

# Rapid parallel adaptive radiations from a single hybridogenic ancestral population

Alan G. Hudson<sup>1,2,\*</sup>, Pascal Vonlanthen<sup>1,2</sup> and Ole Seehausen<sup>1,2</sup>

<sup>1</sup>Division of Aquatic Ecology & Macroevolution, Institute of Ecology and Evolution, University of Bern, Baltzerstrasse 6, CH-3012 Bern, Switzerland

<sup>2</sup>EAWAG Centre of Ecology, Evolution and Biogeochemistry, Department of Fish Ecology and Evolution, Seestrasse 79, CH-6047 Kastanienbaum, Switzerland

The Alpine lake whitefish (*Coregonus lavaretus*) species complex is a classic example of a recent radiation, associated with colonization of the Alpine lakes following the glacial retreat (less than 15 kyr BP). They have formed a unique array of endemic lake flocks, each with one to six described sympatric species differing in morphology, diet and reproductive ecology. Here, we present a genomic investigation of the relationships between and within the lake flocks. Comparing the signal between over 1000 AFLP loci and mitochondrial control region sequence data, we use phylogenetic tree-based and population genetic methods to reconstruct the phylogenetic history of the group and to delineate the principal centres of genetic diversity within the radiation. We find significant cytonuclear discordance showing that the genomically monophyletic Alpine whitefish clade arose from a hybrid swarm of at least two glacial refugial lineages. Within this radiation, we find seven extant genetic clusters centred on seven lake systems. Most interestingly, we find evidence of sympatric speciation within and parallel evolution of equivalent phenotypes among these lake systems. However, we also find the genetic signature of human-mediated gene flow and diversity loss within many lakes, highlighting the fragility of recent radiations.

Keywords: sympatric speciation; postglacial radiation; cytonuclear discordance; AFLPs; Coregonus; parallel evolution

# 1. INTRODUCTION

As human-driven environmental change and species loss gathers pace, understanding how species diversity evolves and how it is genetically and ecologically structured becomes critical. The process of adaptive radiation, which explicitly links environmental heterogeneity with ecological diversification and speciation, potentially provides crucial insights into these questions [1,2]. As populations exploit novel niches, divergent selection drives adaptive divergence and genetic differentiation. Ecologically similar but spatially isolated environments will exert similar selection pressures on the diverging populations, and many adaptive radiations show repeated occurrences of specific ecotypes associated with specific niches. These replicated natural experiments allow testing of hypotheses regarding the drivers of diversification. Young adaptive radiations, where intrinsic post-zygotic reproductive isolation between the constituent taxa is incomplete, provide particularly excellent model systems because the processes leading to speciation are not yet obscured [3].

Many early phylogenetic studies of young adaptive radiations recovered monophyletic gene trees (e.g. [4,5]). This was interpreted as evidence for high rates of diversification from a single ancestral lineage. Many of these systems have been reanalysed with multi-locus nuclear markers and in several cases there is marked

disagreement revealed between these and earlier cytoplasmic gene studies. While phylogenetic conflict may be due to a lack of lineage sorting, there are a number of studies where introgressive hybridization between putative founding taxa [6-8] or between taxa within an adaptive radiation [9-11] is the more parsimonious explanation. At the same time, novel phenotypes and adaptations arising within adaptive radiations, have in several cases been shown to be of hybridogenic origin [12,13]. Interspecific gene flow increases or replenishes standing genetic variation for divergent natural and sexual selection to work with. Also introgression and recombination could overcome constraints caused by genetic covariances between traits, opening the way for new evolutionary trajectories. Thus, the interaction of gene flow and selection could fuel the functional breadth, tempo and duration of adaptive radiations [9,14]. Phylogenetic trees have also been used extensively in testing the hypothesis of sympatric speciation [3,15]. However, the many cases of discordant phylogenetic signals between different genes or genomic compartments suggest that biogeographically complex models of speciation maybe more likely [14,16]. Frequently, an initial period of allopatric isolation may be followed by secondary contact with introgression. The outcome may vary from reinforcement of pre-existing reproductive isolation, to the formation of hybrid swarms, to hybrid speciation. The extent of gene flow on secondary contact will determine genomic similarity among the resultant species, and therefore the magnitude of phylogenetic discord.

The focus of this study, the Alpine whitefish (*Coregonus* sp.) radiation, is native to approximately 40 lakes in three

<sup>\*</sup> Author for correspondence (alan.g.hudson@gmail.com).

Electronic supplementary material is available at http://dx.doi.org/10. 1098/rspb.2010.0925 or via http://rspb.royalsocietypublishing.org.

European drainages: the Rhine, Rhône and Danube [17]. Like other north temperate fish species complexes, colonization of these lakes occurred following the glacial retreat (12-20 kyr BP). Within the Alpine whitefish radiation intralacustrine diversity varies from one to six species. Coexisting species are genetically differentiated [18], and phenotypically differentiated in traits linked to feeding ecology: gill-raker number and density, growth rate, shape and colour and also in traits linked to reproductive ecology: time of spawning, location/depth of spawning, spawning substrate and egg size, with the replicated occurrences of corresponding ecotypes in different lakes [19,20]. Previous phylogeographic studies of the Coregonus lavaretus species complex have shown two very divergent mitochondrial lineages present among the Alpine whitefish [21,22]. However, sampling within the Alpine radiation was taxonomically and geographically limited. It was hence not known how these lineages were structured between different species, recurrent ecotypes or lakes. A statistical association between ecotype and cytoplasmic lineage would be evidence for an allopatric phase in the origin of ecotypes. Microsatellite genetic analysis of the Alpine radiation suggested that species within single large or multiple, closely connected lakes formed monophyletic species flocks, more consistent with multiple origins for each ecotype and sympatric speciation or allopatric-sympatric speciation with high gene flow on secondary contact [18,23]. However, the number of loci was low and the monophyly of the entire radiation has never been tested [18,24]. Therefore, in one of the most species-diverse, young, vertebrate adaptive radiations known, surprisingly little is known about the geographical modes of origin.

We present the first integrated mitochondrial and nuclear genetic study of the Alpine lake whitefish radiation. We use mitochondrial DNA (mtDNA) to identify the historical origins of the colonizing lineages. We then compare the mtDNA phylogenetic signal with genomic AFLP data, combining phylogenetic with population genetic analyses to test two hypotheses: first, the genome-wide monophyly of the Alpine radiation to ascertain whether speciation has taken place within the modern geographical distribution range of the radiation. Second, we characterize the principal structure of genetic diversity to test whether lake flocks have evolved primarily in sympatry, with the parallel evolution of ecotypes between lake flocks.

# 2. MATERIAL AND METHODS

#### (a) Sampling

The species, their geographical occurrence, their ecotype/ chronotype designation and the numbers of individuals genotyped, can be found in electronic supplementary material, table S1. Representative images of species/ecotypic diversity can be found in electronic supplementary material, figure S1. Populations were split into native and non-native (found in lakes where original stocks were extinct or had not previously occurred) according to Steinmann [20] and Kottelat & Freyhof [17]. For ecotype/chronotype designation, species were grouped by mean gill-raker number (feeding ecotype) and spawning season (reproductive chronotype). Feeding ecotype categories were: (i) less than 25 gill-rakers on the left, outer gill arch = very low gill-raker number (VLGR), (ii) 25–30 = low gill-raker (LGR), (iii) 31–35 = medium gill-raker (MGR), (iv) greater than 35 = high gill-raker (HGR). Reproductive chronotype categories were: (i) spawning November–January (Winter) and (ii) spawning July–September (Summer). Where possible all specimens were collected on their respective spawning grounds using overnight gill-netting. Additionally, several species and populations of North European members of the *C. lavaretus* complex, North American (NA) *Coregonus clupeaformis* and European *Coregonus albula* were used in a nested outgroup design.

#### (b) Molecular methods

One thousand, one hundred and seventy-eight basepairs of mitochondrial control region was sequenced for 479 individuals (forward primer L15 926; [25], reverse primer 12S; [26]). Indels longer than one nucleotide were removed, as the mutations could not be modelled [27]. AFLPs were analysed for 274 individuals, nested within the mtDNA dataset. Fourteen different *EcoRI/MseI* primer pairs with three selective bases (ACG-CAC, ACT-CAC, ACC-CAG, ACG-CAG, ACT-CAG, ACT-CAG, ACT-CAG, ATG-CTC, ATG-CTG, AAG-CTT, ACG-CTT) were amplified using a protocol modified from Vos *et al.* [28].

#### (c) Phylogenetic analysis

Unique mtDNA haplotypes were used for phylogenetic reconstruction. Trees were created using different optimization criteria: maximum parsimony (MP) in PAUP\* 4.0b10 [29], Bayesian inference in MRBAYES 3.1 [30] and maximum likelihood (ML) in PHYML [31]. Nucleotide substitution models were obtained using MRMODELTEST 2.3 [32]. MP reconstructions treated single nucleotide gaps as a fifth character and were bootstrap re-sampled 100 times. Both individual-based and population-based neighbourjoining (NJ) trees were reconstructed from AFLPs using the program NEIGHBOR in the software package PHYLIP 3.67 [33]. Individual pairwise genetic distances were calculated from 1050 polymorphic AFLP loci using RESTDIST, in PHYLIP with 1000 bootstrap replicates. Population trees were constructed from Nei's genetic distances between indigenous Alpine populations, calculated in AFLP-SURV 1.0 [34], based on 561 polymorphic AFLP loci data. To investigate their biogeographical origins, this was repeated with the addition of non-native Alpine populations (591 polymorphic loci). The significance of differences between AFLP and mtDNA tree topologies was evaluated using the SH test [35], as implemented in PAUP\*. Only mtDNA haplotypes found in individuals in the AFLP dataset were included. Two ML analyses of mtDNA sequence data were done, one unconstrained, the other one constrained at the node representing monophyly of the Alpine radiation clade.

#### (d) Population genetic analysis

AMOVAs were carried out independently on the mtDNA and AFLP data in ARLEQUIN 3.11 [36]. For both markers, four hierarchical, nested AMOVA models were specified with variation partitioned among and within (i) feeding ecotypes (VLGR, LGR, MGR, HGR), (ii) spawning chronotypes (Winter, Summer), (iii) groups of hydrologically closely connected lakes (hereafter referred to as superlakes) and (iv) individual lakes. The super-lake and single lake AMOVAs were calculated both with and without nonnative populations. The ecotype/chronotype AMOVAs were only carried out with natural populations. Demographic coalescent analyses (mismatch analyses) and neutrality tests





Figure 1. Mitochondrial D-loop phylogeographic results for the Alpine whitefish radiation. (*a*) Bayesian tree of D-loop haplotypes. Clades C and N are coloured red and blue, respectively (following [22]). Numbers across branches refer to the statistical support for nodes from different optimality criteria; (from top to bottom) Bayesian posterior probabilities, PHYML bootstrap and MP bootstrap. (*b*) Geographical distribution of C and N clade D-loop haplotypes. Circles represent all the individuals sampled from populations. Red and Blue colours represent the proportion (%) of individuals belonging to C and N clades, respectively.

were performed on the mtDNA data in ARLEQUIN. As divergence occurred within the nonlinearity period of molecular clock rates (less than 1-2 myr), standard Teleost and Salmonid mtDNA mutation rates could not be applied [37]. To overcome this, mtDNA sequence divergence was compared with the similarly aged Lake Victoria Region cichlid superflock [38]. Using as a calibration point, the Lake Victoria region superflock origin, and comparing the average number of mutational steps between its descendent haplotype lineages with that between the European whitefish haplotype lineages, obtained by mismatch analysis. The divergence between the main European whitefish haplotype lineages could then be tentatively dated. To investigate the number of genetic clusters within the Alpine radiation as well as gene flow between lakes, individual-based Bayesian population assignments were carried out on the AFLP data in STRUCTURE 2.2 [39]. This was first done across the radiation, and then repeated within each identified genetic cluster separately, to identify additional finer-scale genetic structure.

### 3. RESULTS

# (a) Mitochondrial phylogeography: polyphyly and rapid expansion

From the 479 individuals sequenced, 63 different mtDNA haplotypes were recovered (GenBank accession

Proc. R. Soc. B (2011)

numbers: HM634591-HM635074). In all the resulting phylogenetic trees NA and European whitefish formed a monophyletic group, consisting of three major clades (figure 1a). The European whitefish (C. lavaretus complex) contained two very divergent haplotype lineages, and made a paraphyletic group with regard to the NA whitefish (C. clupeaformis). The two major European clades (from here on N for Northern, and C for Central European) overlapped in their geographical distribution and co-occurred within the same populations from Switzerland to the Baltic Sea (figure 1b). N clade haplotypes (blue in figure 1) predominate in the more northerly and eastern populations in Scandinavia and the Baltic Sea region, whereas C clade haplotypes (red) had a higher frequency in the Alpine and North Sea populations.

Phylogenetic relationships between the *C. lavaretus* and *C. clupeaformis* clades varied slightly with the optimality criterion used; model-based ML and Bayesian trees showed relatively weak support (31% bootstrap support, 55% posterior probability, respectively) for a sister group relationship between the European C clade and the NA *C. clupeaformis* clade. The MP consensus tree suggested a polytomy between the three major clades. Little internal structure in the relationships between the constituent haplotypes of either of the two European

(figure 2b). Five major clades were found, centred on

Figure 2. (*Opposite.*) AFLP phylogenies of the Alpine whitefish radiation. (*a*) Individual-based circular NJ tree. Numbers represent bootstrap support (1000 replicates). (*b*) Population-based NJ tree of only Alpine radiation populations. Numbers represent bootstrap support (1000 replicates). Numbers in parentheses show support for nodes recovered using indigenous populations only. Numbers at branch tips refer to population identity: 1. Annecy, 2. Bourget, 3. Geneva, Neuchâtel (4. *C. palaea*, 5. *C. candidus*), Biel (6. *C. palaea*, 7. *C. confusus*), 8. Joux, Thun (9. *C.* sp. 'balchen', 10. *Coregonus alpinus*, 11. *C. albellus*, 12. *C.* sp. 'felchen', 13. *C. fatioi*), Brienz (14. *C.* sp. 'balchen', 15. *C.* sp. 'felchen', 16. *C. albellus*), 17. Sempach, 18. Halwil, Lucerne (19. *C.* sp. 'bodenbalchen', 20. *C. nobilis*, 21. *C. zugensis*, 22. *C. zugensis*?), 23. Alpnach, 24. Sarnern, 25. Lungern, 26. Zug, 27. Aegeri, Zuerich (28. *C. duplex*, 29. *C. heglingus*), Walen (30. *C. duplex* 'river spawner', 31. *C. duplex*, 32. *C. heglingus*, 33. *C. heglingus*?), 34. Greiffen, 35. Pfaffiker, Constance (36. *Coregonus arenicolus*, 37. *C.* sp. 'Alpenrhein', 38. *C.* sp. 'weissfelchen', 39. *Coregonus macrophthalmus*, 40. *Coregonus wartmanni*), 41. Ammer, 42. Woerth, 43. Chiem, 44. Kochel, 45. Starnberg, Maggiore (46. *C.* sp. 'lavarello', 47. *C.* sp. 'bondella'), 48. Lugano. For both figures polygons represent ecotype of sampled populations. Solid polygons, winter spawners; open, summer spawners. Circle, VLGR; triangle, LGR; diamond, MGR; square, HGR; cross, *C. nobilis* (HGR, Summer). Colours refer to the lake/super-lake sampled: Neuchâtel-Biel (pink), Thun-Brienz (orange), Lucerne (blue), Walen-Zuerich (green), Constance (red), Rhône drainage (light blue), Danube drainage (yellow), Po drainage (purple) and Rhine lakes of uncertain affiliation (black).

clades, suggests rapid, recent haplotypic diversification. Also mismatch analyses and neutrality tests suggested recent demographic expansion of each clade from small populations (electronic supplementary material, table S2). European mtDNA clades N and C were eight mutations apart. Mean divergence between mtDNA haplotypes of the cichlid fish superflock of the Lake Victoria/Edward region occurred 4.2 mutations ago, estimated to correspond to between 57 and 489 kyr BP (mean 189 or 273 kyr depending on the calibration model; [38]). Extrapolating from this, estimated divergence dates between the two European *C. lavaretus* clades were between 108 and 929 kyr BP.

AMOVAs of mtDNA data from native Alpine populations, revealed significant structuring of haplotypes by both lake (17.15%, p < 0.001) and super-lake (17.19%, p < 0.001), whereas species did not explain any significant variation either in lake (1.58%, p = 0.10) or super-lake (1.93%, p = 0.056) models. When nonnative populations were included, there was no significant structuring of haplotypic variation either by lake (10.45%, p = 0.14) or super-lake (12.85%, p = 0.06). However, there then was significant structuring appecies within lakes and super-lakes (15.33%, p < 0.001 and 13.2%, p < 0.001, respectively). Finally, levels of genetic variation structured among ecotypes or chronotypes, were not significant (0%, p = 0.32, 5.66%, p = 0.06, respectively).

# (b) Phylogenomic analysis: monophyly and speciation patterns

A total of 1050 variable AFLP loci were amplified for 274 individuals. Averaging across 30 repeated individuals (11% of total), per locus levels of reproducibility, per primer pair were high (greater than 95%). The resulting individual-based NJ tree (figure 2*a*) strongly supported genomic monophyly for the Alpine radiation (100% boot-strap support). The closest relatives were the North German *C. lavaretus* from Lake Drewitz (46%). The most divergent group among the whitefish was *C. clupea-formis* (99%). The SH test showed significant discordance (p = 0.001) between the topologies derived from the mitochondrial sequences and from the nuclear genomes.

Within the Alpine radiation, AFLP genotypes clustered into lake flocks and species in the individual-based tree (figure 2a), however, re-sampling support for these clades was low (less than 50%). Population-based AFLP trees revealed similar, but better supported topologies geographically well-separated lakes or super-lakes, with replicated occurrences of ecotypes between these. Within super-lakes, populations of corresponding ecotype from adjacent lakes are often each other's closest relatives. The five major clades recovered were the West Swiss, Lucerne, Walen-Zuerich, Thun-Brienz and Constance-Danube clades. In the West Swiss clade, phenotypically similar LGR ecotype (Coregonus palaea) populations from connected Biel and Neuchâtel were monophyletic and grouped with Bourget whitefish from the Rhône drainage, to the exclusion of sympatric MGR ecotypes (Coregonus confusus and Coregonus candidus). This West Swiss clade received a bootstrap support of 80 per cent. Within the Lucerne clade, the sympatric species flock of Lake Lucerne was paraphyletic owing to clustering of allopatric whitefish from Lake Sarnen (ca 5 km upstream) with the two HGR Coregonus zugensis chronotypes. Basal support for the Lucerne clade was 68 per cent. The other species included in this clade are the sympatric Coregonus nobilis and C. sp. 'bodenbalchen', and parapatric Lake Alpnach C. sp. 'alpnacherfelchen'. The closely connected lakes Walen and Zuerich formed a well-supported monophyletic clade (80%). The allopatric populations of LGR ecotype (Coregonus duplex) were monophyletic, as were the allopatric populations of HGR ecotype (Coregonus heglingus). The majority of endemic species of lakes Thun and Brienz formed the fourth major monophyletic clade (39%). Again allopatric populations of the same ecotype from sister lakes (i.e. LGR C. sp. 'balchen' and HGR Coregonus albellus) grouped together rather than with sympatric populations of different ecotypes. However, the MGR Coregonus fatioi from Lake Thun grouped basally to the clade formed by species from Lake Constance and Danubian lakes (43% bootstrap support). Populations from Lake Constance and Danubian lakes together made the final major clade (86%). Most populations from the Danubian lakes grouped together as a single divergent lineage within the Constance flock. Including non-native taxa often decreased basal support for the major clades: Constance-Danube (63% instead of 86%), Lucerne (47% instead of 68%) and West Swiss (69% instead of 80%), suggesting some non-native taxa are of hybrid origin.

### (c) Population genetic analysis

In the two geographical AMOVA models, slightly higher levels of genetic variation were explained by





Figure 2. (Caption opposite.)



Figure 3. Individual AFLP assignment results for the Alpine whitefish radiation. (a) STRUCTURE results across native Alpine whitefish populations. Bar plots represent individuals sampled within each lake. Coloured proportions in each column represent probabilities of membership to a specific genetic cluster. Pink, West Swiss; orange, Thun-Brienz; green, Walen-Zuerich; light blue, Lucerne HGR; dark blue, Lucerne LGR; purple, Zug; red, Constance; yellow, Danubian lakes. Vertical black lines within structure profiles divide sympatric species described from left to right. Numbers indicate lake identity: 1. Bourget, 2. Geneva, 3. Neuchâtel (*C. palaea, C. candidus*), 4. Biel (*C. palaea, C. confusus*), 5. Thun (*C. fatioi, C. sp. 'balchen', C. albellus, C. alpinus, C. sp. 'felchen', 6. Brienz (C. sp. 'balchen', C. sp. 'felchen', C. albellus*), 7. Halwil, 8. Zuerich (*C. heglingus, C. duplex, C. duplex, C. duplex 'river spawner'*), 10. Zug, 11. Sempach, 12. Sarnern, 13. Alpnach, 14. Lucerne (*C. zugensis, C. nobilis, C. sp. 'bodenbalchen'*) 15. Constance (*C. macrophthalmus, C. wartmanni, C. sp. 'Alpenrhein', C. sp. 'weissfelchen', C. arenicolus*), 16. Ammer (*Coregonus bavaricus, Coregonus renke*), 17. Woerth, 18. Kochel, 19. Chiem. (b) STRUCTURE 2.2 results using only specific lake flock individuals. Photos represent phenotypes associated with each sampled population. From left to right (i) Walen-Zuerich: Zuerich (*C. duplex, C. heglingus*), Walen (*C. heglingus, C. duplex*). (ii) Lucerne: (*C. nobilis, C. zugensis, C. sp. 'bodenbalchen'*). (iii) Thun-Brienz: Thun (*C. alpinus, C. albellus, C. sp. 'bodenbalchen'*). (iii) Thun-Brienz: Thun (*C. alpinus, C. albellus, C. sp. 'bodenbalchen'*). (iii) Thun-Brienz. Thun (*C. alpinus, C. albellus, C. sp. 'bodenbalchen'*). (iii) Lucerne: (*C. nobilis, C. zugensis, C. sp. 'bodenbalchen'*). (iii) Thun-Brienz: Thun (*C. alpinus, C. albellus, C. sp. 'bodenbalchen'*). (iii) Lucerne: (*C. nobilis, C. zugensis, C. sp. 'bodenbalchen'*). (iii) Thun-Brienz: Thun (*C. alpinus, C. albe* 

super-lake (9.30%, p < 0.001), than by single lake (8.7%, p < 0.001). Lower, but significant amounts of variation resided among species within super-lakes (7.74%, p < 0.001) and lakes (7.96%, p < 0.001), respectively. With non-native populations included, of genetic variation the amount explained by super-lake and lake was lower (7.6%, 7.1%; p <0.001). Concomitantly species explained more (within super-lakes 8.5%, p < 0.001; within lakes 7.96%, p < 0.001). AMOVA models with reproductive chronotype and feeding ecotype as grouping variables, independent of lake or super-lake, showed low but significant amounts of variation explained by chronotype (3.12%, p < 0.001) and ecotype (1.05%, p < 0.05). This was lower than the amount of variation explained species within chronotype/ecotype (14.78%, hv 15.08%; p < 0.001).

Bayesian individual assignment results were congruent with the clades revealed in the population trees (figure 3a). Using all Alpine radiation individuals, a K of 8 had the highest likelihood. West Swiss lakes, Thun-Brienz and Walen-Zuerich were identified as cohesive, major genetic clusters. The Lucerne clade was split into two clusters: one comprising LGR C. sp. 'bodenbalchen', the other HGR C. zugensis and C. nobilis, together with allopatric HGR Sarnern. Lake Zug whitefish formed a distinct genetic cluster. All Constance individuals belonged to a single genetic cluster, also many individuals from Danubian lakes showed high probabilities of genetic input from this cluster, as did Thun Coregonus fatio. However, within certain Danubian populations (Woerth and Ammer) another, distinct cluster predominated. Additional STRUCTURE analyses within lake or super-lake clusters, revealed several

distinct genetic units within Thun-Brienz (five) and Walen-Zuerich (two) (figure 3b). Again, within both systems, similar ecotypes from sister lakes tended to group in the same genetic cluster.

# 4. DISCUSSION

We find in the Alpine lake whitefish radiation marked cytonuclear genealogical discordance. mtDNA sequences reveal the widespread prevalence of two highly divergent clades within most lakes, neither of which is monophyletic in the Alpine lakes. However, AFLP data strongly support the genomic monophyly of the entire Alpine lake whitefish radiation. This suggests that this species-rich adaptive radiation arose in situ from an introgressed ancestral population. AFLP nuclear data reveal seven distinct genetic clusters centred on geographically distinct lakes and super-lake systems. These clusters again constitute smaller scale genomically monophyletic species flocks, nested within the larger radiation. This strongly suggests the parallel origins of ecologically similar species across lake systems and the prevalence of sympatric speciation. This signal of monophyletic lake and super-lake flocks persists despite genetic evidence for historical stocking between lakes and introgressive hybridization.

# (a) Cytonuclear discordance, ancient hybridization and adaptive radiation

The co-occurrence of two divergent mtDNA lineages in the Alpine lake whitefish radiation is consistent with previous phylogeographic studies [21,22,40]. These studies indicate the existence of two geographically distinct European glacial refugia for whitefish, probably located west of the Ural Mountains in Russia (N clade) and, more tentatively, near the mouth of the Rhine (C clade). Both lineages also show strong signals of recent demographic expansion towards the end of the last ice age. Within native Alpine whitefish populations, the lack of structuring of haplotypic variation at the species level or among feeding and reproductive ecotypes suggests against the pre-colonization origins of individual species and entire ecotype categories now found in the Alpine radiation. Thus it is probably that these two mitochondrial lineages were thoroughly admixed within the whitefish stocks colonizing the Alpine region, with some structuring of haplotypic variation among lakes and super-lakes through founder effects, mutation and genetic drift post-colonization. The major alternative hypothesis is the retention of ancestral mtDNA polymorphisms. The lack of resolution in the branching order between the largely sympatric lineages C and N and the NA C. clupeaformis in MP trees, and indeed the placement of the C clade as sister group to the NA clade in Bayesian and ML trees, suggest a rapid divergence of all three lineages from one another. The complete lineage sorting between the American and either European clade then makes retention of ancestral polymorphism a very unlikely explanation for the total lack of lineage sorting between the two European clades.

The AFLP signal of genomic monophyly for the entire Alpine radiation indicates that diversification took place *in situ* from a common ancestral stock. The ancient genetic signature of introgression would be lost in the nuclear data owing to recombination. This suggests that the rich adaptive diversity of the Alpine lake whitefish radiation has sprung from a genetically diverse, hybridogenic source population in the 15 kyr since the end of the last glaciation, an astonishingly short evolutionary time. Similar coincidence of cytonuclear discordance and rapid diversification have been obtained from other adaptive radiations [6,7]. The prevalence of introgression, the probable resulting increase in standing genetic variation and the potential for adaptive novelty [13,16,41], has therefore seen the recent reappraisal of introgression as a key mechanism in adaptive radiation [9,14].

#### (b) Geographical modes of speciation

The consistent sister relationships between sympatric, ecologically divergent species support, at least, a scenario of coarse-grain [42] or mosaic [43] sympatric speciation. The two major alternative explanations of the sympatric co-occurrence of sister taxa are intralacustrine allopatric speciation and regional-scale (between sister lake) allopatric speciation with high gene flow on secondary contact. The first hypothesis is very difficult to disprove [42]. However, within lakes, Alpine whitefish populations of the same species show no significant isolation by distance between spawning aggregates, making within-lake spatial population subdivision unlikely to be important in speciation ([44], B. Lundsgaard-Hansen & P. Vonlanthen, 2009 unpublished data). Regional-scale allopatric speciation would predict that individual species diverge genetically and ecologically in geographically isolated water systems, followed by reciprocal invasion and secondary contact. Despite considerable gene flow at neutral loci between such newly sympatric species, some heritable ecological divergence could be maintained by natural selection on divergent phenotypes, preserving genomic islands of differentiation between now sympatric species [45]. This scenario is difficult to distinguish from sympatric speciation from a hybrid swarm. Both models predict the shared identity of alleles brought to or maintained in dominance by selection in similar ecotypes belonging to different lake flocks. Given the recent common origin of the Alpine radiation, parallel trends in genetic divergence are perhaps probably as similar pools of standing genetic variation are acted on by natural selection in similar lacustrine environments [46]. Therefore, instead of allopatric speciation, this could indicate the same genomic underpinning of traits leading to reproductive isolation and therefore parallel speciation among ecotypes [47]. Indeed AMOVAs indicate low but significant amounts of genetic variation are explained by ecotype and chronotype and further research is being currently carried out to identify and quantify genomic regions under selection. The most conclusive way to rule out regional-scale allopatric speciation followed by gene flow on secondary contact is by showing insufficient numbers of potential refugia to produce the current sympatric species diversity, making the sympatric origin of species diversity most parsimonious. At least within the Lucerne and the Thun-Brienz lake flocks, the high number of coexisting species (at least three and five, respectively) make between lake allopatric speciation unlikely. Therefore while sympatric speciation is considered rare in nature

[15], we identify at least two well-supported putative examples within the Alpine whitefish radiation.

Interestingly in different lakes within super-lake systems, ecomorphologically similar populations from sister lakes are often more closely related to each other than ecologically different species from the same lake. As some of these sister lakes separated relatively early after the initial formation of the super-lakes, this suggests speciation was very rapid, having progressed, by the time the lakes separated, to a stage that the genomic identity of the species was maintained.

# (c) Anthropogenic impacts on diversity

Given the evolutionary youth of the Alpine whitefish radiation, a history of fish stocking and recent major habitat modification through organic pollution, the probability of anthropogenically increased levels of gene flow between and within species flocks is high. We find evidence for partial introgression from non-native lineages in certain lakes (e.g. Thun C. fatioi, most Danubian Lakes). We also find evidence for a loss of species genetic distinctiveness in those lakes that experienced the most severe anthropogenic eutrophication in the recent past (Constance, Neuchâtel, Biel). These lakes lack genetically well-differentiated species, in strong contrast with what historical data imply [20,48]. Overall this suggests that like in other young adaptive radiations [49,50], while speciation has been rapid, the susceptibility of mechanisms structuring species diversity to human-mediated environmental change means that extinction through speciation reversal is much faster [51].

## 5. CONCLUSIONS

Cytonuclear discordance between genetic markers reveals that from a genetically enriched hybridogenic ancestral population, a genomically monophyletic Alpine whitefish radiation arose in less than 15 kyr. Nested within it, are at least five distinct adaptive radiations, each centred on a large lake or system of closely connected lakes, with parallel origination of ecologically similar species. The observation that sympatric species are, with few exceptions, each other's closest relatives suggests that speciation in geographical sympatry is plausible and occurred multiple times in parallel in a short period of time. The combination of phylogenetic and population genetic methods revealed the genetic signature of anthropogenic stocking of whitefish between lakes and suggested the loss of genetic differentiation among sympatric species, highlighting the vulnerability of recent adaptive radiations to speciation reversal.

Many thanks to R. Mueller, D. Bittner, J. Freyhof, M. Kugler, J. Muggli, B. Polli, P. Volta, A. Champigneulle, K. Kahilainen, K. Pohlmann and the Bernatchez Lab for assisting in obtaining samples. Thanks to T. Dammrich for lab assistance, and to K. Lucek, S. Mwaiko, V. Schneider and D. Joyce for technical advice. Suggestions from U. Schliewen, D. Neumann, M. Barluenga, P. Woods, D. Roy, I. Magalhaes and two anonymous reviewers greatly improved this manuscript. This study was funded by the University of Bern and EAWAG Handlungsfeld project Aquadiverse.

### REFERENCES

- 1 Gavrilets, S. & Losos, J. B. 2009 Adaptive radiation: contrasting theory with data. *Science* **323**, 732–737. (doi:10. 1126/science.1157966)
- 2 Schluter, D. 2000 The ecology of adaptive radiation. Oxford series in ecology and evolution. Oxford, UK: Oxford University Press.
- 3 Coyne, J. A. & Orr, H. A. 2004 Speciation Sunderland, MA: Sinauer Associates, Inc.
- 4 Meyer, A., Kocher, T. D., Basasibwaki, P. & Wilson, A. C. 1990 Monophyletic origin of lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. *Nature* 347, 550–553. (doi:10.1038/347550a0)
- 5 Baldwin, B. G., Kyhos, D. W., Dvorak, J. & Carr, G. D. 1991 Chloroplast DNA evidence for a North-American origin of the Hawaiian silversword alliance (Asteraceae). *Proc. Natl Acad. Sci. USA* 88, 1840–1843. (doi:10. 1073/pnas.88.5.1840)
- 6 Barrier, M., Baldwin, B. G., Robichaux, R. H. & Purugganan, M. D. 1999 Interspecific hybrid ancestry of a plant adaptive radiation: allopolyploidy of the Hawaiian silversword alliance (Asteraceae) inferred from floral homeotic gene duplications. *Mol. Biol. Evol.* 16, 1105–1113.
- 7 Seehausen, O., Koetsier, E., Schneider, M. V., Chapman, L. J., Chapman, C. A., Knight, M. E., Turner, G. F., van Alphen, J. J. M. & Bills, R. 2003 Nuclear markers reveal unexpected genetic variation and a Congolese–Nilotic origin of the Lake Victoria cichlid species flock. *Proc. R. Soc. Lond. B.* 270, 2637–2638. (doi:10.1098/ rspb.2003.2001)
- 8 Shaw, K. L. 2002 Conflict between nuclear and mitochondrial DNA phylogenies of a recent species radiation: what mtDNA reveals and conceals about modes of speciation in Hawaiian crickets. *Proc. Natl Acad. Sci. USA* **99**, 16122–16127. (doi:10.1073/pnas. 242585899)
- 9 Grant, B. R. & Grant, P. R. 2008 Fission and fusion of Darwin's finches populations. *Phil. Trans. R. Soc. B.* 363, 2821-2829. (doi:10.1098/rstb.2008.0051)
- 10 Mavarez, J., Salazar, C. A., Bermingham, E., Salcedo, C., Jiggins, C. D. & Linares, M. 2006 Speciation by hybridization in *Heliconius* butterflies. *Nature* 441, 868– 871. (doi:10.1038/nature04738)
- 11 Salzburger, W., Baric, S. & Sturmbauer, C. 2002 Speciation via introgressive hybridization in East African cichlids? *Mol. Ecol.* **11**, 619–625. (doi:10.1046/ j.0962-1083.2001.01438.x)
- 12 Rieseberg, L. H., Archer, M. A. & Wayne, R. K. 1999 Transgressive segregation, adaptation and speciation. *Heredity* 83, 363–372. (doi:10.1038/sj.hdy.6886170)
- 13 Schliewen, U. K. & Klee, B. 2004 Reticulate sympatric speciation in Cameroonian crater lake cichlids. *Front. Zool.* 1, 5. (doi:10.1186/1742-9994-1-5)
- 14 Seehausen, O. 2004 Hybridization and adaptive radiation. *Trends Ecol. Evol.* **19**, 198–208. (doi:10.1016/j. tree.2004.01.003)
- 15 Bolnick, D. I. & Fitzpatrick, B. M. 2007 Sympatric speciation: models and empirical evidence. *Annu. Rev. Ecol. Evol. S.* 38, 459–487.
- 16 Feder, J. L., Xie, X. F., Rull, J., Velez, S., Forbes, A., Leung, B., Dambroski, H., Filchak, K. E. & Aluja, M. 2005 Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in *Rhagoletis. Proc. Natl Acad. Sci. USA* **102**, 6573–6580. (doi:10.1073/ pnas.0502099102)
- 17 Kottelat, M. & Freyhof, J. 2007 Handbook of European freshwater fishes. Switzerland: Publications Kottelat, Cornol.

- 66 A. G. Hudson et al. Hybridogenic radiation in whitefish
- 18 Douglas, M. R., Brunner, P. C. & Bernatchez, L. 1999 Do assemblages of *Coregonus* (Teleostei: Salmoniformes) in the Central Alpine region of Europe represent species flocks? *Mol. Ecol.* 8, 589–603. (doi:10.1046/j.1365-294x. 1999.00581.x)
- 19 Douglas, M. R., Brunner, P. C. & Douglas, M. E. 2005 Evolutionary homoplasy among species flocks of Central Alpine *Coregonus* (Teleostei: Salmoniformes). *Copeia* 2005, 347–358. (doi:10.1643/CG-04-128R)
- 20 Steinmann, P. 1950 Monographie der Schweizerischen Koregonen. Beitrag zum Problem der Entstehung neuer Arten. Spezieller Teil. Schweiz. Z. Hydrol. 12, 340–491. (doi:10.1007/BF02489885)
- 21 Bernatchez, L. & Dodson, J. J. 1994 Phylogenetic relationships among Palearctic and Nearctic whitefish (*Coregonus* sp.) populations as revealed by mitochondrial DNA variation. *Can. J. Fish. Aquat. Sci.* 51, (doi:10. 1139/f94-310)
- 22 Ostbye, K., Bernatchez, L., Naesje, T. F., Himberg, K. J. M. & Hindar, K. 2005 Evolutionary history of the European whitefish *Coregonus lavaretus* (L.) species complex as inferred from mtDNA phylogeography and gill-raker numbers. *Mol. Ecol.* 14, 4371–4387.
- 23 Hudson, A. G., Vonlanthen, P., Mueller, R. & Seehausen, O. 2007 Review: the geography of speciation and adaptive radiation in coregonines. In *Biology* and management of coregonid fishes—2005, vol. 60 (eds M. Jankun, P. Brzuzan, P. Hliawa & M. Luczynski), pp. 111–146. Stuttgart, Germany: E. Schweizerbart'sche Verlagsbuchhandlung.
- 24 Douglas, M. R. & Brunner, P. C. 2002 Biodiversity of Central Alpine *Coregonus* (Salmoniformes): impact of one-hundred years of management. *Ecol. Appl.* **12**, 154– 172. (doi:10.1890/1051-0761(2002)012[0154:BOCACS] 2.0.CO;2)
- 25 Shaw, P. W., Turner, G. F., Idid, M. R., Robinson, R. L. & Carvalho, G. R. 2000 Genetic population structure indicates sympatric speciation of Lake Malawi pelagic cichlids. *Proc. R. Soc. Lond. B.* 267, 2273–2280. (doi:10.1098/rspb.2000.1279)
- 26 Joyce, D. A., Lunt, D. H., Bills, R., Turner, G. F., Katongo, C., Duftner, N., Sturmbauer, C. & Seehausen, O. 2005 An extant cichlid fish radiation emerged in an extinct Pleistocene lake. *Nature* **435**, 90–95. (doi:10. 1038/nature03489)
- 27 Simmons, M. P. & Ochoterena, H. 2000 Gaps as characters in sequence-based 1. phylogenetic analyses. *Syst. Biol.* 49, 369–381. (doi:10.1093/sysbio/49.2.369)
- 28 Vos, P. et al. 1995 AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23, 4407-4414.
- 29 Swofford, D. 2002 PAUP\*: phylogenetic analysis using parsimony (\*and other methods). Sunderland, MA: Sinauer Associates, Inc. Publishers.
- 30 Huelsenbeck, J. P. & Ronquist, F. 2001 MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17, 754–755. (doi:10.1093/bioinformatics/17.8.754)
- 31 Guindon, S., Lethiec, F., Duroux, P. & Gascuel, O. 2004 PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* 33, 557–559.
- 32 Nylander, J. A. A. 2004 MRMODELTEST. Uppsala, Sweden: v2: Evolutionary Biology Centre, Uppsala University.
- 33 Felsenstein, J. 1989 PHYLIP—Phylogeny Inference Package (Version 3.2). *Cladistics* 5, 164–166.
- 34 Vekemans, X. 2002 AFLP-SURV version 1.0. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles. Belgium: Distributed by the author.

- 35 Shimodaira, H. & Hasegawa, M. 1999 Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16, 1114–1116.
- 36 Schneider, S., Roessli, D. & Excoffier, L. 2000 Arlequin: a software for population genetics data analysis. Geneva, Switzerland: Ver 2.000: Genetic and Biometry Lab, University of Geneva.
- 37 Ho, S. Y. W., Phillips, M. J., Cooper, A. & Drummond, A. J. 2005 Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol* 22, 1561–1568. (doi:10.1093/molbev/ msi145)
- 38 Genner, M. J., Seehausen, O., Lunt, D. H., Joyce, D. A., Shaw, P. W., Carvalho, G. R. & Turner, G. F. 2007 Age of cichlids: new dates for ancient lake fish radiations. *Mol. Biol. Evol.* 24, 1269–1282. (doi:10.1093/molbev/ msm050)
- 39 Falush, D., Stephens, M. & Pritchard, J. K. 2007 Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol. Ecol. Notes* 7, 574–578. (doi:10.1111/j.1471-8286. 2007.01758.x)
- 40 Hansen, M. M., Mensberg, K.-L. D. & Berg, S. 1999 Postglacial recolonization patterns and genetic relationships among whitefish (*Coregonus* sp.) populations in Denmark, inferred from mitochondrial DNA and microsatellite markers. *Mol. Ecol.* 8, 239–252. (doi:10. 1046/j.1365-294X.1999.00557.x)
- 41 Stelkens, R. & Seehausen, O. 2009 Genetic distance between species predicts novel trait expression in their hybrids. *Evolution* **63**, 884–897. (doi:10.1111/j. 1558-5646.2008.00599.x)
- 42 Fitzpatrick, B. M., Fordyce, J. A. & Gavrilets, S. 2008 What, if anything, is sympatric speciation? *J. Evol. Biol.* 21, 1452–1459. (doi:10.1111/j.1420-9101.2008.01611.x)
- 43 Mallet, J., Meyer, A., Nosil, P. & Feder, J. L. 2009 Space, sympatry and speciation. *J. Evol. Biol.* 22, 2332–2341. (doi:10.1111/j.1420-9101.2009.01816.x)
- 44 Vonlanthen, P., Roy, D., Hudson, A. G., Largiadèr, C. R., Bittner, D. & Seehausen, O. 2009 Divergence along a steep ecological gradient in lake whitefish (*Coregonus* sp.). *J. Evol. Biol.* 22, 498-514. (doi:10.1111/j.1420-9101.2008.01670.x)
- 45 Wu, C. I. & Ting, C. T. 2004 Genes and speciation. *Nat. Rev. Genet.* 5, 114–122. (doi:10.1038/nrg1269)
- 46 Barrett, R. D. & Schluter, D. 2008 Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23, 38–44. (doi:10.1016/j.tree.2007.09.008)
- 47 Schluter, D. & Nagel, L. M. 1995 Parallel speciation by natural selection. Am. Nat. 146, 292–301. (doi:10. 1086/285799)
- 48 Wagler, E. 1941 Die Lachsartigen (Salmonidae). II.Teil. Coregonen. In *Handbuch der Binnenfischerei Mitteleuropas* vol. IIIA (eds R. Demoll & H. N. Maier), pp. 371–501. Stuttgart, Germany: Schweizerbart.
- 49 Seehausen, O., van Alphen, J. J. M. & Witte, F. 1997 Cichlid fish diversity threatened by eutrophication that curbs sexual selection. *Science* 277, 1808–1811. (doi:10.1126/science.277.5333.1808)
- 50 Taylor, E. B., Boughman, J. W., Groenenboom, M., Sniatynski, M., Schluter, D. & Gow, J. L. 2006 Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus* aculeatus) species pair. Mol. Ecol. 15, 343–355. (doi:10. 1111/j.1365-294X.2005.02794.x)
- 51 Seehausen, O. 2006 Conservation: losing biodiversity by reverse speciation. *Curr. Biol.* **16**, 334–337.