


Complex and divergent histories gave rise to genome-wide divergence patterns amongst European whitefish (*Coregonus lavaretus*)

Marco Crotti¹  | Colin W. Bean^{2,3}  | Andy R. D. Gowans⁴ | Ian J. Winfield⁵ | Magdalena Butowska¹ | Josef Wanzenböck⁶  | Galina Bondarenko⁷ | Kim Præbel⁸  | Colin E. Adams^{1,2}  | Kathryn R. Elmer¹ 

¹Institute of Biodiversity, Animal Health & Comparative Medicine, College of Medical, Veterinary & Life Sciences, University of Glasgow, Glasgow, UK

²Scottish Centre for Ecology and the Natural Environment, University of Glasgow, Glasgow, UK

³NatureScot, Clydebank, UK

⁴Environment Agency, Penrith, UK

⁵Lake Ecosystems Group, UK Centre for Ecology & Hydrology, Lancaster Environment Centre, Bailrigg, UK

⁶Research Institute for Limnology, Mondsee, University of Innsbruck, Mondsee, Austria

⁷Vavilov Institute of General Genetics, Ulitsa Gubkina, Moscow, Russia

⁸Norwegian College of Fishery Science, UiT The Arctic University of Norway, Tromsø, Norway

Correspondence

Kathryn R. Elmer, Institute of Biodiversity, Animal Health & Comparative Medicine, College of Medical, Veterinary & Life Sciences, University of Glasgow, Glasgow, G12 8QQ, UK.
Email: Kathryn.Elmer@glasgow.ac.uk

Funding information

University of Glasgow; Wellcome Trust, Grant/Award Number: 097821/Z/11/Z

Abstract

Pleistocene glaciations dramatically affected species distribution in regions that were impacted by ice cover and subsequent postglacial range expansion impacted contemporary biodiversity in complex ways. The European whitefish, *Coregonus lavaretus*, is a widely distributed salmonid fish species on mainland Europe, but in Britain it has only seven native populations, all of which are found on the western extremes of the island. The origins and colonization routes of the species into Britain are unknown but likely contributed to contemporary genetic patterns and regional uniqueness. Here, we used up to 25,751 genome-wide polymorphic loci to reconstruct the history and to discern the demographic and evolutionary forces underpinning divergence between British populations. Overall, we found lower genetic diversity in Scottish populations but high differentiation ($F_{ST} = 0.433\text{--}0.712$) from the English/Welsh and other European populations. Differentiation was elevated genome-wide rather than in particular genomic regions. Demographic modelling supported a postglacial colonization into western Scotland from northern refugia and a separate colonization route for the English/Welsh populations from southern refugia, with these two groups having been separated for more than *ca.* 50 Ky. We found cyto-nuclear discordance at a European scale, with the Scottish populations clustering closely with Baltic population in the mtDNA analysis but not in the nuclear data, and with the Norwegian and Alpine populations displaying the same mtDNA haplotype but being distantly related in the nuclear tree. These findings suggest that neutral processes, primarily drift and regionally distinct pre-glacial evolutionary histories, are important drivers of genomic divergence in British populations of European whitefish. This sheds new light on the establishment of the native British freshwater fauna after the last glacial maximum.

KEYWORDS

biogeography, population genomics, postglacial history, salmonid

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Journal of Evolutionary Biology* published by John Wiley & Sons Ltd on behalf of European Society for Evolutionary Biology

1 | INTRODUCTION

Understanding the demographic and evolutionary forces shaping differentiation and speciation is a major undertaking in evolutionary biology (Seehausen et al., 2014). This is relevant not only for determining historical processes, but also to generate critical contemporary knowledge on how species are structured across their range, information that is important for conservation (Funk et al., 2012). Genomic differentiation can be rather uniform or localized in specific regions, known as islands of differentiation (Harr, 2006), which are usually considered as important genetic drivers of species divergence. Many factors will influence the distribution and magnitude of such regions (Quilodr an et al., 2020), and detailed knowledge of their patterns can improve the understanding of the evolutionary processes affecting species divergence. For example, divergent selection in populations experiencing gene flow, or secondary contact with heterogeneous gene flow following allopatric divergence, should result in few islands of high differentiation with modest and homogeneous background differentiation (Nosil et al., 2009; Ravinet et al., 2017). On the contrary, allopatric populations with limited or no gene flow may lack genomic islands of differentiation, resulting in a global and random pattern of genome-wide differentiation due to drift (Mattingsdal et al., 2020; but see Cruickshank and Hahn, 2014). Furthermore, although gene flow can swamp genomic regions undergoing divergence, it has been shown that introgression of favourable alleles from divergent lineages or populations can facilitate local adaptation (e.g. Jacobs et al., 2019; Marques et al., 2019). The use of high throughput sequencing techniques enables us to investigate such genome-wide patterns to better understand the evolutionary processes behind population and species divergence.

Because of the continuous expansion and contraction of the glaciers, Pleistocene glaciations are one of the most important events to have shaped the current distribution of flora and fauna in areas that were glaciated (Schmitt, 2007; Stewart & Lister, 2001). Many species were fragmented into lineages in separate refugia, which then experienced secondary contact at the end of the last ice age as they expanded their range into the areas previously covered by ice, leading to complex evolutionary histories (Bernatchez, 2004; Hewitt, 2004). The postglacial colonization of Britain is of particular interest from a biogeographic perspective because, with the exception of southern England, it was completely covered by ice during the Last Glacial Maximum (LGM) (Brooks et al., 2011; Clark et al., 2012).

Amongst temperate freshwater fish species, the genetic diversity and intra-specific structuring of salmonids were notably impacted by Pleistocene glaciations (e.g. Bernatchez & Dodson, 1991; Brunner et al., 2001; Jacobs et al., 2020; Koskinen et al., 2002; Ostberg et al., 2009; Pr ebel et al., 2013; Rougemont & Bernatchez, 2018). The European whitefish, *Coregonus lavaretus*, is such an example, with widespread phylogeographic structuring and the repeated evolution of distinct ecological ecotypes resulting from the interplay between secondary contact of previously allopatric lineages emerging during glacial periods and polygenic selection (Rougeux et al., 2017; Rougeux, Gagnaire, & Bernatchez, 2019; Rougeux, Gagnaire, Pr ebel,

et al., 2019). Furthermore, the same genomic regions responsible for the divergence of European whitefish are behind the parallel evolution of different ecotypes in the sister species lake whitefish *C. clupeaformis* (Rougeux et al., 2017; Rougeux, Gagnaire, Pr ebel, et al., 2019), suggesting the contribution of ancestral polymorphism in the repeated evolution of ecotypes in this genus more generally.

The current understanding is that there are three major mitochondrial lineages of European whitefish that occur in Europe and survived in separate refugia: a northern European clade with contemporary populations ranging from northwest Russia to Denmark, a Siberian clade with populations ranging from the Arctic sea to southwest Norway and a southern European clade with populations ranging from the European Alps to Denmark (Jacobsen et al., 2012; Østbye et al., 2005). In Britain, the European whitefish is only native to seven lakes: Lochs Eck and Lomond in Scotland, Llyn Tegid in Wales, and Haweswater, Red Tarn, Brotherswater and Ullswater in England (Figure 1). Although the relationship between British and continental European populations is unknown, the genetic relationship between British populations has been investigated using allozymes (Beaumont et al., 1995; Ferguson, 1974), mitochondrial restriction fragment length polymorphism (RLFPs) (Hartley, 1995), and more recently with microsatellites (Crotti, Adams, Etheridge, et al., 2020). Although some contradictions exist between these different data types, they generally suggest a separation between populations in Scotland, Wales and England, with the Welsh population being more closely related to the English populations. However, these studies have lacked contextualization by continental Europe populations and were based on few genetic markers, and therefore it was not possible to infer the drivers of differentiation, the role of complex postglacial demography, and whether major genomic structural differences are involved. The European whitefish in the UK is a priority species within the UK Biodiversity Action Plan, and it is protected under Schedule 5 of the Wildlife and Countryside Act 1981 due to its rarity and vulnerability to pressures such as invasive species, water quality and habitat degradation (Maitland & Lyle, 2013; Winfield et al., 2013). Therefore, a better understanding of the relationship between the British populations, the drivers of the population structuring between countries and the evolutionary drivers of distinctiveness is of high importance not only for evolutionary understanding but also for conservation management and policy.

In this study, a high-density genome-wide and mitochondrial dataset for samples from populations across the European whitefish range, including all seven native British populations, was used to reconstruct the evolutionary history of European whitefish in Britain. To do so, we contextualize the contemporary genetic diversity and population structure of British European whitefish with other continental European populations and reconstruct the evolutionary history and demography in a coalescent framework. To infer the role of particular genomic regions in divergence, we investigate genome-wide differentiation and admixture amongst populations. These results shed new light on the establishment of the native British freshwater fauna after the last glacial maximum and highlight the

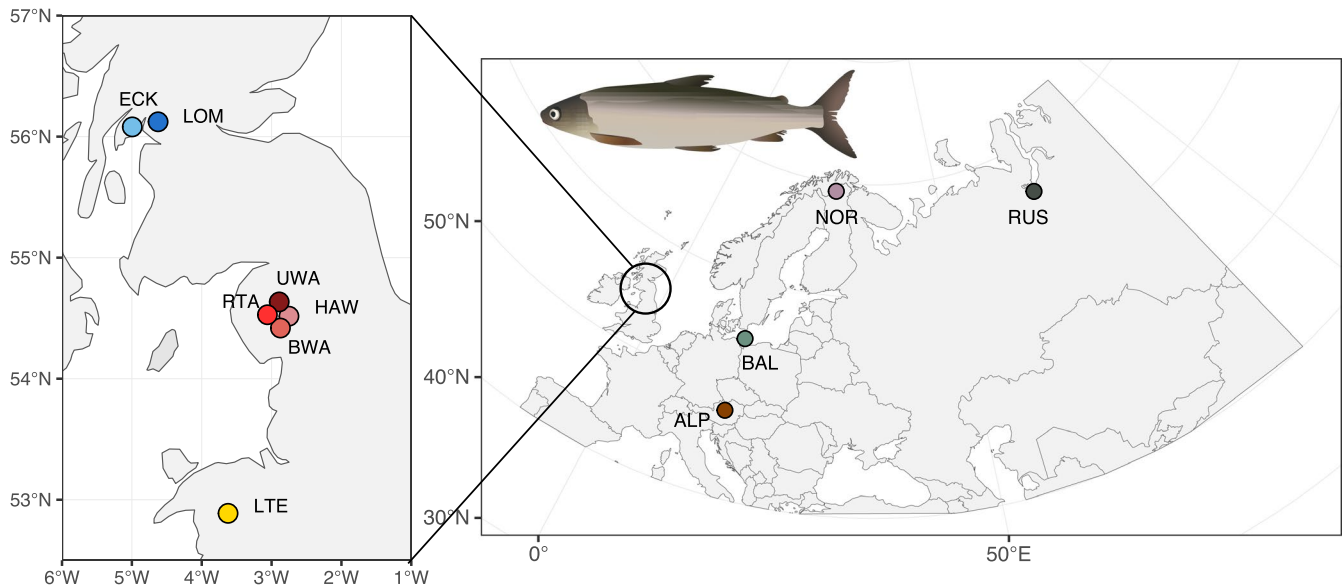


FIGURE 1 Map of the localities of the samples included in our studies. ALP, Koppentraun; BAL, Achterwasser; BWA, Brotherswater; ECK, Loch Eck; HAW, Haweswater; LOM, Loch Lomond; LTE, Llyn Tegid; NOR, Håvgajávri; RTA, Red Tarn; RUS, Ob; UWA, Ullswater. ECK and LOM are in Scotland, LTE in Wales, and RTA, HAW, BWA and UWA in England

value of genome-wide data in reconstructing evolutionary histories for conservation planning.

2 | METHODS

2.1 | Genomic data generation

We obtained a total of 102 *C. lavaretus* samples from 11 populations from across Eurasia (Figure 1): 10 individuals from Koppentraun in the Alpine region in Austria (ALP), 10 individuals from Achterwasser in the Baltic region in Germany (BAL), 10 individuals from Lake Håvgajávri in Norway (NOR), 4 individuals from the River Ob in Russia (RUS), 10 individuals from Loch Lomond (LOM) and 10 from Loch Eck (ECK) in Scotland, 11 individuals from Llyn Tegid in Wales (LTE), and from England, 11 individuals from Red Tarn (RTA), 4 individuals from Haweswater (HAW), 11 individuals from Ullswater (UWA) and 11 individuals from Brotherswater (BWA).

DNA was extracted from fish fin clips using the NucleoSpin Tissue kit (Macherey-Nagel) following the manufacturer's recommendations. The protocol used for the ddRADseq (Peterson et al., 2012) library preparation follows Crotti, Adams, and Elmer (2020), which is modified from Recknagel et al. (2015). Briefly, 1 µg of genomic DNA per sample was double digested using the rare cutting enzyme *Pst*I-HF (CATCAG recognition site) and the common cutting enzyme *Msp*I (CCGG recognition site). Combinatorial barcoded Illumina adapters were then ligated to *Pst*I-HF and *Msp*I overhangs. Samples were size selected using a PippinPrep (Sage Science) at a tight target range of 150–300 bp fragments. To enrich for the selected loci, we performed PCR amplification cycles with the following settings: 30 s at 98°C, 9X (10 s 98°C, 30 s 65°C, 30 s 72°C), 5 min 72°C. After PCR purification, the library was run on a 1.25%

agarose gel stained with SYBR Safe (Life Technologies) to remove any adapter dimers and/or fragments outside the selected size range. DNA was excised manually, cleaned and quantified using the Qubit Fluorometer with the dsDNA BR Assay (Life Technologies) to ensure the final library concentration of >1 ng/µl. The library was composed of 102 samples plus three technical replicates and was sequenced on an Illumina NextSeq500 with 75-bp paired-end reads at Glasgow Polyomics.

2.2 | Data processing

First, raw reads were demultiplexed with *process_radtags* in *Stacks* v.2.4.1 (Catchen et al., 2013; Rochette et al., 2019) and both forward and reverse reads were retained, for a paired-end approach. Reads with quality <20 were removed using *Trimmomatic* (Bolger et al., 2014). Reads were then mapped to a chromosome-level assembly (GCA_902810595.1) of the European whitefish (De-Kayne et al., 2020) using *bwa mem* v.0.7.17 (Li & Durbin, 2009) with default settings and retained if mapping quality was >20 with *samtools* v.1.7 (Li et al., 2009). A total of 10 individuals from eight populations did not pass QC due to low coverage and were excluded from the dataset. We assembled loci using *Stacks* v.2.4.1 and the *ref_map.pl* script. The raw dataset was further filtered depending on the type of analysis performed.

2.3 | Data generation, genetic diversity and population structure

The *populations* programme in *Stacks* was used to generate a vcf file with the following settings: *-p* 9 (minimum number of populations a locus must be present in to be retained), *-r* 0.8 (minimum proportion

of samples in a population required to have a locus), `--max_obs_het` 0.6 (maximum observed heterozygosity for a nucleotide at a locus), `--min_maf` 0.05 (minor allele frequency across populations required for a SNP to be included in the dataset) and `--write_single_snp` (retain one SNP per locus). This file was then filtered further with *vcftools* v.0.1.15 (Danecek et al., 2011), retaining SNPs that fulfilled the following criteria: a minimum sequencing depth of 5 reads per individual, a minimum mean sequencing depth of 10 reads across individuals, a maximum sequencing depth per individual of 30, a maximum mean sequencing depth across individuals of 30 (to remove possible repetitive reads), a minor allele frequency (MAF) of 0.05 and a 15% missing data threshold. We then used the script *filter_hwe_by_pop.pl* (available at https://github.com/jpuritz/dDocent/blob/master/scripts/filter_hwe_by_pop.pl) to filter sites that were out of HWE within populations, although no such sites were detected. Finally, we excluded a sample from the Baltic region which had more than 90% missing data after filtering. The Russian samples had between 68% and 87% missing data but were retained in the dataset only for the population structure and phylogenetic analyses. Through this process, we identified 25,751 RAD loci that were retained for the analysis (i.e. whitelist loci). Summary statistics of genetic diversity (expected heterozygosity [H_E], observed heterozygosity [H_O] and nucleotide diversity [π]) were calculated by the *population* module of *Stacks*. Rarefied allelic richness was calculated in *hierfstat* v. 0.04 (Goudet, 2005), set to a minimum of three individuals (the size of the HAW population). For all these analyses, we included up to 17 SNPs from the retained loci for a total of 41,929 SNPs, as there is no need to account for linkage disequilibrium.

Multiple approaches were used to assess population structure. For most of these analyses, we used the retained loci but only used one SNP from each (to reduce the effect of linkage), with a total of 91 samples, 25,751 SNPs, and an average of 12.4 (\pm 5.5) SNPs per 1 Mb; this is referred as the 'pop-gen dataset'. First, to evaluate population structure at the individual level, we ran a principal component analysis (PCA) with *SNPRelate* v. 1.16 (Zheng et al., 2012) in the R statistical environment (R Core Team, 2019). To investigate whether differentiation along principal components was driven by specific genomic regions, we used the function *snpgdsPCACorr* in *SNPRelate*, which calculates the correlation between a SNP genotype and the eigenvector of each principal component (Zheng et al., 2012). Second, we employed a maximum-likelihood approach for population assignment with *Admixture* v.1.3.0 (Alexander et al., 2009). We ran analyses with a 20-fold cross-validation (CV), and tested K values ranging 2–11, the optimal value was assigned to the one with the lowest CV-error. Third, to complement the model-based *Admixture* with a nonparametric approach (Linck & Battey, 2019), we used the R package *adegenet* v.2.1.1 (Jombart, 2008) to run a Discriminant Analysis of Principal Components (DAPC) (Jombart et al., 2010), which uses k -means clustering and the Bayesian information criterion to identify the most likely number of genetic clusters in the dataset. The *xvalDAPC* function was used to determine the number of PCs to be retained by the DAPC analysis. Fourth, between population Weir and Cockerham F_{ST} (Weir &

Cockerham, 1984) values were calculated in *GenoDive* (Meirns & Van Tienderenn, 2004) and significance was assessed with 10,000 permutations, pairwise D_{xy} distances were calculated with *populations* in *Stacks*, and both were visualized as heatmap matrices. Finally, to infer population structure via shared ancestry, we used the software *fineRADstructure* (Malinsky et al., 2018), which uses haplotype-based statistics, and is a modified version of the *chromopainter/fineSTRUCTURE* package (Lawson et al., 2012) but designed specifically for RADseq-type data. For this, we generated a new dataset specific for *fineRADstructure* with *populations*, excluding the Russian individuals due to high missing data (Malinsky et al., 2018). Starting from all retained loci, we retained all SNPs per locus and filtered loci with the following settings: `-p 9`, `-r 0.9`, `--max_obs_het` 0.6 and `--min_maf` 0.05. The *fineRADstructure* dataset had a total of 24,060 loci and 87 individuals. The analysis was conducted with a burn-in of 300,000 iterations, followed by 300,000 MCMC iterations and tree building with default parameters. The results were visualized with the scripts *fineradstructureplot.r* and *finestructurelibrary.r* (available at <http://cichlid.gurdon.cam.ac.uk/fineRADstructure.html>).

We used *TREEMIX* v.1.13 (Pickrell & Pritchard, 2012) to explore and visualize admixture events across populations using genome-wide allele frequency from the pop-gen dataset. This programme allows modelling migration events by adding discrete mixture events to a bifurcating population tree. Variable numbers of migration edges m , ranging 0 to 11, were tested. To account for linkage disequilibrium, we run the analysis in blocks of 100 SNPs. Because of the lack of outgroup in our dataset, we did not specify a root for the tree.

To understand whether genetic differentiation between the Scottish and other populations was evenly distributed or localized in specific genomic regions, we calculated differentiation across the genome using a window-based approach in *vcftools* with the pop-gen dataset, using a window size of 1 Mb. We compared window-based F_{ST} between the two Scottish and the remaining populations, and between the Baltic and the remaining populations, given its similar level of differentiation from all populations (see Results). Windows with fewer than five SNPs were excluded as they could be unreliable (Lehnert et al., 2020). Furthermore, we looked for regions associated with high divergence and considered as outliers any genomic windows with F_{ST} values ≥ 3 standard deviations from the mean (SD) (Lehnert et al., 2020). For this last analysis, as the aim was to determine outliers across whitefish lineages, we used only one representative population from Scotland (Loch Lomond) and one representative population from England (Red Tarn).

2.4 | Phylogenetic analysis

For phylogenetic analyses, the pop-gen dataset was converted to phylip format using the script *vcf2phylip.py* (Ortiz, 2019). Phylogenetic reconstruction was conducted under a maximum-likelihood framework in *RAXML-NG* (Kozlov et al., 2019), using a

generalized time-reversible accounting for among-site rate variation (GTR+G). Branch support was estimated using 100 bootstrap replicates. The analysis was run on the CIPRES server (<https://www.phylo.org/portal2/home.action>). In addition, we also investigated the relationship of European whitefish populations using unrooted phylogenetic networks as implemented in *SplitsTree* v.4.14.6 (Huson & Bryant, 2006) using the same dataset.

2.5 | Inference of evolutionary history

To distinguish between alternative evolutionary scenarios regarding the colonization of Britain, we used coalescence simulations implemented in *fastsimcoal2* v.2.6 (Excoffier et al., 2013). Six populations were used for this analysis: Norway, Alpine, Baltic, Lomond in Scotland, Llyn Tegid in Wales and Red Tarn in England. Only one population each from Scotland and England were used because the population structure analyses had revealed very close relationship between populations within these two regions. We generated a dataset using the filtering criteria as the pop-gen dataset, the only difference being that no MAF filter was applied to retain informative low-frequency sites (Jacobs et al., 2020). The script *sampleKgenotypesPerPop.py* (available at <https://github.com/marqueda/SFS-scripts>) was then used to down-sample each population to seven individuals, to avoid biases due to different sample sizes, and no missing data were allowed. The filtering resulted in a dataset of 42 individuals and 20,510 SNPs. The minor multidimensional site frequency spectra (MSFS) were generated with the script *easySFS.py* (available at <https://github.com/isaacovercast/easySFS>). To determine absolute values for divergence times and other inferred parameters, we corrected the number of monomorphic sites in the SFS (following Jacobs et al., 2018; Kautt et al., 2016). We used a hierarchical approach (Marques et al., 2019) by firstly optimizing a three populations model (3-Pop) that included the Alpine, Baltic and Norwegian populations from the two major European whitefish lineages and their contact zone (Østbye et al., 2005). The likelihood of each model was maximized from 100 random starting parameters, 40 iterations and 100,000 coalescent simulations. Following previous work, an estimated mutation rate of 1×10^{-8} was used, as no accurate mutation rate for salmonids is available (Rougeux et al., 2017). Once the best 3-Pop model was found, we created four population (4-Pop), five population (5-Pop) and six population (6-Pop) models by sequentially adding the remaining populations. The final model was a 6-Pop model including all populations. As the size of the MSFS increases exponentially when the number of populations is >2 and sample size is large, most of the entries in the SFS will be zero or very low values and be difficult to estimate correctly (Excoffier et al., 2013). Therefore, to reduce the number of zero values, the MSFS for all the models except the 6-Pop were projected down to eight chromosomes per population, whereas the MSFS for the 6-Pop model was projected down to six chromosomes per population (Jacobs et al., 2020). We estimated confidence intervals for the parameters of the 6-Pop model using parametric bootstrap: the point estimates

with the highest likelihood from the best fitting model as identified by *fastsimcoal* were used to simulate 50 MSFS, for which we then run 15 parameter optimizations, starting from the best parameters inferred from the observed data, and the parameter from the run with the highest likelihood for each simulated MSFS was used to compute the confidence intervals. Parameter ranges for all the models are reported in Table S1.

2.6 | Mitochondrial analysis

We sequenced mtDNA for a subset of the populations from this study. From the extracted DNA, the whole NADH dehydrogenase subunit 1 (ND1) mtDNA gene (975 bp) was PCR amplified using the primers ND1_FB1NDF: 5'-TAA GGT GGC AGA GCC CGGTA-3' and ND1_RB1NDR: 5'-TTG AAC CCC TAT TAG CCA CGC-3' (Schenekar et al., 2014), using the PCR conditions described in Schenekar et al., (2014). The PCR product was sent to *DNA Sequencing and Services* (MRC PPU) for Sanger sequencing in the forward direction for all samples but two, which were sequenced in the reverse direction. Overall, we successfully amplified the mtDNA gene ND1 from the Scotland ($n = 17$), Wales ($n = 5$), Baltic ($n = 5$), Alpine ($n = 6$) and Norway ($n = 10$) populations; sequences were not available from the English populations. Contigs were aligned to a reference sequence in *Geneious* v.7.1.9 using the highest sensitivity method. The reference sequence (GenBank JQ661447) came from an individual from Ringkøbing Fjord in Denmark belonging to the southern European mtDNA haplogroup [as defined in Jacobsen et al. (2012)]. In addition to the reference, we downloaded 12 more sequences, three from the Baltic region of Estonia (JQ661382-JQ661384), two from Ringkøbing Fjord (JQ661446, JQ661447), two from Tange lake in Denmark (JQ661433, JQ661434), three from Achterwasser in the Baltic coast of Germany (JQ661394- JQ661396), described by Jacobsen et al. (2012) and two samples of lake whitefish *Coregonus clupeaformis* from Lake Ontario in North America (MH301057, MH301058) (Schroeter et al., 2019) as outgroups. After alignment in *Geneious* v 7.1.9, we constructed a Median Joining Network in PopART (Leigh & Bryant, 2015).

3 | RESULTS

3.1 | Genetic diversity, population structure and admixture

To evaluate the extent to which European whitefish populations evolved independently or underwent significant admixture events, we analysed population structure, differentiation, co-ancestry and admixture within Britain and at the European level. A PCA demonstrated that the Norwegian population was the most distinct from all other populations, separating along PC1, which explained 22.2% of the total variance (Figure 2a). The second principal component (18.8%) separated the two Scottish populations from all

other populations, whereas the Alpine population separated on PC3 (9.77%, Figure 2b). The correlation analysis between SNP genotype and PC eigenvectors revealed a genome-wide effect on the variation explained by the principal components (Figure S1), rather than specific genomic regions, suggesting that divergence between the Scottish and the other populations on PC2 is influenced by genome-wide patterns rather than few elevated or adaptive regions.

All populations were significantly differentiated from each other, although there were substantial differences across geographic scales. We found the Scottish populations to be highly differentiated from the other British populations, with F_{ST} ranging from 0.546 to 0.661 (Table 1, Figure S4). Amongst populations within England, all F_{ST} values were <0.2 , which was slightly lower than differentiation between the two Scottish populations ($F_{ST} = 0.229$). At the European level, the Baltic population was the least differentiated from all populations ($F_{ST} = 0.298$ – 0.502), whereas the population from Norway was the most differentiated ($F_{ST} = 0.457$ – 0.712) (Table 1, Figure S4), consistent with the PCA. The pattern of absolute divergence between populations, D_{xy} , was in agreement with that of F_{ST} (Figure S4), again with the Scottish populations being highly differentiated from other British populations.

The four measures of within-population genetic variation (A_R , H_E , H_O and π) followed the same pattern across all populations (Table 2). Overall, genetic diversity was the lowest in the Scottish populations ($H_O = 0.095$ – 0.135), followed by the Alpine and Norwegian

populations ($H_O = 0.140$ – 0.149), and it was intermediate in the English and Welsh populations ($H_O = 0.157$ – 0.180), and highest in the Baltic population ($H_O = 0.244$).

The ancestry analyses showed that most populations were separated in well-defined clusters (Figure 2c,d). Nevertheless, high levels of haplotype sharing were evident between the two Scottish populations, and between the four English populations as recovered by *fineRADStructure* (Figure 2c). This was paralleled to a certain extent in the DAPC ($K = 7$) (Figure S2) and *Admixture* ($K = 8$) analyses (Figure 2d), which clustered the Scottish populations in one group or in two groups with shared ancestry respectively, and the English populations in two groups formed by Haweswater and Red Tarn and by Brotherswater and Ullswater. Furthermore, the English population of Haweswater showed signals of low shared ancestry with the Welsh population and with the cluster formed by the English Brotherswater and Ullswater populations (Figure 2d). At the European level, the Scottish populations were clustered together with the Norwegian and Baltic populations by *fineRADStructure*, whereas the English and Welsh populations clustered with the Alpine population (Figure 2c). The closer relationship between Scottish and northern European populations was also supported by the *Admixture* analysis for $K = 2$ (Figure S3). However, at $K = 3$, the Scottish populations formed their own ancestry cluster, which was present at a smaller frequency in the Russian, Baltic, Alpine and Welsh populations only (Figure S3).

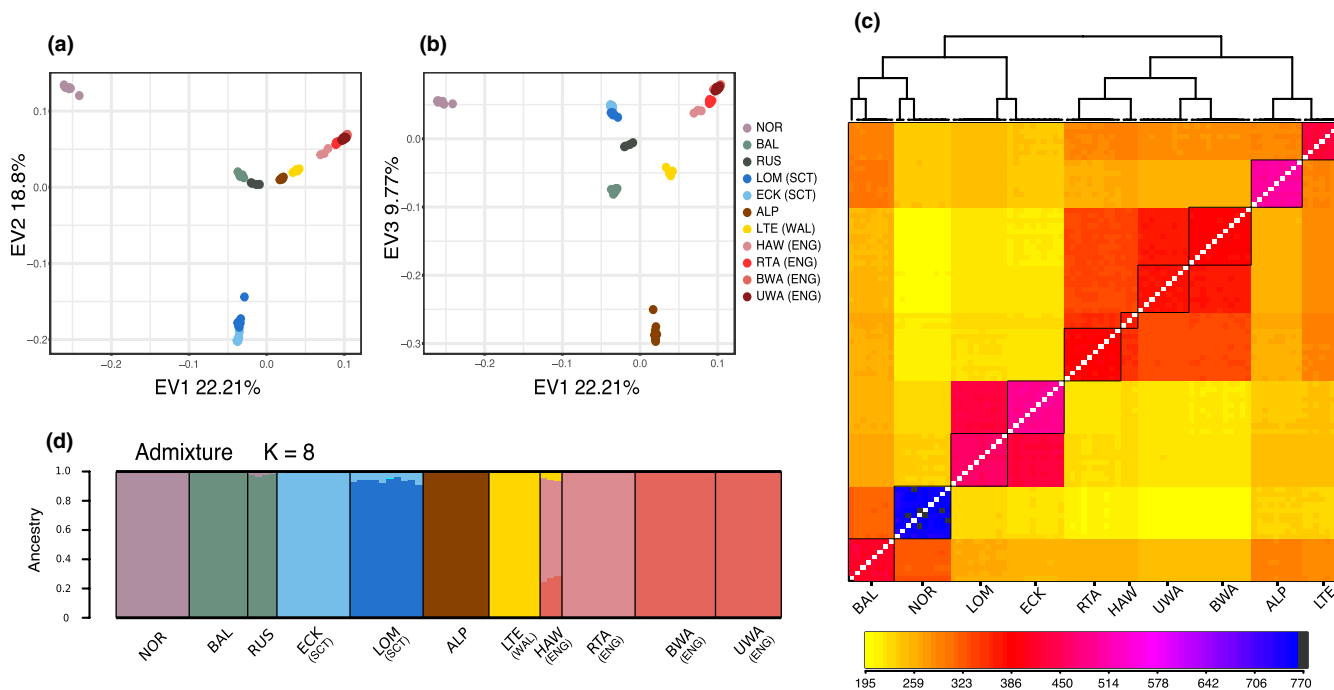


FIGURE 2 Analyses of population structure. (a) Principal component analysis showing PC1 and PC2, with each dot representing an individual and colour representing population. The first axis separates the Norwegian population, whereas the second axis separates the Scottish populations. (b) PCA showing PC1 and PC3. PC3 separates the Alpine population from all the others. (c) Co-ancestry matrix from *fineRADStructure*, with populations delimited by black squares. Colours correspond to the level of low (yellow) and high (blue/black) co-ancestry. The Scottish populations are very closely related to each other and are grouped with the Norwegian, Russian and Baltic population. (d) Ancestry coefficients from *Admixture* analysis. $K = 8$ was identified as the most likelihood number of clusters

TABLE 1 Weir & Cockerham F_{ST} calculated in *GenoDive* with 10,000 bootstrap replicates

	NOR	RUS	BAL	LOM (SCT)	ECK (SCT)	ALP	LTE (WAL)	HAW (ENG)	RTA (ENG)	BWA (ENG)	UWA (ENG)
NOR	-										
RUS	0.568	-									
BAL	0.457	0.314	-								
LOM	0.663	0.615	0.433	-							
ECK	0.712	0.692	0.502	0.229	-						
ALP	0.655	0.590	0.366	0.607	0.670	-					
LTE	0.610	0.515	0.298	0.546	0.616	0.467	-				
HAW	0.626	0.537	0.295	0.578	0.661	0.498	0.314	-			
RTA	0.628	0.549	0.355	0.569	0.626	0.497	0.351	0.094	-		
BWA	0.654	0.585	0.396	0.601	0.656	0.532	0.396	0.175	0.199	-	
UWA	0.647	0.574	0.377	0.593	0.652	0.523	0.380	0.152	0.179	0.035	-

Note: All comparisons were significant. The codes SCT, WAL and ENG identifies the population in Scotland, Wales and England, respectively.

Abbreviations: ALP, Koppentraun; BAL, Achterwasser; BWA, Brotherswater; ECK, Loch Eck; HAW, Haweswater; LOM, Loch Lomond; LTE, Llyn Tegid; NOR, Hávgaðvíri; RTA, Red Tarn; RUS, Ob; UWA, Ullswater.

Lake/river	Code	Country	N	A_R	H_E	H_O	π
Hávgaðvíri	NOR	Norway	10	1.22	0.138	0.149	0.00196
Achterwasser	BAL	Germany	8	1.38	0.235	0.244	0.00338
Ob	RUS	Russia	4	1.16	0.094	0.102	0.00120
Eck	ECK	Scotland	10	1.14	0.088	0.095	0.00125
Lomond	LOM	Scotland	10	1.20	0.124	0.135	0.00177
Koppentraun	ALP	Austria	9	1.21	0.130	0.140	0.00185
Llyn Tegid	LTE	Wales	7	1.27	0.165	0.178	0.00240
Haweswater	HAW	England	3	1.27	0.152	0.180	0.00234
Red Tarn	RTA	England	10	1.26	0.167	0.175	0.00238
Brotherswater	BWA	England	11	1.23	0.150	0.157	0.00212
Ullswater	UWA	England	9	1.24	0.154	0.165	0.00220

Note: All measures were calculated using 41,929 SNPs in 25,751 loci.

TABLE 2 Lakes and rivers sampled, sample abbreviations (Code), location (Country), sample size in the 'pop-gen dataset' (N), rarefied allelic richness (A_R), expected (H_E) and observed (H_O) heterozygosity calculated using only variant sites, nucleotide diversity (π) calculated using both variant and invariant sites

The analysis in *TREEMIX* recovered instances of gene flow amongst most populations (Figure S5). The increase in likelihood due to the addition of migration edges approximately plateaued at $m = 4$ and stabilized at $m = 7$ (Figure S5). We found consistent evidence of gene flow between Norwegian and the Baltic populations, between Welsh and English populations, and between English populations (Figure S5). When including 6 to 11 migration edges, there is also evidence of weak admixture from the base of the Scottish population branch and the English populations.

Visualization of F_{ST} across windows showed a roughly even distribution in differentiation across the genome between Scottish and the Baltic populations (Figure S6), whereas an excess of regions of higher differentiation was observed between the Scottish and the remaining populations, and an excess of regions of low differentiation was observed between the two Scottish populations (Figure S6). The Baltic population showed the highest differentiation with the Scottish populations and an excess of genomic

regions of low differentiation with the Welsh and English populations (Figure S6). When looking for genomic window outliers, we did not find any window with $F_{ST} > 3$ SD (Figure S7), again suggesting that differentiation between Scottish and other populations is driven by genome-wide mechanisms rather than particular genomic regions.

3.2 | Phylogenetic analysis

The unrooted phylogenetic network based on genomic data showed populations to be well defined (Figure 3a), but highlighted the presence of unresolved lineage sorting between the Baltic, Russian and Norwegian populations. The topology of a phylogenetic analysis based on the same data agrees with the result of that phylogenetic network, with a high support for the monophyly of all populations except for Ullswater and Brotherswater in England (Figure S8a).

With midpoint rooting applied, the Norwegian and Russian populations formed a monophyletic group sister to a clade containing all the other populations. The first population to split in this clade was the Baltic, followed by the two Scottish populations (Figure S8a). The four English populations were more closely related to the Welsh population, and these two lineages shared a more recent common ancestor with the Alpine population than with any other population (Figure S8a).

3.3 | Mitochondrial analysis

Scottish populations possessed mitochondrial haplotypes that were predominantly in the northern European mtDNA clade (Figure 3b). The Welsh populations were placed in the southern European clade, being more similar to the Danish populations of Ringkøbing Fjord and Tange Lake (Jacobsen et al., 2012). The Scottish populations showed unique haplotypes, which were more closely related to Baltic ones (Figure 3b). Furthermore, the Norwegian and Alpine possessed the same haplotypes and were central in the haplotype network.

3.4 | Evolutionary history

Using the inference of phylogenetic relationships in the ML tree, we estimated the demographic history of European whitefish using a hierarchical approach. The demographic scenarios for all the models tested are found in Figure 4. The 3-Pop model indicated the oldest divergence to be between the Norwegian and the group formed by the Alpine and Baltic population, with subsequent gene flow from

the Alpine and Norwegian populations into the Baltic one (Figure 4a, Table S2). The 4-Pop model showed the Scottish population of Loch Lomond splitting from the Alpine population, with additional gene flow with the Baltic population (Figure 4b, Table S2). In the 5-Pop model, the Welsh population split from the Alpine population more recently than the Scottish population, and in the 6-Pop model, the English population of Red Tarn shared a more recent common ancestor with the Welsh population (Figure 4c,d, Table S2). In both 5-Pop and 6-Pop models, the Baltic population received gene flow from the ancestor branch of the Welsh and English population (Figure 4c,d).

To infer coalescent times, confidence intervals were calculated for the 6-Pop model using parametric bootstrapping (Table S3, Figure 5). We estimated the first divergence event, between the Norwegian population and the others, occurred 26,247 generations ago (CI: 23,451–26,774), the divergence between the Baltic population and the remaining occurred 23,205 (CI: 20,733–23,671), whereas the Scottish population diverged from the Alpine, Welsh and English populations 18,674 generations ago (CI: 16,685–19,435). On the contrary, the divergence between the Alpine and the Welsh/English populations, and between the Welsh and English populations, were more recent, the former having occurred 5965 generations ago (CI: 5313–6255) and the latter 3887 generations ago (CI: 3337–4259). Furthermore, our suggested inference that the Baltic population received gene flow from all other populations, with admixture proportion into the Baltic ranging from 0.15 from the Lomond population to 0.46 from the Norwegian population, between 6663 (CI: 5953–6934) and 4433 (CI: 3949–4649) generations ago. We also found evidence of low gene flow (admixture proportion = 0.15) from the Lomond to the Welsh population at 3473 (CI: 3093–3641) generations ago.

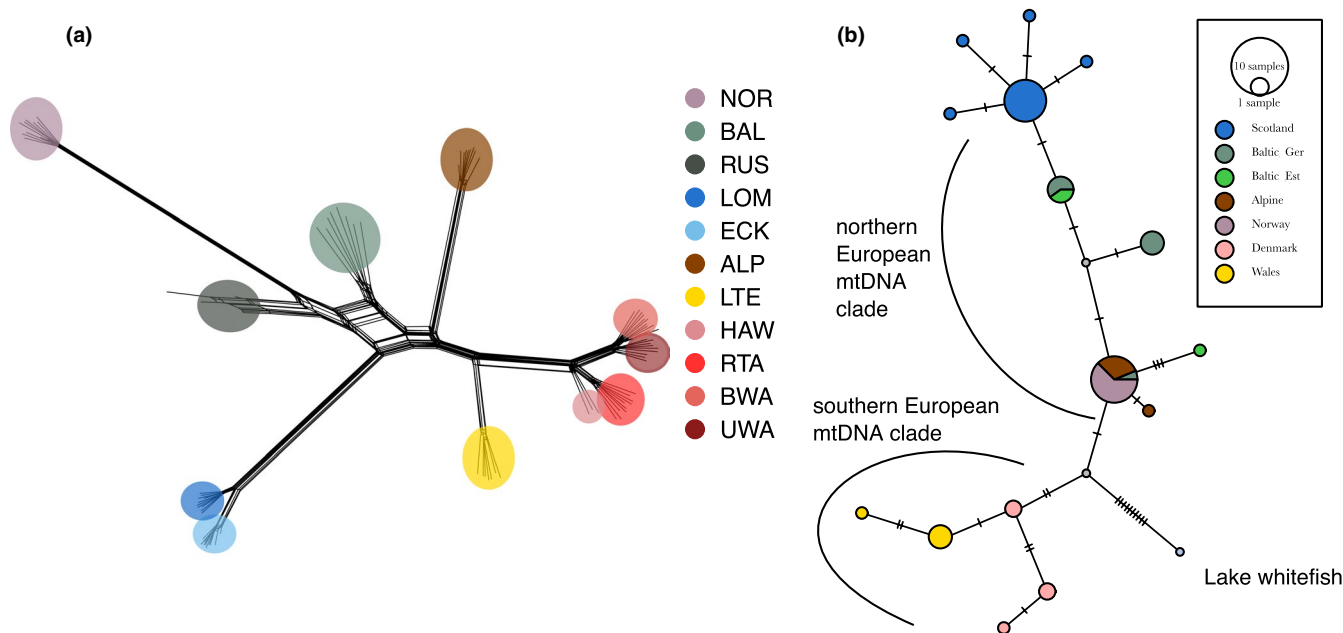
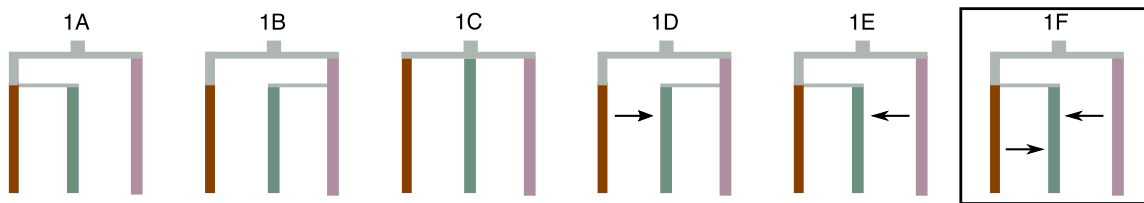
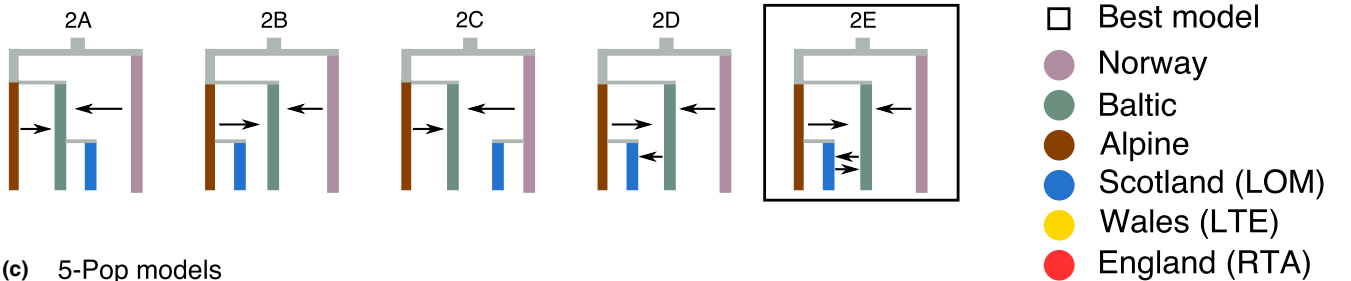


FIGURE 3 Results of phylogenetic and mitochondrial analyses. (a) Unrooted phylogenetic network calculated with SplitsTree using the genome-wide SNPs data. (b) Median-joining network calculated in PopART with $\epsilon = 0$. Colour code same as in previous figures. Hash marks denote mutational steps, and the size of the circle is proportional to the haplotype frequency

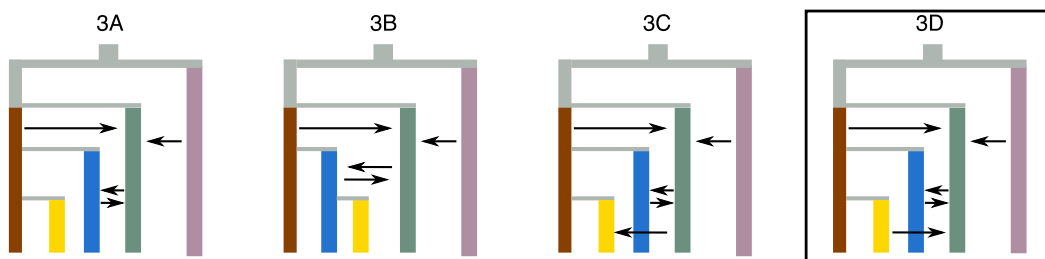
(a) 3-Pop models



(b) 4-Pop models



(c) 5-Pop models



(d) 6-Pop models

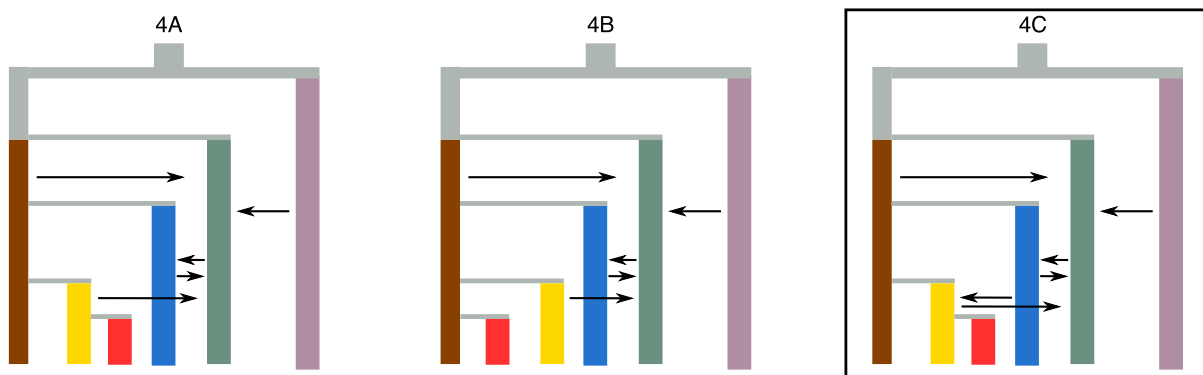


FIGURE 4 Hierarchical demographic modelling approach. (a) First, we established the relationship of three European whitefish populations, specifically the Norwegian, Baltic and Alpine. The data best supported a more recent split between the Baltic and Alpine populations, with the Baltic receiving gene flow from the other two populations. (b) Next, we tested the origin of the Scottish population, with the best supporting model suggesting a split between the Scottish and Alpine population, with genetic contribution from the Baltic population. (c) We then added the Welsh population, for which the modelling indicates a closer relationship with the Alpine population than with the Scottish one. (d) Finally, we tested the English population, which was found to be splitting from the Welsh population

4 | DISCUSSION

In our study to reconstruct the evolutionary and demographic history of European whitefish in Britain, analyses based on genome-wide data support the role of historical genetic isolation due to

demographic history and genetic drift as the main drivers of population differentiation. In fact, we found lower genetic diversity and higher differentiation in the Scottish populations compared with the other British and continental European populations, consistent with a scenario of low effective population size maintained over a long

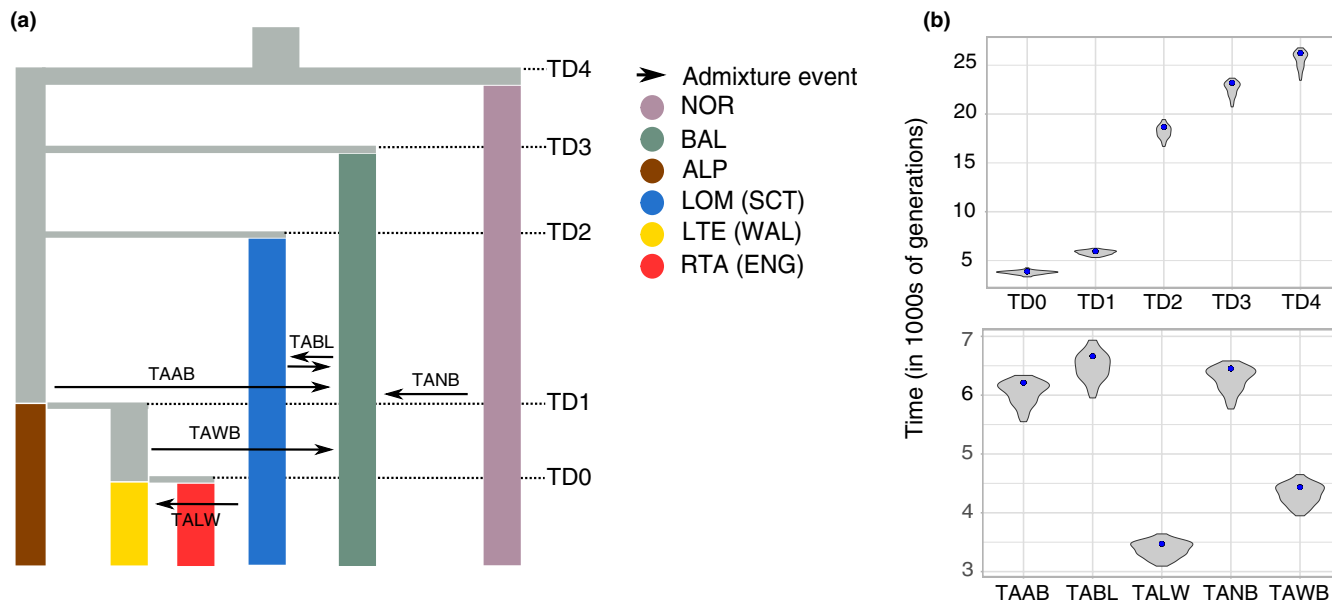


FIGURE 5 Demographic history of European whitefish. (a) The best fitting 6-Pop model, with splitting and admixture events annotated. (b) Confidence intervals for the time of divergence and admixture events obtained by parametric bootstrapping, with maximum-likelihood parameter estimates indicated by blue dots. Time unit is in generations before present

period of time in small glacial refugia (Westergaard et al., 2019). In addition, we found evidence of two separate routes for the colonization of Britain, as the Welsh and English populations shared a more recent common ancestor with an Alpine population than with the two Scottish populations. Finally, we inferred high levels of background genome-wide differentiation between Scottish and remaining populations.

In agreement with the findings of previous studies based on mtDNA and microsatellites (Crotti, Adams, Etheridge, et al., 2020; Hartley, 1995), the Welsh population had the highest genetic variation within Britain and formed its own genetic cluster. The four English populations are very closely related, with high levels of within-population genetic diversity and low between-population genetic differentiation. Interestingly, although allozyme and microsatellite analyses suggested the English population of Red Tarn to be the most distantly related of the four (Beaumont et al., 1995; Crotti, Adams, Etheridge, et al., 2020), our SNPs analyses show a very close relationship between Red Tarn and Haweswater.

We found evidence of elevated genomic differentiation between the Scottish and the remaining populations across the entire genome but did not detect any genomic islands of differentiation. Although in an allopatric model of population divergence, the number of islands is expected to increase with time (Nosil et al., 2009), it has been shown that long periods of isolation and limited gene flow can mask the signals of genomic islands, as it will be swamped by elevated background genetic differentiation due to drift (Mattingsdal et al., 2020; Quilodrán et al., 2020; Wang et al., 2016). The high level of genetic differentiation between Scottish and all other populations could explain why we did not find specific genomic regions involved in divergence. Furthermore, the demographic modelling supports

a deep separation between the Scottish and the remaining British populations. Assuming an average generation time of 3.5 years (Rougeux et al., 2017), our estimates indicate that the Scottish populations split from their common ancestor with the Alpine, Welsh and English populations ca. 65 Kya, and remained isolated from the British populations until ca. 12 Kya, when they experienced low gene flow with the ancestor of the Welsh and English populations. This extended period of isolation for this lineage is also evident from the long branch in the phylogenetic tree and network, and the strong separation on PC2. Nevertheless, our models did not explore potential additional factors that could impact the time of divergence estimation, such as the effect of bottlenecks followed by expansion and heterogeneous gene flow during divergence (Le Moan et al., 2016; Momigliano et al., 2021; Rougeux et al., 2017), so we acknowledge that some caution is needed when interpreting these results. We did not find evidence, in the form of elevated F_{ST} within specific chromosomes, of chromosomal rearrangements between Scottish and the other British and European populations. Despite an allopatric phase of ca. 15 to 60 Ky, an absence of large-scale chromosomal rearrangements has been reported for lake whitefish species pairs (Dion-Côté et al., 2015, 2017), suggesting karyotypic stability within *Coregonus*. However, substantial intrachromosomal polymorphism, particularly in regions associated with heterochromatin and repetitive DNA, was identified between sympatric species pairs of lake whitefish and vendace *Coregonus albula* (Dion-Côté et al., 2017; Symonová et al., 2013), indicating that other mechanisms could be promoting differentiation in *Coregonus* populations.

A possible refugium for the Scottish lineage could have been located in the North Sea, as suggested for brown trout *Salmo trutta* (McKeown et al., 2010), bullhead *Cottus gobio* (Volckaert et al., 2002)

and grayling *Thymallus thymallus* (Gum et al., 2005); alternatively, this lineage could have survived in the putative Lough Hibernia, in the present-day Irish sea (Maitland, 1970). The Welsh and English populations split from the Alpine lineage much more recently, ca. 20 Kya, towards the end of the last glacial maximum (Clark et al., 2012). The Alpine lineage is thought to have had its refuge near the mouth of the river Rhine (Hudson et al., 2011) and colonized the Alpine lakes and the Jutland peninsula through the Rhine and Elbe river systems (Hansen et al., 2008; Østbye et al., 2005). At the time of the split, instead of moving towards central Europe, the ancestral stock that gave rise to the Welsh and English populations must have moved westwards towards Britain. The divergence between the English and Welsh populations at ca. 14 Kya agrees quite well with the retreat of the glaciers from Britain, as parts of the Lake District would have been already free of ice at that time (Brooks et al., 2011; Pinson et al., 2013). The subsequent Loch Lomond Stadial, or Younger Dryas, in Britain (12.9–11.7 Kya) (Bickerdike et al., 2018), a cold period which saw the re-advancement of glaciers in Scotland, the English Lake District and Wales, could have promoted fragmentation of the whitefish populations in England.

Our results support a scenario in which European whitefish established in Britain in two different migratory waves, one formed by Welsh and English populations and one formed by the Scottish populations. Previous hypotheses suggested that *Coregonus* established in British lakes from an ancestral anadromous stock that was inhabiting Lough Hibernia (Maitland, 1970). Then as the glaciers started retreating, British lakes were colonized in different events; in fact, the Scottish lochs are low altitude, whereas the Welsh and English lakes are high altitude, suggesting that the European whitefish established in these water systems at different time points (Wheeler, 1977). This theory partially agrees with our finding of successive waves of colonization by European whitefish. However, we estimate a ca. 50k-year-old divergence and very limited gene flow between Scottish and remaining British populations, suggesting only a short phase of these lineages living in sympatry, or a longer phase of extensive reproductive isolation. This north/south genetic separation is unique amongst salmonids in Britain, which show diverse evolutionary histories. In Britain, the Arctic charr *Salvelinus alpinus* has a similar geographic distribution to whitefish, but to date, there is no evidence of high differentiation between British populations in mtDNA (Brunner et al., 2001; Wilson et al., 2004), and equivalent nuclear analyses have never been conducted. The grayling likely colonized Britain through the North Sea River from the central European lineage, but there is the possibility of a separate colonization for the northern England populations, as they contain both basal and extra haplotypes compared with southern English populations (Dawnay et al., 2011). British populations of brown trout and Atlantic salmon both descend from multiple genetic lineages, with evidence for subsequent hybridization between lineages (Finnegan et al., 2013; García-Marín et al., 1999; McKeown et al., 2010). This suggests that species-specific and geographically broad approaches are needed if we are to reconstruct the postglacial history of fishes in the British Isles.

The ancestors of the Norwegian population diverged from all other populations ca. 92 Ky ago, during a long interglacial stage known as GI 23.1 (Greenland Interstadial, Rasmussen et al., 2014) or MIS (Marine Isotope Stage) 5c (Oppo et al., 2001). During this period, temperatures increased and tree vegetation reached Scandinavia (Hall et al., 2002). The retreat of the glaciers could have allowed the ancestral stock of European whitefish to move westward. Through the early drainages of the ice lakes (Østbye et al., 2005), whitefish entered the Baltic sea around 81 Kya, corresponding to the interstadial GI 21.1 (Rasmussen et al., 2014) or MIS 5a (Oppo et al., 2001). The Baltic population in this study showed the highest genetic diversity. The high genetic diversity could be explained by repeated gene flow between the Baltic population and the other lineages, as suggested by the low differentiation, treemix results and the demographic modelling. These introgression events are also supported by mtDNA analysis (Østbye et al., 2005). In particular, the introgression between the ancestor of the Baltic and Alpine population could explain the presence of northern mtDNA haplotypes in the Alpine whitefish radiation (Hudson et al., 2011; Østbye et al., 2005). Overall, our results add support to the evidence of the Baltic sea as a recognized contact zone between divergent lineages of aquatic species (Johannesson et al., 2020).

Our analysis based on nuclear data does not find a clear distinction between a southern and a northern European whitefish clade, as suggested by mtDNA (Hudson et al., 2011; Jacobsen et al., 2012; Østbye et al., 2005). Instead, the demographic modelling finds a scenario in which three populations, Norwegian, Baltic and Scottish, split one after the other from the common ancestor of European whitefish within a 30 Ky period, from 92 to 65 Kya; then, we observe no split until 20 Ky, when the Alpine lineage splits from the Welsh and English stock. The Baltic population analysed here is from Achterwasser in northern Germany, and it was previously found to possess the northern mtDNA haplogroup (Jacobsen et al., 2012; Winkler et al., 2011). The fact that the demographic modelling shows the Baltic population splitting after the ancestor of the Norwegian population, rather than from it, adds support to the findings that the separation between southern and northern mtDNA haplogroups is not found in nuclear data. Furthermore, in our and previous analyses (Winkler et al., 2011), the northern European mtDNA haplotypes are prevalent in the Alpine population of Koppentraun, Austria. However, the nuclear data show a strong divergence between the Alpine and Norwegian populations, adding another line of evidence to the discrepancy between nuclear and mitochondrial DNA signal for European whitefish. Discordance between nuclear DNA and mtDNA is well documented in the literature (reviewed by Toews & Brelsford, 2012) and might be a common feature of species whose demographic history has been heavily influenced by glacial cycles (Dufresnes et al., 2020). Cyto-nuclear discordance has also been recovered by previous studies on whitefish, in particular in the Alpine whitefish radiation, where nuclear data supported a monophyletic origin of all Alpine populations, whereas mtDNA showed the presence of both southern and northern mtDNA haplogroups (Hudson et al., 2011), and in the Danish and Baltic populations, where F_{ST} differentiation between Baltic (northern mtDNA

haplogroup) and Danish (southern mtDNA haplogroup) was in some cases an order of magnitude higher for mtDNA than from microsatellite data (Hansen et al., 1999).

5 | CONCLUSION

We provide molecular evidence of a deep divergence between, and a dual colonization of, Britain by Scottish and Welsh/English populations of European whitefish, which reflects a distinctive pre LGM pattern at the edge of the species distribution relevant to continental recolonization. Previous studies uncovered a similar structuring between the seven British populations (Beaumont et al., 1995; Crotti, Adams, Etheridge, et al., 2020; Hartley, 1995), but it is the first time it has been possible to estimate the time of divergence, highlighting the increased resolution NGS data can provide for conservation (Funk et al., 2012). Our findings support the delimitation of separate conservation units (CUs) (Funk et al., 2012) for each population and, in particular, highlight the uniqueness of the Scottish lineage which diverged and survived in isolation with limited contact for ca. 50K years. Although we did not detect genomic islands of differentiation between Scottish and remaining populations, future work could look at the adaptive differentiation and demographic history between populations within country, that is between the two Scottish populations and between the four English population, to better understand the adaptive and demographic process operating at the local level (Flanagan et al., 2018). Conservation measures have already been implemented to preserve the genetic stock of most British populations through the use of conservation translocations (Adams et al., 2014; Crotti et al., 2021; Maitland & Lyle, 2013; Thomas et al., 2013; Winfield et al., 2013), and we encourage national conservation bodies to look beyond the simple species-based paradigm as they aim to protect the diversity of European whitefish at the most western edge of its range.

ACKNOWLEDGEMENTS

This work was funded by the University of Glasgow College of Medical, Veterinary and Life Sciences doctoral training programme and Wellcome Trust-Glasgow Polyomics ISSF pilot funding 097821/Z/11/Z. We thank Glasgow Polyomics and J Galbraith for NGS sequencing, M Capstick for support in the laboratory and A Jacobs for analysis advice and comments on the draft manuscript. P Maitland, A Lyle and E Etheridge provided valuable discussion on the value of this study.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

AUTHOR CONTRIBUTIONS

MC, CEA and KRE designed the research. MC, CWB, ARDG, IJW, JW, GB, KP and CEA provided biological samples. MC and MB conducted laboratory work with support from KRE. MC conducted the

analyses. MC wrote the paper with contributions from KRE and CEA. All authors edited/approved the final manuscript.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/jeb.13948>.

DATA AVAILABILITY STATEMENT

Scripts, data and the parameter data files for the demographic modelling in fastsimcoal are archived and made available in the University of Glasgow Enlighten Repository (<http://dx.doi.org/10.5525/gla.researchdata.1059>). Short read data are available at NCBI SRA (BioProject PRJNA648856) and mtDNA at GenBank (accessions MT862810–MT862852).

ORCID

Marco Crotti  <https://orcid.org/0000-0002-8619-7988>

Colin W. Bean  <https://orcid.org/0000-0003-3502-0995>

Josef Wanzenböck  <https://orcid.org/0000-0002-9186-5565>

Kim Præbel  <https://orcid.org/0000-0002-0681-1854>

Colin E. Adams  <https://orcid.org/0000-0003-2470-9754>

Kathryn R. Elmer  <https://orcid.org/0000-0002-9219-7001>

REFERENCES

- Adams, C. E., Lyle, A. A., Dodd, J. A., Bean, C. W., Winfield, I. J., Gowans, A. R. D., Stephen, A., & Maitland, P. S. (2014). Translocation as a conservation tool: case studies from rare freshwater fishes in Scotland. *The Glasgow Naturalist*, 26(1), 17–24.
- Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Beaumont, A. R., Bray, J., Murphy, J. M., & Winfield, I. J. (1995). Genetics of whitefish and vendace in England and Wales. *Journal of Fish Biology*, 46(5), 880–890. <https://doi.org/10.1111/j.1095-8649.1995.tb01611.x>
- Bernatchez, L. (2004). Ecological theory of adaptive radiation: An empirical assessment from coregonine fishes (Salmoniformes). In A. P. Hendry, & S. C. Stearns (Eds.), *Evolution illuminated: Salmon and their relatives* (pp. 175–207). Oxford University Press.
- Bernatchez, L., & Dodson, J. J. (1991). Phylogeographic structure in mitochondrial DNA of the lake whitefish (*Coregonus clupeaformis*) and its relation to Pleistocene glaciations. *Evolution*, 45(4), 1016–1035. <https://doi.org/10.1111/j.1558-5646.1991.tb04367.x>
- Bickerdike, H. L., Evans, D. J. A., Stokes, C. R., & Ó Cofaigh, C. (2018). The glacial geomorphology of the Loch Lomond (Younger Dryas) Stadial in Britain: A review. *Journal of Quaternary Science*, 33(1), 1–54. <https://doi.org/10.1002/jqs.3010>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Brooks, A. J., Bradley, S. L., Edwards, R. J., & Goodwyn, N. (2011). The palaeogeography of Northwest Europe during the last 20,000 years. *Journal of Maps*, 7(1), 573–587. <https://doi.org/10.4113/jom.2011.1160>
- Brunner, P. C., Douglas, M. R., Osinov, A., Wilson, C. C., & Bernatchez, L. (2001). Holarctic phylogeography of arctic charr (*Salvelinus alpinus* L.) Inferred from mitochondrial DNA sequences. *Evolution*, 55(3), 573–586. <https://doi.org/10.1111/j.0014-3820.2001.tb00790.x>

- Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, 22(11), 3124–3140. <https://doi.org/10.1111/mec.12354>
- Clark, C. D., Hughes, A. L. C., Greenwood, S. L., Jordan, C., & Sejrup, H. P. (2012). Pattern and timing of retreat of the last British-Irish Ice Sheet. *Quaternary Science Reviews*, 44, 112–146. <https://doi.org/10.1016/j.quascirev.2010.07.019>
- Crotti, M., Adams, C. E., & Elmer, K. R. (2020). Population genomic SNPs from epigenetic RADs: Gaining genetic and epigenetic data from a single established next-generation sequencing approach. *Methods in Ecology and Evolution*, 11(7), 839–849. <https://doi.org/10.1111/2041-210X.13395>
- Crotti, M., Adams, C. E., Etheridge, E. C., Bean, C. W., Gowans, A. R. D., Knudsen, R., Lyle, A. A., Maitland, P. S., Winfield, I. J., Elmer, K. R., & Præbel, K. (2020). Geographic hierarchical population genetic structuring in British European whitefish (*Coregonus lavaretus*) and its implications for conservation. *Conservation Genetics*, 1, 3. <https://doi.org/10.1007/s10592-020-01298-y>
- Crotti, M., Yohannes, E., Winfield, I. J., Lyle, A. A., Adams, C. E., & Elmer, K. R. (2021). Rapid adaptation through genomic and epigenomic responses following translocations in an endangered salmonid. *Evolutionary Applications*. <https://doi.org/10.1111/EVA.13267>
- Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23(13), 3133–3157. <https://doi.org/10.1111/mec.12796>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158. <https://doi.org/10.1093/bioinformatics/btr330>
- Dawney, N., Dawney, L., Hughes, R. N., Cove, R., & Taylor, M. I. (2011). Substantial genetic structure among stocked and native populations of the European grayling (*Thymallus thymallus*, Salmonidae) in the United Kingdom. *Conservation Genetics*, 12(3), 731–744. <https://doi.org/10.1007/s10592-010-0179-4>
- De-Kayne, R., Zöller, S., & Feulner, P. G. D. (2020). A *de novo* chromosome-level genome assembly of *Coregonus* sp. “Balchen”: One representative of the Swiss Alpine whitefish radiation. *Molecular Ecology Resources*, 20(4), 1093–1109. <https://doi.org/10.1111/1755-0998.13187>
- Dion-Côté, A.-M., Symonová, R., Lamaze, F. C., Pelikánová, Š., Ráb, P., & Bernatchez, L. (2017). Standing chromosomal variation in Lake Whitefish species pairs: The role of historical contingency and relevance for speciation. *Molecular Ecology*, 26(1), 178–192. <https://doi.org/10.1111/mec.13816>
- Dion-Côté, A.-M., Symonová, R., Ráb, P., & Bernatchez, L. (2015). Reproductive isolation in a nascent species pair is associated with aneuploidy in hybrid offspring. *Proceedings of the Royal Society B: Biological Sciences*, 282(1802), 20142862. <https://doi.org/10.1098/rspb.2014.2862>
- Dufresnes, C., Nicieza, A. G., Litvinchuk, S. N., Rodrigues, N., Jeffries, D. L., Vences, M., Perrin, N., & Martínez-Solano, Í. (2020). Are glacial refugia hotspots of speciation and cyto-nuclear discordances? Answers from the genomic phylogeography of Spanish common frogs. *Molecular Ecology*, 29(5), 986–1000. <https://doi.org/10.1111/mec.15368>
- Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, 9(10), e1003905. <https://doi.org/10.1371/journal.pgen.1003905>
- Ferguson, A. (1974). The genetic relationships of the coregonid fishes of Britain and Ireland indicated by electrophoretic analysis of tissue proteins. *Journal of Fish Biology*, 6(3), 311–315. <https://doi.org/10.1111/j.1095-8649.1974.tb04548.x>
- Finnegan, A. K., Griffiths, A. M., King, R. A., Machado-Schiaffino, G., Porcher, J.-P., Garcia-Vazquez, E., Bright, D., & Stevens, J. R. (2013). Use of multiple markers demonstrates a cryptic western refugium and postglacial colonisation routes of Atlantic salmon (*Salmo salar* L.) in northwest Europe. *Heredity*, 111(1), 34–43. <https://doi.org/10.1038/hdy.2013.17>
- Flanagan, S. P., Forester, B. R., Latch, E. K., Aitken, S. N., & Hoban, S. (2018). Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. *Evolutionary Applications*, 11(7), 1035–1052. <https://doi.org/10.1111/eva.12569>
- Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*, 27(9), 489–496. <https://doi.org/10.1016/j.tree.2012.05.012>
- García-Marín, J., Utter, F. M., & Pla, C. (1999). Postglacial colonization of brown trout in Europe based on distribution of allozyme variants. *Heredity*, 82(1), 46–56. <https://doi.org/10.1046/J.1365-2540.1999.00447.X>
- Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes*, 5(1), 184–186. <https://doi.org/10.1111/j.1471-8286.2004.00828.x>
- Gum, B., Gross, R., & Kuehn, R. (2005). Mitochondrial and nuclear DNA phylogeography of European grayling (*Thymallus thymallus*): Evidence for secondary contact zones in central Europe. *Molecular Ecology*, 14(6), 1707–1725. <https://doi.org/10.1111/j.1365-294X.2005.02520.x>
- Hall, A. M., Gordon, J. E., Whittington, G., Duller, G. A. T., & Heijnis, H. (2002). Sedimentology, palaeoecology and geochronology of Marine Isotope Stage 5 deposits on the Shetland Islands, Scotland. *Journal of Quaternary Science*, 17(1), 51–67. <https://doi.org/10.1002/jqs.644>
- Hansen, M. M., Fraser, D. J., Als, T. D., & Mensberg, K.-L.-D. (2008). Reproductive isolation, evolutionary distinctiveness and setting conservation priorities: The case of European lake whitefish and the endangered North Sea houting (*Coregonus* spp.). *BMC Evolutionary Biology*, 8(1), 137. <https://doi.org/10.1186/1471-2148-8-137>
- 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. *Molecular Ecology*, 8(2), 239–252. <https://doi.org/10.1046/j.1365-294X.1999.00557.x>
- Harr, B. (2006). Genomic islands of differentiation between house mouse subspecies. *Genome Research*, 16(6), 730–737. <https://doi.org/10.1101/gr.5045006>
- Hartley, S. E. (1995). Mitochondrial DNA analysis distinguishes between British populations of the whitefish. *Journal of Fish Biology*, 47, 145–155. <https://doi.org/10.1111/j.1095-8649.1995.tb06050.x>
- Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1442), 183–195. <https://doi.org/10.1098/rstb.2003.1388>
- Hudson, A. G., Vonlanthen, P., & Seehausen, O. (2011). Rapid parallel adaptive radiations from a single hybridogenetic ancestral population. *Proceedings of the Royal Society B: Biological Sciences*, 278(1702), 58–66. <https://doi.org/10.1098/rspb.2010.0925>
- Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23(2), 254–267. <https://doi.org/10.1093/molbev/msj030>
- Jacobs, A., Carruthers, M., Eckmann, R., Yohannes, E., Adams, C. E., Behrmann-Godel, J., & Elmer, K. R. (2019). Rapid niche expansion by selection on functional genomic variation after ecosystem

- recovery. *Nature Ecology and Evolution*, 3(1), 77–86. <https://doi.org/10.1038/s41559-018-0742-9>
- Jacobs, A., Carruthers, M., Yurchenko, A., Gordeeva, N. V., Alekseyev, S. S., Hooker, O., Leong, J. S., Minkley, D. R., Rondeau, E. B., Koop, B. F., Adams, C. E., & Elmer, K. R. (2020). Parallelism in eco-morphology and gene expression despite variable evolutionary and genomic backgrounds in a Holarctic fish. *PLoS Genetics*, 16(4), e1008658. <https://doi.org/10.1371/journal.pgen.1008658>
- Jacobs, A., Hughes, M., Robinson, P., Adams, C., & Elmer, K. (2018). The genetic architecture underlying the evolution of a rare piscivorous life history form in brown trout after secondary contact and strong introgression. *Genes*, 9(6), 280. <https://doi.org/10.3390/genes9060280>
- Jacobsen, M. W., Hansen, M. M., Orlando, L., Bekkevold, D., Bernatchez, L., Willerslev, E., & Gilbert, M. T. P. (2012). Mitogenome sequencing reveals shallow evolutionary histories and recent divergence time between morphologically and ecologically distinct European whitefish (*Coregonus* spp.). *Molecular Ecology*, 21(11), 2727–2742. <https://doi.org/10.1111/j.1365-294X.2012.05561.x>
- Johannesson, K., Le Moan, A., Perini, S., & André, C. (2020). A Darwinian laboratory of multiple contact zones. *Trends in Ecology and Evolution*, 35(11), 1021–1036. <https://doi.org/10.1016/j.tree.2020.07.015>
- Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, 11(1), 94. <https://doi.org/10.1186/1471-2156-11-94>
- Kautt, A. F., Machado-Schiaffino, G., & Meyer, A. (2016). Multispecies outcomes of sympatric speciation after admixture with the source population in two radiations of Nicaraguan crater lake cichlids. *PLoS Genetics*, 12(6), e1006157. <https://doi.org/10.1371/journal.pgen.1006157>
- Koskinen, M. T., Nilsson, J., Veselov, A. J., Potutkin, A. G., Ranta, E., & Primmer, C. R. (2002). Microsatellite data resolve phylogeographic patterns in European grayling, *Thymallus thymallus*. *Salmonidae*. *Heredity*, 88(5), 391–401. <https://doi.org/10.1038/sj.hdy.6800072>
- Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453–4455. <https://doi.org/10.1093/bioinformatics/btz305>
- Lawson, D. J., Hellenthal, G., Myers, S., & Falush, D. (2012). Inference of population structure using dense haplotype data. *PLoS Genetics*, 8(1), e1002453. <https://doi.org/10.1371/journal.pgen.1002453>
- Le Moan, A., Gagnaire, P.-A., & Bonhomme, F. (2016). Parallel genetic divergence among coastal-marine ecotype pairs of European anchovy explained by differential introgression after secondary contact. *Molecular Ecology*, 25(13), 3187–3202. <https://doi.org/10.1111/mec.13627>
- Lehnert, S. J., Kess, T., Bentzen, P., Clément, M., & Bradbury, I. R. (2020). Divergent and linked selection shape patterns of genomic differentiation between European and North American Atlantic salmon (*Salmo salar*). *Molecular Ecology*, 29(12), 2160–2175. <https://doi.org/10.1111/mec.15480>
- Leigh, J. W., & Bryant, D. (2015). POPART: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, 6(9), 1110–1116. <https://doi.org/10.1111/2041-210X.12410>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Linck, E., & Battey, C. J. (2019). Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. *Molecular Ecology Resources*, 19(3), 639–647. <https://doi.org/10.1111/1755-0998.12995>
- Maitland, P. S. (1970). The origin and present distribution of *Coregonus* in the British Isles. In C. C. Lindsey, & C. S. Woods (Eds.), *Biology of Coregonid fishes* (pp. 99–114). University of Manitoba Press.
- Maitland, P. S., & Lyle, A. A. (2013). *Ex situ* and *in situ* approaches, including assisted reproduction, for the conservation of native species of charr (*Salmonidae*) and whitefish (*Coregonidae*) in Scotland. *International Zoo Yearbook*, 47(1), 129–139. <https://doi.org/10.1111/j.1748-1090.2012.00192.x>
- Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population inference from RADseq data. *Molecular Biology and Evolution*, 35(5), 1284–1290. <https://doi.org/10.1093/molbev/msy023>
- Marques, D. A., Lucek, K., Sousa, V. C., Excoffier, L., & Seehausen, O. (2019). Admixture between old lineages facilitated contemporary ecological speciation in Lake Constance stickleback. *Nature Communications*, 10(1), 1–14. <https://doi.org/10.1038/s41467-019-12182-w>
- Mattingsdal, M., Jorde, P. E., Knutsen, H., Jentoft, S., Stenseth, N. C., Sodeland, M., Robalo, J. I., Hansen, M. M., André, C., & Blanco Gonzalez, E. (2020). Demographic history has shaped the strongly differentiated corkwing wrasse populations in Northern Europe. *Molecular Ecology*, 29(1), 160–171. <https://doi.org/10.1111/mec.15310>
- McKeown, N. J., Hynes, R. A., Duguid, R. A., Ferguson, A., & Prodöhl, P. A. (2010). Phylogeographic structure of brown trout *Salmo trutta* in Britain and Ireland: Glacial refugia, postglacial colonization and origins of sympatric populations. *Journal of Fish Biology*, 76(2), 319–347. <https://doi.org/10.1111/j.1095-8649.2009.02490.x>
- Meirmans, P. G., & Van Tienderenn, P. H. (2004). Genotype and genotype: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, 4(4), 792–794. <https://doi.org/10.1111/j.1471-8286.2004.00770.x>
- Momigliano, P., Florin, A.-B., & Merilä, J. (2021). Biases in demographic modeling affect our understanding of recent divergence. *Molecular Biology and Evolution*, 38(7), 2967–2985. <https://doi.org/10.1093/MOLBEV/MSAB047>
- Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, 18(3), 375–402. <https://doi.org/10.1111/j.1365-294X.2008.03946.x>
- Oppo, D. W., Keigwin, L. D., McManus, J. F., & Cullen, J. L. (2001). Persistent suborbital climate variability in marine isotope stage 5 and termination II. *Paleoceanography*, 16(3), 280–292. <https://doi.org/10.1029/2000PA000527>
- Ortiz, E. M. (2019). *vcf2phyloip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis*. <https://doi.org/10.5281/ZENODO.2540861>
- Ostberg, C. O., Pavlov, S. D., & Hauser, L. (2009). Evolutionary relationships among sympatric life history forms of dolly varden inhabiting the Landlocked Kronotsky Lake, Kamchatka, and a neighboring anadromous population. *Transactions of the American Fisheries Society*, 138(1), 1–14. <https://doi.org/10.1577/t08-016.1>
- Østbye, K., Bernatchez, L., Næsje, 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. *Molecular Ecology*, 14(14), 4371–4387. <https://doi.org/10.1111/j.1365-294X.2005.02737.x>
- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One*, 7(5), e37135. <https://doi.org/10.1371/journal.pone.0037135>
- Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data.

- PLoS Genetics, 8(11), e1002967. <https://doi.org/10.1371/journal.pgen.1002967>
- Pinson, L. J. W., Vardy, M. E., Dix, J. K., Henstock, T. J., Bull, J. M., & Maclachlan, S. E. (2013). Deglacial history of glacial lake Windermere, UK: Implications for the central British and Irish Ice Sheet. *Journal of Quaternary Science*, 28(1), 83–94. <https://doi.org/10.1002/jqs.2595>
- Præbel, K., Knudsen, R., Siwertsson, A., Karhunen, M., Kahilainen, K. K., Ovaskainen, O., Ostbye, K., Peruzzi, S., Fevolden, S.-E., & Amundsen, P.-A. (2013). Ecological speciation in postglacial European whitefish: Rapid adaptive radiations into the littoral, pelagic, and profundal lake habitats. *Ecology and Evolution*, 3(15), 4970–4986. <https://doi.org/10.1002/ece3.867>
- Quilodr n, C. S., Ruegg, K., Sendell-Price, A. T., Anderson, E. C., Coulson, T., & Clegg, S. M. (2020). The multiple population genetic and demographic routes to islands of genomic divergence. *Methods in Ecology and Evolution*, 11(1), 6–21. <https://doi.org/10.1111/2041-210X.13324>
- R Core Team. (2019). *The R project for statistical computing*. <https://www.r-project.org/>
- Rasmussen, S. O., Bigler, M., Blockley, S. P., Blunier, T., Buchardt, S. L., Clausen, H. B., Cvijanovic, I., Dahl-Jensen, D., Johnsen, S. J., Fischer, H., Gkinis, V., Guillevic, M., Hoek, W. Z., Lowe, J. J., Pedro, J. B., Popp, T., Seierstad, I. K., Steffensen, J. P., Svensson, A. M., ... Winstrup, M. (2014). A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: Refining and extending the INTIMATE event stratigraphy. *Quaternary Science Reviews*, 106, 14–28. <https://doi.org/10.1016/j.quascirev.2014.09.007>
- Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlovi c, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the genomic landscape of speciation: A road map for finding barriers to gene flow. *Journal of Evolutionary Biology*, 30(8), 1450–1477. <https://doi.org/10.1111/jeb.13047>
- Recknagel, H., Jacobs, A., Herzyk, P., & Elmer, K. R. (2015). Double-digest RAD sequencing using Ion Proton semiconductor platform (ddRADseq-ion) with nonmodel organisms. *Molecular Ecology Resources*, 15(6), 1316–1329. <https://doi.org/10.1111/1755-0998.12406>
- Rochette, N. C., Rivera-Col n, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, 28(21), 4737–4754. <https://doi.org/10.1111/mec.15253>
- Rougemont, Q., & Bernatchez, L. (2018). The demographic history of Atlantic salmon (*Salmo salar*) across its distribution range reconstructed from approximate Bayesian computations*. *Evolution*, 72(6), 1261–1277. <https://doi.org/10.1111/evo.13486>
- Rougeux, C., Bernatchez, L., & Gagnaire, P. A. (2017). Modeling the multiple facets of speciation-with-gene-flow toward inferring the divergence history of lake whitefish species pairs (*Coregonus clupeaformis*). *Genome Biology and Evolution*, 9(8), 2057–2074. <https://doi.org/10.1093/gbe/evx150>
- Rougeux, C., Gagnaire, P., & Bernatchez, L. (2019). Model-based demographic inference of introgression history in European whitefish species pairs. *Journal of Evolutionary Biology*, 32(8), 806–817. <https://doi.org/10.1111/jeb.13482>
- Rougeux, C., Gagnaire, P., Pr ebel, K., Seehausen, O., & Bernatchez, L. (2019). Polygenic selection drives the evolution of convergent transcriptomic landscapes across continents within a Nearctic sister species complex. *Molecular Ecology*, 28(19), 4388–4403. <https://doi.org/10.1111/mec.15226>
- Schenekar, T., Lerceteau-K hler, E., & Weiss, S. (2014). Fine-scale phylogeographic contact zone in Austrian brown trout *Salmo trutta* reveals multiple waves of post-glacial colonization and a pre-dominance of natural versus anthropogenic admixture. *Conservation Genetics*, 15(3), 561–572. <https://doi.org/10.1007/s10592-013-0561-0>
- Schmitt, T. (2007). Molecular biogeography of Europe: Pleistocene cycles and postglacial trends. *Frontiers in Zoology*, 4(1), 11. <https://doi.org/10.1186/1742-9994-4-11>
- Schroeter, J. C., Maloy, A. P., Rees, C. B., & Bartron, M. L. (2019). Fish mitochondrial genome sequencing: Expanding genetic resources to support species detection and biodiversity monitoring using environmental DNA. *Conservation Genetics Resources*, 1–14. <https://doi.org/10.1007/s12686-019-01111-0>
- Seehausen, O., Butlin, R. K., Keller, I., Wagner, C. E., Boughman, J. W., Hohenlohe, P. A., Peichel, C. L., Saetre, G.-P., Bank, C., Br nnstr m,  ., Brelsford, A., Clarkson, C. S., Eroukhanoff, F., Feder, J. L., Fischer, M. C., Foote, A. D., Franchini, P., Jiggins, C. D., Jones, F. C., ... Widmer, A. (2014). Genomics and the origin of species. *Nature Reviews Genetics*, 15(3), 176–192. <https://doi.org/10.1038/nrg3644>
- Stewart, J. R., & Lister, A. M. (2001). Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution*, 16(11), 608–613. [https://doi.org/10.1016/S0169-5347\(01\)02338-2](https://doi.org/10.1016/S0169-5347(01)02338-2)
- Symonov , R., Majt nov , Z., Sember, A., Staaks, G. B. O., Bohlen, J., Freyhof, J., R bov , M., & R b, P. (2013). Genome differentiation in a species pair of coregonine fishes: An extremely rapid speciation driven by stress-activated retrotransposons mediating extensive ribosomal DNA duplications. *BMC Evolutionary Biology*, 13(1), 42. <https://doi.org/10.1186/1471-2148-13-42>
- Thomas, R., Fletcher, J. M., James, J. B., & Winfield, I. J. (2013). Assessment and conservation of gwyniad (*Coregonus lavaretus* (L.)) in Llyn Tegid, U.K.: Persistence in the face of eutrophication, water level fluctuations and ruffe (*Gymnocephalus cernuus* (L.)) introduction. *Advances in Limnology*, 64, 363–376. <https://doi.org/10.1127/1612-166X/2013/0064-0016>
- Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21(16), <https://doi.org/10.1111/j.1365-294X.2012.05664.x>
- Volckaert, F. A. M., H nfling, B., Hellemans, B., & Carvalho, G. R. (2002). Timing of the population dynamics of bullhead *Cottus gobio* (Teleostei: Cottidae) during the Pleistocene. *Journal of Evolutionary Biology*, 15(6), 930–944. <https://doi.org/10.1046/j.1420-9101.2002.00469.x>
- Wang, L., Wan, Z. Y., Lim, H. S., & Yue, G. H. (2016). Genetic variability, local selection and demographic history: Genomic evidence of evolving towards allopatric speciation in Asian seabass. *Molecular Ecology*, 25(15), 3605–3621. <https://doi.org/10.1111/mec.13714>
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358–1370.
- Westergaard, K. B., Zemp, N., Bruederle, L. P., Sten ien, H. K., Widmer, A., & Fior, S. (2019). Population genomic evidence for plant glacial survival in Scandinavia. *Molecular Ecology*, 28(4), 818–832. <https://doi.org/10.1111/mec.14994>
- Wheeler, A. (1977). The origin and distribution of the freshwater fishes of the British Isles. *Journal of Biogeography*, 4(1), 1. <https://doi.org/10.2307/3038124>
- Wilson, A. J., Gislason, D., Sk lason, S., Snorrason, S. S., Adams, C. E., Alexander, G., Danzmann, R. G., & Ferguson, M. M. (2004). Population genetic structure of Arctic Charr, *Salvelinus alpinus* from northwest Europe on large and small spatial scales. *Molecular Ecology*, 13(5), 1129–1142. <https://doi.org/10.1111/j.1365-294X.2004.02149.x>
- Winfield, I. J., Bean, C. W., Gorst, J., Gowans, A. R. D., Robinson, M., & Thomas, R. (2013). Assessment and conservation of whitefish (*Coregonus lavaretus*) in the U.K. *Advances in Limnology*, 64, 305–321.
- Winkler, K. A., Pamminer-Lahnsteiner, B., Wanzenb ck, J., & Weiss, S. (2011). Hybridization and restricted gene flow between native and introduced stocks of Alpine whitefish (*Coregonus* sp.) across multiple environments. *Molecular Ecology*, 20(3), 456–472. <https://doi.org/10.1111/j.1365-294X.2010.04961.x>

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., & Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*, *28*(24), 3326–3328. <https://doi.org/10.1093/bioinformatics/bts606>

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Crotti, M., Bean, C. W., Gowans, A. R. D., Winfield, I. J., Butowska, M., Wanzenböck, J., Bondarenko, G., Præbel, K., Adams, C. E., & Elmer, K. R. (2021). Complex and divergent histories gave rise to genome-wide divergence patterns amongst European whitefish (*Coregonus lavaretus*). *Journal of Evolutionary Biology*, *34*, 1954–1969. <https://doi.org/10.1111/jeb.13948>