Received: 4 June 2019 | Revised: 13 July 2020 | Accepted: 16 July 2020

DOI: 10.1111/mec.15566

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Adaptive and maladaptive genetic diversity in small populations: Insights from the Brook Charr (*Salvelinus fontinalis*) case study

Anne-Laure Ferchaud[1] | Maeva Leitwein[1] | Martin Laporte[1] | Damien Boivin-Delisle[1] | Bérénice Bougas[1] | Cécilia Hernandez[1] | Éric Normandeau[1] | Isabel Thibault[2] | Louis Bernatchez[1]

[1]Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC, Canada

[2]Direction de l'expertise Sur la Faune Aquatique, Ministère des Forêts, de la Faune et des Parcs du Québec, Québec, QC, Canada

Correspondence
Anne-Laure Ferchaud, Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, Québec G1V 0A6, Canada.
Email: annelaureferchaud@gmail.com

Funding information
MFFP; Canadian Research Chair in Genomics and Conservation of Aquatic Resources; Ressources Aquatiques Québec (RAQ); Strategic Project Grant; Natural Sciences and Engineering Research Council of Canada

Abstract

Investigating the relative importance of neutral versus selective processes governing the accumulation of genetic variants is a key goal in both evolutionary and conservation biology. This is particularly true in the context of small populations, where genetic drift can counteract the effect of selection. Using Brook Charr (*Salvelinus fontinalis*) from Québec, Canada, as a case study, we investigated the importance of demographic versus selective processes governing the accumulation of both adaptive and maladaptive mutations in closed versus open and connected populations to assess gene flow effect. This was achieved by using 14,779 high-quality filtered SNPs genotyped among 1,416 fish representing 50 populations from three life history types: lacustrine (closed populations), riverine and anadromous (connected populations). Using the PROVEAN algorithm, we observed a considerable accumulation of putative deleterious mutations across populations. The absence of correlation between the occurrence of putatively beneficial or deleterious mutations and local recombination rate supports the hypothesis that genetic drift might be the main driver of the accumulation of such variants. However, despite a lower genetic diversity observed in lacustrine than in riverine or anadromous populations, lacustrine populations do not exhibit more deleterious mutations than the two other history types, suggesting that the negative effect of genetic drift in lacustrine populations may be mitigated by that of relaxed purifying selection. Moreover, we also identified genomic regions associated with anadromy, as well as an overrepresentation of transposable elements associated with variation in environmental variables, thus supporting the importance of transposable elements in adaptation.

KEYWORDS

adaptation, anadromy, deleterious mutations, salmonids, small populations

1 | INTRODUCTION

The fate of wild populations exposed to environmental variation is determined by an interplay between genetic variation and demography (Lande, 1988). The genome carries beneficial but also deleterious variants; for instance, between 500 and 1,200 deleterious variants have been estimated in a single human genome (Fay, Wyckoff, & Wu, 2001; Sunyaev et al., 2001). Deleterious variants are considered putatively harmful since they can lower the fitness of individuals carrying them depending on their dominance and selection coefficient (Charlesworth, 2012). Low genetic diversity has been suspected to be partly responsible for many population and species extinctions because it increases the chances of recessive deleterious alleles to being expressed and decreases the chances for an individual to have a genotype matching the environmental challenges (Fagan & Holmes, 2006; Lynch, Conery, & Buerger, 1995). The potential negative impact that inbreeding will have on population health and reproduction compared to an outbred population is considered a component of "genetic load" (Crow, 1958). According to population genetics theory, the accumulation of deleterious mutations and therefore the genetic load depends on multiple factors such as effective population size, admixture and selection (Whitlock & Bürger, 2004). Although natural selection may increase the frequency of adaptive mutations and remove deleterious alleles at the same time (Whitlock, Griswold, & Peters, 2003), deleterious variants may also hitchhike with advantageous mutations to high frequencies if they have intermediate fitness costs (Chun & Fay, 2011). Additionally, recombination facilitates the removal of maladaptive alleles from a population (Felsenstein, 1974). Consequently, genomic regions of low recombination are more likely to accumulate deleterious alleles because the efficiency of selection is reduced in those genomic regions (Charlesworth, 2009). Inversely, high recombination exposes deleterious mutations to purging. However, when stochastic demographic processes (e.g., genetic drift, bottlenecks) prevail over natural selection, for instance in small isolated populations, deleterious variants may increase in frequency (Lynch et al., 1995) and inbreeding can unmask recessive deleterious alleles affecting fitness, along with a reduced efficacy of natural selection. Thus, at a genomic level, when demographic factors are mainly responsible for the accumulation of deleterious mutations, a random genome-wide distribution of deleterious variants is expected.

Anthropogenic climate change induces an elevation of global temperatures (Kerr, 2007; Marcott, Shakun, Clark, & Mix, 2013; Moss et al., 2010), particularly so in high latitudes (Paaijmans et al., 2013) and its effects are particularly evident in ectotherms. For instance, almost all fishes are ectothermic and thereby unable to adjust body temperature physiologically, which implies that their metabolism, reproduction and growth are strongly influenced by the temperature of the environment (Dickson & Graham, 2004). In fact, thermal tolerance is known to have a genetic determinism (e.g., Dammark, Ferchaud, Hansen, & Sørensen, 2018), and research to date on temperature adaptation and climate change adaptability in fishes has mainly focused on local or environment-dependent temperature-related adaptation in life history traits (Bradbury et al., 2010) particularly in salmonid fish (Hecht, Matala, Hess, & Narum, 2015; Narum, Campbell, Kozfkay, & Meyer, 2010).

The Brook Charr (*Salvelinus fontinalis*, Mitchill), an endemic salmonid of eastern North America (Scott & Crossman, 1998), ranks among the most highly genetically structured animal species (Castric, Bonney, & Bernatchez, 2001; Gyllensten, 1985; Ward, Woodwark, & Skibinski, 1994), with most of its genetic variance partitioned among major drainages (Angers & Bernatchez, 1998; Danzmann & Ihssen, 1995; Ferguson, Danzmann, & Hutchings, 1991; Perkins, Krueger, & May, 1993). Previous phylogeographic studies revealed that a single glacial lineage of Brook Charr colonized most of Northeastern America (Ferguson et al., 1991; Jones, Clay, & Danzmann, 1996, but see Angers & Bernatchez, 1998). Very low effective population sizes have been reported for this species, particularly so in lacustrine populations that are generally isolated with limited gene flow between them (Gossieaux, Bernatchez, Sirois, & Garant, 2019; median *Ne* for the studied populations here is 35, IC 32:38, data not shown). Many Brook Charr populations are also anadromous, a trait shared by many salmonid species which refers to a migratory life cycle whereby individuals are born in freshwater, feed and grow in salt water and return to freshwater to reproduce. Consequently, these populations are more connected by gene flow than resident ones (Castric & Bernatchez, 2003). Previous studies showed that a significant proportion of phenotypic variation in migratory traits is under some genetic control (Hecht, Thrower, Hale, Miller, & Nichols, 2012; Liedvogel, Åkesson, & Bensch, 2011; Pearse, Miller, Abadia-Cardoso, & Garza, 2014), including in Brook Charr (Boulet, Normandeau, Bougas, Audet, & Bernatchez, 2012; Crespel, Bernatchez, Audet, & Garant, 2013; Crespel, Bernatchez, Garant, & Audet, 2013). Freshwater migration harshness, in terms of length and elevation steepness, has a strong effect on the bioenergetic costs of migration in salmonids (Bernatchez & Dodson, 1987), which can drive local adaptation (Eliason et al., 2011; Moore et al., 2017; Pearse et al., 2014). Accordingly, previous studies conducted on the Brook Charr (*Salvelinus fontinalis*, Mitchill) suggested that ancestral populations may have differed in their ability to colonize certain habitats (Fraser & Bernatchez, 2005) and notably that divergence on physical factors affects swimming performance between anadromous and resident populations (Crespel et al., 2017).

In this study, we investigate the importance of demographic versus selective processes governing the accumulation of both adaptive and maladaptive (deleterious) mutations in the Brook Charr. By adaptive, we refer to variations linked to environmental variables, hypothesizing that some mutations are more beneficial in one environment than in another, despite the fact that no fitness values are available to corroborate this association. By maladaptive, we refer to variations having putative damaging effect on the individual fitness inferred from the known functionality of a given mutation (see below). More specifically, based on a GBS (genotype by sequencing) data set from samples of lacustrine, riverine and anadromous Brook Charr populations distributed throughout the province of Quebec (Canada), we aim to (a) assess the extent of genetic diversity

and differentiation among localities and habitats, (b) document the presence of putative adaptive mutations associated with environmental variables as well as maladaptive mutations and (c) estimate the recombination rate along the genome and its potential role in explaining the genome-wide distribution of those mutations. Finally, we combine all results in order to explore the relative roles of neutral and selective processes on the accumulation of maladaptive mutations among salmonid populations. We predict that iflimited gene flow and the occurrence of pronounced genetic drift are the main factors governing the accumulation of deleterious variants, we expect to observe a lower frequency of deleterious mutations among potentially more connected anadromous and river populations compared to more geographically isolated and unconnected lacustrine populations. Moreover, if relaxed purifying selection is the main factor explaining genetic load, we expect to observe a higher accumulation of deleterious mutations in genomic regions of low recombination rate. Arguably, intermediate patterns being observed for the defined variants would suggest an interaction of demographic and selective processes.

2 | METHODS

2.1 | Sampling and study system

A total of 1,416 individuals sampled from 50 sampling locations including 36 lacustrine, seven riverine and seven anadromous populations from 2014 to 2015 (for lakes and rivers) and from 2000 to 2001 for anadromous sites (Castric & Bernatchez, 2003) in Québec, Canada, were successfully genotyped (Figure 1, Table S1). The median number of individuals per site was 28 (Table S1). Fish were sampled by technicians of the Ministère des Forêts, de la Faune et des Parcs du Québec (MFFP) and the Société des Établissements de Plein Air du Québec (SEPAQ) using gillnets or by anglers using fishing rods. Adipose fin clips were preserved in 95% ethanol. Populations were chosen according to their geographic location, the documented absence of stocking and the availability of tissues from which good quality DNA could be extracted (A_{260}/A_{280} ratio between 1.7 and 2.0, and high molecular weight together with no smears while migrated on an agarose gel). Lake size ranged from 3 to 5,591 ha, with a median value of 66 ha (Table S1). Covered latitude ranged from 45.378 to 57.918 and longitude from −79.194 to −57.949. Records of water temperature were unavailable for most of the sites, and so, records of air temperature were used to estimate 10-year averages (2004–2015) of minimum, maximum, mean minimum, mean maximum and mean air temperature for each lake from the database available in BioSIM (Régnière, St-Amant, & Béchard, 2014). Air temperature has been shown to be linearly correlated with growth rate in *Salvelinus* and thus is a good variable to investigate the role of temperature in local adaptation (Black, von Biela, Zimmerman, & Brown, 2013; Ferchaud, Laporte, Laporte, Perrier, & Bernatchez, 2018; Perrier, Ferchaud, Ferchaud, Sirois, Thibault, & Bernatchez, 2017; Torvinen, 2017). Annual average air temperature ranged from −5.04 to 6.29°C (Figure 1, Table S1). Other

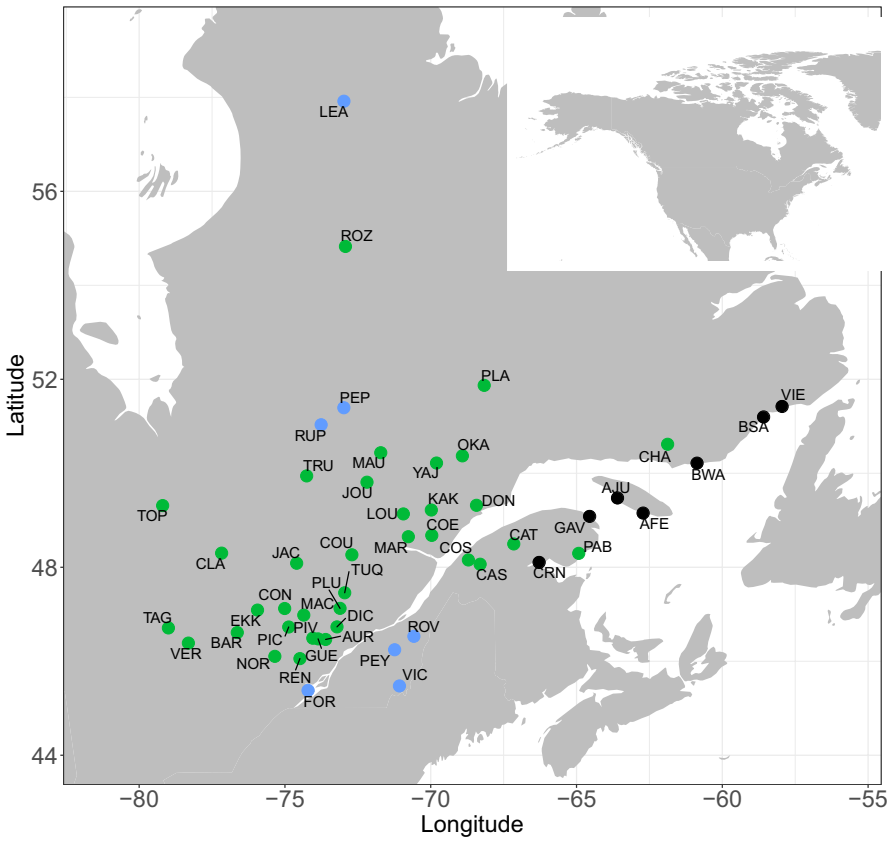

FIGURE 1 Geographic location of the 50 sampling sites studied across the province of Quebec, Canada: 36 lakes (in green), seven rivers (in blue) and seven anadromous sampling locations (in black). Each locality is labelled using a three-letter code displayed in black. The map was drawn using the R package ggmap (Kahle & Wickham, 2013) with R 3.1.0 (R Development Core Team, 2017) [Colour figure can be viewed at wileyonlinelibrary.com]

environmental variables of interest were gathered using BioSIM to get 10-year averages of dew point, frost-free days, total of radiation and growing degree-day (GDD), which was provided by Ministère des Forêts, de la Faune et des Parcs (MFFP) du Québec. Briefly, the basic concept of GDD is that fish growth and development will only occur if the temperature exceeds the so-called base temperature (TBASE). The base temperatures are determined experimentally and are different for each organism. To calculate GDDs, one must (a) estimate the mean temperature for the day (by adding together the high and low temperature for the day and dividing by two); and (b) compare the mean temperature to the TBASE. Then, if the mean temperature is at or below TBASE, the growing degree-day value will be zero. If the mean temperature is above TBASE, the growing degree-day amount will be equal to the mean temperature minus TBASE. For example, if the mean temperature was 25°C, then the GDD amount equals 10 for a TBASE of 15°C. For each locality, we averaged GDD values collected over a 10-year period before the sampling (2004–2015) using a base temperature of 5°C (temperature usually employed for cold water salmonids in our study area, MFFP personal communication). All the environmental variables are reported in Table S1.

2.2 | Molecular analyses

Total DNA was extracted from adipose fin tissue (50 mg) using a slightly modified version of Aljanabi and Martinez (1997) salt extraction protocol (extra ethanol wash). Sample concentration and quality werse checked using 1% agarose gel and a NanoDrop 2000 spectrophotometer (Thermo Scientific). DNA was quantified using the PicoGreen assay (Fluoroskan, Ascent FL; Thermo Labsystems). Genomic DNA was normalized to obtain 20 ng/µl in 10 µl (200 ng) for each individual. The sequencing libraries were created accordingly to Mascher's , Wu, Amand, Stein, and Poland (2013) protocol. Namely, in each sample, a digestion buffer (NEB4) and two restriction enzymes (*PstI* and *MspI*) were added. After a 2-hr incubation period at 37°C, enzymes were inactivated by a 20-min incubation period at 65°C. Then, the ligation of two adaptors was performed using a ligation master mix followed by the addition of T4 ligase and completed for each sample at 22°C for 2 hr. Enzymes were again inactivated by a 20-min incubation period at 65°C. Finally, samples were pooled in 96-plex and QIAquick PCR Purification Kit was used to clean and purify the DNA. After library PCR amplification, sequencing was performed on the Ion Torrent Proton P1v2 chip at the genomic analysis platform of IBIS (Institut de Biologie Intégrative et des Systèmes, Université Laval; http://www.ibis.ulaval.ca/en/home/).

2.3 | Bioinformatics

The program FastQC 0.11.1 (Andrews, 2010) was used to check raw reads for overall quality and presence of adapters. All the bioinformatic steps, options and software versions employed in the subsequent GBS pipeline are detailed in Table S2. Briefly, we used CUTADAPT v1.8.1 (Martin, 2011) to remove the adapter from raw sequences and STACKS v1.40 *process_radtags* to demultiplex the samples, truncate reads (to 80 bp) and do the quality trimming (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013). Since no reference genome is available for the Brook Charr, demultiplexed reads were aligned to a closely related sister species, the Arctic Charr (*Salvelinus alpinus*) reference genome (Christensen et al., 2018) with BWA MEM version 0.7.17 (Li & Durbin, 2009). Then, STACKS v2.40 was executed to call SNPs. In order to do that, the gstacks module was first used to assemble and merge paired-end contigs, call variant sites in the populations and genotypes in each sample (with the --max-clipped 0.1 argument). Second, populations module was executed to filter data (-p 40 and -r 0.5) and export in VCF format.

The output file was also filtered using custom script available in stacks_workflow (https://github.com/enormandeau/stacks_workflow) to retain high-quality SNPs. A genotype was kept if it was present in at least 70% of genotype data per population, if the minimal allele coverage was equal or above four and if at least two samples were harbouring the rare allele. Bad samples were removed for subsequent analysis, and thus, only individuals with less than 20% of missing genotypes and not showing outlier values for relatedness or heterozygosity were kept. Moreover, based on the approach proposed by McKinney, Waples, Seeb, and Seeb (2017) to confidently discriminate SNPs found in duplicated loci and isolate high confidence SNPs (i.e., singletons), we investigated SNPs "anomalies" according to a suite of four parameters (i.e., (a) median of allele ratio in heterozygotes (MedRatio), (b) proportion of heterozygotes (PropHet), (c) proportion of rare homozygotes (PropHomRare) and (d) F_{is}). The 09_classify_snps.py script in stacks_workflow was performed to classify SNPs according to this method (fully described in Dorant et al., 2020), and values of parameters are available in Table S2. Lastly, a final data set was built keeping the genotype calls from the VCF file containing only singleton SNPs, postfiltered by keeping only all unlinked SNPs within each locus using the 11_extract_unlinked_snps.py available in stacks_workflow (again, all details of this approach are presented in Dorant et al., 2020).

2.4 | Population differentiation

A principal component analysis (PCA) was performed using the function glPCA() of ADEGENET 2.1.2 package in R to seek a summary of the diversity among all localities. The R function ggplot2 (Wickham, 2009) was then used to plot the results of this PCA according to (a) the sampling locality origin of each sample and (b) the habitat type (anadromous, lake or river). Genetic structure among populations was also investigated with ADMIXTURE 1.3.0 program (Alexander, Novembre, & Lange, 2009) for K ranging from 2 to 60 across all localities. The cross-validation indices were used to discuss the best values of K across all localities. Both for PCA and Admixture analysis, the data set was converted into genlight object or plink file, respectively. The extent of genetic differentiation among sampling

locations was computed by average F_{ST} estimations with R package STAMPP 1.5.1 (Pembleton, Cogan, and Forster (2013) using the Rscript "StAMPP-fst.R" implemented in https://gitlab.com/YDorant/Toolbox. Those estimations were computed (a) among all locality pairs and then considering the three habitat types separately, (b) among pairs of anadromous locations, (c) among pairs of lacustrine locations and (d) among pairs of riverine locations. Both the pairwise population differentiation (F_{ST}) and the mean pairwise F_{ST} per population were reported. Finally, signals of isolation by distance were tested using a correlation between genetic differentiation (F_{st}/(1−F_{st})) and log(geographic distance) (Rousset, 1997) considering (a) all localities altogether and (b) only lacustrine locations. The correlation between river distance and genetic differentiation was also explored across three different river drainages (i.e., river drainages with at least three localities, i.e., two with four localities and one with nine localities). Results were not different from those obtained with geographic distances using the full data set which are therefore not shown.

2.5 | Population genetic diversity

The minor allele frequency (MAF) of each SNP was calculated within population using the –freq argument in PLINK (Purcell et al., 2007). The number of polymorphic SNPs was reported for each population. We documented the extent of genomic diversity within and among populations by estimating the nucleotide diversity "Π." Π was estimated within each population on a per-site basis using the –site-pi option in VCFtools (Danecek et al., 2011) and then averaged by population. *t* tests were performed in order to verify whether the genetic diversity (Π) was significantly different among the three different types of populations (lacustrine, riverine, anadromous). Furthermore, using linear models in R (lm() function), correlations between average Π (over individuals) and average pairwise F_{ST} (by population) were determined both against latitude and altitude.

2.6 | Proportions of maladaptive (deleterious) mutations

To identify putative deleterious mutations, all defined loci in which singleton SNPs have been called were first used in a BLAST query against the Arctic Charr (*Salvelinus alpinus*) proteome available on NCBI (https://www.ncbi.nlm.nih.gov/assembly/GCF_002910315.2) using blastx. Note that the Brook Charr transcriptome (used below for the annotation analysis) could not be used at this step since we ran a pipeline (see below) requesting a proteome in amino acid and not a transcriptome in nucleotides. All hits that had higher than 25 amino acid similarity and more than 60% similarity between the query read and the reference proteome were retained (those thresholds have been optimally chosen after running several tests with different parameters combinations on the observed data). A comparison of translation results between both variants of each SNP in retained reads was performed to discriminate the synonymous from the non-synonymous mutations. Then, PROVEAN (Protein Variation Effect Analyzer; Choi, Sims, Murphy, Miller, & Chan, 2012) was used to predict the deleterious effect of nonsynonymous mutations. As in previous studies (e.g., Ferchaud et al., 2018; Renaut & Rieseberg, 2015), the default threshold of −2.5 in PROVEAN score was applied to distinguish between nonsynonymous mutations potentially deleterious (≤−2.5) and neutral (>−2.5). In PROVEAN, the deleteriousness of a variant can be predicted based on its effect on gene functioning (such as protein changing, stop-gain, stop-loss), for example by assessing the degree of conservation of an amino acid residue across species. The pipeline used for the entire process is available on GitHub (gbs_synonymy_genome: https://github.com/QuentinRougemont/gbs_synonymy_with_genome). Results of PROVEAN have previously been compared in another salmonid (Lake Trout, *S. namaycush*) to another program PolyPhen-2, and results inferred from the two methods were highly congruent (Ferchaud et al., 2018). For SNPs harbouring a putative deleterious mutation, we hypothesize that the minor allele is the one having a deleterious effect (since it is present in lower frequency at a global level). Because deleterious alleles are selected against, theory predicts that their frequencies should be lower than nondeleterious ones (Fay et al., 2001). Thus, *t* tests were also used to verify whether average "deleterious allele frequency" was lower than the ones of "nonsynonymous, but non-deleterious SNPs" and to that of "anonymous" (SNPs that did not fall into coding regions) and "synonymous SNPs" within each habitat type. Next, the proportion of deleterious mutations (defined as the number of SNPs showing a deleterious mutation in a given population over the number of SNPs harbouring a deleterious mutation across all populations) and the ratio of the proportion of deleterious mutations over the proportion of polymorphic SNPs were estimated and compared among populations. The latter ratio is expected to be lower in populations that more efficiently purge deleterious mutations (see Perrier et al., 2017) and was compared between habitat types using a *t* test.

2.7 | Putative adaptive mutations: gene–environment associations

Redundancy analysis (RDA) was conducted as a multilocus genotype–environment association (GEA) method to detect loci under selection. RDA is an analogue of multivariate linear regression analyses, utilizing matrices of dependent and independent (explanatory) variables. The dependent matrix is represented by the genotypic data (here a matrix of minor allelic frequencies by population), and the explanatory variables are composed into an environmental matrix. Two analyses were performed. First, we aimed to identify SNPs associated with anadromy and then to identify SNPs associated with environmental variables among lacustrine populations.

Forester, Lasky, Wagner, and Urban (2018) recommend to not control by spatial autocorrelation for RDA multilocus outlier identification; however, because the sampled anadromous localities are mainly distributed in the eastern part of our sampling range,

we conducted two RDAs, one with and one without controlling for spatial autocorrelation. Then, we conservatively retained SNPs that were discovered to be significantly linked to anadromy by both approaches. In order to correct for spatial autocorrelation, a distance-based Moran's eigenvector map (db-MEM) based on latitude and longitude, all pretransformed in metres with the function "*geoXY*" of the R package "SoDA" was produced with the "pcnm()" function. The habitat (anadromous vs. resident [lakes + rivers]) constituted the explanatory matrix, and the db-MEMs related to spatial components constituted the conditioning matrix controlling for autocorrelation. The function "*rda*" was used to compute the RDAs on the model (see Laporte et al., 2016; Le Luyer et al., 2017; Marengo et al., 2017 for examples of similar methodology). An analysis of variance (ANOVA; 1,000 permutations) was then performed to assess the global significance of the RDAs, and the percentage of variance explained (PVE) was computed with the function "*RsquareAdj*". When not mentioned, R functions were part of the "VEGAN" package (Oksanen et al., 2019). SNPs linked to anadromy were then defined following instructions from the online tutorial proposed by Brenna Forester (Forester et al., 2018; https://popgen.nescent.org/2018-03-27_RDA_GEA.html). Only SNPs in common between both RDAs (with or without spatial autocorrelation) were considered as candidates underlying putative adaptation to anadromy. A heatmap of the frequency of those SNPs across populations was made using ggplot2 library on R.

Second, we identified SNPs associated with environmental variables across all lacustrine populations following the same instructions (https://popgen.nescent.org/2018-03-27_RDA_GEA.html). First, among all environmental variables, we removed variables with correlated predictors (*r* Pearson > .7) and thus kept only low minimum temperature, total of radiation and the GDD (hereafter, respectively, named LowTmin, AvRad and GDD) to identify candidate SNPs potentially involved in local adaptation. To identify the best model, the function "*ordistep*" was used to select the best explanatory variables. As recommended by Forester et al. (2018), we did not control for spatial autocorrelation among lacustrine locations since simulations showed that this diminishes the power of detection without decreasing false positives in similar demographic scenarios.

2.8 | Gene ontology

In order to investigate whether the putatively adaptive and maladaptive SNPs (defined above) were into coding regions or next to, we first annotated the Arctic Charr genome (used for the SNP calling) with the annotated Brook Charr transcriptome (Pasquier et al., 2016) using GWAN (https://github.com/enormandeau/gawn). Then, we considered 10,000 bp of flanking regions around the nucleotide sequences containing those candidate SNPs to record the annotated regions. Biological functions of those regions were then discussed in the light of the potential harmfulness or the environmental variables putatively associated with SNPs.

2.9 | Effect of recombination rate on the genome-wide distribution of putative adaptive and maladaptive mutations

To estimate genome-wide variation in recombination rate, we took advantage of the Brook Charr high-density linkage map (Sutherland et al., 2016) to retrieve the relative position of our loci (see Leitwein et al., 2017 for the detailed methods). Briefly, we anchored the 3,826 mapped RAD loci of the Brook Charr linkage map (Sutherland et al., 2016) to the Arctic Charr reference genome (Christensen et al., 2018). This was possible after controlling for synteny and collinearity between the Arctic Charr (Nugent, Easton, Norman, Ferguson, & Danzmann, 2017) and the Brook Charr linkage map (Sutherland et al., 2016) using the MAPCOMP pipeline (Sutherland et al., 2016). Then, we were able to create an ordered locus list for each of the 42 Brook Charr linkage groups. The local Brook Charr recombination rate was estimated by comparing the genetic positions (cM) retrieved from the Brook Charr density linkage map (Sutherland et al., 2016), to the physical positions (bp) retrieved from our reconstructed Brook Charr collinear reference genome with MAREYMAP (Rezvoy, Charif, Guéguen, & Marais, 2007). The degree of smoothing was set to 0.9 (*span*) for the polynomial regression method (Loess methods). To infer the recombination rates of the markers not included in the linkage map, we computed the weighted mean recombination rate using the closest marker on each side based on their relative physical positions. Putative adaptive and maladaptive mutations were plotted along the inferred recombination rate across the 42 Brook Charr linkage groups. Differences in mean recombination rate across types of mutations were tested between markers linked to environmental variables versus those that were not (e.g., linked vs. nonlinked to temperature) using a *t* test in R. For the nonsynonymous mutations, the relationship between PROVEAN scores and inferred recombination rate was explored using a linear model in R (lm() function).

3 | RESULTS

3.1 | DNA sequencing and genotyping

The total number of demultiplexed and cleaned reads was 2,989,523,556 with an average of 2,002,360 reads per individual. After DNA control quality and filtering out individuals with more than 20% of missing genotypes (70 individuals), individuals harbouring unusually high relatedness (two individuals) or heterozygosity (five individuals), 1,416 of 1,493 individuals (95%) (with an average number of 2,045,480 reads per individual) were kept for the subsequent analyses. After filtering for quality, 14,779 singleton SNPs distributed in 11,588 different loci were kept (see Table S2 for details at each step and stacks_workfow for reasoning). The approach developed by McKinney et al. (2017) and adapted by Dorant et al (2020) allowed removing 21,284 SNPs among which almost 70% were assigned to "duplicated" SNPs that are thus not correctly defined *via*

the stacks pipeline and might then be removed before conducting any subsequent analysis. In Appendix S1, we provide details and concerns about this filtering step that might be addressed in GBS studies in order to infer accurate biological patterns.

3.2 | Genetic structure and differentiation among localities

A pronounced pattern of population structure was observed among localities from both PCA and Admixture analysis. PCA showed evident clustering of individuals from the same locality (Figure S1a). Although there was some overlap among the three habitat types (Figure S1b), anadromous localities grouped together. Accordingly, analyses with the ADMIXTURE software for $K = 48, 50$ showed that most localities could be assigned to a unique genetic cluster (Appendix S1c, d, e). The most admixed individuals mainly belong to two river populations (VIC and PEY) in the southern part of the study range in accordance with their geographic proximity. Moreover, the seven anadromous sampling locations displayed less pronounced structure, the two localities from Anticosti Island (AFU & AJU) clustered together and the two northern ones (BSA & VIE) presented admixed individuals (Figure S1c). Overall, pairwise population differentiation (F_{ST}) ranged from 0.03 to 0.53 (median = 0.28) (Figure 2, Table S3), whereas mean pairwise F_{ST} per population ranged from 0.25 (VIC) to 0.60 (YAJ) (median = 0.40) (Table S4). Among lake populations, pairwise F_{ST} ranged from 0.12 to 0.82 (median = 0.46, Figure 2), while mean pairwise F_{ST} per population ranged from 0.29 (CON & CLA) to 0.62 (YAJ) (median = 0.45, Figure 2). Among river populations, pairwise population F_{ST} ranged from 0.07 to 0.40 (median = 0.23, Figure 2), while mean pairwise F_{ST} per population ranged from 0.18 (VIC) to 0.34 (FOR) (median = 0.23, Figure 2). Finally, anadromous populations displayed less pronounced pairwise F_{ST}, with pairwise population F_{ST} ranging from 0.03 to 0.19 (median = 0.14, Figure 2) and mean pairwise F_{ST} per population ranging from 0.12 (AFE & AJU) to 0.17 (BWA) (median = 0.13) (Figure 2 and Table S4). Together, these analyses confirmed a general pattern of genetic structure reflecting a situation of more constrained gene flow and

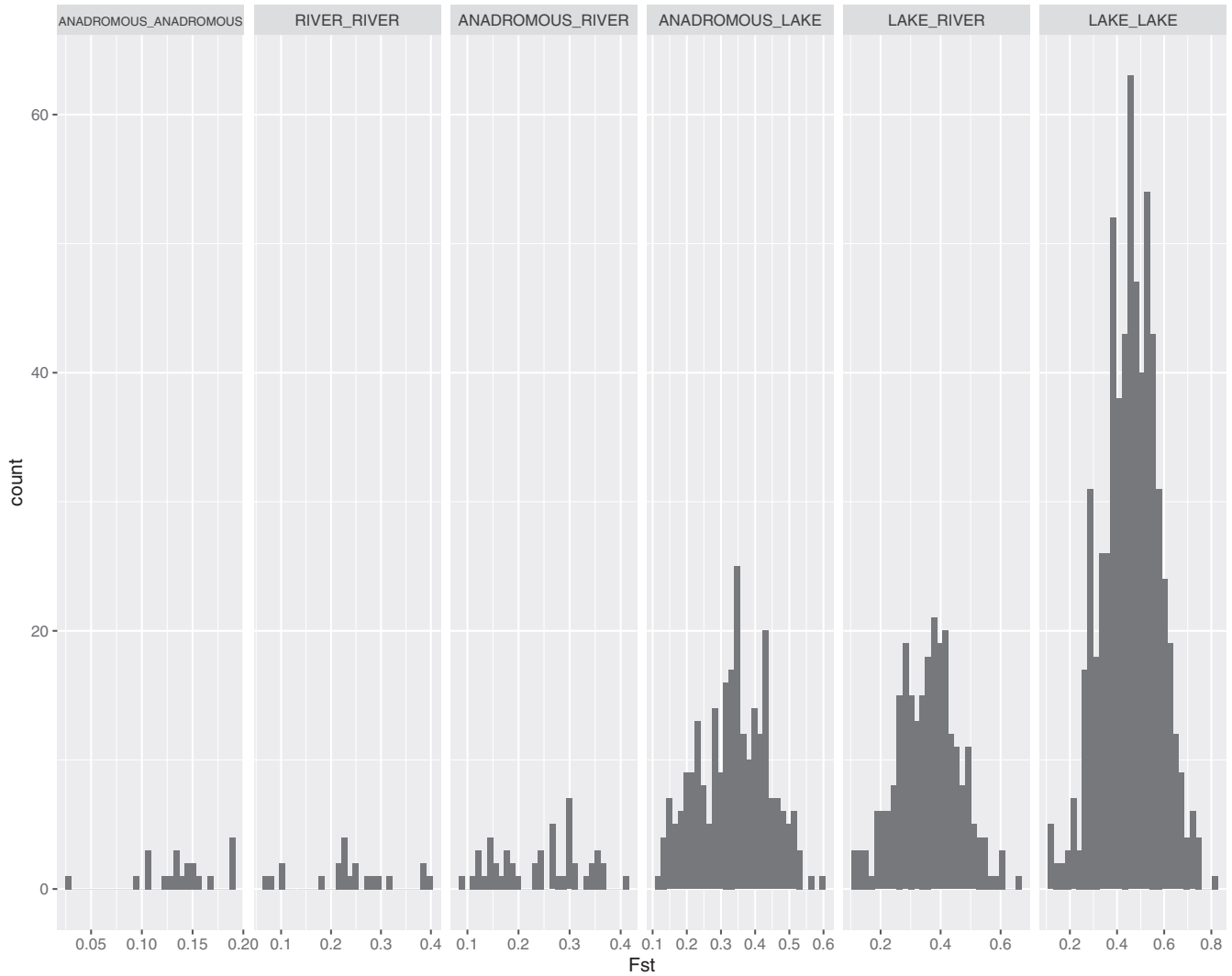

FIGURE 2 Histograms of F_{ST} estimates by categories of pairwise habitat types. The anadromous–anadromous comparisons displayed a distribution of F_{ST} in lower than river–river comparisons or event than lake–lake comparisons which displayed the highest F_{ST} values

more pronounced genetic drift in freshwater (lacustrine and riverine) than anadromous populations. Finally, there was very weak evidence of isolation by distance (IBD) among all populations (adj.R^2 = .004, p = .003) as well as when considering lacustrine populations only (adj.R^2 = .007, p = .001), thus suggesting highly constrained contemporary gene flow between resident populations.

3.3 | Population genetic diversity

Overall, the median proportion of polymorphic SNPs per population was 24%, ranging from 9% (pop YAJ) to 41% (pop CLA, Table S4). Over all 50 populations, the median value of average Π was 0.06, ranging from 0.02 (pop YAJ) to 0.11 (pop CLA, Table S4). Median values of Π were 0.06, 0.07 and 0.09 among lacustrine, riverine and anadromous populations, respectively. The mean Π was significantly lower among lake populations (mean Π_{lakes} = 0.06) than among river populations (mean Π_{rivers} = 0.08, $p_{t.test}$ < .01) and among anadromous populations (mean $\Pi_{anadromous}$ = 0.09, $p_{t.test}$ < .001). Mean Π among river populations was not significantly different from mean Π among anadromous populations ($p_{t.test}$ = .22). No significant correlation was found between nucleotide diversity and latitude (adj.R^2 = .02, p = .87, Figure 3a). However, a negative significant correlation was found between nucleotide diversity and altitude (adj.R^2 = .17, p < .01, Figure 3b). A similar pattern was observed for average F_{ST} which was not correlated to latitude (adj.R^2 = .01, p = .52, Figure 3c), but was found to be positively correlated with altitude (adj.R^2 = .16, p < .01, Figure 3d).

3.4 | Putatively maladaptive (deleterious) mutations

Among the 14,779 genotyped SNPs, 2,056 SNPs had significant BLAST results according to our criteria and were retained to assess synonymy. Synonymous substitutions were identified for 1,690 SNPs and nonsynonymous for 366 SNPs. According to the default deleterious threshold value (−2.5) for PROVEAN, we identified 23% of those nonsynonymous mutations (75/366) as putatively deleterious. The proportion of deleterious mutations within a given population varied from 0.01 (YAJ) to 0.19 (CON, DON & TAG) (Table S4). The median ratio of the proportion of deleterious mutations over the proportion of polymorphic SNPs was 0.41 and ranged from 0.15 (pop YAJ) to 0.6 (pop COE) (Table S4). No significant difference for this ratio was found across populations from the three habitat types (mean$_{lakes}$ = 0.40, mean$_{rivers}$ = 0.44, mean$_{anadromous}$ = 0.43, Table S5). However, the average MAF of

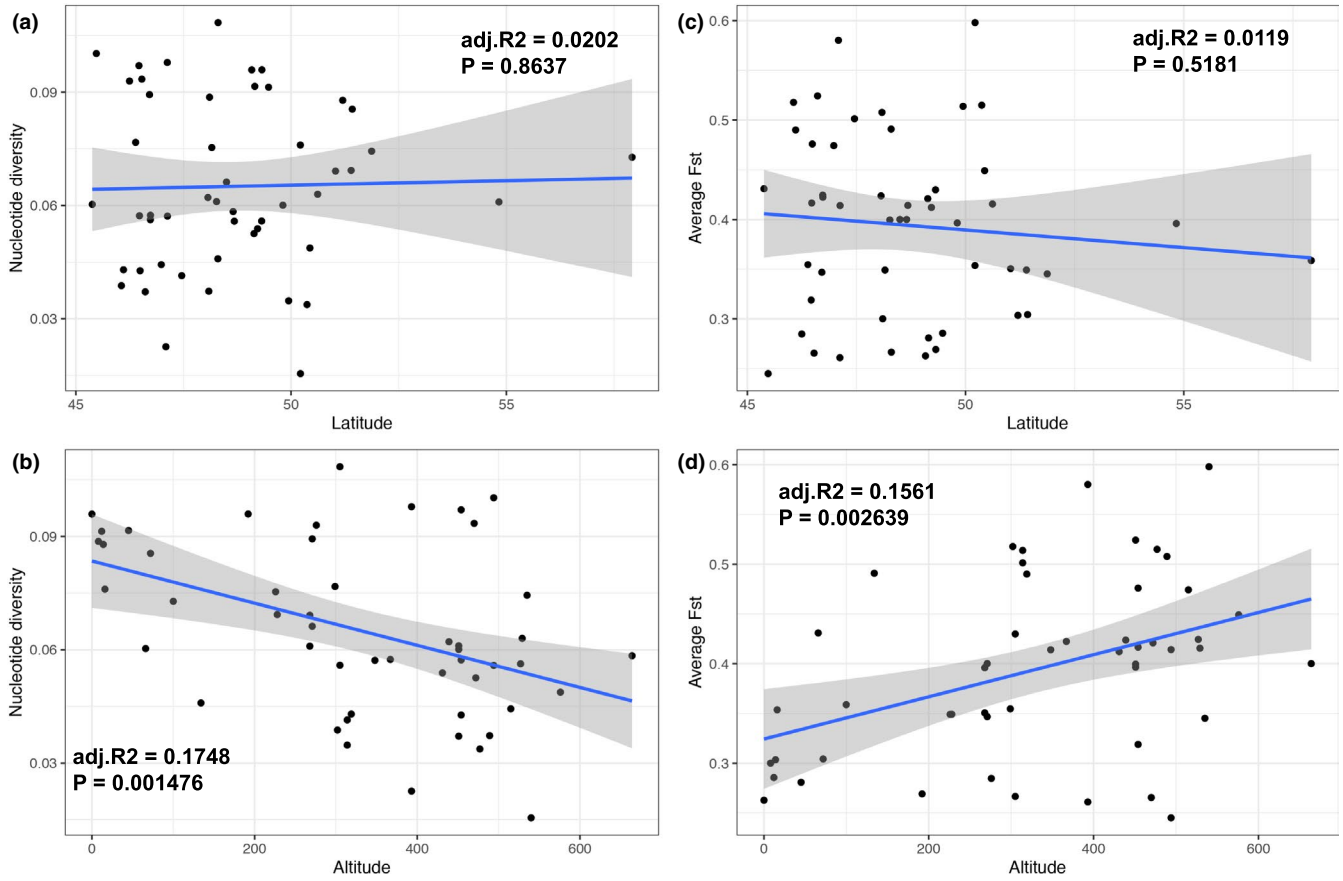

FIGURE 3 Relationships among genomic indices and environmental parameters. Relationship between average nucleotide diversity and (a) latitude and (b) altitude, as well as between mean Fst and (c) latitude and (d) altitude [Colour figure can be viewed at wileyonlinelibrary.com]

putative deleterious SNPs ($mean_{MAF\ deleterious}$ = 0.03) was significantly lower than anonymous ($mean_{MAF\ anonymous}$ = 0.08, $t = -22.909$, $p < .001$), as well as nonsynonymous but nondeleterious ($mean_{MAF\ nsnd}$ = 0.04, $t = -4.6621$, $p < .001$) and synonymous SNPs ($mean_{MAF\ synonymous}$ = 0.07, $t = -18.153$, $p < .001$, Figure S2). Within putative deleterious SNPs, no significant difference was observed between average MAF among populations from the three habitat types.

3.5 | Putatively adaptive mutations

3.5.1 | Anadromy

To correct for spatial autocorrelation, 28 axes were obtained from the db-MEM analysis, and among them, six were significant and therefore were retained as a spatial component *via* an explanatory matrix. A total of 230 SNPs were found to be associated with anadromy when correcting for spatial autocorrelation (Figure S3a), compared to 255 that were discovered without correction (Figure S3b; both analyses with $p_{outlier} < .001$). Forty-three SNPs were shared between these two lists and used as SNPs significantly associated with anadromy for subsequent analyses. The minor allele frequency (MAF) of those SNPs across all localities is represented in Figure 4, revealing a gradual distribution of frequency in populations according to their geographic position. Thus, when the MAF of a given locus is low in anadromous populations, it is generally high in lacustrine populations. However, a group of lacustrine populations exhibited a pattern of MAF more similar to nearby anadromous populations rather than with other lacustrine populations. Interestingly, this cluster of populations corresponds to populations that are located close to the sea with possible connectivity with the nearby anadromous populations (see Figures 1 and 4).

3.5.2 | GEA in lacustrine populations

Among lacustrine populations, the global model retained two out of the three environmental variables as significantly associated with genetic variability (LowTmin and GDD). LowTmin was mostly associated with the second axis of the RDA that explained 5.6% of the variation in allelic frequencies among lacustrine populations ($p = .002$), while GDD was associated with the first axis of the RDA explaining 6.4% of the variation ($p = .002$, Figure S4a). A total of 233 candidate SNPs represented a multilocus set of SNPs associated with GDD, while 261 candidates represented a multilocus set of SNPs associated with LowTmin ($p_{outlier} < .001$; Figure S4b). No SNPs were in common with putative deleterious mutations; accordingly, we retained those 233 and 261 loci, respectively, significantly associated with GDD and LowTmin loci ($p_{outlier} < .001$; Figure S4) and identified them as putative beneficial mutations linked to environmental conditions.

3.6 | Gene ontology

On average, among putative adaptive variants, 42% were located in annotated regions (49% for variants putatively associated with GDD, 35% for those with LowTmin and 42% for the ones linked with anadromy) and a similar proportion was observed for putative maladaptive variants (45% of putative deleterious SNPs were located in annotated regions) as well as for SNPs that were neither positive nor negative (52% of those SNPs were located in annotated regions). Among the DNA sequences successfully annotated, 60% (over all annotations) corresponded to transposons (Table S6), thus reflecting a statistically significant enrichment for transposable elements among adaptive and maladaptive markers relative to the reference transcriptome ($p < .001$). Sixty per cent of those transposons corresponded to Tc1-like class transposons (Table S6). The proportion

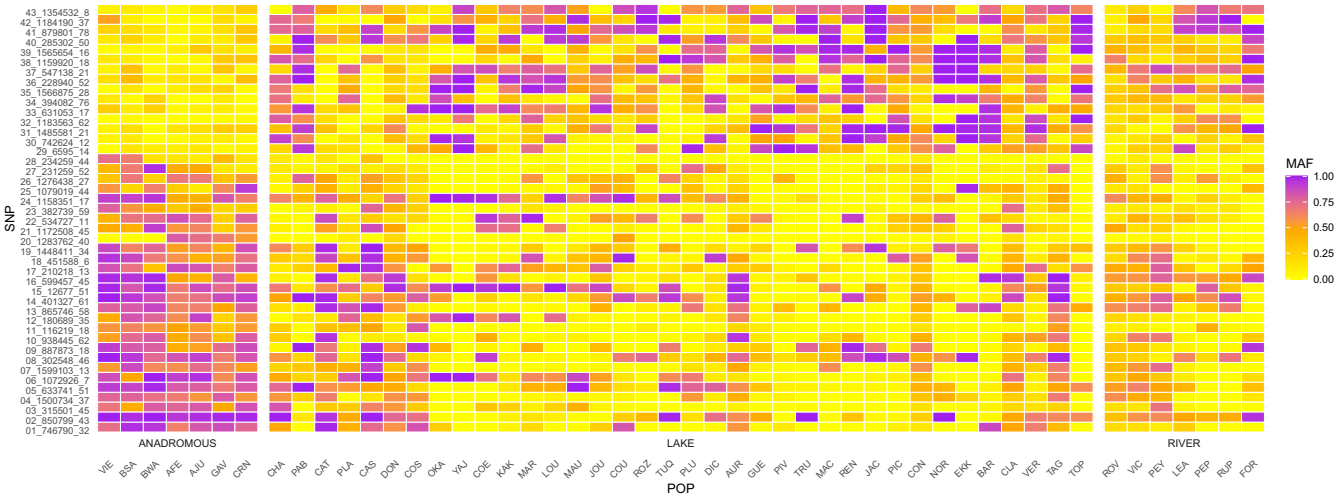

FIGURE 4 Heatmap of the minor allele frequency of the 43 SNPs identified as putatively associated with anadromy across the 50 Brook Charr sampling locations. Localities have been ordered according to their longitude from east to west [Colour figure can be viewed at wileyonlinelibrary.com]

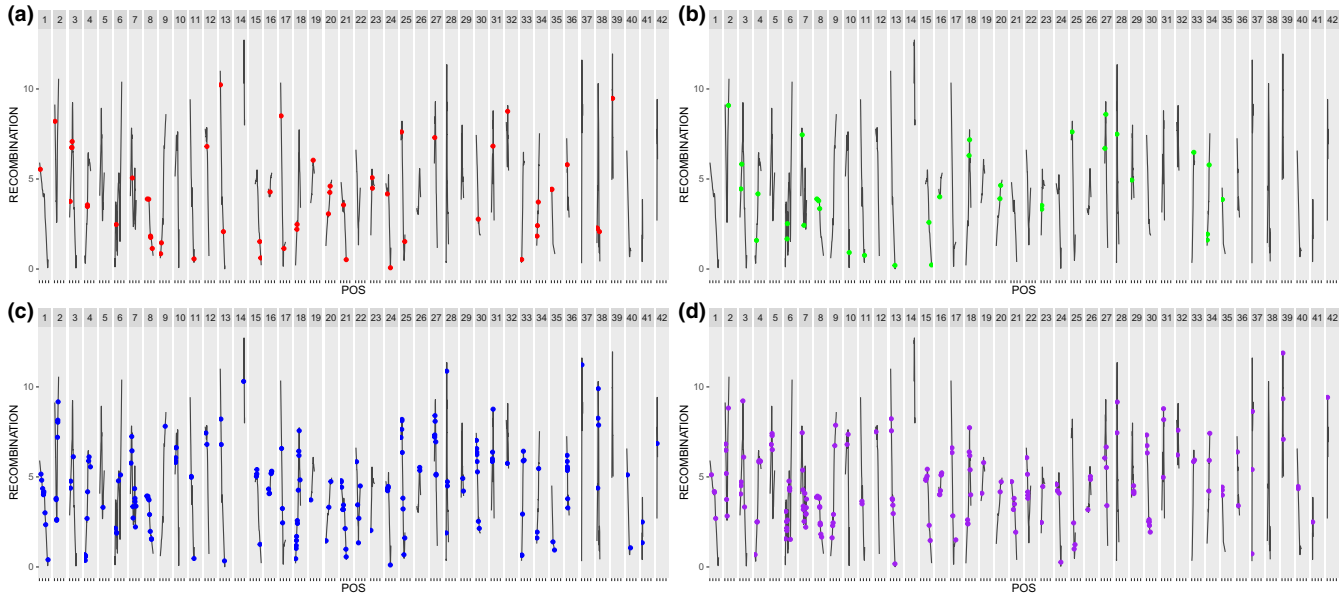

FIGURE 5 Inference of recombination rate along the Brook Charr linkage groups (in cM/Mbp). Both putative deleterious SNPs (a in red dots) and SNPs putatively associated with environmental variables are represented on their genomic localization, (b) orange dots represent the SNPs potentially linked to anadromy, (c) blue dots correspond to SNPs in putative association with LowTmin and the (d) purple ones to GDD [Colour figure can be viewed at wileyonlinelibrary.com]

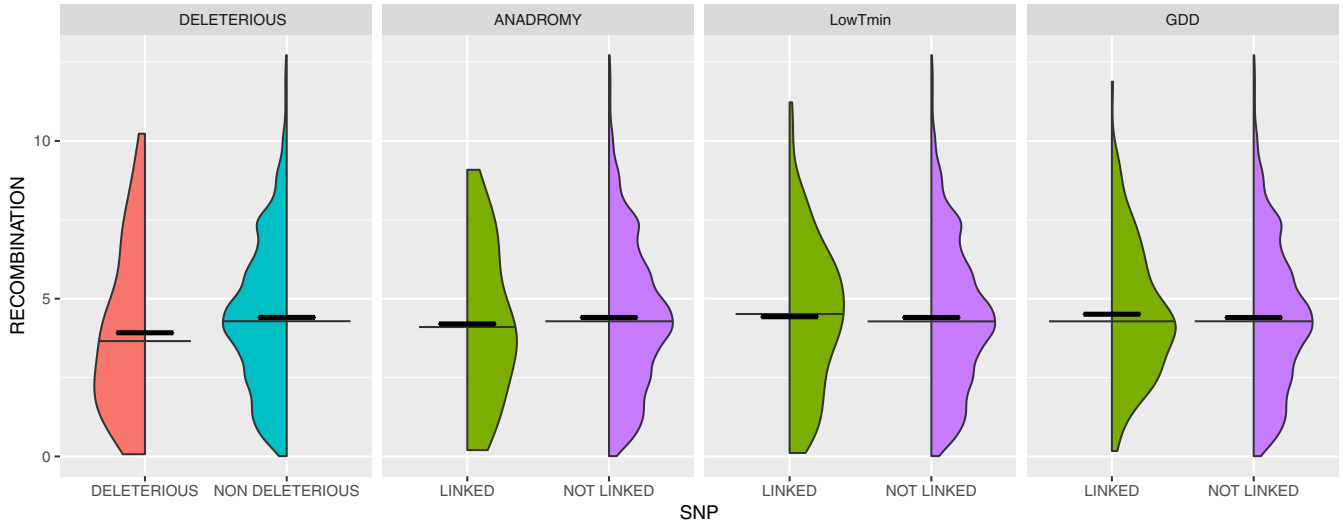

FIGURE 6 Boxplot of recombination rate (in cM/Mbp) across SNPs defined as putatively deleterious versus those that are nondeleterious and SNPs associated with environmental variable versus those not linked to the respective environmental variable [Colour figure can be viewed at wileyonlinelibrary.com]

of transposons was similar for DNA sequences associated with environmental variables (62% LowTmin, 61% for GDD and 67% Anadromy) and for DNA sequences including putatively deleterious variants (63%). Other than transposons, annotated DNA sequences included genes representing multiple biological functions potentially involved in local adaptation. For SNPs associated with anadromy, one gene (Fox-1) involved in nervous system was found as well as one related to thyroid hormone transport (monocarboxylate transporter 8) and one associated with apoptic process (CTCL tumour antigen se20-9) and one related to cell polarity (microtubule-actin cross-linking factor 1) (Table S6). Biological functions such as nervous system development, homeostasis, bone mineralization, cardiac muscle activity and oogenesis characterized genes associated with temperature (LowTmin). Genes involved in fatty acid metabolic processes (e.g., fatty acid desaturase 2) were discovered in genomic regions associated with GDD as well as other functions including retina development (Ran-binding protein 9), response to starvation (estrogen-induced gene 121 protein), hepatic development (glypican-1) or muscle contraction (dystrobrevin alpha). Two variants also involved in fatty acid metabolism were found as putatively

deleterious (fatty acid desaturase 2 and nucleoside diphosphate kinase B). Other biological functions associated with genes carrying putatively deleterious alleles comprised cell division, cytoskeleton, sensory perception of sound, nervous system development or spermatogenesis (Table S6).

3.7 | Effect of recombination rate on the genome-wide distribution of putative adaptive and maladaptive mutations

Variation of inferred recombination rate along the Brook Charr reconstructed genome ranged from 0.01 to 12.72 cM/Mbp (Figure 5). The average recombination rate per chromosome ranged from 2.15 cM/Mbp (CH 35) to 10.85 cM/Mbp (CH 14) with minimal recombination rate ranging from 0.01 cM/Mbp (CH 13) to 7.99 cM/Mbp (CH 14), and maximal recombination rate ranging from 3.86 cM/Mbp (CH 41) to 12.72 cM/Mbp (CH 14). The highest standard deviation of recombination rate within chromosome was in CH 12 (3.20 cM/Mbp), while CH 16 (0.49 cM/Mbp) harboured the lowest standard deviation. Regions encompassing putative deleterious mutations tended to show a lower mean recombination rate than for the rest of SNPs although this difference was not significant (mean $R_{del} = 3.92$, mean $R_{nondel} = 4.41$, $p = .16$, Figure 6). The absence of significant difference was also observed for SNPs associated with anadromy (mean $R_{ana} = 4.20$, mean $R_{nonana} = 4.40$, $p = .63$, Figure 6) and SNPs associated with environmental variables among lacustrine populations (mean $R_{GDD} = 4.51$, mean $R_{non\text{-}GDD} = 4.40$, $p = .51$; mean $R_{LowTmin} = 4.43$, mean $R_{non\text{-}LowTmin} = 4.40$, $p = .33$, Figure 6). Finally, for nonsynonymous variants, no correlation was found between PROVEAN scores and recombination rate (adj.$R^2 = .00$, $p = .68$, Figure S5).

4 | DISCUSSION

The goal of this study was to investigate the relative importance of demography and selection in the accumulation of both putative adaptive and maladaptive mutations in small populations. First, we detected a pronounced genetic structure among Brook Charr lacustrine populations along with a genome-wide reduced genetic diversity compared with river and anadromous populations. The very weak signal of IBD combined with the observed relationship between genetic diversity, differentiation and altitude suggests that colonization history has played an important role in the accumulation of both neutral and putative (mal)adaptive variants in this system. As expected in such systems with an overriding role of demographic processes on genetic variation, a non-negligible accumulation of deleterious mutations was found in Brook Charr. Moreover, our GBS results suggested that deleteriousness may not be modulated by recombination. Finally, functional genes found to be associated with environmental variables and/or anadromy support a significant effect of selection in maintaining local adaptation in the face of the apparent prevailing role of genetic drift in shaping genetic variation in Brook Charr populations.

4.1 | Pronounced genetic structure and reduced genetic diversity in lacustrine populations

Brook Charr ranks among the most highly structured animal species (Ward et al., 1994), and our results confirm it. The high level of population genetic structure observed in our Brook Charr samples has been previously reported (a) at large scale in which genetic variance is partitioned among major drainages or regions associated with distinct refugial origins (Danzmann, Morgan, Jones, Bernatchez, & Ihssen, 1998; Perkins et al., 1993), and (b) at a finer geographical scale (few kilometres) (Angers & Bernatchez, 1998; Angers, Magnan, Plante, & Bernatchez, 1999; Hébert, Danzmann, Jones, & Bernatchez, 2000). A stronger population genetic structure among lacustrine than riverine and anadromous populations highlights the fact that gene flow among Brook Charr populations depends on habitat availability favouring dispersion. The weak evidence of isolation by distance we observed also potentially reflects the highly constrained contemporary gene flow (albeit to a lesser extent in anadromous populations) and a departure from migration–drift equilibrium. Such a scenario was previously proposed to explain genetic differentiation among Brook Charr populations in Maine (USA) in which IBD exists but decreases among the most recently colonized northern populations (Castric et al., 2001). Our results corroborate and strengthen this pattern.

The reduced genetic diversity in lakes compared with rivers and anadromous populations further confirmed the pattern of a highly restricted contemporary gene flow. Moreover, the reduction of genetic diversity with altitude along the increase of differentiation strengthens the impact of colonization history on the observed genetic variability (Castric et al., 2001). Indeed, high-altitude populations are expected to be more physically isolated (and therefore more genetically differentiated), either because of increased probability of physical barriers to gene flow (e.g., impassable waterfalls) or due to more pronounced founder effects, assuming that the number of colonists decreases with altitude (Sagarin & Gaines, 2002).

4.2 | Accumulation of maladaptive (deleterious) mutations

According to population genetics theory, the accumulation of deleterious mutations depends on the relative roles of multiple factors including mutation rate, demographic history and selection (Whitlock & Bürger, 2004). In accordance with theoretical expectations, we observed that putative deleterious mutations were observed at lower frequencies than other types of within a given population. Overall, neither the accumulation of putative deleterious variants nor their predicted level of deleteriousness by PROVEAN was significantly correlated with recombination rate. Arguably, interpretations from

our GBS data set must be made cautiously since we have not sequenced the entire genome, and therefore certainly missed regions with deleterious mutations. Nevertheless, a random distribution of deleterious mutations along the genome prevails and suggests that (a) the role of putative linked selection may be negligible and that (b) genetic drift might be the major factor explaining the randomly potential harmful mutation genomic distribution. Moreover, no significant difference in the accumulation of deleterious mutations was observed among habitat types, suggesting that despite a lower genetic diversity (see above), lacustrine populations may not be subject to more pronounced inbreeding depression compared to riverine or anadromous populations despite higher genetic drift. Globally, albeit to a lower extent than previously reported in isolated lacustrine populations of Lake Trout (60% of nonsynonymous were putatively deleterious, Perrier et al., 2017) we observed a significant accumulation of deleterious mutations in Brook Charr populations (23% of nonsynonymous were putatively deleterious). For comparison, the percentage of nonsynonymous sites estimated to be deleterious in previous studies ranges from 3% in bacterial populations (Hughes, 2005) up to 80% in human populations (Fay et al., 2001). Such relatively abundant levels of putative deleterious mutations in Brook and Lake Trout may be attributed to their colonization history and to their small effective population sizes, which have been revealed to cause the accumulation of deleterious variants overtime (Benazzo et al., 2017; Grossen, Guillaume, Keller, & Croll, 2019). The lower occurrence of deleterious mutations in Brook Charr compared to Lake Trout could be imputable to differences in their life history. Lake Trout is a top predator and a large, long-lived fish restricted to lacustrine populations generally harbouring very small effective population sizes (Wilson and Mandrack, 2003), whereas Brook Charr is a smaller fish with a shorter generation time harbouring both resident and anadromous populations (Scott & Crossman, 1998). Lake Trout is almost entirely restricted to lacustrine environments, whereas Brook Charr is also commonly found in streams and rivers, therefore providing to this species more opportunity for connectivity among populations compared to Lake Trout.

Admittedly, however, our assessment of deleteriousness must be interpreted cautiously. Indeed, we used the functionality of a given mutation based on the degree of conservation of an amino acid residue across species (Choi et al., 2012). However, as in previous studies that applied this approach, those assumptions have not been experimentally validated, and there is no causal relationship available between the occurrence of these putative deleterious mutations and reduced fitness. Despite these caveats, those methods remain the most commonly used to assess the accumulation of deleterious mutations. For instance, they have recently been applied to estimate genetic load in human populations (Peischl et al., 2017) and domesticated plant species (Renaut & Rieseberg, 2015; Zhou, Massonnet, Sanjak, Cantu, & Gaut, 2017) and have been performed to validate theoretical predictions by empirical observations (Balick, Do, Cassa, Reich, & Sunyaev, 2015). Moreover, a very recent infatuation for such approaches has emerged for managing wild populations (Benazzo et al., 2017; Ferchaud et al., 2018; Grossen et al., 2019; Perrier et al., 2017; Zhu et al., 2018). Such studies indicate that in addition to adaptive considerations, management decisions must also adequately account for maladaptation as a potential outcome and even as a tool to bolster adaptive capacity to changing conditions (Derry et al., 2019). Here, we found a significant accumulation of deleterious mutations in populations from all habitat types, indicating that management decisions for Brook Charr must also adequately account for potential maladaptation. However, we did not find that isolated lacustrine trout populations accumulate more deleterious variants than more connected riverine or anadromous ones. This could pertain to the interacting effect between higher genetic drift due to smaller Ne and relaxed purifying selection. This may also suggest that inbreeding depression would not occur more often in lacustrine populations compared to other ones. On the other hand, generally lower genetic diversity suggests that the potential to adapt to new environmental conditions could be lower in lacustrine populations, which could therefore benefit from genetic rescue. Arguably, however, while genetic rescue in the form of supplementation could be beneficial, especially for populations of small effective size, this possible positive asset has to be considered cautiously as the use of the same domestic strain to supplement multiple populations can lead to an homogenization of the population genetic structure (Marie, Bernatchez, & Garant, 2010) and some loss of local adaptation *via* the disruption of co-adapted gene networks (Bougas, Granier, Audet, & Bernatchez, 2010; Lamaze, Sauvage, Marie, Garant, & Bernatchez, 2012).

4.3 | What are the putative adaptive and maladaptive mutations?

Among annotated regions identified for the putative adaptive and maladaptive mutations, a large proportion (63%) shows a statistically significant enrichment of transposable elements. Transposable elements (TEs) are short DNA sequences with the capacity to move between different genomic locations in the genome (Bourque et al., 2018). This ability may generate genomic mutations in many different ways, from subtle regulatory mutations to large genomics rearrangements (Casacuberta & González, 2013). The body of evidence regarding the adaptive role of TEs is growing (Casacuberta & González, 2013). In particular, TEs have been suggested to contribute to adaptation to climate change (Cayuela et al., 2020; Lerat et al., 2019; Rey, Danchin, Mirouze, Loot, & Blanchet, 2016; Shrader & Schmitz, 2019). TEs generate mutations actively by inserting themselves in new genomic locations and passively by acting as targets of ectopic recombination (recombination between homologous sequences that are not at the same position on homologous chromosomes). Also, some TEs contain regulatory sequences that can affect the structure and the expression of the host genes (Chuong, Elde, & Feschotte, 2017; Elbarbary, Lucas, & Maquat, 2016). Salmonid genomes contain a large amount of transcribed mobile elements (Krasnov, Koskinen, Afanasyev, & Mölsä, 2005). In particular, TEs are suspected to play a major role in reshaping genomes and playing

an important role in genome evolution (de Boer, Yazawa, Davidson, & Koop, 2007; Rodriguez & Arkhipova, 2018), especially following genome duplications (Ewing, 2015). Salmonids have undergone a whole-genome duplication event around 60 Mya (Crête-Lafrenière, Weir, & Bernatchez, 2012) followed by rediploidization events (i.e., return to stable diploid states) through genome rearrangements such as fusions and fissions (Sutherland et al., 2016). The majority of TEs we identified in Brook Charr belongs to Tc1-like class transposons (60%). Interestingly, the expression profiles of Tc1-like transposons (the most widespread mobile genetic elements) have been shown to be strongly correlated with genes implicated in defence response, signal transduction and regulation of transcription in salmonids (Krasnov et al., 2005). On the other hand, we also identified TEs associated with putative maladaptive variants (63% of annotated putative deleterious variants), as expected considering that previous studies showed that TEs can have profound deleterious consequences (Goodier, 2016). Previous studies have also evidenced the effect of TEs on deleteriousness in other salmonids. In common garden environments, hybrid breakdown in Lake whitefish (*Coregonus clupeaformis*) has been associated with TE reactivation caused by a genomic shock (genomic incompatibility) and associated hypomethylation in hybrids, leading to a malformed phenotype and aneuploidy (Dion-Côté, Renaut, Normandeau, & Bernatchez, 2014; Dion-Côté et al., 2017; Laporte et al., 2019). Overall, this highlights the importance of increasing our understanding of the (mal)adaptive role of such structural variants compared to SNP variation, especially in small populations (Mérot, Oomen, Tigano, & Wellenreuther, 2020).

Other biological functions identified for the putative beneficial and harmful mutations may be associated with maladaptation (for deleterious ones) or adaptation to anadromy and temperature. Genomic regions putatively associated with temperature corresponded to potentially relevant biological functions such as oogenesis or growth factors, as observed previously in Lake Trout (Perrier et al., 2017). Annotation results concerning the variants potentially associated with growing degree-day (GDD) also strengthen the importance of temperature as a selective agent acting on fish growth processes. Neuheimer and Taggart (2007) showed that fish length (corrected for calendar time and the associated variation of environmental variables such as temperature) was a strong linear function of the GDD metric that explains >92% of growth. However, our study is the first (to our knowledge) showing an association between GDD and genetic variation possibly underlying growth. Indeed, outliers putatively associated with GDD are located in genomic regions with biological functions related to fatty acid metabolic processes, retina development and response to starvation, all relevant in growth processes.

5 | CONCLUSION

The relative roles of demography and selection are keys in understanding the process of the maintenance of adaptive and maladaptive variation in small populations. In particular, understanding the factors that cause harmful mutations to increase in frequency in a genome will facilitate prediction of genetic load in current populations, and in this way contribute to improve conservation practices, for instance by providing a guidance for a best practice in genetic rescue. Indeed, increasing empirical evidence suggests that recently fragmented populations with reduced population sizes may receive demographic benefits from gene flow beyond the addition of immigrant individuals through genetic rescue (Frankham, 2015). However, most empirical studies conducted on wild populations to date are incomplete because of the difficulty of rigorously assessing the impact of detected genotype–environment or genotype–phenotype associations on individual fitness (but see Wells, Bernos, Yates, & Fraser, 2019). As such, our study adds to the increasing realization that in a context of a rapidly changing environment, the risk of maladaptation must be considered in planning conservation strategies.

ACKNOWLEDGEMENTS

We thank biologists and technicians of the Ministère des Forêts, de la Faune et des Parcs du Québec (MFFP) for their implication in the project and their field assistance as well as all the outfitters who contributed to the sampling. We also thank Guillaume Côté and Clément Rougeux for discussion since the preliminary result of this study. We are grateful to the team of the genomic analyses platform of IBIS (Institut de Biologie Intégrative des Systèmes) at Laval University. We are also grateful to the Associate Editor Narum and three anonymous reviewers for their constructive and detailed comments on a first version of this manuscript. This research was funded by the MFFP, the Canadian Research Chair in Genomics and Conservation of Aquatic Resources, Ressources Aquatiques Québec (RAQ) as well as by a Strategic Project Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) to L. Bernatchez.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTION

The study was conceived by L.B., I.T. and A-L.F. The samples were collected by A-L.F., I.T. and C.H. Laboratory work was carried out by B.B. and D.B. Bioinformatics analyses were performed by A-L.F. and E.N. Statistical analyses and writing the manuscript were carried out by A-L.F., M.Le. and M.La. All co-authors critically revised and contributed to edit the manuscript and approved the final version to be published.

DATA AVAILABILITY STATEMENT

Individual read raw sequences have been submitted at the Sequence Read Archive (SRA) (Project Accession Nos.: PRJNA638816 and PRJNA638817), and individual genotypes at all markers have been deposit in Dryad: https://doi.org/10.5061/dryad.n02v6wwv6

ORCID

Anne-Laure Ferchaud https://orcid.org/0000-0002-9577-5508
Maeva Leitwein https://orcid.org/0000-0001-5172-0970

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109

Aljanabi, S. M., Martinez, I. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Res*, *25*(22), 4692-4693. https://doi.org/10.1093/nar/25.22.4692

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*.

Angers, B., & Bernatchez, L. (1998). Combined use of SMM and Non-SMM methods to infer fine structure and evolutionary history of closely related Brook Charr (*Salvelinus fontinalis*, Salmonidae). *Molecular Biology and Evolution*, *15*, 143–159.

Angers, B., Magnan, P., Plante, M., & Bernatchez, L. (1999). Canonical correspondence analysis for estimating spatial and environmental effects on microsatellite gene diversity in Brook Charr (*Salvelinus fontinalis*). *Molecular Ecology*, *8*, 1043–1053. https://doi.org/10.1046/j.1365-294x.1999.00669.x

Balick, D. J., Do, R., Cassa, C. A., Reich, D., & Sunyaev, S. R. (2015). Dominance of deleterious alleles controls the response to a population bottleneck. *PLoS Genetics*, *11*, e1005436–e1005523. https://doi.org/10.1371/journal.pgen.1005436

Benazzo, A., Trucchi, E., Cahill, J. A., Maisano Delser, P., Mona, S., Fumagalli, M., … Bertorelle, G. (2017). Survival and divergence in a small group: The extraordinary genomic history of the endangered Apennine brown bear stragglers. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, E9589–E9597. https://doi.org/10.1073/pnas.1707279114

Bernatchez, L., & Dodson, J. J. (1987). Why do so few anadromous populations minimize the energetic cost of their upstream migrations? *American Fisheries Society Symposium*, *1*, 556.

Black, B. A., von Biela, V. R., Zimmerman, C. E., & Brown, R. J. (2013). Lake trout otolith chronologies as multidecadal indicators of high-latitude freshwater ecosystems. *Polar Biology*, *36*, 147–153. https://doi.org/10.1007/s00300-012-1245-9

Bougas, B., Granier, S., Audet, C., & Bernatchez, L. (2010). The transcriptional landscape of cross-specific hybrids and its possible link with growth in Brook Charr (*Salvelinus fontinalis* Mitchill). *Genetics*, *186*, 97–107.

Boulet, M., Normandeau, É., Bougas, B., Audet, C., & Bernatchez, L. (2012). Comparative transcriptomics of anadromous and resident Brook Charr. *Current Zoology*, *58*, 155–167.

Bourque, G., Burns, K. H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., … Feschotte, C. (2018). Ten things you should know about transposable elements. *Genome Biology*, *19*, 199. https://doi.org/10.1186/s13059-018-1577-z

Bradbury, I. R., Hubert, S., Higgins, B., Borza, T., Bowman, S., Paterson, I. G., … Bentzen, P. (2010). Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 3725–3734. https://doi.org/10.1098/rspb.2010.0985

Casacuberta, E., & González, J. (2013). The impact of transposable elements in environmental adaptation. *Molecular Ecology*, *22*, 1503–1517. https://doi.org/10.1111/mec.12170

Castric, V., & Bernatchez, L. (2003). The rise and fall of isolation by distance in the anadromous Brook Charr (*Salvelinus fontinalis*, Mitchill). *Genetics*, *163*, 983–996.

Castric, V., Bonney, F., & Bernatchez, L. (2001). Landscape structure and hierarchical genetic diversity in the Brook Charr, *Salvelinus fontinalis*. *Evolution*, *55*, 1016–1028. https://doi.org/10.1554/0014-3820(2001)055[1016:LSAHGD]2.0.CO;2

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140. https://doi.org/10.1111/mec.12354

Cayuela, H., Dorant, Y., Mérot, C., Laporte, M., Normandeau, E., Gagnon-Harvey, S., … Bernatchez, L. (2020). Thermal adaptation rather than demographic history drives genetic structure inferred by copy number variants in a marine fish. *bioRxiv*, *17*, 54–25.

Charlesworth, B. (2009). Fundamental concepts in genetics: Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, *10*, 195–205. https://doi.org/10.1038/nrg2526

Charlesworth, B. (2012). The effects of deleterious mutations on evolution at linked sites. *Genetics*, *190*, 5–22. https://doi.org/10.1534/genetics.111.134288

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., & Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS One*, *7*, e46688–e46713. https://doi.org/10.1371/journal.pone.0046688

Christensen, K. A., Rondeau, E. B., Minkley, D. R., Leong, J. S., Nugent, C. M., Danzmann, R. G., … Koop, B. F. (2018). The Arctic charr (*Salvelinus alpinus*) genome and transcriptome assembly (A Palsson, Ed,). *PLoS One*, *13*, 1–30. https://doi.org/10.1371/journal.pone.0204076

Chun, S., & Fay, J. C. (2011). Evidence for hitchhiking of deleterious mutations within the human genome. *PLoS Genetics*, *7*, e1002240. https://doi.org/10.1371/journal.pgen.1002240

Chuong, E. B., Elde, N. C., & Feschotte, C. (2017). Regulatory activities of transposable elements: From conflicts to benefits. *Nature Reviews Genetics*, *18*(2), 71–86.

Crespel, A., Bernatchez, L., Audet, C., & Garant, D. (2013). Strain specific genotype-environment interactions in Brook Charr (*Salvelinus fontinalis*). *G3: Genes, Genomes, Genetics*, *3*, 379–385.

Crespel, A., Bernatchez, L., Garant, D., & Audet, C. (2013). Genetically based population divergence in overwintering energy mobilization in Brook Charr (*Salvelinus fontinalis*). *Genetica*, *141*, 51–64. https://doi.org/10.1007/s10709-013-9705-x

Crespel, A., Dupont-Prinet, A., Bernatchez, L., Claireaux, G., Tremblay, R., Audet, C. (2017). Divergence in physiological factors affecting swimming performance between anadromous and resident populations of Brook Charr *Salvelinus fontinalis*. *Journal of Fish Biology*, *90*, 2170–2193.

Crête-Lafrenière, A., Weir, L. K., & Bernatchez, L. (2012). Framing the Salmonidae family phylogenetic portrait: a more complete picture from increased taxon sampling (D Fontaneto, Ed,). *PLoS One*, *7*, e46662–e46719. https://doi.org/10.1371/journal.pone.0046662

Crow, J. F. (1958). Some possibilities for measuring selection intensities in man. *Human Biology*, *30*(1), 1–13.

Dammark, K. B., Ferchaud, A.-L., Hansen, M. M., & Sørensen, J. G. (2018). Heat tolerance and gene expression responses to heat stress in threespine sticklebacks from ecologically divergent environments. *Journal of Thermal Biology*, *75*, 88–96. https://doi.org/10.1016/j.jtherbio.2018.06.003

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., … Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danzmann, R. G., & Ihssen, P. E. (1995). A phylogeographic survey of Brook Charr (*Salvelinus fontinalis*) in Algonquin Park, Ontario based upon mitochondrial DNA variation. *Molecular Ecology*, *4*, 681–697. https://doi.org/10.1111/j.1365-294X.1995.tb00269.x

Danzmann, R. G., Morgan, R. P., Jones, M. W., Bernatchez, L., & Ihssen, P. E. (1998). A major sextet of mitochondrial DNA phylogenetic assemblages extant in eastern North American brook trout (*Salvelinus fontinalis*): Distribution and postglacial dispersal patterns. *Canadian Journal of Zoology*, *76*, 1300–1318.

de Boer, J. G., Yazawa, R., Davidson, W. S., & Koop, B. F. (2007). Bursts and horizontal evolution of DNA transposons in the speciation of pseudotetraploid salmonids. *BMC Genomics*, *8*, 422. https://doi.org/10.1186/1471-2164-8-422

Derry, A. M., Fraser, D. J., Brady, S. P., Astorg, L., Lawrence, E. R., Martin, G. K., … Crispo, E. (2019). Conservation through the lens of (mal)adaptation: Concepts and meta-analysis. *Evolutionary Applications*, *12*, 1287–1304. https://doi.org/10.1111/eva.12791

Dickson, K. A., & Graham, J. B. (2004). Evolution and consequences of endothermy in fishes. *Physiological and Biochemical Zoology*, *77*, 998–1018. https://doi.org/10.1086/423743

Dion-Côté, A.-M., Renaut, S., Normandeau, É., & 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. https://doi.org/10.1093/molbev/msu069

Dion-Côté, A.-M., Symonová, R., Lamaze, F. C., Pelikánová, S., Ráb, P., Bernatchez, L. (2017). Standing chromosomal variation in Lake Whitefish species pairs: the role of historical contingency and relevance for speciation. *Molecular Ecology*, *26*, 178–192.

Dorant, Y., Cayuela, H., Wellband, K., Laporte, M., Rougemont, Q., Mérot, C., … Bernatchez, L. (2020). Copy number variants outperform SNPs to reveal genotype-temperature association in a marine species. *Molecular Ecology* (in press).

Elbarbary, R. A., Lucas, B. A., & Maquat, L. E. (2016). Retrotransposons as regulators of gene expression. *Science*, *351*, aac7247. https://doi.org/10.1126/science.aac7247

Eliason, E. J., Clark, T. D., Hague, M. J., Hanson, L. M., Gallagher, Z. S., Jeffries, K. M., … Farrell, A. P. (2011). Differences in thermal tolerance among sockeye salmon populations. *Science*, *332*(6025), 109–112. https://doi.org/10.1126/science.1199158

Ewing, A.D. (2015). Transposable element detection from whole genome sequence data. *Mobile DNA*, *6*, 24. https://doi.org/10.1186/s13100-015-0055-3

Fagan, W. F., & Holmes, E. E. (2006). Quantifying the extinction vortex. *Ecology Letters*, *9*, 51–60.

Fay, J. C., Wyckoff, G. J., & Wu, C.-I. (2001). Positive and negative selection on the human genome. *Genetics*, *158*, 1227–1234.

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics*, *78*, 737–756.

Ferchaud, A.-L., Laporte, M., Perrier, C., & Bernatchez, L. (2018). Impact of supplementation on deleterious mutation distribution in an exploited salmonid. *Evolutionary Applications*, *11*, 1053–1065. https://doi.org/10.1111/eva.12660

Ferguson, M. M., Danzmann, R., & Hutchings, J. A. (1991). Incongruent estimates of population differentiation among Brook Charr, *Salvelinus fontinalis*, from Cape Race, Newfoundland, Canada, based upon allozyme and mitochondrial DNA variation. *Journal of Fish Biology*, *39*, 79–85.

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*, *27*, 2215–2233. https://doi.org/10.1111/mec.14584

Frankham, R. (2015). Genetic rescue of small inbred populations: Meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology*, *24*, 2610–2618. https://doi.org/10.1111/mec.13139

Fraser, D., & Bernatchez, L. (2005). Adaptive migratory divergence among sympatric Brook Charr populations. *Evolution*, *59*, 611–624. https://doi.org/10.1554/04-346

Goodier, J. L. (2016). Restricting retrotransposons: A review. *Mobile DNA*, *7*, 16. https://doi.org/10.1186/s13100-016-0070-z

Gossieaux, P., Bernatchez, L., Sirois, P., & Garant, D. (2019). Impacts of stocking and its intensity on effective population size in Brook Charr (*Salvelinus fontinalis*) populations. *Conservation Genetics*, *20*, 729–742. https://doi.org/10.1007/s10592-019-01168-2

Grossen, C., Guillaume, F., Keller, L. F., & Croll, D. (2019). Accumulation and purging of deleterious mutations through severe bottlenecks in ibex. *bioRxiv*, *170*, 56–31.

Gyllensten, U. (1985). The genetic structure of fish: differences in the intraspecific distribution of biochemical genetic variation between marine, anadromous, and freshwater species. *Journal of Fish Biology*, *26*, 691–699. https://doi.org/10.1111/j.1095-8649.1985.tb04309.x

Hébert, C., Danzmann, R. G., Jones, M. W., & Bernatchez, L. (2000). Hydrography and population genetic structure in Brook Charr (*Salvelinus fontinalis*, Mitchill) from eastern Canada. *Molecular Ecology*, *9*, 971–982. https://doi.org/10.1046/j.1365-294x.2000.00965.x

Hecht, B. C., Matala, A. P., Hess, J. E., & Narum, S. R. (2015). Environmental adaptation in Chinook salmon (*Oncorhynchus tshawytscha*) throughout their North American range. *Molecular Ecology*, *24*, 5573–5595.

Hecht, B. C., Thrower, F. P., Hale, M. C., Miller, M. R., & Nichols, K. M. (2012). Genetic architecture of migration-related traits in rainbow and steelhead trout, *Oncorhynchus mykiss*. *G3: Genes*, *Genomes*, *Genetics*, *2*, 1113–1127.

Hughes, A. L. (2005). Evidence for abundant slightly deleterious polymorphisms in bacterial populations. *Genetics*, *169*, 533–538. https://doi.org/10.1534/genetics.104.036939

Jones, M. W., Clay, D., & Danzmann, R. G. (1996). Conservation genetics of brook trout (*Salvelinus fontinalis*): Population structuring in Fundy National Park, New Brunswick, and eastern Canada. *Canadian Journal of Fisheries and Aquatic Sciences*, *53*, 2776–2791.

Kahle, D., & Wickham, H. (2013). ggmap: Spatial visualization with ggplot2. *The R Journal*, *5*, 144–161. https://doi.org/10.32614/RJ-2013-014

Kerr, R. (2007). Global warming is changing the world. *Science*, *316*, 188–190.

Krasnov, A., Koskinen, H., Afanasyev, S., & Mölsä, H. (2005). Transcribed TcI-like transposons in salmonid fish. *BMC Genomics*, *6*, 107.

Lamaze, F. C., 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.

Lande, R. (1988). Genetics and demography in biological conservation. *Science*, *241*, 1455–1460. https://doi.org/10.1126/science.3420403

Laporte, M., Le Luyer, J., Rougeux, C., Côté-Dion, A.-M., Krick, M., & Bernatchez, L. (2019). DNA methylation reprogramming, TE derepression and postzygotic isolation of nascent species. *Science Advances*, *5*(10), eaaw1644.

Laporte, M., Pavey, S. A., Rougeux, C., Pierron, F., Lauzent, M., Budzinski, H., … Bernatchez, L. (2016). RAD sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Molecular Ecology*, *25*, 219–237. https://doi.org/10.1111/mec.13466

Le Luyer, J., Laporte, M., Beacham, T. D., Kaukinen, K. H., Withler, R. E., Leong, J. S., … Bernatchez, L. (2017). Parallel epigenetic modifications induced by hatchery rearing in a Pacific salmon. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 12964–12969.

Leitwein, M., Guinand, B., Pouzadoux, J., Desmarais, E., Berrebi, P., Gagnaire, P.-A. (2017). A dense brown trout (*Salmo trutta*) linkage map reveals recent chromosomal rearrangements in the *Salmo* genus and the impact of selection on linked neutral diversity. *G3 (Bethesda)*, *7*, 1365–1376.

Lerat, E., Goubert, C., Guirao-Rico, S., Merenciano, M., Dufour, A.-B., Vieira, C., & González, J. (2019). Population-specific dynamics and selection patterns of transposable element insertions in European natural populations. *Molecular Ecology*, *28*, 1506–1522. https://doi.org/10.1111/mec.14963

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liedvogel, M., Åkesson, S., & Bensch, S. (2011). The genetics of migration on the move. *Trends in Ecology and Evolution*, *26*, 561–569. https://doi.org/10.1016/j.tree.2011.07.009

Lynch, M., Conery, J., & Buerger, R. (1995). Mutation accumulation and the extinction of small populations. *The American Naturalist*, *146*, 489–518. https://doi.org/10.1086/285812

Marcott, S. A., Shakun, J. D., Clark, P. U., & Mix, A. C. (2013). A reconstruction of regional and global temperature for the past 11,300 years. *Science*, *339*, 1198–1201. https://doi.org/10.1126/science.1228026

Marengo, M., Baudouin, M., Viret, A., Laporte, M., Berrebi, P., Vignon, M., … Durieux, E. D. H. (2017). Combining microsatellites, otholith shape and parasites community data as a holistic approach to assess populations structure of common dentex (*Dentex dentex*) around Corsica Island (NW Mediterranean). *Journal of Sea Research*, *128*, 1–14.

Marie, A. D., Bernatchez, L., & Garant, D. (2010). Loss of genetic integrity correlates with stocking intensity in Brook Charr (*Salvelinus fontinalis*). *Molecular Ecology*, *19*, 2025–2037. https://doi.org/10.1111/j.1365-294X.2010.04628.x

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*, 10–12.

Mascher, M., Wu, S., Amand, P. S., Stein, N., & Poland, J. (2013). Application of genotyping-by-sequencing on semiconductor sequencing platforms: A comparison of genetic and reference-based marker ordering in barley. *PLoS One*, *8*, e76925–e77011. https://doi.org/10.1371/journal.pone.0076925

McKinney, G. J., Waples, R. K., Seeb, L. W., & Seeb, J. E. (2017). Paralogs are revealed by proportion of heterozygotes and deviations in read ratios in genotyping-by-sequencing data from natural populations. *Molecular Ecology Resources*, *17*, 656–669. https://doi.org/10.1111/1755-0998.12613

Mérot, C., Oomen, R. A., Tigano, A., & Wellenreuther, M. (2020). A roadmap for understanding the evolutionary significance of structural genomic variation. *Trends in Ecology and Evolution*, *35*(7), 561–572.

Moore, J.-S., Harris, L. N., Le Luyer, J., Sutherland, B. J. G., Rougemont, Q., Tallman, R. F., … Bernatchez, L. (2017). Genomics and telemetry suggest a role for migration harshness in determining overwintering habitat choice, but not gene flow, in anadromous Arctic Char. *Molecular Ecology*, *26*, 6784–6800. https://doi.org/10.1111/mec.14393

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., … Wilbanks, T. J. (2010). The next generation of scenarios for climate change research and assessment. *Nature*, *463*, 747–756. https://doi.org/10.1038/nature08823

Narum, S. R., Campbell, N. R., Kozfkay, C. C., & Meyer, K. A. (2010). Adaptation of redband trout in desert and montane environments. *Molecular Ecology*, *19*, 4622–4637. https://doi.org/10.1111/j.1365-294X.2010.04839.x

Neuheimer, A. B., & Taggart, C. T. (2007). The growing degree-day and fish size-at-age: The overlooked metric. *Canadian Journal of Fisheries and Aquatic Sciences*, *64*, 375–385. https://doi.org/10.1139/f07-003

Nugent, C. M., Easton, A. A., Norman, J. D., Ferguson, M. M., & Danzmann, R. G. (2017). A SNP based linkage map of the arctic charr (*Salvelinus alpinus*) genome provides insights into the diploidization process after whole genome duplication. *G3: Genes, Genomes, Genetics*, *7*, 543–556.

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., … Wagner, H. (2019). *vegan: Community Ecology Package. R package version 2.5-4*. Retrieved from https://CRAN.R-project.org/package=vegan

Paaijmans, K. P., Heinig, R. L., Seliga, R. A., Blanford, J. I., Blanford, S., Murdock, C. C., & Thomas, M. B. (2013). Temperature variation makes ectotherms more sensitive to climate change. *Global Change Biology*, *19*, 2373–2380. https://doi.org/10.1111/gcb.12240

Pasquier, J., Cabau, C., Nguyen, T., Jouanno, E., Severac, D., Braasch, I., … Bobe, J. (2016). Gene evolution and gene expression after whole genome duplication in fish: The PhyloFish database. *BMC Genomics*, *17*, 368. https://doi.org/10.1186/s12864-016-2709-z

Pearse, D. E., Miller, M. R., Abadia-Cardoso, A., & Garza, J. C. (2014). Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20140012. https://doi.org/10.1098/rspb.2014.0012

Peischl, S., Dupanloup, I., Foucal, A., Jomphe, M., Bruat, V., Grenier, J.-C., … Excoffier, L. (2017). Relaxed selection during a recent human expansion. *Genetics*, *208*, 763–777. https://doi.org/10.1534/genetics.117.300551

Pembleton, L., Cogan, N., & Forster, J. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, *13*, 946–952. https://doi.org/10.1111/1755-0998.12129

Perkins, D. L., Krueger, C. C., & May, B. (1993). Heritage brook trout in Northeastern USA: genetic variability within and among populations. *Transactions of the American Fisheries Society*, *122*, 515–532. https://doi.org/10.1577/1548-8659(1993)122<0515:HBTINU>2.3.CO;2

Perrier, C., Ferchaud, A.-L., Sirois, P., Thibault, I., & Bernatchez, L. (2017). Do genetic drift and accumulation of deleterious mutations preclude adaptation? Empirical investigation using RADseq in a northern lacustrine fish. *Molecular Ecology*, *26*, 6317–6335. https://doi.org/10.1111/mec.14361

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, *81*, 559–575.

R Development Core Team (2007). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Régnière, J., St-Amant, R., & Béchard, A. (2014). *BioSIM 10 User's manual*. Natural Resources Canada, Canadian Forest Service, Information Repost LAU-X-137E. Retrieved from http://cfs.nrcan.gc.ca/pubwarehouse/pdfs/34818.pdf

Renaut, S., & Rieseberg, L. H. (2015). The accumulation of deleterious mutations as a consequence of domestication and improvement in sunflowers and other compositae crops. *Molecular Biology and Evolution*, *32*, 2273–2283. https://doi.org/10.1093/molbev/msv106

Rezvoy, C., Charif, D., Guéguen, L., & Marais, G. A. (2007). MareyMap: an R-based tool with graphical interface for estimating recombination rates. *Bioinformatics*. *23*(16), 2188–2189. https://doi.org/10.1093/bioinformatics/btm315

Rey, O., Danchin, E., Mirouze, M., Loot, C., & Blanchet, S. (2016). Adaptation to global change: A transposable element-epigenetics perspective. *Trends in Ecology and Evolution*, *31*, 514–526. https://doi.org/10.1016/j.tree.2016.03.013

Rodriguez, F., & Arkhipova, I. R. (2018). ScienceDirect transposable elements and polyploid evolution in animals. *Current Opinion in Genetics Development*, *49*, 115–123. https://doi.org/10.1016/j.gde.2018.04.003

Rousset, F. (1997). Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, *145*, 1219–1228.

Sagarin, R., & Gaines, S. D. (2002). The "abundant centre" distribution: To what extent is it a biogeographical rule? *Ecology Letters*, *5*, 137–147. https://doi.org/10.1046/j.1461-0248.2002.00297.x

Scott, W., & Crossman, E. (1998). *Freshwater fishes of Canada* (pp. 1–966). Ottawa: Galt House Publications Ltd.

Shrader, L., & Schmitz, J. (2019). The impact of transposable elements in adaptive evolution. *Molecular Ecology*, *28*, 1537–1549. https://doi.org/10.1111/mec.14794

Sunyaev, S., Ramensky, V., Koch, I., Lathe III, W., Kondrashov, A. S., Bork, P. (2001). Prediction of deleterious human alleles. *Human Molecular Genetics*, *10*, 591–597. https://doi.org/10.1093/hmg/10.6.591

Sutherland, B. J. G., Gosselin, T., Normandeau, E., Lamothe, M., Isabel, N., Audet, C., & Bernatchez, L. (2016). Salmonid chromosome

evolution as revealed by a novel method for comparing RADseq linkage maps. *Genome Biology and Evolution*, *8*, 3600–3617. https://doi.org/10.1093/gbe/evw262

Torvinen, E. S. (2017). *Lake Trout (Salvelinus Namaycush) otoliths as indicators of past climate patterns and growth in Arctic lakes*. ProQuest Dissertations Publishing, University of Alaska Fairbanks.

Ward, R. D., Woodwark, M., & Skibinski, D. O. F. (1994). A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *Journal of Fish Biology*, *44*, 213–232. https://doi.org/10.1111/j.1095-8649.1994.tb01200.x

Wells, Z. R. R., Bernos, T. A., Yates, M. C., & Fraser, D. J. (2019). Genetic rescue insights from population- and family-level hybridization effects in brook trout. *Conservation Genetics*, *20*(4), 851–863. https://doi.org/10.1007/s10592-019-01179-z

Whitlock, M. C., & Bürger, R. (2004). *Fixation of New Mutations in Small Populations*. IIASA Interim Report. Laxenburg, Austria: IIASA. IR-04-064.

Whitlock, M. C., Griswold, C. K., & Peters, A. D. (2003). Compensating for the meltdown: The critical effective size of a population with deleterious and compensatory mutations. *Annales Zoologici Fennici*, *40*, 169–183.

Wickham, H. (2009). *ggplot2*. New York, NY: Springer.

Wilson, C., & Mandrak, N. (2003). History and evolution of lake trout in Shield lakes: Past and future challenges. *Boreal Shield Watersheds*, *2*, 1–15.

Zhou, Y., Massonnet, M., Sanjak, J. S., Cantu, D., & Gaut, B. S. (2017). Evolutionary genomics of grape (*Vitis viniferassp*. vinifera) domestication. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 11715–11720.

Zhu, L., Deng, C., Zhao, X., Ding, J., Huang, H., Zhu, S., … Yang, Z. (2018). Endangered Père David's deer genome provides insights into population recovering. *Evolutionary Applications*, *11*, 2040–2053. https://doi.org/10.1111/eva.12705

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Ferchaud A-L, Leitwein M, Laporte M, et al. Adaptive and maladaptive genetic diversity in small populations: Insights from the Brook Charr (*Salvelinus fontinalis*) case study. *Mol Ecol*. 2020;29:3429–3445. https://doi.org/10.1111/mec.15566