

# Phenotype-environment association of the oxygen transport system in trimorphic European whitefish (*Coregonus lavaretus*) populations

Running Title: Adaptive divergence of oxygen transport system

Melissa L. Evans<sup>1,2,\*</sup>, Kim Præbel<sup>3</sup>, Stefano Peruzzi<sup>3</sup>, Per-Arne Amundsen<sup>3</sup>, and Louis Bernatchez<sup>1</sup>

<sup>1</sup>Institut de Biologie Intégrative et des Systèmes, Pavillon Charles-Eugène-Marchand, 1030 Avenue de la Médecine, Université Laval, Québec, Québec, G1V 0A6, Canada

<sup>2</sup>Present address: Coastal Oregon Marine Experiment Station, Hatfield Marine Science Center, Oregon State University, 2030 SE Marine Science Dr., Newport, OR, 97365, USA

<sup>3</sup>Department of Arctic and Marine Biology, University of Tromsø, N9037 Tromsø, Norway

\*Corresponding author: [melissa.lea.evans@gmail.com](mailto:melissa.lea.evans@gmail.com)

Keywords: phenotype-environment, European whitefish, ecological speciation, oxygen transport, hemoglobin, gene transcription

Word count: 6378; Table count: 7; Figure count: 3

April 22 2014

Document Format: Word

Data archiving: data available from the Dryad Digital Repository: doi:10.5061/dryad.351h3

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/evo.12442.

This article is protected by copyright. All rights reserved.

## ABSTRACT

Replicated adaptive radiation events, typified by phenotypic divergence across resource axes, provide important insight into the eco-evolutionary dynamics that lead to the formation of new species. Here we show that in trimorphic adaptive radiations of European whitefish (*Coregonus lavaretus*) divergence of the physiological oxygen transport system has occurred across the pelagic/littoral (shallow) – profundal (deep) resource axis, and at multiple biological scales. Profundal whitefish exhibited significantly larger red blood cells, a greater proportion of cathodic hemoglobin protein components, and higher hemoglobin transcript abundance in kidney compared to littoral and pelagic morphs. Hemoglobin transcript abundance in brain and gill, but not kidney, and anodic hemoglobin protein component diversity in blood were also linked to variation at an intronic single nucleotide polymorphism (SNP). As the whitefish morphs differ in population genetic structure at this SNP, hemoglobin transcript and protein divergence between profundal and pelagic/littoral morphs is likely being driven by genetic divergence. Our findings, along with our previous work on lake whitefish, highlight the importance of the oxygen transport system to the postglacial colonization of novel lacustrine environments by whitefish throughout the northern hemisphere.

## INTRODUCTION

North temperate populations of freshwater fishes are important systems for integrative studies of the genomic, phenotypic, and ecological factors underlying sympatric speciation processes. These populations are evolutionarily young, having colonized their current distribution following the recession of the last major glaciation (<15,000 ybp) and show widespread evidence of ongoing postglacial ecologically-driven divergence (Bernatchez et al. 2010; Elmer and Meyer 2011). Indeed, sticklebacks (*Gasterosteus* spp.), several species of salmonids (*Oncorhynchus* spp., *Salvelinus* spp., *Coregonus* spp.), and osmerids (*Osmerus* spp.) have evolved sympatric “species” or “eco-morphs” that exhibit relatively well-understood associations with diverging trophic positions and lacustrine environments (Taylor 1999). Furthermore, as evolutionary models, our understanding of these systems at the genomic level has advanced relatively rapidly, allowing for comparative examinations of the traits that have facilitated speciation in the absence of geographical barriers (e.g. Barrett et al. 2008; Renaut et al. 2010, 2011; Bernatchez et al. 2010; DeFaveri et al. 2011).

Considering their evolutionary young age, European whitefish (*Coregonus lavaretus*) from the Fennoscandian region of northern Europe exhibit incredibly polymorphic populations, consisting of one, two, or three sympatric eco-morphs across lakes (Amundsen et al. 2004; Kahilainen and Østbye 2006; Siwertsson et al. 2010). The evolution of polymorphic flocks has been mediated through the postglacial colonization of lakes by a monophyletic clade of whitefish and subsequent adaptation to divergent environments within lakes (Østbye et al. 2005; Østbye et al. 2006; Præbel et al. 2013). Trimorphic flocks, composed of large sparsely rakered (LSR), densely rakered (DR), and small sparsely rakered (SSR) whitefish, are genetically distinct and respectively occupy the littoral, pelagic, and deep profundal lacustrine zones (Præbel et al. 2013). However, outside of gill raker number

and size, which is associated with morph specialization to differing prey communities, relatively little is known about the traits that may have facilitated postglacial colonization of divergent lacustrine environments by whitefish.

The application of genomics approaches to investigations of speciation is refining our understanding of the mechanisms that underlie evolutionary diversification (Abzhanov et al. 2006; Bernatchez et al. 2010; Jones et al. 2012). This work has pointed to genomic regions associated with physiological processes as statistically overrepresented “outliers” between recently diverged freshwater fish species (e.g. Derome and Bernatchez 2006; Salzburger et al. 2008; St-Cyr et al. 2008; Elmer et al. 2010; Renaut et al. 2010, 2011; DeFaveri et al. 2011). While some of this overrepresentation may relate to ascertainment bias, ecophysiological studies of speciation events from the “pre-genomic”-era have similarly highlighted physiological adaptation as a key process associated with postglacial adaptive divergence in freshwater environments (Giles 1983; Bernatchez and Dodson 1985; Trudel et al. 2001). Here, we investigate the role of a candidate physiological pathway, the oxygen transport system, in the adaptive divergence of trimorphic populations of European whitefish. Work on replicated adaptive radiation events of the European whitefish’s North American congener, the lake whitefish (*C. clupeaformis*), has demonstrated the parallel divergence of a suite of traits associated with oxygen uptake and delivery between pelagic “dwarf” and benthic “normal” species (Derome and Bernatchez 2006; Renaut et al. 2010, 2011; Evans and Bernatchez 2012; Evans et al. 2012, 2013), suggesting that adaptation to alternate oxygen environments is central to whitefish postglacial diversification. Furthermore, recent work by Vonlanthen et al. (2012) has shown that oxygen regimes within lakes remain important contemporary drivers of whitefish biodiversity, with lake eutrophication (oxygen depletion) leading to limnetic zone collapse and the loss of whitefish species through demographic decline and homogenization of previously ecologically structured populations.

We investigated the oxygen transport system of European whitefish in two lakes harbouring trimorphic populations, Langfjordvatn and Skrukkebukta, from the Pasvik region of northern Fennoscandia. We examined littoral (LSR), pelagic (DR), and profundal (SSR) morphs for red blood cell (RBC) morphological variation, including cell and nucleus size, blood hemoglobin protein component variation and composition, and variation in hemoglobin gene transcript abundance in the kidney, brain, and gill. Hemoglobin proteins are in continuous contact with external and internal environments, as the primary molecules responsible for transporting oxygen from fish gills to tissues for use in cellular respiration (Weber 1990; Nikinmaa 1997), and many fish species bear diverse hemoglobin protein types adapted to oxygen loading and delivery under fluctuating environmental conditions (Sick 1961; Verde et al. 2006, 2012; Andersen et al. 2009; Wetten et al. 2010; Star et al. 2011). Thus, studies of hemoglobin and hemoglobin carriers – i.e. red blood cells – can provide important insight into the ecological and evolutionary interactions that have shaped contemporary patterns of species diversity.

In the North American lake whitefish, normals show larger RBC cytoplasmic areas and body size standardized gills compared to dwarfs, which may enhance oxygen uptake in the seasonally oxygen depleted benthic lacustrine zones they inhabit (Evans et al. 2012, 2013). In contrast, the high metabolic oxygen demands associated with the more active life history of dwarf compared to normal whitefish may be an important driver of hemoglobin gene upregulation in dwarf brains and differences in hemoglobin protein component composition found between the two morphs (Evans et al. 2012). The littoral LSR and pelagic DR morphs of European whitefish are considered respective phylogenetic replicates of dwarf and normal lake whitefish, and thus, we predict that similar patterns of divergence may have occurred in European whitefish adaptive radiations. Indeed, Jeukens et al. (2009) reported parallel

patterns of divergence in gene transcript abundance at several candidate genes between dwarf/normal lake whitefish and DR/LSR European whitefish sympatric pairs.

We also examined patterns of population genetic variation among the whitefish morphs at a previously characterized single nucleotide polymorphism (SNP), SNP 1544-242, located in intron 1 of the alpha hemoglobin chain, for evidence of selection operating on genomic regions associated with hemoglobin across the three whitefish morphs (Evans et al. 2012). While SNP 1544-242 is located in a putative non-coding region of the genome, we were interested in potential relationships between SNP genotype and hemoglobin gene transcript abundance and protein-level variation, as this polymorphism could be linked to variable regulatory or coding regions for hemoglobin. Previous studies of the whitefish transcriptome have not identified any diverging polymorphisms in coding regions for hemoglobin between sympatric whitefish (Renaut et al. 2010).

## METHODS

### *Whitefish collection*

European whitefish were sampled from Langfjordvatn in September 2009 and 2010 and Skrukkebukta in September 2010. Both lakes are found in the Pasvik region of northern Norway (Fig.1) and have been separated for approximately 9,000 years (Kujansuu et al. 1998). Whitefish were collected from the three principal lake habitats: the littoral zone (< 6m; > 1% of light at surface), the profundal zone (> 25m; < 1% of light at surface) and the pelagic zone (0-6m). Sampling was performed with gillnets consisting of eight 5m sections composed of variable mesh sizes; 10, 12.5, 15, 18.5, 22, 26, 35 and 45mm, knot to knot. Bottom gillnets were used in the littoral and profundal zones whereas the pelagic habitat was sampled using 6m deep floating nets. At each sampling site we collected data on temperature (°C) and oxygen saturation ( $pO_2$ ) at 5m depth intervals. Temperature was determined using a

Sea-bird SBE25 CTD ( $\pm 0.001^\circ\text{C}$ ). Relative oxygen saturation ( $\pm 1\%$   $p\text{O}_2$ ) was measured in water samples collected with a 5L Niskin water bottle using a handheld oxygen meter (OxyGuard® International A/S, Birkerød, Denmark). The measurements of relative oxygen saturation were adjusted for temperature dependency *in situ* using an Acorn® Temp 5 thermistor thermometer (Oakton Instruments, Vernon Hills, IL, USA). For Langfjordvatn we collected these data for the deep north basin, the mid-basin, and the shallow south basin of the lake. For Skrukkebukta, temperature and oxygen saturation data were collected for the deep north basin and the shallow south basin. The environmental data for each of the sampling sites are reported in the Supporting Information Fig. S1.

Whitefish were identified in the field to one of the three morphs: SSR, LSR, or DR. Methods for identifying the three morphs are detailed in Amundsen et al. (2004), Kahilainen and Østbye (2006), and Siwertsson et al. (2010).

We collected the whole brain from each individual and ~30mg of tissue from the kidney and gills of freshly killed fish. The tissues were immediately preserved in RNAlater (Life Technologies, Carlsbad, CA) and later stored at  $-80^\circ\text{C}$ . We also collected between 100-1000 $\mu\text{l}$  of blood by puncturing the caudal vein with either a 21- or 23-gauge needle and syringe containing 2-20 $\mu\text{l}$  of 5000IU/ml heparin. Following collection, the syringes were gently agitated to ensure the blood was completely heparinized. The blood samples were then transferred to a 1.5ml eppendorf tube and stored at  $4^\circ\text{C}$ . A summary of sample sizes obtained for each of the characteristics of the oxygen transport system examined in the study is shown in Table 1. Notably, our sample sizes of DR whitefish from Skrukkebukta are limited due to population declines of this morph related to vendace introductions and competitive exclusion (Sandlund et al. 2013).

### *Red blood cell morphology*

Blood smears were examined for RBC morphological variation among the whitefish morphs. Smears were made from each of the freshly collected blood samples, air-dried, fixed with 95% methanol in the field, and later Giemsa stained in the laboratory. We examined blood cells at 200× magnification under a light microscope, with which a digital photograph was taken. The photograph was then imported into ImageJ (<http://rsbweb.nih.gov/ij/index.html>) where we measured fifty cells per individual for nucleus size (dry elliptical area in  $\mu\text{m}^2$ ) and total cell area.

#### *Hemoglobin protein characterization*

Hemoglobin proteins are heterotetramers composed of two alpha and two beta chains, and in fishes, there are often multiple hemoglobin protein types found within and among individuals (Fyhn and Withler 1991; Quinn et al. 2010; Verde et al. 2012). We examined hemoglobin protein component variation in 88 whitefish from Langfjordvatn and 27 whitefish from Skrukkebukta (Table 1). As in Evans et al. (2012), hemolysate was isolated from RBCs and isoelectric focus (IEF) electrophoresis was used to partition hemoglobin into its protein components. IEF uses a pH gradient across a gel to separate proteins based on overall charge of the molecule. A hemoglobin protein component's isoelectric point, pI, is defined as the position in the gel where the molecule becomes neutrally charged (Sick 1961; Husebø et al. 2004). The hemoglobin concentration of each hemolysate sample was determined spectrophotometrically at 540nm using Drabkin's reagent (D 5941, Sigma, Saint Louis, USA) and a bovine hemoglobin standard (H2500-1G, Sigma) according to the manufacturer's instructions. Novex® IEF gels (pH=3-10, Invitrogen, Life Technologies Co., USA) were used to partition hemoglobin components. Gels were loaded with approximately



50µg of hemoglobin per sample, along with a protein marker (Novex® IEF Marker; Invitrogen), and run for 2.5h at 4°C. The gels were then fixed in 12% trichloroacetic acid containing 3.5% sulfosalicylic acid and stained using Novex® SimplyBlue™ SafeStain (Coomassie® G250) in accordance with the manufacturer's instructions (Invitrogen), and then destained for 7-10h in deionized water. The pI of each protein component was scored with reference to the protein marker. Only gels with regression coefficients ( $R^2$ ) > 0.97, when plotting the pI of the standard proteins against the corresponding migration distance, were included in the study and all gels were scored in duplicate.

This is the first study to examine hemoglobin protein component diversity in European whitefish. However, as the North American lake whitefish shows multiple “anodic” and “cathodic” hemoglobin protein components (Evans et al. 2012), we also identified any shared hemoglobin protein components and banding patterns between the European and previously examined lake whitefish samples as part of this study.

#### *RNA extraction and cDNA synthesis*

Total RNA was extracted from kidney, gill, and whole brain using the Ambion PureLink RNA Mini Kit and according to the manufacturer's instructions (Applied Biosystems, Carlsbad, California). In fishes, kidneys are a primary organ involved in the production of hemoglobin and mature RBCs (Fange 1994; Abdel-Aziz et al. 2010) and gills represent the site of oxygen loading from the external environment, so we predict that differences in hemoglobin gene transcript abundance in this tissue will reflect functional responses to the environmental conditions encountered in each of the lacustrine zones by the whitefish morphs. We examined hemoglobin gene transcript abundance in the brain because previous microarray and qPCR-based studies have shown that the lake whitefish transcribe

hemoglobin genes in the brain, and differentially among sympatric morphs (Whiteley et al. 2008; Evans et al. 2012; Filteau et al. 2013).

RNA extracts were treated with Ambion Superase-In, an RNAase inhibitor (Applied Biosystems). We evaluated RNA purity and concentration by measuring the 260/230nm and 260/280nm absorbance ratios on the Nanodrop 2000 Spectrophotometer (Nanodrop Technologies, Wilmington, Delaware) and RNA integrity was confirmed using the Experion Automated Electrophoresis system (Bio-Rad, Mississauga, Ontario, Canada). We synthesized cDNA from RNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Reverse transcription reactions were performed in 20µl volumes containing 500ng RNA and using random primers.

#### *Hemoglobin gene transcript abundance assays*

We used TaqMan (Applied Biosystems) qPCR assays designed for the evaluation of alpha and beta chain hemoglobin gene transcript abundance in the lake whitefish (Evans et al. 2012). The TaqMan Minor Groove Binder (MGB) probes and primers align with “alpha” and “beta” hemoglobin ESTs annotated as Bohr-effect molecules in the annotated Atlantic salmon (*Salmo salar*) hemoglobin genome (GenBank Accession X97285.1; McMorrow et al. 1996; Quinn et al. 2010). For the alpha hemoglobin assay, we used forward primer 5'-TGGACCCACCAACTTCAA-3', reverse primer 5'-GCGGCAACGACCACAATC-3', and probe 5'-ATCCTGGCTCACAACC-3', sequences that align to portions of exon two (forward primer), exon three (reverse primer), and flank exons two and three (probe) of the Atlantic salmon alpha hemoglobin chain. For the beta hemoglobin assay, we used forward primer 5'-GTGCAGTTTCTCCGAGTGCAT-3', reverse primer 5'-AGAACCTGGATGACATCAAAAACA-3', and probe 5'-ACACTCAGTGCAGTATAG-3', which align with exon two of the Atlantic salmon beta hemoglobin chain. We evaluated the

efficiency of each assay on the European whitefish cDNA using a validation experiment and found that the alpha and beta assays were 96% and 99% efficient, respectively.

### *qPCR*

We utilized the comparative  $C_T$  method ( $\Delta\Delta C_T$  method; Livak and Schmittgen 2001) to quantify hemoglobin cDNA (i.e. transcript abundance) in the kidney, gill, and brain. This method calculates the relative quantity (RQ) of cDNA in each of the “test” samples compared to a common internal reference sample (“calibrator”). RQ values represent fold changes in transcript abundance relative to the calibrator, which has a value of one. To quantify transcript abundance in each of the tissue types (kidney, brain, or gill), we used a calibrator sample derived from the same type of tissue. Specifically, an RNA sample collected from each of the kidney, gill, and brain of a Langfjordvatn littoral (LSR) whitefish was used as a calibrator for each of the kidney, gill, and brain assays, respectively. The use of different calibrators across assays meant that we could not compare hemoglobin transcript abundance among tissues. However, higher transcript abundance was previously demonstrated in kidney compared to gill and brain in the lake whitefish (Evans et al. 2012).

We amplified the hemoglobin gene targets from cDNA in triplicate using the ABI 7500 Real-Time PCR System and TaqMan Universal PCR Master Mix (Applied Biosystems). Real-time quantitative PCR reactions (hereafter qPCR) were performed in 10 $\mu$ l reactions and using the default thermocycler setting for Fast qPCR TaqMan assays. ABI’s Human Euk 18S rRNA gene assay was used as an endogenous control to normalize the total quantity of cDNA loaded into the PCR reaction wells. The Human Euk 18S rRNA is an appropriate target for the normalization of transcription assays in whitefish (Jeukens et al. 2009) and other salmonids (e.g. Olsvik et al. 2005). The efficiencies of our two hemoglobin assays were similar to the 100% efficiency of the Human Euk 18s rRNA endogenous control assay, indicating that the comparative  $C_T$  method is acceptable for use in the evaluation of relative

hemoglobin gene transcript quantification. Transcript abundance for the alpha and beta hemoglobin genes were highly significantly correlated in each of the tissue types (Spearman's rank correlation; kidney:  $\rho=0.812$ ; brain:  $\rho=0.646$ ; gill:  $\rho=0.647$ ; all  $P < 0.001$ ).

### *KASPAR SNP genotyping*

We examined population genetic variation in the Skrukkebukta and Langfjordvatn whitefish morphs at a putative regulatory G/A SNP located in intron 1 of the alpha hemoglobin gene (SNP 1544-242). This SNP was previously identified in lake whitefish and may be associated with alpha hemoglobin gene transcript abundance (Evans et al. 2012). In total, 125 European whitefish, including all whitefish examined for gene transcript abundance, were genotyped at SNP 1544-242 using a KASPar SNP genotyping assay (KBiosciences, Hoddesdon, UK). KASPar assays are composed of fluorescently-labelled SNP-specific primers that enable the determination of homozygous or heterozygous genotypes based on the differential amplification of alleles associated with each of the uniquely tagged primers. The assay was run on the Applied Biosystems 7500 Real-time PCR system under the genotyping application. KASPar genotyping reactions were run in 8 $\mu$ l volumes consisting of 4 $\mu$ l DNA, 4 $\mu$ l 2  $\times$  reaction mix, 0.11  $\mu$ l Assay Mix (KBiosciences aliquot barcode 1035902972), and 0.064  $\mu$ l 50mM  $\text{MgCl}_2$ , as outlined in the manufacturer's instructions (KBiosciences). Observed and expected heterozygosities for the LSR, DR, and SSR morphs in each lake were calculated in Genepop v. 3.4 (Raymond and Rousset 1995). We also tested the observed allele frequencies for heterozygote excess, an indicator of gene paralog amplification, in Genepop.

### *Statistical analyses*

The RBC measurements were approximately normally distributed among individuals (Shapiro-Wilk test:  $P > 0.05$ ). The number of hemoglobin protein components found in

whitefish was not significantly normally distributed (Shapiro-Wilk test:  $P < 0.05$ ) but conformed most closely to a normal distribution; thus, we did not transform these data. Kolmogorov-Smirnov tests revealed that the alpha and beta transcript levels were not normally distributed among the whitefish examined, but conformed to a Log-normal distribution in all tissue types (Kolmogorov's D test:  $P > 0.15$ ); thus, we used Log + 1 transformed transcript abundance data as dependent variables in all models.

Two-way ANOVA was used to partition variation in RBC nucleus area, total cell area, nucleus:total cell area, hemoglobin protein component diversity, and the proportion of cathodic versus anodic hemoglobin protein components found in individuals to lake and morph effects and their interaction (lake×morph). We used a similar model to examine the contribution of lake, morph, and lake×morph effects to variation in alpha and beta hemoglobin transcript abundance (Log + 1 transformed) in kidney, brain, and gill. Posthoc Tukey's tests of honestly significant differences (*HSD*) were used to identify significant differences in mean hemoglobin protein component diversity and transcript abundance among the whitefish groups.

Chi-square tests were used to examine differences in SNP genotype frequencies among the whitefish morphs. Nested ANOVA was used to examine the relationship between SNP 1544-242 genotype (AA, AG, GG) and alpha hemoglobin gene transcript abundance (Log + 1 transformed) in each of the tissues. In these models, morph was nested within SNP genotype, as not all of the SNP genotypes were found in all morphs. Lake was also included in the model to control for population-dependent effects, independent of SNP genotype, on hemoglobin transcript abundance. A Posthoc Tukey's *HSD* test was used to identify significant differences in hemoglobin transcript abundance among whitefish bearing each of the SNP genotypes.

We tested for relationships between SNP genotype and the presence/absence of hemoglobin protein components using Chi-square and Fisher's exact tests. A generalized linear model (GLM) with a Poisson distribution (log link function) was used to further explore the influence SNP genotype on number of anodic hemoglobin protein components found within each of the whitefish morphs. For these tests, all morphs from both lakes were pooled due to sample size limitations.

All statistical analyses were run in JMP v.10 (SAS Institute Inc., Cary, NC). Throughout, means are reported  $\pm 1$  standard deviation (SD), except where indicated, and we used a critical  $\alpha = 0.05$  for all statistical tests. Bonferroni correction was applied to our critical  $\alpha$  ( $P = 0.05/6$ ) when examining relationships between SNP genotype and protein components so as to correct for multiple testing.

## RESULTS

### *Red blood cell morphology*

Both morph and lake were significant predictors of total RBC area (Tables 2-3). Langfjordvatn whitefish exhibited larger cells than their counterparts from Skrukkebukta, and the profundal SSR whitefish from both lakes exhibited larger cells than the littoral LSR and pelagic DR whitefish (Tukey's *HSD*:  $P < 0.05$ ; Table 2-3). Neither morph nor lake was a significant predictor of variation in RBC nucleus area (Table 2-3), though their interaction was, with Skrukkebukta LSR whitefish showing significantly larger nuclei than Skrukkebukta SSR and Langfjordvatn LSR whitefish (Tukey's *HSD*:  $P < 0.05$ ). The ratio of nucleus area to total cell area was larger in Skrukkebukta LSR and DR whitefish compared to SSR whitefish from the same lake and SSR and LSR whitefish from Langfjordvatn (Tukey's *HSD*:  $P < 0.05$ ; Table 2-3).

### *Hemoglobin protein component diversity*

The whitefish morphs exhibited a collective four major anodic (pI=6.1–6.7) and four major cathodic (pI=7.2–8.3) hemoglobin protein components, which were observed across 13 different IEF gel banding patterns (Fig. S2, Table S1-2). An average of  $5.3 \pm 1.1$  (range=3-8) hemoglobin protein components comprised of  $2.6 \pm 0.4$  (range=1-4) cathodic components and  $2.7 \pm 0.8$  (range=1-4) anodic components were found in individual whitefish. In both lakes, the total number of hemoglobin components exhibited by the whitefish morphs differed significantly, with LSR whitefish showing more protein components than DR and SSR whitefish (Tukey's *HSD*:  $P < 0.05$ ; Table 1-2; Table S2). However, SSR whitefish showed a significantly greater proportion of cathodic hemoglobin components compared to LSR whitefish (Tukey's *HSD*:  $P < 0.05$ ; Table S2).

### *Gene transcript abundance variation among morphs and between lakes*

Alpha hemoglobin transcript abundance in kidney varied significantly among the whitefish morphs and between the two lakes (Table 4). Specifically, the profundal SSR morph exhibited greater alpha hemoglobin transcript abundance compared to DR and LSR whitefish (Fig. 2a, Tukey's *HSD*:  $P < 0.05$ ) and the Langfjordvatn whitefish generally showed greater alpha hemoglobin transcript abundance compared to whitefish from Skrukkebukta (Fig. 2a,  $P < 0.05$ ). In contrast, we did not observe a difference in beta hemoglobin transcript abundance in the kidney tissues between the two lakes or among the morphs.

In the brain, we observed greater alpha and beta hemoglobin gene transcript abundance in Langfjordvatn whitefish compared to Skrukkebukta whitefish (Fig. 2b, Student's *t*-test:  $P < 0.05$ ), but no difference among the whitefish morphs at either gene. Similarly in gill, we observed significantly greater alpha hemoglobin gene transcript abundance in Langfjordvatn

compared to Skrukkebukta whitefish (Table 4, Fig. 2c, Student's *t*-test:  $P < 0.05$ ) but no difference among the morphs in hemoglobin transcript abundance.

#### *SNP population genetics and alpha hemoglobin transcript abundance*

There was no evidence of heterozygote excess at SNP 1544-242 in any of the morphs in either lake, which supports the amplification of a single gene as opposed to gene paralogs by our SNP assay (Table 5). For LSR whitefish from both lakes, more and fewer individuals than expected exhibited the GG and AA genotypes, respectively, and we observed a significant heterozygote (AG) deficit in the LSR morph from Langfjordvatn (Table 5). Fewer GG genotypes and a trend towards a greater number of individuals bearing the AA and AG genotypes than expected were observed in SSR whitefish, particularly in Skrukkebukta. We could not compare genotype frequencies patterns in DR whitefish between the two lakes because our sample size was small for Skrukkebukta. However, for the Langfjordvatn DR morph, genotype trends were similar to those seen in the SSR whitefish; i.e. more AG and AA genotypes, and fewer GG genotypes than expected (Table 5).

While SNP genotype and alpha hemoglobin gene transcript abundance were not associated in kidney (Table 6, Fig. 3a), there was a significant effect of SNP genotype on transcript abundance in both the brain and gill (Table 6, Fig. 3b-c). In both tissues, individuals bearing the AA genotype, regardless of morph, exhibited significantly higher alpha hemoglobin gene transcript levels than whitefish bearing the AG or GG genotypes (Tukey's *HSD*:  $P < 0.05$ ; Fig. 3b-c).



### *SNP genotype and hemoglobin protein components*

SNP genotype was a significant predictor of the presence of anodic, but not cathodic, hemoglobin protein components (Table 7). Specifically, protein components A1 and A2 were more frequently observed in individuals bearing the GG genotype than expected. Indeed, essentially all individuals bearing the GG genotype showed the A1 and A2 components. For whitefish bearing the AA genotype, only 1 and 3 individuals exhibited A1 or A2 protein components, respectively, whereas the A4 component was observed in all individuals bearing the AA genotype and most individuals bearing the AG genotype (Table 7). In contrast, fewer whitefish of the GG genotype than expected exhibited the A4 component. Across all whitefish morphs, individuals with the AA genotype showed significantly fewer hemoglobin protein components compared to GG or AG individuals (Poisson GLM:  $\chi^2_8 = 84.3$ ,  $P < 0.001$ ; Fig. S3)

### DISCUSSION

The oxygen transport system of trimorphic European whitefish populations has diverged at multiple biological scales, and primarily between morphs inhabiting the profundal and littoral/pelagic lacustrine zones. Profundal SSR whitefish exhibited larger RBCs compared to the littoral LSR and pelagic DR morphs from Langfjordvatn and Skrukkebukta. The function of RBC size variation is not generally well-understood despite considerable variation in RBC size and morphology across fishes and other organisms (Lay and Baldwin 1999; Snyder and Sheafor 1999). For whitefish, it is possible that RBC size is influenced by differences in activity-energetic associated oxygen requirements of the morphs. Smaller RBCs provide a higher diffusible surface area, which increases gas exchange per unit cell volume when compared to larger cells, and in general, more active fish species exhibit smaller RBCs

(Wilhem et al. 1992; Lay and Baldwin 1999). Previous studies have shown that DR and, to a lesser extent, LSR whitefish face higher predation compared to SSR whitefish (Kahilainen and Lehtonen 2003) and the foraging tactics used by the whitefish morphs also vary (Kahilainen and Østbye 2006; Harrod et al. 2010). Thus, it is possible that behavioural adaptations to predator and prey communities have led to greater metabolic oxygen requirements in the LSR and DR morphs compared to the SSR morph. RBC size has also been shown to positively scale with body size in a fish species (Maciak et al. 2011), but scaling effects cannot explain the variation we observed in RBC size across the whitefish morphs. The DR and LSR whitefish, which are the most divergent in body size of the three morphs (see Østbye et al. 2006), exhibited similar cell sizes, whereas the SSR profundal morph, which is intermediate in size between DR and LSR whitefish, exhibited the largest RBCs. Overall, the parallel divergence of RBC size in two populations of European whitefish suggests that this component of the oxygen transport system has specialized postglacially to divergent physiological challenges across profundal – littoral/pelagic lacustrine environments. Further studies will be necessary to investigate the functional importance of such divergence.

Many fish species exhibit a diversity of hemoglobin proteins adapted to oxygen uptake and transport under fluctuating or variable environmental conditions (Verde et al. 2012). We found a total of four anodic and four cathodic hemoglobin protein components in Skrukkebukta and Langfjordvatn European whitefish, and a collective thirteen distinct hemoglobin protein banding patterns. The diversity of protein components observed here is comparable to that reported by Evans et al. (2012) in the North American lake whitefish. However, between the North American and European species, only five of the observed protein banding patterns are shared, and two of the cathodic components, C2 and C4, show

differing isoelectric points (Table S1) indicating that hemoglobin proteins have diverged between the North American and European species.

Divergence of hemoglobin protein components was also observed among the three European whitefish morphs. The LSR whitefish showed a greater diversity of protein components compared to DR and SSR whitefish in both the Skrukkebukta and Langfjordvatn populations. Lower hemoglobin protein component diversity in SSR and DR whitefish could be driven by founder effects following the colonization of the pelagic and profundal environments by ancestral LSR whitefish. Indeed, a reduction in genetic diversity at microsatellite loci has been observed in SSR morphs when compared to LSR morphs (Präbel et al. 2013). However, as we anticipate that founder effects should affect anodic and cathodic components equally, the maintenance of the full cathodic protein component repertoire in SSR whitefish implicates a possible role for balancing selection as ancestral whitefish colonized the profundal zone. Cathodic protein components may contribute to hemoglobin oxygen binding insensitivity in response to fluctuations in pH (non-Bohr effect hemoglobins; e.g. Powers 1972) or temperature (Brix et al. 2004) and at least in catostomids and trout, are characteristic of populations inhabiting environments that require active swimming behaviours (e.g. streams; Barra et al. 1973; Powers 1980). Studies of lake-dwelling salmonids have, as yet, been unable to pinpoint the ecological factors maintaining hemoglobin protein component diversity, but suggested drivers include low oxygen and thermal stress (Verde et al. 2012). Overall, our results suggest that founder effects, and possibly the adaptive retention of cathodic components in the SSR morph, explain patterns of protein component variation among the whitefish morphs.

Increased hemoglobin protein production can be utilized to meet oxygen uptake and transport demands under hypoxia or other environmental stressors such as exposure to increased temperature (Rutjes et al. 2007; Star et al. 2011). Among the European whitefish

morphs, profundal SSR whitefish showed higher alpha hemoglobin gene transcript abundance in kidney compared to littoral and pelagic whitefish, suggesting that hemoglobin gene upregulation may be involved in whitefish persistence under hypoxia. While increased hemoglobin gene transcription may reflect a strictly plastic response to hypoxia exposure (Kupittayanant and Kinchareon 2011), such a response may be essential to early speciation processes if they facilitate the colonization of, and persistence in, novel environments (Pavey et al. 2010). While the genomic architecture of the whitefish hemoglobin repertoire is currently unknown, a recent study of Atlantic salmon reported a diversity of globin genes (Quinn et al. 2010). It is therefore possible that regulatory differences across gene paralogs could lead to some of the variation observed in hemoglobin transcript abundance and protein component diversity among whitefish morphs. Future work that incorporates the use of RNAseq (Jeukens et al. 2010) and/or BAC libraries to target and define the genomic regions associated with hemoglobin (Jeukens et al. 2011) will be necessary to uncover the precise genetic and transcriptional differences driving hemoglobin transcript abundance and protein divergence in European whitefish.

Our work has uncovered a potential “eQTL” (expression quantitative trait locus) associated with levels of alpha hemoglobin gene transcript abundance in the brain and gill of European whitefish, SNP locus 1544-242, which suggests that differences in hemoglobin transcript abundance among whitefish morphs are at least partially driven by genetic adaptation. Hemoglobin transcript abundance was examined in whole brain and gill, and thus, our results may reflect gene transcription in brain or gill cells (Ullal et al. 2008; Biagioli et al. 2009) or the blood cells that perfuse these tissues. Given that the SNP was also strongly associated with anodic (predicted Bohr-effect) hemoglobin protein components isolated from blood, it is possible that we may have detected a direct link between this SNP, alpha hemoglobin gene transcript abundance, and protein component variation in RBCs present in

gill and brain. SNP 1544-242 is located in a putative non-coding region, so it may not be directly involved in the regulation of gene transcription, but instead linked to regulatory or coding regions within the whitefish hemoglobin genome. Albeit, a growing number of studies have also implicated intronic variation as direct regulators of gene transcription (Croisetière et al. 2010; Barrett et al. 2012). Complete sequencing of the hemoglobin repertoire is needed to further elucidate links between this SNP and the evolution of hemoglobin diversity and divergence among whitefish morphs.

While we did not detect an influence of SNP genotype on hemoglobin transcript abundance in kidney tissue, we observed a strong effect of morph, which may have swamped any signal driven by genetic background. Our sample sizes of whitefish were limited from Skrukkebukta compared to Langfjordvatn, particularly for the DR morph. Further studies that incorporate larger sample sizes of each morph bearing each of the SNP genotypes are required to fully assess the role of this locus as a possible eQTL. Furthermore, hemoglobin transcript abundance in blood should be compared among whitefish bearing each of the genotypes to determine whether hemoglobin gene transcription in RBCs is driving the observed relationship in gill and brain (see Lund et al. 2000).

Interestingly, we have also demonstrated differences in population genetic structure among the morphs at SNP locus 1544-242. Both Skrukkebukta and Langfjordvatn LSR whitefish are predominantly of the GG genotype whereas pelagic DR and profundal SSR whitefish are predominantly of the AA and AG genotypes. While the DR and SSR morphs appear to have lost anodic hemoglobin protein component diversity compared to that observed in the ancestral LSR morph, the AA genotype, which is linked to higher hemoglobin gene transcript abundance in the brain and gill, is more frequent in these morphs, and thus, could represent an alternative genomic mechanism used by whitefish to meet their oxygen transport requirements in the pelagic and profundal zones. Indeed, studies in cod have

shown that hemoglobin gene upregulation may compensate for the low oxygen binding affinities of certain hemoglobin protein types (Star et al. 2011). The observed shift in allele frequencies between the LSR and DR/SSR populations is suggestive of whitefish hemoglobin divergence facilitated by standing genetic variation present in the ancestral LSR population (also see Colosimo et al. 2005; Barrett and Schluter 2008; Jones et al. 2012).

Taken together, our results indicate that divergence of the oxygen transport system in trimorphic populations of European whitefish has primarily occurred across the profundal – littoral/pelagic resource axis, an observation supported by patterns of hemoglobin gene transcript abundance in kidney tissue, hemoglobin protein component composition in blood, and population genetic structure across lacustrine zones. These findings, along with our previous work on lake whitefish, which demonstrated replicated divergence of the oxygen transport system across the limnetic – benthic resource axis in North American lakes, highlights the likely importance of physiological oxygen uptake and transport to the postglacial colonization of novel lacustrine environments by whitefish throughout the northern hemisphere. We hypothesize that the alternate oxygen environments, temperatures, and variable activity energetic demands associated with foraging and predator avoidance in shallow and deep lacustrine zones are important ecological drivers of metabolic trait divergence (Landry et al. 2007; Evans and Bernatchez 2012; Evans et al. 2012, 2013). Experimental studies are now needed to examine the functional significance and fitness consequences of RBC and hemoglobin variants found in European and North American populations.

## ACKNOWLEDGEMENTS

We thank G. Côté and T. L. Hanebrette for assistance with the molecular work. We are grateful to Associate Editor W. Salzburger and three anonymous reviewers for helpful comments on the manuscript. This work was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant and Canada Research Chair Grant to LB, an NSERC PDF award to MLE, and a Norwegian Research Council Grant to PAA (grant no. 186320/V40). The authors report no conflict-of-interest.

## DATA ARCHIVING

doi:10.5061/dryad.351h3

Accepted Article

## REFERENCES

- Abdel-Aziz, E. S. H., T. E. S. Ali, S. B. S. Abdu, and H. F. Fouad. 2010. Chromaffin cells and interrenal tissue in the head kidney of the grouper, *Epinephelus tauvina* (Teleostei, Serranidae): a morphological (optical and ultrastructural) study. *J. Appl. Ichthyol.* 26:522–527.
- Abzhanov, A., W. P. Kuo, C. Hartmann, B. R. Grant, P. R. Grant, and C. J. Tabin. 2006. The calmodulin pathway and evolution of elongated beak morphology in Darwin's finches. *Nature* 442:563–7.
- Amundsen, P.-A., T. Bøhn, and G. H. Våga. 2004. Gill raker morphology and feeding ecology of two sympatric morphs of European whitefish (*Coregonus lavaretus*). *Ann. Zool. Fennici* 41:291–300.
- Andersen, O., O. F. Wetten, M. C. De Rosa, C. Andre, C. Carelli Alinovi, M. Colafranceschi, O. Brix, and A. Colosimo. 2009. Haemoglobin polymorphisms affect the oxygen-binding properties in Atlantic cod populations. *Proc. R. Soc. B* 276:833–41.
- Barra, D., F. Bossa, J. Bonaventura, and M. Brunori. 1973. Hemoglobin components from trout (*Salmo irideus*): determination of the carboxyl and amino terminal sequences and their functional implications. *FEBS Lett.* 35:151–4.
- Barrett, L. W., S. Fletcher, and S. D. Wilton. 2012. Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. *Cell. Mol. Life Sci.* 69:3613–34.
- Barrett, R. D. H., S. M. Rogers, and D. Schluter. 2008. Natural selection on a major armor gene in threespine stickleback. *Science* 322:255–257.
- Barrett, R. D. H., and D. Schluter. 2008. Adaptation from standing genetic variation. *Trends Ecol. Evol.* 23:38–44.
- Bernatchez, L., and J. Dodson. 1985. Influence of temperature and current speed on the swimming capacity of lake whitefish. *Can. J. Fish. Aquat. Sci.* 42:1522–1529.
- Bernatchez, L., S. Renaut, A. R. Whiteley, N. Derome, J. Jeukens, L. Landry, G. Q. Lu, A. W. Nolte, K. Ostbye, S. M. Rogers, and J. St-Cyr. 2010. On the origin of species: insights from the ecological genomics of lake whitefish. *Philos. Trans. R. Soc. B* 365:1783–1800.
- Biagioli, M., M. Pinto, D. Cesselli, M. Zaninello, D. Lazarevic, P. Roncaglia, R. Simone, C. Vlachouli, C. Plessy, N. Bertin, A. Beltrami, K. Kobayashi, V. Gallo, C. Santoro, I. Ferrer, S. Rivella, C. A. Beltrami, P. Carninci, E. Raviola, and S. Gustincich. 2009. Unexpected expression of alpha- and beta-globin in mesencephalic dopaminergic neurons and glial cells. *Proc. Natl. Acad. Sci.* 106:15454–9.
- Brix, O., S. Thorkildsen, and A. Colosimo. 2004. Temperature acclimation modulates the oxygen binding properties of the Atlantic cod (*Gadus morhua* L.) genotypes-HbI\*1/1,



HbI\*1/2, and HbI\*2/2-by changing the concentrations of their major hemoglobin components. *Comp. Biochem. Physiol. A. Mol. Integr. Physiol.* 138:241–51.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Villarreal, M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, and D. M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* 307:1928–33.

Croisetière, S., L. Bernatchez, and P. Belhumeur. 2010. Temperature and length-dependent modulation of the MH class II beta gene expression in brook charr (*Salvelinus fontinalis*) by a cis-acting minisatellite. *Mol. Immunol.* 47:1817–29.

DeFaveri, J., T. Shikano, Y. Shimada, A. Goto, and J. Merilä. 2011. Global analysis of genes involved in freshwater adaptation in threespine sticklebacks (*Gasterosteus aculeatus*). *Evolution* 65:1800–7.

Derome, N., and L. Bernatchez. 2006. The transcriptomics of ecological convergence between 2 limnetic coregonine fishes (Salmonidae). *Mol. Biol. Evol.* 23:2370–2378.

Elmer, K. R., S. Fan, H. M. Gunter, J. C. Jones, S. Boekhoff, S. Kuraku, and A. Meyer. 2010. Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. *Mol. Ecol.* 19 Suppl 1:197–211.

Elmer, K. R., and A. Meyer. 2011. Adaptation in the age of ecological genomics: insights from parallelism and convergence. *Trends Ecol. Evol.* 26:298–306. Elsevier Science Publishers.

Evans, M. L., and L. Bernatchez. 2012. Oxidative phosphorylation gene transcription in whitefish species pairs reveals patterns of parallel and nonparallel physiological divergence. *J. Evol. Biol.* 25:1823–1834.

Evans, M. L., L. J. Chapman, I. Mitrofanov, and L. Bernatchez. 2013. Variable extent of parallelism in respiratory, circulatory, and neurological traits across lake whitefish species pairs. *Ecol. Evol.* 3:546–57.

Evans, M. L., K. Praebel, S. Peruzzi, and L. Bernatchez. 2012. Parallelism in the oxygen transport system of the lake whitefish: the role of physiological divergence in ecological

Fange, R. 1994. Blood-cells, hematopoiesis and lymphomyeloid tissues in fish. *Fish Shellfish Immunol.* 4:405–411.

Filteau, M., S. a Pavey, J. St-Cyr, and L. Bernatchez. 2013. Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. *Mol. Biol. Evol.* 30:1384–96.

Fyhn, U. E. H., and R. E. Withler. 1991. A genetic polymorphism in hemoglobins of chinook salmon, *Oncorhynchus tshawytscha*. *Can. J. Zool.* 69:1904–1910.

Giles, N. 1983. The possible role of environmental calcium levels during the evolution of phenotypic diversity in Outer Hebridean populations of the Three-spined stickleback , *Gasterosteus aculeatus*. *J. Zool.* 199:535–544.

Harrod, C., J. Mallela, and K. K. Kahilainen. 2010. Phenotype-environment correlations in a putative whitefish adaptive radiation. *J. Anim. Ecol.* 79:1057–1068.

Husebø, A., A. K. Imsland, and G. Naevdal. 2004. Haemoglobin variation in cod: a description of new variants and their geographical distribution. *Sarsia* 89:369–378.

Jeukens, J., D. Bittner, R. Knudsen, and L. Bernatchez. 2009. Candidate genes and adaptive radiation: insights from transcriptional adaptation to the limnetic niche among Coregonine fishes (*Coregonus* spp., Salmonidae). *Mol. Biol. Evol.* 26:155–166.

Jeukens, J., B. Boyle, I. Kukavica-Ibrulj, J. St-Cyr, R. C. Levesque, and L. Bernatchez. 2011. BAC library construction, screening and clone sequencing of lake whitefish (*Coregonus clupeaformis*, Salmonidae) towards the elucidation of adaptive species divergence. *Mol. Ecol. Resour.* 11:541–549.

Jeukens, J., S. Renaut, J. St-Cyr, A. W. Nolte, and L. Bernatchez. 2010. The transcriptomics of sympatric dwarf and normal lake whitefish (*Coregonus clupeaformis* spp., Salmonidae) divergence as revealed by next-generation sequencing. *Mol. Ecol.* 19:5389–5403.

Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, E. Birney, S. Searle, J. Schmutz, J. Grimwood, M. C. Dickson, R. M. Myers, C. T. Miller, B. R. Summers, A. K. Knecht, S. D. Brady, H. Zhang, A. a Pollen, T. Howes, C. Amemiya, J. Baldwin, T. Bloom, D. B. Jaffe, R. Nicol, J. Wilkinson, E. S. Lander, F. Di Palma, K. Lindblad-Toh, and D. M. Kingsley. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61.

Kahilainen, K., and H. Lehtonen. 2003. Piscivory and prey selection of four predator species in a whitefish dominated subarctic lake. *J. Fish Biol.* 63:659–672.

Kahilainen, K., and K. Ostbye. 2006. Morphological differentiation and resource polymorphism in three sympatric whitefish *Coregonus lavaretus* (L.) forms in a subarctic lake. *J. Fish Biol.* 68:63–79.

Kujansuu, R., B. Eriksson, and T. Grönlund. 1998. Lake Inarijärvi, northern Finland: Sedimentation and late quaternary evolution. P. 25 in Geological Survey of Finland, Report of Investigation. Espoo.

Kupittayanant, P., and W. Kinchareon. 2011. Hematological and biochemical responses of the flowerhorn fish to hypoxia. *J. Anim. Vet. Adv.* 10:2631–2638.

Landry, L., W. F. Vincent, and L. Bernatchez. 2007. Parallel evolution of lake whitefish dwarf ecotypes in association with limnological features of their adaptive landscape. *J. Evol. Biol.* 20:971–984.

Lay, P. A., and J. Baldwin. 1999. What determines the size of teleost erythrocytes? Correlations with oxygen transport and nuclear volume. *Fish Physiol. Biochem.* 20:31–35.

Livak, K. J., and T. D. Schmittgen. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C_T}$  method. *Methods* 25:402–8.

Lund, S. G., M. C. L. Phillips, C. D. Moyes, and B. L. Tufts. 2000. The effects of cell ageing on protein synthesis in rainbow trout red blood cells. *J. Exp. Biol.* 2228:2219–2228.

Maciak, S., K. Janko, J. Kotusz, L. Choleva, A. Boroń, D. Juchno, R. Kujawa, J. Kozłowski, and M. Konarzewski. 2011. Standard Metabolic Rate (SMR) is inversely related to erythrocyte and genome size in allopolyploid fish of the *Cobitis taenia* hybrid complex. *Funct. Ecol.* 25:1072–1078.

McMorrow, T., A. Wagner, F. Deryckere, and F. Gannon. 1996. Structural organization and sequence analysis of the globin locus in Atlantic salmon. *DNA Cell Biol.* 15:407–414.

Nikinmaa, M. 1997. Oxygen and carbon dioxide transport in vertebrate erythrocytes: An evolutionary change in the role of membrane transport. *J. Exp. Biol.* 200:369–380.

Olsvik, P. A., K. K. Lie, A. E. O. Jordal, T. O. Nilsen, and I. Hordvik. 2005. Evaluation of potential reference genes in real-time RT-PCR studies of Atlantic salmon. *BMC Mol. Biol.* 6:21.

Østbye, K., P. Amundsen, L. Bernatchez, a Klemetsen, R. Knudsen, R. Kristoffersen, T. F. Naesje, and K. Hindar. 2006. Parallel evolution of ecomorphological traits in the European whitefish *Coregonus lavaretus* (L.) species complex during postglacial times. *Mol. Ecol.* 15:3983–4001.

Østbye, K., L. Bernatchez, T. F. Naesje, K. J. M. Himberg, and K. Hindar. 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.

Pavey, S. A., H. Collin, P. Nosil, and S. M. Rogers. 2010. The role of gene expression in ecological speciation. *Ann. N. Y. Acad. Sci.* 1206:110–29.

Powers, D. A. 1972. Hemoglobin adaptation for fast and slow water habitats in sympatric Catostomid fishes. *Science* (80). 177:360–362.

Powers, D. A. 1980. Molecular ecology of teleost fish hemoglobins: strategies for adapting to changing environments. *Am. Zool.* 162:139–162.

Præbel, K., R. Knudsen, A. Siwertsson, M. Karhunen, and K. K. Kahilainen. 2013. Ecological speciation in postglacial European whitefish: rapid adaptive radiations into the littoral, pelagic and profundal lake habitats. *Ecol. Evol.* 3:4970–4986.

Quinn, N. L., K. A. Boroevich, K. P. Lubieniecki, W. Chow, E. A. Davidson, R. B. Phillips, B. F. Koop, and W. S. Davidson. 2010. Genomic organization and evolution of the Atlantic salmon hemoglobin repertoire. *BMC Genomics* 11:539.

Raymond, M., and F. Rousset. 1995. Genepop (version-1.2) - population-genetics software for exact tests and ecumenicism. *J. Hered.* 86:248–249.

Renaut, S., A. W. Nolte, and L. Bernatchez. 2010. Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Mol. Ecol.* 19:115–131.

Renaut, S., A. W. Nolte, S. M. Rogers, N. Derome, and L. Bernatchez. 2011. SNP signatures of selection on standing genetic variation and their association with adaptive phenotypes along gradients of ecological speciation in lake whitefish species pairs (*Coregonus* spp.). *Mol. Ecol.* 20:545–559.

Rutjes, H. A., M. C. Nieveen, R. E. Weber, F. Witte, and G. E. E. J. M. Van Den Thillart. 2007. Multiple strategies of Lake Victoria cichlids to cope with lifelong hypoxia include hemoglobin switching. *Am. J. Physiol. Integr. Comp. Physiol.* 293:R1376–83.

Salzburger, W., S. C. P. Renn, D. Steinke, I. Braasch, H. a Hofmann, and A. Meyer. 2008. Annotation of expressed sequence tags for the East African cichlid fish *Astatotilapia burtoni* and evolutionary analyses of cichlid ORFs. *BMC Genomics* 9:96.

Sandlund, O. T., K. Ø. Gjelland, T. Bøhn, R. Knudsen, and P.-A. Amundsen. 2013. Contrasting population and life history responses of a young morph-pair of European whitefish to the invasion of a specialised coregonid competitor, vendace. *PLoS One* 8:e68156.

Sick, K. 1961. Haemoglobin polymorphism in fishes. *Nature* 192:894.

Siwertsson, A., R. Knudsen, K. K. Kahilainen, K. Praebel, R. Primicerio, and P.-A. Amundsen. 2010. Sympatric diversification as influenced by ecological opportunity and historical contingency in a young species lineage of whitefish. *Evol. Ecol. Res.* 12:929–947.

Snyder, G., and B. Sheafor. 1999. Red blood cells: centerpiece in the evolution of the vertebrate circulatory system. *Am. Zool.* 39:189–198.

Star, B., A. J. Nederbragt, S. Jentoft, U. Grimholt, M. Malmstrøm, T. F. Gregers, T. B. Rounge, J. Paulsen, M. H. Solbakken, A. Sharma, O. F. Wetten, A. Lanzén, R. Winer, J. Knight, J.-H. Vogel, B. Aken, O. Andersen, K. Lagesen, A. Tooming-Klunderud, R. B. Edvardsen, K. G. Tina, M. Espelund, C. Nepal, C. Previti, B. O. Karlsen, T. Moum, M. Skage, P. R. Berg, T. Gjøn, H. Kuhl, J. Thorsen, K. Malde, R. Reinhardt, L. Du, S. D. Johansen, S. Searle, S. Lien, F. Nilsen, I. Jonassen, S. W. Omholt, N. C. Stenseth, and K. S. Jakobsen. 2011. The genome sequence of Atlantic cod reveals a unique immune system. *Nature* 477:207–10.

St-Cyr, J., N. Derome, and L. Bernatchez. 2008. The transcriptomics of life-history trade-offs in whitefish species pairs (*Coregonus* sp.). *Mol. Ecol.* 17:1850–1870.

Taylor, E. B. 1999. Species pairs of north temperate freshwater fishes: evolution, taxonomy, and conservation. *Rev. Fish Biol. Fish.* 9:299–324.

Trudel, M., A. Tremblay, R. Schetagne, and J. B. Rasmussen. 2001. Why are dwarf fish so small? An energetic analysis of polymorphism in lake whitefish (*Coregonus clupeaformis*). *Can. J. Fish. Aquat. Sci.* 58:394–405.

Ullal, A. J., R. W. Litaker, and E. J. Noga. 2008. Antimicrobial peptides derived from hemoglobin are expressed in epithelium of channel catfish (*Ictalurus punctatus*, Rafinesque). *Dev. Comp. Immunol.* 32:1301–12.

Verde, C., M. Balestrieri, D. de Pascale, D. Pagnozzi, G. Lecointre, and G. di Prisco. 2006. The oxygen transport system in three species of the boreal fish family gadidae - Molecular phylogeny of hemoglobin. *J. Biol. Chem.* 281:22073–22084.

Verde, C., D. Giordano, G. Prisco, and Ø. Andersen. 2012. The haemoglobins of polar fish: evolutionary and physiological significance of multiplicity in Arctic fish. *Biodiversity* 13:228–233.

Vonlanthen, P., D. Bittner, a G. Hudson, K. a Young, R. Müller, B. Lundsgaard-Hansen, D. Roy, S. Di Piazza, C. R. Largiader, and O. Seehausen. 2012. Eutrophication causes speciation reversal in whitefish adaptive radiations. *Nature* 482:357–62.

Weber, R. E. 1990. Functional significance and structural basis of multiple hemoglobins with special reference to ectothermic vertebrates. Pp. 58–75 in J. P. Truchot and B. Lahlou, eds. *Animal nutrition and transport processes. 2. Transport, respiration, and excretion: comparative and environmental aspects.* Karger, Basel.

Wetten, O. F., A. J. Nederbragt, R. C. Wilson, K. S. Jakobsen, R. B. Edvardsen, and O. Andersen. 2010. Genomic organization and gene expression of the multiple globins in Atlantic cod: conservation of globin-flanking genes in chordates infers the origin of the vertebrate globin clusters. *BMC Evol. Biol.* 10:315.

Whiteley, A. R., N. Derome, S. M. Rogers, J. St-Cyr, J. Laroche, A. Labbe, A. Nolte, S. Renaut, J. Jeukens, and L. Bernatchez. 2008. The phenomics and expression quantitative trait locus mapping of brain transcriptomes regulating adaptive divergence in lake whitefish species pairs (*Coregonus* sp.). *Genetics* 180:147–164.

Wilhem, D., G. J. Eble, G. Kassner, F. X. Caprario, A. L. Dafre, and M. Ohira. 1992. Comparative hematology in marine fish. *Comp. Biochem. Physiol. A-Molecular Integr. Physiol.* 102:311–321.

#### DATA ACCESSIBILITY

Gene expression, SNP genotypes, red blood cell measurements, and hemoglobin protein component data are available from the Dryad Digital Repository: doi:10.5061/dryad.351h3

#### AUTHOR CONTRIBUTIONS

All authors contributed to study design and data collection. MLE analysed the data and wrote the initial manuscript draft. LB, KP, SP, and PAA contributed to data interpretation and editing of the manuscript.

Table 1. Summary of the number of European whitefish (*Coregonus lavaretus*) individuals examined for hemoglobin gene transcript abundance in each of three tissues, kidney, brain, and gill, hemoglobin protein component variation in blood, RBC morphology, and SNP genotype from two lakes, Langfjordvatn (LF) and Skrukkebukta (SK), in the Pasvik region of northern Norway.

Lake	Morph	Transcript abundance			Hb Protein	RBC Morphology	SNP Genotype
		Kidney	Brain	Gill			
LF	LSR	15	12	16	32	8	37
	DR	13	9	14	24	13	22
	SSR	16	11	18	32	7	27
SK	LSR	6	6	6	12	11	20
	DR	4	4	4	2	4	4
	SSR	5	6	6	13	14	15

Table 2. Red blood cell (RBC) morphological and hemoglobin (Hb) protein component variation found in the littoral large sparsely-rakered (LSR), pelagic densely-rakered (DR), and profundal small sparsely-rakered (SSR) European whitefish (*Coregonus lavaretus*) from two lakes, Langfjordvatn (LF) and Skrukkebukta (SK). For RBCs, the mean nucleus (Nuc) area, total cell area, and nucleus:total cell area (Nuc:Total Area) is indicated  $\pm$  1 SD. For Hb protein components, the total number of anodic (Anod) and cathodic (Cath) components found in each lake is indicated, as is the mean number of each protein component type, total number of hemoglobin protein components, and proportion of cathodic protein components found in individual (Ind.) whitefish  $\pm$  1 SD. The total number of individuals examined per morph (N) is also shown.

Lake	Morph	Nuc Area ( $\mu\text{m}^2$ )	Total Cell Area ( $\mu\text{m}^2$ )	Nuc:Total Area	Anod Hb Lake	Cath Hb Lake	Anod Hb Ind.	Cath Hb Ind.	Total Hb Ind.	Prop Cath Hb Ind.
LF	LSR	25.4 $\pm$ 4.0	150.5 $\pm$ 18.8	0.17 $\pm$ 0.03	4	4	3.1 $\pm$ 0.8	3.0 $\pm$ 0.5	6.0 $\pm$ 1.1	0.49 $\pm$ 0.08
	DR	27.2 $\pm$ 4.8	139.1 $\pm$ 11.0	0.20 $\pm$ 0.04	3	3	2.4 $\pm$ 0.5	2.2 $\pm$ 0.4	4.5 $\pm$ 0.6	0.48 $\pm$ 0.07
	SSR	27.3 $\pm$ 5.2	159.8 $\pm$ 9.9	0.17 $\pm$ 0.03	4	4	2.5 $\pm$ 0.7	2.6 $\pm$ 0.6	5.1 $\pm$ 1.1	0.51 $\pm$ 0.07
SK	LSR	31.3 $\pm$ 3.0	137.1 $\pm$ 7.0	0.23 $\pm$ 0.01	4	4	3.5 $\pm$ 0.5	2.8 $\pm$ 0.6	6.3 $\pm$ 0.9	0.45 $\pm$ 0.06
	DR	29.3 $\pm$ 3.3	130.9 $\pm$ 7.9	0.22 $\pm$ 0.02	3	2	2.5 $\pm$ 0.7	2.0	4.5 $\pm$ 0.7	0.45 $\pm$ 0.07
	SSR	25.7 $\pm$ 2.1	150.9 $\pm$ 8.7	0.17 $\pm$ 0.02	4	3	2.4 $\pm$ 0.9	2.5 $\pm$ 0.5	4.9 $\pm$ 0.8	0.53 $\pm$ 0.11

Table 3. Results of two-way analysis of variance (ANOVA) examining the influence of morph and lake on red blood cell (RBC) morphology and the proportion of cathodic hemoglobin (Hb) protein components in European whitefish (*Coregonus lavaretus*). The ANOVA models examined the contribution of lake, morph, and lake×morph to variation in RBC nucleus area, total cell area, and nucleus:total cell area, and to total number of hemoglobin protein components and the proportion of cathodic (Cath) hemoglobin protein components found per individual. F-statistics, numerator and denominator degrees of freedom (DF) for the model, DF for each fixed effect, and P-values are indicated. P-values falling below the critical  $\alpha = 0.05$  are bolded.

	Nucleus Area			Total Cell Area			Nucleus:Total Cell Area			Total Hb Individuals			Proportion Cath Hb Individuals		
	F	DF	P	F	DF	P	F	DF	P	F	DF	P	F	DF	P
Model	3.53	5,56	<b>0.008</b>	6.94	5,56	<b>&lt;0.001</b>	8.61	5,56	<b>&lt;0.001</b>	10.74	5,109	<b>&lt;0.001</b>	1.89	5,109	0.102
Lake	3.76	1	0.057	10.10	1	<b>0.003</b>	12.32	1	<b>0.001</b>	0.01	1	0.944	4.10	1	0.343
Morph	1.31	2	0.278	13.20	2	<b>&lt;0.001</b>	8.93	2	<b>&lt;0.001</b>	17.90	2	<b>&lt;0.001</b>	3.56	2	<b>0.032</b>
Lake×Morph	4.57	2	<b>0.015</b>	0.29	2	0.747	5.52	2	<b>0.006</b>	0.65	2	0.523	1.72	2	0.185



Table 4. Results of two-way analysis of variance (ANOVA) models examining the contribution of lake, morph, and lake×morph to hemoglobin gene transcript abundance (Log+1 transformed) in the kidney, brain, and gill of the European whitefish (*Coregonus lavaretus*). Results of the ANOVA models are shown for the alpha and beta hemoglobin (Hb) genes. F-statistics, degrees of freedom (DF), including model and error DF for the model, and P-values are indicated. P-values falling below the critical  $\alpha = 0.05$  are bolded.

Tissue		Alpha Hb ANOVA			Beta Hb ANOVA		
		F	DF	P	F	DF	P
Kidney	Model	5.12	5, 53	<b>&lt;0.001</b>	1.59	5, 51	0.177
	Lake	7.58	1	<b>0.008</b>	3.05	1	0.086
	Morph	8.76	2	<b>0.001</b>	1.76	2	0.182
	Lake×Morph	1.13	2	0.331	0.98	2	0.982
Brain	Model	8.42	5, 42	<b>&lt;0.001</b>	3.85	5, 42	<b>0.006</b>
	Lake	38.06	1	<b>&lt;0.001</b>	10.35	1	<b>0.003</b>
	Morph	1.76	2	0.183	2.48	2	0.095
	Lake×Morph	1.59	2	0.216	0.79	2	0.459
Gill	Model	2.87	5, 58	<b>0.022</b>	2.31	5, 58	0.054
	Lake	7.06	1	<b>0.005</b>	2.64	1	0.109
	Morph	4.21	2	0.086	1.52	2	0.227
	Lake×Morph	3.96	2	0.099	3.12	2	0.051

Table 5. Summary of SNP 1544-242 genotype frequencies in the littoral large sparsely-rakered (LSR), pelagic densely-rakered (DR), and profundal small sparsely-rakered (SSR) European whitefish (*Coregonus lavaretus*) from two lakes, Langfjordvatn (LF) and Skrukkebukta (SK). Observed (O) and expected (E) numbers of whitefish bearing each genotype and Hardy-Weinberg tests for heterozygote excess and deficit are shown for each morph within each lake. Associations between the SNP genotype frequencies and whitefish morphs in each lake were estimated using Chi-square ( $\chi^2$ ) and Fisher's exact tests. P-values falling below the critical  $\alpha = 0.05$  are bolded

Lake	Morph	GG		AG		AA		H <sub>0</sub> Excess (P)	H <sub>0</sub> Deficit (P)	$\chi^2$	DF	P ( $\chi^2$ )	P (Fisher)
		O	E	O	E	O	E						
LF	LSR	20	12.9	7	11.2	10	12.9	1.000	< <b>0.001</b>	11.12	4	<b>0.025</b>	<b>0.027</b>
	DR	4	7.7	8	6.6	10	7.6	0.946	0.248				
	SSR	6	9.4	11	8.2	10	9.4	0.916	0.279				
SK	LSR	16	12.3	4	5.1	0	2.6	0.848	1.000	15.51	4	<b>0.003</b>	<b>0.002</b>
	DR	4	2.5	0	1.0	0	0.5	-	-				
	SSR	4	9.2	6	3.8	5	1.9	0.924	0.352				

Table 6. Results of nested analysis of variance (ANOVA) models examining relationships between hemoglobin SNP 1544-242 and levels of alpha hemoglobin gene transcript abundance in European whitefish (*Coregonus lavaretus*) from two lakes, Langfjordvatn and Skrukkebukta. Relationships between SNP genotype (GG, AG, AA) and log+1 transformed levels of gene transcript abundance were examined in kidney, brain, and gill. F-statistics and P-values are reported for the model and each factor included within the model. Degrees of freedom (DF), including model and error DF for the model, are also indicated. P-values falling below the critical  $\alpha = 0.05$  are bolded.

	Kidney			Brain			Gill		
	F	DF	<i>P</i>	F	DF	<i>P</i>	F	DF	<i>P</i>
Model	1.39	8,32	0.238	8.27	7,28	<b>&lt;0.001</b>	3.84	7,38	<b>0.003</b>
Lake	1.06	1	0.311	12.91	1	<b>0.001</b>	0.28	1	0.598
SNP	0.26	2	0.769	8.59	2	<b>0.001</b>	7.92	2	<b>0.001</b>
SNP [Morph]	2.02	5	0.101	1.64	4	0.191	1.58	4	0.199

Table 7. The influence of SNP 1544-242 genotype on hemoglobin protein component variation in European whitefish (*Coregonus lavaretus*). Chi-square ( $\chi^2$ ) tests were used to examine differences in the observed (O) and expected (E) number of individuals of each genotype (GG, AG, AA) bearing each of the protein components. Degrees of freedom (DF) and the significance of each test ( $P$ ) are indicated.  $P$ -values falling below the critical  $\alpha$  after Bonferroni correction ( $P = 0.05/6$ ) are bolded.

Protein Component	GG		AG		AA		$\chi^2$	DF	$P$	$P$ (Fisher's Exact)	
	O	E	O	E	O	E					
C1	0	0	0.4	0	0.3	1	0.4	2.06	2	0.356	0.627
	1	28	27.6	20	19.7	26	26.6				
C2	0	9	10.8	8	7.7	12	10.4	0.90	2	0.636	0.632
	1	19	17.2	12	12.3	15	16.6				
C3	0	0	0	0	0	0	0	-	-	-	-
	1	28	28	20	20	27	27				
C4	0	26	26.1	19	18.7	25	25.2	0.13	2	0.937	1.000
	1	2	1.9	1	1.3	2	1.8				
A1	0	3	17.2	17	12.3	26	16.6	55.56	2	<b>&lt;0.001</b>	<b>&lt;0.001</b>
	1	25	10.8	3	7.7	1	10.4				
A2	0	1	11.9	7	8.5	24	11.5	48.99	2	<b>&lt;0.001</b>	<b>&lt;0.001</b>
	1	27	16.1	13	11.5	3	15.5				
A3	0	0	0	0	0	0	0	-	-	-	-
	1	28	28	20	20	27	27				
A4	0	10	4.1	1	2.9	0	4.0	18.09	2	<b>&lt;0.001</b>	<b>&lt;0.001</b>
	1	18	23.9	19	17.1	27	23.0				

Figure 1. Locations of the two lakes, Langfjordvatn and Skrukkebukta, containing trimorphic populations of European whitefish (*Coregonus lavaretus*).

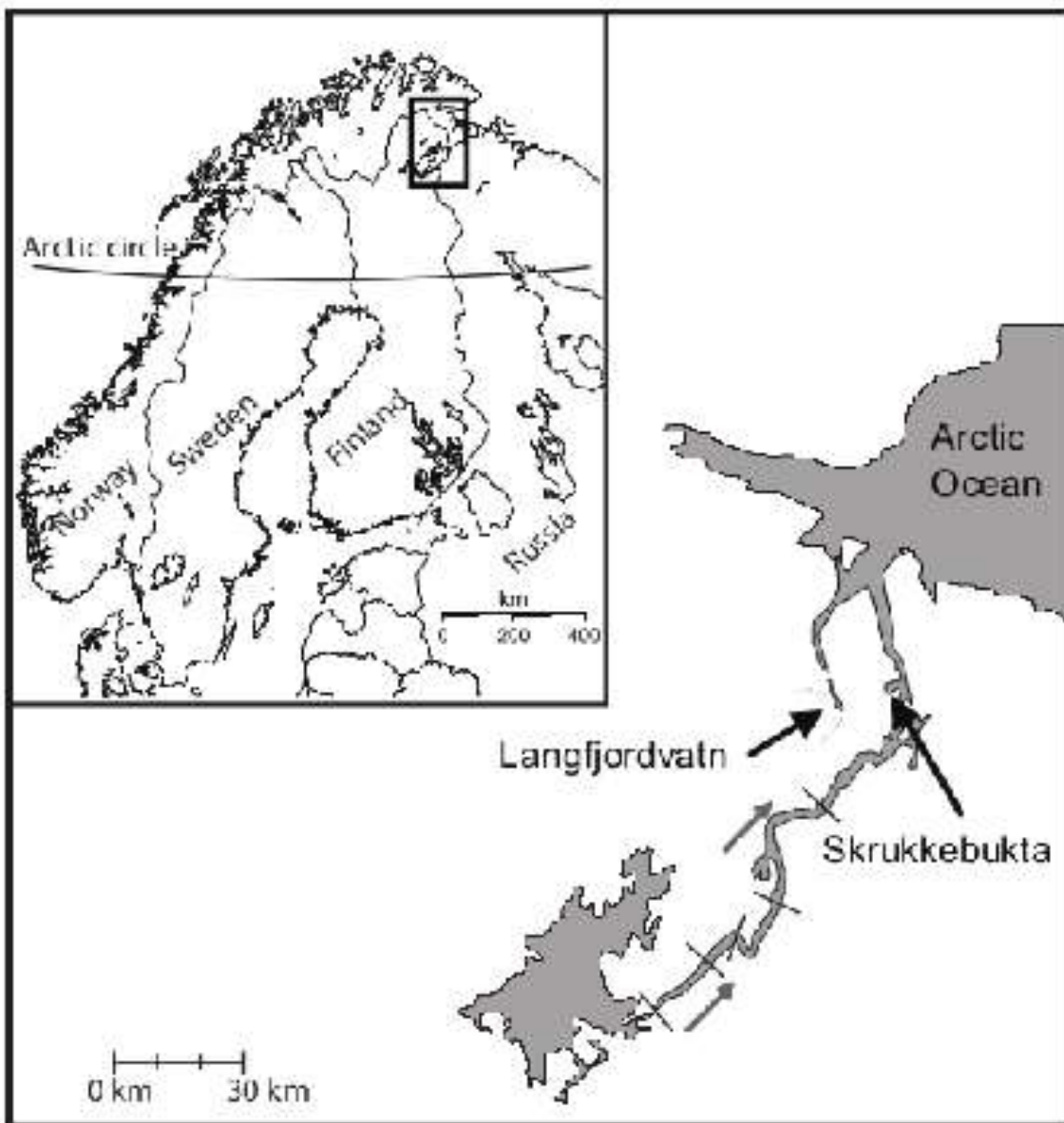


Figure 2. Hemoglobin gene transcript abundance (RQ) exhibited in European whitefish (*Coregonus lavaretus*) from two lakes. Hemoglobin alpha (black bars) and beta (grey bars) RQ levels in the littoral large sparsely-rakered (LSR), pelagic densely-rakered (DR), and profundal small sparsely-rakered (SSR) whitefish morphs from Langfjordvatn (LF) and Skrukkebukta (SK) are shown. Hemoglobin RQ was compared among the morphs and between lakes in the (a) kidney, (b) brain, and (c) gill. The bars show mean RQ  $\pm$  1 SD and sample sizes are shown above the upper error bar. The statistical analyses corresponding to the figure are shown in Table 4.

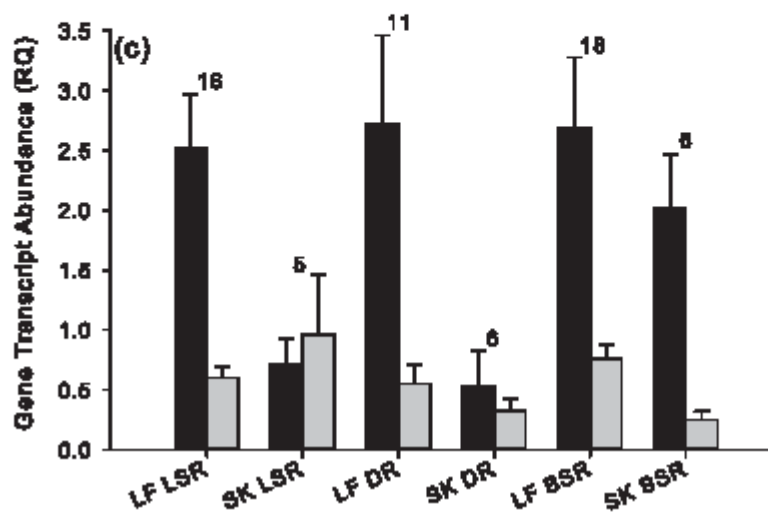
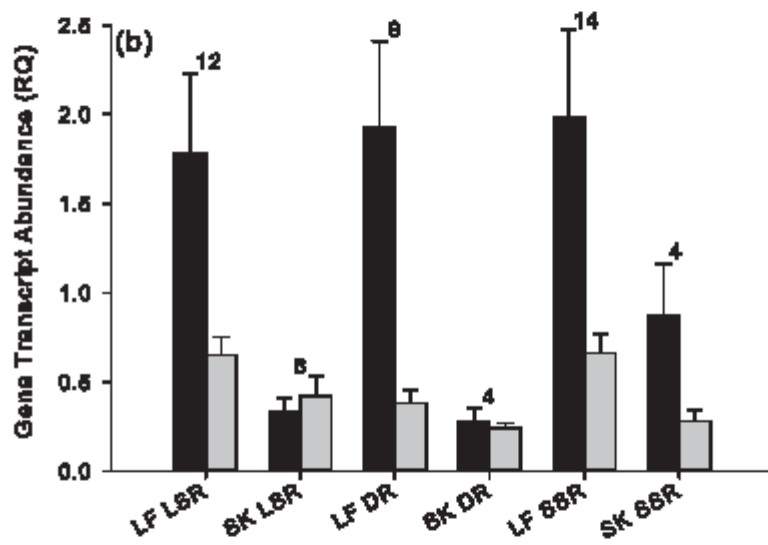
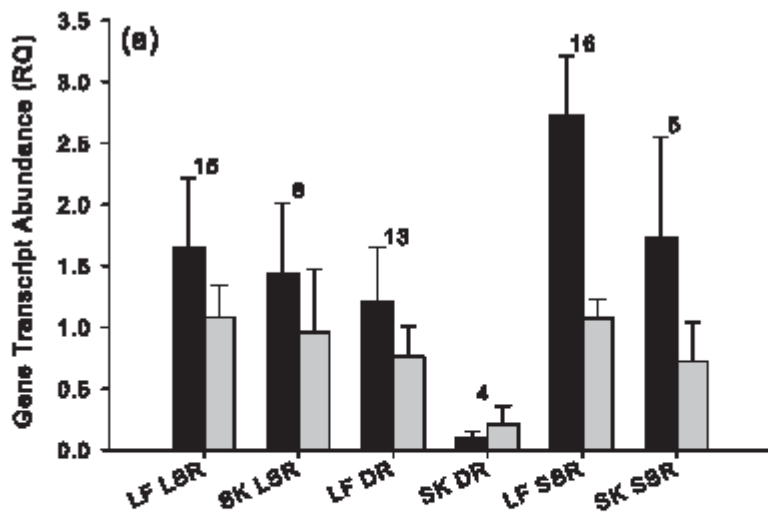


Figure 3. Relationship between hemoglobin SNP 1544-242 genotype and alpha hemoglobin transcript abundance (RQ) in European whitefish (*Coregonus lavaretus*) morphs from two lakes in the Pasvik region of northern Norway. The relationship between the SNP genotype (AA, GA, GG) and RQ in littoral large sparsely-rakered (black bars), pelagic densely-rakered (light grey bars), and profundal small sparsely-rakered (dark grey bars) morph is shown for the (a) kidney, (b) brain, and (c) gill. The bars correspond to the least squares means  $\pm$  1 SE of RQ, which corrects RQ for any lake effects. Sample sizes are reported above the upper error bar. The statistical analyses corresponding to the figure are shown in Table 6.



