depth and feeding niches. Although not statistically significant in our re-analysis (fig. S43), previously observed genetic association for male bower building (a trait important for sexual selection and assortative mating) (*33*) in the inversion regions we describe, is consistent with our observation that genes implicated in social behaviour have enriched signals for adaptation within the large inversions. Combining these observations, we suggest that selective pressures that drive the evolution of sex-linked genetic regions, potentially contributing to sexual selection and reproductive isolation, could at the same time provide the substrate for ecological diversification and consequent speciation (*81*).

The admixture events we identify (Fig. 4B) tend to coincide with eco-morphological diversification of the resulting lineages, the most prominent examples being *benthics* as a whole and the species rich *eukambuzi*, which show exceptional diversity in eco-morphology, body patterning, and colouration (*29*). While our analysis focuses on genetic variants on inversion haplotypes that were likely important in early adaptive divergence, we expect that additional genetic variants specific to particular species groups will prove important for further

diversification. Furthermore, alongside the inversions that identify admixture events, we also see a signal of introgressed material in the rest of the genome, providing a potential substrate for further selection and adaptation. Indeed, while we focussed on large inversions segregating across many species, there are likely to be many more inversions and other structural genetic variants that are smaller or have a more limited taxonomic distribution, as suggested by our whole genome alignments and in Talbi et al. 2024 (*82*). Surely, more such variants with relevance in adaptive diversification are to be found in future studies.

In conclusion, ancestral and inverted haplotypes of five chromosomal scale inversions in the Malawi cichlid adaptive radiation show supergene-like signs of pervasive adaptive evolution and repeated introgression associated with speciation. Together with the transient sex-linked nature of introgressed haplotypes, this provides a substrate for rich evolutionary dynamics around the interactions between natural, sexual, and sexually antagonistic selection.

References and Notes

1. N. Eldredge, J. N. Thompson, P. M. Brakefield, S. Gavrilets, D. Jablonski, J. B. C. Jackson, R. E. Lenski, B. S. Lieberman, M. A. McPeek, W. Miller, The dynamics of evolutionary stasis. *Paleobiology* **31**, 133–145 (2005).

2. D. Berner, W. Salzburger, The genomics of organismal diversification illuminated by adaptive radiations. *Trends Genet.* **31**, 491–499 (2015).

3. T. C. Nelson, W. A. Cresko, Ancient genomic variation underlies repeated ecological adaptation in young stickleback populations. *Evolution Letters* **2**, 9–21 (2018).

4. J. I. Meier, D. A. Marques, S. Mwaiko, C. E. Wagner, L. Excoffier, O. Seehausen, Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun.* **8**, 14363 (2017).

5. D. A. Marques, J. I. Meier, O. Seehausen, A Combinatorial View on Speciation and Adaptive Radiation. *Trends Ecol. Evol.* **34**, 531–544 (2019).

6. D. Roze, N. H. Barton, The Hill–Robertson Effect and the Evolution of Recombination. *Genetics* **173**, 1793–1811 (2006).

7. M. Kirkpatrick, N. Barton, Chromosome inversions, local adaptation and speciation. *Genetics* **173**, 419–434 (2006).

8. A. Akerman, R. Bürger, The consequences of gene flow for local adaptation and differentiation: a two-locus two-deme model. *J. Math. Biol.* **68**, 1135–1198 (2014).

9. R. K. Butlin, Recombination and speciation. *Mol. Ecol.* **14**, 2621–2635 (2005).

10. M. Kirkpatrick, How and why chromosome inversions evolve. *PLoS Biol.* **8** (2010).

11. M. Wellenreuther, L. Bernatchez, Eco-Evolutionary Genomics of Chromosomal Inversions. *Trends Ecol. Evol.* **33**, 427–440 (2018).

12. M. Kirkpatrick, B. Barrett, Chromosome inversions, adaptive cassettes and the evolution of species' ranges. *Mol. Ecol.* **24**, 2046–2055 (2015).

13. E. R. Hager, O. S. Harringmeyer, T. B. Wooldridge, S. Theingi, J. T. Gable, S. McFadden, B. Neugeboren, K. M. Turner, J. D. Jensen, H. E. Hoekstra, A chromosomal inversion contributes to divergence in multiple traits between deer mouse ecotypes. *Science* **377**, 399–405 (2022).

14. F. C. Jones, M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, E. Birney, S. Searle, J. Schmutz, J. Grimwood, M. C. Dickson, R. M. Myers, C. T. Miller, B. R. Summers, A. K. Knecht, S. D. Brady, H. Zhang, A. A. Pollen, T. Howes, C. Amemiya, Broad Institute Genome Sequencing Platform & Whole Genome Assembly Team, J. Baldwin, T. Bloom, D. B. Jaffe, R. Nicol, J. Wilkinson, E. S. Lander, F. Di Palma, K. Lindblad-Toh, D. M. Kingsley, The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55–61 (2012).

15. R. Faria, A. Navarro, Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends Ecol. Evol.* **25**, 660–669 (2010).

16. D. Ayala, R. F. Guerrero, M. Kirkpatrick, Reproductive Isolation and Local Adaptation Quantified for a Chromosome Inversion in a Malaria Mosquito. *Evolution* **67**, 946–958 (2013).

17. M. Joron, R. Papa, M. Beltrán, N. Chamberlain, J. Mavárez, S. Baxter, M. Abanto, E. Bermingham, S. J. Humphray, J. Rogers, H. Beasley, K. Barlow, R. H. ffrench-Constant, J. Mallet, W. O. McMillan, C. D. Jiggins, A conserved supergene locus controls colour pattern diversity in Heliconius butterflies. *PLoS Biol.* **4**, e303 (2006).

18. K. Kunte, W. Zhang, A. Tenger-Trolander, D. H. Palmer, A. Martin, R. D. Reed, S. P. Mullen, M. R. Kronforst, doublesex is a mimicry supergene. *Nature* **507**, 229–232 (2014).

19. C. Küpper, M. Stocks, J. E. Risse, N. Dos Remedios, L. L. Farrell, S. B. McRae, T. C. Morgan, N. Karlionova, P. Pinchuk, Y. I. Verkuil, A. S. Kitaysky, J. C. Wingfield, T. Piersma, K. Zeng, J. Slate, M. Blaxter, D. B. Lank, T. Burke, A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.* **48**, 79–83 (2016).

20. J. Wang, Y. Wurm, M. Nipitwattanaphon, O. Riba-Grognuz, Y.-C. Huang, D. Shoemaker, L. Keller, A Y-like social chromosome causes alternative colony organization in fire ants. *Nature* **493**, 664–668 (2013).

21. P. R. Berg, B. Star, C. Pampoulie, M. Sodeland, J. M. I. Barth, H. Knutsen, K. S. Jakobsen, S. Jentoft, Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Sci. Rep.* **6**, 23246 (2016).

22. M. A. F. Noor, K. L. Grams, L. A. Bertucci, J. Reiland, Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences* **98**, 12084–12088

(2001).

23. R. Castiglia, Sympatric sister species in rodents are more chromosomally differentiated than allopatric ones: implications for the role of chromosomal rearrangements in speciation. *Mammal Review* **44**, 1–4 (2014).

24. D. M. Hooper, T. D. Price, Chromosomal inversion differences correlate with range overlap in passerine birds. *Nature Ecology & Evolution* **1**, 1526–1534 (2017).

25. C.-J. Rubin, E. D. Enbody, M. P. Dobreva, A. Abzhanov, B. W. Davis, S. Lamichhaney, M. Pettersson, A. T. Sendell-Price, C. G. Sprehn, C. A. Valle, K. Vasco, O. Wallerman, B. R. Grant, P. R. Grant, L. Andersson, Rapid adaptive radiation of Darwin's finches depends on ancestral genetic modules. *Sci Adv* **8**, eabm5982 (2022).

26. D. Brawand, C. E. Wagner, Y. I. Li, M. Malinsky, I. Keller, S. Fan, O. Simakov, A. Y. Ng, Z. W. Lim, E. Bezault, J. Turner-Maier, J. Johnson, R. Alcazar, H. J. Noh, P. Russell, B. Aken, J. Alföldi, C. Amemiya, N. Azzouzi, J.-F. Baroiller, F. Barloy-Hubler, A. Berlin, R. Bloomquist, K. L. Carleton, M. A. Conte, H. D'Cotta, O. Eshel, L. Gaffney, F. Galibert, H. F. Gante, S. Gnerre, L. Greuter, R. Guyon, N. S. Haddad, W. Haerty, R. M. Harris, H. A. Hofmann, T. Hourlier, G. Hulata, D. B. Jaffe, M. Lara, A. P. Lee, I. MacCallum, S. Mwaiko, M. Nikaido, H. Nishihara, C. Ozouf-Costaz, D. J. Penman, D. Przybylski, M. Rakotomanga, S. C. P. Renn, F. J. Ribeiro, M. Ron, W. Salzburger, L. Sanchez-Pulido, M. E. Santos, S. Searle, T. Sharpe, R. Swofford, F. J. Tan, L. Williams, S. Young, S. Yin, N. Okada, T. D. Kocher, E. A. Miska, E. S. Lander, B. Venkatesh, R. D. Fernald, A. Meyer, C. P. Ponting, J. T. Streelman, K. Lindblad-Toh, O. Seehausen, F. Di Palma, The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**, 375–381 (2014).

27. E. D. Enbody, A. T. Sendell-Price, C. G. Sprehn, C.-J. Rubin, P. M. Visscher, B. R. Grant, P. R. Grant, L. Andersson, Community-wide genome sequencing reveals 30 years of Darwin's finch evolution. *Science* **381**, eadf6218 (2023).

28. G. Barlow, *The Cichlid Fishes: Nature's Grand Experiment In Evolution* (Perseus Books Group, 2000).

29. A. Konings, *Malawi Cichlids in Their Natural Habitat* (Cichlid Press, 2016).

30. M. Malinsky, H. Svardal, A. M. Tyers, E. A. Miska, M. J. Genner, G. F. Turner, R. Durbin, Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. *Nat Ecol Evol*, doi: 10.1038/s41559-018-0717-x (2018).

31. H. Svardal, F. X. Quah, M. Malinsky, B. P. Ngatunga, E. A. Miska, W. Salzburger, M. J. Genner, G. F. Turner, R. Durbin, Ancestral Hybridization Facilitated Species Diversification in the Lake Malawi Cichlid Fish Adaptive Radiation. *Mol. Biol. Evol.* **37**, 1100–1113 (2020).

32. R. B. Stelkens, C. Schmid, O. Selz, O. Seehausen, Phenotypic novelty in experimental hybrids is predicted by the genetic distance between species of cichlid fish. *BMC Evol. Biol.* **9**, 283 (2009).

33. R. A. York, C. Patil, K. Abdilleh, Z. V. Johnson, M. A. Conte, M. J. Genner, P. T. McGrath, H. B.

Fraser, R. D. Fernald, J. Todd Streelman, Behavior-dependent cis regulation reveals genes and pathways associated with bower building in cichlid fishes. *Proceedings of the National Academy of Sciences* **115**, E11081–E11090 (2018).

34. M. A. Conte, R. Joshi, E. C. Moore, S. P. Nandamuri, W. J. Gammerdinger, R. B. Roberts, K. L. Carleton, S. Lien, T. D. Kocher, Chromosome-scale assemblies reveal the structural evolution of African cichlid genomes. *Gigascience* **8** (2019).

35. O. S. Harringmeyer, H. E. Hoekstra, Chromosomal inversion polymorphisms shape the genomic landscape of deer mice. *Nat Ecol Evol* **6**, 1965–1979 (2022).

36. P. Jay, M. Chouteau, A. Whibley, H. Bastide, H. Parrinello, V. Llaurens, M. Joron, Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms. *Nat. Genet.* **53**, 288–293 (2021).

37. A. Rhie, S. A. McCarthy, O. Fedrigo, J. Damas, G. Formenti, S. Koren, M. Uliano-Silva, W. Chow, A. Fungtammasan, G. L. Gedman, L. J. Cantin, F. Thibaud-Nissen, L. Haggerty, C. Lee, B. J. Ko, J. Kim, I. Bista, M. Smith, B. Haase, J. Mountcastle, S. Winkler, S. Paez, J. Howard, S. C. Vernes, T. M. Lama, F. Grutzner, W. C. Warren, C. Balakrishnan, D. Burt, J. M. George, M. Biegler, D. Iorns, A. Digby, D. Eason, T. Edwards, M. Wilkinson, G. Turner, A. Meyer, A. F. Kautt, P. Franchini, H. William Detrich, H. Svardal, M. Wagner, G. J. P. Naylor, M. Pippel, M. Malinsky, M. Mooney, M. Simbirsky, B. T. Hannigan, T. Pesout, M. Houck, A. Misuraca, S. B. Kingan, R. Hall, Z. Kronenberg, J. Korlach, I. Sović, C. Dunn, Z. Ning, A. Hastie, J. Lee, S. Selvaraj, R. E. Green, N. H. Putnam, J. Ghurye, E. Garrison, Y. Sims, J. Collins, S. Pelan, J. Torrance, A. Tracey, J. Wood, D. Guan, S. E. London, D. F. Clayton, C. V. Mello, S. R. Friedrich, P. V. Lovell, E. Osipova, F. O. Al-Ajli, S. Secomandi, H. Kim, C. Theofanopoulou, Y. Zhou, R. S. Harris, K. D. Makova, P. Medvedev, J. Hoffman, P. Masterson, K. Clark, F. Martin, K. Howe, P. Flicek, B. P. Walenz, W. Kwak, H. Clawson, M. Diekhans, L. Nassar, B. Paten, R. H. S. Kraus, H. Lewin, A. J. Crawford, M. T. P. Gilbert, G. Zhang, B. Venkatesh, R. W. Murphy, K.-P. Koepfli, B. Shapiro, W. E. Johnson, F. Di Palma, T. Margues-Bonet, E. C. Teeling, T. Warnow, J. M. Graves, O. A. Ryder, D. Hausler, S. J. O'Brien, K. Howe, E. W. Myers, R. Durbin, A. M. Phillippy, E. D. Jarvis, Towards complete and error-free genome assemblies of all vertebrate species, *Cold Spring Harbor Laboratory* (2020)p. 2020.05.22.110833.

38. H. Li, P. Ralph, Local PCA Shows How the Effect of Population Structure Differs Along the Genome. *Genetics* **211**, 289–304 (2019).

39. L. S. Stevison, K. B. Hoehn, M. A. F. Noor, Effects of inversions on within- and between-species recombination and divergence. *Genome Biol. Evol.* **3**, 830–841 (2011).

40. J. I. Meier, P. A. Salazar, M. Kučka, R. W. Davies, A. Dréau, I. Aldás, O. B. Power, N. J. Nadeau, J. R. Bridle, C. Rolian, N. H. Barton, W. Owen McMillan, C. D. Jiggins, Y. F. Chan, Haplotype tagging reveals parallel formation of hybrid races in two butterfly species. https://doi.org/10.1101/2020.05.25.113688.

41. D. Bachtrog, Y-chromosome evolution: emerging insights into processes of Y-chromosome

degeneration. *Nat. Rev. Genet.* **14**, 113–124 (2013).

42. C. L. Peichel, S. R. McCann, J. A. Ross, A. F. S. Naftaly, J. R. Urton, J. N. Cech, J. Grimwood, J. Schmutz, R. M. Myers, D. M. Kingsley, M. A. White, Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. *Genome Biol.* **21**, 177 (2020).

43. A. F. Feller, V. Ogi, O. Seehausen, J. I. Meier, Identification of a novel sex determining chromosome in cichlid fishes that acts as XY or ZW in different lineages. *Hydrobiologia* **848**, 3727–3745 (2021).

44. M. D. McGee, S. R. Borstein, J. I. Meier, D. A. Marques, S. Mwaiko, A. Taabu, M. A. Kishe, B. O'Meara, R. Bruggmann, L. Excoffier, O. Seehausen, The ecological and genomic basis of explosive adaptive radiation. *Nature* **586**, 75–79 (2020).

45. A. El Taher, F. Ronco, M. Matschiner, W. Salzburger, A. Böhne, Dynamics of sex chromosome evolution in a rapid radiation of cichlid fishes. *Sci Adv* **7**, eabe8215 (2021).

46. W. J. Gammerdinger, T. D. Kocher, Unusual Diversity of Sex Chromosomes in African Cichlid Fishes. *Genes* **9** (2018).

47. N. Patterson, P. Moorjani, Y. Luo, S. Mallick, N. Rohland, Y. Zhan, T. Genschoreck, T. Webster, D. Reich, Ancient Admixture in Human History. *Genetics* **192**, 1065–1093 (2012).

48. N. Galtier, Half a Century of Controversy: The Neutralist/Selectionist Debate in Molecular Evolution. *Genome Biol. Evol.* **16** (2024).

49. Z. Musilova, F. Cortesi, M. Matschiner, W. I. L. Davies, J. S. Patel, S. M. Stieb, F. de Busserolles, M. Malmstrøm, O. K. Tørresen, C. J. Brown, J. K. Mountford, R. Hanel, D. L. Stenkamp, K. S. Jakobsen, K. L. Carleton, S. Jentoft, J. Marshall, W. Salzburger, Vision using multiple distinct rod opsins in deep-sea fishes. *Science* **364**, 588–592 (2019).

50. A. Brown, S. Thatje, Explaining bathymetric diversity patterns in marine benthic invertebrates and demersal fishes: physiological contributions to adaptation of life at depth. *Biol. Rev. Camb. Philos. Soc.* **89**, 406–426 (2014).

51. H.-S. Yim, Y. S. Cho, X. Guang, S. G. Kang, J.-Y. Jeong, S.-S. Cha, H.-M. Oh, J.-H. Lee, E. C. Yang, K. K. Kwon, Y. J. Kim, T. W. Kim, W. Kim, J. H. Jeon, S.-J. Kim, D. H. Choi, S. Jho, H.-M. Kim, J. Ko, H. Kim, Y.-A. Shin, H.-J. Jung, Y. Zheng, Z. Wang, Y. Chen, M. Chen, A. Jiang, E. Li, S. Zhang, H. Hou, T. H. Kim, L. Yu, S. Liu, K. Ahn, J. Cooper, S.-G. Park, C. P. Hong, W. Jin, H.-S. Kim, C. Park, K. Lee, S. Chun, P. A. Morin, S. J. O'Brien, H. Lee, J. Kimura, D. Y. Moon, A. Manica, J. Edwards, B. C. Kim, S. Kim, J. Wang, J. Bhak, H. S. Lee, J.-H. Lee, Minke whale genome and aquatic adaptation in cetaceans. *Nat. Genet.* **46**, 88–92 (2014).

52. M. Malinsky, R. J. Challis, A. M. Tyers, S. Schiffels, Y. Terai, B. P. Ngatunga, E. A. Miska, R. Durbin, M. J. Genner, G. F. Turner, Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science* **350**, 1493–1498 (2015).

53. C. Hahn, M. J. Genner, G. F. Turner, D. A. Joyce, The genomic basis of cichlid fish adaptation within the deepwater "twilight zone" of Lake Malawi. *Evol Lett* **1**, 184–198 (2017).

54. A. Sur, Y. Wang, P. Capar, G. Margolin, J. A. Farrell, Single-cell analysis of shared signatures and transcriptional diversity during zebrafish development. *bioRxiv*, doi: 10.1101/2023.03.20.533545 (2023).

55. D. Escobar-Camacho, K. L. Carleton, Sensory modalities in cichlid fish behavior. *Curr Opin Behav Sci* **6**, 115–124 (2015).

56. O. Seehausen, Y. Terai, I. S. Magalhaes, K. L. Carleton, H. D. J. Mrosso, R. Miyagi, I. van der Sluijs, M. V. Schneider, M. E. Maan, H. Tachida, H. Imai, N. Okada, Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–626 (2008).

57. O. M. Selz, M. E. R. Pierotti, M. E. Maan, C. Schmid, O. Seehausen, Female preference for male color is necessary and sufficient for assortative mating in 2 cichlid sister species. *Behav. Ecol.* **25**, 612–626 (2014).

58. M. E. Maan, O. Seehausen, J. J. M. Van Alphen, Female mating preferences and male coloration covary with water transparency in a Lake Victoria cichlid fish. *Biol. J. Linn. Soc. Lond.* **99**, 398–406 (2010).

59. I. S. Magalhaes, A. M. Smith, D. A. Joyce, Quantifying mating success of territorial males and sneakers in a bower-building cichlid fish. *Sci. Rep.* **7**, 41128 (2017).

60. S. L. Mederos, R. C. Duarte, M. Mastoras, M. Y. Dennis, M. L. Settles, A. R. Lau, A. Scott, K. Woodward, C. Johnson, A. M. H. Seelke, K. L. Bales, Effects of pairing on color change and central gene expression in lined seahorses. *Genes Brain Behav.* **21**, e12812 (2022).

61. Z. V. Johnson, B. E. Hegarty, G. W. Gruenhagen, T. J. Lancaster, P. T. McGrath, J. T. Streelman, Cellular profiling of a recently-evolved social behavior in cichlid fishes. *Nat. Commun.* **14**, 4891 (2023).

62. A. E. Wright, J. E. Mank, The scope and strength of sex-specific selection in genome evolution. *J. Evol. Biol.* **26**, 1841–1853 (2013).

63. P. Muralidhar, Mating preferences of selfish sex chromosomes. *Nature* **570**, 376–379 (2019).

64. S. H. Smith, K. Hsiung, A. Böhne, Evaluating the role of sexual antagonism in the evolution of sex chromosomes: new data from fish. *Curr. Opin. Genet. Dev.* **81**, 102078 (2023).

65. S. Maezawa, M. Yukawa, K. Hasegawa, R. Sugiyama, M. Iizuka, M. Hu, A. Sakashita, M. Vidal, H. Koseki, A. Barski, T. DeFalco, S. H. Namekawa, PRC1 suppresses a female gene regulatory network to ensure testicular differentiation. *Cell Death Dis.* **14**, 501 (2023).

66. R. P. Piprek, M. Kolasa, D. Podkowa, M. Kloc, J. Z. Kubiak, Cell adhesion molecules expression pattern indicates that somatic cells arbitrate gonadal sex of differentiating bipotential fetal mouse

gonad. *Mech. Dev.* **147**, 17–27 (2017).

67. V. Padovano, I. Lucibello, V. Alari, P. Della Mina, A. Crespi, I. Ferrari, M. Recagni, D. Lattuada, M. Righi, D. Toniolo, A. Villa, G. Pietrini, The POF1B candidate gene for premature ovarian failure regulates epithelial polarity. *J. Cell Sci.* **124**, 3356–3368 (2011).

68. A. Pask, M. B. Renfree, J. A. Marshall Graves, The human sex-reversing ATRX gene has a homologue on the marsupial Y chromosome, ATRY: implications for the evolution of mammalian sex determination. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13198–13202 (2000).

69. Q. Zhou, D. Bachtrog, Chromosome-wide gene silencing initiates Y degeneration in Drosophila. *Curr. Biol.* **22**, 522–525 (2012).

70. R. De-Kayne, O. M. Selz, D. A. Marques, D. Frei, O. Seehausen, P. G. D. Feulner, Genomic architecture of adaptive radiation and hybridization in Alpine whitefish. *Nat. Commun.* **13**, 4479 (2022).

71. H. Svardal, W. Salzburger, M. Malinsky, Genetic Variation and Hybridization in Evolutionary Radiations of Cichlid Fishes. *Annu Rev Anim Biosci* **9**, 55–79 (2021).

72. B. S. Meyer, M. Matschiner, W. Salzburger, Disentangling Incomplete Lineage Sorting and Introgression to Refine Species-Tree Estimates for Lake Tanganyika Cichlid Fishes. *Syst. Biol.* **66**, 531–550 (2017).

73. I. Irisarri, P. Singh, S. Koblmüller, J. Torres-Dowdall, F. Henning, P. Franchini, C. Fischer, A. R. Lemmon, E. M. Lemmon, G. G. Thallinger, C. Sturmbauer, A. Meyer, Phylogenomics uncovers early hybridization and adaptive loci shaping the radiation of Lake Tanganyika cichlid fishes. *Nat. Commun.* **9**, 3159 (2018).

74. J. I. Meier, M. D. McGee, D. A. Marques, S. Mwaiko, M. Kishe, S. Wandera, D. Neumann, H. Mrosso, L. J. Chapman, C. A. Chapman, L. Kaufman, A. Taabu-Munyaho, C. E. Wagner, R. Bruggmann, L. Excoffier, O. Seehausen, Cycles of fusion and fission enabled rapid parallel adaptive radiations in African cichlids. *Science* **381**, eade2833 (2023).

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

76. J. Kitano, S. Ansai, S. Fujimoto, R. Kakioka, M. Sato, I. F. Mandagi, B. K. A. Sumarto, K. Yamahira, A Cryptic Sex-Linked Locus Revealed by the Elimination of a Master Sex-Determining Locus in Medaka Fish. *Am. Nat.* **202**, 231–240 (2023).

77. P. Jay, E. Tezenas, A. Véber, T. Giraud, Sheltering of deleterious mutations explains the stepwise extension of recombination suppression on sex chromosomes and other supergenes. *PLoS Biol.* **20**, e3001698 (2022).

78. C. E. Wagner, L. J. Harmon, O. Seehausen, Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* **487**, 366–369 (2012).

79. M. Plenderleith, C. van Oosterhout, R. L. Robinson, G. F. Turner, Female preference for conspecific males based on olfactory cues in a Lake Malawi cichlid fish. *Biol. Lett.* **1**, 411–414 (2005).

80. K. P. Maruska, U. S. Ung, R. D. Fernald, The African cichlid fish Astatotilapia burtoni uses acoustic communication for reproduction: sound production, hearing, and behavioral significance. *PLoS One* **7**, e37612 (2012).

81. H. D. Rundle, P. Nosil, Ecological speciation. *Ecol. Lett.* **8**, 336–352 (2005).

82. M. Talbi, G. F. Turner, M. Malinsky, Rapid evolution of recombination landscapes during the divergence of cichlid ecotypes in Lake Masoko, *bioRxiv* (2024). https://doi.org/10.1101/2024.03.20.585960.

Acknowledgments

We thank the Department of Fisheries of the Government of Malawi for facilitating Malawi cichlid specimen collection and Martin Genner, Alix Tyers, Mingliu Du, Mexford Mulumpwa, Martin Genner, Gregoire Vernaz, Magnus Nordborg, Karl Svardal, Alexander Hooft van Huysduynen and the staff of the Monkey Bay Fisheries Research Unit for support with sampling, Natalie Van Houtte for support with lab work, Shane McCarthy for bioinformatics support, Walter Salzburger for access to facilities and data, and Patrick Gemmell for input on the project. We are grateful to Daniel Shaw, Regev Schweiger, Victoria Caudill, Peter Ralph, Jonas Lescroart, and Els De Keyzer for useful comments on the manuscript.

Funding

The authors gratefully acknowledge support through the Research Foundation – Flanders (FWO) (G047521N to H.S.), the Wellcome Trust (Wellcome grant 207492 to R.D., Wellcome Senior Investigator Award 219475/Z/19/Z to E.A.M), the Research Fund of the University of Antwerp (BOF) (to H.S.), the Cambridge-Africa ALBORADA Research Fund (to H.S. and B.R.), the German Research Foundation (DFG) (492407022 and 497674620 to A.B.), the European Research Council (Proof-of-Concept Grant 101069219 to M. K. and Y.F.C.), the Max Planck Society (to Y.F.C.), and the Natural Environment Research Council (NERC) (IRF NE/R01504X/1 to M.E.S.). L.M.B. was supported through a Harding Distinguished Postgraduate Scholarship and S.G. through an FWO PhD Fellowship Fundamental Research. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Authors contributions

H.S. and R.D. conceived the study, with support from L.M.B. H.S. organised cichlid collections in Malawi with support from B.R., R.Z., W.S., J.C.G, M.N., M.M., R.D., E.A.M., and G.F.T. G.F.T. undertook final taxonomic assignment. I.B. and V.B. performed DNA extractions. T.L. produced the original variant callset. H.S., L.M.B., V.B., I.A., J.S., and J.C.G. analysed the SNP data. Inversion detection was accomplished by L.M.B. (PCA-based) and I.A. (clusterization,

phylogenetic reconstructions). H.S. performed hybridisation analysis. V.B. made selection tests (together with H.S.), zebrafish-based functional analysis and candidate genes investigation. J.C.G. performed enrichment analysis and candidate gene investigation. J.S. and H.S. performed population history reconstructions. J.E., A.H., M.E.S., and G.F.T. performed the species cross. B.F. designed the PCR essays and provided molecular lab support. V.B. performed PCR typing. C.Z., L.M.B. and B.F. produced assemblies and analysed long read and Hi-C data with support from J.S.. L.M.B. detected sex-linked inversions in the Lake Victoria radiation. M.K. and Y.F.C. performed haplotagging. I.A. analysed haplotagging data. W.S., I.A., and H.S. analysed RNA-seq data. S.L. and F.Y. performed FISH. S.G. produced and analysed the outgroup variant callset with support from H.S. M.M. produced Hi-C data and the ancestral sequence. D.A.J. contributed the *Diplotaxodon* long read samples. A.B. provided support for sex chromosome and adaptation related analyses. L.M.B. designed the main figures, with input from H.S., V.B., J.C. and I.A.. H.S. wrote the initial manuscript. H.S., R.D., L.M.B, V.B., and J.C. edited the manuscript with contributions from all authors. All authors read and approved the final manuscript.

Competing interests

The authors declare that there are no competing interests.

Data and materials availability

Supporting data is made available in the online supplementary on an open-access basis for research use only. ENA Accession numbers of raw sequencing data <*will be added to Supplementary table 1 before final publication*>. VCF are shared in Dryad repository <*DOI added before final publication*>. Data was collected under appropriate ethical and sampling permits and genetic material and sequences are subject to an Access and Benefits Sharing (ABS) agreement with the Government of Malawi. Any person who wishes to use this data for any form of commercial purpose must first enter into a commercial licensing and benefit-sharing arrangement with the Government of Malawi.

Supplementary Materials

Materials and Methods

Supplementary Text

Figs. S1 to S45

Tables S1 to S23

References (82-164)

Data S1 to S22