

## Research



**Cite this article:** Feller AF, Haesler MP, Peichel CL, Seehausen O. 2020 Genetic architecture of a key reproductive isolation trait differs between sympatric and non-sympatric sister species of Lake Victoria cichlids. *Proc. R. Soc. B* **287**: 20200270. <http://dx.doi.org/10.1098/rspb.2020.0270>

Received: 10 February 2020

Accepted: 17 March 2020

**Subject Category:**

Evolution

**Subject Areas:**

evolution, genetics

**Keywords:**

Lake Victoria, cichlids, sympatric speciation, genetic architecture, male coloration

**Author for correspondence:**

Anna F. Feller

e-mail: [anna.feller@eawag.ch](mailto:anna.feller@eawag.ch)

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.4902459>.

# Genetic architecture of a key reproductive isolation trait differs between sympatric and non-sympatric sister species of Lake Victoria cichlids

Anna F. Feller<sup>1,3</sup>, Marcel P. Haesler<sup>1,3</sup>, Catherine L. Peichel<sup>2</sup> and Ole Seehausen<sup>1,3</sup>

<sup>1</sup>Division of Aquatic Ecology and Evolution, Institute of Ecology and Evolution, and <sup>2</sup>Division of Evolutionary Ecology, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland

<sup>3</sup>Department of Fish Ecology and Evolution, Centre of Ecology, Evolution and Biogeochemistry, EAWAG Swiss Federal Institute of Aquatic Science and Technology, 6047 Kastanienbaum, Switzerland

AFF, 0000-0001-5786-7658; CLP, 0000-0002-7731-8944; OS, 0000-0001-6598-1434

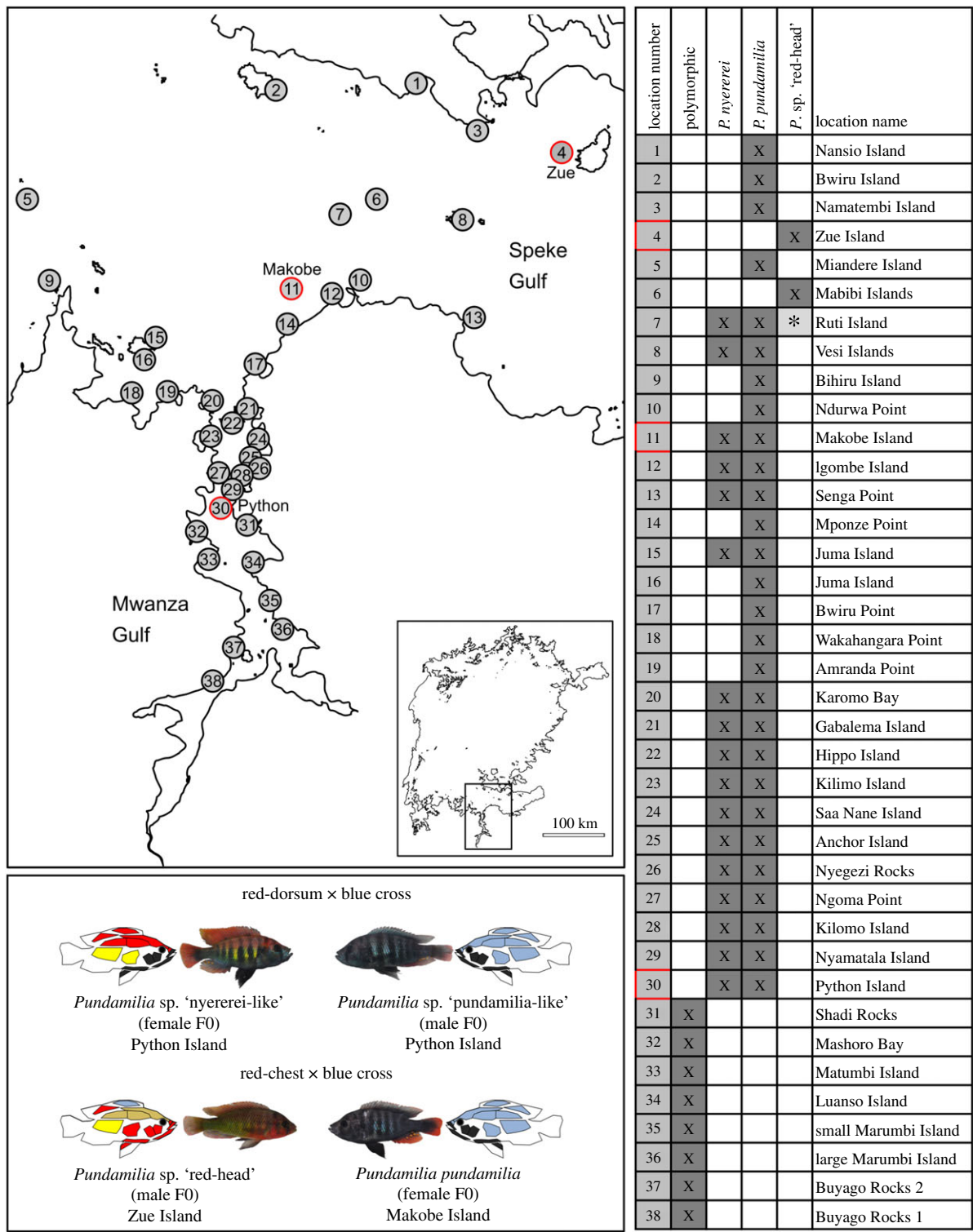
One hallmark of the East African cichlid radiations is the rapid evolution of reproductive isolation that is robust to full sympatry of many closely related species. Theory predicts that species persistence and speciation in sympatry with gene flow are facilitated if loci of large effect or physical linkage (or pleiotropy) underlie traits involved in reproductive isolation. Here, we investigate the genetic architecture of a key trait involved in behavioural isolation, male nuptial coloration, by crossing two sister species pairs of Lake Victoria cichlids of the genus *Pundamilia* and mapping nuptial coloration in the F2 hybrids. One is a young sympatric species pair, representative of an axis of colour motif differentiation, red-dorsum versus blue, that is highly recurrent in closely related sympatric species. The other is a species pair representative of colour motifs, red-chest versus blue, that are common in allopatric but uncommon in sympatric closely related species. We find significant quantitative trait loci (QTLs) with moderate to large effects (some overlapping) for red and yellow in the sympatric red-dorsum × blue cross, whereas we find no significant QTLs in the non-sympatric red-chest × blue cross. These findings are consistent with theory predicting that large effect loci or linkage/pleiotropy underlying mating trait differentiation could facilitate speciation and species persistence with gene flow in sympatry.

## 1. Background

The adaptive radiation of Lake Victoria haplochromine cichlids comprises approximately 500 endemic species that have evolved within the lake in perhaps as little as 15 000 years [1–4] and that are highly diverse in morphology, ecology, colour and behaviour [5–7]. Typically, however, closely related species are similar in morphology and ecology while they differ dramatically in male nuptial coloration [6,8,9]. Male nuptial coloration is considered a trait of key importance in the origin and maintenance of new species in these fish [9–11].

A highly recurrent pattern in male nuptial colour variation in pairs of closely related species of Lake Victoria cichlids is that males in one species are blue-grey on their body with any red colour confined to the fins, whereas males of the other species are yellow-red on the body [9–12]. The red colour can be confined to either dorsal parts of the body ('red-dorsum' type) or to the chest and lower head ('red-chest' type). When closely related species are sympatric, the red form generally has a red-dorsum, whereas many non-sympatric pairs of closely related species involve a red-chest and a blue form [6,8,13].

A representative case of sympatric red-dorsum and blue sister species is the young species pair of *Pundamilia* sp. 'nyererei-like' and *Pundamilia* sp.



**Figure 1.** Distribution and sampling locations of blue, red-dorsum and red-chest *Pundamilias*. The upper left panel shows the Mwanza gulf region in Lake Victoria (the inset figure indicates the location of this region on the full Lake Victoria map) with all sampling locations (circles with numbers). The table on the right indicates which species were recorded at each sampling location (data assembled from [6,8,13,14]). *P. pundamilia* records include records of *P. sp. 'pundamilia-like'*, and *P. nyererei* records include records of *P. sp. 'nyererei-like'* [15,16]. At Ruti island (\*), one single *P. sp. 'red-head'* male was caught by O.S. in 2001, but no other individuals of this species were seen in many hundreds of sampled individuals at this location in 10 years before or after this. Sampling locations 31–38 harbour colour-polymorphic *Pundamilia* populations, where species are not clearly identifiable. Highlighted in red are the catch locations of the four parental species of the two crosses, see lower left panel showing representative male individuals of these four species and their colour phenotypes. Lake Victoria contours were taken from the Lake Victoria Bathymetric map V6 by [17] and smoothed in Inkscape™ 0.92.

'pundamilia-like' at Python Island (figure 1). This pair may have evolved in sympatry from a hybrid population in only a few hundred generations [15,16,18]. *Pundamilia sp. 'red-head'* and *Pundamilia pundamilia* are representatives of a red-chest and blue species with overlapping geographical distributions that are never found together on the same

rocky island, despite islands occupied by the different species lying within dispersal distance of each other (figure 1) [6,8,19]. Male nuptial colour in *Pundamilia* is under intra- and intersexual selection. Males fiercely compete for territories, and they use their bright colours to signal territory ownership to contestants. Negative frequency-dependent selection

generated by own-type aggression biases is likely involved in stabilizing the coexistence of different colour types [11,20–22]. Females, which are cryptically coloured and the sole investors into parental care, exhibit strong preferences for mates with particular nuptial coloration, which generates both directional sexual selection within and assortative mating between species [23–26]. Preference for male colour has been shown to be heritable and likely determined by few major genes or genomic regions [27,28] and to generate disruptive selection on male colour [29].

Despite strong preferences for bright male colour and assortative mating, there is some gene flow in fully sympatric red-dorsum versus blue pairs such as *Pundamilia* sp. ‘nyererei-like’ and *Pundamilia* sp. ‘pundamilia-like’ [15,16]. Nonetheless these two species persist in sympatry. With a genome-wide mean  $F_{ST}$  of 0.053 and hundreds of highly differentiated genomic regions, they show surprisingly strong genetic differentiation considering they have likely evolved in full sympatry in less than 200 generations [15,16]. Red-chest versus blue pairs like *Pundamilia* sp. ‘red-head’ and *Pundamilia pundamilia* experience little gene flow because they usually do not co-occur. Although they are neither strictly geographically isolated nor dispersal limited [6,8,19] (and figure 1), they do not seem to persist as two species in the same island.

Here we ask whether differences in the genetic architecture of the male nuptial colour motifs that differentiate these species could explain the difference in the distribution patterns of the species pairs. We tested the hypothesis that red-dorsum versus blue pairs persisting in sympatry despite some gene flow have an architecture that is robust against potentially homogenizing effects of gene flow, while the absence of such an architecture would make it difficult for red-chest versus blue pairs to persist in the presence of gene flow.

In a sympatric scenario with ongoing gene flow, recombination is expected to break up linkage disequilibrium between favourable combinations of alleles for local adaptation and reproductive isolation [30,31]. This is less likely to occur if divergently selected traits are coded by few genes with large effects, as this reduces the number of targets for recombination to break up, and increases the effectiveness of (correlational) selection because it is concentrated on fewer targets [32]. Indeed, theoretical work has shown that large effect alleles, or groups of tightly linked alleles with smaller effects that then act like a large effect locus, are less likely to be lost when adaptive multilocus phenotypes need to be maintained against gene flow [33–35]. Similarly, the physical linkage of multiple traits (or pleiotropy) could facilitate divergence with gene flow [34,36,37].

Recombinant males between cichlid (*Pundamilia*) species with either red or blue male nuptial coloration might have reduced fitness if they will be less likely to be chosen by females with strong preferences for either colour, as is the case in these species [23–28], or if intermediate coloration makes them targets of territorial aggression by males of both species. We thus predict large effect loci and/or physical linkage of several loci (or pleiotropy) for male nuptial colour to be present in sympatric red-dorsum versus blue cichlid pairs, as such an architecture could both facilitate the establishment of polymorphisms [38] and make it easier to retain phenotypic differentiation in sympatry with ongoing gene flow (see above). Furthermore, theoretical models that investigate the feasibility of sympatric speciation by sexual selection (alone) usually find such speciation feasible when

assuming a simple genetic architecture (i.e. few additive loci with large effects) and that speciation becomes less likely when the number of loci underpinning reproductive isolation traits increases [39–41].

To compare the genetic architecture of red and yellow versus blue male nuptial colour motifs that do or do not persist with gene flow, we crossed *Pundamilia* sp. ‘nyererei-like’ and *Pundamilia* sp. ‘pundamilia-like’ (red-dorsum × blue) and *Pundamilia* sp. ‘red-head’ and *Pundamilia pundamilia* (red-chest × blue) in the laboratory and performed quantitative trait locus (QTL) mapping analyses on male nuptial coloration in the second generation (F<sub>2</sub>) hybrids.

We find several significant QTLs for the presence/absence of red and yellow colour in the sympatric red-dorsum × blue cross, with several traits mapping to the same region, whereas we find no significant QTLs in the non-sympatric red-chest × blue cross. We conclude that the presence of large effect loci, with physical linkage between some traits (or pleiotropy), likely makes up one key element for the rapid evolution of reproductive isolation and species persistence in sympatry despite some gene flow.

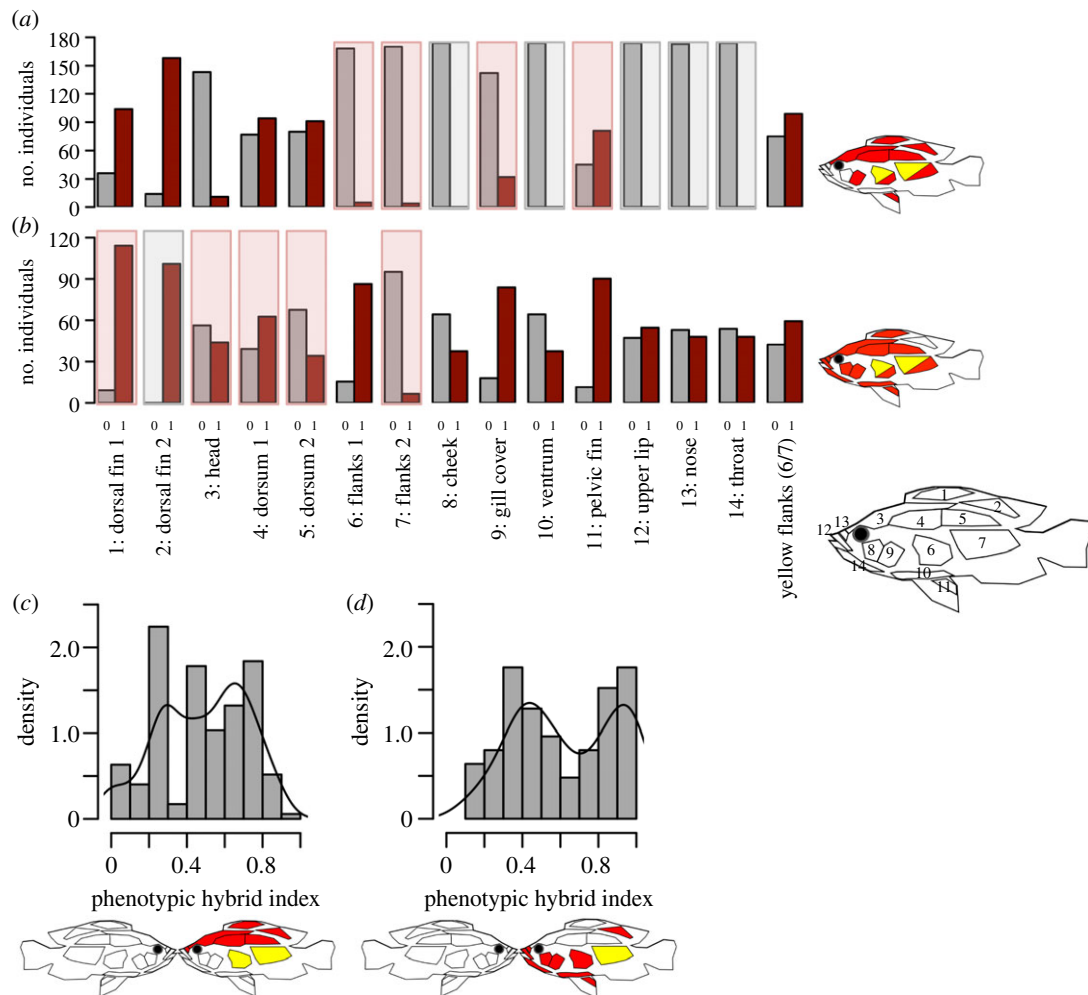
## 2. Material and methods

### (a) Experimental crosses

The red-dorsum × blue cross was started with a *Pundamilia* sp. ‘nyererei-like’ female and a *Pundamilia* sp. ‘pundamilia-like’ male, both from laboratory bred strains established from fishes caught by OS at Python Island in Lake Victoria in 2003 (figure 1). The red-chest × blue cross was started with a *Pundamilia pundamilia* female caught by O.S. at Makobe Island in 2003 and a *Pundamilia* sp. ‘red-head’ male from a laboratory bred strain established from fishes caught by O.S. at Zue Island in 1993 (figure 1). Five to 6 days after spawning, the eggs were removed from the female’s mouth and reared in an egg tumbler until hatching. After yolk sac resorption, the larvae were transferred to rearing aquaria that were part of a large recirculation system. After sexual maturity at the age of 1 to 2 years, two pairs of F<sub>1</sub> individuals of each cross were then allowed to mate, and the eggs and larvae were reared the same way as the F<sub>1</sub> generation. Because the average clutch size is just about 20–30 juveniles, we re-mated each F<sub>1</sub> pair multiple times until we had obtained a total of approximately 300 F<sub>2</sub> individuals in each cross. This procedure of repeatedly re-mating the same pairs took about 2 years. All F<sub>2</sub>s were reared to an age of at least 1 year before they were phenotyped. All four populations from which the grandparents were taken breed true in the laboratory in a common garden environment, hence the differences in coloration are heritable in both crosses.

### (b) Colour photos and scoring of coloration

Sexually mature F<sub>2</sub> males were removed from their holding tank and individually placed into one of six adjacent plexiglas photo cuvette compartments with transparent separations and a grey PVC background, inside a larger aquarium. This set-up induced territoriality in the males who could see each other and hence make them express full colour. Two Walimex pro Daylight 250S lamps were placed on either side of the cuvette and a first colour picture was taken of each fish with a Canon D60 camera equipped with a 50 mm lens (settings; P mode, ISO 200, auto focus). If a male failed to show territorial behaviour within an adjustment time of 1–2 days, it was moved back into its home tank and the process was repeated several weeks later. After good photos were obtained, each fish was sedated in MS222



**Figure 2.** Phenotypic distributions in the two crosses. Distribution of absence (0; grey)/presence (1; red) scores for all traits in the (a) red-dorsum × blue cross and (b) red-chest × blue cross. A grey background indicates lack of variation in the F2 individuals in this trait (i.e. non-mappable traits), a red background indicates traits that are not differentiated between the two parental species but were variable in the F2 hybrids (transgressive traits). The sectors highlighted with colour on the inset cartoon fish are the traits that were mappable in the cross. Half red/half yellow indicates either colour could occur in this trait. (c,d) The distribution of the standardized phenotypic hybrid index, for the red-dorsum × blue cross (c) and for the red-chest × blue cross (d), which was calculated for every individual as the sum of presence scores in all traits differentiating the two parental species divided by the number of differentiating traits (see parts highlighted with colour in the inset fish figures). An index of 0 hence corresponds to the phenotype of the blue species, an index of 1 corresponds to the phenotype of the red species. Note that no red was scored on the nose on any F2 individual in the red-dorsum × blue cross. The comparison of the cartoon fish indicating mappable traits in (a,b) with the cartoon fish showing the parental types in (c,d) highlights which traits are transgressive in the F2 hybrids.

(50 mg l<sup>-1</sup>), euthanized in MS222 (300 mg l<sup>-1</sup>) and immediately put on melting ice for 10 min. This treatment had the effect of relaxation of the melanophores of system I that are under nervous control [42] and made colour better visible. The fish were then pinned down in a dissection bowl containing a wax layer and a ruler with a colour reference, and submerged in water (just covering the fish), and a second colour picture was taken from above with the same camera (same settings), now equipped with an 18–135 mm objective and with one Walimex pro Daylight 250S lamp for lighting. We applied a white balance filter to these second photos in Adobe Photoshop CC 2018 by setting a white point on the white part of the colour bar ruler.

We visually scored the presence/absence of red on 14 sectors of the body (henceforth referred to as traits) that are different between at least two of the parental species (see figures 1 and 2), and the presence/absence of yellow flanks (as 15th trait) from the second photo (when the fish was dead). The first pictures were only used as a cross-reference to confirm uncertain scores. In some individuals (mostly of the red-dorsum × blue cross), some traits could not be scored unambiguously and were scored as NA. In addition, we calculated a phenotypic hybrid index for each individual as the summed up presence scores (NAs treated as 0.5) in the subset of traits that differentiate the two respective parental species divided by the number of

traits in the respective subset (figure 2c,d). The phenotypic hybrid index thus ranges from 0 (all traits like those of the blue species) to 1 (all traits like those of the red species).

### (c) DNA extraction and RAD sequencing

Genomic DNA was extracted from finclips, which had been stored in 98% ethanol, using phenol–chloroform [43]. We prepared restriction-site associated DNA (RAD) sequencing libraries using *SbfI* as restriction enzyme, following [44] with some modifications (see electronic supplementary material, appendix S1). Single end-sequencing (100 bp for all others but 125 bp for the last two libraries, see electronic supplementary material, table S5) was done on an Illumina HiSeq 2500 platform either at the Genomic Technologies Facility of the University of Lausanne or at the Next Generation Sequencing Platform of the University of Bern. Each library was sequenced on a single lane. Bacteriophage PhiX genomic DNA was added to each library (4–12% of reads) to increase complexity in the first 10 sequenced base pairs and for base quality recalibration (see below).

### (d) Sequence processing and genotyping

Demultiplexed, trimmed and filtered reads (see appendix S2 for details) were aligned to the anchored version of the *Pundamilia*

*nyererei* reference genome [45] with Bowtie2 v2.3.2 [46], allowing one mismatch. This was followed by base quality recalibration (see electronic supplementary material, appendix S2 for details) and subsetting to uniquely aligned reads. GATK Unified Genotyper [47] was used for genotyping (minimum base quality score set to 20). The resulting vcf files were filtered with Bcftools implemented in Samtools v.1.8 [48], Vcftools v.0.1.14 [49] and a custom Python script, to obtain bi-allelic SNPs (see electronic supplementary material, appendix S2 for details). In the red-dorsum  $\times$  blue cross, 368 SNPs were homozygous fixed in the F0 and heterozygous in all four F1 parents, and were used in linkage map construction. In the red-chest  $\times$  blue cross, 2358 SNPs were homozygous fixed in the F0 and heterozygous in two F1 parents (the other two F1s had low quality data that was removed during filtering), and were used in linkage map construction.

### (e) Linkage map construction

We used JoinMap 4.0 [50] to build linkage maps for both crosses. We removed loci with extreme segregation distortion ( $p < 0.01$ ), loci with greater than 20% missing genotypes and identical loci (i.e. SNPs within the same RAD locus) (greater than 0.950). Individuals were removed if they had greater than 30% missing data. The linkage maps were generated from 216 F2 individuals (173 males, 43 females) in the red-dorsum  $\times$  blue cross and 171 F2 individuals (115 males, 56 females) in the red-chest  $\times$  blue cross. We identified linkage groups based on an independent logarithm of odds (LOD) threshold of 5. Loci with suspicious linkage (recombination frequency greater than 0.6) were removed. The strongest cross-link (SCL) values in the maps are 4.7 (red-dorsum  $\times$  blue cross) and 4.6 (red-chest  $\times$  blue cross), and unlinked markers were excluded. To build the linkage maps, we used the Kosambi regression mapping algorithm with a LOD threshold of 1.0, a recombination threshold of 0.499, a goodness-of-fit threshold of 5.0 and no fixed order. We performed two rounds of mapping with a ripple after addition of each marker to the map (see [50]).

### (f) QTL mapping

QTL mapping of male nuptial colour traits was performed in R/qlt [51,52]. We mapped the presence/absence of yellow colour on the flanks and of red colour in 14 body and fin locations (traits; see above and figure 2) in F2 males ( $n = 174$  in the red-dorsum  $\times$  blue cross and  $n = 125$  in the red-chest  $\times$  blue cross) as binary traits. Conditional genotype probabilities were calculated using the `calc.genoprob` function with a fixed stepsize of 1 (respectively 3 for `scantwo`; in cM), an assumed genotyping error rate of 0.05, and the Kosambi map function. The `scanone` and `scantwo` functions were used with the EM algorithm, and significance thresholds were determined by permutations ( $n = 1000$ ). For the red and yellow traits, we used the binary model; for a multi-trait hybrid index (see below), we used the normal model. We consider  $p < 0.05$  as significant, and  $p < 0.1$  as marginally significant. For traits showing a significant effect of F1 family, mapping was additionally performed with family as an additive covariate (allowing the average phenotype in the two families to be different) and as both an additive and interactive covariate (additionally allowing the effect of the QTL between the two families to be different). For significant QTLs, approximate Bayesian credible intervals with a 0.95 probability coverage were calculated using the `bayesint` function, and the percentage of variation explained (PVE) of significant QTLs was calculated for each trait individually using the `fitql` function (with the HK algorithm since the EM algorithm was not available in this function for binary models).

The lower number of individuals in the red-chest  $\times$  blue cross could result in reduced power of detecting QTLs of similar effect sizes as compared to the red-dorsum  $\times$  blue cross, and the two linkage maps differ substantially in number and density of markers. To account for this, we repeated the single QTL analyses for the red-

dorsum  $\times$  blue cross after randomly downsampling to 125 (of 174) individuals using the `sample` function in R to randomly pick individuals and then subsetting the genotype–phenotype file to these individuals. For the red-chest  $\times$  blue cross, we repeated the single QTL analyses after randomly downsampling markers on the linkage map to match the number of markers on each linkage group to those in our sparser red-dorsum  $\times$  blue map, again using the `sample` function in R to randomly pick markers within each linkage group and then subsetting them to these markers. The procedure was repeated five times in each cross.

## 3. Results

### (a) Linkage maps

The final map for the red-dorsum  $\times$  blue cross contains 232 markers in 22 linkage groups (corresponding to the number of expected chromosomes in haplo-tilapia cichlids [53,54]) with an average marker distance of 5.3 cM and a total map length of 1117.8 cM. The final map for the red-chest  $\times$  blue cross contains 1198 markers in 22 linkage groups with an average marker distance of 1.2 cM and a total map length of 1360.8 cM.

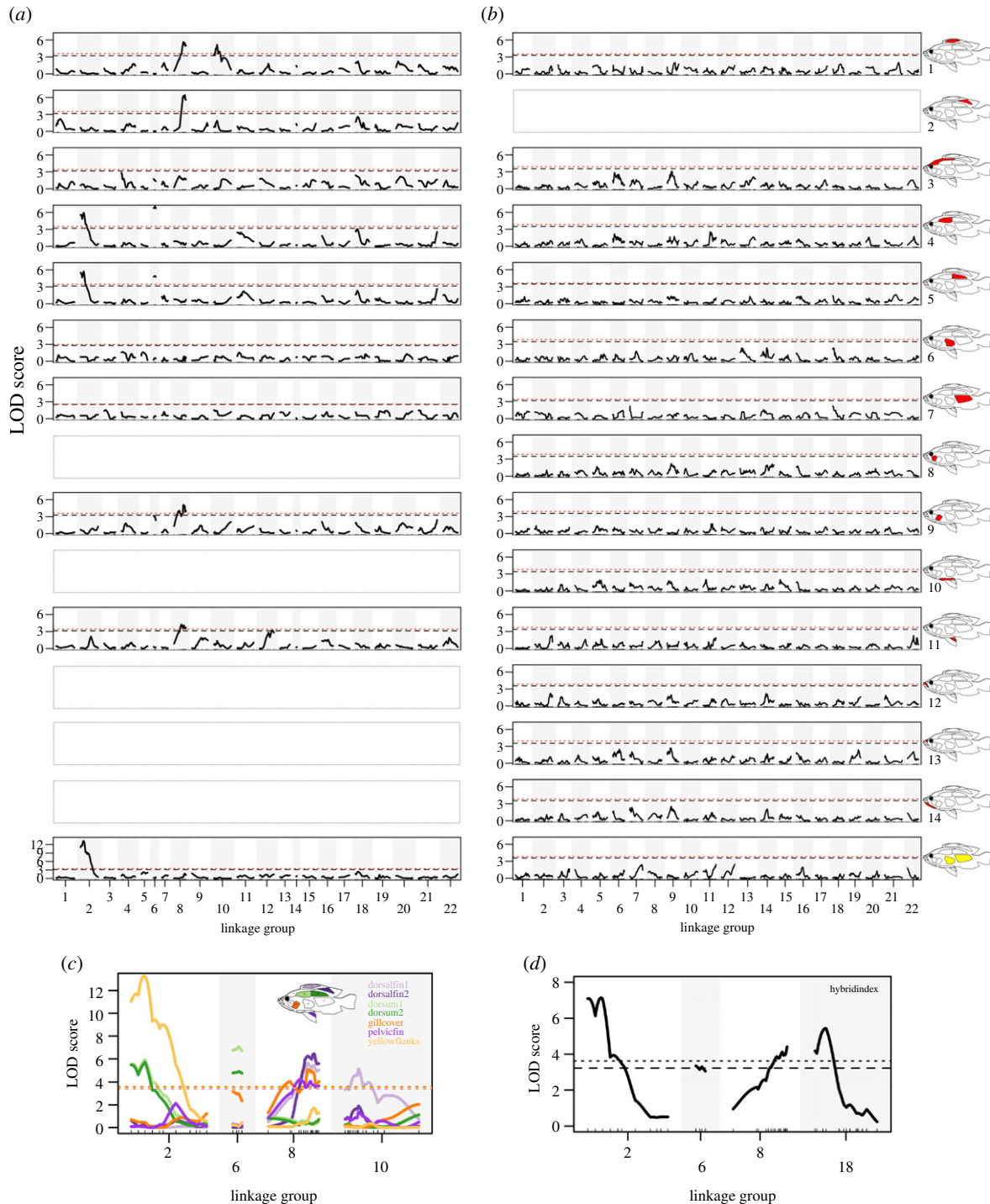
### (b) Phenotypic distributions

Seven traits differentiate the parental species in the red-dorsum  $\times$  blue cross (figure 2c): both parts of the dorsal fin, both parts of the dorsum, head and nose are red versus blue, and both parts of the flank (considered as one trait) are yellow versus blue. However, nose was never scored as red in any F2. Red was also scored in some F2s for both parts of the flank, for gill cover, and for pelvic fin, although neither parental species has red in these traits (figures 1 and 2c). Altogether this resulted in 10 mappable traits for this cross (figure 2a). The F2 phenotypes range from no red or yellow, respectively, in any of the traits that differentiate the parental species (i.e. like *P. sp.* ‘*pundamilia*-like’), to red or yellow, respectively, in all of these traits with the exception of nose (i.e. like *P. sp.* ‘*nyererei*-like’), with most individuals being intermediate in expressing red in some but not all of these traits (figure 2c).

Ten traits differentiate the parental species in the red-chest  $\times$  blue cross (figure 2d): the rear dorsal fin part, the frontal flank part, the pelvic fin and all parts on the head (except for the head part itself, which like the dorsum is greenish versus blue) are red or yellow (flanks) versus blue. The rear dorsal fin part was scored as red in all F2s and can hence not be mapped. Some F2s were scored as red on the front part of the dorsal fin as well as on both parts of the dorsum, on the head and on the rear flank part, even though neither parental species is red there (figures 1 and 2d). Altogether this resulted in 14 mappable traits for this cross (figure 2b). The F2 phenotypes range from no red or yellow, respectively, in any of these traits with the exception of the rear dorsal fin part (i.e. like *P. pundamilia*), to red or yellow, respectively, in all these traits (i.e. like *P. sp.* ‘*red-head*’), with most individuals being intermediate in expressing red in some but not all of these traits. Yet, a large number of these F2 hybrids resemble *P. sp.* ‘*red-head*’ whereas fewer resemble *P. pundamilia*, suggesting more directional dominance effects in this cross than in the other cross.

### (c) QTL mapping results

We found significant QTLs for red and yellow colour for seven out of 10 mappable traits (figure 2a) in the sympatric red-



**Figure 3.** QTL mapping of the presence/absence of red (and yellow) male nuptial colour. LOD scores across the 22 *Pundamilia* chromosomes for all traits with presence scores for red and for yellow flanks (empty plots for non-mappable traits, figure 2) in (a) the red-dorsum × blue cross and (b) the red-chest × blue cross. The black dashed lines represent genome-wide significance thresholds of  $p < 0.1$  for each trait, the red dotted lines for  $p < 0.05$ . (c) LOD scores across the chromosomes containing significant QTLs for red and yellow in the red-dorsum × blue cross (genome-wide significance thresholds of  $p < 0.05$  shown for each trait); (d) shows the chromosomes containing significant QTLs for the hybrid index in the red-dorsum × blue cross. See also electronic supplementary material, table S1 and figure S1.

dorsum × blue cross (electronic supplementary material, table S1 and figure S1; figure 3a,c). For red in both sectors of the dorsal fin, the gill cover and the pelvic fin, we identified QTLs on Pun-LG8 with overlapping 95% confidence intervals, each with a PVE of 12.6–16.8%. A second QTL for red on dorsal fin sector 1 was found on Pun-LG10 with a PVE of 15.5%, and a second marginally significant QTL for red on the pelvic fin was found on Pun-LG12 with a PVE of 10.8%. Two QTLs for red on both dorsum parts were identified, one on Pun-LG2 and one on Pun-LG6, both with a PVE of

12.4–17.4%. Pun-LG2 also contains a QTL for yellow flanks, within the 95% confidence mapping interval for red-dorsum, with a PVE of 29.5%. In this cross, only red in the pelvic fin (a transgressive trait in this cross) showed a significant effect of family ( $F = 26.901$ ,  $p < 0.001$ ). Repeating the mapping with family as covariate for pelvic fin recovered the QTL on Pun-LG8 and revealed an additional marginally significant QTL on Pun-LG22. We also found four QTLs for the phenotypic hybrid index (figure 3d), one each on LG2 (17.2% PVE), LG6 (8.5% PVE), LG8 (10.97% PVE), all of them within the 95%

confidence intervals of the QTLs found for the individual traits and one on LG18 (13.3% PVE).

Two-dimensional two QTL scans revealed several additional putative QTLs in this cross (electronic supplementary material, table S2). Repeating the single QTL mapping with 125 individuals randomly sampled from 174 five times recovered 82% of the expected total of 50 significant QTL results (i.e. 10 significant QTLs in the original dataset times five); 36 were recovered as significant and five as marginally significant (electronic supplementary material, table S3).

We found no significant QTLs for any of the 14 mappable traits (figure 2*b*) nor for the hybrid index in the non-sympatric red-chest × blue cross (electronic supplementary material, table S1; figure 3*b*). Several traits showed a significant effect of family in this cross: cheek ( $F = 16.478$ ,  $p < 0.001$ ), ventrum ( $F = 29.407$ ,  $p < 0.001$ ), upper lip ( $F = 29.079$ ,  $p < 0.001$ ), nose ( $F = 18.232$ ,  $p = 0.001$ ), throat ( $F = 46.281$ ,  $p < 0.001$ ) and yellow flanks ( $F = 5.6352$ ,  $p = 0.025$ ). Repeating the mapping with family as covariate for these traits; however, only detected one marginally significant QTL for red on the throat on Pun-LG7 with a PVE of 8.3%. A QTL for red on the head (a transgressive trait in this cross) reached the 0.1 significance threshold when repeating the mapping with a subsampled linkage map in two out of five such mapping rounds (electronic supplementary material, tables S1 and S4). Two-dimensional two QTL scans for this cross revealed two potentially interacting QTLs each for red on the cheek and throat (Pun-LG1 and Pun-LG4 for both traits) and for yellow on the flanks (Pun-LG3 and Pun-LG11) (electronic supplementary material, table S2).

#### 4. Discussion

We investigated the genetic architecture of a trait complex of key importance to speciation in Lake Victoria cichlid fish, male nuptial colour motifs that feature importantly in behavioural reproductive isolation. In a cross between two sympatric species, representative in their mating trait motifs of many closely related sympatric species pairs, we found significant QTLs with moderate to large effects for red and yellow colour traits, with several traits mapped to the same genomic regions. These results are consistent with genetic architectures predicted to facilitate differentiation and persistence of differentiation in traits contributing to reproductive isolation in sympatry with ongoing gene flow [32–34,36–38]. By contrast, we did not find any significant QTLs in a cross between two species representative in their mating trait motifs of closely related species that are usually seen to occupy different islands but do not occur in sympatry. This is consistent with our hypothesis of a genetic architecture that makes phenotypic differentiation not robust to gene flow. We argue that these differences in genetic architecture of superficially similar trait differences could help explain why species with the red-dorsum nuptial colour motif are often sympatric with blue sister species, whereas those with the red-chest motif seem unable to retain differentiation from their blue relatives in sympatry.

The difference between the two crosses in the presence/absence of QTLs with moderate to large effects cannot simply be explained by a difference in power to detect QTLs: repeatedly and randomly downsampling the red-dorsum × blue cross F2 individuals to match the lower sample size in our red-chest × blue cross (a lower sample

size decreases power to detect QTLs) did not significantly change the results, nor did downsampling the markers on the linkage map in the red-chest × blue cross to match the sparser red-dorsum × blue cross-linkage map (a sparser marker density lowers the significance threshold). Another statistical bias [55], where low sample sizes lead to overestimation of effect sizes or PVE, mainly due to the difficulties of statistically detecting loci with small effects, cannot be ruled out. However, this would be expected to affect both of our crosses and should thus not confound the comparison between the two. The different direction of our two crosses (red female × blue male in the red-dorsum × blue cross, blue female × red male in the red-chest × blue cross) should also not affect our results: the lab-strain populations from which the grandparents were taken have stable male colour. Also, none of the QTLs map to known sex determining chromosomes (Pun-LG10 in the red-chest × blue cross [45]), which is also consistent with earlier studies of experimental crosses of the same red-dorsum versus blue species pair [56–58]. Furthermore, in both crosses, both parental phenotypes (considering traits differentiating the two species, figure 2*c,d*) are recovered (with the exceptions that one trait (nose) was never scored as red in any F2s in the red-dorsum × blue cross, and one trait (dorsal fin 2) was scored as red in all F2s of the red-chest × blue cross). Additionally, most individuals are intermediate in the number of traits in which red is expressed in both crosses, albeit with signs of more red dominance in the red-chest × blue cross.

A previous crossing experiment [56] estimated that the difference in the amount of red on the body (mostly flank and dorsum) of males between *Pundamilia* sp. 'nyererei-like' and *Pundamilia* sp. 'pundamilia-like' (our red-dorsum × blue cross) is likely controlled by at least 2–4 loci (with effects of dominance and epistasis). Furthermore, they estimated one gene with complete dominance for yellow flank and epistatic interaction with red on flank and dorsum. Our results conform to these estimates quite well: we found two significant QTLs each for red on the dorsum (LG2 and LG6) and for red on the dorsal fin (LG8 and LG10), each with a PVE of 12–17%. The interval of the QTL for red on the dorsum on LG2 also contains a major QTL for yellow on the flanks with a PVE of nearly 30%, and the significant QTLs for pelvic fin and gill cover overlap with the interval of the QTL on LG8 for red on the dorsal fin, suggesting either linkage of several loci or pleiotropic effects of a single locus. Two-dimensional 2-QTL scans indicate the presence of additional loci contributing to red/yellow colour. Although our modest sample sizes do not allow us to detect small effect QTLs, they are likely present, as our QTLs do not explain all of the variance in our mapped traits. However, the main contribution to variance in red and yellow male colour in this cross comes from these four QTL regions. A first screen of the genes closest to the QTL peaks has not yet revealed any obvious candidate genes. To follow-up on screening for candidate genes across the mapping intervals, each of which contains many dozens to hundreds of genes, will be a topic of future work.

The presence of moderate to large effect QTLs should make phenotypic differentiation and maintenance of differentiation in sympatry more likely than a more dispersed architecture of many loci with small effects under the opposing effects of disruptive selection and gene flow e.g. [32–34,38]. Furthermore, the theoretical models of sympatric

speciation by sexual selection suggest such a process is more likely when reproductive isolation is based on traits underlain by fewer loci [39–41]. Additionally, we find that several traits (figure 3c) map to the same chromosomal region. Although our current dataset does not allow us to determine whether this is due to a single pleiotropic locus or due to several tightly linked loci, both pleiotropy and physical linkage favour divergence and persistence of phenotypic differentiation despite gene flow [34,36,37]. Linkage (or pleiotropy) of divergent adaptive traits has also been observed in Midas cichlids, which are undergoing rapid sympatric divergence [59]. Our findings are also similar to other systems such as *Heliconius* [60], *Mimulus* [61] or *Drosophila* [62], in which traits involved in reproductive isolation in the presence of gene flow are underpinned by large effect or pleiotropic loci. However, none of these other studies made a direct comparison of the genetic architecture of corresponding key traits for reproductive isolation in species pairs that persist (and probably evolved) in sympatry and others that do not persist in sympatry, as we have done here.

In our cross between the species that do not persist in sympatry, i.e. *Pundamilia* sp. ‘red-head’ and *Pundamilia pundamilia* (red-chest × blue), we cannot directly infer the type of genetic architecture underpinning red, or yellow due to the absence of any significant QTLs. Most likely, the effects of potential QTLs are too small to be detected with our sample size, suggesting the presence of a larger number of small effect loci. In red-chest versus blue species without direct geographical contact and with little gene flow that is restricted to rare (but documented, figure 1) long distance dispersal events, trait differentiation due to more and smaller effect mutations is more likely to evolve and persist because recombination will not erode associations between them.

What we cannot yet resolve is whether the genetic architecture for male colour in our red-dorsum versus blue pair has evolved in the face of gene flow, i.e. through selection for the clustering of small effect loci, for instance through genomic rearrangements [63], or was already present, allowing the two species to speciate (and now persist) in sympatry.

## 5. Conclusion

The presence of large effect loci, and of physical linkage or pleiotropy, underlying traits involved in behavioural reproductive isolation, such as male nuptial coloration, may

enable sister species pairs to differentiate and persist in sympatry despite some gene flow. One hallmark of the East African cichlid radiations is the rapid evolution of strong (behavioural) reproductive isolation that is robust to full sympatry in many closely related species [64]. If the genetic architecture of male nuptial coloration in sympatric *Pundamilia* species we report here is representative for other Lake Victoria cichlid species that live sympatrically, this may help explain how speciation in this system could have led to the rapid emergence of communities with many closely related species that persist in sympatry, and why some phenotypic motifs regularly distinguish sympatric species while others are confined to allopatric species.

**Ethics.** Fish experimentation and euthanasia were authorized by the veterinary offices of the cantons of Lucerne (licence number: LU04/07) and Bern (BE 18/15 & BE 65/18).

**Data accessibility.** Raw read (fastq) files for all genotyped individuals of the red-dorsum × blue cross are deposited in the NCBI short read archive (SRA accession PRJNA612290). The fastq files for the red-chest × blue cross have previously been deposited by [45] (SRA accession PRJNA439430). The filtered vcf files, linkage maps, phenotype-genotype tables, and additional information on sequencing data for both crosses are available from the Dryad Digital Repository at <https://dx.doi.org/10.5061/dryad.5tb2rbp1n> [65].

**Authors' contributions.** A.F.F. scored and analysed colour in both crosses, prepared the RAD sequencing libraries for the red-dorsum × blue cross, processed the sequencing data, generated linkage maps, performed the colour QTL analyses in both crosses, wrote the manuscript together with O.S. M.P.H. reared the fish, took the colour photos, prepared all the red-chest × blue and two of the red-dorsum × blue cross RAD sequencing libraries and provided critical feedback for the manuscript. C.L.P. provided guidance for the analyses and contributed to manuscript writing. O.S. designed and coordinated the study, developed the colour scoring scheme and the colour trait space, collected the species distribution data, wrote the manuscript together with A.F.F.

**Competing interests.** We declare we have no competing interests.

**Funding.** This work was funded by Swiss National Science Foundation (SNSF) grant nos. 31003A\_163338, 31003A\_144046 and 31003A\_118293 to O.S.

**Acknowledgements.** We thank Julia Schwarzer and Philine Feulner for their earlier and preliminary work on the red-chest × blue cross and Oliver Selz for initiating the red-dorsum × blue cross. Many thanks to Salome Mwaiko for assistance in the laboratory, the FishEc team for helpful discussions and feedback throughout the process. Joana Meier for providing the allelic balance filter script. The bioinformatics facilities for sequence data processing were provided by the Genetic Diversity Centre at ETH Zürich. We further thank two anonymous reviewers for their constructive feedback on a previous version of this manuscript.

## References

- Johnson TC, Kelts K, Odada E. 2000 The holocene history of Lake Victoria. *AMBIO A J. Hum. Environ.* **29**, 2–11. (doi:10.1579/0044-7447-29.1.2)
- Stager JC, Johnson TC. 2008 The late Pleistocene desiccation of Lake Victoria and the origin of its endemic biota. *Hydrobiologia* **596**, 5–16. (doi:10.1007/s10750-007-9158-2)
- Seehausen O, Witte F, Katunzi EF, Smits J, Bouton N. 1997 Patterns of the remnant cichlid fauna in southern Lake Victoria. *Conserv. Biol.* **11**, 890–904. (doi:10.1046/j.1523-1739.1997.95346.x)
- Seehausen O. 2002 Patterns in fish radiation are compatible with Pleistocene desiccation of Lake Victoria and 14 600 year history for its cichlid species flock. *Proc. R. Soc. B* **269**, 491–497. (doi:10.1098/rspb.2001.1906)
- Greenwood PH. 1974 *Cichlid fishes of lake Victoria, east Africa: the biology and evolution of a species flock*. Bristol, UK: John Wright and Sons Ltd, Stonebridge Press.
- Seehausen O. 1996 *Lake Victoria rock cichlids: taxonomy, ecology, and distribution*. Zevenhuizen, NL: Verduijn Cichlids.
- Greenwood PH. 1980 Towards a phyletic classification of the ‘genus’ *Haplochromis* (Pisces, Cichlidae) and related taxa. II. The species from Lakes Victoria, Nabugabo, Edward, George and Kivu. *Bull. Br. Mus. nat. Hist.* **39**, 1–101. (doi:10.5962/bhl.part.13268)
- Seehausen O, Lippitsch E, Bouton N, Zwennes H. 1998 Mbibi, the rock-dwelling cichlids of Lake Victoria: description of three new genera and fifteen new species (Teleostei). *Ichthyol. Explor. Freshwaters* **9**, 129–228.
- Seehausen O, Van Alphen JJM, Witte F. 1999 Can ancient colour polymorphisms explain why



- some cichlid lineages speciate rapidly under disruptive sexual selection? *Belgian J. Zool.* **129**, 43–60.
10. Seehausen O, van Alphen JJM, Witte F. 1997 Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* **277**, 1808–1811. (doi:10.1126/science.277.5333.1808)
  11. Seehausen O, Schluter D. 2004 Male–male competition and nuptial-colour displacement as a diversifying force in Lake Victoria cichlid fishes. *Proc. R. Soc. B* **271**, 1345–1353. (doi:10.1098/rspb.2004.2737)
  12. Seehausen O, van Alphen JJM. 1999 Can sympatric speciation by disruptive sexual selection explain rapid evolution of cichlid diversity in Lake Victoria? *Ecol. Lett.* **2**, 262–271. (doi:10.1046/j.1461-0248.1999.00082.x)
  13. Seehausen O. 2009 Progressive levels of trait divergence along a ‘speciation transect’ in the Lake Victoria cichlid fish *Pundamilia*. In *Speciation and patterns of diversity* (eds RK Butlin, J Bridle, D Schluter), pp. 155–176. Cambridge, UK: Cambridge University Press.
  14. Seehausen O. 1997 Distribution of and reproductive isolation among color morphs of a rock-dwelling Lake Victoria cichlid (*Haplochromis nyererei*). *Ecol. Freshw. Fish* **6**, 59–66. (doi:10.1111/j.1600-0633.1997.tb00145.x)
  15. Meier JI, Sousa VC, Marques DA, Selz OM, Wagner CE, Excoffier L, Seehausen O. 2017 Demographic modelling with whole-genome data reveals parallel origin of similar *Pundamilia* cichlid species after hybridization. *Mol. Ecol.* **26**, 123–141. (doi:10.1111/mec.13838)
  16. Meier JI, Marques DA, Wagner CE, Excoffier L, Seehausen O. 2018 Genomics of parallel ecological speciation in Lake Victoria cichlids. *Mol. Biol. Evol.* **35**, 1489–1506. (doi:10.1093/molbev/msy051)
  17. Hamilton SE. 2016 Creation of a bathymetric map of Lake Victoria, Africa. See <http://dx.doi.org/10.7910/DVN/ZSWZIA>.
  18. Seehausen O *et al.* 2008 Speciation through sensory drive in cichlid fish. *Nature* **455**, 620–626. (doi:10.1038/nature07285)
  19. Bouton N, Seehausen O, van Alphen JJM. 1997 Resource partitioning among rock-dwelling haplochromines (Pisces: Cichlidae) from Lake Victoria. *Ecol. Freshw. Fish* **6**, 225–240. (doi:10.1111/j.1600-0633.1997.tb00165.x)
  20. Dijkstra PD, Seehausen O, Groothuis TGG. 2005 Direct male–male competition can facilitate invasion of new colour types in Lake Victoria cichlids. *Behav. Ecol. Sociobiol.* **58**, 136–143. (doi:10.1007/s00265-005-0919-5)
  21. Dijkstra PD, Seehausen O, Pierotti MER, Groothuis TGG. 2007 Male–male competition and speciation: aggression bias towards differently coloured rivals varies between stages of speciation in a Lake Victoria cichlid species complex. *J. Evol. Biol.* **20**, 496–502. (doi:10.1111/j.1420-9101.2006.01266.x)
  22. Dijkstra PD, Hemelrijk C, Seehausen O, Groothuis TGG. 2008 Color polymorphism and intrasexual competition in assemblages of cichlid fish. *Behav. Ecol.* **20**, 138–144. (doi:10.1093/beheco/arn125)
  23. Seehausen O, Van Alphen JJM. 1998 The effect of male coloration on female mate choice in closely related Lake Victoria cichlids (*Haplochromis nyererei* complex). *Behav. Ecol. Sociobiol.* **42**, 1–8. (doi:10.1007/s002650050405)
  24. Selz OM, Pierotti MERR, Maan ME, Schmid C, Seehausen O. 2014 Female preference for male color is necessary and sufficient for assortative mating in 2 cichlid sister species. *Behav. Ecol.* **25**, 612–626. (doi:10.1093/beheco/aru024)
  25. Maan ME, Seehausen O, Söderberg L, Johnson L, Ripmeester EAP, Mrosso HDJ, Taylor MI, Van Dooren TJM, Van Alphen JJM. 2004 Intraspecific sexual selection on a speciation trait, male coloration, in the Lake Victoria cichlid *Pundamilia nyererei*. *Proc. R. Soc. B* **271**, 2445–2452. (doi:10.1098/rspb.2004.2911)
  26. Selz OM, Thommen R, Pierotti MER, Anaya-Rojas JM, Seehausen O. 2016 Differences in male coloration are predicted by divergent sexual selection between populations of a cichlid fish. *Proc. R. Soc. B* **283**, 20160172. (doi:10.1098/rspb.2016.0172)
  27. Haesler MP, Seehausen O. 2005 Inheritance of female mating preference in a sympatric sibling species pair of Lake Victoria cichlids: implications for speciation. *Proc. R. Soc. B* **272**, 237–245. (doi:10.1098/rspb.2004.2946)
  28. Svensson O, Woodhouse K, van Oosterhout C, Smith A, Turner GF, Seehausen O. 2017 The genetics of mate preferences in hybrids between two young and sympatric Lake Victoria cichlid species. *Proc. R. Soc. B* **284**, 20162332. (doi:10.1098/rspb.2016.2332)
  29. Stelkens RB, Pierotti ME, Joyce DA, Smith AM, van der Sluijs I, Seehausen O. 2008 Disruptive sexual selection on male nuptial coloration in an experimental hybrid population of cichlid fish. *Phil. Trans. R. Soc. B* **363**, 2861–2870. (doi:10.1098/rstb.2008.0049)
  30. Felsenstein J. 1981 Skepticism towards Santa Rosalia, or why are there so few kinds of animals? *Evolution* **35**, 124. (doi:10.2307/2407946)
  31. Slatkin M. 1987 Gene flow and the geographic structure of natural populations. *Science* **236**, 787–792. (doi:10.1126/science.3576198)
  32. Gavrilets S. 2004 *Fitness landscapes and the origin of species*. Princeton, NJ: Princeton University Press.
  33. Yeaman S, Otto SP. 2011 Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution* **65**, 2123–2129. (doi:10.1111/j.1558-5646.2011.01277.x)
  34. Yeaman S, Whitlock MC. 2011 The genetic architecture of adaptation under migration–selection balance. *Evolution* **65**, 1897–1911. (doi:10.1111/j.1558-5646.2011.01269.x)
  35. Griswold CK. 2006 Gene flow’s effect on the genetic architecture of a local adaptation and its consequences for QTL analyses. *Heredity* **96**, 445–453. (doi:10.1038/sj.hdy.6800822)
  36. Bürger R, Akerman A. 2011 The effects of linkage and gene flow on local adaptation: a two-locus continent–island model. *Theor. Popul. Biol.* **80**, 272–288. (doi:10.1016/j.tpb.2011.07.002)
  37. Flaxman SM, Wacholder AC, Feder JL, Nosil P. 2014 Theoretical models of the influence of genomic architecture on the dynamics of speciation. *Mol. Ecol.* **23**, 4074–4088. (doi:10.1111/mec.12750)
  38. van Doorn GS, Dieckmann U. 2006 The long-term evolution of multilocus traits under frequency-dependent disruptive selection. *Evolution* **60**, 2226. (doi:10.1554/06-291.1)
  39. Turner GF, Burrows MT. 1995 A model of sympatric speciation by sexual selection. *Proc. R. Soc. B* **260**, 287–292. (doi:10.1098/rspb.1995.0093)
  40. Higashi M, Takimoto G, Yamamura N. 1999 Sympatric speciation by sexual selection. *Nature* **402**, 523–526. (doi:10.1038/990087)
  41. Arnegard ME, Kondrashov AS. 2004 Sympatric speciation by sexual selection alone is unlikely. *Evolution* **58**, 222–237. (doi:10.1111/j.0014-3820.2004.tb01640.x)
  42. Baerends GP, Baerends-Van Roon JT. 1950 *An introduction to the study of the ethology of the cichlid fishes*. Leiden, The Netherlands: E. J. Brill.
  43. Sambrook J, Russell DW. 2001 *Molecular cloning: a laboratory manual*, 3rd edn. New York, NY: Cold Spring Harb. Lab. Press.
  44. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008 Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* **3**, e3376. (doi:10.1371/journal.pone.0003376)
  45. Feulner PGD, Schwarzer J, Haesler MP, Meier JI, Seehausen O. 2018 A dense linkage map of Lake Victoria cichlids improved the *Pundamilia* genome assembly and revealed a major QTL for sex-determination. *G3: Genes Genomes Genetics* **8**, 2411–2420. (doi:10.1534/g3.118.200207)
  46. Langmead B, Salzberg SL. 2012 Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359. (doi:10.1038/nmeth.1923)
  47. McKenna A *et al.* 2010 The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303. (doi:10.1101/gr.107524.110)
  48. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079. (doi:10.1093/bioinformatics/btp352)
  49. Danecek P *et al.* 2011 The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158. (doi:10.1093/bioinformatics/btr330)
  50. van Ooijen JW. 2006 *Joinmap®4, software for the calculation of genetic linkage maps in experimental populations*. Wageningen, The Netherlands: Kyazma BV.
  51. Broman KW, Wu H, Sen S, Churchill GA. 2003 R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**, 889–890. (doi:10.1093/bioinformatics/btg112)
  52. R Core Development team. 2015 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See <http://www.R-project.org/>.

53. Brawand D *et al.* 2014 The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**, 375–381. (doi:10.1038/nature13726)
54. Guyon R *et al.* 2012 A high-resolution map of the Nile tilapia genome: a resource for studying cichlids and other percomorphs. *BMC Genomics* **13**, 222. (doi:10.1186/1471-2164-13-222)
55. Beavis WD. 1995 The power and deceit of QTL experiments: lessons from comparative QTL studies. In *Proc. of the 49th Ann. Corn and Sorghum Industry Research Conf., ASTA, Washington, DC*, pp. 252–268. Alexandria, VA: ASTA.
56. Magalhaes IS, Seehausen O. 2010 Genetics of male nuptial colour divergence between sympatric sister species of a Lake Victoria cichlid fish. *J. Evol. Biol.* **23**, 914–924. (doi:10.1111/j.1420-9101.2010.01960.x)
57. van der Sluijs I, Seehausen O, van Dooren TJM, van Alphen JJM. 2010 No evidence for a genetic association between female mating preference and male secondary sexual trait in a Lake Victoria cichlid fish. *Curr. Biol.* **56**, 57–64. (doi:10.1093/czoolo/56.1.57)
58. van der Sluijs I, Dijkstra PD, Lindeyer CM, Visser B, Smith AM, Groothuis TGG, van Alphen JJM, Seehausen O. 2013 A test of genetic association among male nuptial coloration, female mating preference, and male aggression bias within a polymorphic population of cichlid fish. *Curr. Zool.* **59**, 221–229. (doi:10.1093/czoolo/59.2.221)
59. Fruciano C, Franchini P, Kovacova V, Elmer KR, Henning F, Meyer A. 2016 Genetic linkage of distinct adaptive traits in sympatrically speciating crater lake cichlid fish. *Nat. Commun.* **7**, 1–8. (doi:10.1038/ncomms12736)
60. Merrill RM, Rastas P, Martin SH, Melo MC, Barker S, Davey J, McMillan WO, Jiggins CD. 2019 Genetic dissection of assortative mating behavior. *PLoS Biol.* **17**, 1–21. (doi:10.1371/journal.pbio.2005902)
61. Ferris KG, Barnett LL, Blackman BK, Willis JH. 2017 The genetic architecture of local adaptation and reproductive isolation in sympatry within the *Mimulus guttatus* species complex. *Mol. Ecol.* **26**, 208–224. (doi:10.1111/mec.13763)
62. Noor MAF, Grams KL, Bertucci LA, Reiland J. 2001 Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. USA* **98**, 12 084–12 088. (doi:10.1073/pnas.221274498)
63. Yeaman S. 2013 Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proc. Natl. Acad. Sci. USA* **110**, E1743–E1751. (doi:10.1073/pnas.1219381110)
64. Seehausen O, Wagner CE. 2014 Speciation in freshwater fishes. *Annu. Rev. Ecol. Evol. Syst.* **45**, 621–651. (doi:10.1146/annurev-ecolsys-120213-091818)
65. Feller AF, Haesler MP, Peichel CL, Seehausen O. 2020 Data from: Genetic architecture of a key reproductive isolation trait differs between sympatric and non-sympatric sister species of Lake Victoria cichlids. Dryad Digital Repository. (<https://dx.doi.org/10.5061/dryad.5tb2rbp1n>).