MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 4773–4792 doi: 10.1111/mec.13795

# Investigating genomic and phenotypic parallelism between piscivorous and planktivorous lake trout (*Salvelinus namaycush*) ecotypes by means of RADseq and morphometrics analyses

S. BERNATCHEZ,* M. LAPORTE,* C. PERRIER,*† P. SIROIS‡ and L. BERNATCHEZ*

*Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, Québec Canada, G1V 0A6, †Centre d'Ecologie Fonctionnelle and Evolutive, CNRS, 34293 Montpellier 5, France, ‡Chaire de recherche sur les espèces aquatiques exploitées, Université du Québec à Chicoutimi, Chicoutimi, Québec Canada, G7H 2B1

## Abstract

**Repeated adaptive ecological diversification has commonly been reported in fish and has often been associated with trophic niche diversity. The main goal of this study was to investigate the extent of parallelism in the genomic and phenotypic divergence between piscivorous and planktivorous lake trout ecotypes from Laurentian Shield lakes, Canada. This was achieved by documenting the extent of morphological differentiation using geometric morphometrics and linear measurements as well as the pattern of genomic divergence by means of RADseq genotyping (3925 filtered SNPs) in 12 lakes. Our results indicate that the two ecotypes evolved distinct body shape and several linear measurements in parallel. Neutral genetic differentiation was pronounced between all isolated populations (Mean $F_{ST}$ = 0.433), indicating no or very limited migration and pronounced genetic drift. Significant genetic differentiation also suggested partial reproductive isolation between ecotypes in the two lakes where they are found in sympatry. Combining different outlier detection methods, we identified 48 SNPs putatively under divergent selection between ecotypes, among which 10 could be annotated and related to functions such as developmental processes and ionic regulation. Finally, our results indicate that parallel morphological divergence is accompanied by both parallel and nonparallel genomic divergence, which is associated with the use of different trophic niches between ecotypes. The results are also discussed in the context of management and conservation of this highly exploited species throughout northern North America.**

*Keywords*: adaptive divergence, conservation, genotyping by sequencing, parallel evolution, salmonids, speciation

*Received 18 December 2015; revision received 18 July 2016; accepted 4 August 2016*

## Introduction

A main source of adaptive phenotypic variation is trophic variability (Skúlason & Smith 1995; Smith & Skúlason 1996; Schluter 2000; Nosil 2012). As trophic adaptation is central to the ecological niche concept (MacArthur 1958), investigating its genetic basis is essential to better understand the capacity of a species to adjust to its ecological niche in a new environment (Savolainen *et al.* 2013). Studies of closely related have reported both cases of common genetic basis of independent phenotypic evolution (genetic parallelism; see Conte *et al.* 2012) (Colosimo *et al.* 2005; Foster *et al.* 2007; Hohenlohe *et al.* 2012; Laporte *et al.* 2015) and different genetic pathways underlying similar phenotypic changes (DeFaveri *et al.* 2011; Gagnaire *et al.* 2013; Elmer *et al.* 2014; Laporte *et al.* 2015).

Correspondence: S. Bernatchez, Fax: +1 418 656-7176;
E-mail: simon.bernatchez.1@ulaval.ca

Repeated and independent phenotypic evolution of populations from closely related lineages in similar environments (phenotypic parallelism; see Conte *et al.* 2012) is particularly prevalent in fishes of recently colonized postglacial lakes (Bernatchez & Dodson 1990; Schluter & McPhail 1992; Taylor 1999; Østbye *et al.* 2006; Noakes 2008; Laporte *et al.* 2011). This is generally due to the variation in the availability of trophic niches in these young ecosystems, an important factor believed to promote rapid adaptive divergence (Robinson & Wilson 1994; Smith & Skúlason 1996; Schluter 2000; Willacker *et al.* 2010). Due to their significant impact on foraging efficiency, evolution of locomotion-related traits is particularly important in fishes after the colonization of a new trophic niche (Webb 1982, 1984; Langerhans & Reznick 2010; Walker 2010). In particular, a deeper caudal peduncle allows for faster, more powerful burst swimming, which is advantageous to feed on evasive prey (Webb 1982, 1984; Walker 1997; Blake 2004). In contrast, a more fusiform body shape and a narrower caudal peduncle allow for more steady swimming by reducing drag (Webb 1982, 1984; Robinson & Wilson 1994; Walker 1997; Willacker *et al.* 2010). This is more advantageous for feeding on zooplankton, which often have patchy and scattered distributions within lakes (del Giorgio & Gasol 1995). In addition, gill raker counts often differ between such ecotypes, with planktivorous fish generally having more, and longer gill rakers, facilitating the retention of small prey in the buccal cavity during feeding (Schluter 1993; Taylor & Bentzen 1993; Kahilainen *et al.* 2011; Præbel *et al.* 2013; May-McNally *et al.* 2014). Variation in locomotion-related traits involved in foraging and habitat use in fish has been shown to have a genetic basis (Skúlason *et al.* 1989; Hatfield 1997; Rogers *et al.* 2002; Rogers & Bernatchez 2007; Laporte *et al.* 2015). However, more studies are needed in order to document the repeated evolution of these complex functional traits in other species and to evaluate the extent of parallel genetic changes underlying this phenotypic divergence.

Lake trout (*Salvelinus namaycush*) are predominantly piscivorous, and show a great extent of morphological variation, principally associated with prey consumption (Qadri 1967; Blackie *et al.* 2003; Harvey *et al.* 2003; Zimmerman *et al.* 2007, 2009). The best characterized ecotypes are found in the Laurentian Great Lakes and are referred to as lean, siscowet and humper. The ecotypes differ by their diet (e.g. prey composition), morphology (e.g. head and body shape) and behaviour [see Muir *et al.* (in press) for further details]. In Great Slave Lake, ecotypes similar to the lean and siscowet ecotypes have also been documented and show differences in depth habitats and diet (Zimmerman *et al.* 2006, 2009). Similarly, in Lake Mistassini, a deep-water humperlike and a shallow-water leanlike ecotype have been documented (Zimmerman *et al.* 2007). More recently, in Great Bear Lake, the occurrence of four shallow-water ecotypes has been reported (Chavarie *et al.* 2013). Finally, a distinct planktivorous ecotype that feeds almost exclusively on zooplankton has been reported in several Laurentian Shield lakes (Vander Zanden *et al.* 2000; Houde & Scrosati 2003). This ecotype can be found both in allopatry and in sympatry with the more common piscivorous ecotype. All lake trout populations inhabiting this area are considered to have originated postglacially (less than 15 000 YBP) from a single refugium (Atlantic refugium – Wilson & Hebert 1996, 1998). The presence of lake trout planktivorous ecotype is generally associated with the absence (or low occurrence) of cold-water pelagic forage species in lakes (Martin 1966; Matuszek *et al.* 1990; Houde & Scrosati 2003), suggesting a recent and repeated occurrence of this trophic ecotype. As such, this complex of piscivorous and planktivorous lake trout populations from the Laurentian Shield lakes represents another relevant study system to investigate the extent of genetic parallelism underlying parallel adaptive phenotypic evolution.

Lake trout is also considered one of the most important sport species where it is found, including both piscivorous and planktivorous populations (Legault *et al.* 2001; FOC 2010). This resulted in an overfishing leading to a decline of many populations. In Québec, lake trout stocking has been used for over 100 years and became a major management tool for the provincial government in the last 50 years (FOC 1886; see Valiquette *et al.* 2014 for further precisions). However, until now, only fish from piscivorous populations have been used in these stocking strategies and no study has been conducted on the adaptive divergence between piscivorous and planktivorous ecotypes.

In this context, the main goal of this study was to investigate the extent of genomic and phenotypic parallelism underlying the ecological divergence between piscivorous and planktivorous lake trout ecotypes from the Laurentian Shield lakes. First, we examined the extent of phenotypic divergence between piscivorous and planktivorous ecotypes from different lakes using geometric morphometrics (shape), traditional morphometrics (linear measurements) and the number of gill rakers. Second, we investigated both neutral and putative adaptive divergence between these two ecotypes, using 3925 quality-filtered RAD-SNPs. Third, when possible, we inferred the functions of genes potentially under potential selection towards the goal of identifying the molecular mechanisms underlying phenotypic differentiation of these ecotypes. Finally, we discuss the considerations of these results for management purposes.

## Material and methods

### *Study system and sampling*

Twelve lakes from the Laurentian Shield were sampled between 2007 and 2013 in southern Québec, Canada (Fig. 1; Table 1). These lakes were formed by drainage and isostatic rebound related to glacial retreat following the Wisconsinan glaciation (80 000–10 000 YBP) and were colonized by lake trout around 15 000 years ago (Mandrak & Crossman 1992; Bernatchez & Wilson 1998; Wilson & Mandrak 2004). Samples were collected using either gillnets or fishing rods, and from locations according to the classification of populations in the database of the Ministère des Forêts, de la Faune et des Parcs du Québec (MFFP; see also Houde & Scrosati 2003). The MFFP classification of lake trout populations (planktivorous vs. piscivorous) was based on fish growth rate (the planktivorous ecotype has a much slower growth rate and smaller size at sexual maturity), basic stomach content analysis (predominantly either zooplankton or fish) and the presence or absence of cold-water pelagic prey fishes in the lakes. Six lakes with allopatric piscivorous populations, four with allopatric planktivorous populations and two harbouring sympatric pairs were sampled. For the sympatric lakes, as planktivorous fish are known to reach a smaller maximum length (fork length (FL) <500 mm; Houde & Scrosati 2003), we classified the smaller mature fish (FL $\leq$ 435 mm) as planktivorous and the larger fish (FL $\geq$ 565 mm) as piscivorous (sex and maturity were determined as described in the following section 'Morphometric analyses'). No stocking has been recorded for ten of these lakes, but one stocking event (piscivorous fish) was conducted in Tee Lake in 2008 (but fish sampled in this lakes were older than this stocking event)

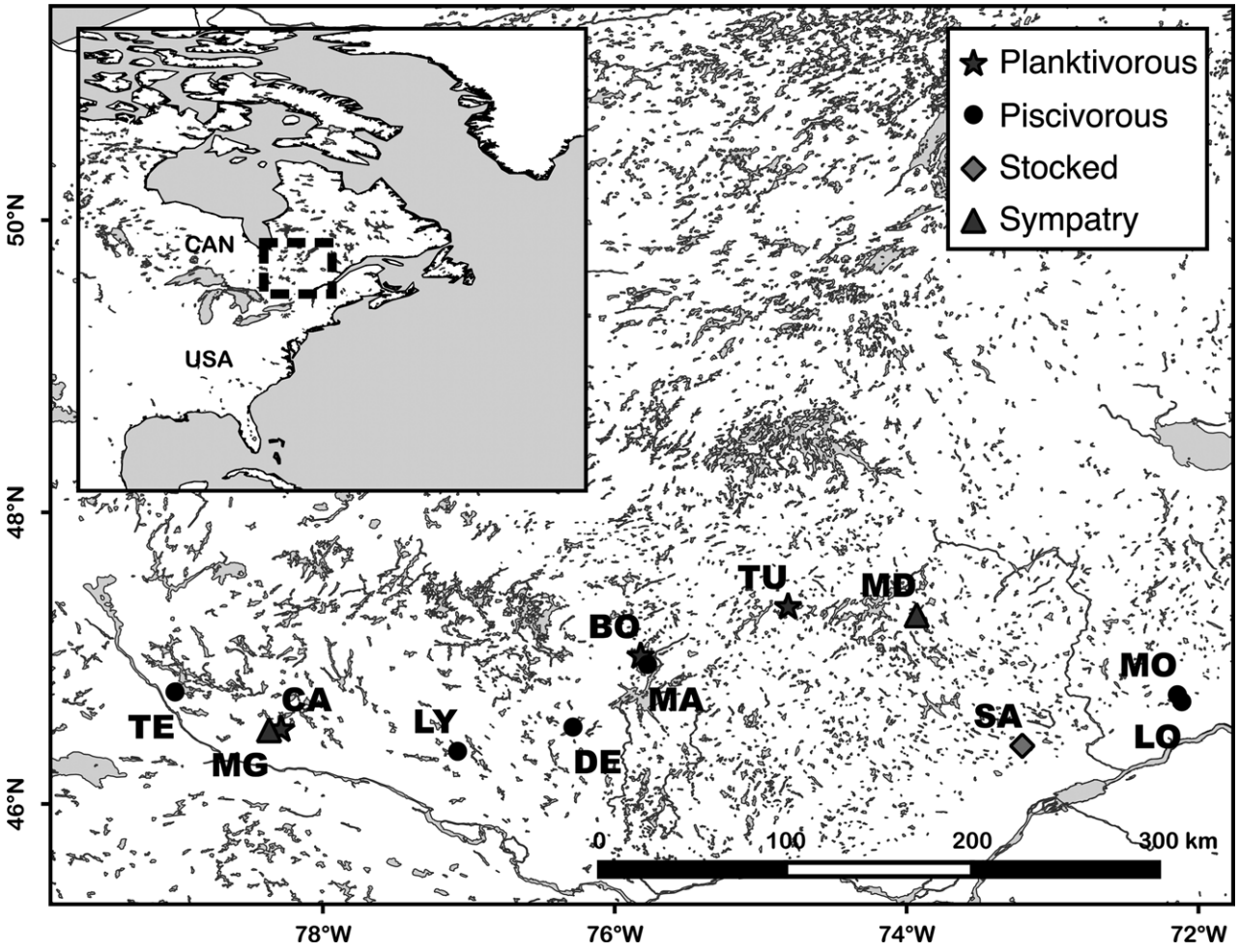


**Fig. 1** Map of studied lakes. BO: Bondy (planktivorous), CA: Caugnawana (planktivorous), DE: Désert (piscivorous), LY: Lynch (piscivorous), LO: Long (piscivorous), MA: Marguerite (piscivorous), MD: Mondonac (sympatric), MG: Maganasipi (sympatric), MO: Montauban (piscivorous), SA: Sacacomie (stocked), TE: Tee (piscivorous), TU: Turnbull (planktivorous).

**Table 1** Lake trout populations analysed in this study with corresponding ecotype, sample size (*n*) available for morphology and genetic analyses, observed heterozygosity ($H_O$), expected heterozygosity ($H_E$) and percentage of polymorphic SNPs

| Lake | Latitude | Longitude | Code | Ecotype | Morphology (*n*) | Genetics (*n*) | $H_O$ over the 3925 SNPs | $H_E$ over the 3925 SNPs | Percentage of polymorphic markers over the 3925 SNPs |
|---|---|---|---|---|---|---|---|---|---|
| Désert | 46.59111 | −76.30667 | DE | Piscivorous | 15 | 24 | 0.12 | 0.13 | 50 |
| Lynch | 46.41139 | −77.09528 | LY | Piscivorous | 17 | 24 | 0.13 | 0.13 | 50 |
| Long | 46.83778 | −72.13833 | LO | Piscivorous | N/A | 22 | 0.10 | 0.10 | 37 |
| Marguerite | 47.02861 | −75.80333 | MA | Piscivorous | 25* | 23 | 0.14 | 0.15 | 56 |
| Montauban | 46.88472 | −72.16917 | MO | Piscivorous | N/A | 17 | 0.09 | 0.10 | 30 |
| Tee[†] | 46.78417 | −79.03833 | TE | Piscivorous | N/A | 23 | 0.15 | 0.16 | 58 |
| Bondy | 47.08389 | −75.85222 | BO | Planktivorous | 22 | 22 | 0.09 | 0.11 | 39 |
| Caugnawana | 46.53972 | −78.30750 | CA | Planktivorous | 10 | 23 | 0.08 | 0.10 | 33 |
| Turnbull | 47.43889 | −74.84667 | TU | Planktivorous | 18 | 24 | 0.08 | 0.09 | 34 |
| Maganasipi | 46.53417 | −78.38972 | MG-1 | Piscivorous | N/A | 21 | 0.07 | 0.07 | 29 |
| | | | MG-2 | Planktivorous | N/A | 25 | 0.08 | 0.09 | 32 |
| Mondonac | 47.39917 | −73.96528 | MD-1 | Piscivorous | N/A | 24 | 0.09 | 0.11 | 40 |
| | | | MD-2 | Planktivorous | N/A | 24 | 0.08 | 0.08 | 30 |
| Sacacomie[‡] | 46.51639 | −73.22528 | SA | Unknown | N/A | 24 | 0.16 | 0.18 | 70 |

*Twenty-five individuals have been used for the geometric morphometrics analysis and 23 individuals for the other morphometric analyses.
[†]One stocking episode (fish from a piscivorous population).
[‡]Six stocking episodes (fish from four piscivorous populations).

and six events (piscivorous fish) in Sacacomie Lake (between 1981 and 1987) (MFFP unpubl. database). One of the lakes containing sympatric ecotypes, Mondonac Lake, is a hydroelectric reservoir with two lake trout populations that became in physical contact in 1944 (Benoît *et al.* 1997). Finally, Sacacomie Lake harbours a population now classified as planktivorous, but that was classified as piscivorous before the crash of the main forage fish (rainbow smelt) in the 1970s (Stéphanie Gagné, MFFP, personal communication). Nonlethal fin clips were obtained for a total of 339 fish (Table 1).

### *Morphometric analyses*

For a subgroup of mature fish ($n = 107$, see Fig. S1, Supporting information), a lateral photograph of the left side of each fish was taken using a digital camera (Nikon Coolpix P7700) fixed on a tripod using a focal length of 50 mm, following the recommendations of Muir *et al.* (2012a). Individuals were positioned in a mesh cradle to minimize the distortion effect caused by the curvature of a fish expanded on a plane surface (Zimmerman *et al.* 2006). All fins were extended, and the anal and dorsal fins were pinned. FL (mm) and weight (mg) were recorded, sex was determined by the examination of gonads, and stage of maturity was evaluated using classification from Nikolsky (1963). For mature specimens, the first left gill arch was collected and stored in a 10% formaldehyde solution. Photographs of the gill arch were taken using a stereo binocular microscope to count the number of gill rakers and measure the length of the longest gill raker. Twenty homologous landmarks (Fig. S2a, Supporting information) were positioned on the images using TPSDIG2 version 2.16 software (http://life.bio.sunysb.edu/morph/) to compare fish shape (Zelditch *et al.* 2012). Landmarks were chosen according to previous studies on lake trout (Zimmerman *et al.* 2006, 2007, 2009; Bronte & Moore 2007; Chavarie *et al.* 2013) and/or due to their functional importance in locomotion and foraging (Webb 1984, 1986a, b). Geometric morphometric analyses were conducted using the software MORPHOJ version 1.06c (Klingenberg 2011). We first produced a Procrustes fit to eliminate variation caused by differences in size and orientation of the fish. A principal components analysis (PCA) was then performed on 107 fish to represent fish shape variation. Using the score of informative principal components (PCs) (informative axes were identified based on a broken-stick distribution; see Legendre & Legendre 1998 for more details) as dependant variables, we conducted a multivariate analysis of variance (MANOVA) with 'lake' nested in the variable 'ecotype', and with 'sex' used as a cofactor. As the 'ecotype by sex' interaction had a significant effect on fish shape variation (see 'Results'), subsequent shape analyses were performed separately for each sex. Thereafter, discriminant function analyses (DFAs; 10 000 permutations) on Procrustes fit were carried out to verify

whether ecotypes could be statistically distinguished based on fish shape (Procrustes distances and parametric *T*-square statistic). Superposition of the mean shape for each condition was performed to visualize the shape difference between ecotypes for each sex.

For each fish, Fulton's condition index was calculated (Neumann *et al.* 2012), the total number of gill rakers on the left gill arch was counted, and the length and width of the longest gill raker were measured using TPSDIG2 version 2.16. In addition, 15 conventional linear measurements (Fig. S2b, Supporting information) were extracted using the software TPSDIG2 version 2.16 and TMORPHGEN6C (IMP software, http://www3.canisius.edu/~sheets/moremorph.html). Linear measurements were also chosen according to previous studies on lake trout and their importance in locomotion and foraging (Zimmerman *et al.* 2006; Chavarie *et al.* 2013). For each linear measurement, we first tested for a significant relationship with fish size. When relationship was present, we used residuals of this relationship for subsequent analyses to control for the effect of fish size on the linear measurements. Nested two-way analysis of variance (ANOVA), with 'lake' nested within 'ecotype', and 'sex' as cofactor was then used to test for significant differences between ecotypes using R software version 3.1.1 (R Development Core Team 2014). Traits with a significant ($\alpha < 0.05$) effect of sex were separated accordingly in a subsequent analysis. False discovery rate adjustment (FDR; $\alpha < 0.05$) was used to control for multiple tests in the nested ANOVAs.

### Genomic analyses

*Library preparation and sequencing*. Genomic DNA was extracted from fin clips using a modified version of the Aljanabi & Martinez (1997) salt extraction protocol (Supplementary file 1). DNA quality and quantification were assessed using agarose gel electrophoresis, NanoDrop spectrophotometer (Thermo Scientific and Quant-iT Picogreen dsDNA Assay Kit (Invitrogen). Genotyping-by-sequencing (GBS) libraries were prepared using a modified version of the Poland *et al.* (2012) two-enzyme (*PstI* and *MspI*) GBS protocol (Supplementary file 2). Single-read 100-bp sequencing was performed on Illumina HiSeq2000 platform at the Genome Quebec Innovation Centre (McGill University, Montreal, Canada).

*Bioinformatics and quality filtering*. Raw sequence data quality was analysed using FASTQC version 0.11.3 (www.bioinformatics.babraham.ac.uk/projects/fastqc/). STACKS software version 1.21 (Catchen *et al.* 2011) was used to identify loci and call genotypes. The libraries were demultiplexed and filtered for overall quality using *process_radtags*. Reads were trimmed to 80 bp to remove barcodes and low-quality bases at the 3′-ends of reads due to increased sequencing error rate (Minoche *et al.* 2011). The formation of RAD loci was allowed with a maximum of two nucleotide mismatches ($M = 2$) among primary reads, and the minimum stacks depth was set to five ($m = 5$). A maximum of one nucleotide difference ($n = 1$) was allowed during the elaboration of the stacks catalogue using *cstacks*. The *population* module was then used to call genotypes. Subsequent steps were applied to filter for quality given the known occurrence of incomplete individual genotypes in RAD-seq data sets. Specifically, a SNP was retained if it met all of the following criteria: (i) genotypes available in a least 66% of the individuals within a population (lake) for at least 75% of the populations (i.e. 9/12); (ii) global minor allele frequency (MAF) ≥0.02 or a local MAF ≥0.10 in a least one population; (iii) heterozygosity ≤0.50 in a least two-thirds of the populations in which two-thirds of the individuals were genotyped; (iv) $F_{IS}$ between −0.15 and 0.30 in at least two-thirds of the populations in which two-thirds of the individuals were genotyped; (v) coverage between 6 and 50× per individual. Only the first SNP of each locus was retained for the analysis. Details on the number of SNPs remaining after each filtering step are shown in Table S1 (Supporting information). Filtering and conversion of the VCF file were performed in R and PYTHON software version 2.7.6 (http://www.python.org/) scripts as well as PGDSPIDER software version 2.0.7.1 (Lischer & Excoffier 2012), VCFTOOLS software version 0.1.11 (Danecek *et al.* 2011) and PLINK software version 1.07 (http://pngu.mgh.harvard.edu/purcell/plink/; Purcell *et al.* 2007).

*Population genetic diversity and structure.* GENODIVE software version 2.0b27 (Meirmans & Van Tienderen 2004) was used to calculate observed ($H_O$) and expected heterozygosity ($H_E$), allele frequencies and counts for the global quality filtered set of loci. The proportion of polymorphic SNPs was calculated for each population (or lake for LO and MO) and ecotype. Hereafter, we refer to allopatric ecotypes of lake trout inhabiting each lake, or sympatric ecotypes within a lake, as a population. Equality of variance between ecotypes for the proportion of polymorphic SNPs was evaluated in R using a Fisher's test. As variances were unequal (Fisher's test: $P = 0.032$), a Welch's *t*-test was used to determine whether there was a significant difference between the means proportion of polymorphic SNPs between the two ecotypes. Allele frequencies were used to calculate mean allele frequency differences between populations of piscivorous and planktivorous ecotypes. Sacacomie Lake was excluded from the analysis because of its

recent ecotype change. ADMIXTURE software version 1.23 (Alexander *et al.* 2009) was used with all loci to determine the number of putative genetic clusters among the 12 lakes (testing $K = 1$ to $K = 15$). If more than one cluster was identified within the two sympatric lakes, clusters were considered as separated groups in further analyses. Pairwise estimates of $F_{ST}$ (Weir & Cockerham 1984 estimator) among populations were obtained using GENODIVE (10 000 permutations). Bootstrapped (1000 bootstraps) neighbor-joining (NJ) phylogenetic matrices were calculated based on $F_{ST}$ values using TREEFIT version 1.2 (Kalinowski 2009). A tree was built with FIGTREE version 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/), and its degree of fit ($R^2$) to the $F_{ST}$ matrix was evaluated in TREEFIT (see Kalinowski 2009). Finally, a hierarchical analysis of molecular variance (AMOVA; 10 000 permutations) (Excoffier *et al.* 1992) was performed for all loci in GENODIVE using ecotype as grouping.

*Discovery of SNPs putatively under divergent selection.* To identify loci putatively under divergent selection between ecotypes, we used two types of approaches. For this we excluded the Sacacomie population, which was stocked intensively. Two genome scan methods were used to detect the presence of outlier loci putatively under divergent selection. First, we used BAYESCAN version 2.1 (Foll & Gaggiotti 2008), specifying a 'prior' odd of 10 000 (Lotterhos & Whitlock 2014; Benestan *et al.* 2015) to minimize type I errors, and after 200 000 iterations, a locus with $\log_{10}$ (Bayes factor) >2 was considered under divergent selection according to the Jeffrey's scale of evidence for Bayes factors (Jeffrey 1961). Second, we used OUTFLANK (Whitlock & Lotterhos 2015) using the default settings.

We then applied two gene–environment association software that controls for strong population structure (Frichot *et al.* 2013; Günther & Coop 2013; Lotterhos & Whitlock 2015). We first used BAYENV2 software (Coop *et al.* 2010; Günther & Coop 2013) that tests for covariance between SNPs allele frequencies and explanatory variables. A covariance matrix was estimated between populations using allele frequencies for all loci to describe covariation across populations and to avoid population-specific effects. The variable tested was the piscivorous or planktivorous status of populations (set to 0 and 1, respectively). The average of five independent runs (100 000 iterations) was used to estimate Bayes factor and Spearman's Rho (absolute values) results (Blair *et al.* 2014). As suggested by Günther & Coop (2013), SNPs from the intersection of the tail of Bayes factor's distribution (top 5%) and the tail of Spearman's Rho distribution (top 5%) were considered to have allele frequencies correlated with trophic status. Gene–environment associations were also tested using LFMM version 1.4, which allows for the inference of background levels of population structure (Frichot *et al.* 2013). Based on the ADMIXTURE analysis (see 'Results'), $K = 13$ was used in the LFMM analysis. FDR adjustment ($\alpha < 0.05$) was used on the $|z|$-scores distribution to control for multiple tests, as suggested in Frichot & François (2015). SNPs with adjusted $P$-values < 0.05 were considered to have allele frequencies correlated with trophic status. Finally, we selected SNPs identified as being putatively under divergent selection between ecotypes using both methods.

To test whether loci putatively under divergent selection between ecotypes were found just by chance, we performed a randomization of the ecotype variable in LFMM. Here, individual genotypes were conserved, but individuals were randomly assigned to one of the two ecotypes. After the randomization step, the LFMM analysis was run exactly as it was in the previous analyses. This procedure was replicated ten times. We did not perform a randomization step in BAYENV2 because the identification of the loci putatively under divergent selection relies on the intersection of the tail of Bayes Factor's distribution (top 5%) and the tail of Spearman's Rho distribution (top 5%). Thus, it would still result in a set of SNPs considered to have allele frequencies correlated with the random ecotypes.

Differences in mean allele frequencies between piscivorous and planktivorous populations (excluding Sacacomie Lake) were calculated for the following sets of markers: (i) all SNPs; (ii) SNPs that were identified either by BAYENV2; or (iii) by LFMM; and (iv) SNPs that were identified by both methods (i.e. the intersection set). These four sets of SNPs were also used in a hierarchical AMOVA using ecotype as grouping to evaluate which SNP discovery method (or combination of methods) would be most efficient at differentiating ecotypes. SNPs from the intersection were also used to build a bootstrapped (1000) NJ phylogenetic tree using $F_{ST}$ values to test whether populations of a same ecotype grouped together.

Finally, to test whether the potentially selected loci were selected in parallel between populations, we performed a 'jackknife-like' procedure using LFMM as described above, but by excluding one population at a time. This procedure was repeated for all populations and, each time, the resulting loci that were significantly correlated with ecotype were used to build a phylogenetic tree (neighbor-joining, $F_{ST}$) to observe whether the population that was excluded from marker selection grouped correctly to its ecotype cluster using this set of loci.

*Gene ontology.* Finally, nucleotide sequences (80 bp) containing SNPs putatively under selection were used for a BLAST query (blastn) against the rainbow trout

(*Oncorhynchus mykiss*) genome database (Berthelot *et al.* 2014). Only hits with *E*-value < $10^{-10}$ and sequences with three or fewer *E*-values of a similar level were considered. For a given sequence, if a single *E*-value was relatively higher than others, the hit was considered to be the most probable association and was kept. If a sequence had three or fewer hits with similarly high *E*-values, all of these hits were kept and considered as probable associations. Sequences were then annotated according Berthelot *et al.* (2014) and using UniProt and Ensembl Gene Ontology database.

## Results

### *Morphometric analyses*

According to the broken-stick distribution, five PCs were retained to represent body shape, collectively representing 80.5% of the shape variation. The effect of ecotype was highly significant ($P < 0.001$), as was the interaction between ecotype and sex ($P = 0.040$). DFAs produced for each sex independently confirmed body shape differentiation between the two ecotypes via both the Procrustes distance method (males: 0.02, $P < 0.001$; females: 0.03, $P < 0.001$; Fig. 2a, b) and the Hotelling's $T^2$ method (males: 828.39, $P < 0.001$; females: 1021.85, $P = 0.018$; Fig. 2a, b). In addition, no overlap between ecotypes was observed in the distribution of DFAs scores (Fig. 2a, b). Finally, mean body shape comparisons showed similar shape differentiation between planktivorous and piscivorous lake trout for both sexes (Fig. 2c, d). In both sexes, the shape differentiation is mostly located in the abdominal and caudal regions, with the piscivorous ecotype having a relative deeper body and caudal region. Planktivorous lake trout also have a relatively larger eye than the piscivorous individuals (Fig. 2c, d).

The interaction between ecotype and sex observed in the MANOVA was explored with two DFAs to separate sexes within each ecotype. DFAs between sexes within ecotypes suggested body shape sexual dimorphism in the planktivorous ecotype only (Hotelling's $T^2$: piscivorous: 60.78, $P = 0.903$; planktivorous: 330.07, $P = 0.043$; Procrustes distance: piscivorous: 0.01, $P = 0.523$; planktivorous: 0.01, $P = 0.051$; Fig. S3, Supporting information). Planktivorous males had a slightly deeper body and a more pointed and longer head than females.

After FDR correction ($\alpha < 0.05$), four of the 19 other morphometric variables were significantly different between ecotypes for both sexes, and two were significantly different for one sex only (Fig. 3). Piscivorous lake trout had longer body ($P < 0.001$; Fig. 3) and caudal fin ($P = 0.001$; Fig. 3), fewer gill rakers ($P = 0.019$; Fig. 3), and a higher condition index ($P < 0.001$; Fig. 3) than planktivorous lake trout. Female piscivorous lake trout had a deeper caudal peduncle ($P = 0.019$; Fig. 3) than female planktivorous lake trout. Male planktivorous lake trout had a longer head ($P = 0.019$; Fig. 3) than male piscivorous lake trout.

### *Genomic analysis*

*Genotyping by sequencing and SNPs discovery.* The total number of raw reads obtained was 4061 220 028, and

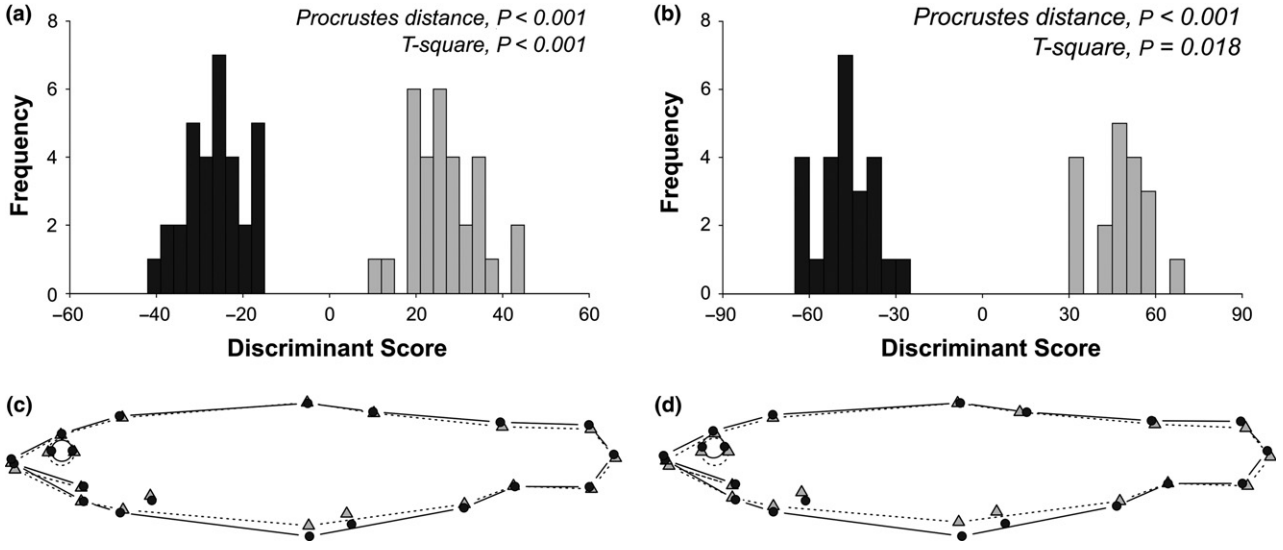


**Fig. 2** Histogram of discriminant scores for (a) male and (b) female planktivorous (light grey) and piscivorous (dark grey) lake trout. Mean shape differences between planktivorous (hatched line, light grey triangles) and piscivorous (continuous line, dark grey circles) for male (c) and female (d) lake trout (scale is magnified by 2× to emphasize differences between ecotypes).

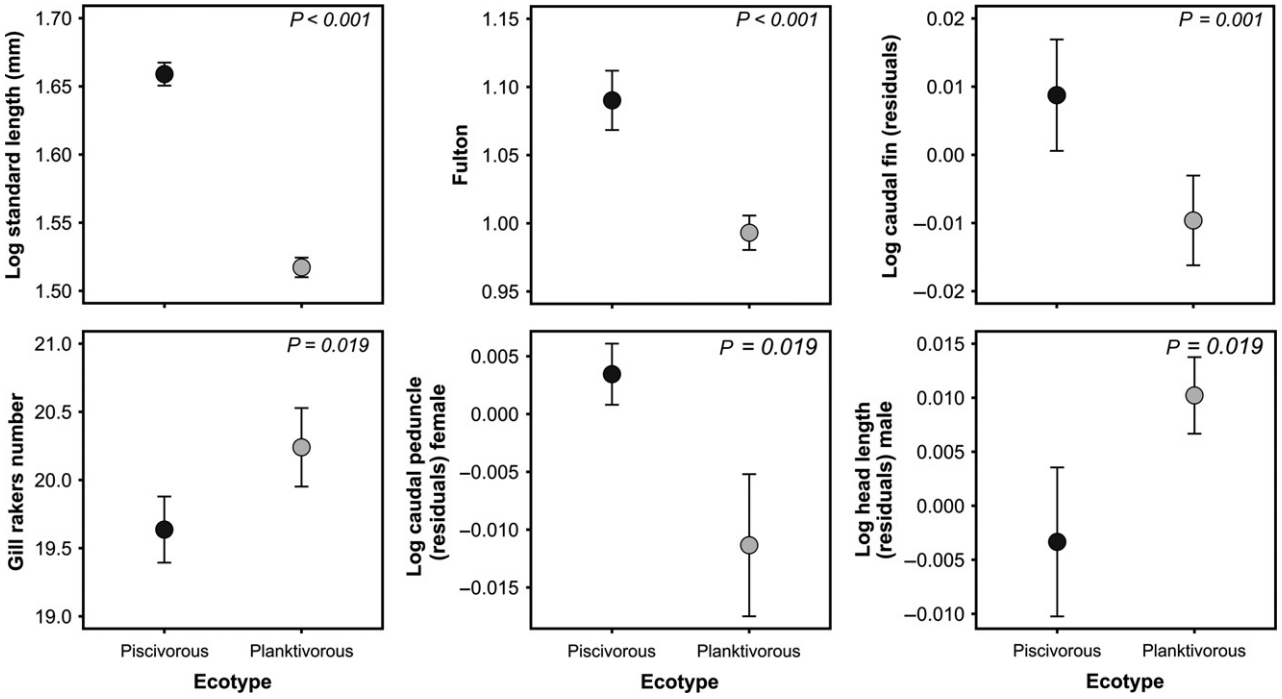


**Fig. 3** Mean (±SE of the mean) of the significantly different phenotypic traits between piscivorous (dark grey) and planktivorous (light grey) lake trout ecotypes. Values were obtained from nested analysis of variance (ANOVA) with the variable 'lake' nested in the variable 'ecotype'. Traits not illustrated here were not statistically different between ecotypes.

the average number of filtered reads per individual was 3172 828. A total of 320 individuals (see Fig. S1, Supporting information) were retained for genomic analyses (Table 1). A total of 4968 SNPs successfully passed quality filters and 3925 SNPs (79.0%) were retained after keeping only one SNP per locus (Table S1, Supporting information).

*Genetic population structure and diversity analysis.* The cluster analysis revealed that the $K = 14$ model was the most likely number of genetically distinct populations. After visualizing the results, $K = 13$ (Fig. S4, Supporting information) was determined to be the most descriptive model taking into account the biological context of the system whereby all allopatric populations formed unique genetic clusters except in LO and MO, the two highly connected lakes, which grouped in the same genetic cluster. Individuals from the stocked lake (SA) formed a cluster with a mixture of several other genetic clusters in different proportions. Both sympatric lakes (MD and MG) formed two genetic clusters within each lake.

Mean observed heterozygosity ($H_O$) and mean expected heterozygosity ($H_E$) were evaluated for each genetic cluster (but with LO and MO kept separate) using all loci, and ranged from 0.07 to 0.16 and from 0.07 to 0.18, respectively (Table 1). Populations showed a very high level of allele fixation with 30–71% of the loci fixed within each population (Table 1). The proportion of polymorphic SNPs were significantly higher in piscivorous populations (44%) than in planktivorous populations (34%) (Welch's *t*-test: $P = 0.039$). The stocked Sacacomie Lake population had the highest proportion of polymorphic SNPs (70%; Table 1).

Pairwise estimates of $F_{ST}$ (3925 SNPs) between populations ranged from 0.028 (the two connected lakes LO and MO) to 0.659 (mean = 0.433), and all comparisons were highly significant ($P < 0.001$; Table 2). Results from the AMOVA using all loci showed that ecotype did not explain any variance (0.0%, $P = 0.434$; Table 3). Accordingly, the NJ tree using $F_{ST}$ calculated with all 3925 loci ($R^2 = 0.995$, Fig. 4a) showed that populations could not be sorted by ecotypes. Instead, results revealed a pronounced divergence among allopatric populations (and sympatric ecotypes), regardless of ecotypes.

*Identification of loci putatively under selection.* BAYESCAN identified two linked SNPs (Pearson's $r = 0.9997$) with very high $F_{ST}$ values as being under divergent selection between ecotypes. These SNPs were fixed in ten of the populations, and allele frequency was >0.71 for one allele in the three others. Although the same locus showed high $F_{ST}$ in multiple population comparisons, the fixed (or nearly fixed) allele is not always the same for each population within a single ecotype. OUTFLANK

**Table 2** Fixation index ($F_{ST}$) calculated for all 3925 SNPs between piscivorous and planktivorous lake trout. Comparisons among planktivorous populations are outlined in black, comparisons among piscivorous populations are highlighted in light grey, and comparisons between ecotypes are highlighted in dark grey

| Populations | BO | CA | TU | MD-2 | MG-2 | DE | LY | MA | LO | MO | TE | MD-1 | MG-1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CA | 0.502 | — | | | | | | | | | | | |
| TU | 0.500 | 0.571 | — | | | | | | | | | | |
| MD-2 | 0.544 | 0.603 | 0.475 | — | | | | | | | | | |
| MG-2 | 0.526 | 0.251 | 0.592 | 0.618 | — | | | | | | | | |
| DE | 0.395 | 0.445 | 0.439 | 0.487 | 0.465 | — | | | | | | | |
| LY | 0.417 | 0.354 | 0.470 | 0.515 | 0.373 | 0.344 | — | | | | | | |
| MA | 0.345 | 0.416 | 0.393 | 0.436 | 0.438 | 0.266 | 0.317 | — | | | | | |
| LO | 0.518 | 0.540 | 0.561 | 0.592 | 0.561 | 0.451 | 0.469 | 0.419 | — | | | | |
| MO | 0.519 | 0.553 | 0.559 | 0.595 | 0.570 | 0.443 | 0.467 | 0.411 | 0.028 | — | | | |
| TE | 0.394 | 0.399 | 0.443 | 0.472 | 0.425 | 0.314 | 0.325 | 0.279 | 0.427 | 0.419 | — | | |
| MD-1 | 0.472 | 0.522 | 0.402 | 0.382 | 0.546 | 0.387 | 0.425 | 0.346 | 0.509 | 0.505 | 0.391 | — | |
| MG-1 | 0.568 | 0.333 | 0.635 | 0.659 | 0.136 | 0.505 | 0.426 | 0.476 | 0.597 | 0.609 | 0.458 | 0.587 | — |
| SA | 0.287 | 0.329 | 0.350 | 0.385 | 0.355 | 0.196 | 0.243 | 0.162 | 0.337 | 0.324 | 0.200 | 0.290 | 0.399 |

**Table 3** Analysis of molecular variance (AMOVA) among planktivorous and piscivorous lake trout ecotype populations using four sets of SNPs (all SNPs, SNPs identified as associated with ecotypes by BAYENV2, by LFMM and by both methods). The numbers in parentheses indicate the number of SNPs for each set

| Set of SNPs | Source of variation | Percentage of variation | $F_{statistic}$ | *P*-value |
|---|---|---|---|---|
| All SNPs (3925) | Among ecotypes | 0.0 | 0.000 | 0.434 |
| | Among populations within ecotypes | 29.9 | 0.299 | <0.001 |
| | Among individuals within population | 70.2 | — | — |
| BAYENV2 (99) | Among ecotypes | 12.1 | 0.121 | 0.003 |
| | Among populations within ecotypes | 26.1 | 0.297 | <0.001 |
| | Among individuals within population | 61.8 | — | — |
| LFMM (191) | Among ecotypes | 12.4 | 0.124 | 0.001 |
| | Among populations within ecotypes | 27.7 | 0.316 | <0.001 |
| | Among individuals within population | 59.9 | — | — |
| BAYENV2∩LFMM (48) | Among ecotypes | 19.2 | 0.192 | 0.001 |
| | Among populations within ecotypes | 24.6 | 0.305 | <0.001 |
| | Among individuals within population | 56.1 | — | — |

identified no loci potentially under divergent selection (data not shown). Given these results, we concluded that BAYESCAN and OUTFLANK were not appropriate to detect loci putatively under selection between ecotypes in the studied system, perhaps due to very pronounced genetic drift. In contrast, SNP–environment association analyses identified 242 putatively selected. Of these, 51 were identified exclusively using BAYENV2, 143 using LFMM and 48 were common to both methods (Fig. S5, Supporting information). To reduce false positives, we only consider these 48 SNPs as candidates being under selection between piscivorous and planktivorous ecotypes. Randomized samples (see 'Methods') did not result in any locus being correlated with ecotype in each replicated trial using LFMM.

The intersection set of SNPs identified by LFMM and BAYENV2 had the highest mean allele frequency difference between allopatric ecotype populations relative to either subset alone (e.g. only LFMM) or the entire set of SNPs and was significantly different from all other set of SNPs (BAYENV2∩LFMM-all SNPs: $P < 0.001$; BAYENV2∩LFMM-BAYENV2: $P = 0.017$; BAYENV2∩LFMM-LFMM: $P = 0.016$). Thus, SNPs common to the BAYENV2 and LFMM analyses are the most differentiated loci between ecotypes. Accordingly, the NJ tree using $F_{ST}$ calculated with loci common to the BAYENV2 and LFMM analyses ($R^2 = 0.932$, Fig. 4b) showed a clearer distinction between ecotypes than the NJ tree constructed using all loci. Natural populations clustered by ecotype except for the piscivorous ecotype from MG Lake that clustered with the planktivorous populations. The Sacacomie population (stocked) clustered with the piscivorous populations. The combination of the 'jackknife-like' procedure using LFMM and the classification step using a neighbor-joining ($F_{ST}$) phylogenetic tree led to different results depending on the population that was left out for selecting the markers and then reclassified

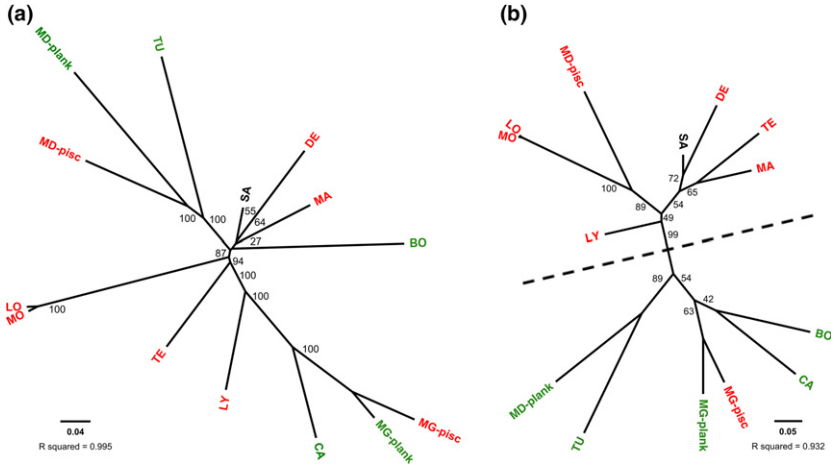


**Fig. 4** (a) Neighbor-joining tree constructed with $F_{ST}$ pairwise values among the 14 lake trout groups based on 3925 markers with bootstrap values based on 1000 replicates, (b) Neighbor-joining tree constructed with $F_{ST}$ pairwise values among the 14 lake trout groups based on 48 potentially selected markers detected using BAYENV2 and LFMM with bootstrap values based on 1000 replicates. Piscivorous populations are presented in red, planktivorous populations in green, and the stocked population in black. The dashed line shows the major split between planktivorous and piscivorous ecotypes in panel (b).

using these. Yet, this procedure resulted in the proper classification of seven allopatric populations out of nine, suggesting that markers selected by LFMM independently of a given population could efficiently reclassify it in its proper ecotype grouping. However, sympatric ecotypes did not classify clearly to their ecotype grouping (Fig. S6, S7 and S8, Supporting information).

Results from the AMOVA using SNPs identified by BAYENV2, LFMM and the intersection between both methods showed that the genetic differentiation among ecotypes was considerably higher than the value obtained with all loci (0.0%) with a proportion of 12.1% ($P$ = 0.003; Table 3), 12.5% ($P$ = 0.001; Table 3) and 19.2% ($P$ = 0.001; Table 3), respectively.

*Gene ontology*. The BLAST analysis of the 48 putatively selected loci against the rainbow trout genome database identified 10 loci ($E$-value < $10^{-10}$) that were associated with 12 scaffolds and 11 annotated genes. Of these, five were located directly in genes (two were nonsynonymous substitutions), and four were located between 1836 and 162 123 bp (mean = 32 549 bp) from a gene (Table 4). These include genes for proteins and transcription factors that are related to cell structure, cell differentiation, development, ionic regulation, neurotransmission and other functions (Table 4).

## Discussion

Our main goal was to investigate the extent of genomic and phenotypic parallelism between piscivorous and planktivorous lake trout ecotypes from Laurentian Shield lakes. This was achieved by documenting the pattern of morphological differentiation using geometric morphometrics and linear measurement, as well as the pattern of genomic divergence by means of RADseq genotyping. The repeated evolution of traits in environments with similar selection pressure (i.e. piscivory and planktivory in this study) is considered as strong evidence for natural selection because it is less likely that genetic drift will lead to shifts in the same direction (Endler 1986; Schluter & Nagel 1995). Also, it is quite clear, both from the mosaic geographic distribution of both ecotypes and the NJ tree built using all loci, that it is highly unlikely that both ecotypes originated from two different ancestral populations. Consequently, our results provide evidence for the parallel evolution of morphotypes associated with the use of distinct trophic niches in terms of locomotion and foraging ability, thus supporting the hypothesis of their adaptive phenotypic divergence. Moreover, despite the high level of genetic differentiation among isolated populations of the studied system, we conservatively identified 48 SNPs putatively under divergent selection between ecotypes and that differentiate essentially all planktivorous from piscivorous populations. These SNPs thus represent the most probable shared genetic basis underlying the ecotypes divergence within our data set, and the fact that the randomization procedure did not result in any SNP being correlated with the ecotypes gives further credibility to our results. The noncircularity of marker selection and correct ecotypic classification was further supported by the jackknife-like procedure whereby seven of nine allopatric populations were independently reclassified correctly in their ecotype cluster.

**Table 4** BLAST hits from sequences containing a SNP found to be putatively correlated with lake trout planktivorous and piscivorous ecotypes using BAYENV2 and LFMM. SNPs are identified by locus ID. Rainbow trout (*Oncorhynchus mykiss*) scaffolds, transcripts and gene names are identified as in Berthelot *et al.* (2014). Statistical significance of the hits is represented by *E*-values. Location (in a gene or not) of the SNPs and distance from the closer gene are indicated. For SNPs located in genes, nature of the nucleotide substitution is indicated

| Locus ID | Scaffold name | Transcript name | *E*-value | In a gene | Nonsynonymous substitution | Distance from gene (bp) | Gene | Principal known functions |
|---|---|---|---|---|---|---|---|---|
| 21160 | scaffold_151 | GSONMT00067876001 | 1E−30 | No | NA | 2120 | Transmembrane protein 150b-like | Membrane protein (Berthelot *et al.* 2014) |
| 31770 | scaffold_614 | GSONMT00044556001 | 3E−31 | No | NA | 5351 | Tubulin-specific chaperone cofactor e-like isoform x1 | Microtubule related (Tian *et al.* 1996; Berthelot *et al.* 2014) |
| 38564 | scaffold_73 | GSONMT00080888001 | 3E−31 | YES | No | N/A | Kinesin-like protein kif26b-like isoform x1 | Development (Uchiyama *et al.* 2010; Terabayashi *et al.* 2012;) |
| 58893* | scaffold_890 | GSONMT00005556001 | 2E−29 | Yes | NA | N/A | Natriuretic peptides a | Cardiovascular hormone (Berthelot *et al.* 2014; De Vito 2014) |
| 67205 | scaffold_1313 | GSONMT00061962001 | 4E−25 | Yes | Yes | N/A | V-type proton ATPase 116-kda subunit a isoform 1-like isoform ×2 | Proton transport, vesicle transport (Berthelot *et al.* 2014) |
| 112391 | scaffold_1739 | GSONMT00037930001 | 3E−31 | No | NA | 10 330 | Contactin-associated 4-like | Neurotransmission, dopaminergic and GABAergic systems, cell adhesion (Berthelot *et al.* 2014; Karayannis *et al.* 2014) |
| | scaffold_1439 | GSONMT00031004001 | 3E−31 | Yes | No | N/A | | |
| 127974 | scaffold_93 | GSONMT00082429001 | 2E−34 | Yes | Yes | N/A | Microtubule-associated serine –threonine protein kinase 4-like isoform ×3 | Phosphorylation (Berthelot *et al.* 2014) |
| | scaffold_157 | GSONMT00068174001 | 7E−23 | Yes | Yes | N/A | Microtubule-associated serine –threonine protein kinase 4-like isoform ×1 | |
| 132231 | scaffold_7069 | GSONMT00032647001 | 2E−19 | No | NA | 13 532 | Solute carrier family 35 member g2-like | |
| 134185 | scaffold_54 | GSONMT00079394001 | 1E−25 | No | NA | 162 123 | Transcription factor sox-11 | Transcription factor, development (Berthelot *et al.* 2014; Tsurusaki *et al.* 2014) |
| 143595 | scaffold_217 | GSONMT00072016001 | 6E−24 | No | NA | 1836 | Homeobox protein aristaless-like 4 | Transcription factor, development (Kayserili *et al.* 2009; Berthelot *et al.* 2014) |

*SNP is located in an intron.

### *Morphological evidence of two trophic ecotypes*

Several lines of evidence confirm the existence of two lake trout ecotypes which were clearly differentiated in terms of condition index, body shape, body size, caudal fin length, caudal peduncle depth (females only), head length (males only) and gill raker number. This differentiation among independent populations of the different ecotypes fits the prediction of resource polymorphism among planktivorous and piscivorous lake trout ecotypes, further suggesting that this pattern of differentiation is adaptive. Thus, planktivorous fishes are typically predicted to have a more streamlined body and a narrower caudal region, which has been shown to result from adaptations to steady pelagic swimming by reducing drag and increasing hydrodynamics (Webb 1982, 1984; Robinson & Wilson 1994; Willacker *et al.* 2010). In fact, planktivorous fish often swim long distances to find productive patches of zooplankton, and thus usually swim more than piscivorous fishes to catch their prey (Webb 1984; del Giorgio & Gasol 1995; Walker 1997; Willacker *et al.* 2010). In accordance with these expectations, planktivorous lake trout exhibit a more slender body shape, a narrower caudal region and a smaller caudal fin relative to the piscivorous fish (body shape and Fulton's condition index). In contrast, a deeper caudal area is predicted to provide a more powerful acceleration burst, a characteristic required to feed on evasive preylike fish (Webb 1982, 1984). In addition, female piscivorous lake trout had a relatively deeper caudal peduncle compared to female planktivorous lake trout. As mentioned for body shape, such differentiation is consistent with morphology linked to burst swimming and steady swimming, respectively, and thus relatable to foraging tactics (Langerhans & Reznick 2010). Male planktivorous Lake Trout also have a longer head relative to male piscivorous lake trout. Despite the fact that head characteristics have often been linked to foraging efficiency, the potential functionality of the relatively longer head of male planktivorous lake trout is less intuitive. In some species, limnetic morphs have been found to have a longer head than benthic morphs (Harrod *et al.* 2010; Præbel *et al.* 2013; Voje *et al.* 2013; Kusche *et al.* 2014). In lake trout, the humper morph in the Great Lakes feeds primarily on a pelagic crustacean, the opossum shrimp (*Mysis relicta*), and has a longer head than the lean morph, which feeds mainly on fish (Moore & Bronte 2001; Eshenroder 2008). Conversely, in the arctic charr (*Salvelinus alpinus*) and the rainbow trout, piscivorous fish have been found to have a longer head than planktivorous or invertivorous fish (Keeley *et al.* 2005, 2007; Janhunen *et al.* 2009). This result should thus be interpreted with caution, but it is possible that the longer head of the planktivorous lake trout provides a hydrodynamic advantage by increasing the fusiform shape of the body (Taylor & Foote 1991; Nicieza 1995). Finally, lake trout is usually considered as a species without sexual dimorphism (Alfonso 2004; Esteve 2005; Esteve *et al.* 2008; Muir *et al.* 2014; but see Muir *et al.* 2012b). However, here we found a different ecotype by sex interaction with body shape and a significant effect of sex on head length and caudal peduncle depth. Differences in body shape between male and female were significant for planktivorous but not for piscivorous fish, suggesting that sexual dimorphism in body shape was developed (at least partly) after the colonization of the planktivorous trophic niche.

In addition, planktivorous lake trout had significantly more gill rakers than piscivorous lake trout. Gill rakers are related to foraging efficiency in limnetic fishes because it increases prey retention efficiency in buccal cavity (Lazzaro 1987; Hessen *et al.* 1988). This trend has been observed between ecotypes of most northern temperate fish species (i.e. three-spine stickleback: Schulter 1993; rainbow smelt (*Osmerus mordax*): Taylor & Bentzen 1993; coregonids (*Coregonus* sp.): Lindsey 1981; Ozerov *et al.* 2015). However, our results are not as markedly different as these previous studies; we observed a difference of less than one gill raker between the two ecotypes. This small difference reflects the low variation in gill raker counts previously reported between lake trout ecotypes (Qadri 1967; Muir *et al.* 2012b). It may hypothetically be related to developmental constraints (Losos 2011), which could hinder the development of numerous gill rakers for planktivorous lake trout in response to the selective pressure imposed by the size of their prey.

### *The genetic basis of trophic adaptation*

We observed pronounced genetic differentiation and high level of fixation among most lake trout populations, reflecting the effect of both very limited (or absence of) gene flow between populations from different lakes and a strong genetic drift. This was observed regardless of the ecotype, as seen commonly in others allopatric populations of fish (Jones *et al.* 2012; Lamaze *et al.* 2012; Deagle *et al.* 2013; Perrier *et al.* 2013). These results are also in agreement with other recent studies on lake trout (Halbisen & Wilson 2009; Valiquette *et al.* 2014). Despite genetic drift, we identified 48 loci that could potentially be under selection between ecotypes. In comparison with the analyses with all loci, these 48 loci have a higher net allele frequency difference between allopatric populations of different ecotypes; they provide a better discrimination of ecotypes than all loci in the NJ trees and an increase in the proportion of

variance explained by ecotype in comparison with the SNPs from either method alone. Moreover, results from the jackknife-like LFMM procedure indicate that seven populations share a common genetic basis correlated with the ecotypic variation. Nonetheless, these SNPs failed to segregate the two sympatric ecotypes present in Maganasipi Lake. It is therefore possible that the differentiation between the two ecotypes in this lake is the result of phenotypic plasticity (West-Eberhard 2003; Proulx & Magnan 2004; Aubin-Horth & Renn 2009; Laporte *et al.* 2016a) or that adaptation occurred via different evolutionary pathways (DeFaveri *et al.* 2011; Gagnaire *et al.* 2013; Elmer *et al.* 2014; Laporte *et al.* 2015). Moreover, in sympatry, gene flow could possibly explain the misclassification of the population excluded from the jackknife-like procedure to its proper ecotype group. Finally, more than 60 years of contact after the impoundment of Mondonac Lake (Houde & Scrosati 2003) did not result in any appreciable level of admixture, which provides evidence for reproductive isolation in relation to the use of the two trophic niches being maintained.

### *Highlights from gene ontology analysis*

Three of the ten loci that were annotated were found near genes (*transcription factor sox-11*, *tubulin-specific chaperone cofactor E* and *homeobox protein aristaless-like 4*) that could be related to development (Tian *et al.* 1996; Meijlink *et al.* 1999; Kayserili *et al.* 2009; Berthelot *et al.* 2014) and a fourth SNP was found near a gene (*contactin-associated 4-like*) that could be related to the maturation of the nervous system (Karayannis *et al.* 2014). In addition, *homeobox protein aristaless-like 4* could also be related to digestive tract development (Ensembl; Dunn *et al.* 1997), suggesting that changes in digestive tract development could be implicated in the observed trophic divergence. Two SNPs were found near or within genes (*transcription factor sox-11* and *natriuretic peptide a*) that could be involved in cardiac development, in addition to regulation of cardiac muscle, blood pressure and salt–water balance (Brenner *et al.* 1990; Abraham *et al.* 2010; De Vito 2014; Ensembl; UniProtKB). Interestingly, in the Lake Whitefish (*Coregonus clupeaformis*), genes associated with cardiac development are also differentially expressed between the planktivorous and the benthivorous ecotypes (Dion-Côté *et al.* 2014). Additionally, a SNP was located in *V-type H+ ATPase*, a transporter known to play a role in osmoregulation of freshwater fishes (Wilson *et al.* 2000; Kirschner 2004; Beyenbach & Wieczorek 2006). Considering that a planktivorous diet involves more active swimming, a higher energetic metabolism and a higher oxygen consumption (Trudel *et al.* 2001; Rogers & Bernatchez 2007; St-Cyr *et al.* 2008; Jeukens *et al.* 2010; Evans *et al.* 2012) and that an increased oxygen consumption also increases ion losses by diffusion in gill cells (Randall *et al.* 1972; Nilsson 1986; Gonzalez & McDonald 1994), it is possible that there may be a difference in ion regulation between planktivorous and piscivorous ecotypes of lake trout. Overall, these findings suggest that the observed divergence may involve parallel developmental and functional changes related to use of different ecological niches.

### *Implications of the presence of multiple ecotypes for the management of lake trout*

Recent improvements in genomics allow biologists to address questions about wild and managed populations with more accuracy (Luikart *et al.* 2003; Allendorf *et al.* 2010; Pool *et al.* 2010; Narum *et al.* 2013; Andrews & Luikart 2014). Here, our results provide insights with implications for management and conservation of lake trout. The pronounced genetic structure observed among the studied populations (regardless of the ecotype) supports the conclusions of Valiquette *et al.* (2014), indicating that these populations should be considered as independent units, particularly for stocking programmes (Ozerov *et al.* 2010; Perrier *et al.* 2013).

Thus, for populations of small effective population size with low genetic diversity resulting from strong genetic drift, it has been suggested that stocking has to be performed cautiously (Ryman 1991; Ryman & Laikre 1991; Jonsson & Jonsson 2001). Particular attention is needed regarding the genetic variability of the stocked fish to maintain local adaptation of populations (Brannon *et al.* 2004; Fraser 2008; Araki & Schmid 2010; Valiquette *et al.* 2014). For instance, it has been documented that stocking with non-native individuals can induce a loss of genetic integrity (Valiquette *et al.* 2014). On the other hand, stocking could in some circumstances also lead to a gain (or regain) of potentially adaptive alleles (Lamaze *et al.* 2012). Accordingly, in highly exploited populations with low genetic diversity and strong genetic drift, it is also plausible that populations could potentially benefit from some gene flow to limit the extent of inbreeding depression and increase fitness of populations via introduction of potential adaptive genetic variants (i.e. genetic rescue; Hedrick *et al.* 2011; Frankham 2015; Whiteley *et al.* 2015). However, if such option is considered, particular care should be taken to avoid outbreeding depression and the disruption of adaptations by the stocking of fish from different ecotypes or domesticated fish. This is particularly important considering that stocking can modify gene expression and physiological condition in introgressed fish (Marie *et al.* 2010; Lamaze *et al.* 2013). Therefore, although potential benefits of genetic rescue should not

be neglected in such contexts, we suggest that caution should be taken to avoid alteration of local adaptation, in particular adaptation related to the trophic niches.

The combination of RADseq and morphometric analyses provides evidence that the observed phenotypic differentiation (planktivorous vs. piscivorous) is, at least partially, genetically determined. In addition, we identified several genes potentially related to the observed phenotypic differentiation, thus supporting an adaptive scenario. Therefore, these findings provide important considerations and a basis for the improvement of management strategies. Future conservation and stocking programmes should take into account this new knowledge, in order to maintain potential adaptations to different trophic niches in piscivorous and planktivorous populations of lake trout. For instance, ongoing research has shown that the growth rates of lake trout are being affected by the extent of introgression from a population of different ecotype origin (O. Morrisette *et al.*, in prep.). Accordingly, each ecotype should be closely monitored and independently managed where they occurred in sympatry. Finally, as the lake trout is susceptible to overexploitation, further studies focusing on population dynamics and life-history traits (i.e. growth rate, age and size at maturity) of these ecotypes are needed.

### *Challenges and caveats*

In highly structured systems, such as lake trout populations of the Laurentian Shield Lakes, signatures of selection can be difficult to detect. In fact, traditional methods used to detect selection (such as BAYESCAN) rely on allele frequency differentiation among populations. However, in systems with no (or low) migration between populations and where strong genetic drift occurs, populations show a high level of neutral genetic differentiation. It is particularly challenging to differentiate adaptive divergence from genetic drift and demographic history (Novembre & Di Rienzo 2009; Frichot *et al.* 2013) as loci showing high $F_{ST}$ values could be the result of strong genetic drift instead of selection. In our study, as a result of the very low or almost absent gene flow, the distribution of locus-specific $F_{ST}$ values does not fit the model of OUTFLANK, which is a very conservative method to detect outliers (François *et al.* 2016).

Recently, concerns have been raised about the use of population differentiation genome scan approaches for the detection of outliers (Le Corre & Kremer 2012; Bernatchez accepted; François *et al.* 2016). As an alternative, ecological association methods such as BAYENV2 and LFMM can be used to detect correlations between allele frequencies and key variables, and thus pinpoint selection linked to specific environmental variables (Lotterhos & Whitlock 2015). A substantial advantage of these methods is that they control for the population structure (Frichot *et al.* 2013; Günther & Coop 2013; Lotterhos & Whitlock 2015). Consequently, in systems with strong genetic drift and low levels of gene flow, pronounced neutral differentiation is taken into account in the detection of outliers. Consequently, these methods may be a more efficient means to detect signals of divergent selection in highly structured systems with limited or no gene flow. In addition, as traits under selection can be highly polygenic and influenced by small changes in allele frequencies (Mackay *et al.* 2009; Atwell *et al.* 2010; Davies *et al.* 2011; Yeaman 2015), $F_{ST}$-based outliers methods may not be necessarily the most powerful methods to detect selection (Le Corre & Kremer 2012; Laporte *et al.* 2016b). Here, the use of traditional methods did not allow the detection of loci potentially under selection, but we were able to identify many potentially selected loci using environmental association methods.

An additional noteworthy issue is the mix of ecological and anthropogenic contexts in the design. Namely, sampling involved lakes with allopatric or sympatric populations of piscivorous and planktivorous lake trout. In addition, one of the allopatric populations has been stocked intensively and one of the sympatric lakes is the result of the mixing of two lakes due to hydroelectric development. Nevertheless, our results show that, even in such a complex systems, the combination of different methods can help to characterize how evolutionary forces may have interacted in the past (and continue to interact) to shape intraspecific biodiversity.

## Acknowledgements

We are grateful to E. Normandeau for his help with the analyses and the bioinformatics part of the study, in addition to L. Benestan and A.-L. Ferchaud for their assistance with the analyses. We thank the field technicians and the biologists of the MFFP (Québec), the sport fishermen and the controlled harvested zones (ZECs) for their contribution to sampling. We also want to thank G. Côté, C. Babin, D. Boivin-Delisle A.-M. Dion-Côté, G. Légaré, C. Rougeux, J.-S. Moore and G. Ouellet-Cauchon who helped for the field and/or laboratory work and T. Gosselin for his help with the photography set-up and the genetic analyses. We are grateful to B.J.G. Sutherland for comments on the manuscript as well as to AE (Paul Bentzen) and two anonymous referees for their useful comments on a previous version of the manuscript. This research was supported by a strategic project grant from the Natural Science and Engineering Research Council of Canada (NSERC) to LB and PS and by Ressources Aquatiques Québec (RAQ). SB was supported by LB (2012–2013 and 2014–2015) an Alexander Graham Bell scholarship from NSERC (2013–2014), a scholarship from FQRNT (2014). ML was supported by a postdoctoral fellowship from FRQNT (2013–2014) and NSERC (2015–2016). CP was supported by a postdoctoral fellowship from RAQ.

## References


Abraham RL, Yang T, Blair M, Roden DM, Darbar D (2010) Augmented potassium current is a shared phenotype for two genetic defects associated with familial atrial fibrillation. *Journal of Molecular and Cellular Cardiology*, **48**, 181–190.

Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, **9**, 1655–1664.

Alfonso NR (2004) Evidence for two morphotypes of lake charr, *Salvelinus namaycush*, from Great Bear Lake, Northwest Territories, Canada. *Environmental Biology of Fishes*, **71**, 21–32.

Aljanabi S, Martinez I (1997) Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, **25**, 4692–4693.

Allendorf FW, Hohenlohe PA, Luikart G (2010) Genomics and the future of conservation genetics. *Nature Reviews Genetics*, **11**, 697–709.

Andrews KR, Luikart G (2014) Recent novel approaches for population genomics data analysis. *Molecular Ecology*, **23**, 1661–1667.

Araki H, Schmid C (2010) Is hatchery stocking a help or harm?: evidence, limitations and future directions in ecological and genetic surveys. *Aquaculture*, **308**, S2–S11.

Atwell S, Huang YS, Vilhjalmsson BJ *et al*. (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature*, **465**, 627–631.

Aubin-Horth N, Renn SCP (2009) Genomic reaction norms: using integrative biology to understand molecular mechanisms of phenotypic plasticity. *Molecular Ecology*, **18**, 3763–3780.

Benestan L, Gosselin T, Perrier C *et al*. (2015) RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). *Molecular Ecology*, **24**, 3299–3315.

Benoît J, Scrosati J, Dumont D (1997) Situation du touladi (*Salvelinus namaycush*) des réservoirs Châteauvert, Kempt, Manouane et Mondonac. Ministère de l'Environnement et de la Faune, Direction régionale Mauricie-Bois-Francs, Trois-Rivières, Rapport Technique, 93 pp.

Bernatchez L (accepted) On the maintenance of adaptive genetic variation to cope with environmental change: considerations from population genomics in fishes. *Journal of Fish Biology*.

Bernatchez L, Dodson JJ (1990) Allopatric origin of sympatric populations of lake whitefish (*Coregonus clupeaformis*) as revealed by mitochondrial-DNA restriction analysis. *Evolution*, **44**, 1263–1271.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of nearctic and palearctic fishes. *Molecular Ecology*, **7**, 431–452.

Berthelot C, Brunet F, Chalopin D *et al*. (2014) The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. *Nature Communications*, **5**, 3657.

Beyenbach KW, Wieczorek H (2006) The V-type H+ ATPase: molecular structure and function, physiological roles and regulation. *Journal of Experimental Biology*, **209**, 577–589.

Blackie CT, Weese DJ, Noakes DLG (2003) Evidence for resource polymorphism in the lake charr (*Salvelinus namaycush*) population of Great Bear Lake, Northwest Territories, Canada. *Ecoscience*, **10**, 509–514.

Blair LM, Granka JM, Feldman MW (2014) On the stability of the Bayenv method in assessing human SNP-environment associations. *Human Genomics*, **8**, 1.

Blake RW (2004) Fish functional design and swimming performance. *Journal of Fish Biology*, **65**, 1193–1222.

Brannon EL, Amend DF, Cronin MA *et al*. (2004) The controversy about salmon hatcheries. *Fisheries*, **29**, 12–31.

Brenner BM, Ballermann BJ, Gunning ME, Zeidel ML (1990) Diverse biological actions of atrial natriuretic peptide. *Physiological Reviews*, **70**, 665–699.

Bronte CR, Moore SA (2007) Morphological variation of siscowet lake trout in Lake Superior. *Transactions of the American Fisheries Society*, **136**, 509–517.

Catchen JM, Amores A, Hohenlohe P *et al*. (2011) Stacks: building and genotyping loci *de novo* from short-read sequences. *G3: Genes, Genomes, Genetics*, **1**, 171–182.

Chavarie L, Howland KL, Tonn WM (2013) Sympatric polymorphism in lake trout: the coexistence of multiple shallow-water morphotypes in Great Bear Lake. *Transactions of the American Fisheries Society*, **142**, 814–823.

Colosimo PF, Hosemann KE, Balabhadra S *et al*. (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, **307**, 1928–1933.

Conte GL, Arnegard ME, Peichel CL, Schluter D (2012) The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B-Biological Sciences*, **279**, 5039–5047.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**, 1411–1423.

Danecek P, Auton A, Abecasis G *et al*. (2011) The variant call format and VCFtools. *Bioinformatics*, **27**, 2156–2158.

Davies G, Tenesa A, Payton A *et al*. (2011) Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Molecular Psychiatry*, **16**, 996–1005.

De Vito P (2014) Atrial natriuretic peptide: an old hormone or a new cytokine? *Peptides*, **58**, 108–116.

Deagle BE, Jones FC, Absher DM, Kingsley DM, Reimchen TE (2013) Phylogeography and adaptation genetics of stickleback from the Haida Gwaii archipelago revealed using genome-wide single nucleotide polymorphism genotyping. *Molecular Ecology*, **22**, 1917–1932.

DeFaveri J, Shikano T, Shimada Y, Goto A, Merilä J (2011) Global analysis of genes involved in freshwater adaptation in threespine sticklebacks (*Gasterosteus aculeatus*). *Evolution*, **65**, 1800–1807.

Dion-Côté A-M, Renaut S, Normandeau E, Bernatchez L (2014) RNA-seq reveals transcriptomic shock involving transposable elements reactivation in hybrids of young lake whitefish species. *Molecular Biology and Evolution*, **31**, 1188–1199.

Dunn NR, Winnier GE, Hargett LK *et al*. (1997) Haploinsufficient phenotypes in *Bmp4* heterozygous null mice and modification by mutations in *Gli3* and *Alx4*. *Developmental Biology*, **188**, 235–247.

Elmer KR, Fan S, Kusche H *et al*. (2014) Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nature Communications*, **5**, 5168.

Endler JA (1986) *Natural Selection in the Wild*. Princeton University Press, Princeton, NJ.

Eshenroder RL (2008) Differentiation of deep-water lake charr *Salvelinus namaycush* in North American lakes. *Environmental Biology of Fishes*, **83**, 77–90.

Esteve M (2005) Observations of spawning behaviour in Salmoninae: Salmo, *Oncorhynchus* and *Salvelinus*. *Reviews in Fish Biology and Fisheries*, **15**, 1–21.

Esteve M, McLennan DA, Gunn JM (2008) Lake trout (*Salvelinus namaycush*) spawning behaviour: the evolution of a new female strategy. *Environmental Biology of Fishes*, **83**, 69–76.

Evans ML, Præbel K, Peruzzi S, Bernatchez L (2012) Parallelism in the oxygen transport system of the lake whitefish: the role of physiological divergence in ecological speciation. *Molecular Ecology*, **21**, 4038–4050.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Fisheries and Oceans Canada (FOC) (1886) Report on the fisheries of Canada for the year 1885. Ottawa.

Fisheries and Oceans Canada (FOC) (2010) Survey of Recreational Fishing in Canada 2010, Economic Analysis and Statistics Strategic Policy. Ottawa.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. *Genetics*, **180**, 977–993.

Foster SA, McKinnon GE, Steane DA, Potts BM, Vaillancourt RE (2007) Parallel evolution of dwarf ecotypes in the forest tree *Eucalyptus globulus*. *New Phytologist*, **175**, 370–380.

François O, Martins H, Caye K, Schoville SD (2016) Controlling false discoveries in genome scans for selection. *Molecular Ecology*, **25**, 454–469.

Frankham R (2015) Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*, **24**, 2610–2618.

Fraser DJ (2008) How well can captive breeding programs conserve biodiversity? A review of salmonids. *Evolutionary Applications*, **1**, 535–586.

Frichot E, François O (2015) LEA: an R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, **6**, 925–929.

Frichot E, Schoville SD, Bouchard G, Francois O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, **30**, 1687–1699.

Gagnaire PA, Pavey SA, Normandeau E, Bernatchez L (2013) The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution*, **67**, 2483–2497.

del Giorgio PAD, Gasol JM (1995) Biomass distribution in freshwater plankton communities. *The American Naturalist*, **146**, 135–152.

Gonzalez RJ, McDonald DG (1994) The relationship between oxygen uptake and ion loss in fish from diverse habitats. *Journal of Experimental Biology*, **190**, 95–108.

Günther T, Coop G (2013) Robust identification of local adaptation from allele frequencies. *Genetics*, **195**, 205–220.

Halbisen MA, Wilson CC (2009) Variable introgression from supplemental stocking in Southern Ontario populations of lake trout. *Transactions of the American Fisheries Society*, **138**, 699–719.

Harrod C, Mallela J, Kahilainen KK (2010) Phenotype-environment correlations in a putative whitefish adaptive radiation. *Journal of Animal Ecology*, **79**, 1057–1068.

Harvey CJ, Schram ST, Kitchell JF (2003) Trophic relationships among lean and siscowet lake trout in Lake Superior. *Transactions of the American Fisheries Society*, **132**, 219–228.

Hatfield T (1997) Genetic divergence in adaptive characters between sympatric species of sticklebacks. *The American Naturalist*, **149**, 1009–1029.

Hedrick PW, Adams JR, Vucetich JA (2011) Reevaluating and broadening the definition of genetic rescue. *Conservation Biology*, **25**, 1069–1070.

Hessen DO, Andersen R, Hindar K, Skurdal J (1988) Food selection and competition in salmonids as reflected by gill-raker number and morphology. *Journal of Applied Ichthyology*, **4**, 121–129.

Hohenlohe PA, Bassham S, Currey M, Cresko WA (2012) Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philosophical Transactions of the Royal Society B-Biological Sciences*, **367**, 395–408.

Houde L, Scrosati J (2003) Le touladi au réservoir Mondonac. État de la population et bilan des mesures de conservation. Société de la faune et des parcs du Québec. Direction de l'aménagement de la faune de la Mauricie, Rapport technique, 25 pp.

Janhunen M, Peuhkuri N, Piironen J (2009) Morphological variability among three geographically distinct arctic charr (*Salvelinus alpinus* L.) populations reared in a common hatchery environment. *Ecology of Freshwater Fish*, **18**, 106–116.

Jeffrey H (1961) *Theory of Probability*. Oxford University Press, New York.

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

Jones FC, Chan YF, Schmutz J *et al.* (2012) A genome-wide SNP Genotyping array reveals patterns of global and repeated species-pair divergence in sticklebacks. *Current Biology*, **22**, 83–90.

Jonsson B, Jonsson N (2001) Polymorphism and speciation in arctic charr. *Journal of Fish Biology*, **58**, 605–638.

Kahilainen KK, Siwertsson A, Gjelland KØ *et al.* (2011) The role of gill raker number variability in adaptive radiation of coregonid fish. *Evolutionary Ecology*, **25**, 573–588.

Kalinowski ST (2009) How well do evolutionary trees describe genetic relationships among populations? *Heredity*, **102**, 506–513.

Karayannis T, Au E, Patel JC *et al.* (2014) Cntnap4 differentially contributes to GABAergic and dopaminergic synaptic transmission. *Nature*, **511**, 236–240.

Kayserili H, Uz E, Niessen C *et al.* (2009) ALX4 dysfunction disrupts craniofacial and epidermal development. *Human Molecular Genetics*, **18**, 4357–4366.

Keeley ER, Parkinson EA, Taylor EB (2005) Ecotypic differentiation of native rainbow trout (*Oncorhynchus mykiss*) populations from British Columbia. *Canadian Journal of Fisheries and Aquatic Science*, **62**, 1523–1539.

Keeley ER, Parkinson EA, Taylor EB (2007) The origins of ecotypic variation of rainbow trout: a test of environmental vs. genetically based differences in morphology. *Journal of Evolutionary Biology*, **20**, 725–736.

Kirschner LB (2004) The mechanism of sodium chloride uptake in hyperregulating aquatic animals. *Journal of Experimental Biology*, **207**, 1439–1452.

Klingenberg CP (2011) MORPHOJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, **11**, 353–357.

Kusche H, Recknagel H, Elmer KR, Meyer A (2014) Crater lake cichlids individually specialize along the benthic–limnetic axis. *Ecology and Evolution*, **4**, 1127–1139.

Lamaze FC, Sauvage C, Marie A, Garant D, Bernatchez L (2012) Dynamics of introgressive hybridization assessed by SNP population genomics of coding genes in stocked brook charr (*Salvelinus fontinalis*). *Molecular Ecology*, **21**, 2877–2895.

Lamaze FC, Garant D, Bernatchez L (2013) Stocking impacts the expression of candidate genes and physiological condition in introgressed brook charr (*Salvelinus fontinalis*) populations. *Evolutionary Applications*, **6**, 393–407.

Langerhans RB, Reznick DN (2010) Ecology and evolution of swimming performance in fishes: predicting evolution with biomechanics. In: *Fish Locomotion: An Eco-Ethological Perspective* (eds Domenici P, Kapoor BG), pp. 200–248. Science Publishers, Enfield, New Hampshire.

Laporte M, Magnan P, Angers B (2011) Genetic differentiation between the blue and the yellow phenotypes of walleye (*Sander vitreus*): an example of parallel evolution. *Ecoscience*, **18**, 124–129.

Laporte M, Rogers SM, Dion-Côté AM *et al.* (2015) RAD-QTL mapping reveals both genome-level parallelism and different genetic architecture underlying the evolution of body shape in Lake Whitefish (Coregonus clupeaformis) species pairs. *G3: Genes, Genomes, Genetics*, **5**, 1481–1491.

Laporte M, Claude J, Berrebi P, Perret P, Magnan P (2016a) Shape plasticity in response to water velocity in the freshwater blenny *Salaria fluviatilis* (Asso 1801). *Journal of Fish Biology*, **88**, 1191–1203.

Laporte M, Pavey SA, Rougeux C *et al.* (2016b) RAD-sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Molecular Ecology*, **25**, 219–237.

Lazzaro X (1987) A review of planktivorous fishes: their evolution, feeding, behaviors, selectivities, and impacts. *Hydrobiologia*, **146**, 97–167.

Le Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, **21**, 1548–1566.

Legault M, Fournier H, Nadeau D, Benoit J (2001) Bilan de la gamme de taille protégée pour la touladi, 1993-1997; État de situation pour le Québec. Société de la faune et des parcs du Québec, Direction de la recherche sur la faune, Direction de l'aménagement de la faune. 76 pp.

Legendre P, Legendre L (1998) *Numerical Ecology*, 2nd edn. Elsevier, Amsterdam.

Lindsey CC (1981) Stocks are chameleons: plasticity in gill rakers of coregonid fishes. *Canadian Journal of Fisheries and Aquatic Sciences*, **38**, 1497–1506.

Lischer HEL, Excoffier L (2012) PGDSPIDER: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, **28**, 298–299.

Losos JB (2011) Convergence, adaptation, and constraint. *Evolution*, **65**, 1827–1840.

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, **23**, 2178–2192.

Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, **24**, 1031–1046.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

MacArthur RH (1958) Population ecology of some warblers of Northeastern coniferous forests. *Ecology*, **39**, 599–619.

Mackay TFC, Stone EA, Ayroles JF (2009) The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics*, **10**, 565–577.

Mandrak NE, Crossman EJ (1992) Postglacial dispersal of freshwater fishes into Ontario. *Canadian Journal of Zoology*, **70**, 2247–2259.

Marie AD, Bernatchez L, Garant D (2010) Loss of genetic integrity correlates with stocking intensity in brook charr (*Salvelinus fontinalis*). *Molecular Ecology*, **19**, 2025–2037.

Martin NV (1966) The significance of food habits in the biology, exploitation, and management of Algonquin Park, Ontario, lake trout. *Transactions of the American Fisheries Society*, **95**, 415–422.

Matuszek JE, Shuter BJ, Casselman JM (1990) Changes in lake trout growth and abundance after introduction of cisco into Lake Opeongo, Ontario. *Transactions of the American Fisheries Society*, **119**, 718–729.

May-McNally SL, Quinn TP, Woods PJ, Taylor EB (2014) Evidence for genetic distinction among sympatric ecotypes of arctic char (*Salvelinus alpinus*) in south-western Alaskan lakes. *Ecology of Freshwater Fish*, **24**, 562–574.

Meijlink F, Beverdam A, Brouwer A, Oosterveen TC, Berge DT (1999) Vertebrate aristaless-related genes. *The Internal Journal of Developmental Biology*, **43**, 651–663.

Meirmans PG, Van Tienderen PH (2004) Genotype and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792–794.

Minoche AE, Dohm JC, Himmelbauer H (2011) Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. *Genome Biology*, **12**, R112.

Moore SA, Bronte CR (2001) Delineation of sympatric morphotypes of lake trout in Lake Superior. *Transactions of the American Fisheries Society*, **130**, 1233–1240.

Muir AM, Vecsei P, Krueger CC (2012a) A perspective on perspectives: methods to reduce variation in shape analysis of digital images. *Transactions of the American Fisheries Society*, **141**, 1161–1170.

Muir AM, Blackie CT, Marsden JE, Krueger CC (2012b) Lake charr *Salvelinus namaycush* spawning behaviour: new field observations and a review of current knowledge. *Reviews in Fish Biology and Fisheries*, **22**, 575–593.

Muir AM, Bronte CR, Zimmerman MS *et al.* (2014) Ecomorphological diversity of lake trout at Isle Royale, Lake Superior. *Transactions of the American Fisheries Society*, **143**, 972–987.

Muir AM, Hansen MJ, Bronte CR, Krueger CC (in press) If arctic charr Salvelinus alpinus is "the most diverse vertebrate", what is the lake charr Salvelinus namaycush? *Fish and Fisheries*, doi:10.1111/faf.12114.

Narum SR, Buerkle CA, Davey JW, Miller MR, Hohenlohe PA (2013) Genotyping-by-sequencing in ecological and conservation genomics. *Molecular Ecology*, **22**, 2841–2847.

Neumann RM, Guy CS, Willis DW (2012) Length, weight and associated indices. In: *Fisheries Techniques*, 3rd edn (eds Zale AV, Parrish DL, Sutton TM), pp. 677–731. American Fisheries Society, Bethesda, MD.

Nicieza AG (1995) Morphological variation between geographically disjunct populations of atlantic Salmon: the effects of ontogeny and habitat shift. *Functional Ecology*, **9**, 448–456.

Nikolsky GV (1963) *The Ecology of Fishes*. Academic Press, London.

Nilsson S (1986) Control of gill blood flow. In: *Fish Physiology: Recent Advances* (eds Nilsson S, Holmgren S), pp. 86–101. Springer, Netherlands, London.

Noakes DLG (2008) Charr truth: sympatric differentiation in *Salvelinus* species. *Environmental Biology of Fishes*, **83**, 7–15.

Nosil P (2012) *Ecological Speciation*. Oxford University Press, Oxford.

Novembre J, Di Rienzo A (2009) Spatial patterns of variation due to natural selection in humans. *Nature*, **10**, 745–755.

Østbye K, Amundsen PA, Bernatchez L *et al.* (2006) Parallel evolution of ecomorphological traits in the European whitefish *Coregonus lavaretus* (L.) species complex during postglacial times. *Molecular Ecology*, **15**, 3983–4001.

Ozerov M, Veselov A, Lumme J, Primmer C (2010) Genetic structure of freshwater Atlantic salmon (*Salmo salar* L.) populations from the lakes Onega and Ladoga of northwest Russia and implications for conservation. *Conservation Genetics*, **11**, 1711–1724.

Ozerov MY, Himberg M, Aykanat T *et al.* (2015) Generation of a neutral $F_{ST}$ baseline for testing local adaptation on gill raker number within and between European whitefish ecotypes in the Baltic Sea basin. *Journal of Evolutionary Biology*, **28**, 1170–1183.

Perrier C, Bourret V, Kent MP, Bernatchez L (2013) Parallel and nonparallel genome-wide divergence among replicate population pairs of freshwater and anadromous Atlantic salmon. *Molecular Ecology*, **22**, 5577–5593.

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE*, **7**, e32253.

Pool JE, Hellmann I, Jensen JD, Nielsen R (2010) Population genetic inference from genomic sequence variation. *Genome Research*, **20**, 291–300.

Præbel K, Knudsen R, Siwertsson A *et al.* (2013) Ecological speciation in postglacial European whitefish: rapid adaptive radiations into the littoral, pelagic, and profundal lake habitats. *Ecology and Evolution*, **3**, 4970–4986.

Proulx R, Magnan P (2004) Contribution of phenotypic plasticity and heredity to trophic polymorphism of lacustrine brook charr (*Salvelinus fontinalis* M.). *Evolutionary Ecology Research*, **6**, 503–522.

Purcell S, Neale B, Todd-Brown K *et al.* (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, **81**, 559–575.

Qadri SU (1967) Morphological comparisons of three populations of the lake char, *Cristivomer namaycush*, from Ontario and Manitoba. *Journal of the Fisheries Research Board of Canada*, **24**, 1407–1411.

R Development Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Randall DJ, Baumgarten D, Malyusz M (1972) The relationship between gas and ion transfer across the gills of fishes. *Comparative Biochemistry and Physiology*, **41**, 629–637.

Robinson BW, Wilson DS (1994) Character release and displacement in fishes: a neglected literature. *The American Naturalist*, **144**, 596–627.

Rogers SM, Bernatchez L (2007) The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp. Salmonidae) species pairs. *Molecular Biology and Evolution*, **24**, 1423–1438.

Rogers SM, Gagnon V, Bernatchez L (2002) Genetically based phenotype-environment association for swimming behavior in lake whitefish ecotypes (*Coregonus clupeaformis* Mitchill). *Evolution*, **56**, 2322–2329.

Ryman N (1991) Conservation genetics considerations in fishery management. *Journal of Fish Biology*, **39**, 211–224.

Ryman N, Laikre L (1991) Effects of supportive breeding on the genetically effective population size. *Conservation Biology*, **5**, 325–329.

Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics*, **14**, 807–820.

Schluter D (1993) Adaptive radiation in sticklebacks: size, shape, and habitat use efficiency. *Ecology*, **74**, 699–709.

Schluter D (2000) *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.

Schluter D, McPhail JD (1992) Ecological character displacement and speciation in sticklebacks. *American Naturalist*, **140**, 85–108.

Schluter D, Nagel LM (1995) Parallel speciation by natural selection. *The American Naturalist*, **146**, 292–301.

Skúlason S, Smith TB (1995) Resource polymorphisms in vertebrates. *Trends in Ecology & Evolution*, **10**, 366–370.

Skúlason S, Noakes DLG, Snorrason SS (1989) Ontogeny of trophic morphology in four sympatric morphs of arctic charr *Salvelinus alpinus* in Thingvallavatn, Iceland. *Biological Journal of the Linnean Society*, **38**, 281–301.

Smith TB, Skúlason S (1996) Evolutionary significance of resource polymorphisms in fishes, amphibians, and birds. *Annual Review of Ecology and Systematics*, **27**, 111–133.

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

Taylor EB (1999) Species pairs of north temperate freshwater fishes: evolution, taxonomy, and conservation. *Reviews in Fish Biology and Fisheries*, **9**, 299–324.

Taylor EB, Bentzen P (1993) Evidence for multiple origins and sympatric divergence of trophic ecotypes of smelt (*Osmerus*) in Northeastern North America. *Evolution*, **47**, 813–832.

Taylor EB, Foote CJ (1991) Critical swimming velocities of juvenile sockeye salmon and kokanee, the anadromous and non-anadromous forms of *Oncorhynchus nerka* (Walbaum). *Journal of Fish Biology*, **38**, 407–419.

Terabayashi T, Sakaguchi M, Shinmyozu K *et al.* (2012) Phosphorylation of Kif26b promotes its polyubiquitination and subsequent proteasomal degradation during kidney development. *PLoS ONE*, **7**, e39714.

Tian G, Huang Y, Rommelaere H *et al.* (1996) Pathway leading to correctly folded β-tubulin. *Cell*, **86**, 287–296.

Trudel M, Tremblay A, Schetagne R, Rasmussen JB (2001) Why are dwarf fish so small? An energetic analysis of

polymorphism in lake whitefish (*Coregonus clupeaformis*). *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 394–405.
Tsurusaki Y, Koshimizu E, Ohashi H *et al.* (2014) De novo SOX11 mutations cause Coffin-Siris syndrome. *Nature Communications*, **5**, 4011.
Uchiyama Y, Sakaguchi M, Terabayashi T *et al.* (2010) Kif26b, a kinesin family gene, regulates adhesion of the embryonic kidney mesenchyme. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 9240–9245.
Valiquette E, Perrier C, Thibault I, Bernatchez L (2014) Loss of genetic integrity in wild lake trout populations following stocking: insights from an exhaustive study of 72 lakes from Québec, Canada. *Evolutionary Applications*, **7**, 625–644.
Vander Zanden MJ, Shuter BJ, Lester NP, Rasmussen JB (2000) Within- and among-population variation in the trophic position of a pelagic predator, lake trout (*Salvelinus namaycush*). *Canadian Journal of Fisheries and Aquatic Sciences*, **57**, 725–731.
Voje KL, Mazzarella AB, Hansen TF *et al.* (2013) Adaptation and constraint in a stickleback radiation. *Journal of Evolutionary Biology*, **26**, 2396–2414.
Walker JA (1997) Ecological morphology of lacustrine threespine stickleback *Gasterosteus aculeatus* L. (Gasterosteidae) body shape. *Biological Journal of the Linnean Society*, **61**, 3–50.
Walker JA (2010) An integrative model of evolutionary covariance: a symposium on body shape in fishes. *Integrative and Comparative Biology*, **50**, 1051–1056.
Webb PW (1982) Locomotor patterns in the evolution of actinopterygian fishes. *Integrative and Comparative Biology*, **22**, 329–342.
Webb PW (1984) Body form, locomotion and foraging in aquatic vertebrates. *American Zoologist*, **24**, 107–120.
Webb PW (1986a) Effect of body form and response threshold on the vulnerability of four species of teleost prey attacked by largemouth bass. *Canadian Journal of Fisheries and Aquatic Sciences*, **43**, 763–771.
Webb PW (1986b) Locomotion and predator-prey relationships. In: *Predator-Prey Relationships* (eds Feder ME, Lauder GV), pp. 24–41. Chicago University Press, Chicago, IL.
Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
West-Eberhard MJ (2003) *Developmental Plasticity and Evolution*. Oxford University Press, Oxford.
Whiteley AR, Fitzpatrick SW, Funk WC, Tallmon DA (2015) Genetic rescue to the rescue. *Trends in Ecology & Evolution*, **30**, 42–49.
Whitlock MC, Lotterhos KE (2015) Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of FST. *The American Naturalist*, **186**, S24–S36.
Willacker JJ, von Hippel FA, Wilton PR, Walton KM (2010) Classification of threespine stickleback along the benthic-limnetic axis. *Biological Journal of the Linnean Society*, **101**, 595–608.
Wilson CC, Hebert PDN (1996) Phylogeographic origins of lake trout (*Salvelinus namaycush*) in eastern North America. *Canadian Journal of Fisheries and Aquatic Sciences*, **53**, 2764–2775.
Wilson CC, Hebert PDN (1998) Phylogeography and postglacial dispersal of lake trout (*Salvelinus namaycush*) in North America. *Canadian Journal of Fisheries and Aquatic Sciences*, **55**, 1010–1024.
Wilson CC, Mandrak NE (2004) History and evolution of lake trout in shield lakes: past and future challenges. In: *Boreal Shield Watersheds: Lake trout Ecosystems in a Changing Environment* (eds Gunn JM, Steedman RJ, Ryder RA), pp. 21–35. CRC Press, Boca Raton, FL.
Wilson JM, Laurent P, Tufts BL *et al.* (2000) NaCl uptake by the branchial epithelium in freshwater teleost fish: an immunological approach to ion-transport protein localization. *Journal of Experimental Biology*, **203**, 2279–2296.
Yeaman S (2015) Local adaptation by alleles of small effect. *The American Naturalist*, **186**, S74–S89.
Zelditch ML, Swiderski DL, Sheets HD (2012) *Geometric Morphometrics for Biologists: A Primer*, 2nd edn. Elsevier/Academic Press, Amsterdam.
Zimmerman MS, Krueger CC, Eshenroder RL (2006) Phenotypic diversity of lake trout in Great Slave Lake: differences in morphology, buoyancy, and habitat depth. *Transactions of the American Fisheries Society*, **135**, 1056–1067.
Zimmerman MS, Krueger CC, Eshenroder RL (2007) Morphological and ecological differences between shallow- and deep-water lake trout in Lake Mistassini, Quebec. *Journal of Great Lakes Research*, **33**, 156–169.
Zimmerman MS, Schmidt SN, Krueger CC, Vander Zanden MJ, Eshenroder RL (2009) Ontogenetic niche shifts and resource partitioning of lake trout morphotypes. *Canadian Journal of Fisheries and Aquatic Sciences*, **66**, 1007–1018.

The study represents S.B.'s master's project. L.B. and P.S. supervised S.B. and designed the study, with S.B.'s help for the field operations. C.P. performed the genotyping and S.B., C.P. and M.L. performed analyses. S.B., M.L., C.P. and L.B. contributed to the interpretation of the results. S.B., M.L., C.P., L.B. and P.S. contributed to the writing of the manuscript. All authors read and approved the manuscript.

## Data accessibility

Genomic data (filtered markers and markers putatively under positive selection) and morphological (shape, linear measurements, gill rakers) data are available on Dryad (doi: 10.5061/dryad.jk680). Raw demultiplexed sequences are available on NCBI SRA (SRP071309).

## Supporting information

Additional supporting information may be found in the online version of this article.

**Supplementary file 1** Modified version of the Aljanabi & Martinez (1997) salt extraction protocol.

**Supplementary file 2** Modified version of the Poland *et al.* (2012) two-enzyme GBS protocol.

**Table S1.** Number of putative SNPs remaining after each step of filtering.

**Fig. S1.** Venn diagram showing the number of Lake Trout used in the genomic and morphometric analyses. The numbers in parentheses indicate the number of planktivorous, piscivorous and stocked (used in genomic analyses exclusively) Lake Trout respectively.

**Fig. S2.** Landmarks and linear measurements used to evaluate morphological divergence between piscivorous and planktivorous Lake Trout. (a) Landmarks: (1) anterior tip of the snout, (2) posterior of neurocranium above tip of opercle, (3) anterior insertion of dorsal fin, (4) posterior insertion of dorsal fin, (5) anterior insertion of adipose fin, (6) dorsal insertion of caudal fin, (7) midpoint of hypural plate, (8) ventral insertion of caudal fin, (9) anterior insertion of anal fin, (10) posterior insertion of anal fin, (11) ventral surface of body below ventral insertion of dorsal fin, (12) ventral surface of head below posterior of neurocranium above tip of opercle, (13) insertion point of pectoral fin, (14) insertion point of pelvic fin, (15) anterior tip of lower jaw, (16) posterior tip of maxilla, (17) ventral surface of head below the tip of maxilla, (18) posterior tip of eye, (19) anterior tip of eye, (20) top of cranium at middle point of eye and (b) linear measurements: (1) standard body length, (2) upper jaw, (3) lower jaw, (4) head depth 1, (5) snout-eye, (6) head length, (7) eye diameter, (8) head depth 2, (9) dorsal fin, (10) caudal fin, (11) caudal peduncle depth, (12) anal fin, (13) pelvic fin, (14) pectoral fin, (15) body depth bellow anterior insertion of dorsal fin.

**Fig. S3.** Histogram of discriminant scores for (a) piscivorous and (b) planktivorous male (dark grey) and female (wavy light grey) Lake Trout. Mean shape differences between male (hatched line, light white triangles) and female (continuous line, black circles) for piscivorous (c) and planktivorous (d) Lake Trout (scale is magnified by 2X to emphasize differences between ecotypes).

**Fig. S4.** Bayesian individual assignment using Admixture for the 14 putative Lake Trout populations and 3925 SNP markers. Planktivorous populations are identified by abbreviations in green, piscivorous populations in red and the stocked population in black. Sympatric ecotypes are shown in black rectangles and separated by a black hatched line. Abbreviations are defined in the legend of Fig. 1.

**Fig. S5.** |Z|-scores from LFMM plotted against Bayes Factors from BAYENV2 for all polymorphic SNPs in planktivorous and piscivorous Lake Trout populations. SNPs showing no association with ecotypes are represented by black circles, SNPs identified as associated with ecotypes by LFMM or BAYENV2 only are represented by blue and orange circles respectively, and SNPs identified by both methods are represented by red circles. The numbers in parentheses indicate the number of SNPs showing no association with ecotypes (3650), SNPs identified as associated with ecotypes by LFMM (143) or BAYENV2 (51) only and SNPs identified by both methods (48).

**Fig. S6.** Neighbour-joining trees constructed with $F_{ST}$ pairwise values among the 14 Lake Trout groups based on potentially selected markers detected using LFMM (excluding one allopatric piscivorous populations at a time to select the markers followed by the replacement of the left out population in the tree) with bootstrap values based on 1000 replicates. Excluded population: a) DE, b) LY, c) LO, d) MO, e) MA and f) TE. Piscivorous populations are presented in red, planktivorous populations in green, the stocked population in black and the population excluded from the LFMM analysis is circled in grey. Abbreviations are defined in the legend of Fig. 1.

**Fig. S7.** Neighbour-joining trees constructed with $F_{ST}$ pairwise values among the 14 Lake Trout groups based on potentially selected markers detected using LFMM (excluding one allopatric planktivorous populations at a time to select the markers followed by the replacement of the left out population in the tree) with bootstrap values based on 1000 replicates. Excluded population: a) BO, b) CA and c) TU. Piscivorous populations are presented in red, planktivorous populations in green, the stocked population in black and the population excluded from the LFMM analysis is circled in grey. Abbreviations are defined in the legend of Fig. 1.

**Fig. S8.** Neighbour-joining trees constructed with $F_{ST}$ pairwise values among the 14 Lake Trout groups based on potentially selected markers detected using LFMM (excluding one sympatric ecotype at a time) with bootstrap values based on 1000 replicates. Excluded ecotype: a) MD-planktivorous, b) MG-planktivorous, c) MD-piscivorous and d) MG-piscivorous. Piscivorous populations are presented in red, planktivorous populations in green, the stocked population in black and the sympatric ecotype excluded from the LFMM analysis is circled in grey. Abbreviations are defined in the legend of Fig. 1.