bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Ecological speciation in European whitefish is driven by a large-gaped predator

Gunnar Öhlund[1,2]✉, Mats Bodin[1,3], Karin A. Nilsson[1,4], Sven-Ola Öhlund[1], Kenyon B. Mobley[1,5,6], Alan G. Hudson[1,7], Mikael Peedu[1], Åke Brännström[3,8], Pia Bartels[1], Kim Præbel[9], Catherine L. Hein[1,10], Petter Johansson[1], Göran Englund[1]

[1]*Department of Ecology and Environmental Science, Umeå University, SE-901 87 Umeå, Sweden*
[2]*Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, SE-971 87 Luleå, Sweden*
[3]*Department of Mathematics and Mathematical Statistics, Umeå University, SE-901 87 Umeå, Sweden*
[4]*Department of Integrative Biology, University of Guelph, Guelph, Ontario N1G2W1, Canada*
[5]*Max Planck Institute for Evolutionary Biology, D-24302 Plön, Germany*
[6]*Organismal and Evolutionary Biology Research Program, Faculty of Biological and Environmental Sciences, PO Box 56, 00014 University of Helsinki, Finland*
[7]*School of Biological Sciences, University of Bristol, Life Sciences Building, Bristol BS8 1TQ, UK*
[8]*Evolution and Ecology Program, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria*
[9]*Norwegian College of Fishery Science, UiT The Arctic University of Norway, N-9037 Tromsø, Norway*
[10]*Climate Impacts Research Centre (CIRC), Abisko Scientific Research Station, SE-981 07 Abisko, Sweden*

✉Email of corresponding author Gunnar Öhlund: guggeoh@gmail.com

Abstract

Lake-dwelling fish that form species pairs/flocks characterized by body size divergence are important model systems for speciation research. While several sources of divergent selection have been identified in these systems, their importance for driving the speciation process remains elusive. A major problem is that in retrospect, we cannot distinguish selection pressures that initiated divergence from those acting later in the process. To address this issue, we reconstructed the initial stages of speciation in European whitefish (*Coregonus lavaretus*) using data from 357 populations of varying age (26-10 000 years). We find that whitefish speciation is driven by a large-growing predator, the northern pike (*Esox lucius*). Pike initiates divergence by causing a largely plastic differentiation into benthic giants and pelagic dwarfs; ecotypes that will subsequently develop partial reproductive isolation and heritable differences in gill raker number. Using an eco-evolutionary model, we demonstrate how pike's habitat specificity and large gape size are critical for imposing a between-habitat trade-off, causing prey to mature in a safer place or at a safer size. Thereby, we propose a novel mechanism for how predators may cause dwarf/giant speciation in lake-dwelling fish species.

Introduction

For several decades, the question of whether speciation can occur in the face of homogenizing gene flow was hotly debated in evolutionary biology. Today, this debate has shifted focus as research has become more occupied with understanding the processes that cause speciation with gene flow in nature[1, 2]. Examples of on-going ecological speciation in sympatry are especially common in lake-dwelling fish, as they have an intriguing propensity to form genetically distinct ecotypes that differ in ecology, morphology, and reproductive biology[3, 4]. There is substantial variation among ecosystems and species as to how far this divergence has progressed, but a common feature is the evolution of large- and small growing ecotypes along resource- and/or

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

habitat gradients in the lake environment. Examples of such ecotypic specialization include threespine sticklebacks[5], African cichlids [6], rainbow smelt[7], Arctic char[8], Dolly Varden[9], *Prosopium sp*.[10] and a number of species belonging to the genus *Coregonus*[11-14]. Although the processes underlying this pattern have been studied intensively during recent decades[12, 15-19], a fundamental question remains largely unanswered; why is divergence initiated in some populations and not in others?

To answer this question, we need to improve our understanding of how ecological mechanisms associated with habitat gradients could drive speciation. It is widely accepted that intense intraspecific competition and/or abundant ecological opportunities can cause divergent selection[20-27]. Other studies suggest that predation[6, 28, 29], spatial variation in temperature[30], environmental stress[31], and reduced habitat- and prey availability[23] can promote divergence. However, the importance of specific selective agents for actually causing speciation still remains elusive. A key problem is that divergence exposes incipient ecotypes to new ecological conditions, and the selective regime can change accordingly over time. For instance, if some ecological mechanism drives individuals to specialize in different habitats, this can cause divergent selection and conspicuous adaptations that are by-products rather than drivers of the initial divergence.

The best way to avoid confounding the causes and consequences of speciation is to study the process at its earliest stages[4, 32-39]. Unfortunately, this approach may lead us to study cases of early population divergence that are unrepresentative of the speciation process, or will never lead to speciation[4, 40]. This problem, in turn, could potentially be avoided by using comparative analyses[21, 23, 25, 26, 41, 42] to identify the environmental conditions under which we can expect a future speciation process to proceed. So far, however, these two approaches have rarely been combined.

We addressed these issues by studying whitefish in Scandinavian lakes, where they form genetically distinct ecotype pairs that differ in body size[12], morphology[12], resource use[12, 43], and time and place of spawning [12]. Besides from being found in large numbers, these ecotype pairs are typically well known among local fishers[12]; opening up the possibility to use interviews as a method for collecting large amounts of spatial comparative data. Moreover, starting in the late 18th century, there is a richly documented history of anthropogenic introductions that gave rise to new whitefish populations[44-46]. Today, the known and variable ages of these young populations provide an excellent opportunity to study how the speciation process initiates and develops over time. In this paper, we present extensive comparative data showing that northern pike is the key driver of ecological speciation in Scandinavian whitefish populations, and use data from populations of different age and modelling to form a hypothesis for why this large-growing predator is so critically important.

Results

Our interviews with local fishers revealed that out of 357 lakes distributed from southern Norway to northern Sweden (Supplementary Fig. 1), 153 harboured ecotype pairs of dwarf and giant whitefish. These ecotype pairs were generally found in lakes with relatively high species-richness of fish, an observation that provides little support for the idea that intraspecific competition and ecological opportunity are the primary drivers of ecological speciation (Supplementary Fig. 2). Instead, our analyses showed that out of 13 analysed biotic and abiotic variables, presence of northern pike together with lake area and maximum depth determine divergence patterns of whitefish populations in our study area. Pike presence induces divergence into dwarfs and giants in lakes that are large and deep enough (Fig. 1, proportion correct predictions=0.90, Cohen's κ=0.85). Smaller/shallower lakes have monomorphic whitefish, but the importance of pike for determining whitefish life histories can still be observed. Interview data on maximum weights showed that when pike are present, these lakes have either dwarf- or giant whitefish (Fig. 1a and Supplementary Fig. 3).

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

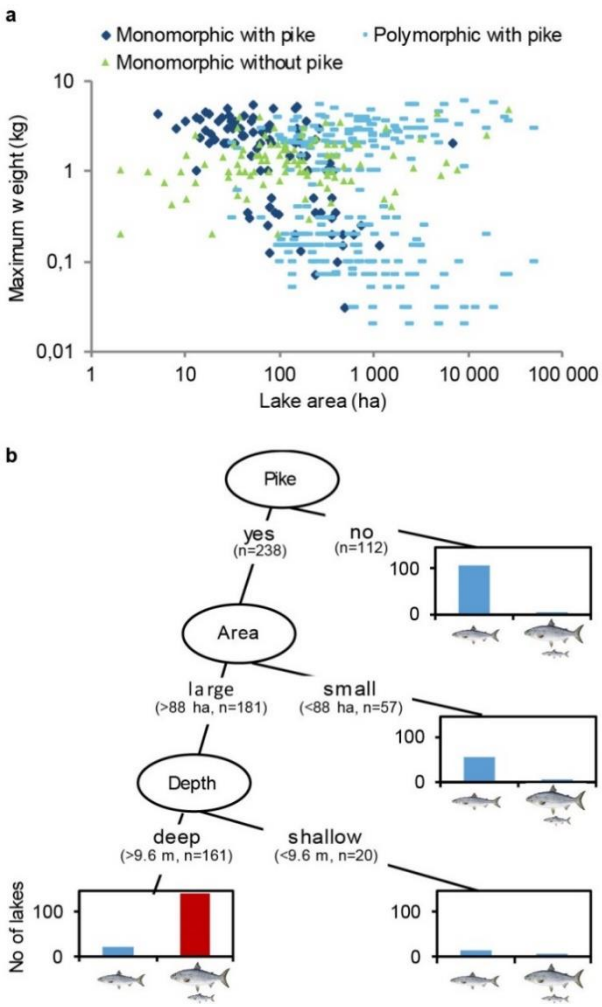

Fig. 1: Pike presence, lake area and maximum depth control the formation of dwarf and giant whitefish ecotypes. a) Maximum weight (kg) of whitefish from populations in lakes with (n=217) and without (n=103) pike as a function of lake area. Light blue symbols represent polymorphic whitefish populations for which each lake has two corresponding observations. **b)** Classification tree (based on 13 explanatory variables, n=350) for the prevalence of polymorphism in whitefish, showing that pike induces co-occurring dwarf and giant ecotypes in lakes that are large and deep enough. The y-axes show the number of lakes. Cohen's kappa for the whole model was 0.85.

Whenever fishers had knowledge about the spawning behaviour of whitefish in a lake with polymorphism, they reported segregation between dwarfs and giants in time and/or space during spawning. As dwarf- and giant whitefish ecotypes typically differ in a range of ecological traits[12], this suggests that the dwarf/giant pairs reported in our interviews represent cases of incipient ecological speciation. To test this hypothesis and validate the fishers' observations of polymorphism, we first looked for signs of reproductive isolation between sympatric dwarfs and giants using microsatellite data from 30 lakes where the interviewees had reported dual ecotypes. F_{ST}-values were significant between dwarf and giant ecotypes in 23 of these lakes, and the non-significant differences between ecotypes were only found in young whitefish populations (introduced between 1825 and 1960, Supplementary Table 1). Second, we tested for differences in gill raker counts, a trait that is under strong genetic control in European whitefish (heritability h^2 = 0.79 [12, 47]), and known to be under divergent selection during resource specialization [43, 48, 49]. When comparing dwarf and giant ecotypes in 70 lakes with reported size polymorphism, we found significant differences in 63 of them (Supplementary Table 2). Again, non-significant differences were only found in young populations. Restricting these analyses to populations that originated before the year 1900, we found that 21 out of 23 ecotype pairs had significant F_{ST}-values (mean global F_{ST}-value = 0.054, non-significant populations were introduced in Bölessjön, 1825 and Sörvikssjön, 1845, Supplementary Table 1), and that 61 out of 62 ecotype pairs differed significantly in gill raker counts (mean difference = 10.6 rakers, the non-significant one was introduced in Bomsjön, 1895, Supplementary Table 2). Except for in the youngest populations, our data thus show that the dwarf and giant ecotypes that fishers report have developed partial reproductive isolation and substantial differences in an ecologically important, heritable trait.

To understand how divergence initiates, we performed standardized gillnet sampling in 38 lakes that have recently introduced whitefish populations (introduced between 1784 and 1985); 23 where pike are present and 15 where they are absent. In order to only include lakes that are suitable for a future speciation process (see Fig. 1b), these sampling efforts were restricted to lakes that are larger than 100 hectares and deeper than 15 meters. First, we scanned the resulting data for signs of initiating body size

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

divergence, and found a strong, rapidly appearing pike effect (Fig. 2). Body size variation among adult whitefish was larger in lakes with pike than in lakes without pike, and increased with time passed since introduction (ANCOVA: pike, t=5.33, p<0.00001, time since introduction, t=2.18, p=0.036, N=38). Separate analyses of the two lake categories showed that size variation increased with time in pike lakes (slope±SE=0.084±0.039, t=2.16, p=0.042, N=23), but not in lakes where pike were absent (slope±SE=-0.0079±0.031, t=0.26, p=0.80, N=15).

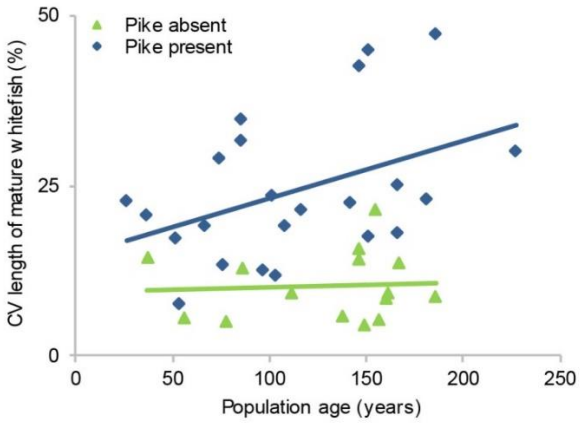

Fig. 2: Pike presence drives rapid body size divergence in whitefish. Coefficient of variation for lengths of mature whitefish in lakes with (n=23) and without (n=15) pike as a function of population age.

Next, we used data from the sampled introduction-lakes to understand how this rapid body size divergence relates to divergence in other traits. Specifically, we compared the timing of divergence in body size and habitat use, which are highly plastic traits, with that of divergence in gill raker counts and neutral genetic markers. In these analyses, we wanted to exclude any pike-presence lakes where introductions of multiple genotypes may have contributed to the observed patterns of divergence. We therefore used microsatellite data (available for 18 lakes with known introduction dates) to exclude populations with signs of introductions of multiple genotypes, a procedure that left us with 11 populations with a putatively sympatric signal (Supplementary Fig. 4).

Going forward with these 11 populations, we first wanted to compare the initial divergence rates of body size and gill raker numbers. In order not to bias the comparison between the two traits, we performed cluster analyses [50] along the two trait axes simultaneously using the individuals caught in our standardized gillnet sampling. To allow comparison with much older populations, we also included samples from nine lakes with native, polymorphic whitefish.

The analyses gave divergent clusters in all populations except one (Lake Murusjøen, where whitefish were introduced in 1975). Analyzing how between-cluster differences in body size and gill rakers depend on population age (excluding Murusjøen), we found that divergence in body size is very rapid and precedes divergence in gill rakers (Fig 3, linear regression, divergence in body size: t=0.67, N=10, p=0.52; divergence in number of gill rakers: t=3.53, N=10, p=0.0076, both regressions excluding native populations). In fact, a large portion of the body size divergence typically seen in native polymorphic populations is expressed within just a few decades (Fig. 3). Moreover, a comparison between benthic and pelagic catches in the underlying gill net samples showed that this early size divergence is accompanied by an equally rapid divergence in habitat use between dwarfs and giants (Fig 4, ANCOVA excluding native lakes: pike, t=6.39, p<0.00001, time since introduction, t=1.17, p=0.26, N=20). Gill rakers on the other hand show very little divergence between the youngest clusters, suggesting that differences in body size form the basis for the initial formation of ecotypes (Fig 3).

Next, we tested if gill raker counts differed between ecotypes within the 10 introduction lakes presented in Fig. 3. As the number of gill rakers could not be used for ecotype assignment in these tests, we classified individuals using body size and spawning site. The results from these analyses were consistent with the pattern resulting from our between-cluster comparisons. While gill raker numbers did not differ between dwarfs and giants in the youngest populations (introduced after year 1900, N=6, all t-values<1.68, all p>0.09, Supplementary Table

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

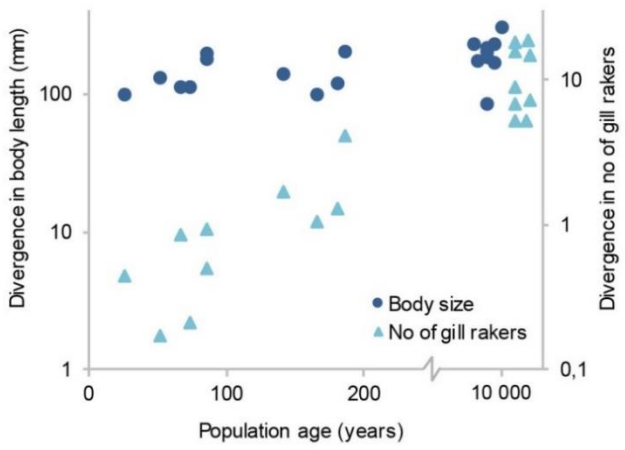

Fig. 3: Rapid body size divergence leads the way to gill raker divergence. Between-cluster differences (based on mature individuals caught in our standardized gillnet surveys, n=19) in average values of body length and gill raker number as a function of population age. The positions of native populations were adjusted along the x-axis to reduce overlap.

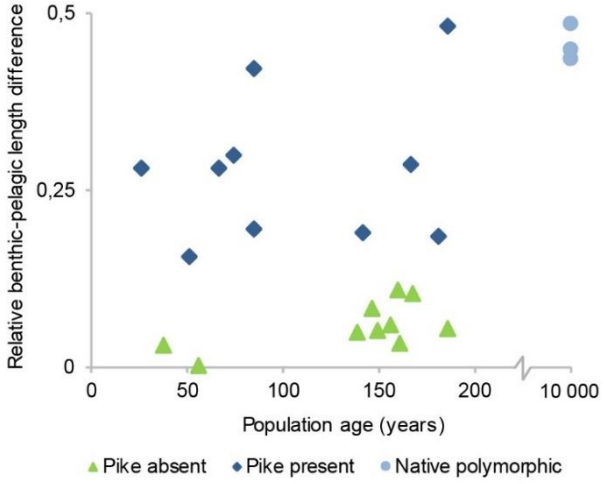

Fig. 4: Body size divergence is associated with formation of benthic and pelagic ecotypes. Average length differences of sexually mature whitefish caught in littoral-benthic and pelagic gillnets as a function of population age. The figure includes the same selection of pike lakes that is presented in Fig. 3 (n=12, see methods section for information about missing data points) and pike-free control lakes (n=10). Length differences were calculated as (mean littoral length-mean pelagic length)/mean littoral length.

2), we found small but significant differences (1.5-2.7 rakers) in 4 out of 4 dwarf/giant pairs in the populations that were introduced during the 1800s (N=4, all t-values>3.22, all p<0.0017, Supplementary Table 2).

The microsatellite data from these lakes showed a similar pattern; no significant population differentiation between ecotypes in the youngest populations but significant F_{ST}-values between ecotypes in 2 out of the 4 older ones (Supplementary Table 1). Hence, the chronosequence of introduced populations suggests a timeline of divergence where the initial formation of dwarf/giant ecotypes is followed by more slowly appearing differences in gill raker numbers and neutral genetic markers (Figs. 3 and 4, Supplementary Tables 1 and 2).

Adding the native populations to the chronosequence, the short-term pike-driven divergence observed in introduced populations and the long-term pike-driven speciation process appear to form a continuum (Figs. 3 and 4). This suggests that we can view divergence in the youngest populations as representing the initial stages of the speciation process. Alternatively, it could be argued that size and habitat divergence may not necessarily lead to heritable differences and reproductive isolation, as has been observed in other fish species that form ecotypes[51, 52]. However, this does not appear to be the case in our study system. Surveying data from large and deep pike lakes with native whitefish, we did not find a single example of a dwarf/giant pair for which divergence had remained restricted to size (significant gill raker differences in 50/50 lakes, average difference =11.7; significant F_{ST}-values in 13/13 lakes, average F_{ST}=0.061). Hence, even though we lack direct experimental evidence, our data suggest that the initially formed dwarf and giant ecotypes with high predictability will continue to diverge along the speciation continuum.

The hypothesis that size differences lead the way to reproductive isolation implies that the spawning habits of whitefish will depend on their body size. In order to assess the validity of this corollary, we collected information (interview data validated with various kinds of sample fishing, see methods section and Supplementary Fig. 5 for details) about the average size of

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

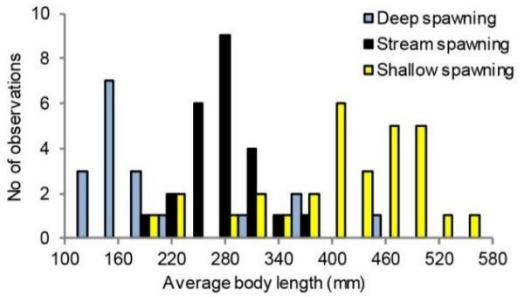

Fig. 5: Whitefish spawning behaviour is related to body size. Histogram showing the distribution of average body lengths for populations that spawn in stream habitat, shallow lake habitat (depth <4 m), or deep lake habitat (depth >4 m) (n=72).

sexually mature individuals in all whitefish populations from our study lakes for which data on both spawning habitat and gill raker numbers were available. The resulting data showed that populations of giants typically spawn in shallow lake habitat, whereas more small-growing populations spawn either in streams, or in deeper water in the lakes (Fig 5). An analysis of this data confirmed that choice of spawning habitat is related to body size but not to gill raker counts (Multinomial logistic regression with stream spawners as reference; body size: stream vs shallow, Z=3.79, p=0.00015, stream vs deep, Z=2.13, p=0.033; gill rakers: stream vs shallow, Z=0.63, p=0.53, stream vs deep, Z=0.0025, p=1.0, N=72, Fig 5).

All our empirical results thus point in the same direction; that the strong pike effect on whitefish divergence comes from a unique ability to induce pelagic dwarfs and benthic giants. To understand why pike have this ability as opposed to other potential predators (e.g. brown trout (*Salmo trutta*), arctic char (*Salvelinus alpinus*), and perch (*Perca fluviatilis*)), we must understand 1) how predation and it's feedbacks on resource competition among prey can drive divergence into pelagic dwarfs and benthic giants, and 2) how this process depends on the characteristics of the focal predator species. Pike are largely restricted to the littoral zone of lakes and stands out by having a gape size large enough to catch relatively large prey [53, 54]. To explore the consequences of the presence of a predator with these characteristics, we developed a size-structured eco-evolutionary model of the pike–whitefish system with whitefish maturation size as the evolving trait (see Supplementary Methods, Supplementary Figs. 6 and 7 and Supplementary Tables 3-5 for a detailed model description). We focused on maturation size because it is an important determinant of growth trajectories [55] that typically differs between sympatric ecotypes in our study system [12].

The model analyses suggest that habitat-specific predation can induce evolutionary divergence into dwarfs and giants by imposing a trade-off that affects life history and habitat choice of prey (Fig. 6a). The presence of pike causes whitefish to either 1) avoid pike in space at the cost of feeding on small pelagic zooplankton that provide limited scope for continued growth[56-58], or 2) grow rapidly to reach a size that is subject to low predation risk by delaying the energy-consuming maturation and using the profitable littoral resource of large benthic invertebrates (Fig. 6c and Fig. 6d). A small-gaped predator does not impose this kind of trade-off (Fig. 6b), a result that corresponds well with our empirical data showing no association between whitefish divergence and the presence of small-gaped predator species such as brown trout, arctic char and perch. The mechanism behind the strong gape size effect is that when predation risk in the littoral habitat is confined to small prey, pelagic whitefish will be able to reach a size that allows them to shift to the littoral habitat without exposing themselves to high predation risk. Thus, two prerequisites for the necessary life history trade-off are 1) that the predator is sufficiently large-gaped to limit the ability of prey to grow out of the predation window when residing in the refuge habitat only, and 2) that prey can potentially reach sexual maturity before obtaining a safe size. Hence, besides prey–resource dynamics, the scope for this kind of predator-induced divergence will depend on a balance between the gape size of the predator and the inherent growth potential and life history of the prey.

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

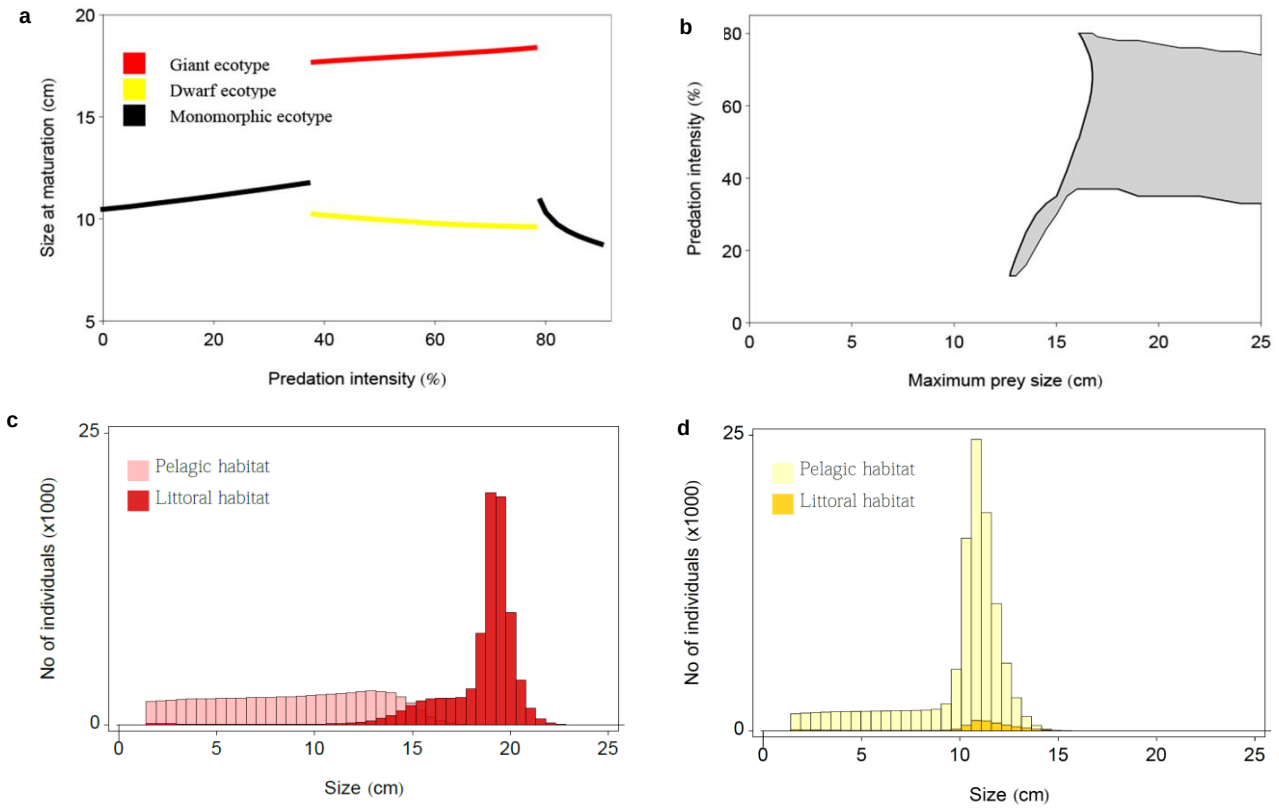

Fig. 6: Large-gaped predators can induce dwarf- and giant prey ecotypes by imposing a habitat choice-growth strategy trade-off. a) Model simulation of maturation size as a function of predation intensity from a littoral predator capable of taking prey up to a maximum size of 18 cm. The red line represents giants that mature in the littoral habitat and the yellow line represents dwarfs that mature in the pelagic zone. b) The range of predation intensities (see Supplementary Table 4 for details) that induce evolutionary divergence at different values of maximum size of prey that can be taken by the predator. c) The distribution of the giant ecotype between the pelagic habitat and the littoral habitat at the evolutionary stable state (ESS) when the littoral predator can take prey up to 18 cm and the predation intensity is 70%. The giants mature at 18.2 cm. d) The corresponding distribution of the dwarf ecotype between the two habitats. Dwarfs mature at 9.7 cm.

Discussion

In this study, we find an answer to the elusive question why benthic-pelagic ecotype pairs develop in some lakes and not in others. Contrary to popular belief, our data shows that ecological speciation along the benthic-pelagic habitat gradient is driven by a large-gaped predator. Recognizing pike's critical role in our study system, we could then target the youngest pike-exposed whitefish populations to study the initial sequence of trait changes, and use a model rich in the necessary type of ecological detail to analyse the underlying mechanisms. The results suggest that pike drives ecological speciation by inducing pelagic dwarfs and benthic giants; a primary ecotypic differentiation that forms the basis for further divergent adaptations to the respective habitats, and at the same time promotes reproductive isolation.

While previous work has shown that both gill raker numbers and body size are under divergent selection during whitefish radiations[49], our data thus suggest that divergent selection on body size and habitat use is the primary route to ecotypic differentiation and subsequent ecological speciation. While body size divergence has been described as an important component of niche differentiation during ecological speciation in other systems[6, 59], the full, ecological implications of size differences have received relatively little attention in studies of ecological speciation in fish. Unlike other morphological traits, body size determines both an individual's potential gain from feeding on a given food type and its

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

exposure to predation risk while doing so[60]. As a consequence, small and large individuals that face between-habitat variation in resource gain and predation risk will often specialize on feeding in different habitats[61, 62]. At the same time, individual growth depends on the density and quality of available resources[60], and feeding on small- or large prey can affect ontogenetic growth trajectories differently[56, 63]. This fundamental property of body size, i.e. that it both determines and is affected by an individual's ecological niche, is a critical component of the trade-off that gives body-size divergence in our model. Hence, our findings are consistent with the idea that phenotypic plasticity is important for speciation[51, 64-66].

Through the plasticity of food-dependent growth, inherently small-growing individuals can be scared into sacrificing growth opportunities, while inherently large-growing individuals can gain access to resources that allow continued, rapid growth. This way, food-dependent growth can greatly enhance the adaptive significance of heritable body size variation. In our model, such variation is represented by differences in maturation size; a major source of growth trajectory variation among fish populations [55], and a typical feature of whitefish radiations [67]. However, any trait variation that affects individual growth could potentially sort individuals along a gradient of size-dependent resource gain/predation risk. Our model should therefore be viewed as the most straightforward representation of a more general idea; that gape limited predation can cause individual prey to either stay in refuge habitat, or maximize growth to reach a safe size, depending on their inherent growth potential.

While our results improve our understanding of how benthic and pelagic ecotypes form, they offer more limited insight into how giants and dwarfs continue to diverge towards speciation; a process that requires assortative mating and some form of heritability that transfers the growth strategies and their spawning behaviour between generations. Predation risk could potentially explain the association between body size and choice of spawning sites in much the same way as with size-dependent habitat choice outside of the spawning season. This remains to be tested, but the size-spawning site association nevertheless provides a plausible explanation for why dwarfs and giants would develop reproductive isolation over time. Our study thus contributes to a growing body of evidence suggesting that differences in body size may be an important driver of reproductive isolation in polymorphic fish populations[68, 69]. When it comes to the inheritance of adult size, a specific mechanism remains to be demonstrated. It could come from genetically controlled differences in maturation time or size, but there are other possible mechanisms by which size differences could be transferred between generations. For example, size-dependent choice of spawning sites could feed back on the hatching time and early growth of offspring because the different spawning habitats have different temperature regimes[15, 70]. Moreover, dwarfs produce smaller eggs than giants[71] and it has been shown that this impedes the initial growth of their offspring[72]. Demonstrating the mechanisms that cause reproductive isolation between dwarfs and giants will be an important challenge for future research.

The phenomenon that fish populations form sympatric, large- and small-growing ecotypes has been repeated in a large number of species, and along all major habitat axes in lakes[5, 8, 73-75]. If this parallelism is mirrored in the underlying mechanisms, our results suggest that predation is heavily underestimated as a driver of intraspecific fish diversity in lakes. While our results apply directly to divergence along the benthic-pelagic resource axis, the type of trade-off that gives divergence in our model could appear along any gradient where small prey fish take refuge in suboptimal growth conditions. Such growth conditions can come from spatial variation in a range of environmental variables, and do not necessarily depend on the presence of discrete, habitat-specific resource types. Hence, predator-induced trade-offs could potentially explain why dwarf- and giant ecotypes form also in situations where diet specialization is less pronounced or even absent[8, 76-78].

To test the hypothesis that predator-induced growth strategies are generally important

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

as a starting point of ecological speciation, we need to disentangle leaders and followers among the selection pressures and diverging traits that are involved when ecotype pairs form. Our study illustrates how this can be achieved by combining comparative and temporal data, as this can allow us to both identify crucial selection pressures and study their effects on populations over time. Applied to a variety of systems and including a wide range of study methods, this approach holds great promise to improve our understanding of how ecology initiates speciation with gene flow.

Methods

We used data from 357 Scandinavian lakes distributed along a south–north gradient from southern Norway (58.99 N, 8.29 E) to northern Sweden (68.17 N, 21.97 E) (Supplementary Fig. 1).

Data collection

Interviews. Local fishers often have detailed knowledge about the habits and spawning sites of whitefish ecotypes in Scandinavian lakes, and co-occurring ecotypes typically have distinct local names[12]. This allowed us to use interviews to assess large-scale patterns of maximum body size and the frequency of polymorphism. We asked local fishers (and other persons with relevant knowledge) if the whitefish in a given lake was indigenous or introduced, if there were one or more ecotypes, and for the maximum weight, spawning site and spawning time of each ecotype. Care was taken to follow the same interview protocol for all lakes. In order to estimate maximum size, we asked about the largest specimen caught in a given lake during the last 25-year period. We used maximum weight as a crude life history metric because fishers tend to remember this figure and because it effectively captures the divergence between dwarfs and giants. We defined polymorphism as the existence of two or more coexisting populations with different maximum sizes. When deciding whether or not whitefish populations were polymorphic, lakes were divided into the following four categories. 1) Fishers report two or more populations with different maximum sizes that use different spawning grounds and/or differ in spawning time. 2) Fishers report two ecotypes that differ in maximum size but could not provide information about spawning. 3) Fishers report indications of polymorphism, such as presence of both large and dwarfed spawners and size-related differences in parasite load, but feel uncertain if these represent different ecotypes. 4) Fishers report that, to the best of their knowledge, there is only one ecotype of whitefish. In the final data set, we defined lakes from categories 1 (n=105) and 2 (n=51) as being polymorphic and lakes from category 4 (n=197) as being monomorphic. Lakes in category 3 (n=22) were excluded, with the exception of four lakes where we had performed standardized sample fishing.

For a subset of lakes (N=72) used to analyze the association between spawning habitat and phenotype (average body size and gill raker number, see below for more details), we also asked fishers about the water depth at the spawning sites and the average size of spawning individuals.

Publications and official records. Data from publications and official records were mainly used to assess the age and origin of populations, and a large proportion of the records of year of introduction in our data set originate from Swedish and Norwegian reports that were published between 1797 and 2013[44, 46, 79, 80]. In cases where the time of introduction was given in a time-span of up to 20 years, we used the middle year. Published data were also included in analyses of differences in gill raker counts and neutral genetic markers (Supplementary Tables 1 and 2). Finally, for the analyses of phenotype-spawning habitat correlation, we used published information about spawning depth[12, 81-83] (N=11) and average body size[76, 81, 83, 84] (N=7) for populations where interviews did not provide this information.

Field sampling. To validate interview data and to catch fish for genetic and phenotypic analyses, we performed standardized gillnet sampling in 51 of the interview lakes, and some form of non-standardized sampling (gillnetting, hand netting

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

or ice fishing) on spawning grounds in 22 of them (or their adjacent streams).

As a standard gillnet setup, we used 24 benthic gillnets (30x1.5 m; 8 of multimesh-type, 4 with panels of 33 mm and 12 with 45 mm mesh size knot to knot) and 8 floating gillnets (two of multimesh type (27x6 m) and six single-meshed nets (30x5m) with mesh sizes 12, 15, 20, 23, 30 and 38 mm). In a subset of the sampled lakes, the standard setup was extended to include two extra floating gillnets with mesh sizes of 33 and 45 mm. Including these mesh sizes in the pelagic setup allowed us to use the combined catches from multimesh, 33mm and 45 mm nets to directly compare the average size of sexually mature whitefish in the benthic and pelagic habitats respectively (Fig. 4). We have performed this extended sampling in 13 of the lakes included in Fig. 3 and in 10 pike-free control lakes. Hence, the data points in Fig. 3 that are missing in Fig. 4 (6/9 of the native populations) are missing because the gillnet setup used in these lakes did not allow the relevant between-habitat comparison of average body size.

Phenotypic data. The number of gill rakers on the first left gill arch were counted under a dissecting microscope. We present gill raker data from ecotype pairs in 72 lakes, out of which 50 had putatively native- and 22 had introduced whitefish populations. In 35 of these lakes, the gill raker counts were based on our own samples, and in the remaining 37, we used published data (Supplementary Table 2). In lakes with more than two ecotypes, we compared the gill raker count of the largest and the smallest ecotype. For the analysis relating average phenotype to spawning habitat (see below), we recorded gill raker means for 10 additional lakes where data were available for only one population. Body length and sexual maturity status were recorded in the field.

Analyses

Our interview-based data set contains data from 357 lakes, and all non-interview based data comes from subsets of these lakes. Populations of recent, monomorphic origin cannot be expected to be polymorphic, and may experience rapidly changing growth conditions. Therefore, we did not include whitefish populations introduced after 1960 (the most recent introduction year that has given rise to a polymorphic population according to our interviews) in Fig. 1, Supplementary Figs. 2 and 3, and the underlying analyses. For all other analyses, we used the maximum number of lakes that was applicable and for which we had relevant data. This means that the number of lakes included in different analyses vary, either because interviews did not result in complete data for all questions, or because non-interview data was not available for all lakes.

For statistical analyses, including linear regression, ANCOVA, and t-test, we scanned residual plots for heteroscedasticity, outliers, and model misspecification. When motivated, we used logarithmic or square root transformations to reduce heteroscedasticity and the influence of outlying observations. For logistic regression analyses, we scanned Pearson and deviance residuals for outliers. No outliers or signs of model misspecification were detected.

Ecological drivers of polymorphism. Relationships between environmental variables and the prevalence of polymorphism were modelled with classification trees, estimated and crossvalidated with the *rpart* module in R[85]. Thirteen variables were used as predictors: the number of fish species co-occurring with whitefish, lake area, maximum depth, altitude, temperature sum (total number of degree days above 6 ^{o}C), and presence/absence of the fish species pike, roach (*Rutilius rutilus*), grayling (*Thymallus thymallus*), burbot (*Lota lota*), Eurasian perch (*Perca fluvialitis*), arctic charr (*Salvelinus alpinus*), brown trout (*Salmo trutta*), and European minnow (*Phoxinus phoxinus*). Optimal tree depth was determined with cross validation and the agreement between data and model predictions was judged with Cohen's κ-statistics[86].

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*Divergence vs population age***.** To analyse the relationship between population age and the degree of divergence in body size and gill raker counts (Fig. 3), we first used the procedure mclustICL in the R module mclust[50] to identify clusters based on body length and gill raker counts from mature whitefish caught in our standardized gillnet sampling. Missing data were imputed with the imputeData command in the mix package[87]. The difference between mean values for the clusters in a lake was then used as a measure of divergence in body size and number of gill rakers. If more than two clusters were identified, we excluded the intermediate ones.

When gill raker comparisons were made between individuals that were preassigned to ecotype, we compared mature small individuals (<25 cm) and large individuals (>35 cm) caught either on their spawning grounds (dwarf sample from six lakes) or from sampling not associated to spawning grounds (dwarf sample from four lakes and giant sample from all 10 lakes). In one lake (Stor-Skirsjön), whitefish rarely grow larger than 35 cm, and we therefore compared the mature dwarfs (average length 182 mm) to fish >275mm.

Phenotype-spawning habitat correlation. Our analysis of the correlation between whitefish phenotype and spawning habitat included populations for which we could get information about spawning habitat and average body size and gill raker number. To ensure that all populations included in the analysis had potential access to all categories of spawning habitat, populations from small and/or shallow lakes (<100 ha, <15 m maximum depth) were excluded. Altogether, 72 whitefish populations from 48 lakes filled these criteria.

Information about spawning depth and habitat were used to categorize populations as stream spawners, shallow lake spawners (depth ≤4 m) or deep lake spawners (depth >4m). The data were then analysed with multinomial regression (multinom procedure in nnet module of R[88]) using average body size and no of gill rakers as predictors and the three spawning categories as response. As the fishers' estimates of average size could be biased by the type of gear they used, we assessed the robustness of this data by comparing individual interview data points to corresponding average sizes from our own samples. This comparison was partly based on the subset of populations that we had targeted with sampling on their spawning grounds (N=22), using non-standardized gillnet sampling (n=5), ice fishing (n=3) or hand netting (n=17, i.e. some populations were sampled with more than one method). We also included average sizes from the standardized sample fishing (not performed on spawning grounds) if the given population/ecotype could be separated from the rest of the catch by visual inspection of size- and gill raker data (n=21). Regardless of sampling method, the interview data correlated well with our sample data (Supplementary Fig. 5).

Genetic analyses. To identify genetic divergence indicative of reproductive isolation and to investigate the structuring of genetic diversity among and within the introduced whitefish populations, we compared neutral microsatellite genotypic data for ecotypes in 32 lakes. We performed population genetic analyses in 30 of these lakes, and extracted data from the published literature for the remaining two [83] (see Supplementary Table 1). 18 of the analysed lakes have whitefish populations originating from introductions between 1784 and 1985. One lake (Valsjön) has conflicting information about the introduction date, and 13 lakes have purportedly native whitefish. Individual fish were assigned to ecotype either through sampling on ecotype-specific spawning grounds or through separation of adult fish based on differences in size and morphology.

Population genetic analyses of sampled whitefish were carried out on genotypes derived from two fully overlapping marker panels comprising nine or 19 polymorphic, di- and tetranucleotide microsatellite loci. Individuals included in the 19 loci data set (36 populations, 16 lakes) formed a fully nested subset within the more extensive nine loci data set (69 populations, 30

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

lakes). The microsatellite loci used in this study were previously developed for the *Coregonus lavaretus* species complex[89-91] (*ClaTet1, ClaTet3, ClaTet5, ClaTet6, ClaTet9, ClaTet1, ClaTet12, ClaTet15, ClaTet18, Cocl-Lav04, Cocl-Lav06, Cocl-Lav10, Cocl-Lav18, Cocl-Lav27, Cocl-Lav52, Cocl-Lav49, BWF2, ClaTet13, C2-157*), and were amplified in four polymerase chain reaction (PCR) multiplexes in 2.5 µl reaction volume following the PCR protocol and conditions in[92]. PCR products were analysed using an ABI 3130XL Genetic Analyzer (Applied Biosystems Inc., Foster City, CA) and fragment lengths were analysed using GENEMAPPER® 4.0 software (Applied Biosystems Inc.). Deviations from linkage equilibrium (LE) and from Hardy-Weinberg equilibrium (HWE) across ecotype samples and across loci were calculated in GENEPOP 4.5.2[93] (10,000 dememorization steps, 100,000,000 Markov chain steps). *P*-values for the LE and HWE tests were corrected with the sequential Bonferroni method[94]. To reduce potential biases introduced into population genetic analyses by the presence of excessively closely related individuals (ECRs), both nine and 19 loci genotype sets were analysed in the R package RELATED[95]. Following simulations, the triadic likelihood estimator [96] was used to identify ECR individuals within each population. One of the individuals in each ECR pair was then excluded from all subsequent analyses.

Genetic differentiation between sympatric ecotypes was quantified as pairwise multilocus estimates of F_{ST}, using ARLEQUIN 3.5.1.2[97], with 1000 permutations to test significance. To investigate the geographic origins of within-lake genetic diversity in the introduced populations (38 ecotypes, 18 lakes), individual assignment analyses were run using STRUCTURE 2.3.4[98]. Parameters used: 50000 burn-in length, 500,000 MCMC chain replicates, admixture model of ancestry, correlated allele frequencies, population specific alpha prior (starting prior of 0.1). For each K, 10 independent STRUCTURE runs were carried out, up to a K of 25. For the STRUCTURE results, the true number of distinct genotypic clusters was estimated by selecting the population grouping (K) with the highest log probability of the data (ln Pr($X|K$)). The STRUCTURE results were summarized and visualized using CLUMPAK[99] and the R package POPHELPER 2.2.5[100]. To corroborate the STRUCTURE results, hierarchical relationships among introduced populations were reconstructed using unrooted neighbor-joining (NJ) trees of Cavalli-Sforza cord distances (D_{CH}), run in PHYLIP 3.695[101]. Support for the recovered tree topology was estimated using 1,000 bootstrap replicates. The resulting tree was visualized in FIGTREE v1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/). For both the STRUCTURE and the PHYLIP analyses, the nine loci genotype set was used for all included populations.

The results of population genetic analyses are summarized in Supplementary Table 1 (where available, only the results for 19 loci are reported). For the nine loci data set, significant locus-specific deviations from HWE were found in 36 out of 692 tests ($p<0.05$). Pairwise tests of linkage disequilibrium (LD) between loci were found to be significant in 79 out of 2630 tests ($p<0.05$). For the 19 loci data set, significant deviations from HWE were found in 45 out of 905 tests ($p<0.05$). LD between loci were found to be significant in 240 out of 7694 tests ($p<0.05$). For all LD and HWE analyses, no tests remained significant following Bonferroni correction.

For the STRUCTURE results, the population grouping with the highest log probability was found to be $K = 16$. At this K, patterns of individual cluster assignment within lakes fell into two broad categories (Supplementary Fig. 4): (i) introduction lakes without a clear signal of secondary introduction (Supplementary Fig. 4b), and (ii) introduction lakes showing signals of secondary contact between distinct genotypic clusters (Supplementary Fig. 4c).

NJ tree-based relationships among ecotypes within lakes were consistent with the patterns of individual genetic cluster assignment (Supplementary Fig. 4a). For the primary divergence lakes, most co-existing species pairs were strongly supported sister species with relatively short branch lengths (bootstrap support

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

100 %). Exceptions were the clades formed by Oxvattensjön/Rissjön, and Hetögeln/Murusjöen, respectively, which showed strong support (100 %) for monophyly. For Oxvattensjön/Rissjön, this reflects the introduction from the same source population [102], and for Hetögeln/Murusjöen, it likely reflects that Hetögeln's dwarf, which spawns in the connecting stream, has spread upstream to Murusjöen. For the lakes included in the secondary contact category, co-existing species generally grouped closest to allopatric populations in other lakes. Only Rosången (50.2 %) and Hökvattnet (77.6 % and 63.1 %) whitefish formed monophyletic groupings, perhaps indicative of more complex secondary contact scenarios with introgression.

The primary divergence lakes were included in the chronosequence of introduced populations used in Figs. 3 and 4. Note that the dwarf/giant ecotype pair in lake Bölessjön was included in the chronosequence even though there is a third, genetically distinct ecotype that was introduced more than hundred years later than the first introduction. The inclusion of this lake was motivated by the apparent lack of introgression between the third ecotype and the original dwarf/giant ecotype pair (Supplementary Fig. 4b, probably explained by spawning segregation in space (stream vs lake) and in the timing of spawning).

The adaptive dynamics model

We investigated the conditions for divergence in whitefish with an adaptive dynamics approach, using a physiologically structured population model (PSPM, see refs[103-106]) in which the population has a continuous size structure and individuals reproduce continuously. Our model contains two habitats – littoral and pelagic – to which whitefish have access at all times. Each habitat has one unique resource type: macroinvertebrates are found in the littoral habitat and zooplankton in the pelagic habitat. An important difference between these resources lies in the way that resource-use efficiencies for whitefish depend on whitefish size. While the feeding efficiency for zooplankton has a hump-shaped relationship to the size of the consumer, it increases almost linearly with whitefish body size for benthic invertebrates (Supplementary Fig. 6a)[57]. Hence, large whitefish generally depend on benthic invertebrates to sustain positive growth.

However, the benefits of shifting to the benthic resource also depend on size- and habitat-specific mortality rates. Both habitats have equal background mortality rates that are unrelated to size. Because pike is a mainly littoral predator[53, 107], pike predation is modelled as an extra, negatively size-dependent mortality rate that affects individuals feeding on the benthic resource (Supplementary Fig. 6b). Individuals allocate their time in each habitat in order to minimize the ratio between mortality rate and prey encounter rate (Supplementary Figs. 6c and d). The intake rate of a given foraging strategy is determined by resource type, resource density and individual size. In order to keep the model structure conservative and simple, there is no genetic or ecological variation among new recruits. Hence, the only way to become different from other individuals of the same size is by acquired changes in the evolving trait, namely maturation size. The model is deterministic and does not include a genetic mechanism. Thus it produces evolutionary divergence under the implicit assumption that assortative mating is present when the population reaches a branching point (or alternatively that reproduction is clonal). A detailed description of the model and parameter values is given in Supplementary Methods, Supplementary Figs. 6 and 7 and Supplementary Table 3-5.

Acknowledgements

We thank André de Roos, Folmer Bokma and Arne Nolte for valuable comments on an earlier version of this manuscript, and the large number of local fishers and fishery managers who reported their observations in interviews and assisted in various ways during field work. We thank Emma Andersson, Jens Andersson, Sophie Bodenstein, Alex Dew, Emilija Emma, Mats-Jerry Eriksson, Johan Fahlman, Björn Fallgren, Olof Filipsson, Erik Forsberg, Tanja L. Hanebrekke, Mikael Johansson, Sandra Kero, Johan Leander, Johan Lidman, Emil Lindberg,

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Sofi Lundbäck, Oscar Lövbom, Martina Magnusson, Sven Norman, Johan Nyberg, Fredrik Olajos, Carmen Secco Perez, Anders Pålsson, Jan Roos and Benjamin Weigel for gillnet sampling and lab work, and Gavin Horsburgh, Lotta Ström and Rachel Tucker for technical support. This research was funded by grants from FORMAS (#2007-1149) to GE, the Swedish Research Council (#2013-5110) to GE, Biodiversa (#2012-1826) to GE and CLH, and from Göran Gustavssons Stiftelse för natur och miljö i Lappland to GÖ and GE.

Author contributions

The study was conceived by GÖ, GE and SOÖ, and planned and coordinated by GÖ and GE. MB developed the eco-evolutionary model with ecological input from GÖ, GE and KAN and ecological/mathematical input from ÅB. MB performed the numerical simulations and the model analysis and wrote the model description with the assistance of KAN and ÅB. KBM, AGH, and KP performed genetic analyses. GÖ and SOÖ conducted interviews and collected information from different archives. GÖ, GE, SOÖ, and MP organized and undertook sample fishing surveys with help from AGH, PB and PJ, and MP and PJ performed analyses of fish in the lab as a part of their candidate- and/or master theses. GÖ, GE and PB analysed non-genetic data, and GÖ wrote the text. All authors read and commented on the manuscript before submission.

Competing financial interests

The authors declare no competing financial interests.

Materials & Correspondence

Correspondence and material requests should be addressed to Gunnar Öhlund.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Bolnick, D.I. & Fitzpatrick, B.M. Sympatric speciation: Models and empirical evidence. *Ann. Rev. Ecol. Evol. Syst.* **38**, 459-487 (2007).
2. Foote, A.D. Sympatric Speciation in the Genomic Era. *Trends Ecol. Evol.* **33**, 85-95 (2018).
3. Skulason, S. & Smith, T.B. Resource polymorphisms in vertebrates. *Trends Ecol. Evol.* **10**, 366-370 (1995).
4. Seehausen, O. & Wagner, C.E. Speciation in Freshwater Fishes. *Ann. Rev. Ecol. Evol. Syst.* **45**, 621-651 (2014).
5. McPhail, J.D. Ecology and evolution of sympatric sticklebacks (*Gasterostous*) - Evidence for a species pair in Paxton lakes. Texada island, Brittish Columbia. *Canadian Journal of Zoology-Revue Canadienne De Zoologie* **70**, 361-369 (1992).
6. Takahashi, T., Watanabe, K., Munehara, H., Ruber, L. & Hori, M. Evidence for divergent natural selection of a Lake Tanganyika cichlid inferred from repeated radiations in body size. *Mol. Ecol.* **18**, 3110-3119 (2009).
7. Taylor, E.B. & Bentzen, P. Evidence for multiple origins and sympatric divergence of trophic ecotypes of smelt (*Osmerus*) in Northeastern North-America. *Evolution* **47**, 813-832 (1993).
8. Sandlund, O.T. *et al.* The arctic charr, *Salvelinus alpinus*, in Thingvallavatn. *Oikos* **64**, 305-351 (1992).
9. Markevich, G., Esin, E. & Anisimova, L. Basic description and some notes on the evolution of seven sympatric morphs of Dolly Varden *Salvelinus malma* from the Lake Kronotskoe Basin. *Ecol. Evol.* **8**, 2554-2567 (2018).
10. White, R.G. Endemic whitefishes of Bear lake, Utah-Idaho: a problem in

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

systematics. All graduate theses and dissertations . 4583. https://digitalcommons.usu.edu/etd/4583 (1974).

11. Schulz, M. & Freyhof, J. *Coregonus fontanae*, a new spring-spawning cisco from Lake Stechlin, northern Germany (Salmoniformes: Coregonidae). *Ichthyological Exploration of Freshwaters* **14**, 209-216 (2019).
12. Svärdson, G. Speciation in Scandinavian *Coregonus*. *Rep. Inst. Fresw. Res, Drottningholm*, 1-95 (1979).
13. Lu, G.Q. & Bernatchez, L. Correlated trophic specialization and genetic divergence in sympatric lake whitefish ecotypes (*Coregonus clupeaformis*): Support for the ecological speciation hypothesis. *Evolution* **53**, 1491-1505 (1999).
14. Mann, G.J. & McCart, P.J. Comparison of sympatric dwarf and normal-populations of least cisco (*Coregonus sardinella*) inhabiting Trout lake, Yukon territory. *Can. J. Fish. Aquat. Sci.* **38**, 240-244 (1981).
15. Skulason, S., Noakes, D.L.G. & Snorrason, S.S. Ontogeny of trophic morphology in 4 sympatric morphs of Arctic charr, *Salvelinus alpinus*, in Thingvallavattnet, Island. *Biol. J. Linnean Soc.* **38**, 281-301 (1989).
16. Rundle, H.D., Nagel, L., Boughman, J.W. & Schluter, D. Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**, 306-308 (2000).
17. Knudsen, R., Klemetsen, A., Amundsen, P.-A. & Hermansen, B. Incipient speciation through niche expansion: an example from the Arctic charr in a subarctic lake. *P R Soc B* **273**, 2291-2298 (2006).
18. Landry, L. & Bernatchez, L. Role of epibenthic resource opportunities in the parallel evolution of lake whitefish species pairs (*Coregonus sp.*). *J. Evol. Biol.* **23**, 2602-2613 (2010).
19. Vonlanthen, P. *et al.* Divergence along a steep ecological gradient in lake whitefish (*Coregonus sp.*). *J. Evol. Biol.* **22**, 498-514 (2009).
20. Bolnick, D.I. Can intraspecific competition drive disruptive selection? An experimental test in natural populations of sticklebacks. *Evolution* **58**, 608-618 (2004).
21. Siwertsson, A. *et al.* Sympatric diversification as influenced by ecological opportunity and historical contingency in a young species lineage of whitefish. *Evol. Ecol. Res.* **12**, 929-947 (2010).
22. Kahilainen, K.K. *et al.* The role of gill raker number variability in adaptive radiation of coregonid fish. *Evol. Ecol.* **25**, 573-588 (2011).
23. Landry, L., Vincent, W.F. & Bernatchez, L. Parallel evolution of lake whitefish dwarf ecotypes in association with limnological features of their adaptive landscape. *J. Evol. Biol.* **20**, 971-984 (2007).
24. Vonlanthen, P. *et al.* Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature* **482**, 357-U1500 (2012).
25. Wagner, C.E., Harmon, L.J. & Seehausen, O. Ecological opportunity and sexual selection together predict adaptive radiation. *Nature* **487**, 366-U124 (2012).
26. Gordeeva, N.V., Alekseyev, S.S., Matveev, A.N. & Samusenok, V.P. Parallel evolutionary divergence in Arctic char *Salvelinus alpinus* complex from Transbaikalia: variation in differentiation degree and segregation of genetic diversity among sympatric forms. *Can. J. Fish. Aquat. Sci.* **72**, 96-115 (2015).
27. Winkelmann, K., Genner, M.J., Takahashi, T. & Ruber, L. Competition-driven speciation in cichlid fish. *Nat. Com.* **5** (2014).
28. Rundle, H.D., Vamosi, S.M. & Schluter, D. Experimental test of predation's effect on divergent

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

selection during character displacement in sticklebacks. *Proc. Nat. Acad. Sci. USA* **100**, 14943-14948 (2003).

29. Andersson, J., Bystrom, P., Claessen, D., Persson, L. & De Roos, A.M. Stabilization of population fluctuations due to cannibalism promotes resource polymorphism in fish. *Am. Nat.* **169**, 820-829 (2007).
30. Ohlberger, J., Staaks, G., Petzoldt, T., Mehner, T. & Holker, F. Physiological specialization by thermal adaptation drives ecological divergence in a sympatric fish species pair. *Evol. Ecol. Res.* **10**, 1173-1185 (2008).
31. Symonova, R. *et al.* Genome differentiation in a species pair of coregonine fishes: an extremely rapid speciation driven by stress-activated retrotransposons mediating extensive ribosomal DNA multiplications. *BMC Evol. Biol.* **13** (2013).
32. Elmer, K.R., Lehtonen, T.K., Kautt, A.F., Harrod, C. & Meyer, A. Rapid sympatric ecological differentiation of crater lake cichlid fishes within historic times. *Bmc Biology* **8** (2010).
33. Kautt, A.F., Machado-Schiaffino, G., Torres-Dowdall, J. & Meyer, A. Incipient sympatric speciation in Midas cichlid fish from the youngest and one of the smallest crater lakes in Nicaragua due to differential use of the benthic and limnetic habitats? *Ecol. Evol.* **6**, 5342-5357 (2016).
34. Marques, D.A. *et al.* Genomics of rapid Incipient speciation in sympatric threespine stickleback. *PLoS Genet.* **12**, 34 (2016).
35. McGee, M.D., Neches, R.Y. & Seehausen, O. Evaluating genomic divergence and parallelism in replicate ecomorphs from young and old cichlid adaptive radiations. *Mol. Ecol.* **25**, 260-268 (2016).
36. Marques, D.A. *et al.* Genomic landscape of early ecological speciation initiated by selection on nuptial colour. *Mol. Ecol.* **26**, 7-24 (2017).
37. Lamichhaney, S. *et al.* Rapid hybrid speciation in Darwin's finches. *Science* **359**, 224-227 (2018).
38. Marques, D.A., Jones, F.C., Di Palma, F., Kingsley, D.M. & Reimchen, T.E. Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation. *Nat. Ecol. Evol.* **2**, 1128-1138 (2018).
39. Moser, F.N. *et al.* The onset of ecological diversification 50 years after colonization of a crater lake by haplochromine cichlid fishes (vol 285, 20180171, 2018). *P R Soc B* **285** (2018).
40. Nosil, P., Harmon, L.J. & Seehausen, O. Ecological explanations for (incomplete) speciation. *Trends Ecol. Evol.* **24**, 145-156 (2009).
41. Vamosi, S.M. The presence of other fish species affects speciation in threespine sticklebacks. *Evol. Ecol. Res.* **5**, 717-730 (2003).
42. Recknagel, H., Elmer, K.R. & Meyer, A. Crater lake habitat predicts morphological diversity in adaptive radiations of chiclid fishes. *Evolution* **68**, 2145-2155 (2014).
43. Præbel, K. *et al.* Ecological speciation in postglacial European whitefish: rapid adaptive radiations into the littoral, pelagic, and profundal lake habitats. *Ecol. Evol.* **3**, 4970-4986 (2013).
44. Nyström, C. *Om fiskfaunan och fiskerierna i Jämtlands län*. (1862).
45. Lundberg, R. On the distribution of Swedish freshwater fishes. *Medd. Kongl. Landtbruksstyr.* **10** (1899).
46. Burman, F. in Källor till Jämtlands och Härjedalens historia utg. digitalt av Landsarkivet i Östersund och Jämtlands läns fornskriftsällskap (2010, www.riksarkivet.se/ostersund, accessed May 2016; 1797).
47. Bernatchez, L. Ecological theory of adaptive radiation: An empirical assessment from coregonine fishes

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Salmoniformes). in *Evolution Illuminated: Salmon and Their Relatives*. (eds. A. Hendry & C.S. Sterns) 175-207 (Oxford University Press, Oxford; 2005).

48. Bolnick, D.I. & Lau, O.L. Predictable patterns of disruptive selection in stickleback in postglacial lakes. *Am. Nat.* **172**, 1-11 (2008).
49. Häkli, K., Østbye, K., Kahilainen, K.K., Amundsen, P.-A. & Præbel, K. Diversifying selection drives parallel evolution of gill raker number and body size along the speciation continuum of European whitefish. *Ecol. Evol.* **8**, 2617-2631 (2018).
50. Scrucca, L., Fop, M., Murphy, T.B. & Raftery, A.E. mclust 5: Clustering, classification and density estimation using Gaussian finite mixture models. *R Journal* **8**, 289-317 (2016).
51. Gislason, D., Ferguson, M., Skulason, S. & Snorrason, S.S. Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (*Salvelinus alpinus*). *Can. J. Fish. Aquat. Sci.* **56**, 2229-2234 (1999).
52. Arbour, J.H., Hardie, D.C. & Hutchings, J.A. Morphometric and genetic analyses of two sympatric morphs of Arctic char (*Salvelinus alpinus*) in the Canadian High Arctic. *Can. J. Zool.* **89**, 19-30 (2011).
53. Vollestad, L.A., Skurdal, J. & Qvenild, T. Habitat use, growth, and feeding of pike (*Esox lucius*) in 4 Norwegian lakes. *Arch. Hydrobiol.* **108**, 107-117 (1986).
54. Mittelbach, G.G. & Persson, L. The ontogeny of piscivory and its ecological consequences. *Can. J. Fish. Aquat. Sci.* **55**, 1454-1465 (1998).
55. Brett, J.R. & Groves, T.D.D. Physiological energetics. Fish Physiology, in *Bio-energetics and Growth*, Vol. 8. (eds. W.S. Hoar, D.J. Randall & J.R. Brett) 279–352 (Academic Press, New York; 1979).
56. Shuter, B.J., Giacomini, H.C., de Kerckhove, D. & Vascotto, K. Fish life history dynamics: shifts in prey size structure evoke shifts in predator maturation traits. *Can. J. Fish. Aquat. Sci.* **73**, 693-708 (2016).
57. Byström, P. & Andersson, J. Size-dependent foraging capacities and intercohort competition in an ontogenetic omnivore (Arctic char). *Oikos* **110**, 523-536 (2005).
58. Persson, A. & Brönmark, C. Foraging capacities and effects of competitive release on ontogenetic diet shift in bream, *Abramis brama*. *Oikos* **97**, 271-281 (2002).
59. Schluter, D. Adaptive radiation in sticklebacks - Size, shape, and habitat use efficiency. *Ecology* **74**, 699-709 (1993).
60. de Roos, A.M. & Persson, L. *Population and Community Ecology of Ontogenetic Development*, Vol. 51. (Princeton University Press, 2013).
61. L'Abee-Lund, J.H., Langeland, A., Jonsson, B. & Ugedal, O. Spatial segregation by age and size in Arctic charr: A trade-off between feeding possibility and risk of predation. *J. Anim. Ecol.* **62**, 160-168 (1993).
62. Werner, E.E. & Hall, D.J. Ontogenetic habitat shifts in bluegill - The foraging rate predation rate trade-off. *Ecology* **69**, 1352-1366 (1988).
63. Werner, E.E. Size, scaling and the evolution of complex life cycles, in *Size-Structured Populations: Ecology and Evolution*. (eds. B. Ebenman & L. Persson) 60-81 (Springer Verlag, Berlin; 1988).
64. West-Eberhard, M.J. Developmental plasticity and the origin of species differences. *Proc. Nat. Acad. Sci. USA* **102**, 6543-6549 (2005).
65. Fitzpatrick, B.M. Underappreciated consequences of phenotypic plasticity for ecological speciation. *Int. J. Ecol.*, 1-12 (2012).
66. Nonaka, E., Svanbäck, R., Thibert-Plante, X., Englund, G. & Brännstrom, Å. Mechanisms by which phenotypic plasticity affects adaptive divergence

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and ecological speciation. *Am. Nat.* **186**, E126-E143 (2015).

67. Kahilainen, K., Lehtonen, H. & Kononen, K. Consequence of habitat segregation to growth rate of two sparsely rakered whitefish (*Coregonus lavaretus* (L.)) forms in a subarctic lake. *Ecol. Freshw. Fish* **12**, 275-285 (2003).
68. Foote, C.J. & Larkin, P.A. The role of male choice in the assortative mating of anadromous and non-anadromous sockeye salmon (*Oncorhyncus nerka*). *Behaviour* **106**, 43-62 (1988).
69. Conte, G.L. & Schluter, D. Experimental confirmation that body size determines mate preference via phenotype matching in a stickleback speciers pair. *Evolution* **67**, 1477-1484 (2013).
70. Lindström, T. Lifehistory of whitefish (*Coregonus*) young in two lake reservoirs. . *Institute of Freshwater Research, Drottningholm. Report* **44**, 113-144 (1952).
71. Olofsson, O. Rommängd och romsvällning hos siken. *Ny Svensk Fiskeritidskift* **24**, 277-279. (1933).
72. Svärdson, G. & Halvarsson, E. Försök med sikhybrider. *Inf. Inst. Freshw. Res. Drottningholm*, 1-11 (1968).
73. Hindar, K. & Jonsson, B. Habitat and food segregation of dwarf and normal arctic char (*Salvelinus alpinus*) from Vagnsvatnet lake, Western Norway. *Can. J. Fish. Aquat. Sci.* **39**, 1030-1045 (1982).
74. Helland, I.P., Freyhof, J., Kasprzak, P. & Mehner, T. Temperature sensitivity of vertical distributions of zooplankton and planktivorous fish in a stratified lake. *Oecologia* **151**, 322-330 (2007).
75. Siwertsson, A. *et al.* Discrete foraging niches promote ecological, phenotypic, and genetic divergence in sympatric whitefish (*Coregonus lavaretus*). *Evol. Ecol.* **27**, 547-564 (2013).
76. Svärdson, G. Postglacial dispersal and reticulate evolution of Nordic coregonids. *Nordic J. Freshw. Res.* **74**, 3-32 (1998).
77. Helland, I.P., Harrod, C., Freyhof, J. & Mehner, T. Co-existence of a pair of pelagic planktivorous coregonid fishes. *Evol. Ecol. Res.* **10**, 373-390 (2008).
78. Adams, C.E., Wilson, A.J. & Ferguson, M.M. Parallel divergence of sympatric genetic and body size forms of Arctic charr, *Salvelinus alpinus*, from two Scottish lakes. *Biol. J. Linnean Soc.* **95**, 748-757 (2008).
79. Huitfeldt-Kaas, H. *Ferskvandsfiskenes utbredelse og indvandring i Norge: med et tillæg om krebsen*. (Centraltrykkeriet, Kristiania., Kristiania; 1918).
80. Filipsson, O. Nya fiskbestånd genom inplanteringar eller spridning av fisk. *Information från Sötvattenslaboratoriet* **2**, 1-65 (1994).
81. Svärdson, G. The Coregonid problem. III. Whitefish from the Baltic, successfully introduced into fresh waters in the north of Sweden. *Rep. Inst. Fresw. Res, Drottningholm* **32**, 79-125 (1951).
82. Fabricius, E. Heterogeneous stimulus summation in the release of spawning activities in fish. *Institute of Freshwater research Dottningholm. Report* **31**, 57-99 (1950).
83. Østbye, K., Næsje, T.F., Bernatchez, L., Sandlund, O.T. & Hindar, K. Morphological divergence and origin of sympatric populations of European whitefish (*Coregonus lavaretus* L.) in Lake Femund, Norway. *J. Evol. Biol.* **18**, 683-702 (2005).
84. Svärdson, G. The Coregonid problem V. Svmpatric whitefish species of the lakes Idsjön, Storsjön and Hornavan. *Rep. Inst. Fresw. Res, Drottningholm* **34**, 141-166 (1953).
85. R development core team R: A language and environment for

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

statistical computing. *R Foundation for Statistical Computing* (2010).

86. Cohen, J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20**, 37-46 (1960).
87. Schafer, J.L. Estimation/multiple Imputation for mixed categorical and continuous data. *https://cran.r-project.org/web/packages/mix/mix.pdf* (2017).
88. Venables, W.N. & Ripley, B.D. *Modern Applied Statistics with S*. (Springer, New York; 2002).
89. Patton, J.C., Gallaway, B.J., Fechhelm, R.G. & Cronin, M.A. Genetic variation of microsatellite and mitochondrial DNA markers in broad whitefish (*Coregonus nasus*) in the Colville and Sagavanirktok rivers in northern Alaska. *Can. J. Fish. Aquat. Sci.* **54**, 1548-1556 (1997).
90. Turgeon, J., Estoup, A. & Bernatchez, L. Species flock in the North American Great Lakes: Molecular ecology of Lake Nipigon Ciscoes (Teleostei : Coregonidae : *Coregonus*). *Evolution* **53**, 1857-1871 (1999).
91. Rogers, S.M., Marchand, M.H. & Bernatchez, L. Isolation, characterization and cross-salmonid amplification of 31 microsatellite loci in the lake whitefish (*Coregonus clupeaformis*, Mitchill). *Molecular Ecology Notes* **4**, 89-92 (2004).
92. Præbel, K. *et al.* A diagnostic tool for efficient analysis of the population structure, hybridization and conservation status of European whitefish (*Coregonus lavaretus* (L.)) and vendace (*C. albula* (L.)), in *Biology and Management of Coregonid Fishes - 2011*, Vol. 64. (eds. J. Wanzenbock & I.J. Winfield) 247-255 (2013).
93. Rousset, F. GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103-106 (2008).
94. Rice, W.R. Analyzing tables of statistical tests. *Evolution* **43**, 223-225 (1989).
95. Pew, J., Muir, P.H., Wang, J. & Frasier, T.R. Related: an R package for analysing pairwise relatedness from codominant molecular markers. *Mol. Ecol. Resour.* **15**, 557-561 (2015).
96. Wang, J. Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetical Research* **89**, 135-153 (2007).
97. Excoffier, L., Laval, G. & Schneider, S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform.* **1**, 47-50 (2005).
98. Falush, D., Stephens, M. & Pritchard, J.K. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587 (2003).
99. Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A. & Mayrose, I. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* **15**, 1179-1191 (2015).
100. Francis, R.M. POPHELPER: an R package and web app to analyse and visualize population structure. *Mol. Ecol. Resour.* **17**, 27-32 (2017).
101. Felsenstein, J. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistic* **5**, 164-166 (1989).
102. Olofsson, O. Några inplanteringar av Lomsjö-sik. *Svensk Fiskeri Tidskrift* **43**, 16-18 (1934).
103. de Roos, A.M. A gentle introduction to physiologically structured population models. *Population and Community Biology Series* **18**, 119-204 (1997).

bioRxiv preprint doi: https://doi.org/10.1101/543744; this version posted February 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

104. Metz, J.A.J. & Diekmann, O. *The dynamics of physiologically structured populations*. (Springer, Berlin; 1986).
105. Claessen, D., De Roos, A.M. & Persson, L. Dwarfs and giants: cannibalism and competition in size-structured populations. *Am. Nat.* **155(2)**, 219-237 (2000).
106. Andersson, J. The development of resource polymorphism - effects of diet, predation risk and population dynamical feedbacks. *Department of Ecology and Environmental Science, Umeå University, Umeå.*, Ph. D. thesis (2005).
107. Chapman, C.A. & Mackay, W.C. Direct observation of habitat utilization by northern pike. *Copeia*, 255-258 (1984).