

<https://helda.helsinki.fi>

Allopatric origin of sympatric whitefish morphs with insights on the genetic basis of their reproductive isolation

Fang, Bohao

2022-08

Fang , B , Momigliano , P , Kahilainen , K K & Merilä , J 2022 , ' Allopatric origin of sympatric whitefish morphs with insights on the genetic basis of their reproductive isolation ' , Evolution , vol. 76 , no. 8 , pp. 1905-1913 . <https://doi.org/10.1111/evo.14559>

<http://hdl.handle.net/10138/347083>

<https://doi.org/10.1111/evo.14559>

cc_by_nc

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Allopatric origin of sympatric whitefish morphs with insights on the genetic basis of their reproductive isolation

Bohao Fang,^{1,2,3}  Paolo Momigliano,^{1,4}  Kimmo K. Kahilainen,^{5,6}  and Juha Merilä^{1,7,8} 

¹Ecological Genetics Research Unit, Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki 00014, Finland

²Department of Organismic and Evolutionary Biology and Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA

³E-mail: bfang@fas.harvard.edu

⁴Department of Biochemistry, Genetics, and Immunology, Universidade de Vigo, Vigo 36310, Spain

⁵Lammi Biological Station, University of Helsinki, Lammi 16900, Finland

⁶Kilpisjärvi Biological Station, University of Helsinki, Kilpisjärvi 99490, Finland

⁷Area of Ecology and Biodiversity, School of Biological Sciences, The University of Hong Kong, Pok Fu Lam, Hong Kong

⁸E-mail: juha.merila@helsinki.fi

Received September 16, 2021

Accepted June 22, 2022

The European whitefish (*Coregonus lavaretus*) species complex is a classic example of recent adaptive radiation. Here, we examine a whitefish population introduced to northern Finnish Lake Tshakal in the late 1960s, where three divergent morphs (*viz.* littoral, pelagic, and profundal feeders) were found 10 generations after. Using demographic modeling based on genomic data, we show that whitefish morphs evolved during a phase of strict isolation, refuting a rapid sympatric divergence scenario. The lake is now an artificial hybrid zone between morphs originated in allopatry. Despite their current syntopy, clear genetic differentiation remains between two of the three morphs. Using admixture mapping, we identify five SNPs associated with gonad weight variation, a proxy for sexual maturity and spawning time. We suggest that ecological adaptations in spawning time evolved in allopatry are currently maintaining partial reproductive isolation in the absence of other barriers to gene flow.

KEY WORDS: Admixture mapping, *Coregonus lavaretus*, demographic modeling, gonad weight, whitefish.

The pace at which evolution and speciation occur, and the mechanisms determining their tempo, has intrigued evolutionary biologists for a long time (Simpson 1944; Bush et al. 1977; Kornfield 1978), and still continue today (Schluter 2000; Nosil 2012; Lescak et al. 2015; Momigliano et al. 2017; Salzburger 2018; Skúlason et al. 2019). Concerns over the organisms' ability to adapt to environmental changes (e.g., Gienapp et al. 2008; Merilä and Hendry 2014), and the realization that adaptation can be very rapid (Thomas et al. 2017; Matthews et al. 2018; Marques et al. 2019b), have sparked new interest toward the rates of evolution. Adaptation from standing genetic variation

can be fast, especially if genetic diversity has been augmented by historical gene flow (Häkli et al. 2018; Salzburger 2018; Jacobs et al. 2019; Marques et al. 2019a). One example is the evolution of freshwater-adapted three-spined sticklebacks (*Gasterosteus aculeatus*) colonizing ponds that were formed during uplift caused by the 1964 Great Alaska Earthquake (Lescak et al. 2015). Similarly rapid adaptation has been observed for sockeye salmon and Darwin's finches when colonizing new environments (Hendry et al. 2000; Lamichhaney et al. 2018). In all these examples, broad ecological opportunities in novel environments likely facilitated rapid phenotypic and genetic differentiation.

Coregonid fishes have undergone extensive recent adaptive radiations but intermittent gene flow among morphs and species is common (Østbye et al. 2005; Bernatchez et al. 2010; Hudson et al. 2011; Vonlanthen et al. 2012). European whitefish (*Coregonus lavaretus*) experienced frequent adaptive radiations throughout their distribution (Svärdson 1979; Østbye et al. 2005; Vonlanthen et al. 2012). Facing different ecological opportunities, *C. lavaretus* evolved partially reproductively isolated, co-existing morphs using different niches (Kahilainen and Østbye 2006; Siwertsson et al. 2010; Præbel et al. 2013). These morphs differ in several phenotypic traits (Harrod et al. 2010; Præbel et al. 2013), such as body size and number of gill rakers, both heritable traits related to feeding ecology (Kahilainen et al. 2011; Häkli et al. 2018). Although rapid whitefish differentiation has been documented in both Nearctic and Palearctic regions (Bernatchez et al. 2010; Hudson et al. 2011), less is known about the genetic basis of their rapid phenotypic differentiation (but see Vonlanthen et al. 2012; Jacobs et al. 2019; Rougeux et al. 2019a).

During the 1960s, the Finnish Fisheries authorities facilitated introductions of whitefish in several high-altitude lakes including Lake Tsahkal, where the species did not previously occur. Sampling in Lake Tsahkal in 2011 revealed three different morphs inhabiting littoral, pelagic, and profundal habitats (Kahilainen et al. 2011; Præbel et al. 2013). There is evidence of both allopatric and sympatric divergence of whitefish morphs (Præbel et al. 2013; Salisbury and Ruzzante 2022), and genetic diversity fueled by historical admixture can facilitate rapid phenotypic divergence when ecological opportunity arises (Jacobs et al. 2019; Frei et al. 2022). It is possible that the whitefish introduced to Lake Tsahkal came from a single ancestral population and gave rise to different morphs in about 50 years (≈ 10 generations). The alternative hypothesis is that two or three morphs were simultaneously introduced in the lake. In such case, the whitefish of Lake Tsahkal would provide a rare opportunity to study the ecological and genetic underpinnings of reproductive barriers following secondary contact.

Lake Tsahkal provides also an ideal system to study the genetic basis of phenotypic traits defining the whitefish morphs via admixture mapping (Gompert et al. 2017). To date, few studies have provided insights on the genetic basis of whitefish phenotypic traits (Vonlanthen et al. 2009; Gagnaire et al. 2013a,b; Feulner and Seehausen 2019; Jacobs et al. 2019). These studies focus mainly on trophic traits, whereas less attention has been paid on traits associated with annual timing of reproduction. Timing of reproduction can act as a reproductive barrier among diverging populations (Hendry and Day 2005), and it is well-known that whitefish morphs have differences in their spawning time and location (e.g., Svärdson 1979; Vonlanthen et al. 2009; Kahilainen et al. 2014; Bitz-Thorsen et al. 2020).

Here, using a demographic modeling framework, we test two possible divergence scenarios of the Lake Tsahkal whitefish morphs: sympatric divergence and allopatric divergence followed by secondary contact. Furthermore, we investigate the genetic architecture of 22 morphological traits, including gonad weight variation, a proxy for spawning time that is a trait expected to be under ecological selection that could lead to reproductive isolation.

Methods

STUDY SYSTEM AND DATA COLLECTION

Lake Tsahkal is located in the tree line of northern Finnish Lapland (69°01'N, 20°50'E, 559 m a.s.l.). The lake is oligotrophic (totP 5 $\mu\text{g/L}$, totN 140 $\mu\text{g/L}$), deep (max 35 m, mean 9 m), clear-water (compensation depth 7.5 m), and has equal habitat distribution (41% pelagic/profundal, 59% littoral; Hayden et al. 2014). Introduced whitefish in the 1960s had broad ecological opportunity as they are more efficient in zooplanktivory and benthivory than two native species in the lake: the brown trout (*Salmo trutta*) and burbot (*Lota lota*) (Siwertsson et al. 2010; Hayden et al. 2014). Reproductive isolation among whitefish morphs could arise via differences in resource use causing differences in spawning times and places (Kahilainen et al. 2014; Taylor and Friesen 2017; Thibert-Plante et al. 2020).

Littoral, pelagic, and profundal zones were identified in the lake (Hayden et al. 2014) (Fig. 1a) and subsequently whitefish were collected from these habitats with gill net series in August 3–17, 2011 (Kahilainen et al. 2011). Caught fish were transported to field laboratory and labeled and frozen (-20°C) for later measurements of fish size, gill raker count, resource use metrics, and morphological traits (for details, see Table S1, Method S1, and Figs. S1, S2, S4). All caught fish were assigned to three different morphs according to their gill raker appearance, head, and body shape (Method S1; Kahilainen and Østbye 2006).

A total of 61 individuals (17 littoral, 22 pelagic, and 22 profundal fish) were sequenced using 2b-RAD sequencing approach (Wang et al. 2012) to obtain 36 bp single-end fragments with a mean coverage $30.2\times$ (13.7–44.0; Table S1) by Illumina HiSeq 4000 at BGI, Hong Kong. DNA extraction and 2b-RAD library preparation followed exactly the protocols detailed in Momigliano et al. (2018, 2021).

POPULATION GENETIC ANALYSES

Raw reads were demultiplexed and PCR duplicates removed as per Momigliano et al. (2018). Reads were aligned to a chromosome-level assembly of the alpine European whitefish (*Coregonus* spp.) publicly available at https://www.ebi.ac.uk/ena/data/view/GCA_9021750 (De-Kayne et al. 2020), using *Bowtie2* (Langmead and Salzberg 2012). SAM files were converted to BAM files and indexed using *SAMtools* (Li and Durbin 2009).

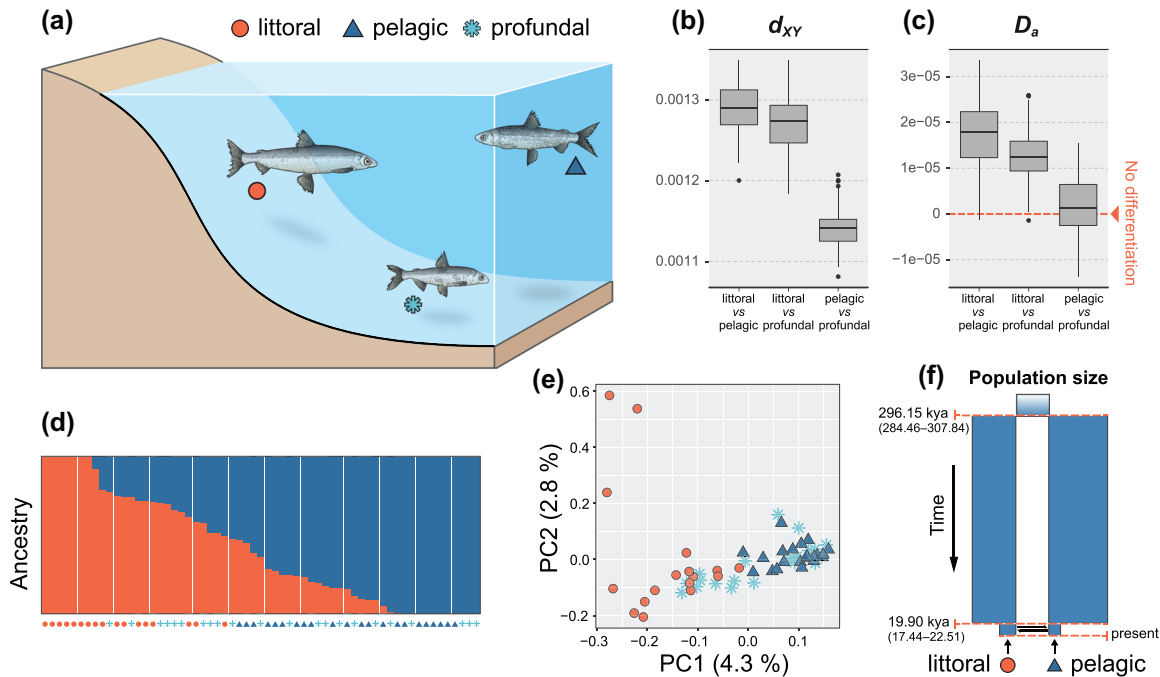


Figure 1. Population structure and demographic history of three morphs of the European whitefish (*Coregonus lavaretus*) in Lake Tsahkal. (a) Graphic presentation of the littoral, pelagic, and profundal morphs inhabiting the lake. (b, c) Genome-wide divergence (absolute divergence, d_{XY} , and net pairwise nucleotide divergence, D_a) between morphs. Boxplots show 100 bootstraps by permuting sites within chromosomes. (d) Individual ancestry reconstructed from *NGSadmix* at $K = 2$. (e) Principal component analysis (PCA) of genotype likelihoods of 15,615 SNPs. (f) Inferred demographic history of the littoral and pelagic morphs: secondary contact with changes in population sizes. Estimated time parameters with 95% confidence intervals are shown. Kya, thousand years ago. Sample identifications of the (d, e) are presented in Figure S3.

Genotype likelihoods were estimated from BAM files using *ANGSD* (Korneliussen et al. 2014), retaining only biallelic loci with $\leq 25\%$ missing data and bases with mapping quality and Phred scores > 20 . We performed a principal component analysis (PCA) using *PCAngsd* (Meisner and Albrechtsen 2018) retaining variants with minimum minor allele frequency (MAF) of 0.02. Individual ancestries were inferred using *NGSadmix* (Skotte et al. 2012) based on genotype likelihoods, assuming one to three ancestral populations. Absolute divergence (d_{XY}) (Nei 1987) and net nucleotide divergence (D_a) (Nei 1987) were estimated based on nonadmixed individuals identified by *NGSadmix* using scripts from Momigliano et al. (2021).

DEMOGRAPHIC MODELING

We compared demographic models using the software package *moments* (Jouanous et al. 2017) based on diffusion approximations of the allele frequency spectrum (SFS). Ten demographic models were tested using nonadmixed littoral and pelagic individuals, as the ancestral pelagic and profundal morphs were not genetically distinct (see *Results*) and we had more pelagic ($n = 8$) than profundal ($n = 5$) nonadmixed individuals. We first defined five competing gene flow scenarios: strict isolation (SI), isolation with migration (IM), secondary contact (SC), ancient migration

(AM), and a two epochs model (2EP) assuming heterogeneous migration rates through time (Roux et al. 2016; Momigliano et al. 2021). Because unaccounted changes in N_e can bias model choice and parameter estimation (Momigliano et al. 2021), we defined five additional models accounting for a change in N_e (increase or decrease) in both daughter populations at time T_2 , based on the above simple models: SI_NeC, AM_NeC, IM_NeC, SC_NeC, and 2EP_NeC; NeC stands for “ N_e Change.” The 10 tested models are visualized in Figure S6. To account for possible effects of linkage in the 2b-RAD data, the best fitting model was chosen using Likelihood Ratio Test (LRT) as outlined in Coffman et al. (2016).

All analyses were performed based on the unfolded two-dimensional SFS (2D-SFS) derived from the sites shared between pairwise morphs. The programs *ANGSD* and *Moments* and custom R and python scripts adopted from Momigliano et al. (2021) and Fang et al. (2021) were used in the analyses. Detailed methods are given in Method S1.

ADMIXTURE MAPPING

Admixture mapping of phenotypic traits was performed with two genome-wide association (GWA) approaches: GEMMA (version 0.98), a genome-wide efficient mixed model association

approach (Zhou and Stephens 2014), and LDna-EMMAX, a linkage disequilibrium (LD) clustering-based approach for association mapping (Kang et al. 2010; Kempainen et al. 2015; Li et al. 2018; Fang et al. 2021). The same loci obtained from the filtering steps above were used in *ANGSD* for GWAs, except that we increased the MAF threshold to 0.1 (-minMaf 0.1), resulting in 12,212 loci. Genotype likelihood data were used in LDna-EMMAX, whereas called genotypes derived from *ANGSD* (-doVcf 1) were fed to GEMMA.

Univariate linear mixed models were conducted for each phenotypic trait in GEMMA, accounting for relatedness between individuals by supplying a relatedness matrix as a covariate. In LDna-EMMAX, we used LD network analyses (LDna; Kempainen et al. 2015) to identify correlated clusters of single-nucleotide polymorphisms (SNPs), followed by GWA by fitting a multilocus mixed model accounting for relatedness as in Li et al. (2018) and Fang et al. (2021). Assuming a linear relationship between trait and body size, phenotypic measurements were corrected to mean total length (22.4 cm) before GWAs (Method S1). Significance tests of the two GWA approaches were performed based on Wald and permutation tests (10,000 permutations), respectively.

We corrected for multiple testing using the false discovery rate (FDR) approach of Benjamini and Hochberg (1995) and Benjamini and Yekutieli (2001) with an FDR threshold of 0.1. Specifically, Q -values (the FDR at which a test is significant) were calculated from P -values using the R-function *p.adjust* and we performed this in two ways. First, for GEMMA, the P -values derived from Wald-tests were corrected for genome-wide 12,212 SNPs; for LDna, we simply adopted P -values estimated by 10,000 permutations. Second, in a more stringent correction, the number of traits (22) was corrected for GEMMA ($12,212 \times 22 = 268,664$ independent tests) and LDna-EMMAX (22 independent tests). To visualize phenotype-genotype correlation, we performed with *PCAngsd* a PCA based on the genotype likelihoods of the loci associated with gonad weight variation, and tested correlations between the first principal component (PC1) and adjusted gonad weight of individuals. This test was also conducted for genome-wide neutral SNPs for comparison. Details and pipelines used to perform GWAs are provided in Method S1.

Results

POPULATION STRUCTURE AND DEMOGRAPHIC HISTORY

Littoral and pelagic whitefish morphs differed significantly in body size, gill raker count, morphology, diet, and parasites, whereas the profundal morph was intermediate (Fig. S4). Stable isotopes and mercury indicated clear differentiation of all three

morphs in their respective habitats (Fig. S4). We obtained genotype likelihoods of 15,615 loci across 61 individuals. *NGSadmix* analyses revealed two ancestral components ($K = 2$), and a clear genetic differentiation between littoral and pelagic/profundal morphs in analysis of nonadmixed individuals (Fig. 1b–d). Overall divergence between nonadmixed pelagic and profundal morphs was low: $d_{XY} = 1.14 \times 10^{-3}$ (95% confidence interval [CI]: 1.08×10^{-3} to 1.21×10^{-3} ; Fig. 1b) and $D_a = 1.81 \times 10^{-6}$ (95% CI: -1.37×10^{-5} to 1.56×10^{-5} ; Fig. 1c). Out of 61 individuals analyzed, 41 were genetically admixed (i.e., having >3% ancestry from a secondary cluster; Fig. 1d). PCA supports these inferences, showing no differentiation of pelagic and profundal morphs, and admixture between littoral and pelagic morphs (Fig. 1e).

The demographic analyses supported secondary contact with changes in N_e (SC_NeC; Figs. 1f, S7, S8) as the best-fit model. Littoral and pelagic morphs diverged 296.15 thousand years ago (kya; 95% CI: 284.46–307.84 kya, i.e., 56.89–61.57 thousand generations), and experienced a secondary contact following the end of the last glaciation 19.9 kya (95% CI: 17.44–22.51 kya, 3.49–4.50 thousand generations) with decreases in N_e for both morphs after the Last Glacial Maximum (LGM; Fig. 1f). For a summary of scaled parameters, see Table S2.

ADMIXTURE MAPPING OF GONADAL SIZE VARIATION

The two GWA analyses collectively revealed five SNPs (Chromosome [Chr] 22: 43,111,579, Chr28: 9,855,214, Chr28 9,763,679, Chr32 37,564,809, and Chr36: 8,058,628) associated with gonad weight variation in the 41 admixed individuals (Fig. 2). Out of these, two SNPs in Chr22: 43,111,579 and Chr36: 8,058,628 were in strong LD ($r^2 > 0.9$) and clustered in LDna (the first step in LDna-EMMAX analyses).

The genotypes of the five SNPs explained 63% of the variation in gonad weight among admixed individuals ($P < 6.5 \times 10^{-10}$; Fig. 2c), 44% among all collected individuals ($P < 7.5 \times 10^{-9}$; Fig. 2d), 44% among all female individuals ($P < 2.3 \times 10^{-5}$; Fig. S8), and 26% among all male individuals ($P < 6.5 \times 10^{-3}$; Fig. S8). The individual SNPs above explained 41%–51% of the variation in gonad weight among admixed individuals (Fig. S9). However, the genome-wide neutral loci did not explain any variation in gonadal weight as shown by nonsignificant phenotype-genotype correlation (Fig. S8). When correcting for multiple testing across 22 traits, two SNPs (Chr 22: 43,111,579 and Chr36: 8,058,628) were identified as significant in the admixture mapping (Fig. S10c–h). No variant was significantly associated with the 21 other traits used to define the morphs (Fig. S12).

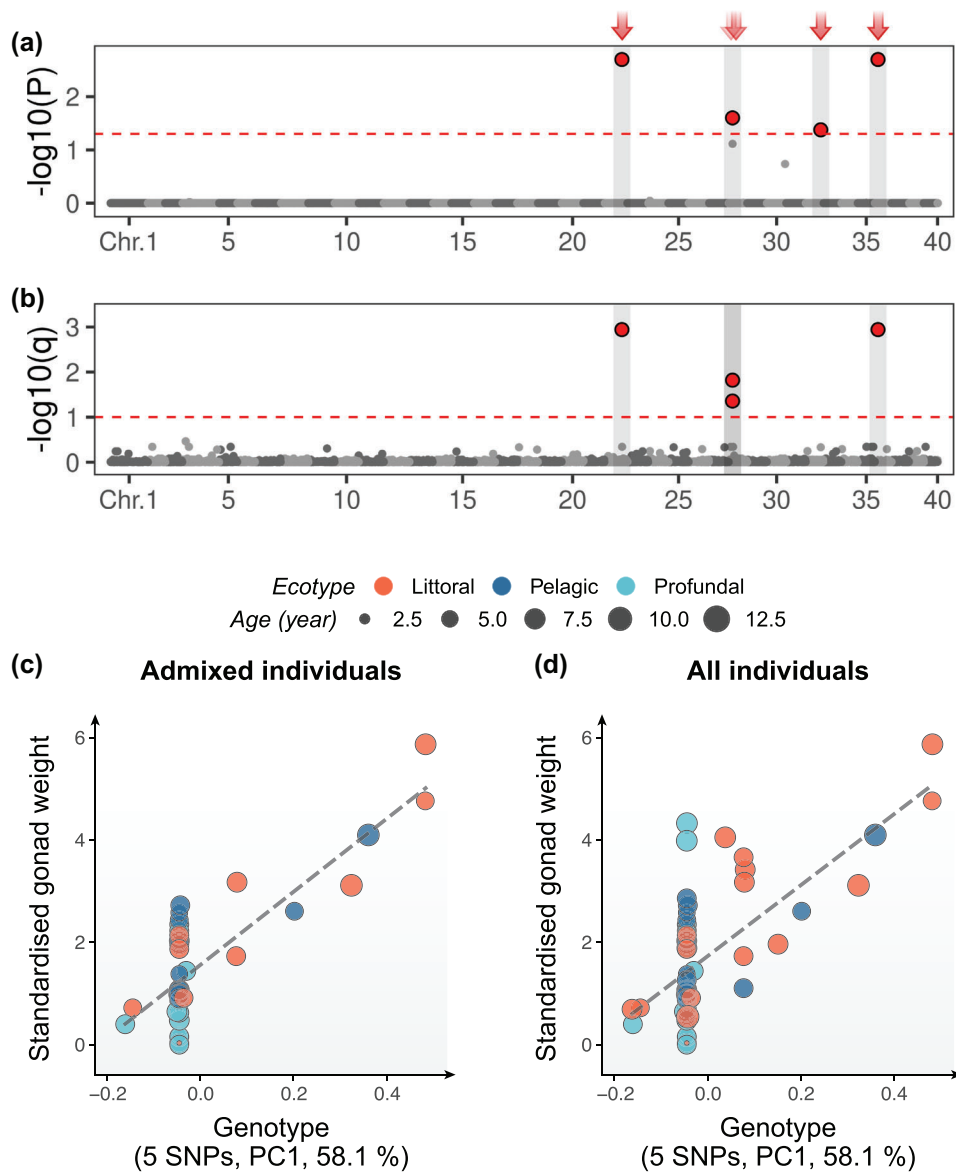


Figure 2. Admixture mapping of the gonad weight variation in European whitefish. Admixture mapping using *LDna-EMMAX* (a) and *GEMMA* (b) suggested five single-nucleotide polymorphisms (SNPs; see results) significantly associated with the gonad weight variation. x-axis depicts genomic position and y-axis the negative logarithm of the association *P*-value (a) and *Q*-value (b; FDR adjusted *P*-value). Dashed red line marks the significance threshold adjusted for multiple testing. Chr = chromosome. (c, d) Correlations between the genotypes of the five gonad weight-associated SNPs and the gonad weight among 41 admixed individuals (c; $y = 67.15x + 1.58$; $R^2 = 0.63$; $P < 6.5 \times 10^{-10}$) and all individuals sampled in Lake Tsahkal (d; $y = 6.93x + 1.73$; $R^2 = 0.44$; $P < 7.5 \times 10^{-9}$).

Discussion

Using demographic models, we refuted the hypothesis of rapid sympatric differentiation of whitefish in Lake Tsahkal. Demographic analyses support an allopatric origin of whitefish morphs clearly predating the transplantation to the lake 50 years before sampling. This is not entirely surprising. While in fishes there are examples of rapid adaptive responses to novel environmental conditions on similar timescales, in most cases, these responses are subtle or have taken place over much longer time, commonly over hundreds of thousands of years (e.g., Albertson et al. 1999;

Danley and Kocher 2001; Bolnick and Fitzpatrick 2007; Hudson et al. 2007, 2011).

Our results suggest that following a long phase of strict isolation, gene flow between divergent whitefish morphs was re-established during the last glacial retreat. Similar cases of secondary contact in post-glacial lakes have been reported from other northern temperate zone whitefish and they often led to hybridization (Bernatchez et al. 2010; Hudson et al. 2011; Rougeux et al. 2017, 2019b). The observed gradient of admixture suggests that the introduction of whitefish to the Lake Tsahkal created an

artificial hybrid zone facilitating gene flow between littoral and pelagic morphs. The absence of genetic differentiation between the pelagic and profundal morphs was somewhat surprising. It is possible that these phenotypic differences are driven by plasticity but given our relatively low coverage of the whitefish genome it is also possible that we have missed important genetic differences between the two morphs underlying an adaptation to the different environments.

This study identifies five SNPs potentially associated with gonad weight variation. Gonad weight is an interesting trait because of its association with spawning time and thereby also with reproductive isolation among whitefish morphs. It is known that whitefish morphs have different spawning times, habitats, and size at maturity (Svärdson 1979; Vonlanthen et al. 2009; Bitz-Thorsen et al. 2020). The fact that the relative size of the gonads was largest in the littoral morph and smallest in the profundal morph (Fig. S5) makes sense as the former is an earlier breeder than the latter (Kahilainen et al. 2014). That differences in relative size of the gonads among the morphs were further mirrored in corresponding differences in frequency of five SNPs potentially associated with gonad size suggests that these SNPs are likely associated with breeding time differences among the morphs. Further weight on this inference is provided by the fact that in the lake whitefish (*Coregonus clupeaformis*), gonadosomatic index maps to chromosome 28 similarly to one of the SNPs identified by us (Gagnaire et al. 2013a). In closely related salmonids, chromosome 28 is related to sex determination in brown trout (*Salmo trutta*), body size in Arctic charr (*Salvelinus alpinus*), and links in multigene cluster with chromosome 22 in Atlantic salmon (*Salmo salar*) (Gharbi et al. 2005; Li et al. 2011; Norman et al. 2011). Furthermore, two of the SNPs associated with gonad weight variation displayed strong long-range LD ($r^2 > 0.9$) across chromosomes (*Chr.* 22 and 36), a pattern that could be indicative of strong selection (Lewontin and Kojima 1960; Nei and Li 1973). Interchromosomal LD among loci under strong selection is not exceptional, as exemplified by strong patterns of interchromosomal LD associated with freshwater adaptation in sticklebacks (Fang et al. 2020). However, the possibility that the reported high long-range LD is due to genome misassemblies cannot be excluded. Selection for different spawning times may generate strong reproductive barriers and maintain reproductive isolation in syntopy as suggested by simulation models (Thibert-Plante et al. 2020) and empirical studies on European and Baltic flounders (Momigliano et al. 2017, 2018) and Atlantic cod (Fevolden and Pogson 1997; Hemmer-Hansen et al. 2013; Berg et al. 2016).

It is noteworthy that we found no SNP associated with gill raker number variation, a highly heritable trait playing a central role in foraging ecology and adaptive radiations of coregonids (Bernatchez 2004; Rogers and Bernatchez 2007; Kahilainen et al.

2011) nor any of the other measured traits. This is possibly a result of the small sample size of 41 admixed individuals and limited number of markers, which lowers the statistical power of GWA analyses (Hong and Park 2012). This could be particularly important if variation in gill raker number is a polygenic trait, which is notoriously difficult to map using GWA (Lin et al. 2021).

In conclusion, the results provide strong evidence that the divergence of the littoral and pelagic morphs in Lake Tsahkal predates their introduction to the lake. Although reproductive isolation among the two morphs is incomplete, their genetic differentiation suggests that spawning time differences have likely evolved in allopatry, and the morphs are currently maintaining reproductive isolation in the absence of other clear barriers to gene flow. The identification of five SNPs likely associated with gonad weight variation, a proxy for spawning time, lends support to this hypothesis and provides a starting point to identify causal loci underlying reproductive isolation. These analyses provide a starting point for further investigations on the genetic architecture of spawning time and the role of this trait in maintaining reproductive isolation, which should focus on expanding sample size, the number of studied populations, and the number of markers used.

ACKNOWLEDGMENTS

We thank T. Holopainen and P. Nieminen for their help in field work, M. Issakainen for help with laboratory work, and P. Kemppainen for advice and technical support in admixture mapping. The research is supported by the Academy of Finland (grants #129662, #134728, and #218343 to JM; #316294 to PM; #1140903 to KK), European Regional Development Fund (#A30205 to KK), the Finnish Cultural Foundation (Säätiöiden post doc-pooli; #00211290 to BF), and the Spanish Agencia Estatal de Investigación and European Union “NextGenerationEU/PRTR” (IJC2020-042611-I/MCIN/AEI/10.13039/501100011033 to PM).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

JM and PM conceived the study. BF and PM conducted analyses. BF, JM, and PM wrote the manuscript. KK found phenotypic divergence, collected samples, measured ecomorphological data, and provided constructive comments on the manuscript. BF performed visualization. All authors approved the final version of the manuscript.

DATA ARCHIVING

Information about samples used is provided in Table S1. Raw sequence data have been uploaded to NCBI (PRJNA761129). The bioinformatic scripts used are deposited at Zenodo (<https://doi.org/10.5281/zenodo.5519633>).

REFERENCES

- Albertson, R., Markert, J., Danley, P. & Kocher, T. (1999) Phylogeny of a rapidly evolving clade: the cichlid fishes of Lake Malawi, East Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 5107–5110.

- Bell, M.A., Aguirre, W.E. & Buck, N.J. (2004) Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution*, 58, 814–824.
- Benjamini, Y. & Yekutieli, D. (2001) The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics*, 29, 1165–1188.
- Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 57, 289–300.
- Berg, P.R., Star, B., Pampoulie, C., Sodeland, M., Barth, J.M., Knutsen, H., Jakobsen, K.S. & Jentoft, S. (2016) Three chromosomal rearrangements promote genomic divergence between migratory and stationary ecotypes of Atlantic cod. *Scientific Reports*, 6, 1–12.
- Bernatchez, L. (2004) Ecological theory of adaptive radiation: an empirical assessment from coregonine fishes (*Salmoniformes*). Pp. 175–207 in A.P., Hendry, & S.C., Stearns, eds. *Evolution illuminated: salmon and their relatives*. Oxford Univ. Press, Oxford, U.K.
- Bernatchez, L., Renaut, S., Whiteley, A.R., Derome, N., Jeukens, J., Landry, L., Lu, G., Nolte, A.W., Ostbye, K., Rogers, S.M., et al. (2010) On the origin of species: insights from the ecological genomics of lake whitefish. *Philosophical Transactions of the Royal Society of London B Biological Sciences*, 365, 1783–1800.
- Bitz-Thorsen, J., Häkli, K., Bhat, S. & Præbel, K. (2020) Allochrony as a potential driver for reproductive isolation in adaptive radiations of European whitefish ecomorphs. *Ecology of Freshwater Fish*, 29, 40–49.
- Bolnick, D.I. & Fitzpatrick, B.M. (2007) Sympatric speciation: models and empirical evidence. *Annual Review of Ecology Evolution and Systematics*, 38, 459–487.
- Bush, G.L., Case, S., Wilson, A. & Patton, J. (1977) Rapid speciation and chromosomal evolution in mammals. *Proceedings of the National Academy of Sciences*, 74, 3942–3946.
- Coffman, A.J., Hsieh, P.H., Gravel, S. & Gutenkunst, R.N. (2016) Computationally efficient composite likelihood statistics for demographic inference. *Molecular Biology and Evolution*, 33, 591–593.
- Danley, P.D. & Kocher, T.D. (2001) Speciation in rapidly diverging systems: lessons from Lake Malawi. *Molecular Ecology*, 10, 1075–1086.
- De-Kayne, R., Zoller, S., & Feulner, P. G. D. (2020). A de novo chromosome-level genome assembly of *Coregonus* sp. “Balchen”: One representative of the Swiss Alpine whitefish radiation. *Molecular Ecology Resources*, 20(4), 1093–1109. Portico. <https://doi.org/10.1111/1755-0998.13187>
- Fang, B., Kempainen, P., Momigliano, P., Feng, X. & Merilä, J. (2020) On the causes of geographically heterogeneous parallel evolution in sticklebacks. *Nature Ecology and Evolution*, 4, 1105–1115.
- Fang, B., Kempainen, P., Momigliano, P. & Merilä, J. (2021) Population structure limits parallel evolution in sticklebacks. *Molecular Biology and Evolution*, 38, 4205–4221.
- Feulner, P.G. D. & Seehausen, O. (2019) Genomic insights into the vulnerability of sympatric whitefish species flocks. *Molecular Ecology*, 28, 615–629.
- Fevolden, S.-E. & Pogson, G.H. (1997) Genetic divergence at the synaptophysin (*Syp I*) locus among Norwegian coastal and north-east Arctic populations of Atlantic cod. *Journal of Fish Biology*, 51, 895–908.
- Frei, D., De-Kayne, R., Selz, O.M., Seehausen, O. & Feulner, P.G. D. (2022) Genomic variation from an extinct species is retained in the extant radiation following speciation reversal. *Nature Ecology and Evolution*, 6, 461–468.
- Gagnaire, P.A., Normandeau, E., Pavey, S.A. & Bernatchez, L. (2013a) Mapping phenotypic, expression and transmission ratio distortion QTL using RAD markers in the lake whitefish (*Coregonus clupeaformis*). *Molecular Ecology*, 22, 3036–3048.
- Gagnaire, P.A., Pavey, S.A., Normandeau, E. & Bernatchez, L. (2013b) The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution*, 67, 2483–2497.
- Gharbi, K., Ferguson, M.M. & Danzmann, R.G. (2005) Characterization of Na, K-ATPase genes in Atlantic salmon (*Salmo salar*) and comparative genomic organization with rainbow trout (*Oncorhynchus mykiss*). *Molecular Genetics and Genomics*, 273, 474–483.
- Gienapp, P., Teplitsky, C., Alho, J., Mills, J. & Merilä, J. (2008) Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology*, 17, 167–178.
- Gompert, Z., Mandeville, E.G. & Buerkle, C.A. 2017. Analysis of population genomic data from hybrid zones. *Annual Review of Ecology, Evolution, and Systematics* 48:207–229.
- Häkli, K., Ostbye, K., Kahilainen, K.K., Amundsen, P.A. & Praebel, K. (2018) Diversifying selection drives parallel evolution of gill raker number and body size along the speciation continuum of European whitefish. *Ecology and Evolution*, 8, 2617–2631.
- Harrod, C., Mallela, J. & Kahilainen, K.K. (2010) Phenotype-environment correlations in a putative whitefish adaptive radiation. *Journal of Animal Ecology*, 79, 1057–1068.
- Hayden, B., Harrod, C. & Kahilainen, K.K. (2014) Lake morphometry and resource polymorphism determine niche segregation between cool- and cold-water-adapted fish. *Ecology*, 95, 538–552.
- Hemmer-Hansen, J., Nielsen, E.E., Therkildsen, N.O., Taylor, M.I., Ogden, R., Geffen, A.J., Bekkevold, D., Helyar, S., Pampoulie, C. & Johansen, T. (2013) A genomic island linked to ecotype divergence in Atlantic cod. *Molecular Ecology*, 22, 2653–2667.
- Hendry, A.P. & Day, T. (2005) Population structure attributable to reproductive time: isolation by time and adaptation by time. *Molecular Ecology*, 14, 901–916.
- Hendry, A.P., Wenburg, J.K., Bentzen, P., Volk, E.C. & Quinn, T.P. (2000) Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science*, 290, 516–518.
- Hong, E.P. & Park, J.W. (2012) Sample size and statistical power calculation in genetic association studies. *Genomics & Informatics*, 10, 117–122.
- Hudson, A.G., Vonlanthen, P., Mueller, R. & Seehausen, O. (2007) Review: the geography of speciation and adaptive radiation in coregonines. *Advances in Limnology*, 60, 111–146.
- Hudson, A.G., Vonlanthen, P. & Seehausen, O. (2011) Rapid parallel adaptive radiations from a single hybridogenic ancestral population. *Proceedings of the Royal Society B: Biological Sciences*, 278, 58–66.
- Jacobs, A., Carruthers, M., Eckmann, R., Yohannes, E., Adams, C.E., Behrmann-Godel, J. & Elmer, K.R. (2019) Rapid niche expansion by selection on functional genomic variation after ecosystem recovery. *Nature Ecology and Evolution*, 3, 77–86.
- Jouganous, J., Long, W., Ragsdale, A.P. & Gravel, S. (2017) Inferring the joint demographic history of multiple populations: beyond the diffusion approximation. *Genetics*, 206, 1549–1567.
- Kahilainen, K. & Østbye, K. (2006) Morphological differentiation and resource polymorphism in three sympatric whitefish *Coregonus lavaretus* (L.) forms in a subarctic lake. *Journal of Fish Biology*, 68, 63–79.
- Kahilainen, K.K., Siwertsson, A., Gjelland, K.O., Knudsen, R., Bøhn, T. & Amundsen, P.A. (2011) The role of gill raker number variability in adaptive radiation of coregonid fish. *Evolutionary Ecology*, 25, 573–588.
- Kahilainen, K.K., Patterson, W.P., Sonninen, E., Harrod, C. & Kiljunen, M. (2014) Adaptive radiation along a thermal gradient: preliminary results of habitat use and respiration rate divergence among whitefish morphs. *PLoS ONE*, 9, e112085.

- Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.-y., Freimer, N.B., Sabatti, C. & Eskin, E. (2010) Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics*, 42, 348-354.
- Kempainen, P., Knight, C.G., Sarma, D.K., Hlaing, T., Prakash, A., Maung Maung, Y.N., Somboon, P., Mahanta, J. & Walton, C. (2015) Linkage disequilibrium network analysis (LDna) gives a global view of chromosomal inversions, local adaptation and geographic structure. *Molecular Ecology Resources*, 15, 1031-1045.
- Korneliusson, T.S., Albrechtsen, A. & Nielsen, R. (2014) ANGSD: analysis of next generation sequencing data. *BMC Bioinformatics*, 15, 356.
- Kornfield, I. (1978) Evidence for rapid speciation in African cichlid fishes. *Experientia*, 34, 335-336.
- Lamichhaney, S., Han, F., Webster, M.T., Andersson, L., Grant, B.R. & Grant, P.R. (2018) Rapid hybrid speciation in Darwin's finches. *Science*, 359, 224-228.
- Langmead, B. & Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9, 357-359.
- Lescak, E.A., Bassham, S.L., Catchen, J., Gelmond, O., Sherbick, M.L., von Hippel, F.A. & Cresko, W.A. (2015) Evolution of stickleback in 50 years on earthquake-uplifted islands. *Proceedings of the National Academy of Sciences of the United States of America*, 112, E7204-7212.
- Lewontin, R. & Kojima, K.I. (1960) The evolutionary dynamics of complex polymorphisms. *Evolution*, 14, 458-472.
- Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25, 1754-1760.
- Li, J., Phillips, R., Harwood, A., Koop, B. & Davidson, W. (2011) Identification of the sex chromosomes of brown trout (*Salmo trutta*) and their comparison with the corresponding chromosomes in Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*). *Cytogenetic and Genome Research*, 133, 25-33.
- Li, Z., Kempainen, P., Rastas, P. & Merilä, J. (2018) Linkage disequilibrium clustering-based approach for association mapping with tightly linked genomewide data. *Molecular Ecology Resources*, 18, 809-824.
- Lin, H., Zhou, Z., Zhao, J., Zhou, T., Bai, H., Ke, Q., Pu, F., Zheng, W. & Xu, P. (2021) Genome-wide association study identifies genomic loci of sex determination and gonadosomatic index traits in large yellow croaker (*Larimichthys crocea*). *Marine Biotechnology*, 23, 127-139.
- Marques, D.A., Lucek, K., Sousa, V.C., Excoffier, L. & Seehausen, O. (2019a) Admixture between old lineages facilitated contemporary ecological speciation in Lake Constance stickleback. *Nature Communications*, 10, 4240.
- Marques, D.A., Meier, J.I. & Seehausen, O. (2019b) A combinatorial view on speciation and adaptive radiation. *Trends in Ecology & Evolution*, 34, 531-544.
- Matthews, B., Best, R.J., Feulner, P.G., Narwani, A. & Limberger, R. (2018) Evolution as an ecosystem process: insights from genomics. *Genome*, 61, 298-309.
- Meisner, J. & Albrechtsen, A. (2018) Inferring population structure and admixture proportions in low-depth NGS data. *Genetics*, 210, 719-731.
- Merilä, J. & Hendry, A.P. (2014) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications*, 7, 1-14.
- Momigliano, P., Jokinen, H., Fraimout, A., Florin, A.B., Norkko, A. & Merila, J. (2017) Extraordinarily rapid speciation in a marine fish. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 6074-6079.
- Momigliano, P., Denys, G.P. J., Jokinen, H. & Merilä, J. (2018) *Platichthys solemdali* sp. nov. (Actinopterygii, Pleuronectiformes): a new flounder species from the Baltic Sea. *Frontiers in Marine Science*, 5, 225.
- Momigliano, P., Florin, A.B. & Merila, J. (2021) Biases in demographic modeling affect our understanding of recent divergence. *Molecular Biology and Evolution*, 38, 2967-2985.
- Nei, M. (1987) Molecular evolutionary genetics. Columbia Univ. Press, New York.
- Nei, M. & Li, W.-H. (1973) Linkage disequilibrium in subdivided populations. *Genetics*, 75, 213-219.
- Norman, J.D., Danzmann, R.G., Glebe, B. & Ferguson, M.M. (2011) The genetic basis of salinity tolerance traits in Arctic charr (*Salvelinus alpinus*). *BMC Genetics*, 12, 81.
- Nosil, P. (2012) Ecological speciation. Oxford Univ. Press, Oxford, U.K.
- Østbye, K., Bernatchez, L., Naesje, T., Himberg, K.J. & Hindar, K. (2005) Evolutionary history of the European whitefish *Coregonus lavaretus* (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers. *Molecular Ecology*, 14, 4371-4387.
- Präbel, K., Knudsen, R., Siwertsson, A., Karhunen, M., Kahilainen, K.K., Ovaskainen, O., Ostbye, K., Peruzzi, S., Fevolden, S.E. & Amundsen, P.A. (2013) Ecological speciation in postglacial European whitefish: rapid adaptive radiations into the littoral, pelagic, and profundal lake habitats. *Ecology and Evolution*, 3, 4970-4986.
- Rogers, S. & Bernatchez, L. (2007) The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp. Salmonidae) species pairs. *Molecular Biology and Evolution*, 24, 1423-1438.
- Rougeux, C., Bernatchez, L. & Gagnaire, P.-A. (2017) Modeling the multiple facets of speciation-with-gene-flow toward inferring the divergence history of lake whitefish species pairs (*Coregonus clupeaformis*). *Genome Biology and Evolution* 9:2057-2074.
- Rougeux, C., Gagnaire, P.-A., Praebel, K., Seehausen, O. & Bernatchez, L. (2019a) Polygenic selection drives the evolution of convergent transcriptomic landscapes across continents within a Nearctic sister species complex. *Molecular Ecology*, 28, 4388-4403.
- Rougeux, C., Gagnaire, P.A. & Bernatchez, L. (2019b) Model-based demographic inference of introgression history in European whitefish species pairs. *Journal of Evolutionary Biology*, 32, 806-817.
- Roux, C., Fraisse, C., Romiguier, J., Anciaux, Y., Galtier, N. & Bierne, N. (2016) Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, 14, e2000234.
- Salisbury, S.J. & Ruzzante, D.E. (2022) Genetic causes and consequences of sympatric morph divergence in salmonidae: a search for mechanisms. *Annual Review of Animal Biosciences*, 10, 81-106.
- Salzburger, W. (2018) Understanding explosive diversification through cichlid fish genomics. *Nature Reviews Genetics*, 19, 705-717.
- Schluter, D. 1996. Ecological speciation in postglacial fishes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 351:807-814.
- Schluter, D.. (2000) The ecology of adaptive radiation. Oxford Univ. Press, Oxford, U.K.
- Simpson, G.G. (1944) Tempo and mode in evolution. Columbia Univ. Press, New York.
- Siwertsson, A., Knudsen, R., Kahilainen, K.K., Präbel, K., Primicerio, R. & Amundsen, P.-A. (2010) Sympatric diversification as influenced by ecological opportunity and historical contingency in a young species lineage of whitefish. *Evolutionary Ecology Research*, 12, 929-927.
- Skotte, L., Korneliusson, T.S. & Albrechtsen, A. (2012) Association testing for next-generation sequencing data using score statistics. *Genetic Epidemiology*, 36, 430-437.
- Skúlason, S., Parsons, K.J., Svanbäck, R., Räsänen, K., Ferguson, M.M., Adams, C.E., Amundsen, P.A., Bartels, P., Bean, C.W. & Boughman, J.W. (2019) A way forward with eco evo devo: an extended theory

- of resource polymorphism with postglacial fishes as model systems. *Biological Reviews*, 94, 1786-1808.
- Svärdson, G.. (1979) Speciation of Scandinavian *Coregonus*. *Reports of the Institute of Freshwater Research, Drottningholm*, 57, 3-95.
- Taylor, R.S. & Friesen, V.L. (2017) The role of allochrony in speciation. *Molecular Ecology*, 26, 3330-3342.
- Thibert-Plante, X., Præbel, K., Østbye, K., Kahilainen, K.K., Amundsen, P.-A. & Gavrillets, S. (2020) Using mathematical modelling to investigate the adaptive divergence of whitefish in Fennoscandia. *Scientific Reports*, 10:1-12.
- Thomas, S.M., Harrod, C., Hayden, B., Malinen, T. & Kahilainen, K.K. (2017) Ecological speciation in a generalist consumer expands the trophic niche of a dominant predator. *Scientific Reports*, 7, 8765.
- Vonlanthen, P., Roy, D., Hudson, A., Largiadèr, C.R., Bittner, D. & Seehausen, O. (2009) Divergence along a steep ecological gradient in lake whitefish (*Coregonus* sp.). *Journal of Evolutionary Biology*, 22, 498-514.
- Vonlanthen, P., Bittner, D., Hudson, A.G., Young, K.A., Müller, R., Lundsgaard-Hansen, B., Roy, D., Di Piazza, S., Largiadèr, C.R. & Seehausen, O. (2012) Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature*, 482, 357-362.
- Wang, S., Meyer, E., McKay, J.K. & Matz, M.V. (2012) 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nature Methods*, 9, 808-810.
- Zhou, X. & Stephens, M. (2014) Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods*, 11, 407-409.

Associate Editor: J. Sachs
 Handling Editor: Dr. T. Chapman

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Detailed sample information about sequencing depth, collection date and depth together with all linear morphometric measures. Supplementary Materials